

T.C.  
DOKUZ EYLÜL ÜNİVERSİTESİ  
SOSYAL BİLİMLER ENSTİTÜSÜ  
EKONOMETRİ ANABİLİM DALI  
YÜKSEK LİSANS TEZİ

**KATEGORİK VERİ ANALİZİNİN İSTATİSTİKSEL  
VERİ ANALİZİ İÇERİSİNDEKİ YERİ VE ÖNEMİ**

**Ayşem Ece YALÇINKAYA**

Danışman

**Yrd.Doç.Dr. Ali Kemal ŞEHİRLİOĞLU**

2008

## Yemin Metni

Yüksek Lisans Tezi olarak sunduğum “**Kategorik Veri Analizinin İstatistiksel Veri Analizi İçerisindeki Yeri ve Önemi**” adlı çalışmanın, tarafımdan, bilimsel ahlak ve geleneklere aykırı düşecek bir yardıma başvurmaksızın yazıldığını ve yararlandığım eserlerin kaynakçada gösterilenlerden oluştuğunu, bunlara atıf yapılarak yararlanılmış olduğunu belirtir ve bunu onurumla doğrularım.

Tarih

.../.../.....

Ayşem Ece

YALÇINKAYA

## YÜKSEK LİSANS TEZ SINAV TUTANAĞI

### Öğrencinin

**Adı ve Soyadı** :Ayşem Ece YALÇINKAYA  
**Anabilim Dalı** :Ekonometri  
**Programı** :Ekonometri  
**Tez Konusu** :Kategorik Veri Analizinin İstatistiksel Veri Analizi İçerisindeki Yeri ve Önemi  
**Sınav Tarihi ve Saati** :..../..../.... .....

Yukarıda kimlik bilgileri belirtilen öğrenci Sosyal Bilimler Enstitüsü'nün ..... tarih ve ..... sayılı toplantısında oluşturulan jürimiz tarafından Lisansüstü Yönetmeliği'nin 18. maddesi gereğince yüksek lisans tez sınavına alınmıştır.

Adayın kişisel çalışmaya dayanan tezini ..... dakikalık süre içinde savunmasından sonra jüri üyelerince gerek tez konusu gerekse tezin dayanağı olan Anabilim dallarından sorulan sorulara verdiği cevaplar değerlendirilerek tezin,

BAŞARILI OLDUĞUNA	<input type="radio"/>	OY BİRLİĞİ	<input type="radio"/>
DÜZELTİLMESİNE	<input type="radio"/>	OY ÇOKLUĞU	<input type="radio"/>
REDDİNE	<input type="radio"/>		

İle karar verilmiştir.

Jüri teşkil edilmediği için sınav yapılamamıştır. \*\*\*  
Öğrenci sınava gelmemiştir. \*\*

\* Bu halde adaya 3 ay süre verilir.  
\*\* Bu halde adayın kaydı silinir.  
\*\*\* Bu halde sınav için yeni bir tarih belirlenir.

Evet

Tez burs, ödül veya teşvik programlarına (Tüba, Fulbright vb.) aday olabilir.   
Tez mevcut hali ile basılabilir.

Tez gözden geçirildikten sonra basılabilir.   
Tezin basımı gerekliliği yoktur.

JÜRİ ÜYELERİ

İMZA

.....	<input type="checkbox"/>	Başarılı	<input type="checkbox"/>	Düzeltilme	<input type="checkbox"/>	Red	.....
.....	<input type="checkbox"/>	Başarılı	<input type="checkbox"/>	Düzeltilme	<input type="checkbox"/>	Red	.....
.....	<input type="checkbox"/>	Başarılı	<input type="checkbox"/>	Düzeltilme	<input type="checkbox"/>	Red	.....

## ÖZET

Yüksek Lisans Tezi

Kategorik Veri Analizinin İstatistiksel Veri Analizi İçerisindeki Yeri ve Önemi

Ayşem Ece YALÇINKAYA

Dokuz Eylül Üniversitesi

Sosyal Bilimler Enstitüsü

Ekonometri Anabilim Dalı

İstatistiksel analiz yöntemleri verilerin ifade ettiği gerçekleri göstermenin bir yolu olarak yıllarca birçok alanda kullanılmıştır. Dünyanın var olan dengesini ispatlamak için fizikçiler, insan davranışlarını ifade edebilmek için davranışçılar, pazarlama dengelerini ve piyasa şartlarını araştırmak için işletmeciler, hastalıkları ve hasta durumlarını incelemek için tıpçılar vb.

Bunlar ve daha birçok alanda 1960'lı yıllara kadar sadece normal dağılıma dayalı analiz yöntemleri üzerinden analizler gerçekleştirilmiştir. Oysaki kullandığımız verilerin büyük bir oranı da kategorik özellik taşımaktadır. Kategorik verilerin dağılımları normal dağılıma uymayacağı gibi, analizlerinde de normal dağılım analizleri kullanılamaz.

Bu alanda en önemli çalışmalar 1990'larda başlamıştır. Bu dönemde birçok yeni analiz yöntemi ve istatistiki prensipler ortaya atılmıştır. Günümüze kadar gelen çalışmalar neticesinde, geç gelişen kategorik veri analiz yöntemleri büyük bir aşama kaydetmiştir.

Türkiye'de sadece kategorik veri analiz yöntemleri ile ilişkin spesifik bir çalışma bulunmamaktadır. Bu alandaki açığın ortadan kaldırılması amacıyla bu çalışma yapılmıştır. Çalışmada kategorik veri analiz yöntemleri ayrıntılandırılmış, uygulamalı gösterimi ise mağazacılık sektöründe bir anket uygulaması ile yapılmıştır. Uygulmanın ikincil hedefi olarak da ordinal veriler üzerinde nominal analiz yöntemlerinin uygulanabileceği ispat edilmiştir.

**Anahtar Kelimeler:** Kategorik Veri, Loglineer Analiz Yöntemi, Sıralı Lojistik Analiz Yöntemi, Multinomial Regresyon Analiz Yöntemi, Rekabet.

## **ABSTRACT**

### **Master Thesis**

## **The Place and The Importance of Categorical Data Analysis Methods in Statistical Data Analysis**

**Ayşem Ece YALÇINKAYA**

**Dokuz Eylül University  
Institute of Social Sciences  
Department Econometrics**

For years, in many areas the statistical analysis methods is used as a way to show what the statistical datas say actually. Physicists to prove the earth's equilibrium, behaviourists to express the human behaviours, administrators to research the market conditions and marketing rules, doctors to examine the illnesses and patients' situations etc.

In these and much more areas, the analysis had been making by normal distribution and its analysis methods until the 1960's. But, the huge proportion of the datas we used is in categorical type. As the distribution categorical datas do not conform with normal distribution, normal distribution analysis can not be used for them, either.

In this area, the most important studies was began in 1990's. In that period of time many statistical methods and statistical principles were invented. Until now as a result of the studies in statistic, it was made progress in categorical datas analysis methods.

There is not a specific study about categorical data analysis methods. This study's goal is to cover this hole. In the reasearch, categorical data analysis methods were shown in detail and its practice was made in the shopping sector by a questionnaire. The second goal of the research is to determine if nominal analysis methods can be used on ordinal data sor not.

**Key Words:** Categorical Data, Loglinear Analysis Methods, Ordinal Logit Analysis Methods, Multinomial Regression Analysis Methods, Competition.

# KATEGORİK VERİ ANALİZİNİN İSTATİSTİKSEL VERİ ANALİZİ İÇERİSİNDEKİ YERİ VE ÖNEMİ

YEMİN METNİ.....	ii
TUTANAK.....	iii
ÖZET.....	iv
ABSTRACT.....	v
İÇİNDEKİLER.....	vi
KISALTMALAR.....	xiii
TABLolar LİSTESİ.....	xiv
ŞEKİLLER LİSTESİ.....	xvii
GİRİŞ.....	1

## BİRİNCİ BÖLÜM

### KATEGORİK VERİLER İÇİN DAĞILIMLAR VE YORUMLAMALAR

1.1. KATEGORİK YANIT VERİLERİ -----	5
1.1.1. Yanıt – Açıklayıcı Değişken Ayrımı -----	6
1.1.2. Nominal – Ordinal Ölçek Ayrımı -----	7
1.1.3. Sürekli – Kesikli Değişken Ayrımı -----	9
1.1.4. Kalitatif – Kantitatif Değişken Ayrımı -----	9
1.2. KATEGORİK VERİLER İÇİN DAĞILIMLAR -----	10
1.2.1. Binom Dağılımı -----	10
1.2.2. Çokterimli Dağılım -----	11
1.2.3. Poisson Dağılımı -----	12
1.2.4. Aşırı Yayılım -----	13
1.2.5. Çokterimli Ve Poisson Dağılımları Arasındaki İlişki -----	15
1.3. KATEGORİK VERİLER İÇİN İSTATİSTİKSEL YORUMLAMA -----	16
1.3.1. Olabilirlik Fonksiyonları Ve En Çok Olabilirlik Tahmini -----	16
1.3.2. Binom Parametresi İçin Olabilirlik Fonksiyonu Ve Ml Tahmini -----	18
1.3.3. Wald - Olabilirlik Oran - Skor Testi Üçlüsü -----	19
1.3.4. Güven Aralıklarının Oluşturulması -----	24

1.4. BİNOM PARAMETRELERİ İÇİN İSTATİSTİKSEL YORUMLAMA -----	25
1.4.1. Bir Binom Parametresi İle İlgili Testler-----	25
1.4.2. Binom Parametresi İçin Güven Aralıkları -----	27
1.4.3. Vejetaryenlerin Oranı İle İlgili Örnek-----	28
1.4.4. Kesin Olasılık Küçük Örneklem Yorumlaması -----	29
1.4.5. Orta P-Değerine Dayanan Yorumlama-----	32
1.5. ÇOK TERİMLİ PARAMETRELER İÇİN İSTATİSTİKSEL YORUMLAMA	34
1.5.1. Belirli Bir Çokterimli İçin Pearson İstatistiği -----	35
1.5.2. Örnek: Mendel’ İn Teorileri-----	36
1.5.3. Ki-Kare Teorik Açıklaması-----	37
1.5.4. Ki-Kare Olabilirlik Oranı -----	39
1.5.5. Tahminlenen Ve Beklenen Frekanslarla Birlikte Test Etme -----	40

## İKİNCİ BÖLÜM

### OLUMSALLIK (KONTENJANS) TABLOLARI

2.1. OLUMSALLIK TABLOLARI İÇİN OLASILIKLARIN YAPISI -----	45
2.1.1. Duyarlılık Ve Seçicilik -----	49
2.1.2. Kategorik Değişkenlerin Bağımsızlığı -----	50
2.1.3. Poisson, Binom Ve Çokterimli Örneklem-----	52
2.1.4. İnceleme Türleri-----	56
2.2. İKİ ORANIN KARŞILAŞTIRILMASI -----	59
2.2.1. Oranların Farkı -----	59
2.2.2. Görel Risk (Relative Risk, Oransal Risk) -----	61
2.2.3. Odds Oranı -----	63
2.2.4. Odds Oranının Özellikleri-----	65
2.2.5. Olgu-Kontrol İncelemeleri Ve Odds Oranı -----	67
2.2.6. Odds Oranı Ve Görel Risk Arasındaki İlişki -----	68
2.2.7. Odds Oranı İle Logaritmik Odds Oranına Göre Yorumlama-----	70
2.3. TABAKALI 2X2 BOYUTLU OLUMSALLIK TABLOLARINDA KISMİ BİRLİKTELİK-----	72
2.3.1. Kısmi Tablolar-----	73

2.3.2. Ölüm Cezası Örneği -----	74
2.3.3. Koşullu Ve Marjinal Odds Oranları -----	77
2.3.4. Koşullu Bağımsızlığa Marjinal Bağımsızlık-----	78
2.3.5. Homojen Birliktelik-----	81
2.3.6. Cochran-Mantel-Haenszel Metotları-----	83
2.3.7. Akciğer Kanserinin Meta Analizi Örneği -----	87
2.3.8. Ortak Odds Oranının Hesaplanması -----	88
2.3.9. Odds Oranlarının Homojenliğinin Testi-----	89
2.3.10. Koşullu Birliktelikler Konusunda Kesin (Exact) Olasılık Yorumlamaları-----	92
2.4. I X J TABLOLARI İÇİN İLAVELER -----	96
2.4.1. I X J Tablolarındaki Odds Oranları-----	96
2.4.2. Birlikteliğin Özet Ölçümleri-----	98
2.4.3. Ordinal Eğilimler: Uyumlu Ve Uyumsuz İkiliiler-----	99
2.4.4. Birlikteliğin Ordinal Ölçümü: Gamma-----	101
2.5. Kİ-KARE BAĞIMSIZLIK TESTLERİ -----	103
2.5.1. Pearson İstatistiği Ve Ki-Kare Dağılımı -----	103
2.5.2. Olabilirlik-Oran İstatistiği-----	104
2.5.3. Bağımsızlık Testleri-----	106
2.5.4. Artıklar -----	107
2.5.5. Ki-Kare Parçalanımı -----	108
2.5.6. Ki-Kare Testlerinin Yorumlanması-----	110
2.6. ORDİNAL VERİLERDE BAĞIMSIZLIĞIN TEST EDİLMESİ -----	111
2.6.1. Bağımsızlık İçin Doğrusal Eğilim (Trend) Alternatifi -----	111
2.6.2. Alkol Ve Sakat Bebek Örneği -----	113
2.6.3. Ordinal Test İle Ekstra Güç Sağlanması -----	115
2.6.4. Skorların Seçimi -----	116
2.6.5. I X 2 Ve 2 X J Tabloları İçin Eğilim Testleri -----	118
2.6.6. Nominal-Ordinal Tablolar -----	120
2.7. KÜÇÜK ÖRNEKLER İÇİN KESİN YORUMLAMA-----	120
2.7.1. Fisher' In Kesin (Exact) Olasılık Testi -----	120
2.7.2. Fisher'ın Çay Denemesi-----	122



2.7.3. P-Değerleri Ve Tip-1 Hata Olasılıkları -----	124
2.7.4. Odds Oranı İçin Küçük-Örnek Güven Aralıkları -----	126
2.7.5. Daha Büyük Tablolarda Bağımsızlığın Kesin Testleri -----	127
2.8. ÇOK YÖNLÜ TABLOLAR VE TABLOLAŞTIRILAMAYAN YANITLAR İÇİN İLAVELER-----	127

## ÜÇÜNCÜ BÖLÜM

### GENELLEŞTİRİLMİŞ DOĞRUSAL MODELLER

3.1. GENELLEŞTİRİLMİŞ LİNEER BİR MODELİN BİLEŞENLERİ-----	131
3.1.1. Rassal Bileşen -----	131
3.1.2. Sistematik Bileşen -----	132
3.1.3. Bağıntı -----	132
3.1.4. Normal Glm-----	134
3.2. İKİLİ VERİLER İÇİN GENELLEŞTİRİLMİŞ LİNEER MODELLER-----	135
3.2.1. Doğrusal Olasılık Modeli -----	135
3.2.2. Lojistik Regresyon Modeli-----	138
3.2.3. Probit Modeller-----	141
3.3. SAYILABİLİR VERİLER İÇİN GENELLEŞTİRİLMİŞ LİNEER MODELLER: POISSON REGRESYON -----	142
3.3.1. Poisson Regresyon -----	144
3.3.2. Oran Verileri İçin Poisson Regresyonu -----	144
3.4. MODEL YORUMU VE MODEL KONTROLÜ -----	145
3.4.1. Wald, Olabilirlik Oranı (Likelihood-Ratio), Skor Testi-----	146
3.4.2. Poisson Model Kontrolü-----	148
3.4.3. Model Artıkları -----	148
3.4.4. Poisson Regresyonunda Aşırı Yayılım -----	150
3.5. GENELLEŞTİRİLMİŞ LİNEER MODELLERİN UYUMU-----	151
3.5.1. Newton-Raphson Algoritması -----	151
3.5.2. Kullanılan Olabilirlik Fonksiyonunun Yorumlanması-----	153
3.5.3. Sapma -----	154
3.5.4. Glm' Lerin Avantajları -----	156

## **DÖRDÜNCÜ BÖLÜM**

### **LOJİSTİK REGRESYON**

4.1. LOJİSTİK REGRESYON MODELİNİN YORUMLANMASI-----	157
4.1.1. Lineer Yaklaşım Yorumları-----	158
4.1.2. Odds Oranının Yorumlanması-----	159
4.1.3. Olgu-Kontrol İncelemeleri İçin Lojistik Regresyon -----	160
4.2. LOJİSTİK REGRESYON İÇİN YORUMLAMA -----	161
4.2.1. Etkiler İçin Güven Aralıkları -----	161
4.2.2. Anlamlılığın Test Edilmesi -----	162
4.2.3. Olasılık Tahminlerinin Dağılımları-----	163
4.3. MODEL KONTROLÜ -----	163
4.3.1. Uyum İyiliği Ve Olabilirlik Oran Model Karşılaştırma Testleri -----	165
4.3.2. Lojit Modeller İçin Artıklar-----	166
4.3.3. Etkinin Teşhis Edilmesindeki Ölçüler-----	167
4.4. KALİTATİF TAHMİN EDİCİLER İÇİN LOJİT MODELLER -----	168
4.4.1. Lojit Modellerde Kukla Değişkenlerin Kullanımı -----	169
4.4.2. Azt Ve Aıds Örneği-----	170
4.4.3. Faktörlerin Anova Tipi Gösterimi-----	171
4.5. ÇOKLU LOJİSTİK REGRESYON-----	174
4.5.1. Birden Fazla Tahmin Ediciye Sahip Model Seçimi -----	175
4.5.2. Tahmin Edicilerin Geriye Doğru Elenmesi-----	176
4.5.3. Tahmin Edicilik Gücünün Korelasyon Özeti-----	177
4.6. LOJİSTİK REGRESYON İÇİN KESİN YORUMLAMA -----	178
4.6.1. Koşullu Maksimum Olabilirlik Yorumlaması -----	178
4.6.2. Olumsuzluk Tabloları İçin Kesin Yorumlama -----	179

**BEŞİNCİ BÖLÜM**  
**OLUMSALLIK TABLOLARI İÇİN LOG-LİNEER MODELLER**

5.1. İKİ YÖNLÜ TABLOLAR İÇİN LOG LİNEER MODELLER -----	182
5.1.1. Bağımsız Model -----	182
5.1.2. Parametrelerin Yorumlanması -----	184
5.1.3. Parametre Kısıtları -----	186
5.1.4. Doygun Model -----	187
5.2. ÜÇ YÖNLÜ TABLOLAR İÇİN LOG LİNEER MODELLER -----	190
5.2.1. Log Lineer Modeller Ve Bağımsızlık Yapısı -----	190
5.2.2. Model Parametrelerinin Yorumlanması -----	192
5.2.3. Alkol, Sigara Ve Marihuana Kullanımına İlişkin Örnek -----	193
5.3. LOG LİNEER MODELLER İÇİN YORUMLAMA -----	197
5.3.1. Log Lineer Modellerin Uygunluğu -----	197
5.3.2. Ki-Kare İyi Uyum Testleri -----	198
5.3.3. Log Lineer Artıkları -----	199
5.3.4. Kısmi Birliktelik Testleri -----	202
5.3.5. Odds Oranları İçin Güven Aralıkları -----	203
5.4. DAHA YÜKSEK BOYUTLAR İÇİN LOG LİNEER MODELLER -----	205
5.5. LOG LİNEER – LOJİT İLİŞKİSİ -----	206
5.5.1. Log Lineer Modelleri Yorumlamak İçin Lojit Modellerin Kullanımı -----	206
5.5.2. Log Lineer Modeller Ve Lojit Modeller Arasındaki Benzerlik -----	208
5.5.3. Model Seçimindeki Stratejiler -----	209

**ALTINCI BÖLÜM**  
**KATEGORİK VERİ ANALİZİ KONUSUNDA YİRMİNCİ YÜZYILDAKİ  
GELİŞMELER**

6.1. PEARSON-YULE BİRLİKTEKİLİK TARTIŞMASI -----	212
6.2. R.A. FISHER' IN KATKILARI -----	214
6.3. LOJİSTİK REGRESYON -----	216

6.4. ÇOK YÖNLÜ OLUMSALLIK TABLOLARI VE LOG LİNEER MODELLER -----	219
6.5. SON (VE GELECEKTEKİ) GELİŞMELER -----	222

## YEDİNCİ BÖLÜM

### MAĞAZACILIK SEKTÖRÜNDE BİR KATEGORİK VERİ ANALİZİ UYGULAMASI

7.1. ANKETLERİN GEÇERLİLİĞİ VE GÜVENİLİRLİĞİ .....	227
7.1.1. Soruların Zorluk Derecesi Testi.....	227
7.1.2. Anket Güvenirlilik Analizi :.....	227
7.1.3. Soru-Bütün Korelasyon Testi.....	228
7.2. ARAŞTIRMANIN EVRENİ VE ÖRNEKLEM.....	228
7.3. PROBLEM CÜMLESİ .....	229
7.4. VERİLERİ TOPLAMA VE ANALİZ YÖNTEMİ .....	229
7.5. ARAŞTIRMA SINIRLILIKLARI.....	230
7.6. ARAŞTIRMA VARSAYIMLARI .....	230
7.7. VERİLERİN ANALİZİ VE BULGULAR .....	230
7.7.1. Kontenjans (Olumsuzluk) Tabloları ve Ki-Kare Analizleri.....	242
7.7.2. Sıralı Lojit Analiz Yöntemi Uygulaması .....	257
7.7.3. Multinomial Logistic Regresyon Analiz Yöntemi Uygulaması.....	260
SONUÇ VE YORUMLAR.....	265
KAYNAKÇA.....	267
EK.....	270

## KISALTMALAR

<b>ASE</b>	: Asymtotic Standart Error
<b>BD</b>	: Breslow-Day
<b>CDA</b>	: Categorical Data Analysis
<b>CDF</b>	: Birikimli Dağılım Fonksiyonu
<b>CMH</b>	: Cochran-Mantel- Haenszel İstatistiği
<b>DF</b>	: Serbestlik Derecesi
<b>EİANTE</b>	: En İyi Asimptotik Normal Tahmin Edici
<b>GLM</b>	: Genelleştirilmiş Lineer Model
<b>MH</b>	: Mantel-Haenszel
<b>MI</b>	: Kalp Kası Bozukluğu
<b>ML</b>	: Maximum Likelihood
<b>OR</b>	: Odds Oranı
<b>SE</b>	: Standart Error

## TABLolar LİSTESİ

Tablo 1.1 İlk Ve İkinci Enfeksiyona Yakalanan Buzağılar -----	41
Tablo 1.2 Hipoteze Göre Olasılık Yapısı -----	42
Tablo 2.1 Cinsiyete Göre Ölümden Sonraki Hayata Olan İnançların Çapraz Sınıflaması -----	45
Tablo 2.2 Tablo-2.1 İçin Notasyonlar -----	48
Tablo 2.3 Göğüs Kanseri Tanılarına Göre Tahminlenen Koşullu Dağılımlar -----	49
Tablo 2.4 Marjinal, Koşullu Ve Ortak Olasılıkların Notasyonları -----	51
Tablo 2.5 Emniyet Kemerinin Kullanımı Ve Otomobil Çarpışmalarının Sonuçları -----	55
Tablo-2.6 Akciğer Kanseri İle Sigara İçiminin Çapraz Sınıflaması -----	56
Tablo 2.7 Kalp Kası Bozukluğu (Myocardial İnfraction) Ve Aspirin Kullanımının Çapraz Sınıflandırılması -----	60
Tablo 2.8 Katilin İrkına Ve Kurbanların İrkına Göre Ölüm Cezası Kararı -----	75
Tablo 2.9 Marjinal Bağımsızlığı İçermeyen Koşullu Bağımsızlık İçin Beklenen Değerlerin Gösterilmesi -----	80
Tablo 2.10 Cochran-Mantel-Haenszel Testi İle İlgili Olarak Çinli Sigara Kullanıcıları Ve Akciğer Kanseri Yakalananların İncelenmesi -----	84
Tablo 2.11 Terfilerle Çalışanların İrklarına İlişkin Olumsuzluk Tablosu -----	93
Tablo 2.12 Gelir Aracılığı İle İş Memnuniyetinin Çapraz Sınıflaması -----	100
Tablo 2.13 Bebek Sakatlığı Ve Annenin Alkol Tüketimi -----	114
Tablo 2.14 Fisher'in Çay-Tatma Deneyi İçin Olumsuzluk Tablosu -----	123
Tablo 2.15 Tablo 2.14'in Marjinlerine Sahip Tablolar İçin Hipergeometrik Dağılım -----	123
Tablo 2.16 Veri Çizelgesine İlişkin Örnek Tablo -----	130
Tablo-3.1 Horlama Ve Kalp Rahatsızlığı Arasındaki İlişki -----	137
Tablo-3.2 İstatistiksel Analizde Model Türleri -----	156
Tablo 4.1 Azt Kullanımına Ve İrka Göre Aids Belirtilerinin Gelişimi -----	170
Tablo 4.2 Tablo 4.1' E Göre Uydurulan Lojit Model İçin Parametre Tahminleri --	173

Tablo-5.1 Ölümünden Sonraki Hayata İnanmanın (Sütunlar) Cinsiyet (Satırlar) İle Olan İlişkisinin Yer Aldığı Tablo 2.1 İçin Uydurulan Bağımsız Modele Göre Parametre Tahminleri-----	184
Tablo-5.2 Doygun Log Linear Model İçin Eşdeğer Birliktelik Parametreleri -----	189
Tablo-5.3 Lise Son Sınıf Öğrencileri İçin Alkol(A), Sigara(C) Ve Marihuana (M) Kullanımı (Kaynak: Agresti, 1996)-----	193
Tablo-5.4 Tablo-5.3' Teki Verilere Uygulanan Log Linear Modeller İçin Beklenen Değerler -----	194
Tablo-5.5 Tablo-5.4' Te Yer Alan Log Linear Modeller İçin Tahmin Edilmiş Odds Oranları -----	196
Tablo 5.6 Alkol (A), Sigara (C) Ve Marihuana (M) Kullanımına İlişkin Oluşturulan Log Linear Modeller İçin Uyum İyiliği Testleri -----	199
Tablo 5.7 İki Log Linear Model İçin Ayarlanmış Artıklar -----	201
Tablo-5.8 İki Sonuçlu Y Yanıt Değişkeni İle Üç Yönlü Bir Tablo İçin Eşdeğer Log Linear Ve Lojit Modeller -----	209
Tablo 7.1 Cinsiyetlerine Göre Frekans ve Yüzde Değerleri .....	231
Tablo 7.2 Çocuk Durumlarına Göre Frekans ve Yüzde Değerleri.....	232
Tablo 7.3 KMO and Bartlett's Test Çıktısı .....	233
Tablo 7.4 Açıklanan Toplam Varyans Çıktısı.....	233
Tablo 7.5 Çevrilmiş Faktör Matrisi (Rotated Component Matrix).....	234
Tablo 7.6 Faktör Grupları ve Ortalama Değerleri Tablosu.....	235
Tablo 7.7 Çeşitlilik-Hizmet ile Müşteri Odaklılık İçin Olumsuzluk Tablosu .....	243
Tablo 7.8 Çeşitlilik-Hizmet Faktörü ile Müşteri Odaklılık İçin Ki-Kare Analiz Tablosu.....	244
Tablo 7.9 Çeşitlilik-Hizmet ile Müşteri Odaklılık Arasındaki Simetrik Ölçüler Tablosu.....	244
Tablo 7.10. Kalite-Müşteri Odaklılık İçin Olumsuzluk Tablosu.....	245
Tablo 7.11 Kalite ve Müşteri Odaklılık Ki-Kare Analiz Tablosu.....	246
Tablo 7.12 Kalite-Müşteri Odaklılık Simetrik Ölçüler Tablosu .....	247
Tablo 7.13 Beden-Müşteri Odaklılık Olumsuzluk Tablosu.....	248

Tablo 7.14 Beden-Müşteri Odaklılık Ki-Kare Analiz Tablosu.....	248
Tablo 7.15 Beden-Müşteri Odaklılık Simetrik Ölçüler Tablosu.....	249
Tablo 7.16 Konum-Müşteri Odaklılık Olumsuzluk Tablosu.....	250
Tablo 7.17 Konum-Müşteri Odaklılık Ki-Kare Analiz Tablosu.....	250
Tablo 7.18 Konum-Müşteri Odaklılık Simetrik Ölçüler Tablosu.....	251
Tablo 7.19 Cinsiyet-Müşteri Odaklılık Olumsuzluk Tablosu.....	252
Tablo 7.20 Cinsiyet-Müşteri Odaklılık Ki-Kare Analiz Tablosu.....	252
Tablo 7.21 Cinsiyet-Müşteri Odaklılık Simetrik Ölçüler Tablosu.....	253
Tablo 7.22 Çocuk Durumu -Müşteri Odaklılık Olumsuzluk Tablosu.....	254
Tablo 7.23 Çocuk Durumu-Müşteri Odaklılık Ki-Kare Analiz Tablosu.....	255
Tablo 7.24 Çocuk Durumu-Müşteri Odaklılık Simetrik Ölçüler Tablosu.....	256
Tablo 7.25 Sıralı Lojit Model Tablosu (1.Deneme).....	258
Tablo 7.26 Sıralı Lojit Model Tablosu (2.Deneme).....	259
Tablo 7.27 Multinomial Lojistik Regresyon Model Tablosu (1.Deneme).....	261
Tablo 7.28 Multinomial Lojistik Regresyon Model Tablosu (2.Deneme).....	262



## ŞEKİLLER LİSTESİ

Şekil-1.1 $H_0 : \beta = 0$ , In Sınanmasında Üç Test Çeşidi İçin Kullanılan Log-Olabilirlik Fonksiyonu Ve Bilgisi. -----	23
Şekil-1.2 N=25 Olduğunda $\pi$ Binom Parametresi İçin Nominal %95 Güven Aralıklarına Göre Kapsama Olasılıklarının Grafiği-----	32
Şekil-2.1 Ölüm Cezası Alanların Yüzdesi-----	75
Şekil 3.1 Kalp Rahatsızlığı Ve Horlama Verileri İçin Modellerin Uyumu-----	138
Şekil 3.2 Lojistik Regresyon Fonksiyonu -----	139
Şekil 3.3 Wald, Olabilirlik Oranı Ve Skor Testlerinde Kullanılan Bilgi -----	154
Şekil 7.1 Cinsiyetlerine Göre Dağılım Grafiği .....	231
Şekil 7.2 Çocuk Durumlarına Göre Dağılım Grafiği .....	232
Şekil 7.3 Sıralı Lojit Modeli .....	263
Şekil 7.4 Sıralı Lojit Model Dörtlü Uygunluk Test Grafiği.....	263

## GİRİŞ

Günümüzde; sosyal bilimler alanında insan düşünce ve davranışlarına etki eden faktörlerin ölçülmesi ve değerlendirilmesinden tıp araştırmalarında özellikle tedavi yöntemlerinin tespit edilmesine kadar çok çeşitli alanlarda sıklıkla “kategorik veri analiz yöntemleri” kullanılmaktadır. Evrendeki doğal dengenin yapı taşı konumunda “normal dağılım” bulunmasına rağmen; belirli kıstaslar altında elde edilebilen “kategorik verileri” analiz edebilmek için farklı analiz yöntemlerine ihtiyaç duyulmaktadır. Uygulamada karşılaşılan veri problemlerinin çoğu kategorik veri içermektedir. Yapısı gereği kategorik özellik gösteren başlıca verilere anket yöntemi ile toplanan veriler örnek olarak gösterilebilir. Pazar araştırmasında kişilere şu şekilde sorular sorulabilir: “Bir ürünü kullanıp kullanmadıkları (Evet, Hayır)”, “Ne sıklıkla kullandıkları (Asla, haftada bir kereden az, ayda 4-10 kez, ayda 10 kereden fazla)”, “Bu üründen memnuniyetlerinin nasıl olduğunu (Kesinlikle kötü, bazen kötü, yansız, bazen memnun edici, kesinlikle memnun edici)”, “Rakip ürünlerden hangisini kullandıkları (A marka, B marka vs.)”. Bu örneklerin sonuncusu ve ilki nominal kategorik değişken, ikinci ve üçüncüsü ordinal kategorik değişkendir.

Kategorik veri analiz yöntemlerinin tarihsel gelişim sürecine bakıldığında, çok uzak bir geçmişi olmadığı görülmektedir. Sürekli verilere uygulanan metotlar yirminci yüzyılın başlarında karmaşık bir seviyeye ulaşırken, kategorik verilere uygulanacak istatistiksel metotların gelişiminde geç kalınmıştır. Bu alanda, İngiliz istatistikçi Karl Pearson tarafından 1900’lü yıllarda etkili çalışmalar yapılmasına rağmen, 1960’lara kadar kategorik verilere uygulanacak istatistik metotlarının gelişimi yavaş kalmıştır. Kategorik veri analizi için istatistiksel metotların geliştirilmesinde; istatistik, biyoistatistik, ekonomi, psikoloji ve sosyoloji gibi farklı alanlardaki bilim adamlarının katkılarından büyük ölçüde yararlanılmıştır. Bu disiplinler arası köken benzer metodolojiler için çoklu yorumlar oluşturmuştur. Sonuç olarak, kategorik veri analizi bilgi olarak zengindir ve çalışma alanı geniştir. Bu disiplinler arası yapı aynı zamanda mevcut tekniklerin sentezini ve birleşimini yapmaktadır. Fakat disiplinler arası terminolojilerin ve uygulamaların farklılığı bu durumu zorlaştırmaktadır.

Kategorik veri analizi ile ilgili istatistiksel metot ve teknikler son 25-30 yılda hızlı bir gelişim göstermiş ve hala göstermektedir. Uygulamalı araştırmalarda bu metot ve tekniklerin kullanılması son yıllarda olağan hale gelmekle beraber özellikle büyük bir bölümü uygun ticari yazılım ve masrafsız bilgisayar hesaplamaları ile yapılmaktadır.

Analizlerde, kategorik veriler için genellikle sayılabilir özet tablolar kullanılmaktadır. Bu tip verilerde, standart normal dağılım verilerinde kullanılan istatistiksel analiz yöntemlerinin kullanılması kesinlikle uygun değildir. Kategorik verilerin analizi için çeşitli istatistik metotların geliştirilmesi gerekmektedir (Simonoff, 2003:3). Bu çalışma ise; kategorik verilerin analiz yöntemlerine ilişkin, yeni gelişen ve disiplinler arası bölünmüş olan belirli metotları ve teorileri özet bir şekilde göstermek ve kategorik verilerin analizi için özel araçları kullanan sosyal bilimcilerin uygulamalarına yardım etmek amacını taşımaktadır.

Çalışmanın literatür kısmı altı bölümden oluşmaktadır. Birinci bölümde; kategorik verilerin önemli türleri ile kullanılan dağılımlar ve istatistiksel yorumlama yöntemleri tanıtılmıştır. Bu bölümde amaç hala tartışma konusu olan kategorik veri türlerinin genel bir tanımlanmasını sağlamak, yorumlayıcı veri analizinin gerektirdiği dağılım varsayımlarını kategorik veriler için ortaya koymak ve parametreler konusunda yorumlamalar yapmak için örneklem verilerinin kullanımına ilişkin metotları incelenmiştir. Bu bölümde, kategorik yanıtlar için “*Binom, Çokterimli ve Poisson dağılımı*” şeklinde üç ana dağılım gözden geçirilmiş ve istatistiksel yorumlama açısından en çok olabilirlik metodu ile bu metoda ilişkin testler ele alınmıştır. Son olarak binom ve çok terimli parametrelerin istatistiksel yorumlama yöntemleri ayrıntıları ile ele alınmıştır.

İkinci bölümde ise, kategorik değişkenler arasındaki ilişkileri gösteren olumsuzluk (kontenjans) tabloları ve birliktelikleri özetleyen parametreler tanımlanmıştır. Ayrıca bu bölümde ki-kare test yöntemlerine de yer verilmiştir. Öncelikle iki kategorik değişken arasında ki birliktelikler ve birlikteliği ifade eden parametreler tanıtılarak bu parametreler için yorumlama metotları gösterilmiştir. Sonra üçüncü bir değişkenin kontrol edilmesi ile alan genişletilmiştir. Çünkü kontrol altında bir birliktelik ciddi olarak değişebilmektedir. Daha sonra nominal ve ordinal çok kategorili değişkenler için parametreler ele alınmıştır.

Üçüncü bölümde, kategorik yanıt değişkenlerin modellenmesi üzerine odaklanılmış ve bir önceki bölümde açıklayıcı değişkenlerin kategorik yanıt değişkenler üzerindeki etkilerini araştırırken kullandığımız analizlerin temelini oluşturan modeller incelenmiştir. Modeller ile birçok açıklayıcı değişkenin etkilerinin eş zamanlı olarak analiz edilmesi gibi daha karmaşık durumlar ele alınmıştır. Modellerdeki açıklayıcı değişkenler sürekli ya da kategorik ya da her iki türde de olabilmektedir. Burada ele alınan modellerin neredeyse tümü genelleştirilmiş doğrusal modellerin özel biçimleridir. Bu modellerin genel bir durumudur ve kategorik yanıt değişkenler için oluşturulan modelleri kapsadığı gibi sürekli yanıt değişkenler için kullanılan sıradan regresyon ve ANOVA modellerini de kapsamaktadır. Ayrıca uyumluluk testlerine de yer verilmiştir.

Kategorik analiz yöntemleri arasında en çok kullanım yeri bulan lojistik regresyon analiz yöntemi dördüncü bölümde, bir diğer önemli analiz yöntemi olan Loglineer analiz yöntemi ise beşinci bölümde ayrıntılandırılmıştır. Altıncı bölümde, kategorik veri analizinin kısa bir tarihsel gelişimi sunulmuştur. Yirminci yüzyılın başlarından günümüze kadar yapılan çalışma ve katkılara değinilerek araştırmacılar için kronolojik bir kaynak listesi verilmiştir. Ayrıca bundan sonraki dönemlerde kategorik veri analizi konusunda günümüzdeki çalışmalara dayanarak hangi yönde gelişim sağlanabileceğine ilişkin tahminleme yapılmıştır.

Çalışmanın son bölümü ise kategorik veriler üzerinden yapılan bir uygulamadan meydana gelmektedir. Uygulama alanı olarak “departman mağazacılığında rekabet olanaklarının geliştirilmesi” maksadıyla bu alanda hizmet veren bir mağaza ve müşterileri seçilmiştir. Bu anlamda mağazaya ait temel süreçlere ilişkin kriterler, belirli bir anket üzerinden müşteriler ile yapılan karşılıklı görüşmeler sonucu belirlenmiştir. Anket sonuçlarının değerlendirilmesi ile iyileştirme sürecinde yapılan analizlerde kategorik veri analizi yöntemlerinden yararlanmıştır. Çalışma sonunda elde edilen anket verilerine uygulanan sıralı lojit analiz yöntemi ile multinominal regresyon analiz yönteminin kullanımının karşılaştırılması hedeflenmiştir.

# BİRİNCİ BÖLÜM

## KATEGORİK VERİLER İÇİN DAĞILIMLAR VE YORUMLAMALAR

İstatistik gelişim sürecinde önceleri doğadaki dengelerin bir dağılım arz ettiği ve bu dağılımın normal dağılım olduğu ifade edilmiştir. O zamanki bilgiler ile her veri grubunun normal dağılım sergilemesi beklenirken, yapılan çalışmalar dengede payı olan dağılım tiplerinin sadece normal dağılım olmadığı ve her veri kütesinin de normal dağılım sergilemediğini göstermiştir. Ancak, bir verinin normal dağılım sergilememesinin de o verilerin belirli bir düzen dahilinde incelenemeyeceği anlamını taşımamaktadır. İşte bu bölümde normal dağılım sergilemeyen veri kütleleri arasında yer alan kategorik veri tipleri ile analizlerinde esas alınan dağılım şekilleri, bu dağılımlar çerçevesinde istatistiksel yorumlama esasları alt ana başlıklar altında irdelenmiştir. Ana başlıklara geçmeden evvel kategorik verilerin önemli türleri tanımlanmıştır.

### 1.1. KATEGORİK YANIT VERİLERİ

Kategorik bir değişken; belirli kategori kümelerinden oluşan bir ölçüm ölçeğine sahiptir. Örnek olarak; “politik felsefe” genellikle “liberal, ılımlı, tutucu” şeklindeki kategorilerle ölçülmektedir ya da memogram çekimlerine dayanarak göğüs kanserine ilişkin teşhisler; “ normal, iyi huylu, olasılıkla iyi huylu, şüpheli, kötü huylu” şeklindeki kategoriler kullanılarak yapılır(Agresti, 2002). Sosyal ve biyomedikal bilimlerdeki araştırma çalışmaları kategorik değişkenlere uygulanacak istatistiksel metotların gelişimini teşvik etmiş ve harekete geçirmiştir. Kategorik ölçekler *sosyal bilimlerde* tutum ve düşüncelerin ölçülmesinde yaygın olarak kullanılmaktadır. *Biyomedikal* bilimlerde ise “medikal bir tedavinin başarılı olup olmaması” şeklindeki sonuçların ölçülmesinde kategorik ölçekler kullanılmaktadır. Kategorik veriler sosyal ve biyomedikal bilimlerde oldukça yaygın biçimde kullanılmasına rağmen kullanım açısından asla bu alanlarla sınırlı değildir.

Kategorik veriler aşağıda sayılan alanlarda da sıklıkla görülmektedir:

- Eğitim alanında bir sınav sorusu için öğrenci yanıtlarının “doğru / yanlış” şeklinde kategorilerle değerlendirilmesi gibi.
- Pazarlama alanında bir ürüne ilişkin lider markalar arasında tüketici tercihlerinin “A markası, B markası ve C markası” şeklindeki kategorilerle ifade edilmesi gibi.
- Kategorik değişkenler aynı zamanda mühendislik bilimi ve endüstriyel kalite kontrol gibi kantitatif alanlarda da görülmektedir. Üretilen parçaların belirli standartlara uygun olup olmaması şeklinde sınıflandırılması, belirli bir kumaşın dokusunun yumuşaklığının ne kadar olması gerektiği, belirli bir yiyeceğin tadının ne kadar iyi olduğu ya da bir çalışanın belirli bir görevi ne kadar kolay yerine getirdiği gibi belirli karakteristiklerin öznel olarak iyileştirilmesi örnek olarak verilebilir.

Uygulamalı bilim dallarında ve mühendislikte bazı araştırma konularında ölçüm veya sayımla ifade edilebilen sürekli ve kesikli değişkenler de standart hale getirilmiş ya da getirilememiş kategoriler halinde belirlenir. Ağrı kesicilerin klinik değerlendirilmelerinin yapıldığı ağrı derecelendirilmesi araştırmasında hastaların süreklilik arz eden ağrıları olmasına karşın bu ağrıların sayısal bir ölçekle değerlendirilmesi oldukça zordur. Ayrıca hastalar az ağrıyla orta derecede ağrıyı ayırt edebilecek durumda olabilirler. Böyle bir çalışmada değerlendirme; 0=ağrısız durum, 1=az ağrı, 2= orta derecede ağrı ve 3= çok ağrı şeklinde yapılabilir. Bu durum sürekli bir değişkenin kesikli değerlerle ifade edildiği sıralama ölçeği olarak tanımlanan bir değerlendirme türüdür. Yukarıdaki örneklerden de anlaşıldığı üzere kategorik değişkenlerin birçok türü vardır. Bu bölümde kategorik değişkenlerin ve diğer değişkenlerin sınıflandırılması üzerinde durulacaktır.

### 1.1.1. Yanıt – Açıklayıcı Değişken Ayrımı

Çoğu istatistiksel analiz, *yanıt (bağımlı)* değişkenler ile *açıklayıcı (bağımsız)* değişkenler arasında ayırım yapmaktadır. Örnek olarak regresyon modelleri, bir evin satış değerinin, açıklayıcı değişkenlerin değerlerine göre (bölge ve alan uzunluğu

değerleri gibi) nasıl değiştiği konusunda olduğu gibi bir yanıt değişkeninin nasıl bir anlamı olduğunu tanımlamaktadır. Bu çalışmada kategorik yanıt değişkenlerine uygulanacak metotlara odaklanılmaktadır. Sıradan regresyonda olduğu gibi açıklayıcı değişkenler herhangi bir türde olabilmektedir.

### 1.1.2.Nominal – Ordinal Ölçek Ayrımı <sup>1</sup>

Kategorik değişkenler iki temel ölçek türüne sahiptir. Değişkenler doğal bir sırada olmayan kategorilere sahipse *nominal* olarak adlandırılır. Bu ölçekte ölçülmüş sayıların anlamı yalnız ve yalnız temsil ettikleri bireyleri, olayları veya durumları tanıtmalarıdır (Ekici ve Karagöz, 2004). Dini inançlar (Katolik, Protestan, Yahudi, Müslüman ve diğerleri), işe gidiş-geliş şekilleri (otomobil, bisiklet, otobüs, metro, yayan), favori müzik türleri (klasik, caz, rock, türkü, halk müziği ), ikamet yeri seçimi ( apartman, villa, diğerleri) şeklinde örnekler verilebilir. Nominal değişkenler için listelenen kategorilerin sırası önemsizdir. İstatistiksel analiz sıraya bağlı değildir.

Birçok kategorik değişken kendiliğinden sıralı kategorilere sahiptir. Bu şekildeki değişkenlere *ordinal* değişken adı verilir. Otomobil büyüklüğü (küçük araba, orta genişlikteki araba, geniş araba ), sosyal sınıf (üst, orta, alt), siyasi görüş (liberal, ılımlı, tutucu), hastalık durumu (iyi, orta, ciddi, kritik) şeklinde örnekler verilebilir. Örneklerde de görüldüğü gibi ordinal değişkenler sıralanmış kategorilere sahiptir. Ancak kategoriler arasındaki mesafe bilinmemektedir.

İlımlı olarak kategorize edilen kişi, tutucu olarak kategorize edilen kişiye göre daha liberaldir. Ancak ne kadar liberal olduğu sayısal bir değer ile ifade edilemez. Ordinal değişkenlere uygulanan metotlarda kategorilerin sıralamasından faydalanılmaktadır.

---

<sup>1</sup> Stevens (1951) ölçüm ölçeklerini ( nominal, ordinal ve aralık) tanımlamıştır. Diğer ölçekler bu türlerin bir karışımından meydana gelmektedir. Örnek olarak; kısmen sıralı ölçekler, bilinmeyen ya da karar verilmemiş kategoriler dışında sıralı kategorilere sahip sorulara verilen yanıtlar söz konusu olduğunda ortaya çıkmaktadır (Oktay,2003,1).



Bir *aralık (interval)* deęişkeni her hangi iki deęer arasında sayısal olarak ifade edilebilen uzaklıklara sahip bir deęişken türüdür. Kan basıncı seviyesi, televizyon cihazının fonksiyonel yaşam uzunluęu, hapisane süresinin uzunluęu, yıllık gelir gibi örnekler verilebilir (Bazen bir aralık deęişkeni, eęer deęerlerin oranları da geçerli ise, *oran deęişkeni* olarak da adlandırılır.).

*Bir deęişkenin ölçüm yöntemi onun sınıfını da belirlemektedir*<sup>2</sup>. Örnek olarak “eęitim” sadece devlet okulu ve özel okul şeklinde ölçülmüşse nominal, eęer “yok, lise, lisans, yüksek lisans doktora” şeklindeki kategoriler kullanılarak elde edilebilen en yüksek dereceye göre ölçülürse ordinal, eęer 0,1,2... gibi tamsayılar kullanılarak eęitim görülen yıllara göre ölçülürse aralık deęişkeni olur. Bir deęişkenin ölçüm ölçeęi hangi istatistiksel metodun uygun olacaęını da belirlemektedir. Ölçüm hiyerarşisinde (aşamalı sistemde); aralık deęişkenleri en yüksekte, sonra ordinal deęişkenler ve en altta da nominal deęişkenler yer almaktadır(Agresti, 2002).

Bir tip deęişkene uygulanacak istatistiksel metotlar daha yüksek seviyedeki deęişkenlere de uygulanabilir ama alt seviyedekilere uygulanamaz. Örnek olarak nominal deęişkenlere uygulanan istatistiksel metotlar kategorilerin sıraları göz ardı edilerek ordinal deęişkenlere de uygulanabilir. Fakat ordinal deęişkenlere uygulanan metotlar nominal deęişkenlere uygulanamaz. Çünkü nominal deęişkenlerin kategorileri anlamlı bir sıra içerisinde deęildir. Genellikle en iyi yol gerçek ölçeęe uygun istatistiksel metotları uygulamaktır.

Bu çalışmada kategorik yanıtlarla uğraşıldığından nominal ve ordinal deęişkenlerin analizi ele alınacaktır. Metotlar aynı zamanda; küçük fark deęerlerine sahip (evlilik süresi gibi) ya da sıralı kategorilere göre gruplandırılmış deęerlere sahip (eęitimin “<10 yıl, 10–12 yıl, >10 yıl” şeklinde ölçülmesi) aralık deęişkenlerine de uygulanabilir.

---

<sup>2</sup> Bazı durumlarda belirli bir ölçekte gözlemlenen bazı deęişkenler ölçek deęiştirilmesi yapılarak başka bir deęişkene dönüşebilir. Aslında sürekli nümerik bir deęişken olan aęırlığın “ zayıf, orta, şişman” gibi nominal şekle dönüştürülmesi buna bir örnektir.

### 1.1.3. Sürekli – Kesikli Değişken Ayrımı

Değişkenler genellikle aldıkları değerlere göre *sürekli ya da kesikli* olarak sınıflandırılır. Belirli bir aralıkta her değeri alabilen ve bu değerleri ölçümler sonucu elde edilebilen değişkenler sürekli niteliktedir. Belirtilen aralıkta elde edilebilecek mümkün gözlem sayısı ölçüm aletinin hassasiyetine bağlıdır.

Yapılan gözlemler sonucunda tam sayısal değerler alabilen ve genelde sayımlarla elde edilen değerler için kullanılan değişkenler kesikli niteliktedir. Bu değişkenler için iki tamsayı arasında ara değerler mümkün değildir. Bütün değişkenlerin gerçek ölçümleri, ölçüm enstrümanlarının hassaslık konusundaki sınırlarından dolayı kesikli bir şekilde meydana gelmektedir. Uygulamada değişkenler arasında Sürekli-Kesikli Sınıflaması değişkenlerin birçok değer alması ya da az miktarda değer alması şeklinde ayırım yapılmasını sağlamaktadır. Örnek olarak; istatistikçiler eğer değerlerin sayısı çok büyükse (test skorları gibi) kesikli aralık değişkenlerine sürekli gibi muamele ederler ve sürekli yanıtlar için kullanılan metotları uygularlar.

Bu çalışmada kesikli bir şekilde ölçülmüş yanıtların belirli tipleriyle ilgilenilecektir:

- Nominal değişkenler
- Ordinal değişkenler
- Göreceli olarak az miktarda değere sahip kesikli aralık değişkenleri
- Az sayıda kategorilere göre sınıflandırılmış sürekli değişkenler

### 1.1.4. Kalitatif – Kantitatif Değişken Ayrımı

Nominal değişkenler **kalitatif (nitel)** yapıdadır- *farklı kategoriler* farklı miktarları değil farklı nitelikleri (kaliteleri) ifade eder.<sup>3</sup> Aralık değişkenleri **kantitatif (nicel)** yapıdadır-*farklı seviyeler* ilgilenilen özellik açısından farklı miktarlara

---

<sup>3</sup> Burada gözlemleri yapılan bireyler kalite sınıflarına ayrılmaktadır.

sahiptir. Kalitatif - kantitatif sınıflandırmasında ordinal değişkenlerin durumu belirsizdir. Analistler sıklıkla ordinal değişkenlere kalitatif olarak muamele ederler ve nominal değişkenlere uygulanan metotları kullanırlar. Ancak bir çok açıdan ordinal değişkenler nominal değişkenlere kıyasla aralık değişkenlerine daha çok benzemektedir. Çünkü ordinal değişkenler önemli nicel (kantitatif) bir özelliğe sahiptir: Her bir kategori diğer bir kategoriye göre özellik açısından daha fazla ya da daha az öneme sahiptir. Ölçülmesi olanaklı olmasa bile esas olarak genellikle sürekli bir değişken olarak ortaya çıkarlar. Siyasi görüşlerin sınıflandırılması (liberal, ılımlı, tutucu) niteliği itibariyle sürekli bir özellik olarak ölçülür. Analistler sıklıkla kategorilere nümerik puanlar (skorlar) atayarak ya da temel bir sürekli dağılım varsayımı yaparak ordinal değişkenlerin kantitatif (nicel) yapısından yararlanırlar. Bu durum, ölçeği kullanan araştırmacıların sağduyusunu, değerlendirmelerini ve rehberliğini gerektirir (Agresti,2002). Veri analizi için mevcut olan metot çeşitlerinde de yarar sağlamaktadır.

## 1.2. KATEGORİK VERİLER İÇİN DAĞILIMLAR

Yorumlayıcı veri analizleri, verilerin üretilmesindeki rastgelelik mekanizması konusunda bazı varsayımları gerekli kılar. Sürekli yanıtlar için kullanılan regresyon modellerinde normal dağılım, merkezi bir rol oynamaktadır. Bu bölümde, kategorik yanıtlar için üç ana dağılım gözden geçirilecektir (Agresti, 2002; Zelderman, 1999): *Binom, Çokterimli ve Poisson dağılımı.*

### 1.2.1 Binom Dağılımı

Birçok uygulamada belirli “n” sayıdaki ikili gözlemlerden yararlanılır.  $y_1, y_2, \dots, y_n$ ; eşit ve birbirinden bağımsız n tane denemenin yanıtlarını gösterebilir. Buna göre  $P(Y_i = 1) = \pi$  ve  $P(Y_i = 0) = 1 - \pi$  şeklindedir. 1 ve 0 şeklindeki sonuçlar “başarı” ve “başarısızlık” şeklindeki sayısal etiketleri ifade etmektedir. *Eşit denemelerin* anlamı  $\pi$  başarı olasılığının her denemede aynı olmasıdır. *Denemelerin bağımsızlığı*;  $\{Y_i\}$  değerlerinin bağımsız rassal değişkenler olduğu anlamına gelmektedir. Bunlar genellikle Bernoulli Denemeleri olarak adlandırılmaktadır.

Başarıların toplam sayısı  $Y = \sum_{i=1}^n Y_i$  ile gösterilmektedir.  $bin(n, \pi)$  şeklindeki ifade ise  $\pi$  parametrelili n indeksli **binom dağılımını** göstermektedir. Y değişkeni için olası y tane sonucu gösteren olasılık yoğunluk fonksiyonu aşağıdaki şekildedir:

$$p(y) = \binom{n}{y} \pi^y (1 - \pi)^{n-y}, y = 0, 1, 2, \dots, n.$$

$$\text{Binom katsayısı: } \binom{n}{y} = \frac{n!}{y!(n-y)!}$$

$E(Y_i) = E(Y_i^2) = 1 \times \pi + 0 \times (1 - \pi) = \pi$  olduğundan dolayı;  $E(Y_i) = \pi$  ve  $\text{var}(Y_i) = \pi(1 - \pi)$  olur.  $Y = \sum_i Y_i$  için binom dağılımının ortalaması ve varyansı ise  $\mu = E(Y) = n\pi$  Ve  $\sigma^2 = \text{var}(Y) = n\pi(1 - \pi)$  olur. Çarpıklık ise  $E(Y - \mu)^3 / \sigma^3 = (1 - 2\pi) / \sqrt{n\pi(1 - \pi)}$  ile ifade edilir. Sabit bir  $\pi$  değeri için n arttıkça binom dağılımı normale yaklaşır.

Ardaşık ikili gözlemlerin birbirinden bağımsız olmasının ya da tek olmasının garantisi yoktur. Bu yüzden ara sıra diğer dağılımlardan yararlanılır ve onlar kullanılır (Agresti, 2002). Sonlu ve ölçülebilir bir popülasyondan yerine koymaksızın yapılan iki sonuçlu örneklem sözü edilen durumlardan biridir (20 kişilik bir sınıftan elde edilen 10 öğrencilik örnekleme cinsiyetlerin gözlemlenmesi). Daha sonra ele alınacak hipergeometrik dağılım daha uygundur. Aşırı yayılım konusunda binom varsayımlarını bozan diğer durumlardan da bahsedilecektir.

### 1.2.2. Çokterimli Dağılım

Bazı denemeler iki olası sonuçtan daha fazlasına sahiptir. n tane birbirinden bağımsız ve eşit olasılıklı denemeden her birinin c kategorili sonuca sahip olduğunu varsayalım. Eğer i. deneme j. kategoride bir sonuca sahip ise  $y_{ij} = 1$  değilse  $y_{ij} = 0$

olur. Öyleyse  $y_i = (y_{i1}, y_{i2}, \dots, y_{ic})$  çokterimli bir denemeyi göstermektedir ve  $\sum_j y_{ij} = 1$  olur.

Örnek olarak (0, 0, 1, 0) şeklindeki durum dört olası kategoriden üçüncü kategoride sonuca sahip olduğunu göstermektedir.  $y_{ic}$  gereksiz bir gösterimdir ve doğrusal bir şekilde diğerlerine bağlı olmaktadır.  $n_j = \sum_i y_{ij}$ , j. kategoride sonuca sahip deneme sayısını temsil etmektedir.  $(n_1, n_2, \dots, n_c)$  sayıları çokterimli dağılıma sahiptir.

$\pi_j = P(Y_{ij} = 1)$  her bir denemede j. kategoride sonuç bulunma olasılığını göstermektedir. Çokterimli olasılık yoğunluk fonksiyonu aşağıdaki şekildedir:

$$p(n_1, n_2, \dots, n_{c-1}) = \left( \frac{n!}{n_1! n_2! \dots n_c!} \right) \pi_1^{n_1} \pi_2^{n_2} \dots \pi_c^{n_c}$$

$\sum_j n_j = n$  olduğu için  $n_c = n - (n_1 + \dots + n_{c-1})$ , (c-1) boyutludur. c=2 çokterimli dağılımın özel bir durumudur ve sonuç binom dağılımıdır. Ayrıca Her bir  $n_j$  için marjinal dağılım, binom dağılımı gösterir. Çokterimli dağılım binom dağılımının genelleştirilmesidir ve beklenen değer, varyans ve kovaryans eşitlikleri aşağıdaki biçimdedir:

$$E(n_j) = n\pi_j, \quad \text{var}(n_j) = n\pi_j(1 - \pi_j), \quad \text{cov}(n_j, n_k) = -n\pi_j\pi_k$$

### 1.2.3. Poisson Dağılımı

Bazen sayılabilir veriler sabit bir deneme sayısından meydana gelmeyebilir. Eğer y= Otoyolda gelecek hafta boyunca otomobil kazalarından dolayı meydana gelen ölüm sayısı ise y için sabit bir n sayıda üst sınır yoktur. y negatif olmayan bir tamsayı olması gerektiği için onun dağılımı, belirlenmiş bir aralıktaki yoğunluğunu vermektedir. Bu şekildeki dağılımlardan en basiti poisson dağılımıdır. Poisson

dağılımı belirli bir zaman ya da aralık boyunca rastgele meydana gelen olayların sayısını belirlemek için kullanılmaktadır. Sonuçlar birbirinden bağımsız olarak parçalanmış periyotlar ya da aralıklar içerisinde meydana geliyorsa poisson dağılımı geçerlidir. Poisson dağılımının olasılıkları, ortalama ( $\mu$ ) şeklindeki tek bir parametreye bağlıdır. Poisson olasılık yoğunluk fonksiyonu:

$$p(y) = \frac{e^{-\mu} \mu^y}{y!}, y = 0,1,2,\dots$$

Beklenen değeri  $E(Y) = \text{var}(Y) = \mu$  şeklindedir.  $\mu$ ' nün tamsayı kısmına eşit bir tepe değerine sahip tek tepeli bir dağılımdır. Çarpıklığı ise;  $E(Y - \mu)^3 / \sigma^3 = 1/\sqrt{\mu}$  ile tanımlanmıştır.  $\mu$  arttıkça dağılım normale yakınsar.

Aynı zamanda;  $\mu = n\pi$  durumunda n yeteri kadar büyükse ve  $\pi$  yeteri kadar küçükse binom dağılımı yerine yaklaşık olarak Poisson Dağılımı kullanılır. Bu nedenle gelecek hafta içerisinde otoyolda 50 milyon sürücüden her birinin o hafta içerisinde ölümcül bir kazada ölme olasılığı 0,000002 iken birbirinden bağımsız bir denemeyi oluşturuyor ise  $Y=\text{ölüm sayısı bin}(50.000.000, 0,000002)$  değişimini ya da yaklaşık olarak  $\mu = n\pi = 50.000.000(0,000002) = 100$  ortalamalı Poisson değişimi göstermektedir. Poisson dağılımın temel özelliği ise varyansının ortalamasına eşit olmasıdır. Örneklem ortalaması arttıkça örneklem sayıları daha da fazla değişim göstermektedir. Haftalık ölümcül kaza sayısının ortalaması 10 iken meydana gelebilecek haftalık gerçekleşme sayısına kıyasla, ortalama 100 olduğunda haftalık gerçekleşme sayısı daha fazla değişim gösterecektir.

#### 1.2.4. Aşırı Yayılım

Uygulamada poisson veya binom ile tahmin edilen gözlemlerin miktarı sıklıkla aşırı değişkenlik göstermektedir. Bu olaya “aşırı yayılım (overdispersion)” adı verilmektedir (Agresti, 2002). Yukarıda, gelecek hafta içerisinde ölümcül bir kazada ölme olasılığının her bir birey için eşit olduğu varsayılmıştı. Daha gerçekçi bir yaklaşıma göre, kişinin emniyet kemeri takıp takmaması, bulunulan coğrafi bölge

ve kişinin araba kullanırken harcadığı süre gibi çeşitli faktörlerden dolayı bu olasılıklar değişkenlik göstermektedir. Bu durum, kaza sayısının “Poisson Modeli” ile tahmin edilen duruma kıyasla daha fazla değişim göstermesine neden olur. Belli bir  $\mu$  için  $Y$ ,  $\text{var}(Y/\mu)$  varyanslı rassal bir değişken iken  $\theta = E(\mu)$  ifadesi mutlak olarak aşağıdaki biçimdedir:

$$E(Y) = E[E(Y/\mu)] , \text{ var}(Y) = E[\text{var}(Y/\mu)] + \text{var}[E(Y/\mu)]$$

$Y$  belirli bir  $\mu$  için koşullu poisson ise  $\theta = E(\mu) = E(Y)$  ve  $\text{var}(Y/\mu) = E(\mu) + \text{var}(\mu) = \theta + \text{var}(\mu) > \theta$  olur. Sayılabilir bir değişken için poisson dağılımı varsayımı aşırı yayılıma neden olan faktörlerden dolayı oldukça basittir. Negatif binom dağılımı sayılabilir verilerle ilişkili bir dağılımdır ve varyansın ortalamayı aşmasına olanak tanır (Agresti, 2002).

Binom dağılımı ya da çokterimli dağılım varsayımı yapılan analizler bazen aşırı yayılımdan dolayı geçersiz olabilmektedir. Ölçülemeyen değişkenlerden dolayı gerçek dağılım parametresi değişkenlik gösteren farklı binom dağılımlarının bir karışımı olduğu zaman oluşabilen bir durumdur (Yeşilova, 2003). Bunu örnekle açıklamak için toksine maruz bırakılan hamile farelere ilişkin bir deneyi ele alalım. Bir hafta sonra sakatlık belirtisi gösteren her bir fare yığınınındaki ceninlerin sayısı gözlemlenmiş olsun.  $n_i$ ; i.fare yığınınındaki ceninlerin sayısını gösterebilir. Fareler aynı zamanda genetik donanımı, genel sağlığı ve ağırlıkları gibi ölçülememiş olan diğer faktörlere göre farklılık göstermektedir.  $\pi$ , sakatlık olasılığı yığından yığına değişkenlik gösterdiğinden dolayı da fazladan ilave değişkenlik ortaya çıkmaktadır. Sakatlıkların görüldüğü her yığındaki cenin sayısının dağılımı 0'a ve  $n_i$ 'e yakın bir şekilde toplanmış olabilir ve  $\pi$  gibi tek bir değere sahip ikiterimli örneklemeden beklenene oranla daha fazla yayılım gösterir. Aşırı yayılım aynı zamanda herhangi bir dağılıma göre  $\pi$ 'nin her bir yığındaki cenin sayısı için değiştiği durumda da ortaya çıkabilmektedir.

---

<sup>4</sup> Biraz evvel tanımlanan ancak ölçülemeyen faktörlerden dolayı ortalama  $\mu$ , kendiliğinden değişim göstermektedir.

Binom ve poisson dağılımlarına ait aşırı yayılmış veriler için geliştirilmiş metotlar söz konusudur. Aşırı yayılıma birçok neden gösterilebilir. Lojistik regresyonda, söz konusu nedenlerden en önemli olanları, başarı olasılıkları arasındaki varyasyon, ikili cevaplar arasındaki korelasyon, yanlış bağlantı fonksiyonunun kullanılması, modelde olması gereken önemli terimlerin olmaması ve örnek büyüklüğünün yetersiz olması şeklinde verilebilir ( Czado, 2003; Lawles, 1987).

### 1.2.5. Çokterimli ve Poisson Dağılımları Arasındaki İlişki

Otoyolda gelecek hafta içerisinde,  $y_1$ : otomobil kazalarında ölen insan sayısı,  $y_2$ : uçak kazasında ölen insan sayısı,  $y_3$ : demiryolu kazasında ölen insan sayısı olsun.  $(Y_1, Y_2, Y_3)$  değişkenleri için bağımsız poisson rassal değişkenleri olarak,  $(\mu_1, \mu_2, \mu_3)$  parametrelerinden oluşan bir poisson modeli söz konusudur.  $\{Y_i\}$  için ortak olasılık yoğunluk fonksiyonu  $p(y) = \frac{e^{-\mu} \mu^y}{y!}, y = 0, 1, 2, \dots$  şeklindeki üç tane olasılık yoğunluk fonksiyonunun çarpımıdır. Toplam  $n = \sum Y_i$  aynı zamanda  $\sum \mu_i$  parametrelili bir poisson dağılımı göstermektedir. Poisson örnekleminde toplam n sayısı belirlenmiş olmayıp oldukça rassaldır. Her bir  $\{Y_i\}$  değeri n sayısını aşamaz. c sayıda bağımsız poisson değişkeni açısından  $E(Y_i) = \mu_i$  ile  $\sum Y_i = n$  için koşullu dağılım elde edilebilmektedir. Bunu sağlayan  $\{n_i\}$  sayısının bir kümesinin koşullu olasılığı şu şekildedir:

$$\begin{aligned} & P \left[ (Y_1 = n_1, Y_2 = n_2, Y_3 = n_3) \mid \sum_j Y_j = n \right] \\ &= \frac{P(Y_1 = n_1, Y_2 = n_2, Y_3 = n_3, \dots, Y_c = n_c)}{P(\sum_j Y_j = n)} \\ &= \frac{\prod_i \left[ \exp(-\mu_i) \mu_i^{n_i} / n_i! \right]}{\exp(-\sum \mu_j) (\sum \mu_j)^n / n!} = \frac{n!}{\prod_i n_i!} \prod_i \pi_i^{n_i} \end{aligned}$$



Burada  $\{\pi_i = \mu_i / (\sum \mu_i)^n\}$  şeklindedir. Bu n örneklem büyüklüğü ve  $\{\pi_i\}$  olasılıkları ile tanımlanan  $(n, \{\pi_i\})$  şeklinde çok terimli bir dağılımdır. Birçok kategorik veri analizinde çokterimli dağılım varsayımı yapılmaktadır. Poisson dağılımı varsayımı yapılan analizler gibi bu şekildeki analizler de olabirlik fonksiyonlarındaki benzerliklerinden dolayı benzer parametre tahminlerine sahiptir.

### 1.3. KATEGORİK VERİLER İÇİN İSTATİSTİKSEL YORUMLAMA

Yanıt değişkeni için uygun dağılımın seçilmesi veri analizinin sadece bir adımıdır. Uygulamada dağılımın parametre değerleri genellikle bilinmemektedir. Bu bölümde parametreler konusunda yorumlamalar yapmak için örneklem verilerinin kullanımına ilişkin metotlar incelenecektir.

#### 1.3.1. Olabirlik Fonksiyonları ve En Çok Olabirlik Tahmini

Bu çalışmada parametre tahmini için en çok olabirlik (ML) kullanılmaktadır. Verilen veriler için olasılık dağılımının seçiminde, olabirlik fonksiyonu bu verilerin olasılığıdır ve bilinmeyen parametrelerin bir fonksiyonu olarak işlem görür. En çok olabirlik<sup>5</sup> (ML) tahmini bu fonksiyonu maksimuma çıkaran parametre değerleridir. Gözlemlenen değerlerden en yüksek gerçekleşme olasılığına sahip parametre değerlerini vermektedir.

Gerçek değerleri içinde barındıran sabit bir boyuta sahip parametre uzayı gibi zayıf koşullar altında en çok olabirlik tahmin edicilerinin sahip olması gereken özellikler aşağıdaki şekildedir (Agresti, 2002):

- Büyük örnek normal dağılımlara sahiptir.
- Asimptotik olarak *tutarlıdır* yani n arttıkça gerçek parametre değerine yakınsarlar.
- Asimptotik olarak *etkindir*.

---

<sup>5</sup> Lojistik regresyon modeliyle kestirim yapılacağı zaman *en çok olabirlik metodu* temel oluşturacaktır.

•Büyük örneklerde meydana gelen standart hatalar diğer tahmin metotlarından daha büyük değildir.

Genel anlamda ML metodu, gözlenen veri kümesini elde etmenin olasılığını maksimum yapan bilinmeyen parametrelerin değerlerini verir ve bu metodu uygulamak için önce en çok olabilirlik fonksiyonunun oluşturulması gerekmektedir(Kartalkanat, 2006). Bu fonksiyon, bilinmeyen parametrelerin bir fonksiyonu olarak, gözlenen verinin olasılığını verir. Bu parametrelerin en çok olabilirlik tahminleri, fonksiyonu maksimum yapacak değerleri bulacak şekilde seçilir. Bu nedenle sonuçta elde edilecek veriler, gözlenen verilerle çok yakın değerlere sahip olacaktır (Hesketh, Skrendal ve Prekles, 2001).

Olabilirlik fonksiyonunu maksimuma çıkaran parametre değeri aynı zamanda fonksiyonun logaritmasını da maksimuma çıkarmaktadır. Log-olabilirliğinin maksimuma çıkarılması daha kolaydır. Çünkü terimlerin toplanması çarpımından daha kolaydır.  $\beta$  genel bir problem için bir parametreyi ifade etsin ve ML tahmini de  $\hat{\beta}$  ile gösterilsin. Olabilirlik fonksiyonu  $\ell(\beta)$  ve log olabilirlik fonksiyonu ise  $L(\beta) = \log[\ell(\beta)]$  ile ifade edilir. Birçok model için  $L(\beta)$  iç bükey bir şekle sahiptir ve  $\hat{\beta}$  değeri türevin alınıp 0'a eşitlendiği noktadır(Agresti, 2002). Başka bir deyişle ML tahmini  $\partial L(\beta)/\partial \beta = 0$  şeklinde olabilirlik eşitliğinin çözümüdür. Genellikle  $\beta$  çok boyutludur ve  $\hat{\beta}$  olabilirlik eşitliklerinin bir kümesinin çözümüdür.

SE ,  $\hat{\beta}$ 'nin standart hatasını ve  $\text{cov}(\hat{\beta})$  ise  $\hat{\beta}$ 'nin asimptotik kovaryans matrisini göstermektedir.  $\text{cov}(\hat{\beta})$ , bilgi matrisinin tersidir. (j,k) bilgi matrisinin elemanı olup aşağıdaki şekilde hesaplanır:

$$-E\left(\frac{\partial^2 L(\beta)}{\partial \beta_j \partial \beta_k}\right)$$

Standart hatalar, bilgi matrisinin tersinin alınmasıyla oluşan kovaryans matrisinin köşegen elemanlarının karekökleridir(Agresti, 2002).

Log-Olabilirliğinin çarpıklığı daha fazladır ancak standart hataları daha küçüktür. Bu mantıklıdır çünkü büyük çarpıklık, log-olabilirliğin hızlı bir şekilde azaldığını belirtir ( $\beta$ ,  $\hat{\beta}$ 'den uzaklaşması). Bu nedenle veriler  $\beta$ ,  $\hat{\beta}$ 'ye yakın değerler aldıkça daha yüksek olasılığa sahip olabilirler.

### 1.3.2. Binom Parametresi için Olabilirlik Fonksiyonu ve ML Tahmini

Bir olabilirlik fonksiyonun parametreleri kapsayan kısmına “çekirdek (kernel)” adı verilmektedir. Olabilirliğin maksimizasyonu parametreler ile ilgili olduğundan geri kalanı konu dışıdır. Örnekle açıklamak için,

$p(y) = \binom{n}{y} \pi^y (1 - \pi)^{n-y}$ ,  $y = 0, 1, 2, \dots, n$ . binom dağılımını ele alalım. Binom

katsayısının  $\binom{n}{y}$ ,  $\pi$  ile ilgili olarak maksimumun sağlanması üzerinde hiçbir etkisi

yoktur. Bu yüzden binom katsayısı dikkate alınmaz ve olabilirlik fonksiyonu olarak çekirdek kısmı kullanılır. *Binom log-olabilirliği:*

$$L(\pi) = \log[\pi^y (1 - \pi)^{n-y}] = y \log(\pi) + (n - y) \log(1 - \pi)$$

$\pi$ 'ye göre türev alınırsa aşağıdaki eşitlik elde edilir:

$$\partial L(\pi) / \partial \pi = y / \pi - (n - y) / (1 - \pi) = (y - n\pi) / \pi(1 - \pi)$$

Bu eşitlik 0'a eşitlenir ve  $\hat{\pi} = y/n$  (n deneme için başarıların örneklem oranı) çözümünü sağlayan olabilirlik fonksiyonu elde edilir.

$\partial^2 L(\pi) / \partial \pi^2$  ifadesi hesaplanıp, beklenen değeri alınır ve terimler sadeleştirilirse aşağıdaki eşitliği elde ederiz:

$$-E[\partial^2 L(\pi) / \partial \pi^2] = E\left[\frac{y}{\pi^2} + \frac{(n-y)}{(1-\pi)^2}\right] = n[\pi(1-\pi)]$$

Bu nedenle  $\hat{\pi}$ 'nin asimptotik varyansı  $\pi(1-\pi)/n$  şeklindedir.  $E(Y) = n\pi$  ve  $\text{var}(Y) = n\pi(1-\pi)$  olduğundan  $\hat{\pi} = Y/n$ 'nin dağılımının ortalaması ve standart hatası:  $E(\hat{\pi}) = \pi$ ,  $\text{var}(\hat{\pi}) = \sqrt{\frac{\pi(1-\pi)}{n}}$  olur.

### 1.3.3. Wald - Olabilirlik Oran - Skor Testi Üçlüsü

Büyük örneklem yorumlamalarını gerçekleştirmek için olabilirlik fonksiyonunun kullanımına ilişkin üç standart yöntem söz konusudur. Sıfır hipotezinin ( $H_0 : \beta = \beta_0$ ) anlamlılık testi için bunlar tanıtılacak ve aralık tahminlemesi ile ilişkileri ele alınacaktır. Bu yöntemler, ML tahmin edicilerinin büyük örneklem normalliğini kullanmaktadır.

$\beta$  katsayılarının anlamlılık testleri için sıfır hipotezi  $H_0 : \beta = 0$  şeklinde ifade edilir. İstatistiksel modeldeki (GLM gibi)  $\beta$  katsayılarının anlamlılığını test etmek amacıyla kullanılan üç yöntem aşağıdaki şekildedir:

- Wald testi
- Olabilirlik oran testi
- Skor testidir

$\beta$ 'ların anlamlılığını test etmede kullanılan wald test istatistiği, t veya z istatistiğinin genelleştirilmiş şekli olup, en çok olabilirlik (ML) tahmini ve hipotezlenen değer arasındaki farkın bir fonksiyonudur. Bu fark, en çok olabilirlik tahmininin standart hatasının tahmini değeri ile normalleştirilmiştir (Theureau ve

Grambsch, 2000.).  $\hat{\beta}$ ' nin sıfır olmayan standart hatası (SE) ile birlikte test istatistiği şu şekilde hesaplanır:  $z = (\hat{\beta} - \beta_0) / SE$  Bu test istatistiği,  $\beta = 0$  olduğu zaman standart normal dağılıma yaklaşık bir dağılım gösterir. z değerleri için bir ya da iki yanlı P-değerlerini içeren standart normal dağılım tablosu kullanılır. Eşit bir şekilde iki yanlı dönüşüm için,  $z^2$  değerleri 1 serbestlik dereceli (df) sıfırdan farklı bir ki-kare dağılımı göstermektedir. P-değeri, gözlemlenmiş değerler üzerinde, sağ kuyruktaki ki-kare olasılığıdır. Sıfırdan farklı standart hata kullanımını içeren bu türdeki istatistiğe *Wald İstatistiği* adı verilir (Garson, 1999). Wald istatistiği değişkenlere ilişkin bir değerlendirme ölçütüdür. Wald' ın 2' den büyük değerler için önemli olduğu kabul edilmektedir (Wald istatistiğinin hesaplandığı değişkenin anlamlı olduğunu göstermektedir; özellikle lojistik regresyon modelinde değişkenlerin anlamlılığının araştırılmasında tercih edilir) ve bununla beraber Wald değeri büyüdükçe olasılık değeri (p) küçülür (Bircan ve diğerleri, 2004).  $H_0 : \beta = \beta_0$ ' ın Wald Testi çok değişkenli durumlar için genişletilir ve şu şekilde bir test istatistiği kullanılır:

$$W = (\hat{\beta} - \beta_0)' [\text{cov}(\hat{\beta})]^{-1} (\hat{\beta} - \beta_0)$$

(İlk terim bir vektör ya da matrisin transpozisini ifade etmektedir.)

Sıfırdan farklı kovaryans ise daha önce ele alınan,  $\hat{\beta}$  için log olabilirliğin eğimine  $(-E\left(\frac{\partial^2 L(\beta)}{\partial \beta_j \partial \beta_k}\right))$  dayandırılmıştır.  $\hat{\beta}$  için asimptotik çok değişkenli normal dağılım, W için asimptotik ki-kare dağılımı anlamına gelmektedir.

Serbestlik derecesi (df) ,  $\text{cov}(\hat{\beta})$ ' nin rankına ( $\beta$  matrisinde gereksiz olmayan parametre sayısını veren) eşittir. Genel amaçlı ikinci metot ise iki maksimum oran arasındaki olabilirlik fonksiyonunu kullanır:

1.  $H_0$  altında olası parametre değerleri üzerinde maksimumluk.
2.  $H_0$  veya  $H_a$  alternatif hipotezlerinden doğru olanın kabul edilmesini sağlayan daha büyük bir parametre değerleri kümesinin üzerinde maksimumluk.

$\ell_0$ ;  $H_0$  altında olabilirlik fonksiyonunun maksimize edilmiş değerini gösterebilir,  $\ell_1$ ; genel olarak maksimize değerini gösterir (yani  $H_0 \cup H_1$  altındaki maksimize değeri). Örnek olarak parametre vektörü  $\beta = (\beta_0, \beta_1)'$  ve  $H_0 : \beta_0 = 0$  için  $\ell_1$ , en yüksek olasılığa sahip olabilecek verilerden oluşan  $\beta$  değerine göre hesaplanmış olabilirlik fonksiyonudur ve  $\ell_0$ ;  $\beta_0 = 0$  iken en yüksek olasılığa sahip olabilecek verilerden oluşan  $\beta_1$  değerine göre hesaplanmış olabilirlik fonksiyonudur.  $\ell_0$ , parametre değerlerinin sınırlanmış bir kümesi üzerinde maksimizasyondan meydana geldiğinden dolayı  $\ell_1$  her zaman en azından  $\ell_0$  kadar büyüktür (Agresti,2002).

$\Lambda = \ell_0 / \ell_1$ ; maksimize edilmiş olabilirliklerin oranını göstermekte ve 1' i aşmamaktadır. Wilks (1935,1938) ;  $n \rightarrow \infty$  giderken  $-2 \log \Lambda$  değerinin sıfır ile sınırlandırılmış bir ki-kare dağılımına sahip olduğunu göstermiştir. Serbestlik derecesi  $H_0$  ve  $H_0 \cup H_1$  altındaki parametre uzaylarının çeşitli boyutları için farklıdır. Olabilirlik-oran test istatistiği aşağıdaki gibidir:

$$-2 \log \Lambda = -2 \log(\ell_0 / \ell_1) = -2(L_0 - L_1)$$

$L_0, L_1$ ; maksimize edilmiş log-olabilirlik fonksiyonlarını ifade eder. Çok sayıda örnek için olabilirlik oran istatistiği, tahmin edilecek parametre sayısına eşit serbestlik derecesi ile ki-kare dağılımı gösterir (Yay, Çoker ve Uysal,2007:139).

Üçüncü metotta ise R.A. Fisher ve C.R. Rao' ya göre "skor (puan) istatistiği" kullanılmaktadır. Skor testi;  $\beta_0$  için  $L(\beta)$  log-olabilirlik fonksiyonun beklenen eğriliğine ve eğimine dayanmaktadır(Agresti,2002).  $\beta_0$ ' a göre değerlendirilen

$u(\beta) = \partial L(\beta) / \partial \beta$  şeklindeki skor (puan) fonksiyonu kullanılır.  $u(\beta_0)$  değeri ;  $\hat{\beta}$ ,  $\beta_0$ ’ dan uzaklaştıkça mutlak değerde daha büyük olma eğilimindedir.  $-E(\partial^2 L(\beta) / \partial \beta^2)$  (yani bilgi);  $u(\beta_0)$  aracılığıyla  $\beta_0$ ’ a göre değerlendirildiğini göstermektedir. Skor (puan) istatistiği, sıfıra yakın önemsenmeyecek kadar küçük “SE” ( $[u(\beta_0)]^{1/2}$ ) ile  $u(\beta_0)$  oranlanmasına dayanmaktadır. Bu yaklaşık bir standart normal dağılım göstermektedir. Skor (puan) istatistiğinin ki-kare yapısı ise şu şekildedir:

$$\frac{[u(\beta_0)]^2}{\iota(\beta_0)} = \frac{[\partial L(\beta) / \partial \beta_0]^2}{-E[\partial^2 L(\beta) / \partial \beta_0^2]}^6$$

Kısmi türev notasyonu;  $\beta_0$ ’a göre değerlendirilmiş  $\beta$ ’ ya göre türevleri ifade etmektedir. Skor testi, birden fazla değişken ve sürekli değişkenler için log-sıra testinin genelleştirilmiş şeklidir. Skor test istatistiği, serbestlik derecesi tahmin edilen parametre sayısı olmak üzere dağılımına uyar (Özdamar, 1999:475) S için hesaplanan değer ki-kare tablo değeri ile karşılaştırılarak istatistiğin anlamlılığı test edilir.

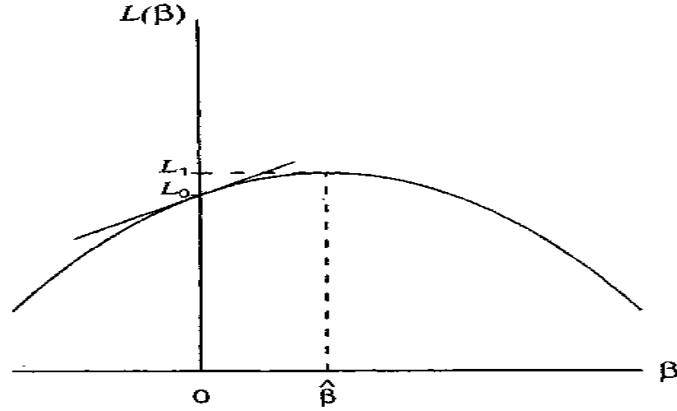
Çok parametrelili durumda skor istatistiği,  $\beta$  ile ilgili log-olabilirliğin kısmi türevlerinin vektörüne ve bilgi matrisinin tersine<sup>7</sup> bağlı olan kuadratik bir formdadır. Şekil-1.1, tek değişkenli bir durum için log-olabilirliğinin ( $L(\beta)$ ) genel bir grafiğidir. Bu grafik  $H_0 : \beta = 0$ ’ ın üç testini göstermektedir Wald testi ki-kare formuna ( $(\hat{\beta} / SE)^2$ ) sahip ML tahmini olan  $\hat{\beta}$ ’ ya göre  $L(\beta)$ ’ den faydalanır.  $\hat{\beta}$ ’ nın SE’ i;  $\hat{\beta}$ ’ya göre  $L(\beta)$ ’ nin eğriliğine bağlıdır. Skor testi ise  $\beta = 0$ ’ a göre

<sup>6</sup> İki değerli değişkenin gözlenme oranı “p” olsun. Bu orana göre ilgilenilen değişkenin r kez gözlenme olasılığını kullanarak maksimum olabilirlik fonksiyonu:  $L = \log(l) = r \log p + (n - r) \log (1 - p)$  şeklinde yazılır. Skor test istatistiğini elde etmek için, L’nin ‘p’ ye göre alınan birinci ve ikinci türevleri sıfıra eşitlenerek çözülür ve :  $S = (L')^2 / (L'')$  şeklindeki oran elde edilir (Yay, Çoker ve Uysal, 2007: 140).

<sup>7</sup> Her ikisi de  $H_0$  tahminlerine yani  $\beta = \beta_0$  varsayımına göre değerlendirilmiştir.

$L(\beta)$ 'nin eğriliğine ve eğimine bağlıdır. Olabilirlik oran testi  $\hat{\beta}$  ve  $\beta_0 = 0$  değerlerinin her ikisine göre  $L(\beta)$  ile ilgili bilgileri birleştirir. Bu,  $\hat{\beta}$ 'ya göre belirlenen  $L_1$  ve  $\beta_0 = 0$ 'a göre  $L_0$  şeklindeki log-olabilirlik değerlerini ki-kare istatistiğini,  $-2(L_0 - L_1)$ , kullanarak karşılaştırmaktadır. Şekil-1.1' de bu istatistik iki defa hesaplanmıştır ve  $\hat{\beta}$  ile 0'a göre ayrı ayrı hesaplanan  $L(\beta)$  değerleri arasındaki dikey uzaklık gösterilmiştir. Bir anlamda bu istatistik üç test istatistiğinin bilgisinin çoğunu kullanmaktadır ve çok yönlüdür.  $n \rightarrow \infty$  için, Wald, olabilirlik oranı ve skor testleri belirli asimptotik eşitliklere sahiptir. Orta büyüklükteki örneklem için olabilirlik oran testi Wald testinden daha güvenilirdir (Agresti, 2002).

**Şekil-1.1**  $H_0 : \beta = 0$ , in Sınanmasında Üç Test Çeşidi İçin Kullanılan log-Olabilirlik Fonksiyonu ve Bilgisi.



Skor metodu  $\hat{\beta}$ 'yi kullanmamaktadır. Bu nedenle,  $\beta$  modelin parametresi ise, uygun bir model olmaksızın  $H_0 : \beta = \beta_0$ 'in test edilmesi için skor istatistiği hesaplanmalıdır. Açıklayıcı bir analiz için uygun birkaç model oluşturmak ve model uygunluğu hesaplamak oldukça yoğun ve karışıkta bu yararlıdır. Skor ve olabilirlik oran metodlarının bir avantajı da  $|\hat{\beta}| = \infty$  olduğunda bile uygulanabilir olmasıdır. Böyle bir durumda, Wald istatistiği hesaplanamaz. Wald metodunun bir diğer dezavantajı ise; sonuçlarının parametrelere bağlı olması ve çıkarsamanın da  $\hat{\beta}$ 'ya dayanmasıdır (Agresti, 2002).



### 1.3.4. Güven Aralıklarının Oluşturulması

Uygulamada parametre değerleri ile ilgili hipotezlerin test edilmesinden çok parametrelere ilişkin güven aralıklarının oluşturulması daha fazla bilgi verici niteliktedir.

Üç test metodundan herhangi biri için bir güven aralığı testin tersine çevrilmesinden meydana gelmektedir. Örnek olarak,  $H_0 : \beta = \beta_0$  'ın test edilmesi amacıyla;  $\beta$  için %95 güven aralığı  $\beta_0$  'ın kümesidir ve 0.05'i aşan bir P-değerine sahiptir.  $z_a$  ; sağ kuyruk olasılığı a olan bir standart normal dağılımdan elde edilen z-Skor (puan)larını göstermektedir ki bu dağılımın  $100(1-a)$  şeklinde yüzde birliğini oluşturmaktadır.  $\chi^2_{df}(a)$  ise "df" serbestlik dereceli ki-kare dağılımının yüzde birliğidir,  $100(1-a)$ . Asimptotik normalliğe bağlı olarak  $100(1-\alpha)\%$  güven aralıkları,  $z_{\alpha/2}$  değerlerini kullanır. Örnek olarak, %95 güven düzeyi için  $z_{0.025} = 1,96$  olur.

**Wald güven aralığı** ise  $\frac{|\hat{\beta} - \beta_0|}{SE} < z_{\alpha/2}$  şeklinde  $\beta_0$  'ın kümesidir. Bu ,  $\hat{\beta} \pm z_{\alpha/2}(SE)$  aralığını verir. **Olabilirlik oranına bağlı güven aralığı** ise  $-2[L(\beta_0) - L(\hat{\beta})] < \chi^2_1(\alpha)$  şeklinde  $\beta_0$  'ın kümesidir ( $\chi^2_1(\alpha) = z^2_{\alpha/2}$  olduğu hatırlanmalıdır.).  $\hat{\beta}$  normal dağılımlı olduğu zaman, log-olabilirlik fonksiyonu parabolik bir şekle sahip olur (yani ikinci dereceden polinom).

Kategorik veriler için küçük örneklerde,  $\hat{\beta}$  normallikten uzak olabilir ve log-olabilirlik fonksiyonu da parabolik şekilli simetrik bir eğriden oldukça uzaklaşmış olur (Agresti, 2002). Bir model parametrelerin çoğunu kapsıyorsa, bu durum büyük örneklerde daha ılımlı hale gelir. Böyle durumlarda,  $\hat{\beta}$  'nın asimptotik normalliğine bağlı olarak yapılan yorumlama yetersiz bir performansa sahip olabilmektedir. Wald ve olabilirlik oran sonuçlarındaki yorumlamada belirgin

bir uzaklaşma,  $\hat{\beta}$ ' nin dağılımının normale yakın olmadığını gösterebilir. Bu şekildeki birçok durumda, yorumlama yerine kesin olasılık küçük örneklem dağılımı ya da “daha yüksek sıra” şeklindeki asimptotik metotlar kullanılır. Bu metotlar basit normallik ile ilerlemektedir.

Wald güven aralığı uygulamada en yaygın olanıdır. Çünkü istatistiksel yazılımlar aracılığı ile ML tahminleri ve standart hatalar kullanılarak kurulması oldukça basittir. Olabilirlik oran temelli aralık ise yazılımlarda çok daha yaygın olarak bulunmaktadır ve küçük örneklemler için daha fazla tercih edilmektedir. En iyi bilinen istatistiksel modelde, (normal bir yanıt için regresyon) yorumlamanın üç türünde de ister istemez aynı sonuçlar elde edilir.

#### 1.4. BİNOM PARAMETRELERİ İÇİN İSTATİSTİKSEL YORUMLAMA

Bu bölümde; n tane birbirinden bağımsız denemede y başarı sayısını veren  $\pi$  binom parametresi için güven aralıkları ve testlerin gösterilmesi aracılığı ile kategorik verilere ilişkin yorumlama metotları açıklanacaktır. Daha önceki kısımlarda  $\pi$ ' nin ML tahmin edicisini ( $\hat{\pi} = y/n$ ) ve olabilirlik fonksiyonunu ele almıştık.

##### 1.4.1. Bir Binom Parametresi İle İlgili Testler

$H_0 : \pi = \pi_0$  hipotezini dikkate alalım.  $H_0$  tek bir parametreye sahip olduğundan dolayı Wald ve skor test istatistiklerinin ki-kare yerine normal formları kullanılmaktadır. Normal formlar çift yanlı alternatifler kadar tek yanlıların testlerine de fırsat vermektedir. Wald istatistiği;

$$z_w = \frac{\hat{\pi} - \pi_0}{SE} = \frac{\hat{\pi} - \pi_0}{\sqrt{\hat{\pi}(1 - \hat{\pi})/n}}$$

$\pi_0$  için  $-E[\partial^2 L(\pi)/\partial \pi^2] = E[y/\pi^2 + (n-y)/(1-\pi)^2] = n/[\pi(1-\pi)]$  eşitliğindeki bilgi ve  $\partial L(\pi)/\partial \pi = y/\pi - (n-y)/(1-\pi) = (y-n\pi)/\pi(1-\pi)$  şeklindeki binom skor ölçümleri aşağıdaki eşitliklerle sağlanır:

$$u(\pi_0) = \frac{y}{\pi_0} - \frac{n-y}{1-\pi_0}, \quad \iota(\pi_0) = \frac{n}{\pi_0(1-\pi_0)}$$

**Skor istatistiğinin** normal formu aşağıdaki şekilde sadeleştirilir:

$$z_s = \frac{u(\pi_0)}{[\iota(\pi_0)]^{1/2}} = \frac{y-n\pi_0}{\sqrt{n\pi_0(1-\pi_0)}} = \frac{\hat{\pi} - \pi_0}{\sqrt{\pi_0(1-\pi_0)/n}}$$

Wald istatistiği,  $z_w$ ;  $\hat{\pi}$  için hesaplanmış standart hatayı kullanırken, skor istatistiği  $z_s$ ;  $\pi_0$  için hesaplanmış standart hatayı kullanır. Skor istatistiği daha iyidir. Çünkü bir tahmindense etkisiz gerçek SE' yi kullanır (Agresti, 2002). Onun etkisiz örneklem dağılımı Wald istatistiğine göre standart normal dağılıma daha yakındır. **İkiterimli (binom) log-olabilirlik fonksiyonu**<sup>8</sup>:

$$H_0 \text{ altında: } L_0 = y \log(\pi_0) + (n-y) \log(1-\pi_0)$$

Daha genel olarak şu şekilde yazılır:

$$L_1 = y \log(\hat{\pi}) + (n-y) \log(1-\hat{\pi})$$

**Olabilirlik oran test istatistiği** ise aşağıdaki şekilde basitleştirilir:

---

<sup>8</sup>  $L(\pi) = \log[\pi^y (1-\pi)^{n-y}] = y \log(\pi) + (n-y) \log(1-\pi)$  daha önce verilmişti.

$$-2(L_0 - L_1) = 2 \left( y \log \frac{\hat{\pi}}{\pi_0} + (n - y) \log \frac{1 - \hat{\pi}}{1 - \pi_0} \right)$$

$$-2(L_0 - L_1) = 2 \left( y \log \frac{y}{n\pi_0} + (n - y) \log \frac{n - y}{n - n\pi_0} \right)$$

Gözlemlenen başarı ve başarısızlık miktarları ile uydurulmuş miktarları aşağıdaki eşitlik aracılığıyla kıyaslanır:

$$2 \sum \text{gözlemlenen} \log \frac{\text{gözlemlenen}}{\text{uygun}}$$

Bu formül Poisson ve çokterimli parametreleri ile ilgili testlerde de kullanılmaktadır.  $H_0$  altında bilinmeyen parametreler meydana gelmezse ve biri  $H_a$  altında oluşursa bu formül  $df=1$  serbestlik derecesine sahip asimptotik bir ki-kare dağılımına sahip olur.

#### 1.4.2. Binom Parametresi için Güven Aralıkları

Bir anlamlılık testi sadece belirli bir  $\pi$  değerinin ( $\pi = 0.5$  gibi) olası olup olmadığını göstermektedir. Olası değerlerin dağılımını (alanını) belirleyen bir güven aralığı kullanılarak daha fazla bilgi edinilebilir. **Wald test istatistiğini**<sup>9</sup> ters çevirerek  $\pi_0$  'ın aralığı bulunur:  $|z_w| < z_{\alpha/2}$  ya da  $\hat{\pi} \pm z_{\alpha/2} \sqrt{\hat{\pi}(1 - \hat{\pi})/n}$ . Tarihsel olarak bu aralık herhangi bir parametre için kullanılmış ilk güven aralıklarından biridir (Agresti, 2002). Ne yazık ki “n” yeteri kadar büyük olmadıkça yetersiz bir uygulama olur. Gerçek kapsama olasılığı nominal güven katsayısının altına düşer ve  $\pi$ , 0 ya da 1'e yaklaştıkça daha da aşağısında kalır. Bu formülü kullanmadan önce her bir örneklem sınıfına  $1/2z^2_{\alpha/2}$  kadar gözlem eklenerek basit bir düzeltme yapılır ve uygulama daha iyi hale gelir.

---


$$^9 z_w = \frac{\hat{\pi} - \pi_0}{SE} = \frac{\hat{\pi} - \pi_0}{\sqrt{\hat{\pi}(1 - \hat{\pi})/n}}$$

**Skor güven aralığı** ise  $\pi_0$  değerlerini içine almaktadır,  $z_s < z_{\alpha/2}$ . Skor güven aralığının uç noktaları;  $(\hat{\pi} - \pi_0) / \sqrt{\pi_0(1 - \pi_0) / n} = \pm z_{\alpha/2}$  şeklindeki eşitlik ile sağlanan  $\pi_0$  çözümleridir ve bu çözümler kuadrattır. İlk olarak E.B. Wilson (1927) ele almıştır, bu aralık aşağıdaki gibidir:

$$\hat{\pi} \left( \frac{n}{n + z^2_{\alpha/2}} \right) + \frac{1}{2} \left( \frac{z^2_{\alpha/2}}{n + z^2_{\alpha/2}} \right) \pm z_{\alpha/2} \sqrt{\frac{1}{n + z^2_{\alpha/2}} \left[ \hat{\pi}(1 - \hat{\pi}) \left( \frac{n}{n + z^2_{\alpha/2}} \right) + \left( \frac{1}{2} \right) \left( \frac{1}{2} \right) \left( \frac{z^2_{\alpha/2}}{n + z^2_{\alpha/2}} \right) \right]}$$

Aralığın orta noktası ( $\tilde{\pi}$ )  $\hat{\pi}$ 'nin ağırlıklı bir ortalamasıdır ve  $n/(n + z^2_{\alpha/2})$  ağırlığı  $1/2$  olur (n arttıkça  $\hat{\pi}$ ' de artar.). Terimler birleştirilirse, orta nokta  $\tilde{\pi} = (y + z^2_{\alpha/2} / 2) / (n + z^2_{\alpha/2})$  şeklindeki denkleme eşit olur. Bu düzeltilmiş (her bir örneklem sınıfının yarısı kadar olan  $z^2_{\alpha/2}$  kadar gözlem eklenerek ) bir örneklem için örneklem oranıdır. Bu formüldeki  $z_{\alpha/2}$ 'li katsayının karesi;  $\pi = \hat{\pi}$  ve  $\pi = 1/2$  durumundaki örneklem oranının ağırlıklandırılmış ortalamasıdır (n yerine  $(n + z^2_{\alpha/2})$  şeklindeki düzeltilmiş/ayarlanmış örneklem büyüklüğü kullanılmıştır.). Bu aralık Wald aralığından daha iyi bir performansa sahiptir (Agresti, 2002). **Olabilirlik oranına dayanan güven aralığı** hesaplama açısından daha karmaşık olmasına rağmen ilke olarak daha basittir.  $\alpha$  değerini aşan P-değerine sahip olabilirlik oran testine göre  $\pi_0$ ' in kümesidir.

### 1.4.3. Vejetaryenlerin Oranı İle İlgili Örnek

Giriş niteliğindeki tanıtıcı istatistik kursunda verileri toplamak için, Alan Agresti öğrencilerine bir anket vermiş ve ilk olarak kız ya da erkek öğrencilere bir vejetaryen olup olmadıklarını sormuştur. n=25 öğrenci olmak üzere evet cevabı için y=0 değerini almaktadır.

Bu belirli bir populyasyondan yapılan rasgele bir örneklem değildir. Fakat bu veriler; binom parametresi  $\pi$  için %95 güven aralıklarının örnekle açıklanması

amacıyla kullanılacaktır.  $y = 0$  olduğundan  $\hat{\pi} = 0/25 = 0$  olur.  $\pi$  için %95 güven aralığı için Wald yaklaşımı kullanılırsa;  $0 \pm 1.96\sqrt{(0.0) \times 1.0/25}$ , yada  $(0,0)$  şeklinde bir aralık elde edilir. Örneklem uzayının sınırlarına düşen gözlemler olduğunda, genellikle Wald metotları mantıklı cevaplar sağlanamaz. Çelişkili olarak; %95 **skor güven aralığı** ise  $(0.0,0.133)$  şeklindedir. Bu daha ikna edici bir çıkarımadır.  $H_0 : \pi = 0.5$  için skor istatistiği,  $z_s = (0 - 0.5) / \sqrt{(0.5 \times 0.5)/25} = -5.0$  şeklinde bulunur ki 0.5 değerinin aralığın içine düşmediği görülmektedir. Bu durumun aksine  $H_0 : \pi = 0.10$  için:  $z_s = (0 - 0.10) / \sqrt{(0.10 \times 0.10)/25} = -1.67$  olur ve böylece 0.10 değerinin bu aralığın içine düştüğü görülür.  $y=0$  ve  $n=25$  olduğunda, olabilirlik **fonksiyonunun çekirdeği (kernel)**:  $\ell(\pi) = \pi^y (1 - \pi)^{n-y} = (1 - \pi)^{25}$  şeklindedir.

$L(\pi) = \log[\pi^y (1 - \pi)^{n-y}] = y \log(\pi) + (n - y) \log(1 - \pi)$  şeklinde ifade edilen **log-olabilirlik fonksiyonu** ise  $L(\pi) = 25 \log(1 - \pi)$  şeklinde olur.  $L(\hat{\pi}) = L(0) = 0$  olduğuna dikkat edilmelidir. %95 olabilirlik oran güven aralığı, aşağıda gösterilen olabilirlik oran istatistiğine göre  $\pi_0$ ' in kümesidir:

$$-2(L_0 - L_1) = -2[L(\pi_0) - L(\hat{\pi})] = -50 \log(1 - \pi_0) \leq \chi_1^2(0.05) = 3.84$$

Üst sınır;  $1 - \exp(-3.84/50) = 0.074$  ve güven aralığı  $(0.0, 0.074)$  şeklindedir. [Burada doğal logaritma kullanılmaktadır, bu yüzden tersi (devriği) üssel bir fonksiyondur,  $\exp(x) = e^x$ ]. Üç büyük örneklem metodu ile oldukça farklı sonuçlar elde edilir.  $\pi$ , 0 değerine yakın olduğunda,  $\hat{\pi}$ ' in örneklem dağılımı son derece sağa çarpıktır (Küçük n için). Bu asimptotik yaklaşımları gerektirmeyen alternatif metotların dikkate alınmasının uygun olacağını göstermektedir.

#### 1.4.4. Kesin Olasılık Küçük Örneklem Yorumlaması

Modern hesaplama yöntemlerinin gücü ile birlikte,  $\hat{\pi}$  şeklindeki istatistiklerin dağılımlarının büyük örneklem yaklaşımlarına dayanması gerekliliği ortadan kalkmıştır. Testler ve güven aralıklarında normal yaklaşımdan daha çok

doğrudan binom dağılımı kullanılmaktadır. Bu şekildeki yorumlamalar küçük örnekler için doğal bir şekilde gerçekleşmektedir. Modern hesaplama yöntemleri olduğu için süreklilik düzeltmesi yapılarak sağlanan yakınsama yerine doğrudan uygulama yapılır. Vejetaryenlikle ilgili araştırma sonuçları için ( $y=0$  ve  $n=25$ );  $H_a : \pi \neq 0.5$  alternatif hipotezine karşı,  $H_0 : \pi = 0.5$  sıfır hipotezinin test edilmesi aracılığı ile durumu açıklayalım. Skor istatistiği  $z=-5.0$  şeklinde dikkate alındı. Bu istatistik için kesin P-değeri  $\text{bin}(25,0.5)$  şeklindeki dağılıma ve sıfır hipotezine dayanmaktadır ve aşağıdaki şekilde bulunur:

$$P(|z| \geq 5.0) = P(Y = 0 \text{ yada } Y = 25) = 0.5^{25} + 0.5^{25} = 0.00000006$$

%100(1- $\alpha$ ) güven aralığı, kesin olasılık binom testlerinde  $\alpha$  değerini aşan P-değerlerine göre bütün  $\pi_0$  değerlerinden oluşmaktadır. En iyi bilinen aralıkta (Clopper ve Pearson 1934), güven aralıklarının oluşturulmasında kuyruk metodu kullanılır. Bu  $\alpha/2$ 'yi aşan tek yanlı P-değerlerinin her birini gerektirmektedir. Üst ve alt sınırlardaki uç noktalar aşağıdaki eşitliklerde  $\pi_0$  için yapılacak çözümlerle bulunur:

$$\sum_{k=y}^n \binom{n}{k} \pi_0^k (1 - \pi_0)^{n-k} = \alpha/2 \text{ ve } \sum_{k=0}^y \binom{n}{k} \pi_0^k (1 - \pi_0)^{n-k} = \alpha/2$$

$y=n$  olduğunda üst sınırın 1 olması ve  $y=0$  olduğunda alt sınırın 0 olması durumu haricinde bu çözümler dikkate alınır.

$y=1,2,\dots,n-1$  olduğunda binom toplamları ve tamamlanmamış beta fonksiyonu arasındaki ilişkilerden ve F ile beta dağılımlarının birikimli dağılım fonksiyonlarından yararlanılarak güven aralıkları aşağıdaki ifadeye eşittir:

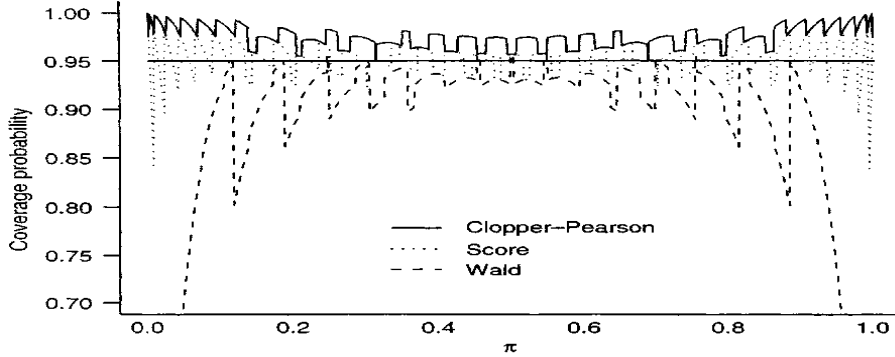
$$\left[ 1 + \frac{n-y+1}{y F_{2y, 2(n-y+1)}(1-\alpha/2)} \right]^{-1} < \pi < \left[ 1 + \frac{n-y}{(y+1) F_{2(y+1), 2(n-y)}(\alpha/2)} \right]^{-1}$$

$F_{a,b}(c)$ ; a ve b serbestlik derecesi ile F dağılımının 1-c sıklık derecesini göstermektedir. n=25 için y=0 olduğunda,  $\pi$  için Clopper-Pearson %95 güven aralığı ; (0.0 , 0.137) şeklinde bulunur. İlke olarak bu yaklaşım ideal gözükmektedir. Fakat ciddi bir karışıklık söz konusudur. Kesiklilikten dolayı, herhangi bir  $\pi$  için gerçek olasılık kapsamı en azından nominal güven aralığı kadar büyüktür (Agresti, 2002) hatta daha büyükte olabilmektedir.

Benzer şekilde  $H_0 : \pi = \pi_0$  hipotezinin, 0.05 gibi istenilen sabit bir  $\alpha$  büyüklüğünde test edilmesinde, o büyüklüğe ulaşmak genellikle olası değildir. Olası örneklem sayısı sınırlıdır ve bu yüzden olası P-değerlerinin sayısı da sınırlı olacaktır ve 0.05 bu değerlerden biri olmayabilir. Sabit bir  $\pi_0$  ile  $H_0$  ' ın test edilmesinde bir P-değeri şeklinde belirli bir  $\alpha$  değeri seçilmelidir. Fakat aralık tahmininde, bu tercih edilmez. Çünkü ;  $H_0 : \pi = \pi_0$  için  $\pi_0$  değerlerinden oluşan tam bir bölgenin ters çevrilmesine karşılık gelen aralık oluşturulmaktadır ve her farklı  $\pi_0$  değeri kendi olası P-değerlerinin kümesine sahiptir; bu da bir teste tek bir anlamsız  $\pi_0$  değerinin olmadığını göstermektedir. Herhangi belirli (sabit) bir parametre değeri için gerçek kapsama olasılığı nominal güven düzeyinden daha büyük olabilir. n=25 olduğunda Şekil-1.2; Clopper-Pearson Metodu, Skor Metodu ve Wald Metodu için  $\pi$  ' nin bir fonksiyonu olarak kapsama olasılıklarının grafiklerini göstermektedir.



**Şekil-1.2**  $n=25$  olduğunda  $\pi$  Binom Parametresi İçin Nominal %95 Güven Aralıklarına Göre Kapsama Olasılıklarının Grafiği



Belirli bir metot ile sabit bir  $\pi$  değerinde, kapsama olasılığı;  $\pi$  değerini içine alan aralık sonucunu veren tüm örneklemelerin ikiterimli(binom) olasılıklarının toplamıdır. Burada 26 olası örneklem ve 26 eş güven aralığı vardır. Bu nedenle, 0 ve 26 binom olasılıklarının arasındaki bir toplamı vermektedir.  $\pi$  0' dan 1' e doğru ilerlerse, bu kapsama olasılığı ;  $\pi$  bu aralıkların içine ya da dışına doğru hareket ettiği zaman yukarı ya da aşağı doğru atlar. Şekil-1.3' te; Wald metodu için kapsama olasılıklarının aşırı düşük olduğu, buna rağmen Clopper-Pearson metodu için zıt yönde bir hata söz konusu olduğu görülmektedir. Yine şekle bakıldığında Skor metodu, 0 veya 1'e yakın bazı  $\pi$  değerleri haricinde çok iyi hareket etmektedir. Skor metodunun kapsama olasılıkları nominal seviyeye yakın olma eğilimindedir, sürekli tutarlı bir biçimde ölçülü ya da serbest değildir.  $\pi$ , 0' a ya da 1' e çok yakın olmadıkça skor metodu iyi bir metottur. Küçük örneklem dağılımlarını kullanan kesikli problemlerde, daha kısa güven aralıkları genellikle; iki tek yanlı testten daha çok, bir tane iki yanlı testin ters çevrilmesinden ileri gelmektedir. Öyleyse aralık; iki yanlı bir testin  $\alpha$  değerini aşan P-değerlerine göre parametre değerlerinin kümesidir.

#### 1.4.5. Orta P-Değerine Dayanan Yorumlama

Küçük örneklem dağılımlarında kesikliliğin değerlendirilmesi orta p-değerine göre yapılan yorumlamaya dayanmaktadır (Agresti,2002). Gözlemlenen değer  $t_0$  ve  $H_a$  tek yanlı alternatif hipotezi ile T gibi bir test istatistiğine göre(T büyükse  $H_0$ ' ın

aksi iddia edilir.) sıfır dağılımından hesaplanmış olasılıklarla orta P-değeri aşağıdaki şekildedir:

$$orta - P - de\ddot{g}. = \frac{1}{2} P(T = t_0) + P(T > t_0)$$

Bu nedenle, orta P-değeri gözlenen sonucun olasılığının yarısı aracılığı ile elde edilen sıradan P-değerinden daha düşük bir değerdedir. Sıradan P-değeri ile karşılaştırıldığında; orta P-değeri daha çok sürekli bir dağılıma sahip bir test istatistiği için kullanılan P-değeri gibi hareket etmektedir. Tek yanlı iki tane P-değerinin toplamı 1.0'a eşittir. Kesikli olmasına rağmen,  $H_0$  altında orta-P-değerinin dağılımı daha çok, sürekli durumlar için oluşturulan tekdüze (Uniform) dağılıma benzemektedir. Örnek olarak; orta-P-değerinin 0.5'in sıfır hipotezine göre beklenen bir değeri vardır, oysa bu beklenen değer kesikli bir test istatistiği için oluşturulan sıradan P-değerine göre 0.5' i aşmaktadır.

Sıradan P-değeri ile bir kesinlik (doğru) testinin aksine; orta P-değeri kullanılarak yapılan bir test, birinci tip hatanın olasılığının nominal bir değerden daha büyük olmayacağını garanti edemez. Fakat bu biraz ılımlı bir şekilde genellikle iyi işlemektedir. Sıradan kesinlik testine kıyasla daha az ölçülü ve ılımlıdır (daha az risk içermektedir.). Benzer şekilde orta P-değeri ile doğru dağılım kullanılarak yapılan testler ters çevrilerek oluşturulan güven aralıkları da aynı özelliğe sahiptir ( yani %95 güven aralıkları; orta P değerinin 0.05' i aşmasına göre parametre değerlerinin bir kümesini göstermektedir.).

Vejeteryanların oranı hakkındaki örnekte;  $H_a : \pi \neq 0.5$  alternatif hipotezine karşı  $H_0 : \pi = 0.5$  sıfır hipotezinin test edilmesinde, n=25 ve y=0 olmak üzere; gözlemlenen sonuç son derece olasıdır. Bu yüzden orta-P-değeri sıradan P-değerinin yarısıdır ya da 0.00000003 olur. orta P-değeri ile binom kesinlik testinin Clopper-Pearson' a göre alınan tersini kullanılmasıyla  $\pi$  için % 95 güven aralığı (0.000, 0.113) şeklinde elde edilir. Sıradan Clopper-Pearson aralığına göre bu (0.000, 0.137) şeklinde bulunur.

Orta-P-değeri kesiklilik problemlerinin ortadan kaldırılması için gereksiz rasgele seçimlerin uygulanması ile aşırı derecede ılımlı (ölçülü) çıkarsamaların yapılması arasında ara bulucu, uzlaştırıcı bir hassasiyete sahip gözükmektedir. Çok önemli kesikli dağılımlar ile ilgili güven aralıklarının oluşturulması ve testlerin yapılması konusunda orta-P-değerinin kullanılması önerilmektedir.

## 1.5.ÇOK TERİMLİ PARAMETRELER İÇİN İSTATİSTİKSEL YORUMLAMA

Bu kısımda  $\{\pi_j\}$  şeklinde ifade edilen çokterimli parametreler için çıkarsama metotları tanıtılacaktır. n gözlemede,  $n_j$ ; n. kategorideki gözlemi ifade etmektedir ve  $j=1,\dots,c$  şeklindedir. İlk olarak;  $\{\pi_j\}$  için ML tahminleri elde edilir.  $\{\pi_j\}$ ' nin bir fonksiyonu şeklinde çokterimli olasılık yoğunluk fonksiyonunun çekirdek (kernel) kısmı ile orantılıdır. Çokterimli olasılık yoğunluk fonksiyonu aşağıdaki şekildedir:

$$p(n_1, n_2, \dots, n_{c-1}) = \left( \frac{n!}{n_1! n_2! \dots n_c!} \right) \pi_1^{n_1} \pi_2^{n_2} \dots \pi_c^{n_c} \text{ }^{10}$$

Çekirdeği:

$$\prod_j \pi_j^{n_j} \text{ where all } \pi_j \geq 0 \text{ and } \sum_j \pi_j = 1$$

Bu eşitliği en yükseğe çıkararak (maksimize eden)  $\{\pi_j\}$  değerleri ML tahminleridir. Çokterimli log-olabilirlik fonksiyonu ise;  $L(\pi) = \sum_j n_j \log \pi_j$  şeklindedir. Gereksiz oluşumları ortadan kaldırmak amacıyla; L'yi,  $\{\pi_1, \dots, \pi_{c-1}\}$ ' in

---

<sup>10</sup> Hatırlatma:  $\sum_j n_j = n$  olduğu için, bu  $n_c = n - (n_1 + \dots + n_{c-1})$  ile (c-1) boyutludur. c=2 için çokterimli dağılımın özel bir durumudur ve sonuç binom dağılımıdır. Çokterimli dağılım için;  $E(n_j) = n\pi_j$ ,  $\text{var}(n_j) = n\pi_j(1 - \pi_j)$ ,  $\text{cov}(n_j, n_k) = -n\pi_j\pi_k$  şeklindedir.

bir fonksiyonu olarak ele alırız<sup>11</sup>. Böylece  $L(\pi)$ ' nin  $\pi_j$ ' ye göre türevi alınırsa olabilirlik eşitliği elde edilir ve sıfıra eşitlenerek çözülür:

$$\partial \pi_c / \partial \pi_j = -1, j = 1, \dots, c - 1$$

$$\frac{\partial \log \pi_c}{\partial \pi_j} = \frac{1}{\pi_c} \frac{\partial \pi_c}{\partial \pi_j} = -\frac{1}{\pi_c}$$

$$\frac{\partial L(\boldsymbol{\pi})}{\partial \pi_j} = \frac{n_j}{\pi_j} - \frac{n_c}{\pi_c} = 0$$

ML çözümü ise aşağıdaki eşitliği verir:

$$\hat{\pi}_j / \hat{\pi}_c = n_j / n_c.$$

$$\sum_j \hat{\pi}_j = 1 = \frac{\hat{\pi}_c \left( \sum_j n_j \right)}{n_c} = \frac{\hat{\pi}_c n}{n_c},$$

$\hat{\pi}_c = n_c / n$  olduğuna göre  $\hat{\pi}_j = n_j / n$  olur. Bu çözüm olabilirliği en yükseğe çıkartır.  $\{\pi_j\}$ ' nin ML tahminleri örnek oranlarıdır.

### 1.5.1. Belirli Bir Çokterimli İçin Pearson İstatistiği

1900 yılında ünlü İngiliz istatistikçi Karl Pearson ilk yorumlayıcı metotlardan biri olarak bir hipotez testi ortaya koymuştur. Bu kategorik veri analizi üzerinde özellikle birlikteliklerin tanımlanmasına yoğunlaşılması ile devrim niteliğinde bir etki oluşturmuştur (Agresti, 2002). Pearson' nun testi, çokterimli parametrelerin kesin belirlenmiş değerlere eşit olup olmadığını ölçmekte ve değerlendirmektedir. Bu testi geliştirirken Pearson' un asıl hareket noktası; belirli bir Monte Carlo rulet çarkı üzerindeki olası sonuçların eşit olasılıklara sahip olup olmadığını analiz etmektir.

---

<sup>11</sup>  $\pi_c = 1 - (\pi_1 + \dots + \pi_{c-1})$  olduğu için bu mümkündür.

(Stigler, 1986).  $H_0 : \pi_j = \pi_{j0}, j = 1, \dots, c$ <sup>12</sup> (  $\sum_j \pi_{j0} = 1$  ) dikkate aldığımızda;  $H_0$  doğru ise  $\{n_j\}$ ' nin beklenen değerleri (beklenen frekanslar olarak da ifade edilebilmektedir.)  $\mu_j = n\pi_{j0}, j = 1, \dots, c$  şeklinde oluşur. Pearson aşağıdaki test istatistiğini ortaya koymuştur:

$$X^2 = \sum_j \frac{(n_j - \mu_j)^2}{\mu_j}$$

Sabit bir n değerine göre, daha büyük  $\{n_j - \mu_j\}$  farkları daha büyük  $X^2$  değerlerine neden olur.  $X^2_0; X^2$ ' nin gözlemlenmiş değerini ifade etmektedir.

P-değeri ise  $P(X^2 \geq X^2_0)$ ' in sıfır değeridir. Bu;  $X^2 \geq X^2_0$  ile birlikte n' in toplamını oluşturmaktadır ve sıfır hipotezi altında çokterimli olasılıkların toplamına eşittir. Büyük örnekler için  $X^2$  istatistiği yaklaşık olarak  $df=c-1$  serbestlik dereceli ki-kare dağılımına sahiptir. P-değeri;  $P(\chi^2_{c-1} \geq X^2_0)$ ' ya yaklaşmaktadır.  $\chi^2_{c-1}$  ise  $df=c-1$  serbestlik dereceli bir ki-kare rassal değişkenini ifade etmektedir.

### 1.5.2. Örnek: Mendel' in Teorileri

Birçok uygulamanın arasında; Pearson' in testi; doğal kalıtım ile ilgili Mendel Teorilerinin test edilmesi için genetik alanında kullanılmıştır. Mendel sadece yeşil türdeki bezelye bitkileri ile sadece sarı türdeki bezelye bitkilerini melezlemiştir (çaprazlamıştır.). Mendel' in tahmini ise; ikinci kuşak melez döllerin %75' i sarı ve %25' i yeşil olabilir, sarı baskın yapıdadır. Bir denemede  $n=8023$  melez üretilmiştir ki bunların  $n_1 = 6022$  tanesi sarı ve  $n_2 = 2001$  tanesi yeşil olmuştur.  $H_0 : \pi_{10} = 0.75, \pi_{20} = 0.25$ ' e göre beklenen frekanslar;  $\mu_1 = 8023(0.75) = 6017,25$  ve  $\mu_2 = 2005,75$  şeklindedir. Pearson istatistiği  $X^2 = 0.015(df = 1)$   $P=0.90$ ' lı bir P-değerine sahiptir. Bu Mendel' in iddiasının aksini iddia etmemektedir. Mendel bu tür denemelerden birkaç kere daha yapmıştır.

<sup>12</sup> “  $H_a$ : En az bir tanesi farklıdır” şeklinde alternatif hipotezi kurulur.

<sup>13</sup> Pearson Ki-Kare istatistiği olarak adlandırılmaktadır.

1936’ da, R.A.Fisher, Mendel’ in elde ettiği sonuçları özetlemiştir. Fisher ki-kare’ nin yeniden üretilebilir özelliğini kullanmıştır: Eğer  $X^2_1, \dots, X^2_k$ ;  $v_1, \dots, v_k$  serbestlik dereceli birbirinden bağımsız ki-kare istatistikleri ise;  $\sum_i \chi^2_i$ ;  $\sum_i v_i$  serbestlik derecesine sahip bir ki-kare dağılımıdır. Fisher, df=84 serbestlik dereceli 42’ ye eşit özet bir ki-kare dağılımı elde etmiştir. df =84 serbestlik dereceli bir ki-kare dağılımının ortalaması 42 ve standart sapması  $(2 \times 84)^{1/2} = 13.0$  şeklindedir<sup>14</sup>. 42’nin üstündeki sağ kuyruk olasılığı P=0.99996 olur. Başka bir deyişle ki-kare istatistiği oldukça küçüktür ve uyum oldukça iyi gibi gözükmektedir.

Fisher’in yorumu: “ Mendel’ in beklentileri ile rapor sonuçları arasındaki uyumun genel seviyesi; binlerce yinelemeden en iyisi için beklenebilecek duruma daha yakındır.” Özet olarak uyum iyiliği testleri, sadece uyumun yetersiz olduğunu ortaya koymaz aynı zamanda beklediğimiz rassal dalgalanmalardan daha iyi bir uyum olduğunu da ortaya koyacaktır.

### 1.5.3. Ki-Kare Teorik Açıklaması

Pearson istatistiğinin neden sınırlandırılmış bir ki-kare dağılımı olduğu ana hatları ile açıklanacaktır. n genişliğinde  $(n_1, \dots, n_c)$  şeklinde çokterimli bir örnekleme göre;  $n_j$ ’ nin marjinal dağılımı  $bin(n, \pi_j)$  şeklinde bir binom dağılımıdır. n büyük olduğunda, binomun normal yaklaşımına göre  $n_j$  ( ve  $\hat{\pi}_j = n_j / n$  ) normal dağılıma yaklaşır. Daha genel olarak “Merkezi Limit Teoremi” yoluyla örneklem oranları çok değişkenli normal dağılıma yaklaşır:

$$\hat{\pi} = (n_1/n, \dots, n_{c-1}/n)$$

<sup>14</sup> Ki-kare dağılımının ortalaması df, varyansı 2df ve çarpıklığı  $(8/df)^{1/2}$  şeklindedir. df büyüdükçe normale yakınsar.

$\Sigma_0$  ;  $\sqrt{n\hat{\pi}}$  sıfır kovaryans matrisini ifade etsin ve  $\pi_0 = (\pi_{1,0}, \dots, \pi_{c-1,0})'$  olsun.  $H_0$  altında,  $\sqrt{n(\hat{\pi} - \pi_0)}$ ;  $N(0, \Sigma_0)$  şeklinde normal bir dağılıma yakınsadığından dolayı, kuadratik formu aşağıdaki biçimde  $df=c-1$  serbestlik dereceli ki-kare' ye yakınsayan bir dağılıma sahiptir:

$$n(\hat{\pi} - \pi_0)' \Sigma_0^{-1} (\hat{\pi} - \pi_0)$$

$\sqrt{n\hat{\pi}}$  kovaryans matrisinin elemanları:

$$\sigma_{jk} = \begin{cases} -\pi_j \pi_k & \text{if } j \neq k \\ \pi_j(1 - \pi_j) & \text{if } j = k \end{cases}$$

$j \neq k$  olduğunda;  $\Sigma_0^{-1}$  matrisinin (j,k)' ninci elemanı  $1/\pi_{c0}$ ,  $j = k$  olduğunda  $(1/\pi_{j0} + 1/\pi_{c0})$  olur (Bu durum;  $\Sigma_0 \Sigma_0^{-1}$  ifadesinin birim matrise eşit olduğunu göstermek aracılığı ile kanıtlanabilir.). Bu ikame ile doğrudan yapılacak hesaplama  $\sigma_{jk}$  formülünü  $X^2$  için kolaylaştıracaktır.

Bu konu 1900 yılında Pearson' un ortaya attığı iddia ile benzeşmektedir. R.A. Fisher (1922) daha basit bir gerekçe vermiştir, ana fikir şudur ki;  $(n_1, \dots, n_c)$ , nin  $(\mu_1, \dots, \mu_c)$  ortalamalı bağımsız Poisson rassal değişkenleri olduğunu varsayalım,  $\{\mu_j\}$ , nin büyük olmasına göre standartlaştırılmış  $\{z_j = (n_j - \mu_j) / \sqrt{\mu_j}\}$  değerleri, yaklaşık bir normal dağılıma sahiptir. Bu nedenle,  $\sum_j z_j^2 = X^2$  ifadesi c serbestlik dereceleri ile yaklaşık bir ki-kare dağılımına sahiptir.  $\sum_j (n_j - \mu_j) = 0$  şeklinde tek bir doğrusal sınırlama eklenir ve böylece Poisson dağılımı çok terimliye dönüştürülerek bir serbestlik derecesi kaybedilir (Agresti, 2002).  $c=2$  olduğunda ise, Pearson  $X^2$  değeri normal skor istatistiğini basitleştirir<sup>15</sup>. Mendel'in verilerine

---

<sup>15</sup> Skor istatistiğinin,  $z_s = \frac{u(\pi_0)}{[u(\pi_0)]^{1/2}} = \frac{y - n\pi_0}{\sqrt{n\pi_0(1-\pi_0)}} = \frac{\hat{\pi} - \pi_0}{\sqrt{\pi_0(1-\pi_0)/n}}$  olduğunu hatırlayalım.

göre;  $X^2 = (0.123)^2 = 0.015$  değeri için  $\hat{\pi}_1 = 6022/8023, \pi_{10} = 0.75, n = 8023$  ve  $s = 0.123$  bulunur. Sonuç olarak “c” için Pearson istatistiği; çok terimli parametreler için skor istatistiğidir.

#### 1.5.4. Ki-kare Olabilirlik Oranı

Çok terimli parametreler için alternatif bir test olarak olabilirlik-oran testinden yararlanır. Çokterimli olabilirliğin çekirdeği ise tüm  $\pi_j \geq 0$  ve  $\sum_j \pi_j = 1$  olduğunda  $\prod_j \pi_j^{n_j}$  şeklindedir.  $H_0$  altında olabilirlik;  $\hat{\pi}_j = \pi_{j0}$  olduğunda en yükseğe çıkar. Genel bir durum için;  $\hat{\pi}_j = n_j / n$  olduğunda maksimize edilir. Olabilirliklerin oranı aşağıdaki ifadeye eşittir:

$$\Lambda = \frac{\prod_j (\pi_{j0})^{n_j}}{\prod_j (n_j / n)^{n_j}}$$

Olabilirlik oran istatistiği  $G^2$  ile ifade edilir:

$$G^2 = -2 \log \Lambda = 2 \sum n_j \log(n_j / n \pi_{j0})$$

Bu istatistik<sup>16</sup> ki-kare olabilirlik oran istatistiğidir. Daha büyük  $G^2$  değerleri  $H_0$ , a karşı daha büyük kanıt oluşturur. Genel olarak; parametre uzayı  $\sum_j \pi_j = 1$ , e bağlı olarak  $\{\pi_j\}$ , den oluşmaktadır. Bu yüzden boyutu da c-1'dir.  $H_0$  altında;  $\{\pi_j\}$  tamamen belirlenmiştir ve boyutu ise 0 olur. Büyük n durumunda  $G^2$ , df=c-1 serbestlik dereceli ki-kare sıfır dağılımıdır.

---

<sup>16</sup>  $2 \sum \text{gözlemlenen} \log \frac{\text{gözlemlenen}}{\text{uygun}}$  şeklindedir.



$H_0$  geçerli olduğunda, Pearson  $X^2$  ve  $G^2$  olabilirlik oranının her ikisi de,  $df=c-1$  serbestlik dereceli asimptotik ki-kare dağılımlarına sahiptir. Gerçekte  $X^2 - G^2$  olasılığı sıfıra yaklaştığında her ikisi de asimptotik olarak eşittir.  $H_0$  yanlış olduğunda ise  $n$ 'e göre orantısal olarak artma eğilimindedirler; fakat çok büyük  $n$  değerleri olsa dahi benzer değerler almak zorunda değillerdir. Belirli, sabit bir  $c$  değerine göre,  $n$  arttıkça  $X^2$ 'nin dağılımı genellikle  $G^2$  dağılımından daha hızlı bir şekilde ki-kare dağılımına yakınsar. Ki-kare yakınsaması;  $n/c < 5$  için  $G^2$  açısından oldukça zayıftır.  $c$  büyükse ve tablo hem çok küçük hem de çok büyük frekansları göstermiyorsa;  $n/c$  değeri 1 kadar küçük olduğunda  $X^2$  için uygun olabilir. Alternatif olarak bu test istatistikleri için kesin olasılık dağılımlarını oluşturmak amacıyla çokterimli olasılıklar kullanılabilir (Agresti, 2002).

#### 1.5.5. Tahminlenen ve Beklenen Frekanslarla Birlikte Test Etme

$$X^2 = \sum_j \frac{(n_j - \mu_j)^2}{\mu_j}$$

şeklindeki Pearson  $X^2$ ; örneklem dağılımı ile  $\{\pi_{j0}\}$

hipotetik değerini karşılaştırmaktadır. Bazı uygulamalarda,  $\{\pi_{j0} = \pi_{j0}(\theta)\}$ ; bilinmeyen  $\theta$  parametresinin daha küçük bir kümesinin fonksiyonlarıdır.  $\theta$ 'nin  $\hat{\theta}$  şeklindeki ML tahminleri,  $\{\pi_{j0}\}$ , in  $\{\pi_{j0}(\hat{\theta})\}$  şeklindeki ML tahminlerini ve  $X^2$  için beklenen sıklıkların ML tahminlerini  $(\{\hat{\mu}_j = n\pi_{j0}(\hat{\theta})\})$  belirler.  $\{\hat{\mu}_j\}$  tahminlerinin  $\{\mu_j\}$  yerine kullanılması  $X^2$ 'nin dağılımını etkiler.  $\dim(\theta) = p$  olduğunda gerçek serbestlik derecesi  $df=(c-1)-p$  olur. Bu konuda Pearson başarısız olmuştur. Şimdi tahminlenmiş beklenen değerler ile birlikte uyum iyiliği testi gösterilecektir. Florida'da günlük buzağı doğumlarından 156 tane örneklem alınmış ve doğumlarından itibaren 60 gün içinde akciğer iltihaplanmasına (zatürree) yakalanıp yakalanmadıklarına göre sınıflandırılmıştır (Agresti, 2002).

Zatürree enfeksiyonuna yakalanan buzağular aynı zamanda ilk enfeksiyon bulunduktan ve temizlendikten sonra iki hafta içinde ikinci kez enfeksiyon kapıp kapmamalarına göre de sınıflandırılır. Tablo-1.1’ de veriler gösterilmiştir. İlk enfeksiyona yakalanmayan buzağular ikinci enfeksiyona da yakalanmazlar. Böylece hiçbir gözlem ilk enfeksiyona yakalanma için “hayır”, ikinci kez enfeksiyona yakalanma için “evet” şeklinde bir kategorinin içine düşmez. Böyle bir kombinasyon “yapısal sıfır” olarak adlandırılır.

**Tablo-1.1** İlk Ve İkinci Enfeksiyona Yakalanan Buzağular

İlk Enfeksiyon	İkinci Enfeksiyon*	
	Evet	Hayır
Evet	30(38.1)	63(39)
Hayır	0(-)	63(78.9)

Kaynak: Agresti, 2002

\* Beklenen değerler parantez içinde yer almaktadır.

Bu çalışmanın amacı; ilk enfeksiyona yakalanan buzağı için; ilk enfeksiyonun olasılığının ikincil enfeksiyonun koşullu olasılığı ile aynı olup olmadığını test etmektir. Başka bir ifade ile, eğer  $\pi_{ab}$ , bu tablonun b’ ninci sütunu ve a’ nıncı sırasında bir buzağının sınıflandırılmış olması olasılığını gösterirse, sıfır hipotezi;  $H_0 = \pi_{11} + \pi_{12} = \pi_{11} / (\pi_{11} + \pi_{12})$  ya da  $\pi_{11} = (\pi_{11} + \pi_{12})^2$  şeklindedir.  $\pi = \pi_{11} + \pi_{12}$  ilk enfeksiyon olasılığını gösterir. Sıfır hipotezi,  $(\pi^2, \pi(1-\pi), 1-\pi)$  ifadesini sağlayan, birinci ve ikinci enfeksiyona göre oluşturulan kategoriler (evet-evet, evet-hayır, hayır-hayır) için üç terimli olasılıkları gösteren Tablo-1.2’ de yapıyı sağlayan olasılıkları belirlemektedir.

$n_{ab}$ ; (a,b) şeklindeki kategori içindeki gözlem sayısını göstermektedir.  $\pi$ ’ nin ML tahmini çokterimli olabilirliğin çekirdeğini maksimize eden değerdir ve  $(\pi^2)^{n_{11}} (\pi - \pi^2)^{n_{12}} (1 - \pi)^{n_{22}}$  şeklinde ifade edilir. Log-olabilirliği ise aşağıdaki gibidir:

$$L(\pi) = n_{11} \log \pi^2 + n_{12} \log(\pi - \pi^2) + n_{22} \log(1 - \pi)$$

$\pi$ ' ye göre türev alınırsa olabirlik eşitliği elde edilir:

$$\frac{2n_{11}}{\pi} + \frac{n_{12}}{\pi} - \frac{n_{12}}{1 - \pi} - \frac{n_{22}}{1 - \pi} = 0.$$

Çözüm ise aşağıdaki gibidir:

$$\hat{\pi} = (2n_{11} + n_{12}) / (2n_{11} + 2n_{12} + n_{22})$$

**Tablo-1.2** Hipoteze Göre Olasılık Yapısı

İlk Enfeksiyon	İkinci Enfeksiyon		
	Evet	Hayır	Toplam
Evet	$\pi^2$	$\pi(1 - \pi)$	$\pi$
Hayır	-	$1 - \pi$	$1 - \pi$

Tablo-1.1' e göre,  $\hat{\pi} = 0.494$  olur.  $n=156$  olduğundan dolayı tahminlenmiş beklenen frekanslar aşağıdaki şekilde hesaplanır. Tablo-1.1 bunları göstermektedir:

$$\hat{\mu}_{11} = n\hat{\pi}^2 = 38.1, \hat{\mu}_{12} = n(\hat{\pi} - \hat{\pi}^2) = 39.0 \text{ ve } \hat{\mu}_{13} = n(1 - \hat{\pi}) = 78.9$$

Pearson istatistiği ise  $X^2 = 19.7$  olur.  $c=3$  olduğunda, olası yanıtlar beklenen frekansları belirleyen  $p=1$  ve  $(\pi)$  parametresine sahiptir ve serbestlik derecesi,  $df=(3-1)-1=1$  olur<sup>17</sup>.  $H_0(P = 0.00001)$  hipotezine karşı güçlü bir kanıt oluşturur.

<sup>17</sup>  $\dim(\theta) = p$  olduğunda gerçek serbestlik derecesi  $df=(c-1)-p$  olduğunu hatırlayalım.

Tablo-1.1' in incelendiğinde  $H_0$  ' in öne sürdüğünden çok fazla buzağının birincil enfeksiyona yakalandığı fakat ikinci enfeksiyona yakalanmadığını görülür. Araştırmacılar, birincil enfeksiyonun bağışıklık kazandıran bir etkisi olduğunu bu nedenle ikincil enfeksiyona yakalanma olabilirliğinin azaldığı sonucuna varırlar.

## İKİNCİ BÖLÜM

### OLUMSALLIK (KONTENJANS) TABLOLARI

Çoğu çok değişkenli analizin özünde birlikteliğin analizi yer almaktadır. Bu bölümde kategorik değişkenler arasındaki ilişkileri gösteren tablolar ve birliktelikleri özetleyen parametreler tanımlanacaktır. Öncelikle iki kategorik değişken arasında ki birliktelikler ve birlikteliği ifade eden parametreler tanıtılarak bu parametreler için yorumlama metotları gösterilecektir. Bu parametreler sonuç kategorilerdeki yanıtların oranlarına göre grupların karşılaştırılmalarında kullanılır. Parametreler içinde bahis oranı ya da odds oranı (odds ratio) özel bir öneme sahiptir ki modellerde önemli bir parametre olarak görülmektedir. Sonra üçüncü bir değişkenin kontrol edilmesi ile alan genişletilecektir. Kontrol altında bir birliktelik ciddi olarak değişebilmektedir. Daha sonra nominal ve ordinal çok kategorili değişkenler için parametreler gösterilecektir.

Tablo 2.1' de Amerikalılardan oluşan bir örnek için ölümden sonraki hayata hakkındaki düşüncelerine ve cinsiyetlerine göre çapraz sınıflandırma yapılmıştır. Örnekte yer alan bayanlardan 435 kişi inandığını, 147 kişi ise ya inanmadığını ya da kararsız kaldığını söylemiştir. Bu şekildeki verilerde, cinsiyet ile ölümden sonraki hayata olan inanç arasında bir ilişkinin (birlikteliğin) olup olmadığı incelenebilir. Bir cinsiyetin diğerine kıyasla ölümden sonraki hayata inanması çok daha olası mıdır? Ya da ölümden sonraki hayata inanma cinsiyetten bağımsız mıdır? Birçok uygulama iki kategoride gözlemlenen bağıntılı gözlem sayılarına göre iki grubun karşılaştırmasını gerektirmektedir. Tablo 2.1' e göre ölümden sonraki hayata inanan kadın ve erkelerin oranları karşılaştırılabilir.

**Tablo 2.1** Cinsiyete Göre Ölümden Sonraki Hayata Olan İnançların Çapraz Sınıflaması

<b>Ölümden sonraki hayata inanma</b>		
Cinsiyet	Evet	<i>Hayır ya da Kararsız</i>
Kadın	435	147
<i>Erkek</i>	375	134

Kaynak: Agresti,1996, s:17

Orantıların oranlarının ve farklılıklarının analiz edilmesi için çeşitli metotlar ile odds oranı (odds ratio) tanıtılacaktır ki birçok metot için kilit rol oynamaktadır. Daha sonra iki kategorik değişken arasında bir birlikteliğin olup olmadığı konusunda büyük örneklem anlamlılık testleri anlatılacaktır. Sonra nominal değişkenler için testler ve ordinal değişkenler için alternatif testler ortaya konulacaktır.

## 2.1. OLUMSALLIK TABLOLARI İÇİN OLASILIKLARIN YAPISI

Kategorik veri, yanıt kategorileri içinde meydana gelen gözlemlerin frekans sayılarından meydana gelmektedir (Agresti, 1996). I tane düzeye sahip X ve J tane düzeye sahip Y gibi iki kategorik değişken olsun. X' in kategorileri için I tane sıra ve Y' nin kategorileri için J tane sütuna sahip dikdörtgen şeklindeki tabloda sonuçların IJ tane olası kombinasyonu gösterilir. Tablonun hücreleri IJ tane olası sonucu göstermektedir.

Bu şekilde, sonuçların frekans sayılarını içeren hücelere sahip tabloya “Çapraz Tablo” , “Olumsuzluk Tablosu” ya da “Kontenjans Tablosu” (bu ifade Karl Pearson tarafından 1904 yılında ortaya konulmuştur.) adı verilmektedir<sup>18</sup>. Olumsuzluk

---

<sup>18</sup> Olumsuzluk tablolarından daha çok az sayıda seçenek içeren kategorik değişkenler için ya da sınıflara bölünerek kodlanmış ve k sayıda sınıfa indirgenmiş aralık ölçekli verilerin gösteriminde yararlanılır. Değişkenin/değişkenlerin kategorilerinin sayısının değişken sayısından az olduğu durumlarda, kategorik verileri tablolar halinde özetlemek uygun bir yaklaşımdır. Kategorik değişkenlerin ya da **kategorik** hale dönüştürülebilen değişkenlerin özetlenmesinde genellikle iki

tabloları sosyal bilimlerde en eski ve en yaygın kullanıma sahip istatistiksel araçlardır (Power ve Xie, 2000:87). Onların bu kadar popüler olmasının en önemli sebebi basit oluşudur. Diğer bir sebebi ise tabloların parametrik olmaması ya da çok zayıf parametrik varsayımlar (dağılımla ilgili) gerektirmesidir. Araştırmacı bir olumsuzluk tablosu aracılığı ile doğrudan tanımlayıcı istatistikleri yorumlar ve kesin bir modellemeye başvurmaksızın önemli sonuçlar elde eder.

Goodman (1981) üç ideal tür tablo ortaya koymuştur (Goodman, 1981):

- İki açıklayıcı değişkenin ortak dağılımı (Boy ve kilo gibi)
- Bir açıklayıcı değişken ile bir sonuç değişkeninin nedensel ilişkisi (Sigara içimi ve akciğer kanseri)
- İki sonuç değişken arasındaki ilişki (Evlilik öncesi ilişki konusundaki tutum ile kürtaj konusunda ki tutum)

Bu üç tür olumsuzluk tablosu arasındaki fark kavramsaldir. Çünkü aynı şekilde hazırlanırlar. Gerçekte olumsuzluk tabloları için kullanılan istatistiksel modeller, genelleştirilmiş bir doğrusal modeldeki bir bağımlı değişkeni açıklayacak bağımsız değişkene bağlı tahminlemeden öte sıklıklara bağlı olarak oluşturulmaktadır (Powers, ve Xie, 2000, s:87). Korelasyon olması halinde çapraz tablolar doğal olarak simetrikdir.

Olumsuzluk tablosu iki değişkenin çapraz sınıflamasından oluşuyorsa “iki-yönlü tablo”, üç değişkenli değişkenlerin çapraz sınıflamasından oluşuyorsa “üç-yönlü tablo” adı verilir. İki yönlü tablo I tane sıra ve J tane sütuna sahip olup  $I \times J$  tablosu olarak ifade edilir. Tablo 2. 1,  $2 \times 2$  şeklinde bir tablodur. İki kategorik değişken arasındaki ortak dağılım onların ilişkilerini belirler. Bu dağılım aynı zamanda marjinal ve koşullu olasılıkları da belirlemektedir.  $\pi_{ij} = P(X = i, Y = j)$ ; i’ ninci sıra ve j’ ninci sütunda yer alan hücrede (X,Y)’ nin bulunma olasılığıdır.

---

yönlü tablolar kullanılmaktadır. İki yönlü tablolarda ilişkiyi belirlemede yaygın olarak kullanılan istatistikler arasında; Fisher’in Kesin Olasılık Testi, Ki-kare Testi, Oran testi, G testi, Logaritmik Doğrusal modeller sayılabilir(Eyduran, Özdemir ve Küçük, 2005, s:1-3).

$\{\pi_{ij}\}$  değerleri X ve Y' nin ortak dağılımını meydana getirmektedir ve bu değerler  $\sum_{i,j} \pi_{ij} = 1$  ifadesini sağlamaktadır. Marjinal dağılımlar; sıra ve sütun toplamlarıdır, ortak olasılıkların toplamından meydana gelirler. Bu toplamaları göstermek için sıra değişkeni için  $\{\pi_{i+}\}$  ve sütun değişkeni için  $\{\pi_{+j}\}$  kullanılacaktır, “+” indisi bulunduğu hücreye ait tüm indeksin toplamıdır:  $\{\pi_{i+}\} = \sum_j \pi_{ij}$  ve  $\{\pi_{+j}\} = \sum_i \pi_{ij}$ .

$2 \times 2$  boyutlu olumsallık tablosu için;  $\pi_{1+} = \pi_{11} + \pi_{12}$  ve  $\pi_{+1} = \pi_{11} + \pi_{21}$  şeklindedir. Örneklem için;  $\pi$  yerine p kullanılır. O zaman,  $\{p_{ij}\}$  örnek ortak dağılımını ifade eder ve bu değerler de örnek hücre oranlarını oluşturur. Hücredeki denek sayısı  $\{n_{ij}\}$  ile gösterilir ve  $n = \sum_{i,j} n_{ij}$  ise toplam örneklem büyüklüğünü

ifade eder.  $p_{ij} = \frac{n_{ij}}{n}$  ile hücre oranları ve hücredeki denek sayısı ilişkilendirilir.

Marjinal sıklıklar sıra toplamlarından  $\{\pi_{i+}\}$  ve sütun toplamlarından  $\{\pi_{+j}\}$  oluşur. Bu ifadeler  $\sum_j \pi_{+j} = \sum_i \pi_{i+} = \sum_i \sum_j \pi_{ij} = 1.0$  eşitliğini sağlamaktadır. Marjinal dağılımlar tek değişken için bilgi verir. Çoğu olumsallık tablosunda Y şeklindeki yanıt değişkeni sütun değişkeni ve diğer X açıklayıcı değişkeni ise satır (sıra) değişkenidir. Bu X' in her bir seviyesine göre Y' nin tek başına olasılık dağılımının düzenlenmesinde aydınlatıcı ve bilgi verici niteliktedir. Böyle bir dağılım belirli bir X seviyesine göre Y' nin koşullu olasılıklarından oluşmaktadır ve koşullu dağılım olarak adlandırılmaktadır (Agresti, 1996).

X rassal değil belirlenmiş sabit bir değere sahipse artık X ve Y için ortak dağılım notasyonu anlamlı değildir. Fakat (X,Y) şeklinde belirlenmiş kategori bir olasılık dağılımına sahiptir ki X' in kategorisinin değişmesine karşın bu dağılımın nasıl değiştiğini gösteren çalışma ile ilgilenilir (Agresti, 2002). X' in i' ninci sırasında sınıflandırılmış belirli bir denek için  $\pi_{j/i}$ , Y' nin j' ninci sütununda sınıflandırılma olasılığını ifade etmektedir ( $j=1, \dots, J$ ) ve ayrıca  $\sum_j \pi_{j/i} = 1$  olduğuna



dikkat edilmelidir.  $\{\pi_{1|i}, \dots, \pi_{j|i}\}$  olasılıkları X' in i kategorisinde, Y' nin koşullu dağılımını oluşturmaktadır. Birçok çalışmanın ana amacı açıklayıcı değişkenin çeşitli seviyelerinde Y' nin koşullu dağılımlarını karşılaştırmaktır (Power ve Xie, 2000). Olumsallık tabloları için olasılık dağılımları örnekleme düzenine bağlıdır. İlk olarak iki-yönlü çapraz tablolar için temel olasılık türleri ele alınacaktır. İlk varsayımımız bir örnekteki deneğin ilgilenilen populyasyondan rastgele çekildiğidir ve sonra ki varsayımımız ise bu deneğin, iki kategorik yanıtı göre sınıflandırılmış olduğudur. Tablo 2.1, 2×2 boyutlu olumsallık tablosu olup, ölümden sonraki hayata olan inançlar ile cinsiyetlere göre yapılan “1991 Genel Sosyal Anketi” ne cevap veren n=1091 kişi çapraz olarak sınıflandırılmıştır. Tablo 2.2 ise bu veriler için hücre sayısının notasyonlarla gösterilmiş halidir. Örnek olarak  $n_{11} = 435$  ve uygun örnek ortak oranı ise  $p_{11} = 435/1091 = 0.399$  olur.

**Tablo 2.2** Tablo–2.1 İçin Notasyonlar

<b>Ölümden sonraki hayata inanma</b>			
Cinsiyet	Evet	Hayır ya da Kararsız	<i>Toplam</i>
	$n_{11} = 435$	$n_{12} = 147$	$n_{1+} = 582$
Erkek	$n_{21} = 375$	$n_{22} = 134$	$n_{2+} = 509$
<i>Toplam</i>	$n_{+1} = 810$	$n_{+2} = 281$	$n = 1091$

Tablo 2.1’ de ölümden sonraki hayata inanma yanıt değişkenidir ve cinsiyet ise açıklayıcı değişkendir. Onun için belirli bir cinsiyete göre ölümden sonraki hayata inanmanın koşullu olasılığı incelenecektir. Bayanlar için, “evet” yanıtlarının oranı  $435/810 = 0.747$ , “hayır” yanıtlarının oranı ise  $147/582 = 0.253$  şeklindedir. (0.747, 0.253) kadınların ölümden sonraki hayata inanmalarının örneklem koşullu dağılımıdır. Erkekler için de örneklem koşullu dağılımı (0.737, 0.263) şeklindedir.

### 2.1.1. Duyarlılık ve Seçicilik

Tablo 2.3’ de yer alan sonuçlar, göğüs kanserini teşhis etmek amacıyla çeşitli deneme metotları konusunda yazılmış bir makaleden alınmıştır. Literatür araştırmasına bağlı olarak, yazarlar klinik göğüs kontrolü (muayenesi) ile birlikte mamografi kullanımının etkisine ilişkin sonuçları rapor etmişlerdir.

**Tablo 2.3** Göğüs Kanseri Tanılarına Göre Tahminlenen Koşullu Dağılımlar

Tanı Testi			
Göğüs kanseri	Pozitif	Negatif	Toplam
Evet	0.82	0.18	1.0
Hayır	0.01	0.99	1.0

Kaynak: Agresti, 2002

X= Gerçek hastalık durumu (bir kadının gerçekten göğüs kanseri olup olmadığı) ve Y= Teşhisler ( pozitif, negatif) olsun<sup>19</sup>. Tablodaki tahminlenmiş olasılıklar belirli bir X değerine göre Y’ nin koşullu olasılıklarıdır. Bir hastalık için teşhis testlerinde eğer denek hasta ise **pozitif test sonucu** ve eğer hasta değilse **negatif test sonucu** elde edilir. Bu iki teşhiste doğrudur. Denek gerçekten hasta ve teşhis testi pozitif ise koşullu olasılık “*duyarlılık*” olarak adlandırılır. Denek hasta değil ve teşhis negatif ise koşullu olasılık “*seçicilik*” olarak adlandırılır.

---

<sup>19</sup> Pozitif sonuç bir kadının göğüs kanseri olduğunun belirlenmesidir.

İdeal olanı her ikisinin de yüksek olmasıdır<sup>20</sup>. Çünkü testin geçerliliği artar. Ancak bu uygulamada her zaman olmaz. Bu nedenle de eğer amaç sağlamları ayırmak ise seçicilik yüksek olmalı ama eğer amaç patolojik olanları ayırmak ise duyarlılık yüksek olmalıdır (Uğurman ve Berktaş, 2003:236). Tablo 2.3 şeklinde oluşturulan  $2 \times 2$  boyutlu olumsuzluk tablosuna göre; duyarlılık= $\pi_{11}$  ve seçicilik= $\pi_{22}$  şeklindedir. Klinik kontrolü ve mamografinin birleştirilmiş tahmini duyarlılığı; 0.82 şeklinde bulunur. Göğüs kanseri olan kadınların %82' si doğru bir şekilde teşhis edilmiştir. Seçicilik ise 0.99' dur ve kanser olmayan kadınların %99' u doğru bir şekilde teşhis edilmiştir. Bu test sağlamları ayırma amacına daha iyi hizmet etmektedir.

### 2.1.2. Kategorik Değişkenlerin Bağımsızlığı

Eğer X' in her seviyesine göre bulunan Y koşullu dağılımları özdeş ise iki değişkenin birbirinden *“istatistiksel olarak bağımsız”* olduğu söylenir. İki değişken bağımsızsa, herhangi bir sütun yanıtının (j) olasılığı her bir sıra için aynıdır. Ölümden sonraki hayata inanmak cinsiyetten bağımsız ise ölümden sonraki hayata inanmanın olasılığı 0.740 olur (hem erkekler hem de kadınlar için eşittir.). Eğer iki değişken de yanıt değişkeni ise aralarındaki ilişki ortak dağılımlar kullanılarak ya da belirli bir X değeri için Y' nin koşullu dağılımı ya da belirli bir Y için X' in koşullu dağılımı kullanılarak belirlenebilmektedir. İstatistiksel bağımsızlık, bütün ortak olasılıkların, marjinal olasılıkların çarpımına eşit olmasıdır ve aşağıdaki şekilde gösterilir:

$$\pi_{ij} = \pi_{i+} \pi_{+j} ; i = 1, \dots, I \text{ ve } j = 1, \dots, J$$

X'in i' ninci sıraya düşmesi ve Y' nin de j' ninci sütuna düşmesi olasılığı her ikisinin ayrı ayrı olasılıklarının çarpımına eşittir. Belirli bir X değeri için Y' nin

---

<sup>20</sup> Gerçek pozitifler: a; gerçekten hasta olana hasta tanısında bulunmak.  
Gerçek negatifler: d; hasta olmayanlar için hasta değil tanısında bulunmak.  
Yalancı pozitifler: b; hasta olmadığı halde tanı testinin hasta olarak saptanması  
Yalancı negatifler: c; hasta olduğu halde tanı testinin saptayamadığı durumlar.  
Bu durumda: **Duyarlılık = a/a+c**  
**Seçicilik = d/d+b** şeklinde ifade edilir (Uğurman ve Berktaş, 2003: 237).

koşullu olasılığı  $\pi_{j|i} = \pi_{ij} / \pi_{i+}$  şeklindeki eşitlik ile belirtilen ortak dağılımla ilişkilidir (eşitlik tüm i ve j' ler için). Eğer tüm ortak olasılıklar değişkenlerin marjinal olasılıklarının çarpımına eşit ise iki kategorik yanıt değişkeni de bağımsız olarak tanımlanır.

**Tablo 2.4** Marjinal, Koşullu ve Ortak Olasılıkların Notasyonları

Satır	Sütun		Toplam
	1	2	
1	$\pi_{11}$ ( $\pi_{11}$ )	$\pi_{12}$ ( $\pi_{12}$ )	$\pi_{1+}$ (1.0)
2	$\pi_{21}$ ( $\pi_{21}$ )	$\pi_{22}$ ( $\pi_{22}$ )	$\pi_{2+}$ (1.0)
<b>Toplam</b>	$\pi_{+1}$	$\pi_{+2}$	(1.0)

X ve Y bağımsız ise,  $\pi_{j|i} = \pi_{ij} / \pi_{i+} = (\pi_{i+} \pi_{+j}) / \pi_{i+} = \pi_{+j}$  eşitliği söz konusudur (Y' nin her bir koşullu dağılımı marjinal dağılımı ile özdeştir.). Bu nedenle,  $\pi_{j|1} = \dots = \pi_{j|I}, j = 1, \dots, J$  olduğu zaman iki değişkende bağımsızdır. Yani herhangi bir sütun yanıtının olasılığı her sıra için aynıdır. Bu durum Y bir yanıt ve X bir açıklayıcı değişken diye tanımlanırsa bağımsızlığı tanımlamak için daha doğal bir yoldur. Bağımsızlık genellikle koşullu dağılımların homojenliğine dayanmaktadır. Tablo 2.4,  $2 \times 2$  durumu için ortak, koşullu ve marjinal dağılımların notasyonlarını göstermektedir. Örneklem dağılımları için de aynı notasyonlar kullanılır. Ancak  $\pi$  yerine,  $\hat{\pi}$  ya da p kullanılır.

Hücre frekansları ise;  $\{n_{ij}\}$  gösterilir ve  $n = \sum_{i,j} n_{ij}$  toplam örnek büyüklüğüdür.  $\{p_{ij}\}$ , örneklem ortak dağılımını göstermektedir:  $p_{ij} = \frac{n_{ij}}{n}$ . i' ninci

sıradaki deneklerin verdiği yanıt j iken örneklem oranı  $p_{j/i} = p_{ij} / p_{i+} = n_{ij} / n_{i+}$  ve  $n_{i+} = np_{i+} = \sum_j n_{ij}$  şeklindedir.

### 2.1.3. Poisson, Binom ve Çokterimli Örneklem

Burada örneklem modelleri olumsuzluk tablolarında yer alan hücre miktarları açısından ele alınacaktır. Örnek olarak  $2 \times 2$  tablosu için Poisson örneklem modeli dört hücredeki miktarların her birini  $\{Y_{ij}\}$  ve  $\{\mu_{ij}\}$  parametrelerinden oluşan bağımsız Poisson rassal değişkenleri olarak işleme sokar<sup>21</sup>. Olumsuzluk tablosunun sıraları farklı gruplardan oluşuyor ise bu gruplar için örneklem büyüklükleri genellikle *örneklem tasarımı* ile belirlenir. Rassal değil belirlenmiş X seviyeleri için marjinal toplamlar alındığında, X ve Y için ortak dağılım artık bir anlam içermemektedir. Fakat her bir X seviyesi için koşullu olasılıklar anlam kazanmaktadır. *İki yanıt kategorisi* var ise her bir *sıradaki örneklem* için bir *binom dağılımı* varsayımı yapılır (deneme sayıları, belirlenmiş sıra toplamına eşittir.). Her bir sıradaki örneklem ayrı rassal örneklem şeklinde bağımsız ise bu örneklem düzenine “bağımsız binom örneklemesi” adı verilir (Agresti, 1996, s:18). Tablodaki *toplam örnek büyüklüğü belirli* ise (fakat sıra ve sütun toplamları belli değilse) olası sonuçları veren hücreler için *çokterimli bir örneklem modeli* kullanılır.

IJ hücreleri olası sonuçları göstermektedir. Hücre miktarlarının olasılık yoğunluk fonksiyonu ise çokterimli formdadır:

$$[n! / (n_{11}! \dots n_{IJ}!)] \prod_i \prod_j \pi_{ij}^{n_{ij}}$$

---

<sup>21</sup> Potansiyel sonuçların ( $\{n_{ij}\}$ ) ortak olasılık yoğunluk fonksiyonu; IJ hücrelerine göre  $P(Y_{ij} = n_{ij})$  şeklindeki Poisson olasılıklarının çarpımıdır ya da  $\prod_i \prod_j \exp(-\mu_{ij}) \mu_{ij}^{n_{ij}} / n_{ij}!$  şeklindedir (Agresti, 1996).

Genellikle bir yanıt değişkeni üzerindeki gözlemler, X açıklayıcı değişkenin her bir düzeni için farklı bir şekilde gerçekleşir. Bu durumda normal olarak sıra toplamları sabit alınır ve basitleştirilir.  $n_i = n_{i+}$  notasyonu kullanılır (Agresti, 1996, s:19). X' in i dizisi için Y üzerinde  $n_i$  tane gözlem bağımsızdır ve her bir sıranın olasılık dağılımı  $\{\pi_{1|i}, \dots, \pi_{J|i}\}$  şeklindedir.  $\{n_{ij}, j = 1, \dots, J\}$  miktarları için elde edilen  $\sum_j n_{ij} = n_i$  ifadesi çok terimli bir şekle sahiptir:

$$\frac{n_i!}{\prod_j n_{ij}!} \prod_j \pi_{j|i}^{n_{ij}}$$

X' in farklı dizilimlerinde örneklem bağımsız ise, tüm veri setine göre ortak olasılık fonksiyonu çeşitli dizilimlerden oluşan çokterimli fonksiyonların (yukarıda verilen notasyon) çarpımıdır. Bu örnekleme düzeni *bağımsız çokterimli örneklemedir* ve aynı zamanda çokterimli örneklem çarpımı olarak adlandırılır (Agresti, 1996, s:19). Örnek olarak tablo 2.1' de rassal örneklem ile oluşturan 1091 tane denek ölümden sonraki hayata olan inançlarına ve cinsiyetlerine göre çapraz sınıflandırılmıştır. Dört hücredeki miktarlar dört kategoriye sahip çokterimli bir dağılımdan elde edilen örneklem değerleridir. Bir olumsuzluk tablosunun hücreleri aracılığı ile yapılan birçok çokterimli örneklem için, sütunlar yanıt değişkeni ve sıralar ise açıklayıcı değişkendir.

Verileri tanımlamak için, sıra toplamları ve yanıt üzerindeki koşullu dağılımlar aracılığı ile hücre miktarlarını belirlemek oldukça hassastır. Bu işlemden, sıra toplamları sabit olarak alınır ve eğer veriler ayrı bağımsız örneklerden oluşmuş ise aynı yolla analiz edilir (Agresti, 1996). Tablo 2.1' de kadınlar için sonuçlar bir binom örnekleme şeklinde işleme sokulur (ölümden sonraki hayata olan inanç bakımından “evet” ve “hayır ya da kararsız” şeklinde sonuç kategorilerinden oluşan bir binom örnekleme). Erkekler içinde aynı yanıt değişken için bağımsız binom örnekleme olarak sonuçlar dikkate alınır. Eğer iki yanıt kategorisinden daha fazla kategori olsaydı, (“evet”, “hayır”, “kararsız” şeklinde) örneklem bağımsız çokterimli örneklemler şeklinde işleme dahil edilirdi.

Bağımsız çokterimli örneklem aynı zamanda şu koşullar altında sonuçlanır (Agresti,1996): Varsayalım ki;  $\{n_{ij}\}$ , ya  $\{\mu_{ij}\}$  ortalamalı Poisson örneklemeden ya da  $\{\pi_{ij} = \mu_{ij} / n\}$  olasılıkları ile IJ hücreleri için çokterimli örneklemden meydana gelmiş olsun. X açıklayıcı bir değişken ise;  $\{\sum_j n_{ij} = n_i\}$  toplamlarına bağlı olarak istatistiksel çıkarsama yapmak oldukça hassastır (değerleri örneklem tasarımı aracılığı ile belirlenmemiş olsa bile).  $\{\sum_j n_{ij} = n_i\}$  ' ye bağlı olarak hücre miktarları ( $\{n_{ij}, j = 1, \dots, J\}$ ) aşağıdaki şekilde çokterimli dağılıma sahiptir:

$$\frac{n_i!}{\prod_j n_{ij}!} \prod_j \pi_{j|i}^{n_{ij}}$$

Yanıt olasılıkları  $\{\pi_{j|i} = \mu_{ij} / \mu_{i+}, j = 1, \dots, J\}$  şeklindedir. Farklı sıralardaki hücre miktarları birbirinden bağımsızdır. Bu şartlarla, sıra toplamları sabit olarak işleme tabi tutulur ve eğer veriler ayrı bağımsız örneklerden oluşuyorsa analiz edilir (Agresti, 2002). Çoğu analize göre, hangi örneklem modelinin en çok hassaslığı verdiği konusunda endişe duymaya gerek yoktur. Burada başlıca yorumlama metotlarına göre aynı sonuçlar poisson, çokterimli ve bağımsız binom/çokterimli örneklem modelleri için de meydana gelmektedir.

Bazen sıra ve sütun paylarının her ikisi de doğal olarak sabittir. O zamanda uygun örneklem dağılımı *hipergeometrik dağılım* olur. Bu dağılım daha sonra ele alınacaktır. Massachusetts Otoyolu Departmanındaki araştırmacılar; Massachusetts Paralı Yolunda meydana gelen kazalardaki sürücüler için emniyet kemeri kullanımı (evet, hayır) ile otomobil çarpışması sonucunun (ölümcül, ölümcül değil ) arasındaki ilişki incelemek için plan hazırlamışlar (Agresti,2002). Sonuçlar Tablo 2.5' da görülen formatta özetlenmiştir. Daha sonraki yıl için paralı yol da bütün kazaların listesi hazırlanıp, her biri bu değişkenlere göre sınıflandırılmıştır. Toplam örneklem büyüklüğü rassal bir değişkendir. Çarpışma sonucu ile emniyet kemeri kullanımının dört kombinasyonu için gözlem sayısı, bağımsız Poisson rassal değişkeni

(bilinmeyen ortalamalar ile,  $\{\mu_{11}, \mu_{12}, \mu_{21}, \mu_{22}\}$  )olarak işleme dahil edilebilir. Bunun yerine varsayalım ki; geçen yıl içinde otoyol üzerinde gerçekleşen çarpışma sonuçlarına ilişkin polis kayıtlarından 200 tanesini rassal bir şekilde alsınlar ve yine aynı değişkenlere göre sınıflandırsınlar. Bu incelemeye göre, örneklem büyüklüğü n belirlenmiştir (sabittir). O zaman araştırmacılar dört hücredeki miktarı n=200 denemeden oluşan ve bilinmeyen ortak olasılıklardan,  $\{\pi_{11}, \pi_{12}, \pi_{21}, \pi_{22}\}$  meydana gelen çokterimli rassal bir değişken olarak işleme dahil edebilirler.

**Tablo 2.5** Emniyet Kemerini Kullanımı Ve Otomobil

Çarpışmalarının Sonuçları

Çarpışma Sonucu		
Emniyet Kemerini Kullanımı	Ölümcül	Ölümcül Değil
Evet		
Hayır		

Bunun yerine de varsayalım ki; ölümlü sonuçlanan kazalar için polis kayıtları ayrı bir şekilde dosyalanmıştır. Araştırmacılar ölümcül kazalardan rassal olarak 100 örneklem ve ölümcül olmayanlardan da rassal olarak 100 örneklem seçmiş olsunlar.

Bu yaklaşım tablo-2.5' da ki sütun toplamlarını belirler. O nedenle de, araştırmacılar tablodaki her bir sütunu, bağımsız binom örnekleme olarak dikkate alabilirler. Hatta diğer yaklaşım ise; klasik deney tasarımıdır. 200 denek alınır ve emniyet kemeri takmak için 100 tanesini rassal olarak belirlenir, sonra 200 kişinin tümü kaza yapmaya zorlanır. Kayıt edilen sonuçlar, her bir sıra için bağımsız binom örnekleri olabilir (belirlenmiş sıra toplamı ile birlikte). Açıktır ki bazı deneysel bilimlerde oldukça yaygın olan klasik tasarımlar etik ve insancıl olamayabilir. Özellikle tıbbi çalışmalarda bu bir gerçektir.



#### 2.1.4. İnceleme Türleri

Tablo 2.6 akciğer kanseri ve sigara içimi arasındaki bağı ilişkin, Richard Doll ve A.Bradford Hill tarafından 1950 yılında yapılan ilk çalışmalardan birinden alınmıştır (Agresti, 2002). Londra-İngiltere’ deki 20 hastanede, önceki yıl kabul edilen kanserli hastaların sigara içimi konusundaki tutumları sorgulanmıştır. Kabul edilen 709 hastadan her biri için araştırmacılar aynı hastanede yaş bakımından aynı 5 yıllık gruplar içinden aynı cinsiyette kanser olmayan bir hastanın da sigara içme tutumunu incelemişler. Tablodaki ilk sırada 709 kişilik hasta grubu ve ikinci sütununda ise kanser olmayan 709 kişilik kontrol grubu yer almaktadır. Sigara içen en azından bir yıl için günde en az bir tane sigara içen kişi olarak tanımlanmıştır.

**Tablo-2.6** Akciğer Kanseri İle Sigara İçimin Çapraz Sınıflaması

<b>Akciğer Kanseri</b>		
<b>Sigara İçen</b>	<b>Hastalıklı Grup</b>	<b>Kontrol Grubu</b>
Evet	688	650
Hayır	21	59
Toplam	709	709

Kaynak: Agresti, 2002

Normal olarak; akciğer kanserinin meydana gelip gelmemesi yanıt değişken ve sigara ile ilgili tutum ise açıklayıcı değişkendir. Fakat bu çalışmada akciğer kanserinin marjinal dağılımı örnekleme tasarımı ile belirlenmiştir ve ölçülen sonuç ise deneğin her zaman bir sigara içicisi olup olmadığıdır. “*Geçmişe bakmak*” ile *geriye dönük tasarım (retrospektif)* kullanan inceleme; *olgu-kontrol incelemesi* olarak adlandırılır. Olgular, hastalar ve hasta olmayanlar (kontrol) olarak ikiye ayrılır ve zamanda geriye doğru gidilerek şüphelenilen risk faktörü ile maruz kalıp kalmadıkları araştırılır (Uğurman ve Berktaş, 2003: 237). Genellikle bu tip bir çalışmada iki örneklem eşleştirilmiş örneklemelerdir. Bu şekildeki çalışmalar genellikle incelenen hastalığa sahip olan yeterli ve uygun büyüklükte örnek sağlamak amacıyla sağlıkla ilgili uygulamalarda oldukça yaygın olarak kullanılmaktadır.

Bazen olgular ve kontrollerin örnekleme, eşleştirmeye kıyasla bağımsız olur. Örnek olarak daha önceden akciğer kanseri ve sigara içimi üzerine yapılan olgu-kontrol incelemesinde kullanılan denekler 1950 ya da 1951 yılında bazı kanser türlerinden ölenlerin durumları ile ilgili raporlar ele alınabilir ve gözlemler kanser tipine ve denekler sigara tutumlarına göre çapraz sınıflandırılabilir. Akciğer kanserine yakalanma oranı için, sigara içenlerle içmeyenler karşılaştırmak istenebilir. Bu oranlar, belirlenmiş sigara içme tutumuna göre akciğer kanserinin koşullu dağılımına dayanmaktadır. Bunun yerine olgu-kontrol incelemeleri, açıklamanın ters çevrilmesi ile oluşan oranları sağlamaktadır (yani belirlenmiş akciğer kanseri hastaları için sigara içme tutumlarının koşullu dağılımına göre ). Bunlara göre tablodaki akciğer kanseri olanlar için sigara içme oranı  $688/709=0.970$ , kontrol grubu için ise  $650/709=0.917$  olarak bulunur.

Akciğer kanseri olan popülasyonun oranı biliniyorsa, asıl ilgilenilen konu için örneklem koşullu dağılımını hesaplamak amacıyla *Bayes Teoremi* kullanılır (Agresti,2002). Diğer taraftan, geriye dönük bir örnekleme kullanarak, sigara içimi ile ilgili tutumun her bir kategorisi için akciğer kanseri olasılığı tahminlenemez. Tabloya göre akciğer kanserinin olduğu popülasyon bilinmemektedir ve kansere yakalanan hastalar, genel popülasyonda ortaya çıkma durumundan fazla bir orana göre örneklenmiştir.

Çelişki olarak, bir çalışma düşünelim, genç bir popülasyondan denekler örnekleme alınsın ve 60 yıl sonra sigara içenler ve içmeyenler için akciğer kanseri oranları ölçülsün. Böyle bir örnekleme tasarımı ise *ileriye dönük tasarım* (prospektif) olarak adlandırılır. İki tür ileriye dönük tasarım vardır. *Klinik denemelerde* denekler rassal olarak sigara içen ve içmeyen gruplara paylaştırılır. *Kohort çalışmalarında* ise, denekler sigara içip içmeme konusunda kendi tercihlerini yaparlar ve gelecek zaman boyunca kimin kanser olduğu gözlemlenir. Bir diğer yaklaşımda *çapraz parçalı tasarımdır* Bu durumda denekler örnekleme alınır ve her iki değişkene göre eş-zamanlı olarak sınıflandırılır. Başlangıçta sağlıklı olan ve şüphelenilen risk faktörlerine farklı düzeylerde maruz kalmış bireylerden oluşan *kohort* (incelediğimiz özelliğin dışındaki belirli bir ortak noktaya sahip bireyler

topluluğu) zamanda ileri doğru takip edilir. Bir süre sonra her bir grupta araştırılan hastalığın insidansı (Yayılma hızı: Bir zaman aralığında bir bireyin hasta olması olasılığıdır.) saptanır. Grupların risk faktörüne maruziyetin dışında diğer bütün özellikler açısından benzer olması gerekir.

Sonuçların değerlendirilmesi ikili değişkenlerde olduğu gibi 2 x 2 tablo yardımıyla yapılır. Burada sütun kısmında yine hastalık yer alır ama satır bölümünde maruz kalma durumu belirtilir. Maruz kalan ve kalmayanlardan oluşan grupların büyüklüğü baştan bellidir. Ancak her iki grupta ne kadar hasta ortaya çıkacağını zaman belirler. Her iki gruptaki hastalık insidansları karşılaştırıldığında istatistiksel olarak anlamlı ölçüde fazla (yani farklılık şansa bağlı olamayacak kadar büyük) insidans eğer etkene maruz kalan grupta görüldüyse bu durum risk faktörü ile hastalık arasındaki ilişkinin kanıtı kabul edilir. Tersine hastalık insidansı daha düşük ise etken (risk faktörü) hastalıktan koruyor demektir<sup>22</sup>(Uğurman ve Berktaş, 2003: 237).

*İleriye dönük incelemeler* genellikle X' in kategorilerine göre bulunan toplama bağlıdır ve J' nin her sırasındaki sayılar, Y için bağımsız çokterimli örneklem ile elde edilir. *Geriye dönük incelemeler* genellikle Y değişkenini,  $\{n_{+j}\}$  toplamı belirlenmiş (sabit) olarak işleme dahil eder ve I' nın her sütunundaki sayılar X için çokterimli örneklem ile elde edilir. *Çapraz parçalı incelemelerde* ise toplam örneklem büyüklüğü belirlenmiştir. Fakat sıra ile sütun toplamaları belirlenmemiştir ve IJ hücre sayıları çokterimli örneklem ile elde edilir. Olgu-kontrol, kohort ve çapraz parçalı incelemeler “*gözlemsel incelemeler*” olarak adlandırılır. Onlar açık bir şekilde, ilgilenilen sonuca sahip denekleri ve deneklerin grup seçimlerini gözlemlerler. Aksine klinik bir deneme, *deneysel bir çalışmadır*, araştırmacı deneysel kontrolün avantajına sahiptir (yani her işleme alınan denekler üzerinde kontrol avantajı vardır.). Gözlemsel incelemeler oldukça yaygındır ancak diğer

---

<sup>22</sup> Eğer veri sayımla elde edildiyse (sayımla elde edilen veriler hasta sayısı, ölenlerin sayısı gibi 3, 5, 70 şeklindeki tam sayılardır. Süreklilik gösteren 3.5, ya da 2.75 gibi sayılar bu amaçla kullanılamaz) risk faktörü ile hastalık arasındaki ilişkinin gücü göreceli risk ya da odds oranı ile gösterilebilir. Ancak eğer veri sürekli sayısal değişkenlerden oluşuyorsa (mg/dl, boy, ağırlıkta olduğu gibi 2,376 şeklinde bütün ara değerleri alabiliyorsa) ilişkinin gücü korelasyon ve regresyon analizleri ile belirlenir.

inceleme türlerindeki sapmalara kıyasla daha fazla potansiyele sahiptir. Bazı dezavantajlara sahiptir: Rassal olmayabilir ve ölçülmüş değişkenler için hesap yapılır ancak genelleme yapılamaz.

## 2.2. İKİ ORANIN KARŞILAŞTIRILMASI

İki kategoriye sahip yanıt değişkenler ikili değişkenler olarak adlandırılır. Örnek olarak, ölümden sonraki hayata inanma durumu “evet, hayır” kategorileri kullanılarak ölçülürse ikili bir değişkendir. Birçok çalışmada ikili yanıt (Y) göre iki grup karşılaştırılır. Veriler  $2 \times 2$  boyutlu olumsuzluk tablolarında gösterilir. Sıralar iki grubu ifade ederken sütunlar Y'nin yanıt seviyelerini ifade eder (Sıralar gruplardır ve sütunlar ise Y'nin kategorileridir.). Bu kısımda ikili yanıtlara göre iki grubun karşılaştırılmasındaki ölçümler yani grupların karşılaştırılmasındaki parametreler tanıtılacaktır.

### 2.2.1. Oranların Farkı

İkili değişken için yanıt kategorilerini başarı ve başarısızlık şeklindeki genel terimleri kullanarak ifade edelim. 1' inci sıradaki denekler için,  $\pi_1$  başarının olasılığı ve  $1 - \pi_1$  başarısızlığın olasılığıdır.  $(\pi_1, 1 - \pi_1)$  olasılıkları 1' inci sıradaki Y'nin koşullu dağılımını şekillendirir. 2' inci sıradaki denekler için,  $\pi_2$  başarının olasılığıdır.  $\pi_1 - \pi_2$  şeklindeki oranların farkı iki sıradaki başarı olasılıklarının karşılaştırılmasıdır. Bu fark -1 ile +1 arasında değer alır. Eğer  $\pi_1 = \pi_2$  ise bu fark sıfır olur ki yanıt iki gruba göre bağımsızdır.  $p_1$  ve  $p_2$ , iki sıra için başarıların örneklem oranlarını göstermektedir. Örneklem farkı  $p_1 - p_2$  ile  $\pi_1 - \pi_2$  tahminlenir. Örnek olarak tablo 2.2' nin birinci sırası için;  $p_1 = n_{11} / n_{1+} = 435 / 582 = 0.747$  örneklem büyüklüğünün birinci sıraya bölünmüş “evet” yanıtlarının sayısının birinci sıradaki örneklem oranıdır.  $p_2 = n_{21} / n_{2+} = 375 / 509 = 0.737$  ikinci sıra için örneklem oranıdır. Oranların örneklem farkı:  $0,747 - 0,737 = 0,010$  şeklinde bulunur. Kolaylık sağlamak amacıyla, iki grup için örneklem büyüklüklerini (yani sıra

toplamlarını:  $n_{1+}$  ve  $n_{2+}$  )  $N_1$  ve  $N_2$  ile gösterelim. İki sıradaki sayılar bağımsız binom örnekleme ise  $p_1 - p_2$  'nin tahminlenmiş standart hatası:

$$\hat{\sigma}(p_2 - p_1) = \sqrt{\frac{p_1(1-p_1)}{N_1} + \frac{p_2(1-p_2)}{N_2}}$$

Örnekleme büyüklükleri arttıkça standart hata azalır ve böylece  $\pi_1 - \pi_2$  'nin tahmini düzelir.  $\pi_1 - \pi_2$  için büyük örnekleme  $\%100(1-\alpha)$  güven aralığı:  $(p_1 - p_2) \pm z_{\alpha/2} \hat{\sigma}(p_1 - p_2)$  ile elde edilir.  $z_{\alpha/2}$  ;  $\alpha/2$  'ye eşit sağ-kuyruk olasılığına sahip standart normal yüzde birliğidir (  $\%95$  aralık için,  $\alpha=0.05$ ,  $z_{\alpha/2} = z_{0.025} = 1.96$  ).

**Tablo 2.7** Kalp Kası Bozukluğu (Myocardial İnfraction) Ve Aspirin Kullanımının Çapraz Sınıflandırılması

Kalp Kası Bozukluğu (MI)			
Grup	Evet	Hayır	Toplam
Yalancı ilaç	189= $n_{11}$	10,845= $n_{12}$	11,034
Aspirin	104= $n_{21}$	10,933= $n_{22}$	11,037

Kaynak: Agresti, 1996, s:20

Tablo 2.7' deki iki sıra bağımsız binom örneklemleri olarak dikkate alınır.  $N_1 = 11,034$  yalancı ilaç alanların sayısı olup 189 kişi inceleme süresince MI' ye yakalanmıştır. MI' ye yakalanma oranı  $p_1 = 189/11,034 = 0.0171$ .  $N_2 = 11,037$  aspirin alanların sayısı olup, 104 kişi MI' ye yakalanmıştır ve MI' ye yakalanma oranı  $p_2 = 104/11,037 = 0.0094$ . Oranların örnekleme farkı:  $0,0171 - 0,0094 = 0,0077$  olur. Bu farkın tahmin edilmiş standart hatası ise aşağıdaki şekilde bulunur:

$$\sqrt{\frac{(0.0171)(0.9829)}{11,034} + \frac{(0.0094)(0.9906)}{11,037}} = 0.0015$$

$\pi_1 - \pi_2$  şeklindeki gerçek farkın %95 güven aralığı:  $0.0077 \pm 1.96(0.0015)$  ya da  $0.008 \pm 0.003$  yada  $(0.005, 0.011)$ . Bu aralık sadece pozitif değerleri içerdiği için; sonuç olarak  $\pi_1 - \pi_2 > 0$  yani  $\pi_1 > \pi_2$ , şeklindedir ve aspirin alımı MI riskini azaltıyor gibi gözükmemektedir.  $i$ ' ninci sıradaki denekler için,  $\pi_{1|i}$  1.kategori ("başarı") için istenilen sonuca sahip yanıtların olasılığıdır. Sadece iki olası sonuca göre;  $\pi_{2|i} = 1 - \pi_{1|i}$  ve  $\pi_{1|i}$  için daha basit bir notasyon olan  $\pi_i$  ile gösterilsin.

Başarıların oranlarının farkı,  $\pi_1 - \pi_2$ ; iki sıranın ana karşılaştırmasıdır. Başarısızlıkların karşılaştırılması da aynı şekildedir (Agresti,2002):  $(1 - \pi_1) - (1 - \pi_2) = \pi_2 - \pi_1$ . Oranların farkı -1 ve +1 değerleri arasındadır. Sıralar eğer özdeş koşullu dağılımlara sahipse fark sıfırdır. Yanıt (Y),  $\pi_1 - \pi_2 = 0$  olduğunda sıraların sınıflandırılması istatistiksel olarak bağımsızdır. İki değişkende yanıt değişkeni ise, koşullu dağılımlar her ikisi içinde uygulanmalıdır. 1. sıradaki oranlar arasındaki fark aracılığıyla iki sütunda karşılaştırılabilir. Bu genellikle sıraları karşılaştıran  $\pi_1 - \pi_2$  farkına eşit olmaz.

### 2.2.2. Göreli Risk (Relative Risk, Oransal Risk)

$\pi_i$ , diğerlerine göre<sup>23</sup> 0 ya da 1'e yakınsa belirlenmiş bir  $\pi_1 - \pi_2$  fark değeri daha büyük bir öneme sahip olur. İlaç kullandığında ters reaksiyon gösteren deneklerin oranına bağlı olarak iki ilacın karşılaştırılmasını dikkate alalım. Ölen deneklerin oranına ilişkin iki denemenin karşılaştırılması şeklindeki bu incelemede 0.010 ve 0.001 arasındaki fark, 0.0410 ve 0.401 arasındaki farktan daha çok dikkate değer olabilir (her ikisi içinde farkın 0.009' a eşit olmasına rağmen). Bu şekildeki durumlarda oranların oranı aynı zamanda açıklayıcı niteliktedir (Agresti, 1996,s:21).

<sup>23</sup> Diğerleri aralığın ortasına daha yakındır.

Risk faktörüne maruz kalındığında ve kalınmadığında ki hastalık geliştirme risklerini (insidanslarını) karşılaştırır. Sadece kohort tipi gözlemsel çalışmalardan ve deneysel çalışmalardan hesaplanabilir. Asla geriye dönük çalışmalardan bulunamaz. İleriye dönük çalışmalar riskin doğrudan tahminin sağlayan tek gözlemsel çalışma tasarımıdır. Bu bakımdan deneysel çalışmalara en yakın olandır. Ancak yürütülmesi birçok zorluk içerir (Uğurman ve Berktaş, 2003: 238).

Görel risk, risk faktörüne maruz kalan hastalık insidansının, risk faktörüne maruz kalmayanlarda hastalık insidansına oranıdır. Dört gözlü tablodan şöyle bulunur<sup>24</sup>:

$$\text{Görel Risk} = \frac{a/(a+b)}{c/(c+d)}$$

Görel risk  $\frac{\pi_1}{\pi_2}$  oranı şeklinde tanımlanır<sup>25</sup>. Herhangi negatif olmayan gerçek bir sayı olmalıdır. Görel riskin 1 olması bağımsızlığın karşılığıdır<sup>26</sup>. Biraz evvel verilen oranlar için görel riskler;  $0.010/0.001=10.0$  ve  $0.410/0.401=1.02$  şeklindedir. İkinci yanıt kategorisine (başarısızlık) göre sıraların karşılaştırılması farklı bir görel risk değeri verir ( $\frac{1-\pi_1}{1-\pi_2}$ ). Bazen bu daha açıklayıcı nitelikte olabilir. Sonuçlardan hangisinin olasılığı küçükse normal olarak sonuca ilişkin olasılıkların oranı için o hesaba katılır (Agresti, 2002).  $p_1$  ve  $p_2$  şeklindeki örneklem oranları ile birlikte iki grup;  $\frac{p_1}{p_2}$  şeklinde örneklem görel risk oranına sahiptir. Örneklem dağılımı örneklem büyüklüğü yeteri kadar büyük olmadıkça oldukça çarpık olabilir. Bu nedenle güven aralığı formülü oldukça karmaşıktır. Tablo 2.7 için örneklem görel

24

25  $2 \times 2$  boyutlu olumsuzluk tablosunda, iki grup için başarı olasılıklarının oranına görel risk denir.

26  $\pi_1 = \pi_2$  ise görel risk=1 olur. Bu yanıt gruptan bağımsızsa gerçekleşir.

risk oranı;  $\frac{p_1}{p_2} = \frac{0.0171}{0.0094} = 1.82$ . MI durumlarının örneklem oranı, yalancı ilaç alanlar

açısından %82 daha fazladır.

Yazılımlar (SAS) kullanılarak gerçek görel risk oranı için %95 güven aralığı bulunabilir(1.43,2.30). %95 güvenle beş yıl sonra, MI durumlarının yalancı ilaç alanlara göre oranı aspirin alan MI oranının 1.43 ve 2.30 arasında bir katı kadar olur (Agresti, 2002). Görel risk için güven aralığı şunu göstermektedir ki, MI riski yalancı ilaç alanlar için en azından %43 daha fazladır. Oranların farkı için güven aralığı (0.005, 0.011) ise iki grup arasındaki farkı sanki önemsizmiş gibi göstermektedir. Fakat görel risk ise farkın halk sağlığı açısından önemli olabileceğini göstermektedir. İki grubu karşılaştırmak amacıyla yalnızca oranların farkını kullanmak, iki oranın her ikisi de sıfıra yaklaştığında oldukça yanıltıcı olabilir.

### 2.2.3. Odds Oranı

2×2 boyutlu olumsuzluk tablolarında birlikteliğin bir diğer ölçüm şeklide odds oranıdır (odds ratio). Daha sonraki bölümlerde ele alınacak modeller için temel bir parametredir(Agresti,1996,s:22). Başarının olasılığı ( $\pi$ ) için odds şu şekilde tanımlanır:

$$\Omega = \frac{\pi}{1-\pi} \quad 27$$

Başarı, başarısızlığa göre daha fazla olasılığa sahip olduğunda; odds negatif olamayan bir değerdir,  $\Omega > 1.0$ .

27

---

$$\text{Odds} = \frac{\text{Olayın olma olasılığı (p)}}{\text{Olayın olmama olasılığı (1-p)}}$$



$\pi = 0.75$  iken,  $\Omega = \frac{0.75}{0.25} = 3.0$  olur; başarı başarısızlığa göre üç kat daha

olasılıklıdır ve her bir başarısızlık için üç başarının gerçekleşmesi beklenir.  $\Omega = \frac{1}{3}$

ise, bir başarısızlık, başarıya kıyasla üç kat daha olasılıklıdır. Ters olarak,

$\pi = \Omega / (\Omega + 1)$  eşitliğine göre  $\Omega = \frac{1}{3}$  iken,  $\pi = 0.25$  olur. Tekrar  $2 \times 2$  boyutlu

olumsallık tablosuna dönelim.  $i$ ' ninci sıra içerisinde, başarısızlığın yerine başarının

olasılığı (odds):  $\Omega_i = \frac{\pi_i}{1 - \pi_i}$  şeklindedir. İki sıraya göre ,  $\Omega_1$  ve  $\Omega_2$  şeklinde ifade

edilen oddsların oranı; odds oranı (odds ratio) olarak adlandırılır:

$$\theta = \frac{\Omega_1}{\Omega_2} = \frac{\pi_1 / (1 - \pi_1)}{\pi_2 / (1 - \pi_2)}$$

Hücre olasılıklarının ( $\{\pi_{ij}\}$ ) ortak dağılımına göre,  $i$ ' ninci sıradaki odds için

eşdeğer bir tanımlama;  $\Omega_i = \frac{\pi_{i1}}{\pi_{i2}}$  ,  $i=1,2$  şeklinde yapılır. Böylece odds oranı

aşağıdaki gibi gerçekleşir:

$$\theta = \frac{\pi_{11} / \pi_{12}}{\pi_{21} / \pi_{22}} = \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}}$$

Her iki değişkende yanıt değişkeni ise odds oranı ortak olasılıklar kullanılarak belirlenir. Bu şekilde elde edilen  $\theta$ 'ya “*çapraz-çarpım oranı*” da denmektedir.

Çünkü hücrelerin çapraz olarak, karşılıklı çarpımlarının ( $\pi_{11}\pi_{22}$  ve  $\pi_{12}\pi_{21}$ ) oranına eşittir (Agresti,2002).

Sıralar için koşullu dağılımlar özdeşse (yani;  $\pi_1 = \pi_2$ ), oddslar  $\Omega_1 = \Omega_2$  şeklindedir. Bu nedenle değişkenler bağımsızdır<sup>28</sup>. Görelî risk ilgilenilen bir özelliğe ilişkin iki olasılığın oranı iken oddsların bir oranıdır.

#### 2.2.4. Odds Oranının Özellikleri

Odds oranı herhangi bir negatif sayıya eşit olamaz. X ve Y birbirinden bağımsız ise ( $\pi_1 = \pi_2$ );  $\Omega_1 = \Omega_2$  ve  $\theta = \Omega_1 / \Omega_2 = 1$  olur ve bu durum bağımsızlığın bir karşılığı olup, karşılaştırma için bir ölçü olarak kullanılır (Agresti, 1996, s:23).  $1 < \theta < \infty$  ise, birinci sıradaki denekler ikinci sıradakilere nazaran başarı açısından daha yüksek bir olasılığa sahiptir,  $\pi_1 > \pi_2$ . Örnek olarak  $\theta = 4$  için birinci sıradaki başarının odds'u; ikinci sıradakine göre tam dört kat daha fazladır. Bu  $\pi_1 = 4\pi_2$  olduğu anlamına gelmez.  $\pi_1 = 4\pi_2$  ifadesi görelî riskin 4 olması durumundaki yorumdur.  $0 < \theta < 1$  ise,  $\pi_1 < \pi_2$  olur. Bir hücre sıfır olasılığına sahipse  $\theta$ ,  $\infty$ ' a ya da  $0$ ' a eşit olur. Belirlenmiş bir yönde  $1.0$ ' den daha yüksek  $\theta$  değerleri daha güçlü birlikteliği (ilişkiyi) göstermektedir. Biri diğerinin tersi olan iki değer aynı birlikteliği gösterir fakat ters yönlü bir birlikteliktir. Örnek olarak;  $\theta = 0.25$  olduğunda; birinci sıradaki başarının oddsu ikinci sıradakinin  $0.25$  katıdır ya da aynı şekilde ikinci sıradaki başarının oddsu birinci sıradakinin  $1/0.25=4.0$  katıdır. Sıraların yerleri ters çevrildiğinde ya da sütunların yerleri değiştiğinde  $\theta$ ' nin değeri orijinal değerinin tersi olur.

Yorum yapmak için  $\log \theta$  kullanmanın uygun olduğu göz önüne alınmalıdır. Bağımsızlığın varlığı ancak  $\log \theta = 0$  olduğunda söz konusudur. Sıra ve sütunların

---

<p><sup>28</sup></p> <p>Riske maruz kalanlarda hastalık odds'u</p> <p>Riske maruz kalmayanlarda hastalık odds'u ise</p>	$= \frac{a/(a+b)}{b/(a+b)} = a/b$ $= \frac{c/(c+d)}{d/(c+d)} = c/d$
---	---

Bu oranlar formülde pay ve paydaya konduğunda:

$$OR = \frac{a/b}{c/d} = \frac{axd}{bxc} \text{ olarak yazılabilir.}$$

“Olgu-kontrol” çalışmaları için geçerli şartlı olasılıklar, prospektif ya da deneysel çalışmalar için geçerli olanlardan farklı olsa da basitleştirilmiş hesaplama şekli olan  $a \cdot d / b \cdot c$  her ikisi için de doğrudur. Eğer bağımlı gruplar söz konusu ise  $OR = b / c$  olur (Uğurman, ve Berktaş, 2003: 237).

ters çevrilmesi ile deęerin iřaretinin deęiřmesi sonucunu ortaya ıkartır ki dolayısı ile log-odds oranı simetriktir(Agresti,1996,s:24). rnek olarak iki log  $\theta$  deęeri (iřaretleri farklı ama deęerleri aynı;  $\log 4=1.39$  ve  $\log 0.25=-1.39$  řeklinde) birliktelik aısından aynı gtedir. Tablonun yn ters evrildięinde (Stunları yanıt deęiřkeni ve sıraları aıklayıcı deęiřken olarak iřleme alsak ya da sıraları yanıt deęiřken ve stunları aıklayıcı deęiřken olarak iřleme alsak) bile bu deęer deęiřmemektedir. nk odds oranı deęiřkenleri simetrik bir biimde iřleme tabi tutar. Bylece ařaęıdaki řekilde de ifade edilen odds oranının simetrik olduęu aıka grlmektedir:

$$\theta = \frac{\pi_{11} / \pi_{12}}{\pi_{21} / \pi_{22}} = \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}}$$

$\theta$  kullanıldıęı zaman hesaplama yapmak amacıyla yanıt deęiřken řeklindeki sınıflandırmanın tanımlanması gerekmemektedir. Bu durumun aksine greli risk bunu gerektirmektedir ve deęeri de birinci ya da ikinci yanıt kategorisinin kullanımına baęlı olmaktadır. Gerekte odds oranı;  $\theta = \frac{\Omega_1}{\Omega_2} = \frac{\pi_1 / (1 - \pi_1)}{\pi_2 / (1 - \pi_2)}$  řeklinde<sup>29</sup> olduęunda sınıflandırma tanımlanmıř olmasına raęmen, kořullu olasılıkların ters evrilmesi ile sadece biri tanımlanır. Ortak daęılım ile kořullu daęılımlar her bir yn iinde mevcuttur ve ařaęıdaki gibi ifade edilir(Agresti,2002):

$$\begin{aligned} \theta &= \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}} = \frac{P(Y = 1 | X = 1) / P(Y = 2 | X = 1)}{P(Y = 1 | X = 2) / P(Y = 2 | X = 2)} \\ &= \frac{P(X = 1 | Y = 1) / P(X = 2 | Y = 1)}{P(X = 1 | Y = 2) / P(X = 2 | Y = 2)} \end{aligned}$$

Aslında ileriye dnk ( prospective), geriye dnk (retrospective) ya da apraz paralı rneklem tasarımları aısından aynı durum geerlidir ki rneklem odds oranı her durum iin kullanılan parametrelerin tahminlenmesidir. Hcre sayıları

---

<sup>29</sup>  $\pi_i = P(Y = 1 | X = i)$  kullanan odds iin.

$\{n_{ij}\}$  için örneklem odds oranı<sup>30</sup>  $\hat{\theta} = n_{11}n_{22} / n_{12}n_{21}$  şeklindedir. Her hangi bir sıradaki hücre sayıları sıfırdan farklı bir sabit ile çarpıldığında ya da herhangi bir sütundaki hücre sayıları sıfırdan farklı bir sabitle çarpıldığında yukarıdaki şekilde ifade edilen tahmini odds oranı değişmez. Bir anlamı da şudur ki örneklem orantısız olarak bir değişkenin marjinal kategorilerinden küçük ya da büyük olsa bile örneklem odds oranı benzer özelliği ( $\theta$ ) tanımlar.

Aşı yaptırma ile belirlenmiş grip türüne yakalanma arasındaki birlikteliğin (ilişkinin) geriye dönük incelenmesinde örneklem odds oranı; 100 gribe yakalanmış ve 100 yakalanmamış kişiden oluşan rassal örneklem ya da 40 kişilik gribe yakalanmış ve 160 yakalanmamış kişilerden oluşan rassal bir örneklem ile aynı karakteristiği tahminler. Oranların farkının ve görel riskin (relative risk) örneklemden hesaplanması sıralar içindeki sayıların bir sabit ile çarpılması durumunda değişmez. Fakat sütunlardaki çarpım ya da sıra- sütun değişimi ile değişmektedir.

### 2.2.5. Olgu-Kontrol İncelemeleri ve Odds Oranı

Geriye dönük örnekleme tasarımları ile (olgu-kontrol incelemeleri gibi)  $P(X=i|Y=j)$  şeklindeki koşullu olasılıkların tahmin edilmesi mümkündür. Genellikle ilgilenilen bir sonucun olasılığının,  $P(Y=j|X=i)$  ya da ilgilenilen sonuç için oranlar farkının ya da görel riskin tahminlenmesi mümkün olmayabilir.

Ancak odds oranının tahmin edilmesi mümkündür. Her iki yön içindeki koşullu olasılıklar aracılığı ile belirlenir. Örnekle açıklamak amacıyla, daha önce ele aldığımız akciğer kanserine ilişkin örneğe geri dönelim.  $X$ = sigara içme tutumu ve  $Y$ = akciğer kanseri şeklinde belirlenmişti. Veriler, belirli  $Y$  seviyelerine göre  $X$ ' e ilişkin iki binom örneklemeinden elde edilmiştir. Böylece deneğin akciğer kanseri

---


$$^{30} \hat{\theta} = \frac{p_1 / (1 - p_1)}{p_2 / (1 - p_2)} = \frac{n_{11} / n_{12}}{n_{21} / n_{22}} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$$

Bu aynı zamanda standart örneklem düzenine göre; gerçek odds oranının ML tahmincisidir.

olup olmamasına ilişkin belirli bir sonuç için deneğin sigara içme olasılığını tahmin edebiliriz; hasta grup (olgular) için 688/709 ve kontrol gurubu için 650/709 şeklindedir.

Sigara içip içmemesine ilişkin belirli bir sonuca göre akciğer kanseri olma olasılığını tahmin edemeyiz. Bu nedenle de akciğer kanseri olasılıklarının oranları ya da farklarını da hesaplayamayız. Oranların farkları ve görel risk, sigara içme olasılıklarının karşılaştırılması ile sınırlıdır. Fakat çapraz çarpım oranına (koşullu olasılıkların çarpımı) benzer olarak örneklem için odds oranı hesaplanabilir:

$$\frac{(688/709)/(21/709)}{(650/709)/(59/709)} = \frac{688 \times 59}{650 \times 21} = 3.0.$$

Üstelik inceleme geriye dönük olsa bile bu formül aracılığı ile hesaplama sonucuna göre yapılan yorumlamalar ilgilenilen yönde fayda sağlar: Sigara içenler için akciğer kanseri olmanın tahminlenmiş olasılığı (odds) sigara içmeyenlerin akciğer kanseri olma olasılığının 3.0 katıdır.

### 2.2.6. Odds Oranı ve Görel Risk Arasındaki İlişki

Odds oranı ve görel risk için verilen formüllerden<sup>31</sup> aşağıdaki sonuç elde edilir (Agresti, 1996):

$$OddsOranı = \frac{p_1/(1-p_1)}{p_2/(1-p_2)} = GörelRisk \times \left( \frac{1-p_2}{1-p_1} \right)$$

$$OddsOranı = GörelRisk \times \left( \frac{1-\pi_2}{1-\pi_1} \right)$$

---

<sup>31</sup> Görel Risk;  $\frac{\pi_1}{\pi_2}$  oranı şeklinde tanımlanır.

Odds oranı ise:  $\theta = \frac{\Omega_1}{\Omega_2} = \frac{\pi_1/(1-\pi_1)}{\pi_2/(1-\pi_2)}$

Odds oranının 1.83 olması  $p_1$ ' in  $p_2$ ' nin 1.83 katı olduğu anlamına gelmemektedir. Bu ancak bir görel risk yorumudur. Çünkü ölçüm olasılıklardan (odds) daha çok oranlarla ilgilidir.  $\hat{\theta}=1.83$  değeri; olasılık (odds) değeri  $p_1/(1-p_1)$ 'in, olasılık değeri  $p_2/(1-p_2)$ 'nin 1.83 katı olduğunu göstermektedir. İlgilenilen sonucun olasılığı  $\pi_i$  (başarı oranları) her iki grup içinde sıfıra yakın olduğunda her iki oranında büyüklüğü birbirine benzer. Kalp kası bozukluğu ile ilgili örnek bu benzerliği göstermektedir. Her bir grup için, kalp kası bozukluğu durumlarının örneklem oranı sıfıra yakındır (kalp krizi oranının her grup için 0.02' den daha küçük olması durumunda). Bu nedenle de örneklem odds oranı (1.83), örneklem görel risk (1.82) ile aynıdır. Böyle bir durumda odds oranı 1.83;  $p_1$ ' in  $p_2$ ' nin 1.83 katı olduğu anlamına gelebilmektedir.

Odds oranı ile görel risk arasındaki bu ilişki oldukça faydalıdır. Çünkü bazı veri setleri için görel riski hesaplamak mümkün olmayabilir. Bu durumda odds oranı hesaplanarak görel riskin yaklaşık değeri olarak kullanılır. Her bir  $\pi_i$  değeri küçük olduğunda sağlanan bu benzerlikten dolayı, olgu-kontrol incelemelerinde olduğu gibi (Agresti, 2002) görel risk dolaysız olarak tahmin edilemediğinde odds oranı görel riskin kabaca bir ölçümünü elde etmemizi de sağlar. Belirli Y değerlerine göre X'in koşullu dağılımını kullandığımızda odds oranı, belirli X değerlerine göre Y' nin dağılımının kullanılması ile aynı değeri alır. Bu nedenle değişkenler simetrik bir şekilde dikkate alınır. Odds oranı, her iki yönde de koşullu dağılımlar aracılığıyla belirlenmektedir ve her bir Y seviyesine göre X' teki yanıtın (tepkinin) ölçülmesi amacıyla bir inceleme planımız olsa dahi hesaplanabilmektedir. X'in her değeri için Y=1 olasılığı küçük olduğunda odds oranının ve görel riskin benzer değerler aldığına dikkat edilmelidir. Yalnızca belirli Y değerleri için X' in koşullu olasılıklarını tahmin etsek bile, eğer  $P(Y=1|X)$  değerinin küçük olması bekleniyorsa, örneklem odds oranını görel riskin kaba bir ölçümü olarak kullanabiliriz.

Tablolarda her bir sütun, her bir sıranın aksine bir binom örneklemi olarak işleme dahil edilmektiğinde eşleştirmeden dolayı (olgu-kontrol incelemelerinde oluşturulan) iki sütundaki binom örnekleri bağımsızlıktan daha ziyade bağımlı

nitelikte olur. Birinci sütundaki her bir gözlem doğal olarak ikinci sütunda yer alan iki gözlemle eşleştirilmiştir.

### 2.2.7. Odds Oranı ile Logaritmik Odds Oranına Göre Yorumlama

Küçük örnek hacimleri için odds oranı oldukça çarpıktır (n yeteri kadar büyük olmadığından).  $\theta = 1$  olduğunda  $\hat{\theta}$  değeri  $\theta$  değerinden daha küçük bir değer alamaz (Çünkü  $\hat{\theta} \geq 0$ ). Fakat ihmal edilemeyecek bir olasılıkla daha büyük bir değer alabilir. Bu çarpıklıktan dolayı odds oranı için istatistiksel yorumlamada alternatif fakat eş değer bir ölçü kullanılmaktadır ki bu ölçü odds oranının doğal logaritmasının alınması ile elde edilir,  $\log(\theta)$  (Agresti,1996, s:24).

Bağımsızlık  $\log(\theta) = 0$  olduğunda söz konusu olmaktadır. Odds oranının 1 olması, log-odds oranının 0 olmasıdır. Odds oranı 2 ise log-odds oranı 0.7 olur. Log-odds oranı sıfır etrafında simetriktir ki bu nedenle de sıra ya da sütunların ters çevrilmesi durumunda işareti değişir.  $\log(2.0) = 0.7$  ve  $\log(0.5) = -0.7$  şeklinde iki log odds oranı işaretleri dışında aynı değeri almaktadır. İşaretleri farklı olsa da aynı ilişki derecesini ifade ederler. Bir log-odds oranının iki katı; odds oranının karesinin karşılığıdır. Örnek olarak; log-odds oranı 0.7 için  $2(0.7)=1.4$  ve -0.7 için  $2(-0.7)=-1.4$  değerleri odds oranı açısından sırasıyla  $2^2=4$  ve  $0.5^2=0.25$  değerlerinin karşılığıdır. Örnek odds oranının logaritmik dönüşümü,  $\log(\hat{\theta})$ , daha az çarpık bir örneklem dağılımına sahiptir ki bu dağılım normalliğe daha çabuk yakınsar.

$\log(\hat{\theta})$  için tahmin edilen standart hata asimptotik standart hata olarak adlandırılır ve ASE ile gösterilir (Agresti, 1996,s:24):

$$ASE(\log \hat{\theta}) = \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}$$

Hücrelerdeki sıklık sayıları arttıkça ASE değeri azalır. Çünkü örneklem dağılımı normalliğe daha yakındır. Buna dayanarak,  $\theta$  için güven aralığının

oluşturulması için en iyi yol  $\log(\theta)$  için güven aralıklarının kurulması, sonra bulunan değerlerin anti-logaritmasının alınmasıdır (Üssel fonksiyon kullanılarak anti-log alınır).  $\log(\theta)$  için büyük örnek güven aralığı  $\log(\hat{\theta}) \pm z_{\alpha/2} ASE(\log \hat{\theta})$  şeklindedir. Bu güven aralığının sınır değerlerinin üssü alınarak  $\theta$  için bir güven aralığı elde edilir. Daha önce oranların farkı konusunda örnek olarak kullanılan aspirin kullanımı ve kalp kası bozukluğu arasındaki ilişkileri gösteren tablo 2.7' deki verilere göre;  $\hat{\theta}$ 'nın doğal logaritması,  $\log(1.83)=0.605$  değerine eşittir.

$\log(\hat{\theta})$ 'nin ASE değeri,  $(1/189+1/10.933+1/10.845+1/104)^{1/2} = 0.123$  olarak bulunur. Bu örnek, popülasyonu;  $\log(\theta)$  için %95 güven düzeyinde  $0.65 \pm 1.96(0.123)$  yani  $(0.365, 0.845)$  güven aralığında temsil etmektedir.  $\theta$  için güven aralığı ise  $[\exp(0.365), \exp(0.846)] = (e^{0.365}, e^{0.846}) = (1.44, 2.33)$  şeklinde bulunur.  $\theta$  için hesaplanan güven aralığı 1 değerini içine almadığından dolayı MI'nın gerçek olasılığı (odds'u) iki grup arasında farklılık göstermektedir. Bu aralığa göre şu tahminleme yapılır: Yalancı ilaç alan deneklerin MI'ya yakalanma olasılığı (odds'u), aspirin alan deneklerin yakalanma olasılığından en az %44 daha fazladır.  $\hat{\theta}$ 'nin örnekleme dağılımının sağa çarpık olmasından dolayı; aralığın sınır değerleri  $\hat{\theta} = 1.83$  değerinden eşit uzaklıklarda değildir.

Herhangi bir hücredeki örnek hacmi 0 ise,  $n_{ij} = 0$  örnek odds oranı  $\hat{\theta}$  ya "0" ya da " $\infty$ " değerine eşit olur. ayrıca bir sıra ya da sütundaki kayıtlardan her ikisi de sıfır ise bu oran tanımsızdır. Her bir hücredeki örnek hacmine  $\frac{1}{2}$  eklenerek tahmin edicinin çok az düzeltilmesi her hangi bir probleme neden olmayacaktır (Agresti, 2002):

$$\tilde{\theta} = \frac{(n_{11} + 0.5)(n_{22} + 0.5)}{(n_{12} + 0.5)(n_{21} + 0.5)}$$



Hücre sayıları çok küçük olması ya da herhangi bir hücre sayısının sıfır olması durumunda bu düzeltmenin kullanılması önerilir. Bu durumda ASE formülünde  $\{n_{ij}\}$  yerine  $\{n_{ij} + 0.5\}$  yazılır. Tekrar tablo-2.8' ye dönersek  $\tilde{\theta} = (189.5)(10,933.5)/(10,845.5)(104.5) = 1.828$  olarak bulunur ki bu değer  $\hat{\theta} = 1.83$  değerine yakındır. Bunun nedeni de hiçbir hücre sayısının küçük olmamasıdır.

### **2.3. TABAKALI 2x2 BOYUTLU OLUMSALLIK TABLOLARINDA KISMI BİRLİKTELİK**

Çoğu incelemenin önemli bir kısmı (özellikle gözlemsel incelemelerin) kontrol değişkenlerinin seçiminden oluşmaktadır. X' in Y üzerindeki etkisini incelerken ilişkiyi etkileyen her hangi bir ortak değişken kontrol edilmelidir(Agresti,1996,s:53). Ortak değişken sabitinin elde edilmesi bazı tekniklerin kullanılmasını gerektirmektedir. Diğer taraftan X' in Y üzerinde gözlemlenen etkisi gerçekte hem Y hem de X için ortak bir değişkenin etkisini yansıtabilir. Bu durumda X ve Y arasındaki ilişki “etki değiştirici” (confounding) şeklinde bir durumu gösterir.

Gözlemsel incelemeler sapmaya (hata ya da bilerek veya bilmeden taraf tutma) daha açıktır. Örneğin, yaş ve cinsiyet yönünden benzer iki grup oluşturduğumuzu ancak daha sonra bir grupta siyah ırktan olanların daha fazla ve bu grupta ayrıca hastalık şiddetlerinin daha ağır olduğunu fark ettiğimizi varsayalım. Buradaki “ırk” ve “hastalık şiddeti” değişkenleri bizim izlemediğimiz değişkenlerdir. Ancak sonuçta gözlenen etkinin nedenleri arasında bu iki değişkenin de rolü olabilir. Bunlara etki değiştirici değişkenler denir. Deneysel incelemeler, X' in farklı seviyeleri için denekleri rassal olarak belirleme yoluyla “etki değiştirici değişkenlerin” etkilerini ortadan kaldırır. Fakat gözlemsel incelemelerde bu mümkün değildir<sup>32</sup>. Varsayalım ki incelemede pasif sigara içiminin etkileri (sigara

---

<sup>32</sup> Gözlemsel incelemelerde Olguların hangi grupta yer alacakları kendiliğinden bellidir (biz seçemeyiz). Risk faktörlerine değişik düzeyde maruz kalan bu grupların bu maruziyete karşı yanıtları gözlemlenir (bu sürece araştırmacının aktif müdahalesi söz konusu değildir). İleriye yönelik

içenlerle yaşayan sigara içmeyen kişiler üzerindeki etkiler) dikkate alınsın. Pasif sigara içiminin akciğer kanseri ile ilişkili olup olmadığını analiz etmek için, çapraz parçalı bir inceleme ile sigara içmeyenler (eşleri sigara içen ve içmeyenler şeklinde gruplanırlar) arasındaki kanser oranları karşılaştırılabilir.

İncelemede yaş, sosyoekonomik durum ya da diğer faktörler (sigara içen eş ve akciğer kanseri gelişiminin her ikisi ile de ilişkili olabilecek) kontrol altına alınabilir. Ayrıca sonuçlar sınırlı fayda sağlayabilir. Sigara içmeyenlerin eşleri, sigara içenlerin eşlerinden daha genç olma eğiliminde olabilir ve daha genç kişiler kansere yakalanma açısından daha az bir olasılığa sahiptir. Sigara içmeyenlerin eşleri arasında akciğer kanserine yakalananların oranının daha düşük olması sadece onların daha düşük yaş ortalamasına sahip olduğunu yansıtabilir. Burada etki değiştirici nitelikte olabilen Z gibi bir değişken kontrol edilirken X ve Y gibi kategorik değişkenlerin arasındaki birlikteliğin (ilişkinin) analiz edilmesi ele alınacaktır. Basitleştirmek için tek bir kontrol değişkenine sahip örnekler incelenecektir.

### 2.3.1. Kısmi Tablolar

Z' nin belirli seviyelerinde XY ilişkisini incelemek aracılığı ile Z kontrol altına alınır. Üç yönlü çapraz tabloların iki yönlü çapraz parçalı kısımları, Z' nin ayrı kategorilerinde X ve Y' yi çapraz olarak sınıflandırır. Bu çapraz bölümler **kısmi tablolar** olarak adlandırılır. Kısmi tablolar Z' nin etkisini, Z' nin değerini sabit tutarak ortadan kaldırırken aynı zamanda XY ilişkisini de göstermektedir. Kısmi tabloların birleştirilmesi ile elde edilen iki yönlü çapraz tabloya **XY marjinal tablosu** denmektedir. Marjinal tablodaki her bir hücre sayısı, kısmi tablolarda aynı bölgedeki sayıların toplamıdır. Marjinal tablo Z hakkında bir bilgi içermemektedir. X ve Y ile ilgili olarak iki yönlü tablo oldukça sadedir. Fakat Z' nin X ve Y üzerindeki etkilerini de yansıtabilmektedir. Kısmi tablolardaki birliktelikler **koşullu birliktelikler** olarak adlandırılır. Çünkü herhangi bir seviyede sabitlenen Z koşuluna bağlı olarak, X' in

---

(prospektif) incelemeler, kohort incelemeleri, olgu-kontrol çalışması, geriye dönük (retrospektif) incelemeler “gözlemsel” incelemelerdir.

Y üzerindeki etkisine dayanmaktadır. Kısmi tablolardaki koşullu birliktelikler marjinal tablolardakinden oldukça farklı olabilmektedir. Gerçekten çok yönlü bir tablonun yalnızca marjinal tablolarını analiz etmek yanıltıcı olur (Agresti, 1996, s:54). Aşağıdaki örnekler bu durumu açıklamaktadır.

### 2.3.2. Ölüm Cezası Örneği

Tablo-2.8 katil olduğu kanıtlanmış kişilerin ölüm cezası alıp almaması konusunda ırksal özelliklerin etkisinin incelendiği bir makaleden alınan  $2 \times 2 \times 2$  boyutlu bir olumsuzluk tablosudur (iki sıra, iki sütun ve iki tabakadan oluşmaktadır) (Agresti, 1996, s:54-57). Bu şekildeki tablolar kontrol değişkeninin dahil edildiği tablolardır. Tabloda sınıflandırılan 674 denek 1976 ve 1987 yılları arasında Florida’da birçok cinayet ile ilgili delilleri kapsayan iddianamelerde yer alan katillerdir. Tablodaki değişkenler; Y= Ölüm cezası kararı (“evet, hayır” şeklinde kategorilere sahiptir), X= Katillerin ırkı ve Z= Kurbanların ırkı ( X ve Z değişkenlerinin her biri “beyaz, siyah” şeklinde iki kategoriye sahiptir.) şeklindedir. Kurbanların ırkı bir kontrol değişkeni olarak işleme tabi tutulmakta ve ölüm cezası kararı üzerinde katilin ırkının etkisi incelenmektedir. Tablo, kurbanların ırk kategorilerinin her birinde ölüm cezası kararı ve katilin ırkı ile ilgili bir  $2 \times 2$  kısmi tabloya sahiptir. Kurbanların ırkı ve katillerin ırkı arasındaki her bir kombinasyona (birleşime) göre, ölüm cezası alan katillerin yüzdesi tabloda listelenmiş ve şekil-2.1’ de gösterilmiştir. Bu yüzdeler koşullu birliktelikleri tanımlamaktadır.

Kurbanlar beyaz olduğunda, siyah katiller için ölüm cezası beyaz katillerden %22.9-%11.3=%11.6 oranında daha sık uygulanmıştır. Kurbanlar siyah olduğunda, beyaz katiller için ölüm cezası siyah katillerden %2.8 oranında daha sık uygulanmıştır. Kurbanların ırkı sabitlenerek kontrol altına alındığında ise ölüm cezası, siyah katillere beyaz katillerden daha az uygulanmıştır. Tablonun alt kısmında marjinal tablo gösterilmiştir. Bu marjinal tablo kurbanların ırkının iki kategorisine göre hücre sayılarının toplanmasından meydana gelmektedir ki nitekim iki kısmi tablonun birleşimidir ( $11+4=15$  gibi). Beyaz katillerin %11.0’ i

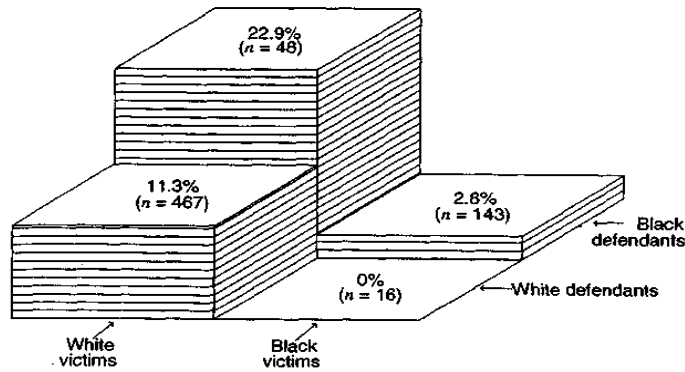
( $53/53+430=53/483=0.109$ ) ve siyah katillerin %7.9' u ( $15/176+15=15/191=0.0785$ ) ölüm cezası almıştır.

**Tablo 2.8** Katilin Irkına Ve Kurbanların Irkına Göre Ölüm Cezası Kararı

		Ölüm Cezası		
Kurbanların Irkı	Katillerin Irkı	Evet	Hayır	Evet Yüzdesi
Beyaz	Beyaz	53	414	11.3
	Siyah	11	37	22.9
Siyah	Beyaz	0	16	0.0
	Siyah	4	139	2.8
Toplam	Beyaz	53	430	11.0
	Siyah	15	176	7.9

Kaynak: Agresti, 1996, s:54

**Şekil-2.1** Ölüm Cezası Alanların Yüzdesi



Kurbanların ırkını göz ardı ettiğimizde, ölüm cezası siyah katillere daha az uygulanmıştır (beyaz katillerden daha az). Birliktelik kısmi tablolarla karşılaştırıldığında zıt yöndedir. Kurbanların ırkı önemsenmediğinde, birliktelik

neden bu kadar değişmektedir? Bu kurbanların ırkı ile diğer değişkenlerin her biri arasındaki birlikteliğin yapısı ile ilgilidir.

*İlk olarak*, kurbanların ırkı ile katilin ırkı arasındaki birliktelik son derece güçlüdür. Bu değişkenlerle ilgili marjinal tablo için odds oranı;  $(467 \times 143) / (48 \times 16) = 87.0$ <sup>33</sup>. *İkinci olarak*; tablo 2.8 katillerin ırkını dikkate almaksızın, kurbanların beyaz olması durumunda verilen ölüm cezasının(64), siyah olması durumunda verilen ölüm cezasından (4) daha fazla olasılıkla gerçekleşmiş olduğunu gösterir. Buna göre beyazlar, beyazları öldürme eğiliminde olacaktırlar ve beyazları öldürenlerin ölüm cezası alması olasılığı daha fazladır. Buna göre marjinal birliktelik, koşullu birlikteliklere kıyasla beyaz katillerin ölüm cezası alma konusunda daha büyük bir eğilime sahip olduğunu gösterir. Tablo 2.8 bu yapıya sahiptir.

Sonuç olarak: Marjinal birliktelik, koşullu birlikteliklerin her birinden farklı bir yöne sahiptir. Buna ***Simpson Paradoksu*** (çelişkisi) adı verilir (Simpson 1951, Agresti, 1996, s:57). Bu durum kategorik değişkenler kadar kantitatif değişkenler için de kullanılır. İstatistikçiler genel olarak, X ile Y'nin birlikteliğinden kaynaklanan etkilerin verilmesinde dikkat etmek için bunu kullanırlar. Örnek olarak, doktorlar akciğer kanseri ile sigara içimi arasındaki güçlü olasılık oranlarını gözlemlemek ile başlarken, R.A. Fisher gibi istatistikçiler, kontrol altında ortadan kalkan birlikteliği var eden bazı değişkenlerin (genetik faktör gibi) olabileceği konusunda uyarıda bulunurlar. Fakat J.Cornfield gibi bazı istatistikçiler; kontrol altında meydana gelen değişimin ya da yok olmanın etkisini görebilmek için çok güçlü bir XY birlikteliği ile birlikte, Z etki değiştirici değişkeni ve X ile Y' nin her ikisi arasında çok güçlü bir birlikteliğin var olması gerektiğini savunur (Breslow ve Day 1980, Böl-3).

---

<sup>33</sup>  $414+53=467$  beyaz kurbanları öldüren beyaz katillerin toplam sayısı (ceza alanlar ile almayanların toplamı) ve  $139+4=143$  siyah kurbanları öldüren siyah katillerin sayısı (ceza alanlar ile almayanların toplamı)  $11+37=48$  beyaz kurbanları öldüren siyah katillerin toplam sayısı (ceza alan ve almayanların toplamı)  $16+0=16$  siyah kurbanları öldüren beyaz katillerin toplam sayısıdır.

### 2.3.3. Koşullu ve Marjinal Odds Oranları

Odds oranları marjinal ve koşullu birliktelikleri tanımlamaktadır.  $2 \times 2 \times K$  tabloları için durumu örnekle açıklayalım.  $\{\mu_{ijk}\}$ , binom, çokterimli ya da Poisson örnekleme şeklindeki bazı örnekleme modellerinde beklenen hücre frekanslarını göstermektedir.  $Z$ ' nin belirli bir  $k$  kategorisi içerisinde, odds oranı aşağıdaki gibidir ve  $k$  kısmi tablosu içinde  $XY$  koşullu birlikteliğini gösterir (Agresti, 1996, s:57):

$$\theta_{XY(k)} = \frac{\mu_{11k}\mu_{22k}}{\mu_{12k}\mu_{21k}}$$

$k$  tane kısmi tablo için elde edilen odds oranları,  $XY$  koşullu odds oranları olarak adlandırılır ve marjinal odds oranlarından oldukça farklıdır.  $XY$  marjinal tablosu,  $\{\mu_{ij+} = \sum_k \mu_{ijk}\}$  şeklindeki beklenen frekansları verir.  $XY$  marjinal odds oranı aşağıdaki gibidir (Agresti, 1996, s:57):

$$\theta_{XY} = \frac{\mu_{11+}\mu_{22+}}{\mu_{12+}\mu_{21+}}$$

$\theta_{XY(k)}$  ve  $\theta_{XY}$ ' nin örneklem değerleri beklenen frekansların yerine hücre sayıları ile benzer formüllerden yararlanılarak bulunurlar. Tablo 2.8' deki ölüm cezası ile katillerin ırkları arasındaki birlikteliği açıklayalım. İlk kısmi tabloda, kurbanların ırkları beyazdır ve buna göre marjinal odds oranının tahmini aşağıdaki gibidir:

$$\hat{\theta}_{XY(1)} = \frac{53 \times 37}{414 \times 11} = 0.43$$

Ölüm cezası alan beyaz katillerin örneklem oddsları siyah katillerin örneklem oddslarının %43' üdür. İkinci kısmi tabloya göre, kurbanların ırkı siyahtır ve tahminlenmiş odds oranı  $\hat{\theta}_{XY(2)} = (0 \times 139) / (16 \times 4) = 0.0$  şeklindedir. Çünkü siyah

kurbanları öldüren beyaz katillere hiçbir zaman ölüm cezası verilmemiştir. Marjinal odds oranının tahmininde tablo 2.8' de yer alan  $2 \times 2$  boyutlu marjinal tablodan yararlanılır (kurbanların ırkları aracılığı ile çökertilir.) ya da  $(53 \times 176)/(430 \times 15) = 1.45$  kullanılır. Ölüm cezasının örneklem oddsu beyaz katiller açısından siyah katillere kıyasla %45 daha yüksektir. Oysa kurbanların kategorilerinin her biri içerisinde bu oddslar beyaz katiller için daha azdır. Kurbanların ırkları kontrol altına alındıktan sonra birlikteki bu terse dönme durumu Simpson Paradox' unu örnekleyerek açıklamaktadır.

### 2.3.4. Koşullu Bağımsızlığa Marjinal Bağımsızlık

X, I kategorilerine ve Y' de J kategorilerine sahiptir.  $I \times J \times K$  tablosu, Z' nin kontrol altına alınması ile X ve Y arasındaki ilişkiyi ifade etmektedir ve Y değişkeni k tane kısmi tablo içerisinde bağımsızdır ve buna göre X ve Y, Z' nin k seviyesinde koşullu olarak bağımsızdır denir. Y bir yanıt ise, bunun anlamı aşağıdaki gibidir:

$$P(Y = j | X = i, Z = k) = P(Y = j | Z = k), \text{ tüm } i, j' \text{ ler için}$$

Daha genel olarak, X ve Y belirli bir Z değerine göre koşullu olarak bağımsızdır diye ifade edilir. Yani, bütün k' lar için yukarıdaki formül hesaplanır. Belirli bir Z için Y, X' e bağlı değildir. Varsayalım ki, tek bir çokterimli  $\{\pi_{ijk} = P(X = i, Y = j, Z = k)\}$ , ortak olasılıkları ile birlikte üç yönlü bir tabloya uygulansın.

Sonra  $\pi_{ijk} = P(X = i, Z = k)P(Y = j | X = i, Z = k)$  eşitliği belirli bir Z için, X ve Y' nin koşulu bağımsızlığı halinde aşağıdaki formüle eşit olmaktadır:

$$= \pi_{i+k}P(Y = j | Z = k) = \pi_{i+k}P(Y = j, Z = k) / P(Z = k)$$

Böylece koşullu bağımsızlık aşağıdaki formüle eşit olmaktadır (Agresti, 1996, s:59):

$$\pi_{ijk} = \pi_{i+k} \pi_{+jk} / \pi_{++k} \quad \text{Tüm } i, j \text{ ve } k' \text{ lar için}$$

Koşullu bağımsızlık, marjinal bağımsızlık anlamına gelmemektedir. Örnek olarak, son eşitlik k üzerinden toplanarak aşağıda ki eşitlik sağlanır:

$$\pi_{ij+} = \sum_k (\pi_{i+k} \pi_{+jk} / \pi_{++k})$$

Toplamdaki üç terimin hepside k' yı içermektedir ve bu eşitlik  $\pi_{ij+} = \pi_{i++} \pi_{+j+}$  şeklindeki marjinal bağımsızlığın elde edilmesi amacıyla sadeleştirilemez.  $2 \times 2 \times K$  tablolarına göre, X ve Y arasındaki odds oranı Z' nin her bir kategorisi için 1' e eşit olduğunda; “X ve Y koşullu olarak bağımsızdır” denir. Tablo 2.9'daki frekanslar ( $\{\mu_{ijk}\}$ ), Y=Yanıt (başarı, başarısızlık), X= İlaç tedavisi (A,B) ve Z=Klinik (1,2) için bu ilişkiyi göstermektedir. Z değişkeninin kontrol altına alınmasına bağlı olarak X ve Y arasındaki gerçek ilişki dikkate alınmaktadır (Tabloda beklenen frekanslar üç değişken arasında var olan varsayımsal bir ilişkiyi göstermektedir.). X ve Y her bir kısmi tabloda bağımsızdır ve belirli bir Z değişkenine göre X ve Y arasında koşullu bir biçimde bağımsızlık olduğu söylenmektedir. X ve Y arasındaki tüm koşullu odds oranları “1” e eşit olur.

Belirli bir Z değişkenine göre X ve Y'nin koşullu bağımsızlığı X ve Y' nin marjinal bağımsızlığı ile uyumlu değildir. X ve Y arasındaki odds oranları Z' nin her bir seviyesinde “1” e eşit olduğunda, marjinal odds oranı “1” den farklı olabilmektedir.



**Tablo 2.9** Marjinal Bağımsızlığı İçermeyen Koşullu Bağımsızlık İçin Beklenen Değerlerin Gösterilmesi

		<b>Yanıt</b>	
<b>Klinik</b>	Tedavi	Başarılı	Başarısız
1	A	18	12
	B	12	8
2	A	2	8
	B	8	32
<b>Toplam</b>	<b>A</b>	<b>20</b>	<b>20</b>
	<b>B</b>	<b>20</b>	<b>40</b>

Kaynak:Agresti, 1996, s:58

Marjinal ve koşullu odds oranı konusunda XY koşullu birlikteliğini gösteren formülden, koşullu XY odds oranları aşağıdaki şekilde hesaplanır (Z değişkeninin iki seviyesine göre X ve Y arasındaki koşullu birliktelikler aşağıdaki odds oranlarına göre belirlenir):

$$\theta_{XY(1)} = \frac{18 \times 8}{12 \times 12} = 1.0, \quad \theta_{XY(2)} = \frac{2 \times 32}{8 \times 8} = 1.0.$$

Belirli bir klinik için yanıt ile tedavi şekli koşullu olarak bağımsızdır. Marjinal tablo iki klinik türüne göre tabloları birleştirir.

Marjinal tablonun odds oranı aşağıdaki gibidir ve bu yüzden, değişkenler marjinal olarak bağımsız değildir:

$$\theta_{XY} = (20 \times 40)/(20 \times 20) = 2.0.$$

Klinik türleri önemsenmediğinde, tedavi A için bir başarının olasılığı neden tedavi B' nin iki katıdır? XZ ve YZ koşullu odds oranları bir ipucu verir. Z ile Y ya da X' den biri arasındaki odds oranı, diğer değişkenin belirlenen her bir kategorisine göre, 6.0' ya eşitti. Örnek olarak, Y' nin ilk kategorisi için XZ odds oranı,  $(18 \times 8) / (12 \times 2) = 6.0$ . Belirli bir yanıtı göre klinik 1' de A tedavisini görenlerin koşullu oddsu, klinik 2' de A tedavisini alanların altı katıdır. Klinik 1, tedavi A' yı daha sık kullanma ve bu tedaviden daha çok başarı elde etme eğilimindedir. Örnek olarak, klinik 1' deki hastalar daha genç olma eğiliminde ve sağlıkları klinik 2' de bulunan hastalardan daha iyi ise, belki de görülen tedaviyi dikkate almaksızın daha iyi başarı oranlarına sahip oldukları söylenebilir. Sadece marjinal tabloyu incelemek yanıltıcı olabilmektedir. Sonuç olarak, başarılar tedavi A ile birlikte daha çok olasılıkla elde edilir. Belirli bir klinikte bulunan deneklerin, bütün örneğe nazaran daha homojen olmaları daha olasıdır ve yanıt her bir klinikteki tedavinin bağımsızlığından oluşmaktadır.

### 2.3.5. Homojen Birliktelik

$2 \times 2 \times K$  boyutlu bir olumsuzluk tablosu için aşağıdaki eşitlik gerçekleştiği zaman “*homojen XY birlikteliği*” söz konusu olur:

$$\theta_{XY(1)} = \theta_{XY(2)} = \dots = \theta_{XY(K)}$$

O zaman, X' in Y üzerindeki etkisi Z' nin her bir kategorisi için aynı olur (X ve Y arasındaki koşullu odds oranı Z' nin her bir seviyesinde özdeştir.).

Bu nedenle tek bir sayı ile X-Y koşullu birliktelikleri ifade edilir. X ve Y' nin koşullu bağımsızlığı Z değişkenin her bir seviyesinde odds oranının  $\theta_{XY(k)} = 1.0$  olması durumunda ortaya çıkan özel bir durumdur. Homojen XY birlikteliği durumunda, X' in iki kategorisi ile Z' nin iki kategorisi arasındaki koşullu odds oranı Y' nin her bir kategorisine göre özdeştir. Odds oranına göre, homojen birliktelik simetrik özellik taşımaktadır. Üçüncü değişkenin kategorileri karşısında incelenen ikili (eşleştirilmiş) değişkenlerin her hangi biri ile kullanılabilir. Bu gerçekleştiğinde, “diğer değişkenler üzerinde etkileri bakımından iki değişken

arasında etkileşim olmadığı” söylenir. Etkileşimler var olduğunda, herhangi bir ikili değişkenin koşullu odds oranı üçüncü değişkenin kategorileri karşısında değişmektedir.

X= Sigara içmek ( evet, hayır), Y=Akciğer kanseri (evet, hayır) ve Z= Yaş (<45, 45-65, >65) şeklindeki değişkenler için varsayalım ki,  $\theta_{XY(1)} = 1.2$ ,  $\theta_{XY(2)} = 3.9$  ve  $\theta_{XY(3)} = 8.8$  olsun. O halde sigara içimi, genç insanlar için akciğer kanseri üzerinde zayıf bir etkiye sahiptir, fakat etki yaşla birlikte şiddetlenmektedir. Yaş, **etki değiştirici** olarak adlandırılır ve sigara içiminin etkisi yaşın değerine bağlı olarak değişir. Tablo 2.8’deki ölüm cezası verilerine göre,  $\hat{\theta}_{XY(1)} = 0.43$  ve  $\hat{\theta}_{XY(2)} = 0.0$  şeklindedir. Değerler yakın değildir. Fakat sıfır hücre sayısından dolayı ikinci tahmin tutarlı değildir. Her bir hücreye  $\frac{1}{2}$  eklenerek,  $\hat{\theta}_{XY(2)} = 0.94$  bulunur. Çünkü  $\hat{\theta}_{XY(2)}$  değişkendir ve daha fazla değişim örneklem değişkenliğinden meydana gelebilir. Bu kısmi tablolar bir popülasyondaki homojen birliktelik ile ister istemez çelişmektedir. Daha sonra örneklem verilerinin homojen birliktelik ile tutarlı olması ya da koşullu bağımsızlık ile tutarlı olması durumunda nasıl analiz edilecekleri gösterilecektir.

Bir  $I \times J \times K$  tablosunda homojen X-Y birlikteliği şu anlama gelmektedir: X ve Y’nin iki seviyesini de kullanılarak oluşturulan her hangi bir odds oranı Z’ nin her bir seviyesi için aynıdır.

X-Y koşullu odds oranları Z’ nin her bir seviyesinde özdeş olduğunda, diğer birliktelikler için aynı özellikler göz önünde tutulur.  $2 \times 2 \times 2$  boyutlu tablolarda homojen birliktelik simetrik özelliğe sahiptir. Dolayısıyla XY koşullu odds oranı YZ koşullu odds oranına eşittir. Bu boyuttaki tablolarda birlikteliğin incelenmesi amacıyla ayrıca üç test istatistiği de hesaplanmaktadır: Mantel-Haenszel Testi, Cochran-Mantel-Haenszel Testi ve Breslow-Day Testi.

### 2.3.6. Cochran-Mantel-Haenszel Metotları

Bu bölümde üç yönlü tablolar için yorumsal analizler ele alınacaktır.  $2 \times 2 \times K$  boyutlu tablolarda K tane koşullu odds oranına göre homojen birlikteliğin ve koşullu bağımsızlığın test edilmesi gösterilecektir. Aynı zamanda kısmi birlikteliğin tek bir özet ölçümüne göre K tane kısmi tabloda örnek odds oranlarının nasıl birleştirileceği de ele alınacaktır. Koşullu birliktelik analizleri, çok değişkenli verilere sahip çoğu uygulama için uygun niteliktedir. Örnekle açıklamak amacıyla aşağıdaki tabloda gösterilen verilerin analizini ele alalım. Tabloda sigara kullanımı ve akciğer kanseri konusunda Çin’de yapılan sekiz ayrı incelemenin özetleri yer almaktadır.

X değişkeninin grup sınıflaması sigara içenler ve içmeyenler şeklindedir, akciğer kanseri açısından “evet, hayır” şeklinde iki olası çıktı söz konusudur ve Y yanıt değişkeninin seviyelerini göstermektedir. Z kontrol değişkeninin seviyelerini ise farklı şehirler oluşturmaktadır.

**Tablo 2.10** Cochran-Mantel-Haenszel Testi İle İlgili Olarak Çinli Sigara Kullanıcıları Ve Akciğer Kanseri Yakalananların İncelenmesi

		Akciğer Kanseri				
Şehirler	Sigara Kullanımı	Evet	Hayır	Odds Oranı	$\mu_{11k}$	$\text{var}(n_{11k})$
Pekin	Evet	126	100	2.20	113.0	16.9
	Hayır	35	61			
Shanghai	Evet	908	688	2.14	773.2	179.3
	Hayır	497	807			
Shenyang	Evet	<b>913</b>	<b>747</b>	<b>2.18</b>	<b>799.3</b>	<b>149.3</b>
	Hayır	<b>336</b>	<b>598</b>			
Nanjing	Evet	<b>235</b>	<b>172</b>	<b>2.85</b>	<b>203.5</b>	<b>31.1</b>
	Hayır	<b>58</b>	<b>121</b>			
Harbin	Evet	<b>402</b>	<b>308</b>	<b>2.32</b>	<b>355.0</b>	<b>57.1</b>
	Hayır	<b>121</b>	<b>215</b>			
Zhengzhou	Evet	<b>182</b>	<b>156</b>	<b>1.59</b>	<b>169.0</b>	<b>28.3</b>
	Hayır	<b>72</b>	<b>98</b>			
Taiyuan	Evet	<b>60</b>	<b>99</b>	<b>2.37</b>		<b>9.0</b>
	Hayır	<b>11</b>	<b>43</b>			
Nanchang	Evet	<b>104</b>	<b>89</b>	<b>2.00</b>	<b>96.5</b>	<b>11.0</b>
	Hayır	<b>21</b>		<b>36</b>		

Kaynak: Liu, Z. Smoking and Lung Cancer in China. Intern. J. Epidemiol., 21:197-2 (1992).

Denekler; sosyoekonomik durum şeklindeki karakteristiklere bağlı olarak şehirlerarasında farklılıklar gösterebilmektedir. Buna bağlı olarak sigara içme oranları ve akciğer kanseri bakımından şehirlerarasında heterojenliğe neden olabilmektedir. Bu durumdan dolayı Z değişkeni kontrol altına alınarak, X ve Y değişkenleri arasındaki birliktelik araştırılır.

$2 \times 2 \times K$  tabloları için sıfır hipotezi belirlenmiş bir  $Z$  değişkenine göre  $X$  ve  $Y$  değişkenlerinin koşullu olarak bağımsızlığını ifade eder. Bunun anlamı her bir kısmi tabloda  $X$  ve  $Y$  arasındaki koşullu odds oranının,  $\theta_{XY(k)}$ , “1” e eşit olmasıdır. Standart örnekleme modelleri, hücre sayılarını dört farklı biçimde işleme tabi tutmaktadır (Agresti, 2002):

1. Bağımsız Poisson değişkenleri
2. Tüm örneklem hacminin sabit olduğu çok terimli sayılar
3. Her bir kısmi tablo için örneklem hacminin sabit olduğu çok terimli sayılar (farklı kısmi tablolarda sayılar birbirinden bağımsızdır.)
4. Sabit satır toplamlarına sahip olan her bir kısmi tabloda bağımsız binom örnekleri

$k$  tane kısmi tabloda; satır toplamları,  $\{n_{1+k}, n_{2+k}\}$  ve sütun toplamları,  $\{n_{+1k}, n_{+2k}\}$  şeklindedir. Bu toplamların her ikisi de belirlenmişken, tüm örnekleme tanımlamalarına göre, ilk satır ve sütunun kesiştiği hücre sayısına ( $n_{11k}$ ) göre hipergeometrik bir dağılım sağlanmaktadır. Bu hücre sayısına göre kısmi tablodaki diğer hücre sayıları belirlenmektedir. Test istatistiğinde ise her bir kısmi tablodaki bu hücre sayısı kullanılmaktadır. Sıfır hipotezi altında  $n_{11k}$ 'in ortalaması ve varyansı aşağıdaki şekilde hesaplanmaktadır (Agresti, 1996,s:61):

$$\mu_{11k} = E(n_{11k}) = \frac{n_{1+k}n_{+1k}}{n_{++k}}, \quad \text{var}(n_{11k}) = \frac{n_{1+k}n_{2+k}n_{+1k}n_{+2k}}{n_{++k}^2(n_{++k} - 1)}$$

Gerçek odds oranı ( $\theta_{XY(k)}$ )  $k$  tane kısmi tabloda “1.0” değerini aştığı zaman,  $\{n_{11k} - \mu_{11k}\} > 0$  sonucunun oluşmasını bekleriz. Test istatistiği,  $K$  tane kısmi tablonun tümü için gözlemlenen bu farkları birleştirmektedir. Her bir kısmi tabloda odds oranı “1.0” değerini aştığında bu şekildeki farkların toplamı göreceli olarak büyük pozitif bir sayı olma eğilimindedir; her bir tabloda odds oranı “1.0” değerinden küçük olduğunda ise bu şekildeki farkların toplamı göreceli olarak negatif büyük bir sayı olma eğilimindedir.

Test istatistiđi, K tane kısmi tablodaki bilgileri ařađıdaki formül yardımıyla özetlemektedir:

$$CMH = \frac{\left[ \sum_k (n_{11k} - \mu_{11k})^2 \right]}{\sum_k Var(n_{11k})}$$

Bu test istatistiđi *Cochran-Mantel-Haenszel İstatistiđi* olarak adlandırılır (Agresti, 1996, s:61). Bu istatistik df=1 serbestlik derecesine sahip büyük örneklem ki-kare dağılımına sahiptir. Yanıt deđiřkeninin açıklayıcı deđiřkenden kořullu olarak bađımsız olup olmadığını test eder. Ortak bir odds oranı varsayımı yapılır.  $H_0$  : Belirli bir kontrol deđiřkenine (Z) göre X ve Y kořullu olarak bađımsızdır. Bir deđiřken ilave edildiđinde diđer iki deđiřken arasındaki birliktelikler eklenen deđiřkenin etkisi kontrol altına alınarak sınıranır. CMH testi kontrol deđiřkeninin her bir kategorisinde, diđer iki deđiřkenin kořullu odds oranının “1” e eřit olup olmadığını test eder ( $p > 0.05$  ise odds oranı istatistiksel olarak önemli deđildir). CMH’ nin küçük p deđerine sahip olması ( $p < .05$ ); Z deđiřkeni kontrol altında iken, X ve Y deđiřkenlerinin kořullu olarak bađımsız olmadığını gösterir.  $(n_{11k} - \mu_{11k})$  farkı, bazı tablolarda pozitif bazı tablolarda negatif olmasının dıřında tüm tablolarda tutarlı bir řekilde negatif ya da pozitif olduđunda CMH istatistiđi daha büyük deđerler alır. Kısmi tablolar arasında birliktelik ađısından ciddi anlamda farklılıklar varsa bu test uygun deđildir.

Ancak her bir kısmi tabloda X-Y birlikteliđi benzer nitelikte olduđunda bu test en iyi řekilde işlemektedir. CMH istatistiđi karřılıklı kısmi tablolardaki bilgileri birleřtirir. Her bir tablodaki birliktelik aynı ise, bu test her bir tablo iđerisinde ayrı ayrı kullanılan testlerden daha güçlüdür. Teste göre  $2 \times 2$  boyutlu tek bir marjinal tablo řeklindeki kısmi tabloların birbirlerine eklenmesi yoluyla sonuçların birleřtirilmesi belirsizlik oluřturmaktadır. Simpson Paradoksu (daha önce tanımlanmıřtır: Marjinal birliktelik, kořullu birlikteliklerin her birinden farklı bir yöne sahiptir.) üç yönlü tabloların parçalanmasındaki tehlikeleri ortaya koymaktadır.

### 2.3.7. Akciğer Kanserinin Meta Analizi Örneği

Yukarıda ele alınan tabloda sigara kullanımına ve akciğer kanserine ilişkin Çin’ de yapılan sekiz tane olgu-kontrol incelemesinin özet bilgileri yer almaktadır. Her bir incelemede kanser olmayan kontrol grupları ile kansere yakalanan olgu grupları eşleştirilmiştir ve denekler sigara içip içmemelerine göre kayıt altına alınmıştır. Her bir kısmi tabloda, sütunlardaki sayılar binom örnekleme olarak işleme tabi tutulur. Ayrıca sütun toplamı belirlenmiş ve sabittir. Burada şehirlere göre sigara kullanımı ile akciğer kanserine yakalanma arasındaki koşullu bağımsızlık hipotezi test edilmektedir. Bu hipoteze göre (sıfır hipotezi) gerçek odds oranının her bir şehir için “1” e eşit olması gerekmektedir. Tabloda, her bir tablo için örnek odds oranları ve bu hipotez altında sigara içenler (her bir tabakadaki ilk satır ve ilk sütunun kesiştiği hücredeki sayı) arasında akciğer kanserine yakalanan olguların varyansları ve beklenen değerleri yer almaktadır. Her bir kısmi tabloda örnek odds oranı orta derecede bir pozitif ilişkiyi göstermektedir. Bu nedenle CMH istatistiği bakımından sonuçlar birleştirilebilir. Aşağıdaki hesaplamalar yapılarak CMH istatistiği elde edilir:

$$\sum_k n_{11k} = 2930 \quad \sum_k \mu_{11k} = 2562.5 \quad \sum_k Var(n_{11k}) = 482.1$$

$$CMH = (2930 - 2562.5)^2 / 482.1 = 280.1 \text{ ve } df=1.$$

Koşullu bağımsızlığın aksi yönde çok güçlü bir kanıt oluşturmaktadır ( $P < 0.0001$ ). Bu şaşırtıcı değildir, çünkü birleştirilen incelemeler için örnek hacmi büyüktür ( $n=8419$ ). Birçok incelemeden elde edilen bilgilerin birleştirilmesi ile yapılan istatistiksel analize “Meta Analizi” adı verilmektedir. Aynı konu üzerinde farklı yer ve zamanlarda yapılmış çalışmalarını bir araya getirerek araştırılan konu hakkında daha güvenilir ve daha doğru bir sonuç elde etmek amacıyla uygulanan istatistiksel bir yöntem olup sosyal bilimlerde tıp alanında kullanıldığı görülmektedir (Akçil ve Karaağaoğlu, 2001). Yukarıdaki tabloda yer alan meta analizi herhangi tek bir kısmi tabloya kıyasla birliktelik açısından daha güçlü bir kanıt sağlamaktadır.



### 2.3.8. Ortak Odds Oranının Hesaplanması

Birlikteliğin gücünün tahmin edilmesi açısından bu yöntem, birlikteliğin gücüne ilişkin oluşturulan hipotezin test edilmesinden daha fazla bilgi verici niteliktedir. Kısmi tablolar açısından birliktelik tutarlı ise K tane gerçek odds oranı için varsayılan ortak bir değer tahmin edilebilir ve birlikteliğin gücünün ölçülmesini sağlayabilir.  $2 \times 2 \times K$  boyutlu bir tabloda,  $\theta_{XY(1)} = \dots = \theta_{XY(K)}$  olduğunu varsayalım. *Mantel-Haenszel* tahmin edicisi ortak bir değerdir ki aşağıdaki şekilde hesaplanır (Akçıl ve Karağaoğlu, 2001):

$$\hat{\theta}_{MH} = \frac{\sum_k (n_{11k} n_{22k} / n_{++k})}{\sum_k (n_{12k} n_{21k} / n_{++k})} \quad 34$$

$2 \times 2$  boyutlu tablolarda, OR, birinci satırdaki başarının odds' nu ikinci satırdaki başarı odds'una göre açıklamaktadır. Diğer yandan  $2 \times 2 \times 2$  boyutlu tablolarda ise iki tane odds ratio vardır ve bundan dolayı da birlikteliğin gücünün ölçülmesi için ortak bir odds oranının hesaplanması gerekmektedir. MH testi, açıklayıcı ve yanıt değişken arasında ortalama bir koşullu birlikteliğin hesaplanmasına dayanmaktadır. Farklı tablolardaki OR' lerin ağırlıklı ortalamasıdır (Akçıl ve Karağaoğlu, 2001)

---

<sup>34</sup> Her çalışmaya ait ya da iki boyutlu tüm kısmi tablolara ait OR aşağıdaki gibidir:

$$\theta_i = \frac{a_i d_i}{c_i b_i}$$

Tüm çalışmaların ya da kısmi tabloların birleştirilmesi durumunda ortak odds oranı:

$$\theta = \frac{\sum_k (\theta_i W_i)}{\sum_k W_i}$$

k: Çalışma sayısı ya da tabaka(kontrol değişkeninin kategori) sayısıdır.

$W_i$ : i. varyansın tersidir:

$$W_i = \frac{1}{V_i} \quad V_i = \frac{n}{(b_i \cdot c_i)}$$

$\log(\hat{\theta}_{MH})$  için hesaplanan standart hata karışık bir formüle sahiptir (Agresti,1990, p.236). Bundan dolayı burada ele alınmayacaktır. SAS-PROC FREQ gibi bazı yazılımlar standart hatayı hesaplamaktadır ve aynı zamanda da logaritmik modellerin kullanıldığı standart hatayı ve bir tahmini de sağlamaktadır. Bu yöntem ikili (binary) veri kümesinin sonuçlarını birleştirmek için kullanılıp, etki ölçütünün oran ve sabit etki modeline dayalı olduğu durumlarda kullanılan bir yöntemdir.

Tabloda özetlenen Çinli sigara kullanıcıları için, Mantel-Haenszel (MH) odds oranı tahmini aşağıdaki şekilde hesaplanır:

$$\hat{\theta}_{MH} = \frac{(126)(61)/(322) + \dots + (104)(36)/250}{(35)(100)/(322) + \dots + (21)(89)/(250)} = 2.17$$

$\log(\hat{\theta}_{MH}) = \log(2.17) = 0.777$  'in tahmin edilen standart hatası 0.046'dır.

Ortak log odds oranı için yaklaşık olarak %95 güven aralığı şu şekildedir:  $0.777 \pm 1.96 \times 0.046$  ya da (0.686, 0.868). Bu aralığın logaritması ( $\exp(.686)$ ,  $\exp(868))=(1.98, 2.38)$ . Sigara içenlerin akciğer kanserine yakalanma oddsu, sigara içmeyenlerin kansere yakalanma odds'unun yaklaşık olarak iki katıdır.

Bu şekildeki odds oranları Batı toplumlarında tipik olarak daha büyüktür. Son zamanlarda Çin'de kısmen daha küçük olabilir. Çünkü daha uzun gövdelere sahip pipolar sigaralardan daha yaygındır. Eğer gerçek odds oranları özdeş değil fakat büyük ölçüde de farklılık göstermiyorsa,  $\hat{\theta}_{MH}$ , K tane koşullu birlikteliği temsil eden bir özet sağlar. Benzer şekilde örneklem birliktelikleri tek bir yönde değişiyor ise, CMH testi koşullu bağımsızlık hipotezine karşı güçlü bir kanıtın özetini oluşturmaktadır.

### 2.3.9. Odds Oranlarının Homojenliğinin Testi

“X ve Y arasındaki odds oranı Z' nin her bir seviyesinde aynıdır.” şeklindeki hipotez ( $H_0 : \theta_{XY(1)} = \dots = \theta_{XY(K)}$ ) test edilir ki bu  $2 \times 2 \times K$  tablolarında

birlikteliğin homojenliğinin testidir (Agresti, 1996, s:63).  $\{\hat{\mu}_{11k}, \hat{\mu}_{12k}, \hat{\mu}_{21k}, \hat{\mu}_{22k}\}$ ; gözlemlenen veriler kadar aynı marjinal toplamlara sahip kısmi tablolarda k' nıncı kısmi tablodaki tahmin edilen beklenen frekansları ifade etmektedir. Odds oranı ortak bir odds oranının Mantel-Haenszel tahminine ( $\hat{\theta}_{MH}$ ) eşittir. Buradaki test istatistiği Breslow-Day İstatistiği olarak adlandırılır ve Pearson formuna sahiptir(Agresti, 1996, s:63):

$$\sum \frac{(n_{ijk} - \hat{\mu}_{ijk})^2}{\hat{\mu}_{ijk}}$$

Bu test istatistiği odds oranının homojenliği için oluşturulan sıfır hipotezinin test edilmesinde kullanılmaktadır. Z kontrol değişkeninin farklı kategorileri için X-Y arasındaki odds oranının özdeş olup olmadığını test eder.

Yüksek P değeri ile BD istatistiği ( $p > 0.05$  ise  $H_0$ :Kabul Edilir.) odds oranının homojenliğinin ifade edildiği sıfır hipotezini kabul ederse tek bir odds oranı ile koşullu birlikteliğin özetlenmesi mümkün olmaktadır. BD istatistiği düşük P değerine sahip ise kontrol değişkenin kategorilerine göre odds oranlarında homojenlik yoktur ( $p < 0.05$  ise  $H_0$ :Red Edilir.). Bundan dolayı da sadece bir tane odds oranı ile birlikteliğin özetlenmesi mümkün olmaz. Değerler toplam tablodaki tüm hücreler üzerinden alınmaktadır. Hücre sayıları birbirine yaklaştıkça daha küçük istatistik değerine sahip olur ve  $H_0$ ' a karşı daha az kanıt oluşturur. Ortak bir odds oranını sağlayan  $\{\hat{\mu}_{ijk}\}$ 'nin hesaplanması oldukça karmaşıktır ve burada ele alınmayacaktır. Ancak SAS yazılımındaki PROC FREQ gibi standart yazılımlar bu istatistiği hesaplamaktadır (Agresti, 1996). Breslow-Day istatistiği  $df=K-1$  serbestlik derecesine sahip yaklaşık bir ki-kare dağılımıdır. Örnek hacmi her bir kısmi tabloda kısmen büyük olmalıdır, en azından hücrelerin %80'inde  $\{\hat{\mu}_{ijk} \geq 5\}$  olmalıdır (Agresti, 2002).

Akciğer kanseri ve sigara kullanımına ilişkin yapılan meta analizine göre, yazılım aracılığı ile Breslow-Day istatistiği 5.2' ye eşittir ve serbestlik derecesi,  $df=7$  olarak bulunur ( $P=0.64$ ). Bu sonuç, eşit odds oranları hipotezini reddetmek için kanıt oluşturmamaktadır. Sekiz kısmi tablonun tümü için tek bir odds oranının kullanımı

ile koşullu olasılığın özetlenmesi değerlendirilebilir. Örnek hacmi arttıkça Breslow-Day istatistiği dağılımının Ki-kare'ye yakınsamasını sağlamak amacıyla, R. Tarone (Biyometri, 72: 91-95, 1985) aşağıdaki eşitlik aracılığı ile bir düzeltme yapılması gerektiğini ortaya koymuştur:

$$\frac{\left[ \sum_k (n_{11k} - \hat{\mu}_{11k}) \right]^2}{\sum_k \left[ \frac{1}{\hat{\mu}_{11k}} + \frac{1}{\hat{\mu}_{12k}} + \frac{1}{\hat{\mu}_{21k}} + \frac{1}{\hat{\mu}_{22k}} \right]^{-1}}$$

Çin' de yapılan araştırmalardan oluşan tablodaki her bir kısmi tablo bir olgu-kontrol incelemesidir. Akciğer kanserine sahip her bir denek, akciğer kanseri olmayan bir ya da daha fazla kontrolle eşleştirildiğinden dolayı iki sütunda gerçekten birbirinden bağımsız binom örneklerinden oluşmamaktadır. Uygulamada olgunun (akciğer kanseri olan denekler) sigara içicisi olup olmaması, kontrolün sigara içicisi olup olmamasından bağımsızdır. Bundan dolayı her bir kısmi tablodaki sütunlar bağımsız binom örnekleri ile aynı olmalıdır. Eğer her bir olgu-kontrol ikilisi için her bir deneğin sigara kullanıcısı olup olmadığı konusunda da bilgi edinilseydi, kontrol deneğinin sigara içip içmemesi (E,H) ile olgu deneğinin sigara içip içmemesine (E,H) göre  $2 \times 2$  yönlü bir tablo oluşturulurdu. Sonra sigara içme oranlarına göre olgular ve kontrollerin karşılaştırılmasını sağlayacak metotlar kullanılırdı ki potansiyel bağımlılığı da dikkate almaktadırlar. Oranlar arasında pozitif bir korelasyon söz konusu olduğunda bağımsız örnekle elde edilen P-değerinden daha küçük P-değerleri elde edilir. Daha sonra logaritmik doğrusal modellere bağlı olarak odds oranlarının homojenliğinin test edilmesinde alternatif testler ele alınacaktır.

Kısaca ikiden fazla açıklayıcı değişken söz konusu olursa ( X: Açıklayıcı değişken, Z: Kontrol değişkeni olur.), X-Y değişkenleri arasındaki birlikteliğe ilişkin üç test kullanılır:

- \* X-Y koşullu bağımsızlığın test edilmesi için CMH testi
- \* X-Y birlikteliğinin gücünün belirlenmesi için MH testi
- \* Odds oranının homojenliğinin test edilmesi için BD testi.

Aynı zamanda meta analizinde çeşitli çalışmaların birleştirilmesi amacıyla da bu istatistiksel yöntemler kullanılmaktadır. Çoğunlukla olgu-kontrol çalışmalarında iki boyutlu olumsuzluk tablolarında Z kontrol değişkeni şeklinde bir değişkeninde dahil edilmesi durumunda kullanılmaktadır ([www.stat.com](http://www.stat.com)).

### **2.3.10. Koşullu Birliktelikler Konusunda Kesin (Exact) Olasılık Yorumlamaları**

Daha önceki bölümlerde  $I \times J$  boyutlu tablolarda bağımsızlığın sınanmasında kullanılan ki-kare testleri büyük örnek testleriydi. Örnek hacmi artıkça gerçek örnekleme dağılımı da ki-kare dağılımına yakınsamaktadır. Örnek hacminin ne kadar büyük olması konusunda genel bir ilke ortaya koymak oldukça zordur. Bu testlerin yeterliliği, ayrı ayrı kısmi tablolardaki sayılardan daha çok iki yönlü tablolardaki marjinal toplamlara bağlıdır. Örnek olarak CMH istatistiği için, kısmi tablolardaki hücre sayıları küçük olabilir (K büyük olduğunda bu gerçekleşmektedir). Fakat X-Y marjinal toplamları göreceli olarak büyük olmalıdır. Uygulamada küçük örnek hacmi genellikle sorun yaratmaktadır. Bundan dolayı koşullu birliktelikler konusunda kesin olasılık (exact) testlerine dayalı yorumlama yol gösterici olur. Örnek olarak koşullu bağımsızlığın kesin olasılık testleri  $2 \times 2$  boyutlu tablolarda kullanılan Fisher Kesin Olasılık Testinin genelleştirilmesine dayanmaktadır (Agresti, 1996, s:64).

#### **2 x 2 x K Boyutlu Tablolarda Koşullu Bağımsızlığın Kesin Olasılık Testi:**

$2 \times 2 \times K$  boyutlu tablolar için, her bir kısmi tablodaki marjinal toplamlar üzerindeki koşula göre, CMH koşullu bağımsızlık testi  $\sum_k n_{11k}$  ile hesaplanan hücre sayılarına bağlıdır.  $2 \times 2$  boyutlu tablolar için Fisher Kesin Olasılık Testinde  $n_{11}$  kullanılırken buradaki kesin olasılık testlerinde de  $\sum_k n_{11k}$  kullanılır. Her bir kısmi tablodaki hipergeometrik dağılımlar  $\{n_{11k}, k = 1, \dots, K\}$  için olasılıkların belirlenmesini sağlamaktadır. Bunlar toplamların dağılımını belirlemektedir. Koşullu bağımsızlığın ifade edildiği sıfır hipotezi : “Tüm koşullu odds oranları,  $\{\theta_{XY(k)}\}$ , “1”

e eşittir.” şeklindedir. Pozitif bir koşullu birliktelik; tek yanlı alternatif hipotezi oluşturmaktadır,  $\theta_{XY(k)} > 1$  şeklinde ifade edilir.

P-değeri sağ kuyruk olasılığına eşittir ve sabit marjinal toplamlar için  $\sum_k n_{11k}$  en azından gözlemlenen kadar büyüktür. Tek yanlı alternatif hipotezi  $\theta_{XY(k)} < 1$  şeklinde ise P-değeri sol kuyruk olasılığıdır ve  $\sum_k n_{11k}$  gözlemlenenen daha büyük olamaz. İki yanlı alternatif hipotezlerinde ise sonuçların iki kuyruktaki olasılıkları kullanılır ki sonuçlar gözlemlenenen daha fazla olamaz. Koşullu bağımsızlığın kesin olasılık testleri hesaplama açısından oldukça karışıktır. Uygulamada yazılım kullanımını gerektirmektedir. U.S. Hükümetinde aynı kidede sahip bilgisayar uzmanları GS-13 seviyesinden GS-14 seviyesine kadarki sınıflamada terfileri bakımından dikkate alınmış ve elde edilen bulgular aşağıdaki tabloda özetlenmiştir(Agresti, 1996). Tabloda personelin ırkına göre verilen terfi kararı üç ayrı ay için çapraz olarak sınıflandırılmıştır. Terfi kararı ile personel ırkı arasında koşullu bağımsızlık test edilecektir. Tabloda küçük sayılar söz konusudur. Toplam örnek hacmi küçük değildir, n=74. Fakat bir marjinal toplam sifıra eşittir (kararın aylara göre parçalanmasından dolayı). Bundan dolayı da CMH testini kullanmaktan sakınabiliriz.

**Tablo 2.11** Terfilerle Çalışanların Irklarına İlişkin Olumsuzluk Tablosu

Aylara ve ırklara göre verilen terfi kararları	Terfi Durumu (Evet/Hayır)					
	Temmuz Terfileri		Ağustos Terfileri		Eylül Terfileri	
	E	H	E	H	E	H
Siyah	0	7	0	7	0	8
Beyaz	4	16	4	13	2	13

Kaynak: Agresti, 1996, s:65 (E: Evet, H: Hayır)

İlk olarak 1'den küçük olasılık oranını ifade eden tek yanlı alternatif kullanılır ki bu siyah personele karşı bir ayırım yapıldığını göstermektedir (Beyaz personele kıyasla siyah personelin daha az terfi alma olasılığıdır). Her bir kısmi tablodaki sütun ve satır toplamları sabittir ve testte her birindeki ilk hücre sayısı ( $n_{11k}$ ) kullanılmaktadır. Kısmi tabloların marjinleri için,  $n_{111}$  değeri 0 ile 4 arasında,  $n_{112}$  değeri de 0 ile 4 arasında,  $n_{113}$  değeri ise 0 ile 2 arasında yer almaktadır.  $\sum_k n_{11k}$  ise toplamlarıdır ve 0 ile 10 arasında değer almaktadır. Örnek veriler, üç durumun her biri için en aşırı olasılığı vermektedir. Gözlemlenen  $\sum_k n_{11k}$  değeri ise "0" a eşittir. P-değeri de bu sonucun sıfır olasılığını vermektedir ve yazılım aracılığı ile 0.026 olarak bulunmuştur. İki yanlı P-değeri ise 0.056 olarak bulunmuştur (gözlemlenen tablodan daha büyük olmayan olasılıklara sahip tüm kısmi tabloların olasılıklarının toplamına dayanmaktadır.). Bu durumda terfi kararının ırkla ilişkili olduğuna dair kanıt oluşmaktadır.

### **Ortak Odds Oranı için Kesin Olasılık Güven Aralığı**

Küçük örneklerdeki kesiklilik kesin olasılık testlerinin tutuculuğunu içermektedir. Örnek olarak sıfır hipotezi doğru olduğunda, P-değeri %5 değerinden (0.05) daha aşağıda bulunabilir. Orta-P-değeri kullanılarak tutuculuk hafifletilir. Daha uçtaki tabloların olasılığı ile gözlemlenen sonucun olasılığının yarısının toplanması ile elde edilir.  $\{\theta_{XY(k)}\}$  yerine  $\theta$  gibi varsayılan ortak bir değer için aynı zamanda da "kesin olasılık" güven aralıkları kurulmaktadır. Kesiklilikten dolayı bu güven aralıkları da tutucu niteliktedir. %95 bir "kesin olasılık" güven aralığı için, gerçek güven düzeyi en azından 0.95 kadar büyüktür, fakat tam değeri bilinmemektedir. Daha olumlu ve yararlı bir %95 güven aralığı;  $H_0 : \theta = \theta_0$  hipotezinin test edilmesinde 0.05 değerini aşan orta-P-değerlerine sahip  $\theta_0$  değerlerini kapsamaktadır.

Orta P-değerine bağlı güven aralıklarında, gerçek güven düzeyinin en azından nominal düzey kadar olması garanti edilemese de bu güven aralıkları daha dardır ve genellikle kesin ya da büyük örnek aralıklarına kıyasla nominal seviyeye daha

yakındır. Yaklaşımların hepsi hesaplama açısından çok karışıktır ve yazılım kullanımını gerektirir. Terfi kararları konusunda ele alınan örneğe geri dönelim. Örneklem sonucu en uç nokta olasılığı olduğundan varsayılan ortak bir odds oranı için kullanılan Mantel-Haenszel tahmin edicisi;  $\hat{\theta}_{MH} = 0.0$  olarak bulunur. StatXact yazılımı ile ortak bir odds oranı için elde edilen %95 kesin güven aralığı; (0, 1.01) şeklindedir (Agresti,1996,s:66). %95 güvenle gerçek odds oranının bu aralık içine düştüğünü söyleriz. Orta P-değeri kullanılarak yapılan testlere ilişkin %95 güven aralığı ise (0, 0.78) şeklindedir. %95 güvenle yaklaşık olarak şu yorum yapılır: Siyahlar için verilen terfilerin olasılığı (odds), beyazlara verilen terfilerin olasılığından (%78=odds) daha fazla değildir.

### **Odds Oranlarının Homojenliğinin Kesin Olasılık Testi**

Odds oranlarının homojenliğinin test edilmesinde kullanılan Breslow-Day testi aynı zamanda büyük örneklem testidir. Her bir kısmi tablodaki örnek hacmi nispeten büyük olduğunda uygulanmaktadır. Toplam örnek hacmi küçük olduğunda ya da toplam örnek hacmi büyük fakat kısmi tablo sayısı (K) büyük ve ayrı ayrı tablolardaki örnek hacimleri küçük olduğunda bu test geçerli olmamaktadır. Odds oranlarının homojenliğinin kesin olasılık testi (çoğunlukla “*Zelen Exact Testi*” olarak adlandırılır.) bu şekildeki durumlarda kullanılmaktadır (Agresti,1996).

Kesin olasılık dağılımı ise  $2 \times 2 \times K$  boyutlu tablolar kullanılarak hesaplanmaktadır ki bu tablolar gözlemlenen tablo gibi benzer iki yönlü marjinal tablolara sahiptir. P-değeri ise bu tabloların olasılıklarının toplamıdır ki gözlemlenen tablodan daha fazla olasılığa sahip değildir.

Yine terfilere ilişkin örneğe geri dönersek;  $\{\hat{\mu}_{ijk}\}$  değerleri (bu değerlere göre ortak odds oranı için Mantel-Haenszel tahmini,  $\hat{\theta}_{MH} = 0.0$  elde edilmiştir) her bir kısmi tabloda gözlemlenen sayılarla özdeştir. Breslow-Day test istatistiği 0/0 şeklindeki terimleri içermektedir ve dolayısı ise test istatistiği sonucu belirsizdir. Ek olarak gözlemlenen tablodan başka hiçbir tablo bu tablonun iki yönlü marjinal toplamalarının hepsini içermemektedir. Bundan dolayı Zelen Exact Testi  $P=1<0$  değeri için niteliğini yitirir ve bozular. Bu şekildeki veriler için, gerçek odds



oranlarının farklı olup olmadığı konusunda anlamlı her hangi bir bilgi sağlanamamaktadır (Agresti, 2002). Kesin yorumlamanın sakıncası şudur ki: Küçük örnek koşullu dağılımı genelde oldukça kesiklidir. Birçok durumda bu şekilde tek bir noktaya odaklanılır. Küçük örnek hacimleri ile ancak bu kadarı yapılabilmektedir.

## 2.4. I x J TABLOLARI İÇİN İLAVELER

$2 \times 2$  tablolarına göre odds oranı şeklindeki tek bir değer birlikteliği özetlemektedir<sup>35</sup>.  $I \times J$  tabloları için bazı bilgi kayıpları olmaksızın tek bir değer ile birlikteliğin özetlenmesi çok nadir olarak mümkün olabilmektedir. Fakat odds oranlarının bir kümesi ya da başka bir özet gösterge, birlikteliğin kesin ve tam özelliklerini tanımlayabilmektedir.

### 2.4.1. I x J Tablolarındaki Odds Oranları

Odds oranları sütunların  $\binom{J}{2} = J(J-1)/2$  tane çiftin her birini birleştiren sıraların  $\binom{I}{2} = I(I-1)/2$  tane çiftini kullanmaktadır. a ve b sırası ile c ve d sütunlarına göre odds oranı,  $(\pi_{ac}\pi_{bd})/(\pi_{bc}\pi_{ad})$  şeklindedir ve dikdörtgen şeklindeki tabloda ki dört hücreyi kullanır.

Odds oranlarının bu kümesi çok fazla gereksiz bilgi içermektedir.  $(I-1)(J-1)$  şeklindeki *bölgesel odds oranlarının* ( local odds ratios) bir altkümesi dikkate alınırsa aşağıdaki eşitlik elde edilir(Agresti, 2002):

---

<sup>35</sup> Goodman (2000)  $2 \times 2$  tablosu içinde birlikteliğin incelenmesinin beş farklı yöntemini göstermiştir ve bu yöntemlerin hepsini içeren genel bir ölçüm önermiştir (Agresti, 2002).

$$\theta_{ij} = \frac{\pi_{ij}\pi_{i+1,j+1}}{\pi_{i,j+1}\pi_{i+1,j}}, \quad i = 1, \dots, I-1, \quad j = 1, \dots, J-1.$$

Bu (I-1)(J-1) şeklindeki *bölgesel odds oranları* sütun çiftlerinden ve sıra çiftlerinden oluşmuş bütün odds oranlarını belirler. Örnekle açıklamak amacıyla tablo-2.2’deki örneği ele aldığımızda, ilk iki sütun için örneklem bölgesel odds oranı, 2.08 ve ikinci ve üçüncü sütunlar için 1.74 şeklinde bulunur. Her bir durumda, yalancı ilaç alan gruba göre daha genel ciddi sonuçlar vardır. Birinci ve üçüncü kolonlar için hesaplanan odds oranının çarpımı ise 3.63 olmaktadır. (I-1)(J-1) şeklindeki *bölgesel odds oranlarının* bir altkümesi dikkate alınarak oluşturulan eşitliğin, odds oranlarını en basite indirgeyen (en az) bir kümesi için oluşturulması tek yol değildir. Başka bir temel (basit) küme ise aşağıdaki gibi ifade edilir:

$$\alpha_{ij} = \frac{\pi_{ij}\pi_{IJ}}{\pi_{ij}\pi_{iJ}}, \quad i = 1, \dots, I-1, \quad j = 1, \dots, J-1.$$

Bu eşitlik ise, en son sıra ve en son sütundaki hücre ile j’ ninci sütun ve i’ ninci sıradaki hücre aracılığı ile belirlenen hücrelerin dikdörtgen şekliinden faydalanır. Belirlenmiş  $\{\pi_{i+}\}$  ve  $\{\pi_{+j}\}$  şeklindeki marjinal dağılımlara göre,  $\{\pi_{ij} > 0\}$  olduğunda, yukarıdaki sözü geçen formüllerdeki odds oranları setindeki olasılıkların yer değiştirmesi bilgiyi atmamaktadır.

Hücre olasılıkları odds oranlarını belirler ve belirli marjinallere göre odds oranları hücre olasılıklarını belirlemektedir. Bu anlamda, (I-1)(J-1) tane parametre, bir  $I \times J$  tablosundaki herhangi bir birlikteliği ifade etmektedir. Bağımsızlık ise, tüm (I-1)(J-1) odds oranlarının 1.0’ a eşit olması ile sağlanmaktadır. Üç yönlü  $I \times J \times K$  olumsuzluk tablolarına göre, kısmi tablolarda yer alan odds oranı kümeleri koşullu birlikteliği ifade etmektedir. Homojen XY birlikteliği ise şu anlama gelmektedir; X’in ve Y’ nin iki kategorisi kullanılarak oluşturulan herhangi bir koşullu odds oranı Z’ nin her kategorisi için aynıdır.

## 2.4.2. Birlikteliğin Özet Ölçümleri

Birlikteliği tanımlamak için alternatif bir yöntem ise tek bir özet göstergesi kullanmaktır. Bu konu öncelikle nominal değişkenler için daha sonra ordinal değişkenler için ele alınacaktır. Nominal değişkenler için en çok yorumlanabilen göstergeler, aralık değişkenleri için kullanılan  $R^2$  ile benzer yapıya sahiptir. Nominal değişkenler için kullanılan gösterge ve sıralar arası korelasyon katsayısı ile korelasyon oranı (Kendall ve Stuart 1979), belirli bir X açıklayıcı değişkenine göre Y' nin koşullu dağılımları ile Y yanıtının marjinal dağılımından dolayı ortaya çıkan varyans bakımından oransal indirgemeyi ifade etmektedir (Agresti, 2002).  $V(Y)$ ; Y' nin  $\{\pi_{+j}\}$  marjinal dağılımına göre değişimin (varyasyonun) ölçümünü ve  $V(Y|i)$ ; X' in i' ninci setine göre Y' nin  $\{\pi_{1i}, \dots, \pi_{ji}\}$  şeklindeki koşullu dağılımı için hesaplanan ölçümü göstermektedir. Varyasyon ölçümü için oransal bir indirgeme aşağıdaki şekildedir (Agresti, 2002):

$$\frac{V(Y) - E[V(Y|X)]}{V(Y)}$$

$E[V(Y|X)]$ , X' in dağılımına göre koşullu değişimin (varyasyonun) beklenen değeridir. X' in  $\{\pi_{i+}\}$  şeklindeki marjinal dağılımı için koşullu değişimin beklenen değeri aşağıdaki gibidir:

$$E[V(Y|X)] = \sum_i \pi_{i+} V(Y|i)$$

Nominal bir yanıt için, Theil (1970)  $V(Y) = \sum \pi_{+j} \log \pi_{+j}$  şeklinde bir varyasyon (değişim) ölçümünü kullanan bir gösterge önermiştir ki buna **entropi** adı verilir. Bu türdeki ölçümler ile ilgili bir zorluk; büyük bir değer ne kadar güçlü bir birliktelik oluşturduğu ile ilgili bir önsezi geliştirmesidir. Bu ne anlama gelmektedir? Örnek olarak, entropi içinde %30 indirgeme var mıdır? Özet ölçümler, bütün sınıflandırmalar ordinal olduğunda, yorumlama açısından daha kolay, daha kullanışlı

ve yararlı gözükmetedir. Olumsallık tablolarına göre entropi için oransal indirgeme aşağıdaki ifadeye eşittir (Agresti, 2002):

$$U = \frac{\sum_i \sum_j \pi_{ij} \log(\pi_{ij} / \pi_{i+} \pi_{+j})}{\sum_j \pi_{+j} \log \pi_{+j}}$$

Buna **belirsizlik katsayısı** (uncertainty coefficient) adı verilir. Bu ölçüm,  $\{\pi_{+j}\}$  değerlerinden bir tanesinden daha fazlası;  $\pi_{+j} > 0$  şeklinde olduğu zaman en iyi şekilde belirlenir. 1 ile 0 arasında değer alır: U=0 durumu X ve Y' nin bağımsızlığını göstermektedir. U=1 ise koşullu varyasyonun bir gereksinimine karşılık gelmektedir, bu anlamda her bir i' ye göre, bazı j' ler için  $\pi_{ji} = 1$  olur.

#### 2.4.3. Ordinal Eğilimler: Uyumlu ve Uyumsuz İkili

Tablo 2.12' da yer alan değişkenler, gelir ve iş memnuniyetidir, (U.S) ulusal örneklem içinde siyah erkekler için ölçümlenmiştir. Her iki sınıflandırmada ordinal niteliktedir. İş memnuniyetinin kategorileri “çok memnuniyetsiz(VD), az memnuniyetsiz (LD), orta derecede memnun (MS) ve çok memnun (VS)” şeklindedir. X ve Y ordinal ise monoton (tekdüze) bir eğilim birlikteliği yaygındır. X' in seviyesi arttıkça, Y üzerindeki yanıtlar daha yüksek seviyelere doğru artma eğilimindedir. Örnek olarak, muhtemelen iş memnuniyeti gelir arttıkça artacaktır. Tek bir parametre bu eğilimi tanımlamaktadır. Korelasyona benzer ölçümler tek düze (monoton) olan bir ilişkinin derecesini tanımlamaktadır. Bazı ölçümler, uyumlu ya da uyumsuz şeklinde deneklerin her bir ikilisinin (çiftinin) sınıflandırılmasına bağlıdır(Oktay,2003).

Bir **ikilinin uyumlu** olması, deneğin X' in daha yüksek seviyelerde sıralanmış iken aynı zamanda Y' nin de daha yüksek seviyelerinde sıralanmış olmasıdır. Eğer denekler X' in daha yüksek seviyelerinde sıralanmış denekler Y' nin daha düşük seviyelerinde sıralanırsa **uyumsuz ikili** oluştururlar (Oktay,2003). Bir ikili, eğer denekler X ve/veya Y üzerinde aynı sınıflandırmaya sahipse bağlıdır.

**Tablo 2.12** Gelir Aracılığı İle İş Memnuniyetinin Çapraz Sınıflaması

Gelir (\$)	VD	LD	MS	VS
<15,000	1	3	10	6
15,000-25,000	2	3	10	7
25,000-40,000	1	6	14	12
>40,000	0	1	9	11

Kaynak: Agresti, 2002

Tablo-2.14' e göre açıklayalım. Deneklerin bir ikilisini dikkate alalım, (<15, VD) hücresinde yer alan denek ve (15–25, LD) hücresinde yer alan denekten oluşsun. Bu ikili uyumludur. Çünkü hem gelir hem iş memnuniyetine göre ikinci denek birincisine göre daha yüksek bir seviyede sıralanır. (<15, VD) hücresinde ki denek, (15–25, LD) hücresinde sınıflandırılmış üç denegin her biri ile eşleştirildiğinde uyumlu ikilileri oluşturur. Böylece bu iki hücre  $1 \times 3 = 3$  tane uyumlu ikili sağlar. (<15, VD) hücresinde ki denek aynı zamanda, her iki değişkenin daha yüksek seviyelerinde sıralanmış diğer deneklerin (10+7+6+14+12+1+9+11) her biriyle eşleştirildiğinde de uyumlu çiftin bir parçasını oluşturur. Benzer şekilde, (15-25, LD) hücresindeki üç denek, her iki değişkenin daha yüksek seviyelerinde sıralanmış denekler (10+7+14+12+9+11) ile eşleştirildiğinde uyumlu ikililerin bir parçası olurlar. Uyumlu ikililerin toplam sayısı C ile ifade edilir ve aşağıdaki gibi hesaplanarak bulunur(Agresti, 2002):

$$\begin{aligned} C = & 1(3 + 10 + 7 + 6 + 14 + 12 + 1 + 9 + 11) \\ & + 3(10 + 7 + 14 + 12 + 9 + 11) + 10(7 + 12 + 11) \\ & + 2(6 + 14 + 12 + 1 + 9 + 11) + 3(14 + 12 + 9 + 11) \\ & + 10(12 + 11) + 1(1 + 9 + 11) + 6(9 + 11) + 14(11) = 1331. \end{aligned}$$

Gözlemlerin uyumsuz ikililerinin toplam sayısı da şu şekildedir:

$$D = 3(2 + 1 + 0) + 10(2 + 3 + 1 + 6 + 0 + 1) + \dots + 12(0 + 1 + 9) = 849.$$

Bu örnekte  $C > D$  olur. Bu, yüksek iş memnuniyetinin yüksek gelir ve düşük iş memnuniyetinin de düşük gelir meydana getirme eğiliminde olduğu konusunda fikir vermektedir.

$\{\pi_{ij}\}$  şeklindeki bir ortak olasılık dağılımıyla iki bağımsız gözlemi dikkate alalım. Bu ikili için, uyum ve uyumsuzluk olasılıkları aşağıdaki gibidir:

$$\Pi_c = 2 \sum_i \sum_j \pi_{ij} \left( \sum_{h>i} \sum_{k>j} \pi_{hk} \right), \quad \Pi_d = 2 \sum_i \sum_j \pi_{ij} \left( \sum_{h>i} \sum_{k<j} \pi_{hk} \right).$$

Burada toplamaların içinde yer alan  $i$  ve  $j$  belirlenmiştir. İki faktör ortaya çıkar: Birinci gözlem  $(i,j)$  hücresinde ve ikinci gözlem de  $(h,k)$  hücresinde olabilir ya da tam tersi. Ordinal değişkenler için birbirinden farklı birliktelik ölçümlerinde  $\Pi_c - \Pi_d$  farkından yararlanılır.

#### 2.4.4. Birlikteliğin Ordinal Ölçümü: Gamma

İki değişken içinde çözülmüş bir ikili verildiğinde,  $\Pi_c / (\Pi_c + \Pi_d)$  ifadesi uyumun olasılığını ve  $\Pi_d / (\Pi_c + \Pi_d)$  ifadesi de uyumsuzluğun olasılığını göstermektedir. Bu olasılıklar arasındaki fark aşağıdaki eşitlikle ifade edilir<sup>36</sup>:

<sup>36</sup>

$$\pi_{ij}^{(c)} = \sum_{a<i} \sum_{b<j} \pi_{ab} + \sum_{a>i} \sum_{b>j} \pi_{ab} \quad \pi_{ij}^{(d)} = \sum_{a<i} \sum_{b>j} \pi_{ab} + \sum_{a>i} \sum_{b<j} \pi_{ab}$$

Burada  $i$  ve  $j$  toplamlar içerisinde belirlenmiştir.

$$\Pi_c = \sum_i \sum_j \pi_{ij} \pi_{ij}^{(c)} \quad \text{ve} \quad \Pi_d = \sum_i \sum_j \pi_{ij} \pi_{ij}^{(d)}$$

$$\gamma = \frac{\prod_c - \prod_d}{\prod_c + \prod_d}$$

Bu eşitliğe **gama** adı verilir (Goodman ve Kruskal 1954) ve örneklem için ise;  $\hat{\gamma} = (C - D)/(C + D)$  şeklinde hesaplanır<sup>37</sup> (Agresti, 2002). Korelasyon gibi gama da değişkenleri simetrik bir biçimde ele alır. Bir yanıt değişkeni şeklinde bir sınıf belirlemek gerekmemektedir. Ayrıca korelasyon gibi gamada -1 ile 1 aralığındadır (  $-1 \leq \gamma \leq 1$  ). Bir değişkenin kategori sıralamasının ters çevrilmesi gamanın işaretinin değişmesine neden olur. Oysa X ve Y tamamen doğrusal olduğunda korelasyonun mutlak değerinin 1 olmasına karşın,  $|\gamma|=1$  için monotonlaştırma gerekmektedir.

Eğer  $\prod_d = 0$  ise  $\gamma = 1$  ve  $\prod_c = 0$  ise  $\gamma = -1$  olur. Bağımsızlık için  $\gamma = 0$  olması gerekmektedir. Fakat  $\gamma = 0$  olduğunda bağımsızlık vardır denemez. Örnek olarak, U-şeklindeki ortak dağılımda;  $\prod_d = \prod_c$  olur dolayısı ile  $\gamma = 0$ ’ dır. Tablo 2.12’ e göre, C=1331 ve D=849 olarak hesaplanır. Buna göre;  $\hat{\gamma} = (1331 - 849)/(1331 + 849) = 0,221$  şeklinde bulunur. Gelir arttıkça iş memnuniyetinin artması ile ilgili sadece zayıf bir eğilim bulunmaktadır. Çözülmüş ikililer için, uyumlu ikililerin oranı uyumsuz ikililerin oranından 0.221 daha fazladır.

Agresti (1984, Bölüm-9 ve 10) ve Kruskal (1958) birlikteliğin ordinal ölçümlerini araştırmışlardır (Agresti, 1984). Y ordinal ve X nominal olduğunda,  $I > 2$  için, bu bölümde ele alınan hiçbir ölçüm yardımcı olamamaktadır. Ordinal modelleme yaklaşımları X’in her bir kategorisi için bir parametre kullanmaktadır;

---

<sup>37</sup> Sürekli değişkenler için, örneklem tam olarak sıralanmıştır ( yani her hangi bir bağ

oluşmaz). Bu nedenle de  $(C + D) = \binom{n}{2}$  ve  $\hat{\gamma} = (C - D) / \binom{n}{2}$  olur. Bu Kendall tau olarak adlandırılır.

parametrelerin karşılaştırılması ile  $X'$  in ikili kategorileri için ordinal yanıt karşılaştırılır.

## 2.5. Kİ-KARE BAĞIMSIZLIK TESTLERİ

Bu bölümde hücre olasılıklarının belirli  $\{\pi_{ij}\}$  değerlerine eşit olduğunu gösteren  $H_0$  sıfır hipotezinin nasıl test edileceği gösterilecektir. Hücre miktarları  $\{n_{ij}\}$  olan  $n$  hacimli bir örneklemede  $\{\mu_{ij} = n\pi_{ij}\}$  değerleri beklenen frekanslar olarak adlandırılır ve  $H_0$  doğru olduğunda da beklenen değerleri  $\{E(n_{ij})\}$  ifadesi temsil eder. Bu notasyon iki yönlü tablolarda kullanılmaktadır. Fakat çok yönlü tablolar ya da tek bir kategorik değişkenin değişik değerlerinden oluşan bir küme için de benzer notasyonlar kullanılmaktadır.

$n$  kere havaya para atma deneyi için, her bir atışta  $\pi$  tura gelme olasılığını ve  $1 - \pi$  ise yazı gelme olasılığını göstermektedir. Sıfır hipotezi  $1 - \pi = .5 = \pi$  şeklindedir. Turanın beklenen sıklığı ise  $\mu = n\pi = n/2$  olur ve aynı zamanda bu değer yazının beklenen sıklığını da göstermektedir. Eğer  $H_0$  hipotezi doğruysa, gözlemlerin yarısının yazı yarısının tura olması beklenir.

Verilerin sıfır hipotezi ile ters düşüp düşmediğini değerlendirmek için beklenen sıklıklar ile örnek hücre sayıları karşılaştırılır. Daha büyük  $\{n_{ij} - \mu_{ij}\}$  farkları sıfır hipotezine karşıt olarak daha güçlü bir kanıt oluşturur. Bu şekildeki karşılaştırmaları yapmak için kullanılan test istatistikleri büyük örnek ki-kare dağılımlarını oluşturur.

### 2.5.1. Pearson İstatistiği ve Ki-Kare Dağılımı

$H_0$  hipotezinin test edilmesi için kullanılan *Pearson Ki-Kare İstatistiği* aşağıdaki şekilde ifade edilmektedir (Agresti, 2002; Everit, 1992):



$$X^2 = \sum \frac{(n_{ij} - \mu_{ij})^2}{\mu_{ij}}$$

1900 yılında İngiliz İstatistikçi Karl Pearson tarafından oluşturulmuş ve ileri sürülmüştür. Bir çok katkısının arasında Pearson Çarpım-Momenti Korelasyonu da yer almaktadır. Tüm değerler için  $n_{ij} = \mu_{ij}$  eşitliği sağlanırsa Pearson Ki-Kare istatistiği sıfır olur ki bu minimum değeridir. Belirli bir örneklem hacmi için  $n_{ij}$  ve  $\mu_{ij}$  arasındaki fark arttıkça  $X^2$  değeri de artmaktadır ve sıfır hipotezine karşıt olarak daha güçlü bir kanıt oluşturmaktadır. Daha büyük  $X^2$  değerleri sıfır hipotezi için daha fazla karşıtlık oluşturduğu için, testin P-değeri sıfır olasılığını vermektedir ki  $X^2$  en azından gözlemlenen değer kadar büyük olur.  $X^2$  istatistiği büyük örneklem hacimleri için yaklaşık olarak bir ki-kare dağılımı gösterir. Burada büyük kavramı ile ne anlatıldığını açıkça belirtmek oldukça zordur. Fakat  $\{\mu_{ij} \geq 5\}$  olması yeterlidir. P-değeri ise gözlemlenen  $X^2$  değerinin üzerinde kalan ki-kare dağılımı sağ kuyruk olasılığıdır.

Ki-kare dağılımı serbestlik dereceleri ile belirlenmektedir. Serbestlik derecesi “df” ile gösterilir. Ki-kare dağılımının ortalaması  $df$  ve standart sapması ise  $\sqrt{2df}$  şeklindedir. Serbestlik derecesi arttıkça, dağılım daha büyük değerler etrafında yoğunlaşır ve daha fazla yayılma gösterir. Sadece negatif olmayan değerler için tanımlıdır ve sağa çarpıktır. Fakat df arttıkça daha çok çan-eğrisi (normal) şeklini alır. Serbestlik derecesi alternatif ve sıfır hipotezlerindeki parametre sayıları arasındaki farka eşittir.

### 2.5.2. Olabilirlik-Oran İstatistiği

Sıfır hipotezinin test edilmesinde kullanılan alternatif bir istatistik ise anlamlılık<sup>38</sup> testlerinde kullanılan olabilirlik oran metodundan meydana gelmektedir (Agresti, 2002). Test ile sıfır hipotezinin doğru olduğu varsayımı altında olabilirlik

---

<sup>38</sup> İstatistikte boş hipotezdeki parametreye ilişkin koşulun örnekleme dalgalanmalarından ileri gelmemesidir. Eğer istatistiksel sınama gerçek yaşamda önemli olan bir sonuç veriyor ise anlamlılık aynı zamanda önemliliği de getirir.

fonksiyonunu maksimize eden parametre değerleri belirlenir. Aynı zamanda da sıfır hipotezinin doğru olması ya da olmaması şeklinde daha genel bir koşul altında maksimize eden değerler belirlenir. Test istatistiği maksimize edilmiş olabirliklerin oranına dayanmaktadır:

$$\Lambda = \frac{\text{Parametrelerin sıfır hipotezini sağlaması durumunda}}{\text{Parametreler kısıtlanmadığı zaman maksimum olabirlik}}$$

Bu oran 1'i aşamaz. Parametreler sıfır hipotezini sağlamak için zorlanmamışken eğer maksimuma çıkarılan olabirlik çok daha büyükse,  $\Lambda$  oranı 1'in altındadır ve sıfır hipotezine karşı güçlü bir kanıt söz konusu olur (Agesti, 2002). Olabirlik oran testine göre test istatistiği,  $-2 \log(\Lambda)$  şeklinde ifade edilir. Bu değer negatif değildir ve  $\Lambda$ 'nın küçük değerleri büyük  $-2 \log(\Lambda)$  değerlerini verir. Log-dönüşümü nedeni ile test istatistiği yaklaşık bir ki-kare örneklem dağılımını verir. İki yönlü olumsuzluk (kontenjans) tabloları için aşağıdaki şekilde sadeleştirilir(Eyduran, 2005):

$$G^2 = 2 \sum n_{ij} \log \left( \frac{n_{ij}}{\mu_{ij}} \right)$$

Bu şekilde elde edilen  $G^2$  istatistiği, *olabirlik-oran ki-kare istatistiği* olarak adlandırılır. Aynı Pearson istatistiği gibi  $G^2$  istatistiği de tüm değerler için  $n_{ij} = \mu_{ij}$  ise minimum sıfır değerini alır ve daha büyük değerler sıfır hipotezine karşı daha güçlü kanıt sağlar. Pearson  $X^2$  ve olabirlik oranı  $G^2$  farklı test istatistiklerini sağlasa da birçok ortak özelliği paylaşırlar ve genellikle aynı sonuçları verirler(Eyduran,2005). Sıfır hipotezi doğru ve hücrelerdeki sayılar büyük ise iki istatistikte aynı ki-kare dağılımını verir ve aynı sayısal değerleri alırlar. Ki-kareye yakınsama  $X^2$  için  $G^2$ 'ye kıyasla daha hızlıdır.  $n / IJ < 5$  olduğunda  $G^2$  açısından bu yaklaşım genellikle yetersiz kalmaktadır. I ya da J büyük olduğunda, bazı beklenen frekanslar 1 kadar küçük olur ve çoğu beklenen frekansta 5'i aşarsa  $X^2$  için bu yaklaşım uygun ve yeterli olacaktır(Agesti, 2002). Küçük örnek metotları "n" değerinin yeteri kadar büyük olup olmadığı konusunda şüphe duyulduğu zaman

geçerli ve kullanışlıdır. Her bir istatistiğin avantaj ve dezavantajları vardır ki bunlara modellerin kurulması ve uygulanmasında değinilmektedir.

### 2.5.3. Bağımsızlık Testleri

İki yönlü olumsuzluk tablolarında, iki yanıtın istatistiksel bağımsızlığını ifade eden sıfır hipotezi tüm  $i$  ve  $j$ ' ler için aşağıdaki şekildedir:

$$H_0 : \pi_{ij} = \pi_{i+} \pi_{+j}$$

İlk önce marjinal olasılıklar ile ortak olasılıklar belirlenir. Sıfır hipotezini test etmek için beklenen frekans  $\mu_{ij} = n\pi_{ij} = n\pi_{i+}\pi_{+j}$  şeklinde hesaplanır. Burada  $\mu_{ij}$ , bağımsızlık varsayımı altında  $n_{ij}$ 'nin beklenen değeridir. Genellikle  $\{\pi_{i+}\}$  ve  $\{\pi_{+j}\}$  beklenen değer olduğundan dolayı bilinmemektedir.

Bilinmeyen olasılıklar için örnek oranları kullanarak beklenen frekanslar tahmin edilir ve aşağıdaki eşitlik kullanılır:

$$\hat{\mu}_{ij} = n p_{i+} p_{+j} = n \frac{n_{i+}}{n} \frac{n_{+j}}{n} = \frac{n_{i+} n_{+j}}{n}$$

$\{\hat{\mu}_{ij}\}$ , tahminlenen beklenen frekanslar olarak adlandırılır. Bu değerler aynı sıra ve sütunda gözlemlenen sayıların toplamıdır ve bağımsızlığın modellenmesinde kullanılırlar.  $I \times J$  boyutlu olumsuzluk tablolarında bağımsızlığın test edilmesinde kullanılan Pearson ve Olabilirlik oran istatistikleri aşağıdaki şekilde hesaplanır (Agresti, 2002):

$$X^2 = \sum \frac{(n_{ij} - \hat{\mu}_{ij})^2}{\hat{\mu}_{ij}}, \quad G^2 = 2 \sum n_{ij} \log \left( \frac{n_{ij}}{\hat{\mu}_{ij}} \right)$$

Bu istatistiklerin büyük örnek ki-kare dağılımları,  $df=(I-1)(J-1)$  şeklinde serbestlik derecesine sahiptir. Bu durum şu anlama gelmektedir: “Sıfır hipotezi altında,  $\{\pi_{i+}\}$  ve  $\{\pi_{+j}\}$  ile hücre olasılıkları belirlenir.” (I-1) tane gereksiz olmayan sıra olasılığı vardır ve toplamları da 1’ e eşittir. İlk (I-1) tanesi ile  $\pi_{I+} = 1 - (\pi_{1+} + \dots + \pi_{I-1,+})$  ifadesi kullanılarak sonuncusu belirlenir. (I-1)+(J-1) tane parametrenin toplamına ulaşmak için benzer şekilde J-1 tane sütun olasılığı hesaplanır. Alternatif hipotez ile IJ tane hücre olasılığı belirlenemez. Sadece toplamlarının 1 olması gerektiğinden dolayı (IJ-1) tane gereksiz olmayan parametre söz konusudur. Serbestlik derecesinin değeri ise alternatif ve sıfır hipotezleri altında parametre sayıları arasındaki farktır ve aşağıdaki şekilde hesaplanır:

$$(IJ-1)-[(I-1)+(J-1)]=IJ-I-J+1=(I-1)(J-1)$$

$\{\pi_{i+}\}$  ve  $\{\pi_{+j}\}$ ’ nın boyutları,  $\sum_i \pi_{i+} = \sum_j \pi_{+j} = 1$  sınırlamalarını yansıtmaktadır. R.A. Fisher (1992), Pearson’ un hatasını düzeltmiştir. Fisher makalesinde serbestlik derecesi kavramını tanıtmıştır (Pearson ki-kare dağılımları ailesinin listelenmesi ile ilgilenmiş fakat açık bir şekilde serbestlik dereceleri ile ilgilenmemiştir).

#### 2.5.4. Artıklar

Bir test istatistiği ve onun P-değeri açık bir şekilde sıfır hipotezine karşı bir kanıtı ifade etmektedir. Gözlemlenen ve tahmin edilen beklenen frekansların hücre hücre karşılaştırılması kanıtın yapısını daha iyi anlamamız açısından yardımcı olur. Hücreler için  $n_{ij}$  ve  $\hat{\mu}_{ij}$  arasında daha büyük farkların oluşması söz konusudur ki hücreler daha büyük beklenen frekanslara sahip ise sadece  $n_{ij} - \hat{\mu}_{ij}$  yeterli değildir. Bağımsızlığın test edilmesi için, hücre artıklarının kullanılması faydalıdır (Agresti, 2002) ve aşağıdaki şekilde hesaplanır:

$$\frac{n_{ij} - \hat{\mu}_{ij}}{\sqrt{\hat{\mu}_{ij}(1 - p_{i+})(1 - p_{+j})}}$$

Bu hücre artıkları “Ayarlanmış Artıklar” olarak adlandırılır (Agresti,2002). Sıfır hipotezi doğru ise her bir ayarlanmış artık, büyük örnek standart normal dağılıma sahiptir. Mutlak değer olarak hemen hemen 2 ya da 3’ü aşan bir ayarlanmış artık o hücrede sıfır hipotezinin uygun olmadığını göstermektedir. Gözlenen değerler ve tahmin edilen beklenen frekanslar aynı sıra ve sütun toplamalarına sahip olduğunda belirli bir sütun için, eğer bir hücrede  $n_{ij} > \hat{\mu}_{ij}$  ise diğer hücre için bu durumun aksi söz konusu olur.

$n_{1j} - \hat{\mu}_{1j}$  ve  $n_{2j} - \hat{\mu}_{2j}$  şeklindeki farklar aynı önem ve büyüklüğe sahiptir fakat işaretleri farklıdır. Bu o hücrelere ait ayarlanmış artıklar için aynı örneklerin kullanılmasından kaynaklanmaktadır.

### 2.5.5. Ki-Kare Parçalanımı

Ki-kare istatistikleri üretken bir özelliğe sahiptir. Eğer bir ki-kare istatistiği için  $df=df_1$  iken ayrı ve bağımsız bir diğer ki-kare istatistiği için de  $df=df_2$  ise bu iki istatistiğin toplamı bir ki-kare eder ve serbestlik derecesi de  $df= df_1 + df_2$  olur. Örnek olarak üniversite eğitimi almış ve almamış katılımcılar için tablo-2.15 gibi ayrı birer tablomuz olsun. Bu iki tablo için ayrı ayrı bulunan  $X^2$  değerlerinin toplamı ya da  $G^2$  değerlerinin toplamı yine bir ki-kare istatistiği olacaktır ve serbestlik derecesi de  $df=2+2=4$  olacaktır. Benzer şekilde, serbestlik derecesi 1’den büyük ( $df>1$ ) ki-kare istatistikleri daha az serbestlik derecelerine sahip bileşenlere ayrılabilir. Örnek olarak serbestlik derecesi  $df=2$  olan bir istatistik, serbestlik derecesi  $df=1$  olan iki bağımsız bileşene ayrılabilir.

Bağımsızlığın test edilmesi için kullanılan başka bir ilave yöntem ise ki-kare test istatistiğinin parçalanmasına dayanmaktadır (Agresti, 2002). Böylece bileşenler birlikteliğin (ilişkinin) belirli yönlerini temsil ederler. Parçalanım öncelikle belirli kategoriler ya da kategori grupları arasındaki farkları yansıtan bir ilişkiyi de

gösterebilmektedir.  $2 \times J$  tablolarında bağımsızlığın test edilmesi için kullanılan  $G^2$  istatistiğinin parçalanımını örnekle açıklayalım.

Test istatistiğinin serbestlik derecesi  $df=(J-1)$ 'dir ve biz bu istatistiği  $(J-1)$  tane bileşene parçalarız.  $j$ ' ninci bileşen bir  $2 \times 2$  tablosu için bağımsızlığın test edilmesinde kullanılan  $G^2$  istatistiğidir.  $2 \times 2$  tablosundaki ilk sütun orijinal tablodaki 1' den  $j$ ' ye kadar tüm sütunların birleştirilmesinden meydana gelmektedir.

İkinci sütun ise orijinal tabloda yer alan  $j+1$ ' inci sütundur. Bu nedenle bir  $2 \times J$  tablosunda bağımsızlığın test edilmesi için kullanılan  $G^2$  istatistiği şu şekilde hesaplanır: "İlk iki sütunun karşılaştırılması için kullanılan  $G^2$  istatistiği, ilk iki sütunun birleştirilip üçüncü sütun ile karşılaştırılması için oluşturulan bir  $2 \times 2$  tablosunda kullanılan  $G^2$  istatistiği ve aynı şekilde ilk  $J-1$  sütunun birleştirilip  $J$ ' ninci yani son sütunla karşılaştırılmasının yapıldığı bir  $2 \times 2$  tablosu için kullanılan  $G^2$  istatistiğine kadar tüm  $G^2$  istatistikleri hesaplanır ve toplanır.  $G^2$  istatistiğinin her bir bileşeninin serbestlik derecesi  $df=1$  olacaktır."

Ayrı ayrı  $2 \times 2$  tabloları için  $G^2$  istatistiğinin hesaplanması daha doğal görülebilir. Birçok veri setinde birlikteliğin (ilişkinin) araştırılması için bu yol mantıklı gözükse dahi bileşen şeklindeki bu istatistikler bağımsız değildir ve dolayısı ile tüm tabloyu temsil eden  $G^2$  istatistiğinin hesaplanması için toplanamaz (Agresti, 2002). Belirli kurallar aracılığı ile tabloların oluşturulma yolları ve buna bağlı olarak ki-kare parçalanımları belirlenir(Agresti, 2002). Fakat bu kurallar buradaki konumuzun dışında kalmaktadır. Gerekli bir koşul olarak, bileşen konumundaki tablolar için hesaplanan  $G^2$  istatistiklerinin toplamı orijinal tablo için hesaplanan  $G^2$  istatistiğini vermelidir.  $G^2$  istatistiği belirli parçalanımlara konu olurken Pearson  $X^2$  istatistiği için aynı durum söz konusu değildir (Agresti, 2002). Yani bir parçalanım da ayrı ayrı tablolar için hesaplanan  $X^2$  istatistiklerinin toplamı tüm tablo için hesaplanan Pearson  $X^2$  istatistiğini vermemektedir. Ayrı ayrı tablolar için hesaplanan  $X^2$  istatistiklerinin kullanımı geçerlidir. Fakat tüm tablo için hesaplanan  $X^2$  istatistiğinin belirli cebirsel parçalanımını açıkça sağlamazlar.

### 2.5.6. Ki-Kare Testlerinin Yorumlanması

Ki-kare bağımsızlık testleri her anlamlılık testi gibi, ciddi sınırlamalara sahiptir. Bir birlikteliğin (ilişkinin) olduğuna dair kanıtın derecesini gösterirler. Sahip olduğumuz veri setine ilişkin sorulabilecek tüm soruların cevaplanması konusunda nadiren yeterli ve uygun olurlar. Sadece bu testlerin sonuçlarına güvenmekten öte birlikteliğin yapısının incelenmesinde de kullanılmalıdır.

Ki-karenin bileşenlere ayrıştırılması, artıkların incelenmesi ve birlikteliğin gücünü ifade eden odds oranları gibi parametreleri tahmin etmek bu anlamda mantıklıdır.  $X^2$  ve  $G^2$  şeklindeki ki-kare istatistikleri aynı zamanda, uygulanabilir durumdaki çeşitli veri setleri için sınırlamalara sahiptir. Örnek olarak, bu test istatistikleri için büyük örneklemler gerekmektedir.  $X^2$  ve  $G^2$  istatistiklerinin örneklem dağılımları, tüm IJ hücrelerine ait toplam örneklem büyüklüğü “n” arttıkça ki-kareye daha da yaklaşır. Bu yakınsama  $G^2$  istatistiğine nazaran  $X^2$  istatistiği için daha hızlı bir şekilde gerçekleşir. “ $n / IJ < 5$ ” olduğunda ki-kare yakınsaması  $G^2$  istatistiği için genellikle zayıftır. I ya da J büyük olduğunda, beklenen frekanslardan bazıları 1 kadar küçük olsa dahi  $X^2$  istatistiğine göre hesaplamak daha uygundur. “n” örneklem hacminin yeterli büyüklükte olmadığı durumlarda küçük örneklem yöntemleri söz konusudur.

$X^2$  ve  $G^2$  istatistikleri için kullanılan  $\{\hat{\mu}_{ij} = n_{i+}n_{+j} / n\}$  değerleri sıra ve sütunlara ait marjinal toplamlara dayanmaktadır. Fakat listelenmiş bulunan sıra ve sütunların sırasına bağlı değildir. Bundan dolayı  $X^2$  ve  $G^2$  istatistiklerinin değerleri sıra ve sütunların keyfi olarak tekrar sıralanması neticesinde değişmez. Bunun anlamı, bu testler tüm sınıflandırmaları nominal olarak işleme tabi tutar. Ancak ordinal sınıflandırmalar arasında bağımsızlığın test edilmesi için bu istatistikleri kullanırken bazı bilgileri önemsemeyiz. En azından bir değişken ordinal olduğunda, bağımsızlık açısından daha güçlü testler söz konusudur. Bundan sonraki bölümde bu türdeki testler ele alınacaktır.

## 2.6. ORDİNAL VERİLERDE BAĞIMSIZLIĞIN TEST EDİLMESİ

$X^2$  ve  $G^2$  istatistiklerinin kullanıldığı ki-kare bağımsızlık testlerinde tüm sınıflar nominal olarak işleme tabi tutulmaktadır ki ordinal sınıflamalar arasındaki bağımsızlığın test edilmesinde bazı bilgileri yansıtamamaktadır. Sıra ve/ve ya sütunlar ordinal ise test istatistiklerinin sıralamanın sağlanmasında kullanımı daha uygundur.

### 2.6.1. Bağımsızlık için Doğrusal Eğilim (Trend) Alternatifi

Sıra değişkeni  $X$  ve sütun değişkeni  $Y$  ordinal ise “trend (eğilim)” birlikteliği oldukça yaygın olarak kullanılır (monoton eğilimin sıralı bir ölçümü kullanılır) (Agresti, 2002; Oktay, 2003 ).  $X$ 'in seviyesi arttıkça,  $Y$ ' deki etkileri daha düşük seviyelere doğru azalma ya da daha yüksek seviyelere doğru artma şeklinde olmaktadır. Bu şekildeki bir ordinal eğilim birlikteliğini tanımlamak için tek bir parametre kullanılır. En çok kullanılan analiz, kategorilere skorlar (puanlar) atanması ve doğrusal eğilim ya da korelasyon derecesinin ölçülmesi şeklindedir. Daha sonra  $X$  ve  $Y$  arasındaki ilişkide negatif ya da pozitif doğrusal eğilimlerin hesaplanmasında oldukça duyarlı bir test istatistiği gösterilecektir. Bu istatistik verilerdeki korelasyona ilişkin bilgiden yararlanmaktadır.

$u_1 \leq u_2 \leq \dots \leq u_I$  sıralara atanan skorları ve  $v_1 \leq v_2 \leq \dots \leq v_J$  ise sütunlara atanan skorları ifade etsin. Skorlar kategori seviyeleri şeklinde benzer sıralamalara sahiptir ve *monoton* niteliktedir. Skorlar kategoriler arasındaki uzaklıkları ifade etmektedir. Kategoriler arasındaki daha büyük mesafeler daha uzak kısımlar olarak işleme tabi tutulur. Gerçekleşen frekanslar yoluyla skorların çapraz çarpımının ağırlıklandırılmasında kullanılan  $\sum_{i,j} u_i v_j n_{ij}$  toplamı  $X$  ve  $Y$  arasındaki kovaryasyon ile ilişkilidir. Seçilen skorlara göre  $X$  ve  $Y$  arasındaki Pearson Çarpım-Momenti Korelasyonu bu toplamın standartlaştırılması ile elde edilir (Agresti, 2002):



$$r = \frac{\sum_{i,j} u_i v_j n_{ij} - (\sum_i u_i n_{i+})(\sum_j v_j n_{+j})/n}{\sqrt{\left[ \sum_i u_i^2 n_{i+} - \frac{(\sum_i u_i n_{i+})^2}{n} \right] \left[ \sum_j v_j^2 n_{+j} - \frac{(\sum_j v_j n_{+j})^2}{n} \right]}}$$

“r” için alternatif formüller vardır. Sıra sınıflarında her bir denek için atanan skorlar ile sütun sınıflarında her bir denek için atanan skorlar bilgisayara girilerek standart yazılım aracılığı ile bu değer hesaplanabilmektedir. Korelasyon -1 ve +1 arasında değişmektedir. Değişkenler arasındaki bağımsızlıktan söz edilebilmesi için bu değerın sifira eşit olması gerekmektedir. Mutlak değeri alınan daha büyük bir korelasyon değeri, bu doğrusal ölçü içerisinde verilerin bağımsızlıktan daha da uzaklaştığını göstermektedir. Sıfırdan farklı tam korelasyon olduğu ileri sürülen iki-yanlı alternatif hipotezine karşı bağımsızlığın ifade edildiği sıfır hipotezinin test edilmesinde kullanılan bir istatistik söz konusudur ve aşağıdaki şekilde hesaplanır:

$$M^2 = (n-1)r^2$$

Bu istatistiğın değeri örnek korelasyonu “r” ve örnek hacmi “n” büyüdükçe artmaktadır. Büyük örnekler için, bu istatistik yaklaşık olarak, serbestlik derecesi(df)=1 olan bir ki-kare dağılımı göstermektedir (Mantel 1963). Büyük değerler bağımsızlıktan uzaklaşmaya neden olmaktadır. Bundan dolayı  $X^2$  ve  $G^2$  istatistiklerinde olduğu gibi P-değeri, sağ kuyrukta gözlemlenen değerin üstündeki olasılığı göstermektedir. Sadece  $H_0$ ’ a karşı güç oluşturmaya yardımcı birliktelik için doğrusal bir birleşenin varlığını araştırılmasını sağlar. Bu istatistiğın karekökü ise,  $M = \sqrt{n-1}r$ , yaklaşık olarak standart bir normal sıfır dağılımı göstermektedir. Sınıflar arasında pozitif korelasyonun söz konusu olması durumunda bu istatistik yön belirten bir alternatif olarak uygulanmaktadır.  $M^2$  istatistiğinin kullanıldığı testler değişkenleri simetrik olarak işleme tabi tutmaktadır. Bir  $I \times J$  tablosunda, sütunlar ile sıralar ve skorlar arasında bir değişim söz konusu olursa,  $M^2$  istatistiği  $J \times I$  tablosuna uygun olarak özdeş bir değer alır.

## 2.6.2. Alkol ve Sakat Bebek Örneği

Tablo 2.13 doğuştan gelen sakatlıklar ile anne tarafından alkol kullanımına ilişkin ileriye dönük bir incelemeye aittir. Hamileliğin ilk üç ayından sonra örnekte yer alan kadınlara alkol tüketimine ilişkin bir anket uygulanmıştır. Çocuğun doğumu izlenerek, doğuştan cinsel organ sakatlıklarının olup olmadığı kayıt altına alınmıştır. Alkol tüketimi, her gün ortalama olarak içilen miktar olarak ölçülmüştür. Sakatlıklar yanıt değişkenidir ve nominal niteliktedir.

Bir değişken nominal fakat sadece iki kategoriye sahip ise  $M^2$  istatistiğindeki gibi istatistiklerde değişkenin ordinal olarak işleme tabi tutulması daha uygundur. Örnek olarak sakatlığı ordinal olarak dikkate alırız ve sakatlığın olmamasını(yok) “düşük”, sakatlığın olmasını(var) “yüksek” olarak adlandırarak işleme tabi tutarız. İki skordan herhangi birinin seçilmesi durumunda yine aynı  $M^2$  değerine ulaşmamızı sağlar. Burada “yok” seçeneği için “0” değerini, “var” seçeneği için “1” değerini kullanırız. Tablo 2.13; çok küçük, ılımlı ve son derece büyük değerlerin bir karışımından meydana gelmektedir. Örneklem büyüklüğü yeteri kadar büyük olsa dahi ( $n=32,574$ ) bu gibi durumlarda  $X^2$  ya da  $G^2$  istatistiklerinin gerçek örneklem dağılımları ki-kareye yakın olmayabilir. Bu verilere göre;  $df=4$ ,  $G^2=6.2$  ( $P=.19$ ) ve  $X^2=12.1$  ( $P=.02$ ) değerleri bulunur. Böylece bu verilerle karma sonuçlar elde edilir. Nasıl olursa olsun alkol tüketiminin sıralanabilirliği göz ardı edilir.

Tablo 2.13’de, alkol tüketiminin her seviyesi için var olan sakatlık durumlarının yüzdeleri de yer almaktadır. Bu yüzdeler tahmini olarak, artan bir eğilim olduğunu göstermektedir. İlk iki yüzdesel değer ile sonraki iki yüzdesel değer birbirlerine çok yakındır. Fakat son üç yüzdesel değere bakıldığında, her hangi bir sakatlık durumunun silinmesi ya da eklenmesi ile ciddi bir şekilde değişim olacağı görülmektedir.

**Tablo 2.13** Bebek Sakatlığı ve Annenin Alkol Tüketimi

Alkol İçimi	Sakatlık				
	Yok	Var	Toplam	%Var	Ayar. Artık
0	17,066	48	17,114	0.28	-0.18
<1	14,464	38	14,502	0.26	-0.71
1-2	788	5	793	0.63	1.84
3-5	126	1	127	0.79	1.06
≥ 6	37	1	38	2.63	2.71

Kaynak: Agresti,2002

Bu tabloda aynı zamanda, “var” kategorisine göre ayarlanmış artıklarda rapor edilmiştir. Bu değerler alkol tüketiminin en düşük seviyeleri için negatif, en yüksek seviyeleri için pozitifdir. Aynı zamanda verilerdeki ufak değişimler karşısında oldukça fazla değişen değerlerdir. Örnek yüzdeleri ve ayarlanmış artıklara göre; daha yüksek alkol tüketim seviyelerinde sakatlıklar açısından olası bir eğilimin söz konusu olduğu ileri sürülebilir. Ordinal test istatistiği  $M^2$  için, alkol tüketim seviyelerine atanacak skorların belirlenmesi gerekmektedir. Bu amaçla kategorilerin orta noktalarının skor olarak kullanılması mantıklı gözükmemektedir. Buna göre sütunlar için skorlar aşağıdaki biçimde belirlenir.

$$v_1 = 0, v_2 = 0.5, v_3 = 1.5, v_4 = 4.0, v_5 = 7.0$$

Son skor değeri az çok keyfi olarak belirlenmiştir.  $M^2$  ve  $r$  değerleri, yazılım kullanılarak hesaplanabilmektedir (SAS’ ta PROC FREQ komutu) (Stokes, Davis ve Koch, 1991). Alkol tüketimi ile sakatlıklar arasındaki örnek korelasyon değeri  $r=.014$  ve  $M^2=(32,573)(.014)^2=6.6$  olarak bulunur.

P-değeri=0.01; sıfırdan farklı bir korelasyonun olduğu konusunda güçlü bir kanıt oluşturmaktadır. Pozitif bir korelasyonun olduğunun ileri sürüldüğü tek-yanlı alternatif hipotezine göre; standart normal istatistik,  $M=2.56$  ve bu istatistiğe ait P-değeri de  $P=0.005$  olarak bulunur. Seçilen skora göre korelasyonun 0.014 olması zayıf bir ilişkinin var olduğunu yansıtmaktadır. Fakat tablolarda tanımlayıcı bir ölçü

olarak “r” sınırlı kullanımlara sahiptir. Çünkü bu tablolar oldukça kesiklidir ve dengesizlik içerirler. Gelecek bölümlerde model-temelli analiz kısmında  $M^2$  şeklindeki testler ele alınacaktır. Model-temelli yaklaşımlar hücre olasılıklarının düzeltilmiş tahminleri kadar, etkilerin büyüklüklerinin de tahminlerini elde etmemizi sağlamaktadır. Bu tahminler zayıf anlamlılık testlerine kıyasla daha fazla bilgi verici niteliğe sahiptir(Agresti, 1996).

### 2.6.3. Ordinal Test ile Ekstra Güç Sağlanması

Bağımsızlığın test edilmesi amacıyla kullanılan  $X^2$  ve  $G^2$  istatistikleri en genel alternatif hipoteze dayanmaktadır ki burada hücre olasılıkları yoluyla istatistiksel bağımsızlık türü ortaya koyulur.  $(I-1)(J-1)$  şeklindeki serbestlik derecesi ise alternatif hipotezi yansıtmaktadır. Alternatif hipotez, sıfır hipotezine kıyasla  $(J-1)(I-1)$  tane daha fazla parametreye sahiptir. Gereksiz olmayan odds oranları da birlikteliği tanımlamaktadır. Bu istatistikler eklenen parametreler için herhangi bir model türünün belirlenebilmesi amacıyla tasarlanmışlardır. Bu genellemeyi sağlamak amacıyla da bazı modellerin belirlenmesinde kullanılan bu testlerde duyarlılık ihmal edilmektedir.

Sıra ve sütun değişkenleri ordinal olduğu zaman, ilave bir parametre kullanılarak ilişki açıklanmaya çalışılır. Birliktelikte gerçekten pozitif ya da negatif bir eğilim söz konusu ise,  $M^2$  istatistiğinin kullanıldığı ordinal test yöntemi,  $X^2$  ve  $G^2$  istatistiklerine dayanan diğer testlerden daha güçlü bir avantaj sağlamaktadır.

Serbestlik derecesi (df) ki-kare dağılımının ortalaması olduğu için, serbestlik derecesi  $df=1$  olan  $M^2$  istatistiğinin görece olarak büyük bir değer alması sonucunda, serbestlik derecesi  $df=(I-1)(J-1)$ 'e sahip  $G^2$  veya  $X^2$  istatistiklerinin değerlerine kıyasla sağ kuyrukta daha uzak yerlere düşer. Gerçekten doğrusal bir eğilim söz konusu olduğunda  $M^2$  istatistiği,  $X^2$  veya  $G^2$  ile benzer değerler alır. Bu nedenle daha küçük P-değerlerini sağlayarak daha fazla güce sahip olmaktadır.

Her hangi bir bağımlılık durumunun belirlenmesinde  $X^2$  ve  $G^2$  istatistikleri, gerçekten bağımlılık söz konusu iken belirli bir bağımlılık çeşidinin ortaya çıkarılması amacıyla tasarlanmış istatistiklerin gücünü sağlayamazlar. Küçük  $df$  değerlerine sahip ki-kare testlerinin bir diğer avantajı da ki-kare yakınsamalarının doğruluğu (kesinliği) ile ilgilidir. Küçük ve ılımlı örnek hacimleri için,  $df$  daha küçük değerler aldıkça örneklem dağılımları ki-kareye daha da yakınlaşmaktadır. Bazı hücrelerdeki sayılar küçük olduğunda ki-kare yakınsaması  $M^2$  istatistiğine kıyasla,  $X^2$  ve  $G^2$  için daha kötü olabilecektir (Agresti, 2002).

Ordinal değişkenler açıkça belirlenmiş bir metriğe sahip değildir. Bağımsızlık için doğrusal bir eğilim belirlemek amacıyla  $X$  ve  $Y$ ’ ye puanlar atanması gerekmektedir. Puanlar atandıktan sonra değişkenler aralık değişkenleri şeklinde işleme tabi tutulur. Başka bir yöntem olarak katı bir ordinal analizde birlikteliğin ordinal bir ölçümü kullanılır (Gama gibi; birlikteliğin ordinal ölçümü). Büyük rassal örnekler için, örnek gama yaklaşık olarak normal bir örneklem dağılımına sahiptir. Standart hata (SE) delta metodu ile takip edilir. Gama;  $z = \hat{\gamma} / SE$  test istatistiğini kullanan, bağımsızlığın ordinal testinin temelini oluşturmaktadır. Güven aralığı ise negatif ya da pozitif monoton bir ilişkinin gücünü tanımlamaktadır. Gelir ve iş memnuniyeti üzerine oluşturulan Tablo–2.14 için,  $\hat{\gamma} = 0.221$  olduğu gösterilmişti. Örnekte daha yüksek gelir seviyelerinde iş memnuniyetinin daha yüksek olması eğilimi oldukça zayıftır.

#### 2.6.4. Skorların Seçimi

Çoğu veri seti için, skorların seçimi sonuçlar üzerinde az da olsa bazı etkilere neden olur. Monoton skorların farklı biçimlerdeki seçimleri genellikle benzer sonuçlar vermektedir (Agresti, 2002). Ancak verilerin dağılımında dengesizlikler söz konusu olduğunda benzer sonuçlar vermeyebilir. Buradaki dengesizlik bazı kategorilerin diğer kategorilere kıyasla daha fazla gözleme sahip olmasından kaynaklanır. Tablo 2.13’ te bu durum görülmektedir. Eşit uzaklıktaki satır skorlarına

(1, 2, 3, 4, 5 şeklinde) göre test istatistiği  $M^2 = 1.83$  olarak bulunmuştur ve çok daha zayıf bir hüküm vermektedir ( $P=0.18$ ).

$r$  ve  $M^2$  değerlerinin büyüklükleri, skorların dönüşümleri ile değişmediği için buradaki skor dönüşümleri kategoriler arasında yine aynı mesafeleri sağlamaktadır. Örnek olarak (1, 2, 3, 4, 5) şeklindeki skorlar, (0, 1, 2, 3, 4) ya da (2, 4, 6, 8, 10) ya da (10, 20, 30, 40, 50) şeklindeki skorlar ile aynı korelasyonu sağlamaktadır. Alternatif bir yaklaşımda ise skorların seçimi ile ilgili sorumluluktan kaçınılmaktadır ve veriler kullanılarak otomatik şekilde skorlar oluşturulmaktadır. Özellikle, denekler için sınıflar (dereceler) belirlenir ve kategorilerin skorları olarak kullanılır. Bir kategorideki tüm denekler için, örneklemin 1'den  $n$ 'e kadar tam bir şekilde sıralanmasına bağlı olarak sıraların ortalaması skor olarak atanır. Bunlara sıra ortaları adı verilir. Tablo-2.13' teki alkol tüketim seviyelerine göre sıra orta noktalarının atanmasını bir örnekle açıklayalım. Alkol tüketiminin sıfır olduğu seviyedeki 17,114 tane denek 1'den 17,114'e kadar elde edilen sıraları paylaşmaktadır. Bu deneklerin her birine bu sıraların ortalaması atanır ki burada sıra orta noktası  $(1+17,114)/2=8,557$  şeklindedir. Alkol tüketim seviyesi  $<1$  olan 14,502 tane denek 17,115'den  $17,114+14,502=31,616$ 'ya kadar elde edilen sıraları paylaşmaktadır ve sıra orta noktası  $(17,115+31,616)/2=24,365.5$  şeklinde bulunur. Benzer şekilde son üç kategori için sıra orta noktaları sırasıyla 32,013.0, 32,473.0 ve 32,555.5 şeklindedir. Bu skorlarla  $M^2 = 0.35$  değeri elde edilir ve zayıf bir sonuçtur ( $P=.55$ ).

*Niçin bu şekilde oldu?* Göreli olarak daha az gözleme sahip bitişik kategoriler aynı sıra orta noktalarına sahiptir. Örnek olarak, Tablo-2.13' e göre (8,557, 24,365, 32,013, 32,473, 32,555) şeklindeki sıra orta noktaları için son üç kategori için neredeyse aynıdır. Çünkü bu kategoriler dikkate değer biçimde ilk iki kategoriden biraz daha fazla gözleme sahiptir. Bir sonuç olarak, bu skora düzenine göre alkol tüketim seviyesi 1-2 (üçüncü kategori) alkol tüketim seviyesi 0'a (birinci kategoriye) kıyasla tüketim seviyesi  $\geq 6$ 'ne (beşinci kategoriye) çok daha yakın bir şekilde işleme tabi tutulmuştur. Bu uygun gözükmemektedir. Kategoriler arasındaki farkları yansıtabilecek skorların seçilmesi vasıtasıyla karar verilmesi daha iyi olur. Bu seçim

konusunda herhangi bir belirsizlik söz konusu olduğunda bir duyarlılık analizi oluşturulur. İki ya da üç “duyarlı” seçim yapılır ve kontrol edilir ki her biri için sonuçlar benzer olmalıdır. Eğer kategori etiketleri açık bir seçime olanak vermiyorsa, eşit uzaklıklara sahip skorların seçilmesi genellikle mantıklı bir uyuşma sağlamaktadır (Agresti, 2002). Bu kategorilere örnek olarak “liberal, ılımlı, muhafazakar” şeklindeki politik düşünce kategorileri verilebilir. X ve Y değişkenlerinden her ikisi de ordinal nitelikte ise, her biri için orta sıra skorları kullanılır. Bu durumda  $M^2$ , ye bağlı korelasyon *spearman ro* olarak adlandırılır. Spearman Ro Korelasyonunun parametrik olmayan formunun sıfırdan farklı değerlerinin saptanmasında  $M^2$  istatistiği duyarlıdır.

$J \times I$  tabloları için kullanılan alternatif ordinal testlerde diğer ordinal ilişki ölçülerinden faydalanılmaktadır. Örnek olarak Gamma ve Kendall Tau-b ölçüleri (Oktay, 2003) Kendall Tau adı verilen ordinal ölçüsünün olumsuzluk tabloları için genelleştirilmeleridir. Bağımsızlığın test edilmesinde örnek standart hatasına bölünen bu şekildeki her hangi bir ölçünün örneklem değeri büyük örneklem standart normal dağılımına sahiptir. Ayrıca istatistiğin karesi de serbestlik derecesi(df)=1 olan ki-kare dağılımı göstermektedir.  $M^2$  istatistiğine dayanan testler gibi, bu testler de potansiyel güç avantajına sahiptir ve ilişkinin(birlikteliğin) tanımlanmasında tek bir parametrenin kullanılmasından meydana gelmektedirler.

#### 2.6.5. I x 2 ve 2 x J Tabloları için Eğilim Testleri

X ya da Y değişkenleri sadece iki seviyeli olduğunda  $M^2$  istatistiğinin hesaplanmasında örnek verilerinin nasıl kullanıldığı incelenecektir. Varsayalım ki, sıra değişkeni X açıklayıcı değişken ve sütun değişkeni Y yanıt değişkeni (tepki değişkeni) olsun. X ikili değişken olduğunda, tablo  $2 \times J$  boyutlu olur. Bu boyuttaki tablolar, iki grubun karşılaştırılması amacıyla oluşturulur (satırlar iki işlemi temsil ettiğinde). Bu durumda X'in seviyeleri için ( $u_1 = 0, u_2 = 1$ ) skorları kullanılarak,  $M^2$  istatistiğine dayanan  $\sum_{i,j} u_i v_j n_{ij}$  şeklindeki kovaryasyon ölçüsünü  $\sum_j v_j n_{2j}$  şeklinde basitleştiririz (Agresti,2002). Bu ifade ikinci satırdaki tüm denekler için Y

değişkenine atanan skorların toplamıdır. Bu ifade ikinci satırdaki denek sayısına bölünerek ilgili satır için ortalama skor elde edilir.

Gerçekte sütunlar (Y) ordinal nitelikte ve  $\{v_j\}$  skorlarına sahip ise  $2 \times J$  boyutlu tablolar için  $M^2$  istatistiği Y değişkenine bağlı olarak skorların iki satır ortalaması arasındaki farklılıkların belirlenmesi amacıyla uygulanır.  $M^2$  istatistiği kullanılarak bağımsızlığın test edilmesinde, küçük P-değerleri elde edildiğinde satır ortalamaları arasında sıfırdan farklı gerçek bir fark olduğu kabul edilir. Y için sıra orta skorları kullanıldığında,  $2 \times J$  tablolarının testi, iki satır için sıra ortalamalarındaki farklılıklara duyarlıdır. Bu teste *Wilcoxon ya da Mann-Whitney Testi* adı verilir. Parametrik olmayan istatistik kaynaklarının çoğunda, tam olarak derecelenmiş (sıralanmış) yanıt verileri için bu test önerilmekte ve kullanılmaktadır. Oysa  $2 \times J$  tablosu genişletilmiş bir durumdur. Y' nin aynı seviyesindeki denek kümelerine bağlıdır ve sıra ortaları kullanılır. Büyük örneklem için parametrik olmayan testlerde standart normal z istatistiği kullanılmaktadır. z istatistiğinin karesi ise  $M^2$  istatistiğine eşit olmaktadır. Satırlar için keyfi olarak (0,1 gibi) skorlar ve sütunlar içinde sıra ortaları kullanılır.  $I \times 2$  boyutlu tablolarda iki şıklı açıklayıcı değişkenden öte ikili yanıt değişken söz konusudur. X'in farklı seviyelerine karşılık Y şeklindeki belirli bir yanıt kategorisindeki değişimde, oranların nasıl sınıflandırıldığına odaklanılması doğal olacaktır.

Monoton satır skorlarına sahip ordinal nitelikteki X değişkenine göre (iki sütun için keyfi skorlar atanmıştır)  $M^2$  istatistiği bu orantıdaki doğrusal eğilimin belirlenmesine odaklanmaktadır ve genelleştirilmiş doğrusal modellerle ilişkilidir.  $M^2$  istatistiği kullanılarak bağımsızlığın test edilmesinde, küçük P-değerleri sonucunda bu doğrusal eğilime göre eğimin sıfırdan farklı olduğu ileri sürülür.  $I \times 2$  boyutlu tablolarda Ordinal teste Cochran-Armitage Trend Testi adı verilmektedir (Agresti, 2002).



## 2.6.6. Nominal-Ordinal Tablolar

$M^2 = (n-1)r^2$  şeklinde hesaplanan test istatistiği tüm sınıfları ordinal olarak işleme tabi tutar. Bir değişken ( $X$  gibi) nominal nitelikte ama sadece iki kategoriye sahip olsa bile yine bu istatistik kullanılır. Ancak  $X$  ikiden fazla kategoriye sahip nominal bir değişken ise bu istatistik uygun olmaz ve farklı bir istatistik kullanılır. Satır ortalamaları arasındaki varyasyonun dikkate alınmasına ve her bir satırdaki ordinal değişken için ortalama bir yanıtın hesaplanmasına dayanmaktadır. Hesaplama açısından oldukça karışıktır<sup>39</sup>. Bu istatistik serbestlik derecesi,  $df=(I-1)$  değerine sahip büyük örneklem ki-kare dağılımına sahiptir.  $I=2$  için  $M^2$  istatistik değeri ile eş değerdir ve iki satır ortalamasının karşılaştırılmasında kullanılmaktadır.

## 2.7. KÜÇÜK ÖRNEKLER İÇİN KESİN YORUMLAMA

Şu ana kadar ele alınan güven aralıkları ve testler büyük örneklem için geçerliydi.  $n$  örneklem hacmi büyüdükçe, hücre sayıları artar ve  $X^2$ ,  $G^2$  ve  $M^2$  gibi “ki-kare” istatistikleri ki-kareye daha yakın dağılımlara sahip olur.

Örneklem hacmi küçük olduğunda, büyük örneklem yaklaşımlarından daha ziyade “kesin (exact) olasılık” dağılımları kullanılarak yorumlama yapılmaktadır. Bu bölümde iki yönlü olumsuzluk (kontenjans) tabloları için kesin yorumlama tartışılacaktır.

### 2.7.1. Fisher’ın Kesin (Exact) Olasılık Testi

İlk olarak  $2 \times 2$  durumu incelenecektir. Bağımsızlığı ifade eden sıfır hipotezi,  $\theta = 1$  şeklindeki odds oranı ile uyusmaktadır (Agresti, 2002). Hücre sayıları için küçük-örneklem olasılık dağılımı, gözlemlenen verilere göre aynı satır ve sütun toplamalarına sahip tablo setine göre tanımlanmaktadır. Hücre sayılarına göre kabul

---

<sup>39</sup> Bhapkar (1968) ve Yates (1948)  $M^2$ , ye benzer istatistikler önermiştir. Graubard ve Korn (1987)  $2 \times J$  tabloları için 14 istatistik listelemiştir ve korelasyon türündeki bir istatistikten yararlanmışlardır.

edilen Poisson, Binom ya da Çok terimli örneklem varsayımları<sup>40</sup> altında satır ve sütun toplamlarının sabit olduğu (marjinal toplamların her ikisinin de belirlenmesi durumunda) sınırlandırılmış tablolara uygulanan dağılıma Hipergeometrik Dağılım adı verilir. Belirli satır ve sütun marjinal toplamlarına göre,  $n_{11}$  değeri diğer üç hücre sayısını belirler. Bundan dolayı hipergeometrik dağılım formülü sadece  $n_{11}$  değeri bakımından dört hücre sayısına göre olasılıkların ifadesidir.  $\theta = 1$  iken, belirli bir  $n_{11}$  değerinin olasılığı aşağıdaki şekildedir:

$$P(n_{11} = t) = \frac{\binom{n_{1+}}{n_{11}} \binom{n_{2+}}{n_{+1} - n_{11}}}{\binom{n}{n_{+1}}}$$

Binom katsayıları:

$$\binom{a}{b} = \frac{a!(a-b)!}{b!}$$

Bağımsızlığın test edilmesi için P-değeri, alternatif hipotez için arzulanan ve gözlemlenen çıktılara göre hesaplanan hipergeometrik olasılıkların toplamıdır.  $H_a : \theta > 1$  için bunu açıklayalım. Belirli marjinal toplamlara göre, daha büyük  $n_{11}$  değerlerine sahip tablolar aynı zamanda daha büyük örnek odds oranlarına ( $\hat{\theta} = (n_{11}n_{22})/(n_{12}n_{21})$ ) sahiptir. Bu yüzden alternatif hipotez lehinde daha güçlü bir kanıt sağlayacaktır. Böylece P-değeri,  $P(n_{11} \geq t_0)$  şeklinde olup sağ-kuyruk hipergeometrik olasılığına eşittir ve  $n_{11}$  en azından gözlemlenen değer kadar büyüktür.  $t_0$ ,  $n_{11}$ 'in gözlemlenen değerini ifade etmektedir.  $2 \times 2$  boyutlu tablolar için bu test Fisher'in Kesin (*Exact*) Olasılık Testi olarak adlandırılır ve 1934 yılında ünlü İngiliz İstatistikçi R.A. Fisher tarafından önerilmiştir (Agresti, 2002).

---

<sup>40</sup> Poisson örnekleme için hiçbir şey belirlenmiş değildir. Çok terimli örneklem için sadece n belirlidir. İki sıralı bağımsız binom örnekleme için sıraların marjinal toplamları belirlenmiştir.

### 2.7.2. Fisher'in Çay Denemesi

Fisher'in 1935'te yayınlanan deney tasarımına ilişkin kaynağında bu test açıklanmıştır. Fisher şu deneyi tasarlamıştır (Agresti, 2002): Londra yakınlarındaki Rothamsted Deney Merkezinde Fisher'in bir meslektaşı bir konuda şikayetini dile getirmiştir. Çay içerken, fincana önce çay mı süt mü koyulduğunu ayırt edebilmektedir. Bunu test etmek için Fisher arkadaşına sekiz fincan çay tattırarak bir deney tasarlamıştır. Dört fincana önce süt diğer dört fincana da önce çay eklenmiştir. Arkadaşına da her iki tür içinde dörder fincan olduğu söylenmiştir. Buna göre arkadaşını ilk olarak süt eklenmiş dört fincanı seçmeye çalışmalıydı. Fincanlar da rastgele bir sırayla verilmiştir.

Tablo-2.18 deneyin mümkün sonuçlarından birini göstermektedir.  $H_a : \theta > 1$  hipotezine karşı  $H_0 : \theta = 1$  hipotezinin test edilmesinde Fisher'in Kesin Olasılık Testini kullanalım. Sıfır hipotezinde Fisher'in meslektaşının tahmini ile dökülme sırasının bağımsız olduğu ifade edilmektedir.

Alternatif hipotez onun şikayetini yansıtmaktadır ki meslektaşının tahmini ile gerçek dökülme sırası arasında pozitif bir birlikteliğin olduğu ileri sürülmektedir. Bu deneysel tasarıma göre arkadaşını ilk olarak süt eklenen dört tane fincan olduğunu bildiğinden dolayı sütun marjınları satır marjınları ile özdeştir (4,4). Tüm marjinal dağılımlar doğal olarak belirlenmiştir.  $n_{11}$ ' in sıfır dağılımı, satır ve sütun marjınları (4,4) şeklinde olan tüm  $2 \times 2$  boyutlu tablolar için tanımlanmış hipergeometrik dağılıma uymaktadır.  $n_{11}$ ' in potansiyel değerleri (0.1.2.3.4) şeklindedir. Gözlem değerlerinin yer aldığı tabloya göre sıfır olasılık aşağıdaki şekildedir (ilk süt eklenen dört fincanın üçü doğru tahmin edilmiştir):

$$P(3) = \frac{\binom{4}{3} \binom{4}{1}}{\binom{8}{4}} = \frac{[4!/(3!)(1!)] [4!/(1!)(3!)]}{[8!/(4!)(4!)]} = \frac{16}{70} = .229$$

**Tablo 2.14** Fisher'in Çay-Tatma Deneyi için Olumsuzluk Tablosu

Tahminler(ilk dökülen için)			
İlk Dökülen	Süt	Çay	Toplam
Süt	3	1	4
Çay	1	3	4
<b>Toplam</b>	4	4	

Kaynak: Agresti, 2002

Dört doğru tahminden meydana gelen tablo  $H_a : \theta > 1$  alternatif hipotezi için aşırı uçtır ve  $n_{11} = n_{22} = 4$  ve  $n_{12} = n_{21} = 0$  değerleri söz konusu olup olasılığı aşağıdaki şekildedir:

$$P(4) = \frac{\binom{4}{4} \binom{4}{0}}{\binom{8}{4}} = \frac{1}{70} = .014$$

**Tablo 2.15** Tablo 2.14'in Marjinlerine Sahip Tablolar İçin

Hipergeometrik Dağılım

$n_{11}$	Olasılık	P-değeri	$X^2$
0	.014	1.000	8.00
1	.229	.986	2.0
2	.514	.757	0.0
3	.229	.243	2.0
4	.014	.014	8.0

Tablo 2.15’ de ise  $n_{11}$  için tüm mümkün değerler ve olasılıklar özetlenmektedir. Tek yanlı alternatif hipotezi  $H_a : \theta > 1$  için P-değeri sağ kuyruk olasılığına eşittir ki  $n_{11}$  en azından gözlemlenen değer kadar büyüktür.  $P=P(3)+P(4)=0.243$  değeri ( $P(n_{11} \geq 3) = 0.243$ ) bağımsızlığı ifade eden sıfır hipotezi için yeterli bir kanıt oluşturmamaktadır. Başka bir deyişle deney gerçek dökülme sıraları ile tahminler arasında bir birliktelik oluşturmamıştır. Tabi ki bu şekilde küçük bir örnekle etkileri göstermek oldukça zordur.

Eğer çayın tadına bakan tüm fincanları doğru bir şekilde tahmin etmiş olsaydı (yani  $n_{11}=4$ ), gözlemlenen sonuç hipergeometrik dağılımın sağ kuyruğundaki en uç olasılığı olurdu. Potansiyel  $n_{11}$  değerlerine göre tablo 2.15’de  $H_a : \theta > 1$  alternatif hipotezi için P değerleri gösterilmektedir.

### 2.7.3. P-Değerleri Ve Tip-1 Hata Olasılıkları

İki yanlı alternatif hipotezi ( $H_a : \theta \neq 1$ ) ki-kare testlerinde kullanılan istatistiksel bağımlılık şeklindeki genel alternatife dayanmaktadır. Bu hipotezin kesin (exact) P-değeri genellikle şu şekilde belirlenir: Gözlemlenen tablodan çok daha farklı olmayan tabloların iki kuyruktaki olasılıklarının toplamıdır. Bunu hesaplamak amacıyla,  $P(y) \leq P(n_{11})$  ( $n_{11}$  gözlemlenen sayıdır) eşitliğine göre ilk hücre sayısı için tüm  $y$  sonuçlarının hipergeometrik olasılıkları toplanır. Tablo 2.15’ a göre,  $P(3)=0.229$  şeklindeki gözlemlenen tablo olasılığından daha büyük olmayan tüm olasılıkların toplamı,  $P=P(0)+P(1)+P(3)+P(4)= 0.014+0.229+0.229+0.014=0.486$  olarak bulunur. Satır ya da sütun marjin toplamları eşit ise, hipergeometrik dağılım simetriktir ve iki yanlı P-değeri tek yanlı P-değerinin iki katıdır.

Büyük örneklem ki-kare dağılımı yerine  $X^2$ ’ nin kesin olasılık küçük örneklem dağılımı kullanılır. Tablo-2.19’ da, tablo-2.18’ teki marjinlere sahip beş ayrı tablo için hesaplanan  $X^2$  değerleri yer almaktadır.  $X^2$ ; 0.514 olasılıkla 0.0 değerini, 0.458 olasılıkla 2.0 değerini ve 0.028 olasılıkla 8.0 değerini almaktadır. Gözlemlenen tabloda,  $X^2 = 2.0$  olup ve P-değeri de sıfır olasılık değeri kadar ya da

daha büyüktür ya da  $0.458+0.028=0.486$  şeklindedir. Bu verilere göre,  $X^2$  istatistiğine dayanan bu P-değeri olasılıklara dayanan P-değeri ile özdeşdir. Hipergeometrik dağılım için hesaplamalar daha karmaşıktır.

Bu dağılım bir kenara ayrılır ve “Yates Süreklilik Düzeltmesinin” kullanıldığı Pearson İstatistiği, ki-kare dağılımından elde edilen bir P-değerini kapsayan  $X^2$  için kesin (exact) P-değerine yaklaştırılır. Ancak bu yaklaştırmayı kullanmanın bir mantığı yoktur. Çünkü modern yazılımlar Fisher Kesin Olasılık Testinin, olasılık ölçütüne ya da  $X^2$  istatistiğine dayanan hipergeometrik P-değerleri ile birlikte büyük örneklem için dahi hesaplanmasını olanaklı kılmaktadır. Test istatistiği kesikli olduğunda 0.05 gibi bir seviyede karar vermek uyumsuzluk oluşturmaktadır. Sürekli bir dağılıma sahip test istatistikleri için P-değeri [0,1] aralığında tek düze bir dağılım göstermektedir. Bu nedenle P, “0 ve 1” arasında herhangi bir yere düşmektedir. Bundan dolayı belirli bir  $\alpha$  seviyesinin üstüne düşen P değerine göre olasılık  $\alpha$ ’ ya eşit olup P’ nin beklenen değeri de 0.5 değerini alır. Kesikli dağılımlara sahip test istatistikleri için, P-değerinin sıfır dağılımı kesiklidir ve 0.5’ ten daha büyük bir değer alması beklenir. Örnek olarak, çay-tatma verilerine göre tek yanlı test için, P-değeri 0.014’ e eşit olduğunda olasılık  $P(0)=.014$ , 0.243’ e eşit olduğunda olasılık  $P(1)=0.229$  şeklindedir. Tablo–2.19’ a göre P-değerinin beklenen değeri aşağıdaki şekildedir:

$$\sum P \times \text{Pr ob}(P) = .014(.014) + .243(.229) + .757(.514) + .986(.229) + 1.0(.014) = .685$$

Bu ortalamaya göre, kesikli dağılımlar için P-değerleri aşırı büyük olma eğilimindedir. Kesikli veriler için testlerin tutuculuğunu azaltmak amacıyla P-değerinin biraz farklı bir tanımı kullanılır: Orta-P-değeri gözlemlenen sonuç olasılığının ve aşırı uç sonuçların olasılığının yarısına eşittir (Agresti, 2002). Çay tatma deneyinden elde edilen verilere göre,  $n_{11}$  için gözlemlenen değer 3’tür. Buna göre tek yanlı orta P- değeri şu şekilde hesaplanır:  $P(3)/2+P(4)=0.229/2+0.014=0.129$ <sup>41</sup>. Hesaplanan sıradan P-değeri 0.243 şeklindedir.

---

<sup>41</sup> *orta-P-değeri* =  $(1/2)P(n_{11} = 3) + P(n_{11} > 3) = 0.129$

$X^2$  istatistiğine dayanan iki yanlı test için orta-P-değeri ise  $P(X^2=2)/2+P(X^2=8)=0.257$  olarak bulunur.

Sıradan P-değeri için ise 0.486 olarak bulunmuştu. Marjinal toplamların hepsi belirlenmiş olduğunda yüksek bir kesiklilik seviyesinin varlığı kaçınılmazdır ve Fisher'ın kesin olasılık testi en iyi yöntemdir (Agresti, 2002). Hangi marjinlerin belirlendiğine bakılmaksızın, orta-P-değerinin kullanılması kesikliliğin tutucu etkilerinin azaltılmasında yardımcı olmaktadır. Tablo 2.15'da tüm marjinler doğal olarak sabittir. Sadece bir grup sabit olduğunda (bağımsız binom örneklerine göre satır toplamlarının sabit olması şeklinde) alternatif kesin testler Fisher Kesin Olasılık Testine kıyasla daha az tutucudur. Bu testler buradaki konunun dışında kalmaktadır.

#### 2.7.4. Odds Oranı için Küçük-Örnek Güven Aralıkları

Kesin yorumlama testlerle sınırlı değildir. Aynı zamanda odds oranı için küçük örnek güven aralıkları da oluşturulur. Bu güven aralığı  $H_0 : \theta = \theta_0$  şeklinde keyfi bir değeri test eden Fisher Kesin Olasılık Testinin genelleştirilmesine karşılık gelir. %95 güven aralığı  $P>0.05$  durumunu sağlayan  $H_0 : \theta = \theta_0$  hipotezinin kesin olasılık testine göre tüm  $\theta_0$  değerlerini içine almaktadır ve 0.05 seviyesinde sıfır hipotezinin reddedilemeyeceği aralıktır. Kesin testler oluşturulduğunda kesiklilik bu güven aralıklarının tutucu olmasına neden olmaktadır.

Örnek hacmi küçük olduğunda nominal ve gerçek düzeyler arasındaki fark önemli derecede büyük olmaktadır. Tutuculuğu azaltmak için orta P-değerinin kullanıldığı teste uygun bir şekilde aralık oluşturulur. Güven aralığı 0.05 düzeyini aşan orta P-değerine sahip tüm  $\theta_0$  değerlerinden meydana gelmektedir. Bu aralık daha kısadır. Gerçek güven düzeyinin en azından nominal düzeyde olmasının garantisi olmasa da o düzeye yakın olma eğilimindedir.

Bu tür güven aralıklarından herhangi birine göre yapılan hesaplamalar karışıktır ve özel yazılımların kullanılmasını gerektirir (Örnek olarak; StataXact, Cytel Software, Cambridge,MA). Çay tatma deneyinden elde edilen verilere göre %95 “kesin(exact)” güven aralığı (0.21, 626.17) şeklindedir. Orta P-değerinin kullanıldığı teste bağlı güven aralığı ise (0.31, 308.55) şeklindedir (Agresti, 2002). Her iki aralıkta örnek hacminin küçük olmasından dolayı çok geniştir.

### **2.7.5. Daha Büyük Tablolarda Bağımsızlığın Kesin Testleri**

$2 \times 2$  boyutlu olumsallık tablolarından daha büyük tablolarda bağımsızlığın kesin olasılık testlerinde hipergeometrik dağılımın çok değişkenli biçimi kullanılır. Bu dağılım aynı zamanda gözlemlenen tablodaki gibi aynı satır ve sütun marjinlerine sahip tüm tablolarda uygulanabilmektedir. Kesin olasılık testleri el ya da hesap makinesi ile hesaplama yapılması açısından pratik değildir. Ancak bilgisayar kullanımı ile hesaplamalar yapılabilmektedir. İlk olarak  $H_0$  hipotezine göre gözlemlenen veriler arasındaki mesafeleri tanımlayan bir test istatistiği seçilir. Daha sonra test istatistiğinin en azından gözlemlenen kadar büyük olması şartına bağlı olarak tabloların olasılıkları hesaplanır. Örnek olarak nominal değişkenler için test istatistiği olarak  $X^2$  kullanılmaktaydı. Bu istatistiğin değeri en azından gözlemlenen değer kadar büyüktür ve P-değeri sıfır olasılıktır (sıfır hipotezinin kabul edildiği olasılık değeri). Çünkü büyük örnek ki-kare dağılımından öte kesin olasılık dağılımı kullanılarak hesaplama yapılmaktadır. Son zamanlarda geliştirilen yazılımlar, büyük örnek yaklaşımlarının geçerli olmadığı tablolarda kesin olasılık testlerinin uygulanmasına olanak sağlamaktadır. StatXact yazılımında kategorik verilerin kesin yorumlamalarının çoğu yapılmaktadır (Agresti, 2002).

## **2.8. ÇOK YÖNLÜ TABLOLAR ve TABLOLAŞTIRILAMAYAN YANITLAR İÇİN İLAVELER**

Bu bölüm boyunca anlatılan metotlar çok yönlü tablolar için genişletilebilir. Örnek olarak iki-yönlü tablolar için kullanılan bağımsızlık testleri, üç yönlü tablolarda kullanılan koşullu bağımsızlık testleri şeklinde genişletilmiştir. Gelecek



bölümlerde, ilgili parametrelerin ve onların istatistiklerinin tanımlanması için temel oluşturan modellerle birlikte bu şekildeki metotlar ele alınmaktadır. Şimdiye kadar verilen örneklerde, kategorik veriler olumsuzluk tabloları ile sunulmuştu. Kategorik yanıt değişkenleri için oluşturulan modeller kategorik açıklayıcı değişkenler kadar sürekli açıklayıcı değişkenlere de uygulanabilir. Bütün değişkenler ya da değişkenlerin çoğu kategorik olduğu zaman, kaynak veri dosyaları genellikle olumsuzluk tabloları değildir. Fakat her bir denek için oluşturulmuş bir veri çizelgesine sahiptirler (Agresti, 2002).

Cinsiyet, ırk, eğitim ( 1=liseden düşük, 2=lise ya da benzeri, 3=üniversite mezunu) açısından ölçülmüş deneklerin incelenmesinde yanıtları ve eşcinsellik konusundaki görüşleri (1=hoşgörülü, 2=homofobik) kapsayan bir veri dosyasının içindeki ilk üç veri çizelgesi aşağıdaki şekilde olabilir:

**Tablo 2.16** Veri Çizelgesine İlişkin Örnek Tablo

Denek	Cinsiyet	İrk	Eğitim	Görüş
1	K	B	2	1
2	E	S	3	1
3	E	B	1	2

Kaynak: Agresti, 2002

Yazılım bu türdeki veri dosyasını çözer ve sonra analizlere rehberlik eder (hatta olumsuzluk tablolarının düzenlenmesini de içerebilir).Şu ana kadar incelediğimiz bütün metotlar kategorik veriler için kullanılan modellerin içinde yer alan parametreler için istatistiksel yorumlamaları sağlamaktadır.

Paket programlar, olumsuzluk tablolarını oluşturduktan sonra burada ele alınan birçok istatistiği ve ölçüyü hesaplamaktadır. Bunlar, tablo tipine ve içerdiği verinin ölçeğine göre aşağıda özet olarak toplu halde verilmiştir:

- \*  $I \times J$  tablolarında ( $I \geq 2, J \geq 2$ ); Bağımsızlık Analizi (ki-kare bağımsızlık testi), Loglinear Analiz yapılır, Korelasyon katsayısı ve Spearman Korelasyon katsayısı hesaplanır. Ki-kare analizinde; Pearson ki-kare, Düzeltilmiş (Yates) kikare, Benzerlik Oranı (G test) ve Mantel-Haenszel kikare analizi sonuçları verilir.  $2 \times 2$  boyutlu tablolarda ise uygun koşullar oluştuğunda Fisher kesin kikare analizi yapılır.
- \* Nominal veriler içeren tablolarda; Kontenjans katsayısı, Phi katsayısı ve Gramer V katsayısı, Lambda katsayısı, Goodman-Kruskal Tau katsayısı ve Belirsizlik katsayısı hesaplanır.
- \* Ordinal veriler içeren tablolarda; Kruskal Gamma katsayısı, Somer d katsayısı, Kendal'ın Tau b ve Tau c katsayıları hesaplanır.
- \* Aralıklı ölçeklerden birleştirme ile oluşturulan nominal veriler içeren tablolarda Eta katsayısı hesaplanır.
- \* Kohort araştırma ya da Olgu kontrol araştırma sonucu oluşturulan tablolarda Relatif Risk ve Odds Ratio katsayıları hesaplanır.
- \* Karesel  $I \times J$  boyutlu tablolarda; iki gözlemcinin X olgusunu k kategoriye göre değerlendirmeleri durumunda bu değerlendirmenin uyumluluğunu belirlemek için Cohen Kappa katsayısı hesaplanır, iki gözlemci X olgusunu k kategoriye göre değerlendirir. Bulgular  $I \times J$  ( $I=J$ ) biçiminde eşit sıra ve sütunlu çapraz tablo biçiminde düzenlenir. Gözlemciler arasındaki değerlendirmenin tutarlılığı (concordance, uyumluluk) Cohen Kappa katsayısı aracılığı ile belirlenir.

## ÜÇÜNCÜ BÖLÜM

### GENELLEŞTİRİLMİŞ DOĞRUSAL MODELLER

Daha önceki bölümlerde iki yönlü ve üç yönlü olumsuzluk tablolarında birlikteliklerin analiz edilmesi için kullanılan metotlar ele alınmıştır. Bu metotlar, açıklayıcı değişkenlerin kategorik yanıt değişkenler üzerindeki etkilerini incelememize yardımcı olmaktadır. Bu bölümde ise bu tür analizlerin temelini oluşturan modellerin kullanımı incelenecektir. İyi uyumlu (good-fitting) bir model birçok fayda sağlamaktadır: Modelin yapısal şekli etkileşim ve birliktelik türlerini tanımlamaktadır. Model parametreleri için yapılan yorumlamalar ile olası etki değiştirici değişkenlerin (Confounding variables) etkileri kontrol altına alınarak açıklayıcı değişkenlerin yanıt değişken üzerindeki etkileri ortaya konulmaktadır. Tahmin edilen model parametrelerinin büyüklükleri, etkilerin önemini ve gücünü belirlemektedir.

Daha önce ele alınan yorumlamalar aynı zamanda belirli modellerdeki etkilerin analiz edilmesinde kullanılmaktadır. Modeller ile birçok açıklayıcı değişkenin etkilerinin eş zamanlı olarak analiz edilmesi gibi daha karmaşık durumlar ele alınmaktadır. Ek olarak, model oluşturma konusunda parametre tahminlerine odaklanılmaktadır ki sadece anlamlılığın test edilmesinden daha bilgi verici nitelikteki etkilerin belirlenmesi amaçlanır. Kategorik yanıt değişkenlerin modellenmesi bundan sonraki konularda ana konu olacaktır. Modellerdeki açıklayıcı değişkenler sürekli ya da kategorik ya da her iki türde de olabilmektedir.

Burada ele alınan modellerin neredeyse tümü genelleştirilmiş doğrusal modellerin özel biçimleridir. Bu modellerin genel bir durumudur ve kategorik yanıt değişkenler için oluşturulan modelleri kapsadığı gibi sürekli yanıt değişkenler için kullanılan sıradan regresyon ve ANOVA modellerini de kapsamaktadır. Bu bölümde kategorik ve diğer kesikli yanıt değişkenler için oluşturulan genelleştirilmiş doğrusal modeller ele alınmaktadır.

Genelleştirilmiş doğrusal modeller için “GLM” kısaltması kullanılmaktadır. İlk olarak tüm genelleştirilmiş doğrusal modeller için ortak olan üç bileşen

incelenmektedir. Sonra ikili yanıt deęişkenler için kullanılan modeller tanıtılmaktadır ki bu modeller binom verileri için uygulanmaktadır. Bu modellerin önemli bir türü olan lojistik regresyon modeli ayrıntıları ile ele alınmaktadır.

### 3.1. GENELLEŐTİRİLMİŐ LİNEER BİR MODELİN BİLEŐENLERİ

Tüm genelleŐtirilmiş doęrusal modeller üç bileŐene sahiptir (Agresti, 1996): *Rassal bileŐen* yanıt deęişkeni  $Y$ ’ nin tanımlanması ve  $Y$  deęişkeni için bir olasılık dağılımı varsayımının yapılmasıdır. *Sistemantik bileŐen*, modelde tahmin edici olarak kullanılan açıklayıcı deęişkenlerin belirlenmesidir. *Baęıntı* ise sistemantik bileŐen ile rassal bileŐenin beklenen deęeri (ortalama) arasındaki fonksiyonel iliŐkinin tanımlanmasıdır. GLM, doęrusal forma sahip bir eŐitlięin tahmin edilmesi aracılıęı ile açıklayıcı deęişkenlerden oluŐan bir fonksiyon ile iliŐkilidir.

#### 3.1.1. Rassal BileŐen

$N$  örnek hacmi için  $(Y_1, \dots, Y_N)$  şeklinde ifade edilen  $Y$  yanıt deęişkenine iliŐkin gözlemlerin ortaya konulmasıdır. Bu çalışmada ele alınan GLM’ ler,  $(Y_1, \dots, Y_N)$  deęerlerini baęımsız olarak işleme tabi tutmaktadır. Bir GLM’ nin rassal bileŐeni  $Y$  yanıt deęişkeninin belirlenmesinden ve  $(Y_1, \dots, Y_N)$  için bir olasılık dağılımının sečilmesinden meydana gelmektedir. Birçok uygulamada, her bir  $Y_i$  gözlemi için potansiyel sonuçlar “başarı” ya da “başarısızlık” şeklinde ikililerden meydana gelmektedir. Daha genel olarak her bir  $Y_i$  gözlemi, belirli sabit bir deneme sayısında ortaya çıkan başarıların sayısı olabilir. Sonra rassal bileŐene göre binom dağılımı varsayımı yapılır.

Bazı uygulamalarda ise her bir yanıt gözlemi kontenjans tablosunda yer alan hücre sayısı gibi negatif olmayan bir sayıdır ve buna göre rassal bileŐen için Poisson dağılımı varsayılabilir. Eđer gözlemler, bir beslenme araştırmasında deneklerin aęırlıkları gibi sürekli nitelikte ise normal bir rassal bileŐen varsayımı yapılabilir.

### 3.1.2. Sistematik Bileşen

Y değişkeninin beklenen değeri belirlenirken olasılık dağılımından yararlanılmaktadır ve  $\mu = E(Y)$  ile gösterilmektedir. GLM' de  $\mu$ ' nün değeri açıklayıcı değişkenin seviyelerine göre değişmektedir. GLM' nin sistematik bileşeni açıklayıcı değişkenlerin belirlenmesidir. Bu değişkenler eşitlik modelinin sağ tarafına tahmin edici olarak yerleştirilmektedir. Bundan dolayı da sistematik bileşen değişkenlerin belirlenmesidir ki aşağıdaki formülde  $\{x_i\}$  şeklinde yer almaktadır:

$$\alpha + \beta_1 x_1 + \dots + \beta_k x_k$$

Açıklayıcı değişkenlerin bu doğrusal bileşimi, doğrusal tahmin edici olarak adlandırılmaktadır. Bazı  $\{x_i\}$  değerleri modeldeki diğer değişkenlere bağlı olabilmektedir. Örnek olarak Y değişkeni üzerindeki etkileri bakımından  $x_1$  ve  $x_2$  arasında karşılıklı etkilerin(etkileşimin) ortaya konulduğu  $x_3 = x_1 x_2$  şeklinde ya da  $x_1$ ' in eğrisel etkisinin sağlandığı  $x_3 = x_1^2$  şeklinde olabilmektedir.

### 3.1.3. Bağntı

GLM' in üçüncü bileşeni ise rassal ve sistematik değişkenler arasındaki bağıntıdır. Doğrusal tahmin edici bakımından  $\mu = E(Y)$  'nin açıklayıcı değişkenler ile nasıl bir ilişki içinde olduğunu belirlemektedir.

$\mu$  parametresi ya doğrudan modellenmektedir ya da  $g(\mu)$  şeklinde monoton bir fonksiyon olarak modellenmektedir. Model aşağıdaki şekilde formüle edilir:

$$g(\mu) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$$

$g(\cdot)$  fonksiyonu *bağıntı fonksiyonu* (link function) olarak adlandırılır (Agresti, 1996). En basit olasılık bağıntı fonksiyonu  $g(\mu) = \mu$  formundadır. Burada ortalama

doğrudan modellenmektedir ve *özdeşlik bağıntısı* (identity link) olarak adlandırılmaktadır. Ortalama yanıtı göre doğrusal bir modelin belirlenmesini sağlamaktadır:

$$\mu = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$$

Bu sürekli yanıtlar için oluşturulan sıradan regresyon modellerinin formudur. Diğer bağıntılar, tahmin edicilerin doğrusal olmayan bir şekilde oluşturulmasına olanak tanımaktadır. Örnek olarak  $g(\mu) = \log(\mu)$  bağıntı fonksiyonu ile ortalamanın logaritması modellenir. MLogaritmik fonksiyon pozitif sayılara uygulanmaktadır. Bundan dolayı da “log bağıntısı”;  $\mu$ ’nin negatif olmadığı durumlarda tercih edilir. Log bağıntısının kullanıldığı GLM’ e *log-doğrusal model* adı verilir ve aşağıdaki formda ifade edilir:

$$\log(\mu) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$$

$g(\mu) = \log[\mu/(1 - \mu)]$  şeklindeki bağıntı fonksiyonu ile bir odds’un logaritması modellenmiştir. Buna lojit bağıntı adı verilmektedir.  $\mu$  değeri olasılık gibi “0” ile “1” arasında olduğu zaman uygundur.

Lojit bağıntısının kullanıldığı GLM türüne *lojit model* adı verilmektedir. Rassal bileşen için belirlenebilecek her bir mümkün olasılık dağılımı ortalamaya bağlı özel bir fonksiyona sahiptir ki ortalama burada *doğal parametre* (natural parameter) olarak adlandırılmaktadır. Normal dağılıma göre ortalamanın kendisidir. Poisson dağılımına göre doğal parametre ortalamanın logaritmasıdır. Binom dağılımına göre doğal parametre başarı olasılığının lojitidir. GLM’ de  $g(\mu)$  şeklindeki doğal parametrenin kullanılması ile oluşturulan bağıntı fonksiyonu *Kanonik Bağıntı* adını alır (Agresti,1996). Örnek olarak,  $\mu = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$  şeklindeki model formülü normal bir şekilde dağılmış yanıt için kanonik bağıntı ile oluşturulan GLM formudur. Poisson yanıt

verileri için kanonik bağıntılı GLM ise  $\log(\mu) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$  şeklindedir. Uygulamada en yaygın olanı kanonik bağıntılardır.

### 3.1.4. Normal GLM

Sürekli değişkenler için kullanılan sıradan regresyon ve ANOVA modelleri de GLM'lerin özel durumlarıdır. Rassal bileşene göre normal dağılım varsayımı yapılır ve ortalama,  $g(\mu) = \mu$  şeklindeki özdeşlik bağıntısı kullanılarak doğrudan modellenir. GLM, sıradan regresyon modellerini iki yöntemle geliştirir (Garson, 1999, s.1): İlk olarak, rassal bileşenin normal dağılımdan başka bir dağılıma sahip olmasına olanak tanınmaktadır. İkinci olarak ta ortalamanın birçok fonksiyonuna göre modelleme yapılmasına olanak tanınmaktadır. Bu iki geliştirmede kategorik veriler açısından önem taşımaktadır. Normal olmayan verilerin analiz edilmesinde kullanılan klasik yöntemde yanıt değişkeninin dönüşümü yapılmaktadır. Bundan dolayı da yaklaşık olarak sabit varyanslı normal dağılım göstermektedir. Daha sonra en küçük kareler yönetiminin kullanıldığı sıradan regresyon metotları uygulanabilmektedir.

Ancak sabit varyans sağlayan bir dönüşüm normalliği sağlamayabilir ya da açıklayıcı değişkenler için kullanılan basit doğrusal modeller bu ölçek için zayıf ölçüde uygunluk gösterebilmektedir (Agresti, 2002; Power ve Xie, 2000). GLM oluşturma sürecinde, rassal bileşen seçimimize göre maksimum olabilirlik metotlarından yararlanılmaktadır ve bu seçimde sadece normallik ile sınırlı olunmamaktadır (Le, 1998). Ek olarak geliştirilmiş doğrusal modellerde bağıntı seçimi rassal bileşen seçiminden ayrı tutulmaktadır. Eğer bir bağıntı etkilerin toplamsallığını sağlıyorsa (yani eğer doğrusal bir model bu bağıntıyı kullanıyorsa) normalliğin sağlanması için durağanlaştırılmış varyansa gerek duyulmamaktadır (Agresti, 1996).

GLM'ler çok çeşitli istatistiksel metotların birleştirilmesine göre ortaya konmaktadır. Regresyon, ANOVA ve kategorik veriler için oluşturulan modeller özel durumlar arasında en iyi modellerdir. Gerçekte aynı uygunluktaki metotlar tüm

GLM' ler için parametrelerin en çok olabilirlik tahminlerini (ML) vermektedir. Bu metot GLM' lerin uygunluğu için kullanılan GLIM ve SAS (PROC GENMOD) şeklindeki yazılımın temelini oluşturmaktadır (Agresti, 2002). Bundan sonraki iki bölümde, kategorik yanıt değişkenler için kullanılan en önemli iki genelleştirilmiş doğrusal model ele alınarak üç GLM bileşeni açıklanmaktadır.

### 3.2. İKİLİ VERİLER İÇİN GENELLEŞTİRİLMİŞ LİNEER MODELLER

Birçok kategorik yanıt değişkeni sadece iki kategoriye sahiptir. Örnek olarak, bir seçimde verilen oy (Demokrat, Cumhuriyetçi), otomobil seçimi (yerli, yabancı) ya da bir kadının göğüs kanseri olup olmamasına ilişkin koyulan tanı (var, yok). İkili yanıt değişkeni  $Y$  ile gösterilsin. 1 ve 0 şeklinde iki olası sonucu olsun (ya da sonuçlar daha genel olarak başarı ve başarısızlık şeklinde ifade edilir). İkili bir yanıt sıklıkla Bernoulli değişkeni olarak adlandırılır.

Dağılımı ise başarı  $P(Y=1)=\pi$  ve başarısızlık  $P(Y=0)=(1-\pi)$  olasılıkları yoluyla belirlenmektedir. Bu dağılımın ortalaması  $E(Y)=\pi$  ve varyansı  $Var(Y)=\pi(1-\pi)$  şeklindedir.  $\pi$  parametresi ile ikili bir yanıt üzerindeki  $n$  tane bağımsız gözlem için başarıların sayısı  $n$  ve  $\pi$  ile gösterilen binom dağılımına sahiptir. Bu bölümde ikili yanıt verileri için GLM' ler ele alınmaktadır. Modeldeki rassal bileşen bir binom dağılımına sahiptir. Genelleştirilmiş doğrusal modeller çoklu yanıt değişkenlere sahip olmalarına rağmen basitleştirmek amacıyla burada  $X$  gibi tek bir açıklayıcı değişken kullanılarak açıklanmaktadır.  $\pi$ ' nin değeri  $X$ ' in değeri  $x$  değiştikçe değişim göstermektedir.  $X$  değerine bağımlı olduğunu yansıtmak amacıyla  $\pi$  yerine  $\pi(x)$  notasyonu kullanılmaktadır.

#### 3.2.1. Doğrusal Olasılık Modeli

$X$  değişkeninin etkisinin modellenmesinde kullanılan bir yaklaşımda sıradan regresyon formu kullanılmaktadır ki  $Y$ ' nin beklenen değeri  $X$  değişkeninin doğrusal bir fonksiyonudur. Model aşağıdaki şekildedir ve *doğrusal olasılık modeli* olarak adlandırılmaktadır (Agresti, 2002; Le, 1998):



$$\pi(x) = \alpha + \beta x$$

$\beta$  parametresi  $x$  deęerindeki bir birimlik deęişim karşısında olasılıktaki deęişimi göstermektedir. Bu model binom rassal bileşeni ve özdeş baęıntı fonksiyonuna sahip genelleştirilmiş bir doğrusal modeldir. Ne yazık ki, bu model çok önemli yapısal bir kusura sahiptir. Olasılıklar 0 ve 1 arasında yer almaktadır. Oysa doğrusal fonksiyonlar tam gerçek doğru üzerinde deęerler almaktadır. Bu model ile yeterli derecede büyük ya da küçük  $x$  deęerlerine göre  $\pi(x) < 0$  ve  $\pi(x) > 1$  şeklinde tahminlemeler yapılmaktadır. Bu model sınırlı bir aralıkta yer alan  $x$  deęerleri için geçerlidir (Agresti,2002). Fakat çoęu uygulamada daha karmaşık bir model formu gerekli olmaktadır (özellikle çok sayıda tahmin ediciye sahip uygulamalarda).

$\pi(x) = \alpha + \beta x$  formu sıradan regresyon modeline benzemesine rağmen model parametrelerinin en küçük kareler tahmin edicileri optimal deęildir. Her bir denek için ikili sonucun varyansı,  $\text{Var}(Y)=\pi(x)[1-\pi(x)]$ , tüm  $x$  deęerleri için sabit deęildir.  $\pi(x)$  üzerindeki etkisi nedeniyle  $x$  deęerine baęlıdır. Sabit olmayan (deęişen) varyans nedeniyle maksimum olabilirlik tahmin edicileri (ML) en küçük kareler tahmin edicilerinden daha küçük standart hataya sahip olur. Bu modelde ML tahminleri (çoęu GML' deki gibi) kapalı bir forma sahip deęildir. Bu nedenden dolayı da onları temsil eden bir formül yoktur. GLM' ler için bilgisayar yazılımı olabilirlik fonksiyonunun maksimum deęerinin sayısal olarak belirlenmesi yoluyla ML tahminlerini hesaplayabilmektedir (Agresti,1996). Tahmin edilen deęerler gözlemlenen bir çok  $x$  deęeri için (0,1) aralığının dışında yer alıyorsa ML uygunluk yöntemi başarısız olur. Fakat hala en küçük kareler tahminleri hesaplanabilmektedir.

Doęrusal olasılık modelini açıklamak amacıyla, Tablo-3.1' de yer alan verileri dikkate alalım. Tabloda yer alan bilgiler Kalp rahatsızlığı için horlamanın olası bir risk olarak incelenmesi amacıyla 2484 denekten oluşan epidemiolojik bir araştırmaya dayanmaktadır. Araştırmaya dahil edilen denekler eşlerinin ne kadar horladıkları konusunda verdikleri raporlara göre sınıflandırılmıştır. Model şu şekilde oluşturulmuştur: Kalp rahatsızlığının olasılığı  $\pi(x)$ , doğrusal olarak  $x$  horlama

seviyesi ile ilişkilendirilmiştir. Tablodaki satırlar bağımsız binom örnekleri olarak işleme tabi tutulmaktadır. Parametre olarak bu örneklere dayalı olasılıklar kullanılmaktadır. horlama kategorileri için (0,2,4,5) skorları kullanılmıştır. Son iki seviye diğer bitişik ikililerden daha yakın bir şekilde işleme alınmıştır.

GLM için kullanılan yazılım aracılığı ile modelin ML uygunluğu  $\hat{\pi} = 0.0172 + 0.0198x$  şeklinde hesaplanmıştır. Örnek olarak, horlamayanlar ( $x=0$ ) için kalp rahatsızlığına sahip olan deneklerin tahmin edilen orantısı (proportion)  $\hat{\pi} = 0.0172 + 0.0198(0) = .0172$  şeklinde olur. Uydurulmuş değerler olarak bir GLM için  $E(Y)$ ' nin tahmin edilen değerlerine başvurulur.

Tablo-3.1' de doğrusal olasılık modeline göre uydurulmuş (fitted) değerler gösterilmektedir. Şekil-3.1' de ise örneklem değerleri ile uydurulmuş değerlerin grafiği verilmiştir. Tablo ve grafiğe bakıldığında modelin bu veriler için uygun olduğu söylenir. Modelin yorumlaması basittir. Kalp rahatsızlığına sahip olmanın tahmin edilen olasılığı horlamayanlar için yaklaşık olarak 0.02 (tam olarak, 0,0172) olarak bulunur, arada sırada horlayanlar için  $0,0198(0)=0.04$  kadar artış gösterir. Neredeyse her akşam horlayanlar için bu değer 0.04, her zaman horlayanlar için de 0.02 olarak bulunur. Bu etki önemlidir. 0.0198 şeklinde tahmin edilen eğim katsayısının standart hatası 0.0028 olarak bulunur.

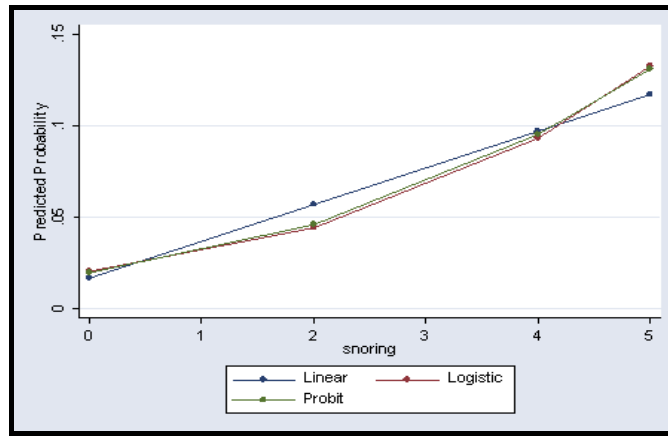
**Tablo-3.1** Horlama ve Kalp Rahatsızlığı Arasındaki İlişki

Horlama	Kalp Rahatsızlığı					
	Evet	Hayır	Evet Oranı	Doğrusal Uyum	Lojit Uyum	Probit Uyum
Asla	24	1355	0.017	0.017	0.021	0.020
Bazen	35	603	0.055	0.057	0.044	0.046
Neredeyse Her Gece	21	192	0.099	0.096	0.093	0.095
Her Gece	30	224	0.118	0.116	0.132	0.131

Kaynak: <http://www.ats.ucla.edu/stat/stata/examples/icda/snoring>

Bu veriler için skorlardaki herhangi bir artış şu şekilde sonuçlanır: Horlama seviyesi arttıkça kalp rahatsızlığına yakalanma değişimi de artmaktadır. Eğer 2484 tane veriyi 0 ve 1 şeklinde ikili gözlemler şeklinde girseydik ve ML yerine sıradan en küçük kareler yöntemini kullanarak uygun modeli oluştursaydık şu sonucu elde ederdik:  $\hat{\pi} = 0.0169 + 0.0200x$ . Uygulamada hangi modelin iyi uyum gösterdiği konusunda en küçük kareler ve ML tahminleri genellikle benzer sonuçlar vermektedir.

**Şekil 3.1** Kalp Rahatsızlığı ve Horlama Verileri için Modellerin Uyumu



### 3.2.2. Lojistik Regresyon Modeli

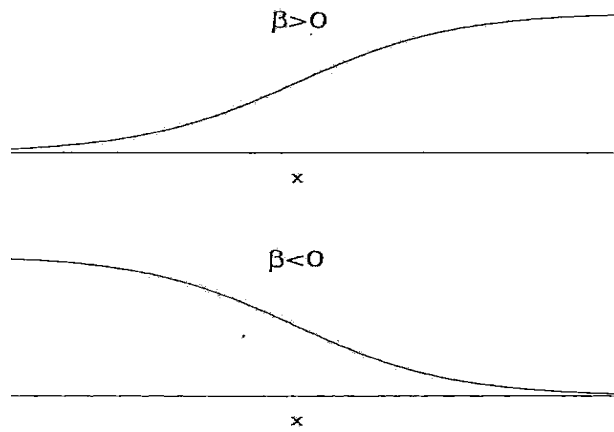
$\pi(x)$  ve  $x$  arasındaki ilişkiler genel olarak doğrusal değildir.  $X$  değişkeninde meydana gelen sabit bir değişim,  $\pi$  değerinin aralığın ortasına yakın olması durumuna kıyasla “0” ya da “1” değerine yakın olması durumunda daha az bir etkiye sahip olabilir. Örnek olarak bir otomobil satın alımında yeni ya da ikinci el arasında bir tercih yapılması söz konusudur.  $\pi(x)$ ;  $x$ =yıllık aile geliri iken yeni araba seçme olasılığını gösterebilir. Yıllık aile gelirinde 50,000\$ tutarında bir artış;  $x=1,000,000$ \$ iken ( $\pi$  değeri 1’ e yakın olduğunda),  $x=50,000$ \$ olduğundan daha az etkiye sahiptir. Uygulamada  $\pi(x)$  ve  $x$  arasında var olan doğrusal olmayan ilişkiler sıklıkla tekdüzedir:  $x$  artarken  $\pi(x)$  değeri de sürekli olarak artmaktadır ya da  $x$  artarken  $\pi(x)$  değeri de sürekli olarak azalmaktadır. S-şeklindeki eğriler Şekil-3.2’ de yer

almaktadır ki ilişki açısından gerçekçi şekillerdir. Bu şekle sahip en önemli fonksiyon aşağıdaki model formuna sahiptir:

$$\log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \alpha + \beta x$$

Bu fonksiyona lojistik regresyon fonksiyonu adı verilmektedir. Modelin formülüne göre yorumlamalar açık değildir ve daha sonra ayrıntıları ile ele alınacaktır. Lojistik regresyon modelleri genelleştirilmiş doğrusal modellerin özel durumlarıdır. (başarı, başarısızlık) şeklindeki belirlemelere göre rassal bileşen binomdur. Bağntı fonksiyonu ise  $\pi$ ' nin  $\log[\pi/(1-\pi)]$  şeklindeki lojit dönüşümüdür ve  $\text{lojit}(\pi)$  ile gösterilmektedir. Lojistik regresyon modelleri genellikle *lojit modeller* olarak adlandırılmaktadır. Lojit model, bağımlı değişkenin tahmini değerlerini olasılık olarak hesaplayarak olasılık kurallarına uygun sınıflama yapma imkanı veren, tablolaştırılmış ya da ham veri setlerini analiz eden bir istatistiksel yöntemdir (Özdamar, 1999:476).

**Şekil 3.2** Lojistik Regresyon Fonksiyonu



Lojit, binom dağılımının doğal parametresidir. Bundan dolayı lojit bağıntı kanonik bağıntıdır (Agresti, 1996).  $\pi$  değeri (0,1) aralığı ile sınırlı olsa bile lojiti herhangi bir reel sayı olabilmektedir (Gujarati,1995: 555). Reel sayılar aynı zamanda  $\alpha + \beta x$  şeklindeki doğrusal tahmin ediciler için potansiyel aralığı oluşturmaktadır ki genelleştirilmiş bir doğrusal modelin sistematik bileşenini şekillendirmektedirler. Bu nedenle bu model doğrusal olasılık modelinin sahip olduğu yapısal sorunlara sahip değildir (Agresti,1996).

Yapılacak olan çalışmada eğer bağımsız değişkenlerin bazısı sürekli veya uygun (ilgili) sınıflar içine ayrıştırılamazsa, o zaman logaritmik doğrusal analiz yerine lojistik regresyon kullanılmalıdır. Lojit modelin yukarıda verilen formülündeki  $\beta$  parametresi eğrideki artış ya da azalış oranını belirlemektedir.  $\beta > 0$  iken  $\pi(x)$  değeri x arttıkça artmaktadır,  $\beta < 0$  iken  $\pi(x)$  değeri x arttıkça azalmaktadır ki şekil-4.2' de görülmektedir.

$\beta$  parametresinin büyüklüğü eğrinin ne kadar hızla azaldığını ya da arttığını göstermektedir. Tablo-3.1' deki horlama ve kalp rahatsızlığı verilerine göre lojistik regresyon modeli için ML uygunluğu yazılım aracılığı ile şu şekilde bulunmuştur:

$$\text{lojit}[\hat{\pi}(x)] = -3.87 + 0.40x$$

$\hat{\beta} = 0.40$  şeklindeki pozitif değer horlamanın daha yüksek seviyelerinde kalp rahatsızlığı olasılığının arttığını göstermektedir. Örnek olarak model uyumu için tahmin edilen olasılıkların nasıl hesaplandığı gösterilmektedir. Tablo-4.1' de aynı zamanda bu uydurulmuş değerler de yer almaktadır. Şekil-4.1' dede bu uygunluk gösterilmektedir. Uygunluk, tahmin edilen olasılıkların dar bir aralığından öte doğrusallığa yakındır. Sonuçlar, doğrusal olasılık modeline göre bulunan sonuçlar ile aynıdır.

### 3.2.3. Probit Modeller

F, normal dağılımın kümülatif dağılım fonksiyonu (cdf) ise  $\pi(x) = F(x)$  model formu *probit model* olarak adlandırılmaktadır. Modeldeki bağıntı fonksiyonu ise *probit bağıntı* adını alır.

Probit modelin alternatif bir ifadesi ise aşağıdaki şekildedir:

$$probit[\pi(x)] = \alpha + \beta x$$

Probit bağıntı uygulanarak  $\pi(x)$  olasılığı standart normal z-skorlarına dönüşümü sağlar ki sol kuyruk olasılığı  $\pi(x)$  ' a eşittir. Örnek olarak,  $probit(.05)=-1.645$ ,  $probit(.05)=0$ ,  $probit(.95)=1.645$  ve  $probit(.975)=1.96$ . probit model binom rassal bileşenine ve probit bağıntıya sahip genelleştirilmiş bir doğrusal modeldir. Horlama ve kalp rahatsızlığı verilerini kullanarak bu durumu örnekleyelim. Probit modelin ML uygunluğu, horlama seviyesi için  $\{0,2,4,5\}$  şeklindeki skorların kullanılmasıyla aşağıdaki şekilde bulunur:

$$probit[\hat{\pi}(x)] = -2.061 + 0.188x$$

$x=0$  horlama seviyesinde uydurulmuş probit değeri;  $-2.61+0.188(0)= - 2.06$  şeklinde bulunur. Uydurulmuş (fitted) olasılık ,  $\hat{\pi}(0) -2.06$  değeri için standart normal dağılımın sol kuyruk olasılığına eşittir ve değeri de 0.020 şeklindedir.  $x=5$  horlama seviyesi için, uydurulmuş probit değeri;  $-2.61+0.188(5)= - 1.12$  şeklinde bulunur. Bu değer 0.131 şeklindeki uydurulmuş olasılık değerine benzemektedir. Uydurulmuş değerler, Tablo-3.1 ve Şekil-3.1' de gösterilmiştir. Doğrusal olasılık ile lojistik regresyon modelleri ile sağlanan değerler ile benzerdir. Pratik amaçlara göre probit ve lojistik regresyon eğrileri benzer niteliktedir. Lojistik bir regresyon modelinin uygun olduğu probit modelin zayıf kaldığı ya da karşıt olduğu veriler ara sıra söz konusu olur ve büyük örnek hacimlerini gerektirir. Parametre tahminleri iki modelde de farklıdır. Çünkü modellerdeki bağıntılar farklı ölçeklere sahiptir. Tüm

modellerin uygunluğu söz konusu olduğunda ise lojistik regresyon modellerindeki eğim katsayısının tahminleri, probit modellerdeki tahminlerin yaklaşık olarak 1.6-2.0 katıdır.

Probit dönüşümü  $\pi(x)$  değerlerine uygulanır ve  $\pi(x)$  için (ya da  $\beta < 0$  iken  $1 - \pi(x)$  için) regresyon eğrisi; ortalaması  $\mu = -\alpha / \beta$  ve standart sapması  $\sigma = 1/|\beta|$  olan normal bir cdf gibi görünmektedir. Horlama ve kalp rahatsızlığı verilerine göre probit uyumu, ortalaması  $-\hat{\alpha} / \hat{\beta} = 2.061 / 0.188 = 11.0$  ortalamalı ve  $1/|\hat{\beta}| = 1 / 0.188 = 5.3$  standart sapmalı bir normal cdf' in karşılığıdır. Probit modelin lojit modelden tek farkı, tercih olasılıklarını tanımlamada kullanılan birikimli dağılım fonksiyonudur (cdf). Lojistik CDF' den türetilen lojit modelin aksine probit model normal CDF' den türetilmektedir. Probit model 1934' te toksikolojideki modeller için ortaya çıkarılmıştır (Agresti, 1996). Lojistik regresyon modeli bu tarihten yaklaşık on yıl sonrasına kadar incelenmemiştir. Fakat şu anda probit modelden daha popülerdir. Lojit model probit modele göre daha iyi sonuçlar üretmektedir (Gujarati, 1995:563). Odds oranlarının kullanıldığı lojistik regresyonda, etkilerin yorumlanması söz konusu olduğundan dolayı popülerlik kazanmıştır (Agresti,2002). Bundan dolayı olgu-kontrol incelemelerinden elde edilen veriler için bu modeller oldukça uygundur. Çünkü bu şekildeki veriler için odds oranları tahmin edilmektedir. Probit ve lojit modeller, doğrusal olasılık modelinin bazı zafiyetleri nedeniyle uygulamada yaygın olarak kullanılır. Probit modellerin bir özelliği, bağımsız değişkenlerin bağımlı değişken üzerindeki etkilerinin doğrusal olmamasıdır. Logit ve probit gibi doğrusal olmayan modeller, bir olayın gerçekleşme olasılığı ile ilgilendikleri için eğim katsayılarının yorumlanmasında dikkatli olunması gerekir.

### **3.3. SAYILABİLİR VERİLER İÇİN GENELLEŞTİRİLMİŞ LİNEER MODELLER: POISSON REGRESYON**

Çoğu kesikli yanıt değişkeninin olası sonuçları sayılardan meydana gelmektedir. Örnek olarak, dünya üzerindeki şehirlerden oluşan bir örnekte her bir

gözlem 2007 yılında meydana gelen otomobil hırsızlıklarının sayısı olabilir ya da bilgisayar mikro kartlarının üretiminde kullanılan silisyum pullardan oluşan bir örnekte bir pul üzerindeki hata sayıları olabilir. Bu bölümde sayılabilir veriler için geliştirilmiş doğrusal modeller incelenmektedir.

Sayılar kadar Poisson değişkenleri de negatif olmayan herhangi bir tamsayı değeri alırlar. Yanıt verileri, iki ya da daha fazla kategorik yanıt değişkenine göre deneklerin çapraz olarak sınıflandırılması ile elde edilen hücre sayılarıdır. Bu bölümde alternatif bir uygulama kullanılarak poisson regresyon türü model ele alınacaktır. Tek bir yanıt değişkene göre ya oran verileri ya da sayılabilir veriler modellenecektir. Kategorik veya kesikli verilerin analizinde, normal dağılış varsayımını sağlamak için kullanılan dönüşümler çoğunlukla yetersiz kalmaktadır. Bundan dolayı üstel (exponential) dağılış ailesini esas alan Poisson regresyon analizi kullanılmaktadır (Agresti, 1996). Ayrıca normal dağılış varsayımlarını gerekli bulmadığından dolayı yoğun olarak kullanılmaktadır.

Poisson regresyon analizi, açıklayıcı değişkenler ile sayıma dayalı olarak elde edilen cevap değişkeni arasındaki ilişkiyi açıklamaktadır. Poisson regresyonunda açıklayıcı değişkenlerin doğrusal yapısını yanıt değişkenin beklenen değerine bağlayan bağıntı fonksiyonu logaritmik dönüşüm ile verilmektedir. Poisson dağılışının en temel özelliklerinden biri ortalama ile varyansının birbirine eşit olmasıdır. Ancak veri kümelerinin analizinde bu özelliğin sağlanması oldukça güçtür. Böyle durumlarda aşırı yayılım (overdispersion) ve nadir de olsa az yayılım (underdispersion) ile karşılaşılmaktadır. Aşırı yayılım, örnek varyansının örnek ortalamasında büyük olduğu durumlarda ortaya çıkmaktadır. Poisson ortalaması olan  $\mu$ 'nün logaritmasının bağımsız değişkenlerin bir doğrusal fonksiyonu olduğu varsayılmaktadır.  $\mu$ , bağımsız değişkenlerin üstel bir fonksiyon olmaktadır. Bağımlı değişkenin 0, 1, 2, 3, ... gibi kesikli değer aldığı fakat kategorik olmadığı durumlar vardır. Bu tür değişkenlere, doğalgaz boruları üzerinde kazaların sayısı, verilen patentlerin sayısı, yazlıklarda çıkan yangınların sayısı gibi örnekler gösterilebilir.



### 3.3.1. Poisson Regresyon

Poisson dağılımı pozitif bir ortalama değerine sahiptir. Özdeşlik bağıntısı kullanılarak oluşturulan genelleştirilmiş doğrusal modellerde Poisson ortalamasının modellenmesine rağmen ortalamanın logaritmik dönüşümünün modellenmesi daha yaygındır.  $\alpha + \beta x$  şeklinde ifade edilen doğrusal tahmin edici gibi ortalamanın logaritması da herhangi bir reel değer almaktadır. Log-ortalama poisson dağılımının doğal parametresidir ve logaritmik bağıntı, Poisson rassal bileşenine göre oluşturulan genelleştirilmiş doğrusal model için kanonik yapıdadır (Agresti, 1996). Bir Poisson Log-doğrusal modeli genelleştirilmiş bir doğrusal modeldir ki Y değişkeni için Poisson dağılımı varsayımı yapılır ve logaritmik bağıntı kullanılır.  $\mu$ , Y gibi bir Poisson değişkeninin beklenen değerini ve X değişkeni de açıklayıcı değişkeni temsil etsin. Buna göre Poisson Log-Doğrusal Modeli aşağıdaki forma sahiptir:

$$\log \mu = \alpha + \beta x$$

Bu modele göre ortalama üssel bir ilişki sağlanmaktadır:  $\mu = \exp(\alpha + \beta x) = e^\alpha (e^\beta)^x$  X değişkenindeki 1 birimlik bir artış,  $\mu$  üzerinde  $e^\beta$  şeklinde çarpımsal bir etkiye sahiptir. x+1 için Y değişkenin ortalaması x tane  $e^\beta$  için bulunan Y değişkenin ortalamasına eşittir.  $\beta=0$  ise  $e^\beta = e^0 = 1$  ve çarpımsal faktörü de 1 olur. Buna göre X değişkeni değiştikçe Y değişkeninin ortalaması değişmemektedir. Eğer  $\beta>0$  ise  $e^\beta > 1$  olmaktadır ve X arttıkça Y değişkeninin ortalaması da artmaktadır. Eğer  $\beta<0$  ise X arttıkça Y değişkeninin ortalaması azalmaktadır.

### 3.3.2. Oran Verileri için Poisson Regresyonu

Belirli türdeki olaylar eğer bir zaman diliminde ya da alanda gerçekleşmiş (ya da diğer büyüklük indekslerine göre) ise, gerçekleşmiş olaylardaki oranın modellenmesi ile ilgilenilir. Örnek olarak şehirlerden oluşan bir örnekleme göre 2007 yılında meydana gelen oto hırsızlıklarının sayısının modellenmesinde, şehir nüfusunun hacmine göre hırsız sayısının bölünmesi aracılığıyla her bir şehirdeki oran

belirlenebilirdi. Model ile lise mezunu nüfusun yüzdesi, ortalama gelir ve işsizlik oranı gibi açıklayıcı değişkenlere göre oranın nasıl oluşturulacağı tanımlanabilir. Y yanıt sayısı populasyon büyüklüğü şeklinde t gibi bir indekse sahiptir. Çıktılara ilişkin örneklem oranı Y/t olur. Oranın beklenen değeri  $\mu/t$  olur. Beklenen orana göre oluşturulan log-doğrusal model aşağıdaki forma sahiptir:

$$\log(\mu/t) = \alpha + \beta x$$

Bu model aşağıdaki gösterime eşittir:

$$\log \mu + \log t = \alpha + \beta x$$

Denklemin sol tarafındaki ortalamanın logaritmik bağıntısında yer alan “logt” şeklindeki düzeltme terimi *denkleştirme*(ofset) olarak adlandırılır(Agresti, 1996). Standart GLM yazılımı ile *denkleştirmelere* sahip model uyumları elde edilir.  $\log(\mu/t) = \alpha + \beta x$  şeklindeki modele göre çıktıların beklenen sayısı aşağıdaki şekilde elde edilir:

$$\mu = t \exp(\alpha + \beta x)$$

Ortalama t indeksi ile orantılıdır (açıklayıcı değişkenin değerine bağlı orantısallık sabiti ile birlikte). x’ in sabit bir değerine göre, t populasyon hacminin iki misli aynı zamanda  $\mu$  oto hırsızlarının beklenen sayısının da iki katıdır.

### 3.4. MODEL YORUMU VE MODEL KONTROLÜ

Kategorik veriler için ortaya konulan iki önemli GLM açısından istatistiksel yorumlama ve model kontrolü konularına odaklanılacaktır. İlk önce Poisson rassal bileşenine sahip GLM için açıklama yapılacaktır. İlerleyen bölümlerde ise binom rassal değişkenine sahip GLM’ ler için aynı metotlar ele alınacaktır. Kategorik yanıtlar için kullanılan çoğu GLM için, ML parametre tahminlerini elde etmek hesaplama açısından oldukça karmaşıktır. GLM yazılımı nümerik bir metot kullanmaktadır. ML tahminleri, büyük örnekler için yaklaşık olarak normal dağılım

göstermektedir. Bu nedenle model parametresi  $\beta$  için oluşturulan bir güven aralığı şu şekilde ifade edilir:  $\hat{\beta} \pm z_{\alpha/2} ASE$  (ASE:  $\hat{\beta}$ 'nin asimptotik standart hatasıdır.).

### 3.4.1. Wald, Olabilirlik Oranı (likelihood-ratio), Skor Testi

GLM'lerde yer alan parametrelere ilişkin  $H_0 : \beta = 0$  hipotezinin anlamlılık testlerinin oluşturulması için üç metot söz konusudur. En basit metotta ML tahminlerinin büyük örneklem normalliğinden yararlanılmaktadır. Test istatistiği aşağıdaki şekildedir ve  $\beta = 0$  olduğunda (yani sıfır hipotezi doğruyken) yaklaşık bir normal dağılıma sahiptir:

$$z = \hat{\beta} / ASE$$

İki yanlı ya da tek yanlı P-değerleri ile z için standart normal dağılım tablosu kullanılır.

İki yanlı alternatif hipotezi için  $z^2$ , df=1 serbestlik dereceli bir ki-kare dağılımına sahiptir ve P-değeri ise gözlemlenen değerin üzerinde kalan sağ kuyruk ki-kare olasılığıdır. Parametre tahminin standart hatasına bölünmesi ile elde edilen bu türdeki istatistiğe *Wald İstatistiği* adı verilmektedir (Agresti, 1996). İkinci metotta ise maksimize edilen iki oran arasındaki olabilirlik fonksiyonu kullanılmaktadır: (1) Sıfır hipotezi doğru iken olası parametre değerleri arasındaki maksimum oran. (2) tüm değişkenlerin dahil olduğu tam modele göre elde edilebilecek daha fazla parametre değerleri arasındaki maksimum oran (sıfır ya da alternatif hipotezinin doğru olmasını sağlayan değerlerdir).

$\ell_1$ ; tam model için olabilirlik fonksiyonunu maksimize eden değeri ve  $\ell_0$ ; sıfır hipotezi ile temsil edilen daha basit olan model için maksimize eden değeri gösterebilir. Örnek olarak, doğrusal tahmin edici  $\alpha + \beta x$  şeklinde ve sıfır hipotezi de  $H_0 : \beta = 0$  ise,  $\ell_1$  en büyük olasılıkla elde edilebilecek verilere göre

oluşturulabilecek  $(\alpha, \beta)$  kombinasyonu için hesaplanan olabilirlik fonksiyonudur.  $\ell_0$ ;  $\beta = 0$  olduğunda elde edilmesi en yüksek olasılığa sahip verilere göre  $\alpha$  ile hesaplanan olabilirlik fonksiyonudur.  $\ell_1$  her zaman en azından  $\ell_0$  kadar büyüktür. Çünkü  $\ell_0$ ;  $\ell_1$ 'in elde edildiği parametre değerlerinin sınırlandırılması aracılığı ile maksimize etmeye dayanmaktadır. Olabilirlik oran (Likelihood-ratio) test istatistiği aşağıdaki şekildedir:

$$-2 \log(\ell_0 / \ell_1) = -2[\log(\ell_0) - \log(\ell_1)] = -2(L_0 - L_1)$$

$L_0$  ve  $L_1$ ; maksimum log-olabilirlik fonksiyonlarıdır.  $\ell_0$  ve  $\ell_1$ ' in bu şekilde dönüştürülmesi ile bir ki-kare istatistiği elde edilir.  $H_0 : \beta = 0$  altında, bu istatistik aynı zamanda  $df=1$  serbestlik dereceli büyük örneklem ki-kare dağılımına sahiptir.

GLM' ler için kullanılan çoğu yazılımda maksimum log-olabilirlik değerleri ve  $-2(L_0 - L_1)$  şeklindeki olabilirlik oran istatistiği hesaplanmaktadır (Agresti,1996). Hatta bazı yazılımlarda, skor istatistiğinin kullanıldığı üçüncü bir ki-kare testi de hesaplanmaktadır ve bazen yeterli ya da etkin skor istatistiği olarak da adlandırılmaktadır.

Uygulamada kullanılan örnek büyüklükleri için olabilirlik oran testi Wald testine kıyasla daha güvenilirdir. Üç istatistiğin değerlerinde belirlenen bir ıraksama;  $\hat{\beta}$  dağılımının normallikten uzak olabileceğine işaret etmektedir. Bu durumda büyük örnek metotlarından ziyade küçük örnek metotları daha uygun olur. En iyi bilinen GLM, özdeşlik bağıntısının kullanıldığı normal verilere göre oluşturulan regresyondur ve bu regresyon için yukarıda sözü edilen üç test istatistiği de özdeş sonuçlar sağlamaktadır.

### 3.4.2. Poisson Model Kontrolü

Uyum iyiliği istatistiklerinin özetlenmesi ve artıklar; bir GLM uyumunun yeterliliğinin ve uygunluğunun araştırılmasında yardımcı olmaktadır. Poisson rassal değişkeni ile birlikte oluşturulan GLM' ler için bu durumu açıklayalım. N tane açıklayıcı değişkenden  $i$ 'si  $y_i$  ile gösterilen gözlem sayısıdır ve  $\hat{\mu}_i$  ile de uydurulmuş değerdir. Uyum iyiliğinin Pearson ve olabilirlik oran istatistikleri aşağıdaki şekildedir:

$$X^2 = \sum \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i}, \quad G^2 = 2 \sum y_i \log \left( \frac{y_i}{\hat{\mu}_i} \right)$$

$\{\hat{\mu}_i\}$  şeklinde uydurulmuş değerler göreceli olarak büyükse (5 değerini aşarsa) ve N sayısı sabit ise bu test istatistikleri yaklaşık ki-kare dağılımı göstermektedir. df serbestlik derecesi yanıt sayısından model parametre sayısının çıkarılması ile elde edilir. Bu df değeri model için *artık df* olarak adlandırılır. P-değeri sağ kuyruk olasılığıdır; büyük test istatistikleri ve küçük P-değerleri zayıf bir model uyumu göstermektedir. Fakat bu istatistikler için büyük örnek ki-kare teorisi iki açıdan geçerli olmamaktadır (Agresti,1996): İlki, gözlemlenen toplam değerlerin bir çoğu küçüktür ve bazıları sifıra eşittir. İkincisi, ki-kare yaklaşımı; N tane sabit sayıdaki açıklayıcı değişkene göre Poisson uyumlu değerleri arttırmaktadır. Böylece  $X^2$  ve  $G^2$  değerleri burada güvenilir nitelikte olmamaktadır (uyumsuzluğun ölçülmesi şeklinde). Uydurulmuş değerler her bir kategori için  $\hat{\mu}_i \geq 5$  değerini sağlamalıdır. Daha basit bir yaklaşım ise gruplandırılmış değerler ile modelin tekrar oluşturulmasından meydana gelmektedir. Bu durumda kategorilere skorların atanması gerekmektedir.

### 3.4.3. Model Artıkları

Herhangi bir GLM için, uyum iyiliği istatistikleri sadece modellerin verilere en iyi nasıl uydurulacağını özetlemektedir. Uydurulmuş ve gözlemlenen değerlerin

kıyaslanması ile daha fazla bilgi elde edilmektedir.  $i$ . gözleme göre gözlemlenen ve uydurulmuş sayı arasındaki artık farkı  $y_i - \hat{\mu}_i$  sınırlı bir fayda sağlamaktadır. *Pearson artığı* (*Pearson residual*) bu farkın standartlaştırılmasıdır ve aşağıdaki şekilde belirlenir (Agresti, 1996):

$$PearsonArtığı = \frac{(gözlemlenen - uydurulan)}{\sqrt{\hat{Var}(gözlemlenen)}}$$

Poisson GLM' lere göre  $i$ . değere göre şu şekilde basitleştirilir:

$$e_i = \frac{y_i - \hat{\mu}_i}{\sqrt{\hat{\mu}_i}}$$

Bu farkın tahmin edilen Poisson standart sapmasına bölünmesi ile standartlaştırılmaktadır. Bu artıklar  $\sum e_i^2 = X^2$  aracılığı ile belirlenen Pearson uyum iyiliği istatistiği ile ilişkilidir. Daha büyük artıklara sahip değerler model uyumunun test edilmesinde genel  $X^2$  değerine daha fazla katkı yapmışlardır. Pearson artık değerleri, yaklaşık olarak sıfırın etrafında dalgalanmaktadır.  $\mu_i$  büyük olduğunda bu dalgalanma yaklaşık olarak normal bir dağılımı izlemektedir. Uydurulmuş değer örnek verileri ile belirlendiğinden dolayı  $y_i - \hat{\mu}_i$  farkı  $y_i - \mu_i$  farkından daha küçük olma eğilimindedir. Tahmin edilen standart hataya bölünmesi ile elde edilen Pearson artığı *ayarlanmış artık* (*Adjustment Residual*) olarak adlandırılmaktadır (Agresti, 1996).  $\mu_i$  büyük olduğunda bu yaklaşık bir normal dağılıma sahip olur. Bu nedenle ayarlanmış artıklara göre bir sapmanın  $y_i - \hat{\mu}_i$  "büyük" olduğunu söylemek daha kolaydır. Kategori sayısı büyük olduğunda bu hacimdeki bazı değerlerin gerçekleşmesi beklense de, mutlak değer olarak yaklaşık olarak 2'den daha büyük ayarlanmış artıklar dikkat gerektirmektedir.

Ayarlanmış artıklar Pearson artıklarına kıyasla daha fazla tercih edilmektedir. Daha önceki bölümde iki yönlü kontejans tablolarında bağımsızlığın test edilmesi için ayarlanmış artıklar açıklanmıştır. Poisson GLM' ler için ayarlanmış artığın genel formu aşağıdaki şekildedir:

$$\frac{(y_i - \hat{\mu}_i)}{\sqrt{\hat{\mu}_i(1-h_i)}} = \frac{e_i}{\sqrt{(1-h_i)}}$$

$h_i$ :  $i$ . gözlemin kaldıraç gücü (leverage) olarak adlandırılmaktadır. Kaldıraç gücü için oluşturulan formül oldukça karışıktır. Kaldıraç gücü arttıkça ilgili gözlemin, model uyumunu etkilemede daha fazla potansiyele sahip olduğu şeklinde kabaca bir yorum yapılabilir. Çoğu GLM yazılımında (SAS'ta PROC GENMOD komutu) GLM' ler için ayarlanmış artıklarda raporlanmaktadır (Stokes, Davis ve Koch, 1991). Regresyon modellemesinde kullanılan diğer teşhis araçları aynı zamanda GLM uyumlarının değerlendirilmesinde de yardımcı olmaktadır. Örnek olarak bir gözlemin tüm modelin uyumu üzerindeki etkisini belirlemek amacıyla öncelikle bu gözlem silinerek model tekrar oluşturulur. Bu sadece ilgili gözleme dayanan bir parametreye sahip daha karışık bir modelin uygunluğuna karşılık gelmektedir. İlk olarak uyum iyiliğindeki değişim için parametre tahminleri ile uydurulmuş değerlerdeki değişimler incelenir.

#### 3.4.4. Poisson Regresyonunda Aşırı Yayılım

Eğer yanıt dağılımı gerçekten Poisson ise genellikle sayılabilir veriler, yanıt değerlerinde beklenenden daha fazla değişkenlik göstermektedir. Örnek olarak varyanslar ortalamadan daha büyük olabilir. Oysa Poisson dağılımlarında ortalama ve varyans aynıdır. Rassal bileşene göre tahmin edilenden daha fazla değişkenliğe sahip bir GLM olayına “Aşırı Yayılım (overdispersion)” adı verilmektedir (Agresti, 1996). Aşırı yayılımın en yaygın nedeni denekler arasındaki heterojenliktir. Eğer tüm ilgili değişkenler kontrol altına alınırsa varyans ortalamaya eşit olur. Sadece biri kontrol altına alınırsa varyans ortalamayı aşar. Böyle durumlarda bilinen Poisson regresyonu uygulanamamaktadır. Çünkü GLM ile elde edilen parametre değerleri sapmalı olmaktadır. Bunun yerine extra- Poisson değişimini açıklayan yayılım parametresini içeren regresyon analizleri ya da Quasi-olabilirlik yöntemleri kullanılmalıdır. Söz konusu regresyon analizleri, negatif binomial karışık Poisson regresyon ve karışıklı Poisson regresyon analizleri olmaktadır (Yeşilova,2003).

Aşırı yayılım normal bir şekilde dağılmış rassal bir bileşene göre oluşturulan sıradan regresyon modellerinde herhangi bir sorun oluşturmamaktadır. Çünkü normal dağılım değişkenliğinin tanımlanması amacıyla ortalama dan ayrı bir parametreye sahiptir (yani varyans,  $\sigma^2$  ). Fakat Binom ve Poisson dağılımları için varyans ortalamanın bir fonksiyonudur. Dolayısı ile Poisson ve binom sayılarının modellenmesinde aşırı yayılım ortak bir sorundur. Kesikli verilerdeki aşırı yayılım ile ilgili birçok metot söz konusudur. Ancak bu metotlar konumuz dışında kalmaktadır (Başvuru kaynakları; Collet(1992) ve Morgan (1992)). Sadece temel bir yönteme dikkat çekilecektir. Model kabul edildiğinde serbestlik derecesi yaklaşık olarak  $X^2$ ' nin beklenen değeridir. Gerçek yanıt dağılımı Poissondan daha değişken ise açıklayıcı değişkenler için beklenen değere göre oluşturulan model formu doğru olmamaktadır. Eğer yanıt varyansı ortalama ile orantılı (Poissonda gerektiği gibi eşit olmaktan öte) ise  $X^2/s.d.$  ile orantısallık sabiti tahminlenir (Agresti, 2002).

### **3.5. GENELLEŞTİRİLMİŞ LİNEER MODELLERİN UYUMU**

Genelleştirilmiş doğrusal modeller için model uyumunu ele alarak bu bölümü bitireceğiz. İlk olarak model parametrelerinin ML tahminlerini hesaplamak için kullanılan bir algoritma tanıtılacaktır. Olabilirlik fonksiyonunun kullanıldığı temel yorumlamanın nasıl yapıldığı konusunda daha fazla detay verilecektir. Son olarak GLM' lerin uyumunu özetleyen ve *sapma* adı verilen bir ölçü tartışılacaktır.

#### **3.5.1. Newton-Raphson Algoritması**

Birçok GLM için ML parametre tahminlerinin belirlenmesinde kullanılan eşitlikler doğrusal olmayıp tahminler kapalı formda bir ifadeye de sahip değildir. Yazılımlar doğrusal olmayan eşitliklerin çözümü için yinelemeli bir algoritmanın kullanıldığı tahminleri hesaplamaktadır. Algoritmada öncelikle olabilirlik fonksiyonunu maksimize eden parametre değerleri için bir başlangıç tahmini gerekmektedir.



Algoritma ile elde edilen ardışık yaklaşımlar ML tahminlerine daha yakın olma eğilimindedir. Bunu yapmak için kullanılan popüler bir algoritma söz konusudur ki “Fisher Skorlaması” adı verilmektedir (Agresti,1996). İlk olarak R.A. Fisher tarafından probit modellerin uygunluğu amacıyla orta konmuştur. Binomial Lojistik Regresyon ve Poisson Log-doğrusal modeller için, Fisher Skorlaması Newton-Raphson Algoritması olarak adlandırılan genel amaçlı bir metodu basitleştirmektedir. *Newton-Raphson Algoritması*; daha basit bir polinomial fonksiyon aracılığı ile yapılan ilk tahminin bir komşusu bölgesinde log-olabilirlik fonksiyonuna yaklaşmaktadır ki içbükey parabol bir şekle sahiptir (tepe-şeklinde). İlk tahminde Log-olabilirlik fonksiyonunda olduğu gibi aynı eğime ve çarpıklığa sahiptir. Polinomial yaklaşımın maksimum bölgesini belirlemek kolaydır. İlgili bölge ML tahminlerinin ikinci tahmininden oluşmaktadır. Başka bir iç bükey parabolik fonksiyon aracılığı ile yapılan ikinci tahminin komşu bölgesinde de log-olabilirliğe yaklaşma söz konusudur. Üçüncü tahmin ise maksimum bölgedir.

Ardışık yaklaşımlar genellikle birkaç döngüde hızlı bir şekilde ML tahminlerine yakınsamaktadır. GLM’ ler için kullanılan yazılım ilk tahmini sağlamak konusunda kullanıcıya gerek duymamaktadır. *Newton-Raphson* metodunda yer alan her bir döngü ağırlıklandırılmış en küçük kareler uygunluğunun bir türünü temsil etmektedir (Agresti, 2002). Bu sıradan en küçük karelerin bir genelleştirilmesidir ve GLM’ lerde Y değişkeninin sabit olmayan varyansını dikkate almaktadır. Ağırlıklar her döngüde az da olsa değişim göstermektedir. Bu nedenle de ML tahminleri ve varyans tahminleri için yaklaşım düzeltilir. GLM’ ler için ML tahmini arada sırada ***tekrarlı yeniden ağırlıklandırılmalı en küçük kareler*** olarak adlandırılmaktadır. Newton-Raphson metodunda *bilgi matrisi* adı verilen bir matris kullanılmaktadır ki parametre tahminleri için ASE değerlerinin elde edilmesini sağlamaktadır. Standart hatalar, bilgi matrisinin tersi alınarak elde edilen yeni matristeki köşegen elemanların karekökünden meydana gelmektedir. Log-olabilirliğin daha fazla çarpıklığa sahip olması durumunda daha küçük standart hatalar elde edilir. Yazılım aracılığıyla bilgi matrisi ve ilgili standart hatalar hesaplanabilmektedir.

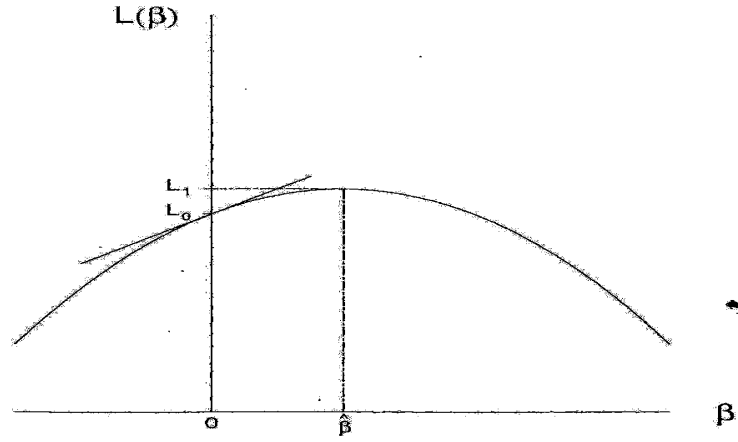
### 3.5.2. Kullanılan Olabilirlik Fonksiyonunun Yorumlanması

Daha önce bir GLM model parametresi için oluşturulan  $H_0 : \beta = 0$  hipotezinin test edilmesi amacıyla üç metot tanıtılmıştı: **Wald testleri, Olabilirlik Oran testleri ve Etkin Skor testleri**. Şekil-3.3' de  $\beta$  parametresinin bir fonksiyonu olarak log-olabilirliğin (L) genel bir çizimi gösterilmiştir ve üç test örneklenmiştir. **Wald testi**, ML tahmini olan  $\hat{\beta}$  için log-olabilirlik fonksiyonunun davranışına dayanmaktadır ve  $(\hat{\beta} / ASE)^2$  şeklinde bir ki-kare formuna sahiptir.  $\hat{\beta}$ 'nin ASE' si log olabilirlik fonksiyonunun maksimize olduğu noktada log-olabilirlik fonksiyonunun çarpıklığına bağlıdır.

**Etkin Skor Testi** ise  $\beta=0$  şeklindeki sıfır değerinde (sıfır hipotezi altında) log-olabilirlik fonksiyonunun davranışına dayanmaktadır. Burada log-olabilirlik fonksiyonunun türevinin (eğimdir) büyüklüğü kullanılmaktadır. Poisson log-doğrusal modeller ve binomial lojistik modeller de dahil bazı GLM' ler için log-olabilirlik fonksiyonu iç bükey bir şekle sahiptir (Agresti, 2002).  $\hat{\beta}$ , ML tahmini türevin sıfıra eşit olduğu noktadır.

Skor istatistiği, türevin ASE' ye oranının karesidir (Garson, 1999). Aynı zamanda 1 serbestlik dereceli bir ki-kare yaklaşımına da sahiptir. Skor istatistikleri için kullanılan genel formül burada gösterilmeyecektir. Genellikle skor istatistikleri model uygunluğu olmaksızın hesaplanmaktadır. Bir diğer avantajı ise  $\hat{\beta}$ , ML tahmini sonsuz olduğunda da hesaplanabilmesidir. Ancak bu durumda Wald istatistiği hesaplanamamaktadır. **Olabilirlik Oran testi**;  $\beta=0$  şeklindeki sıfır değeri ile  $\hat{\beta}$  değerinin her ikisi için hesaplanan log-olabilirlik fonksiyonuna ilişkin bilgileri birleştirmektedir (Agresti, 2002).

Şekil 3.3 Wald, Olabilirlik Oranı ve Skor testlerinde Kullanılan Bilgi



$\hat{\beta}$  için hesaplanan log-olabilirlik değeri ( $L_1$ ) ile  $\beta=0$  için hesaplanan log-olabilirlik değerini ( $L_0$ ) karşılaştırmaktadır ve  $-2(L_1 - L_0)$  şeklindeki ki-kare istatistiğini kullanmaktadır. Şekil-4.3’ de bu istatistik,  $\beta=0$  ve  $\hat{\beta}$  için hesaplanan log-olabilirlik fonksiyon değerleri arasındaki dikey uzaklığın iki katıdır. Bu durumda bu istatistik üç tür istatistikten elde edilen bilginin çoğunu kullanmaktadır ve genellikle en güvenilir istatistiktir.

### 3.5.3. Sapma

Normal-yanıt modelleri için iç-içe geçmiş regresyon modellerinin karşılaştırılmasında verilerdeki değişkenliği temsil eden kareler toplamının bir parçalanımından yararlanılmaktadır. Bu Varyans Analizi GLM’ ler için “Sapma Analizinin” geliştirilmesidir (Agresti, 1996).

$L_M$  ilgilenilen model için maksimize edilmiş log-olabilirlik değerini ifade etmektedir.  $L_S$  en karışık model (her bir açıklayıcı değişken kümesinde ayrı bir parametre değeri alan) için maksimize edilen log-olabilirlik değerini ifade etmektedir ki burada sözü edilen model *doygun* modeldir (Agresti, 1996; Garson, 1999). Bir model için sapma aşağıdaki şekilde tanımlıdır:

$$sapma = -2[L_M - L_S]$$

Sapma doygun model ile M modelinin karşılaştırılmasında kullanılan olabilirlik oran istatistiğidir ki doygun modeldeki tüm parametrelere ilişkin hipotezin test edilmesinde kullanılmaktadır. Bu kaynaktaki modeller için sapma M modelinde uyum iyiliği istatistiği olan  $G^2$  olabilirlik oranı ile aynı formadadır. Birçok GLM için sapma yaklaşık olarak bir ki-kare dağılımı göstermektedir. Örnek olarak başarı ve başarısızlığa ilişkin nispeten büyük sayılar ve açıklayıcı seviyelerin sabit bir değeri ile birlikte iki sonuçlu regresyon modelleridir. Bu şekildeki modeller için model uyumunun test edilmesinde sapma kullanılmaktadır (Hosmer ve Lemeshow, 2000). Her iki durumda da, “Artık Serbestlik Derecesi = Yanıt Sayısı - Gereksiz Olmayan Parametre Sayısı” şeklindedir. Yanıt sayısı Poisson değerleri ya da binom başarı toplamlarıdır. Sapma bileşenleri *sapma artıkları* (deviance residual) olarak adlandırılmakta ve her bir gözlemin uyumsuzluğunun teşhisi için ölçümler sağlamaktadır (Agresti, 2002). Bunlar ayarlanmış artıklara ve Pearson artıklarına alternatiftir. Pearson artıkları gibi sapma artıkları da yaklaşık olarak normal dağılmaktadır (Agresti, 2002). İki model için varsayalım ki  $M_0$ ;  $M_1$ ' in özel bir durumu olsun. Daha karışık model söz konusu olduğunda daha basit modelin testi için olabilirlik oran istatistiği aşağıdaki gibidir:

$$-2[L_0 - L_1] = -2[L_0 - L_S] - \{-2[L_1 - L_S]\} = Sapma_0 - Sapma_1$$

Model sapmalarının karşılaştırılması aracılığı ile modeller karşılaştırılır. Büyük örnekler için bu yaklaşık bir ki-kare istatistiğidir ve serbestlik derecesi ayrı modellerin artık serbestlik dereceleri arasındaki farka eşittir. Bu serbestlik derecesinin değeri  $M_1$ ' e gereksiz olmayan ilave parametre sayısına eşittir. Bu olabilirlik oran testi; normal dağılımlı yanıtlar için oluşturulan regresyon modellerinin karşılaştırılmasında kullanılan F testinin GLM benzetimidir.

### 3.5.4. GLM' lerin Avantajları

Son 20 yıldır GLM teorisinde sağlanan gelişim sürekli ve kategorik yanıt değişkenleri için kullanılan önemli modellerin birleştirilmesi ile sağlanmıştır. Teorik mantığa göre, bir GLM' in tanımlanmasında kullanılan rassal bileşen bir üstel dağılım ailesine sahip olmalıdır. Bu sınırlama önemli değildir. Çünkü bu dağılım ailesi, Poisson, binom, normal ve gamma gibi çok önemli dağılımları da kapsamaktadır.

Tablo-3.2'te pratik bir uygulama için bazı popüler GLM' ler listelenmiştir. GLM' lerin en iyi özelliği model uyumu algoritmasıdır (Fisher skorlaması) ve her hangi bir GLM için aynıdır. Bağlantı fonksiyonunun seçimi ya da rassal bileşenin dağılımının seçimine bağlıdır. Bu nedenle SAS' taki GLIM ve PROC GENMOD gibi bir GLM yazılımı kullanışlı modellerin çok geniş bir alanına uygulanabilmektedir ((Stokes, Davis ve Koch, 1991).

**Tablo-3.2** İstatistiksel Analizde Model Türleri

Rassal (Random Component)	Bağlantı (Link)	Sistemantik Bileşen (Systematic Component)	Model
Normal	Özdeşlik	Sürekli	Regresyon
Normal	Özdeşlik	Kategorik	Varyans Analizi
Normal	Özdeşlik	Karışık	Kovaryans Analizi
Binom	Lojit	Karışık	Lojistik Regresyon
Poisson	Log	Karışık	Log-doğrusal
Multinomial	Genelleştirilmiş Lojit	Karışık	Multinomial Yanıt

## DÖRDÜNCÜ BÖLÜM

### LOJİSTİK REGRESYON

Bu bölümde iki sonuçlu yanıt değişkenlerin istatistiksel modellemesi daha yakından incelenecektir; her bir denek için yanıt ölçümü başarılı ya da başarısız şeklindedir. İki sonuçlu veriler kategorik verilerin kullanılan en yaygın şeklidir ve bu bölümdeki metotlar temel anlamda öneme sahiptir. İki sonuçlu veriler için kullanılan en popüler model lojistik regresyondur. Daha önceki bölümlerde binom rassal bileşenine göre oluşturulan genelleştirilmiş lineer bir model olarak ortaya konmuştur. Bu bölümde ise lojistik regresyon uygulaması daha detaylı incelenecektir.

#### 4.1. LOJİSTİK REGRESYON MODELİNİN YORUMLANMASI

Y gibi ikili bir yanıt değişken ve X gibi kantitatif bir açıklayıcı değişken için  $\pi(x)$ ; X değişkeni x değerini aldığı anda başarı olasılığını göstermektedir. Lojistik regresyon modeli bu olasılığın lojiti için lineer bir forma sahiptir (Agresti, 1996; Hosmer ve Lemeshow; 2000:6; Le, 1998):

$$\text{lojit}[\pi(x)] = \log\left(\frac{\pi(x)}{1 - \pi(x)}\right) = \alpha + \beta x$$

Formülde; x, S-şeklinde bir fonksiyon olduğundan dolayı  $\pi(x)$  artar ya da azalır. Lojistik regresyon için alternatif bir formül ise doğrudan başarı olasılığına dayanmaktadır. Bu formülde  $\exp(x) = e^x$  şeklinde üstel fonksiyon kullanılmaktadır ve aşağıdaki şekildedir (Pampel, 2000):

$$\pi(x) = \frac{\exp(\alpha + \beta x)}{1 + \exp(\alpha + \beta x)}$$

Bu kısımda bu modellerin yorumlanma yöntemleri gösterilmektedir<sup>42</sup>.

---

<sup>42</sup> Lojistik regresyon analizi çeşitli varsayım bozulmaları ( normal dağılıma, ortak kovaryansa sahip olmama gibi) durumunda diskriminant ve olumsuzluk tablolarına bir alternatif olurken, bağımlı

#### 4.1.1. Lineer Yaklaşım Yorumları

$\beta$  parametresi ile S-şeklindeki eğrideki artış ya da azalışlar belirlenir.  $\beta$  parametresinin işareti eğrinin artışı ya da azalışı konusunda bilgi vermektedir.  $|\beta|$  arttıkça değişim oranı da artmaktadır.  $\beta=0$  için model sağlandığında ikinci formüldeki sağ taraf bir sabit ile sadeleşir ve  $\pi(x)$  değeri tüm  $x$  değerleri için özdeş olur. Bu nedenle de eğri yatay bir doğru şeklini alır.  $Y$  ikili yanıt değişkeni  $X$  değişkeninden bağımsızdır.  $\pi(x)$  için lojistik regresyon modeli lineer olmaktan öte bir eğri şeklindedir. Bu nedenle de ikinci formülde yer alan fonksiyon ile  $x'$  teki bir birimlik değişim sonucunda  $\pi(x)$  ' te meydana gelen değişim ifade edilmektedir. Eğriye belirli bir  $x$  değerine göre düz bir çizgi çizilerek tanjant elde edilir. Bu o noktada ki değişim oranını tanımlamaktadır.  $\beta$  şeklindeki lojistik regresyon parametresine göre doğru;  $\beta\pi(x)[1-\pi(x)]$  hesaplamasına eşit bir eğime sahiptir. Olasılıklar 1' e ya da 0' a yaklaştıkça eğimde sifira yaklaşmaktadır.  $\pi(x)=0.5$  ' e göre  $x$  değerinde eğrinin en üstteki eğimi gerçekleşmektedir ki  $x$  değeri  $x = -\frac{\alpha}{\beta}$  ' dır (İkinci formülde  $x$  yerine  $-\frac{\alpha}{\beta}$  koyularak bu noktada  $\pi(x)=0.5$  bulunur veya ilk formülde de  $\pi(x)=0.5$  konularak  $x$  için çözümlene yapılırsa  $-\frac{\alpha}{\beta}$  değeri elde edilir). Bu  $x$  değeri genellikle *medyan etkili seviye* (Median effective level) olarak adlandırılır ve  $EL_{50}$  ile gösterilir (Agresti, 1996). Bu her bir sonuçta seviyenin % 50 şansa sahip olduğunu göstermektedir.

Lojistik regresyon modellerinin uygunluğu için Maksimum Olabilirlik (ML) hesaplamaları karışıktır. Fakat istatistiksel yazılımlar kullanılarak oluşturmak kolaydır (Agresti, 1996; Stokes, Davis ve Koch,1991).  $Y=1$  için daha yüksek  $x$  değerlerinde bulunma eğilimi söz konusudur. Fakat  $Y$  değişkeni sadece 0 ve 1

---

değişkenin 0 ve 1 gibi ikili ya da ikiden çok düzey içeren kesikli değişken olması durumunda normallik varsayımının bozulması nedeniyle doğrusal regresyon analizine alternatif olmaktadır (Bircan ve diğerleri, 2004). Lojistik regresyon modeliyle kestirim yapılacağı zaman *en çok olabilirlik metodu* (ML) temel oluşturacaktır.

değerini aldığından dolayı  $x'$  e karşı  $Y$  değerlerinin çizilmesi aracılığı ile lojistik regresyon modelinin uygun olup olmadığını belirlemek oldukça zordur.

Alternatif olarak bazı yazılımlar verileri düzleştirmektedir (Stokes, Davis ve Koch,1991). İlişki açısından belirli bir fonksiyonel form varsayımı yapmaksızın genel bir eğilim ortaya koymaktadır. **Genelleştirilmiş toplamsal (additive) modellere** dayanan düzleştirme metotları GLM' lerde daha genel yapısal form sağlayarak bunu gerçekleştirmektedir (Agresti, 1996). Örnek olarak, en iyi tahmin edicilerin bir türü olarak hizmet eden açıklayıcı değişkenlerin kompleks fonksiyonlarını ortaya çıkarırlar. Yorumlama için en basit model lineer olasılık modelidir;  $\pi(x) = \alpha + \beta x$ . Veriler için bu GLM' e göre bazı tahminlenen değerler bir binom parametresi için belirlenen resmi aralığın dışına düşer ise ML uyumu başarısızdır (Agresti,1996).

#### 4.1.2. Odds Oranının Yorumlanması

Lojistik regresyon modelinin bir diğer yorumlamasında ise odds ve odds oranı kullanılmaktadır. İlk verilen formüle<sup>43</sup> göre yanıt 1' in oddsu (yani başarının olasılığı) şu şekildedir:

$$\frac{\pi(x)}{1 - \pi(x)} = \exp(\alpha + \beta x) = e^{\alpha} (e^{\beta})^x$$

Bu üssel ilişki  $\beta$  için bir yorum sağlamaktadır:  $x'$  teki her bir birimlik artış için  $e^{\beta}$  nedeniyle, odds çarpımsal bir şekilde artmaktadır.

$x+1$  seviyesindeki odds,  $x$  seviyesindeki olasılıkla  $e^{\beta}$  değerinin çarpımına eşittir.  $\beta=0$  olduğunda,  $e^{\beta}=1$  olur ve odds  $x$  değerindeki değişmeler karşısında değişme göstermez. Odds' un logaritması ( $\pi(x)$ ' in lojit dönüşümüdür) ilk formülle belirtilen doğrusal ilişkiyi vermektedir. Bu modelin lojit ifadesidir ve daha önce açıklanmıştır.  $x$  değerinde meydana gelen her bir birimlik artış için lojit,  $\beta$  kadar

---

<sup>43</sup>  $lojit [\pi(x)] = \log\left(\frac{\pi(x)}{1 - \pi(x)}\right) = \alpha + \beta x$



artış gösterir. Doğal olarak lojit bir ölçekle düşünmek zor olduğundan bu yorumlama sınırlı kullanıma sahiptir<sup>44</sup>.

#### 4.1.3. Olgu-Kontrol İncelemeleri İçin Lojistik Regresyon

Lojistik regresyon modelinin bir diğer özelliği ise Y yanıt değişkeninden öte X açıklayıcı değişkeninin rassal olduğu durumlarla ilgilenilmesidir. Bu en yaygın olarak, olgu-kontrol incelemeleri şeklindeki (Bölüm-2’de ele alınmıştır.) geriye dönük örneklem tasarımları ile gerçekleşmektedir (Agresti,1996). Y=1 (Olgular) değerini ve Y=0 (Kontroller) değerini alan denekler için X değeri gözlemlenmektedir.

Eğer olgular ve kontroller arasında X değerlerinin dağılımı farklılık göstermekte ise bu durum X ve Y arasında bir ilişkinin bulunduğu konusunda kanıt oluşturmaktadır. Birçok biyomedikal incelemede (özellikle epidemioloji incelemelerinde) olgu-kontrol tasarımları kullanılmaktadır. Geriye dönük verilere göre bölüm-2’de gösterildiği üzere ilk olarak odds oranları tahmin edilir. Lojistik regresyon parametreleri için olasılık (odds) ve odds oranına başvurulmaktadır. Bundan dolayı geriye dönük veriler için bu şekildeki modeller kullanılır ve olgu-kontrol incelemelerinde ki etkiler tahmin edilir. Bu iki sonuçlu yanıtlar için

---

<sup>44</sup> İki değişkenli lojistik regresyon modeli,

$$P(Y) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} = \frac{1}{1 + \exp(-\beta_0 - \beta_1 X)}$$

Lojistik regresyon modeli yazıldıktan sonra modeldeki katsayılar,

$$\ln \frac{P(Y)}{Q(Y)} = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p$$

$$\frac{P(Y)}{Q(Y)} = e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p} = e^{\beta_0} \dots e^{\beta_p X_p}$$

şeklinde hesaplanır.

Burada Q(Y), Q(Y)=1 – P(Y) şeklinde hesaplanır. Odds oranının OR= P(Y) / Q(Y) şeklinde hesaplandığı hatırlanacak olursa, her bir parametrenin exp(β) değerleri olasılık oranları olarak ele alınırlar. Böylece exp(β<sub>p</sub>), Y değişkeninin X<sub>p</sub> değişkeninin etkisi ile kaç kat daha fazla ya da yüzde kaç oranda fazla gözlenme olasılığına sahip olduğunu belirtir (Özdamar,1999:487). Modelde P(Y) / Q(Y), örnek olarak alkol kullanım olasılığının, kullanmama olasılığına oranını gösteren olasılık oranını ifade etmektedir. Bu ifadenin doğal logaritması ise logit olarak bilinir. Lojistik regresyon modelindeki parametreler diğer bağımsız değişkenler sabit kalmak kaydıyla bağımsız değişkenden bir birim değişimin odds oranı üzerindeki etkisini belirtir (Işığışık, 2003).

kullanılan diğer modeller için doğru olmayabilir. Çünkü odds oranı; etkilerin tanımlanmasında o modellerin doğal ölçümünü oluşturmamaktadır. Probit gibi bağıntılar arasında odds oranı, lojit bağıntının önemli bir avantaj sağlamasına neden olmaktadır ve lojit modelin popülerlik açısından diğer modellerden üstün olmasının en önemli nedenini oluşturmaktadır(Agresti, 1996).

## 4.2. LOJİSTİK REGRESYON İÇİN YORUMLAMA

Bir lojistik regresyon modelinin nasıl bir uyum sağladığı konusundaki incelemeler, bir tahmin edicinin iki sonuçlu bir yanıt değişkeni üzerindeki etkilerinin belirlenmesinde yardımcı olmaktadır. Mevcut yazılımlar, model uyumu konusunda sağladığı diğer bilgiler kadar parametre tahminlerini ve bunlara ilişkin standart hataları da geniş ölçüde sağlamaktadır (Stokes, Davis ve Koch,1991).

### 4.2.1. Etkiler için Güven Aralıkları

$lojit[\pi(x)] = \alpha + \beta x$  şeklindeki lojistik regresyon modelinde yer alan  $\beta$  parametresi için büyük örnek güven aralığı şu şekildedir:  $\hat{\beta} \pm z_{\alpha/2}(ASE)$ . Bu aralığın uç noktalarının üstü alınarak,  $e^{\beta}$  şeklinde bir değer elde edilir ki bu X değişkenindeki 1 birimlik artışın olasılık (odds) üzerindeki çarpımsal etkisidir (Agresti, 1996; Powers ve Xie, 2000: 41).

Örnekle açıklamak maksadıyla,  $\beta$  için %95 güven aralığının şu şekilde hesaplanmış olsun:  $0.497 \pm 1.96(0.102) \rightarrow (0.298, 0.697)$ : X' teki her bir artışın olasılık (odds) üzerindeki etkisi için güven aralığı:  $(e^{0.298}, e^{0.697}) = (1.35, 2.01)$ . Şu şekilde yorumlanır: X'teki her bir birimlik artış Y değişkeninin gerçekleşme olasılığında(odds'unda) en fazla iki katlık ve en az %35'lik bir etkiye sahiptir. Daha önce lojistik regresyon eğrisine lineer yaklaşımlardan meydana gelen daha basit yorumlamalar ele alınmıştı.  $\beta\pi(1-\pi)$  terimi x' teki bir birimlik değişimin olasılıkta meydana getirdiği değişimi ifade etmektedir. Örnek olarak,  $\pi = 0.5$  için değişimin tahmin edilen oranı;  $0.25\hat{\beta} = 0.124$  olarak bulunur.  $0.25\beta$  için %95 güven aralığı,

$\beta$  için oluşturulan aralığın uç noktalarının 0.25 katına eşittir:  $[0.25(0.298), 0.25(0.697)]=(0.074, 0.174)$ . Eğer lojistik regresyon modeli  $\pi = 0.5$  ' e karşılık gelen x değerleri için kurulursa, X' te meydana gelen her bir birimlik artışta Y' nin gerçekleşme olasılığındaki artış oranı yaklaşık olarak 0.07 ile 0.17 arasında düşmektedir.

#### 4.2.2. Anlamlılığın Test Edilmesi

İki sonuçlu yanıt değişkeni üzerinde X değişkeninin etkisi için anlamlılık testleri açıklanacaktır. Lojistik regresyon modeline göre sıfır hipotezi ( $H_0 : \beta = 0$ ) başarımın olasılığı X değişkeninden bağımsızdır. Büyük örnekler için test istatistiği:

$$z = \frac{\hat{\beta}}{ASE}$$

Bu test istatistiği  $\beta = 0$  olduğunda standart normal bir dağılıma sahiptir. İlk olarak genel anlamda, z değerlerinin tek yanlı ya da iki yanlı bir P değerini elde etmek amacıyla standart normal tablosuna bakılır.

Aynı şekilde  $\beta \neq 0$  şeklindeki iki yanlı alternatif hipotez için  $(\hat{\beta} / ASE)^2$  eşitliği, serbestlik derecesi 1 olan bir büyük örnek kıkare dağılımı gösteren Wald istatistiğidir. Wald testi çok büyük örnekler için çok iyi işlese de olabilirlik oran testi uygulamada örnek büyüklükleri için daha güvenilir ve daha güçlü bir testtir (Agresti, 1996; Bircan, 2004: 185-208). Test istatistiği,  $\beta = 0$  için (yani  $\pi(x)$ , tüm değerler için özdeş olmaya zorlanmaktadır) log-olabilirlik fonksiyonunun maksimumu,  $L_0$  ile sınırlanmamış  $\beta$  için log-olabilirlik fonksiyonunun maksimumunu ( $L_1$ ) karşılaştırmaktadır.  $-2(L_0 - L_1)$  şeklindeki test istatistiği aynı zamanda serbestlik derecesi 1 olan bir büyük örneklem ki-kare dağılımına sahiptir. Lojistik regresyon için çoğu yazılım maksimize edilmiş log-olabilirlikleri,  $L_0$  ve  $L_1$  ve bu maksimizasyondan türetilen olabilirlik oran istatistiğini hesaplamaktadır(Stokes, Davis ve Koch,1991).

### 4.2.3. Olasılık Tahminlerinin Dağılımları

X' in belirli bir x düzeni için Y=1 şeklinde tahmin edilen olasılığı aşağıdaki eşitlik aracılığı ile elde edilir:

$$\hat{\pi}(x) = \frac{\exp(\hat{\alpha} + \hat{\beta}x)}{1 + \exp(\hat{\alpha} + \hat{\beta}x)}$$

Lojistik regresyon açısından yazılımların çoğu gerçek olasılıklar için güven aralıklarını hesapladığı kadar bu şekildeki tahminleri de hesaplamaktadır(Stokes, Davis ve Koch,1991). Lojistik regresyon modeli doğru bir şekilde kurulduğunda, bir olasılığın model-temelli tahmin edicisi örneklem oranından mantıksal açıdan daha iyidir. Uygulamada her hangi bir model,  $\pi(x)$  ve x arasında var olan gerçek ilişkiyi kesin olarak temsil edemez. Bu nedenle örneklem hacmi arttıkça model temelli tahmin edici, olasılığın gerçek değerine yakınsamayabilir.

Fakat bu uygulamada örneklem oranının gerçekten daha iyi bir tahmin edici olduğu anlamına da gelmemektedir. Eğer model gerçek olasılıklara yeterince yaklaşırsa o modelin tahmin edicisi, gerçek değere örneklem oranından daha yakın olma eğiliminde olur(Agresti, 1996). Özet olarak, eğer lojistik regresyon modeli x üzerinde  $\pi(x)$  'in gerçek bağımlılığına yaklaşırsa, x' e bağlı olarak  $\pi(x)$  ' in nokta ve aralık tahmin edicileri oldukça fayda sağlamaktadır.

### 4.3. MODEL KONTROLÜ

Şimdiye kadar lojistik regresyon modeli, iki sonuçlu yanıtlar üzerinde tahmin edicilerin etkileri konusunda yorumlama ve tanımlama için kullanılmıştır. Fakat bu şekilde belirli bir modelin uygun olduğu ya da veriler için iyi bir uyuma sahip olduğu konusunda bir garanti yoktur. Bu bölümde model uyumunun kontrol edilmesi için çeşitli yöntemler ele alınmaktadır. Sıfır hipotezinin testi; model kurulduktan sonra Pearson  $X^2$  veya olabilirlik oran  $G^2$  test istatistikleri kullanılarak gözlemlenen ve uydurulan değerler karşılaştırılarak yapılır. Belirli bir sayıdaki kümeye göre

uydurulan değerlerin çoğu en azından yaklaşık olarak 5 ise  $X^2$  ve  $G^2$  istatistikleri yaklaşık olarak bir ki-kare dağılımı gösterir. Serbestlik derecesi, model için artık serbestlik derecesi (residual df) olarak adlandırılır ve örnek lojitlemin sayısı ile (yani açıklayıcı değişken kümelerinin sayısı) modeldeki parametre sayısı arasındaki farka eşittir (Agresti, 1996). Doğal olarak,  $X^2$  ve  $G^2$  istatistiklerinin büyük değerlere sahip olması uygunsuzluğun bir kanıtını oluşturmaktadır ve P-değeri gözlemlenen değer üzerindeki sağ kuyruk olasılığıdır. Uyum zayıf olduğunda, artıklar ve diğer teşhis ölçümleri, bireysel gözlemlerin model uyumu üzerindeki etkilerini tanımlamaktadır ve uygunsuzluğun nedenlerini belirtmektedir. Açıklayıcı değişkenler sürekli nitelikte olduğunda, her hangi bir gruplandırma türü kullanılmaksızın uyumsuzluğun analiz edilmesi oldukça zordur. Gruplandırmanın alternatif bir yolu ise gözlemlenen ve uydurulmuş değerlerin tahminlenen olasılıklara bağlı olarak şekillendirilmesidir. Modelde kaç tane parametre olursa olsun gözlemlenen ve uydurulmuş değerler tahminlenen olasılıklara göre parçalanabilir.

En yaygın yaklaşım parçalınımına göre gruplamadır. Bu nedenle de bu gruplar yaklaşık olarak eşit büyüklüktedir. Her bir grup için bir sonucun uydurulmuş değeri; ilgili gruptaki tüm gözlemlerin tahmin edilen olasılıklarının toplamıdır. Bu yapı Hosmer ve Lemeshow<sup>45</sup> (1989) için bir testin temelini oluşturmaktadır. Onların Pearson benzeri istatistikleri gerçekte bir ki-kare dağılımına sahip değildir. Fakat bu istatistiğin dağılımı, s.d=g-2 olan bir ki-kare dağılımına yaklaştırılarak benzetimleri gösterilmiştir(g burada grup sayısını ifade etmektedir)<sup>46</sup>. Daha karışık modeller ile kullanılan modelin karşılaştırılmasında kullanılan olabilirlik oran testi aracılığı ile aynı zamanda uygunsuzlukta ortaya çıkarılabilmektedir. Örnek olarak, kantitatif tahmin ediciler ya da etkileşim ifadeleri için lineer olmayan etkileri (kuadratik terimler şeklindeki) içeren daha karışık modeller dikkate alınabilir.

---

<sup>45</sup> Katsayıların önem kontrolü yapıldıktan sonra katsayıların yorumlanması odds oranları ya da Hosmer ve Lemeshow tarafından geliştirilen daha kapsamlı uyum iyiliği testleri kullanılarak yapılmaktadır (Arabacı, 2002). Lojistik regresyon analizinin kullanım amaçlarından en önemlisi sıklıkla karşılaşılan bağımlı değişkenin iki ya da daha çok düzey içerdiği, bağımsız değişkenlerin ise hem kesikli hem de sürekli olabildiği durumlarda, verilerin ait oldukları gruplara en doğru şekilde atayacak ve konuya ilişkin risk faktörlerini belirleyebilecek modeli kurmaktır.

<sup>46</sup> Kurulan modelin uyum iyiliği testi Hosmer-Lemeshow'un hem onlu risk grupları hem de sabit kesim noktası yöntemine göre hesaplanmaktadır.

Eğer daha iyi bir uyum sağlayan daha karışık bir model bulunamazsa uydurulan modelin mantıklı ve uygun olduğu konusunda güven elde ederiz. Bu yaklaşım bilimsel bir perspektiften daha fazla yarar sağlamaktadır (Agresti, 1996). Büyük bir uyum iyiliği istatistiği bazı uyumsuzluklar olduğuna işaret etmektedir. Fakat yapısı konusunda her hangi bir bilgi vermemektedir.

#### 4.3.1. Uyum İyiliği ve Olabilirlik Oran Model Karşılaştırma Testleri

Daha önce bir modelde belirlenen parametrelerin sıfıra eşit olup olmadığını test etmek amacıyla  $-2(L_1 - L_0)$  şeklindeki olabilirlik oran istatistiği gösterilmiştir. Kurulan model için maksimum log- olabilirlik( $L_1$ ) ile daha basit bir model için (bu modelde parametreler silinir) log-olabilirliğini ( $L_0$ ) karşılaştırmaktadır.

$M_1$ ; uydurulmuş modeli ve  $M_0$ ; uydurulmuş modeldeki parametrelerin sıfıra eşit olduğu daha basit bir modeli ifade etsin.  $M$  şeklindeki bir lojistik regresyon modelinin uyumunun test edilmesi için  $G^2$  şeklindeki uyum iyiliği istatistiği olabilirlik oran istatistiğinin özel bir durumudur ( $M_0 = M$  ve  $M_1$ ; en karışık model). Bu karışık model her bir lojit için ayrı bir parametre değerine sahiptir ve örneklem lojitleri için mükemmel bir uyum sağlamaktadır. Bu modele *doymuş model* adı verilmektedir (Agresti, 1996; Bircan, 2004).  $M'$  in uygun olup olmadığını test ederken, doymuş modelde yer alan ama  $M'$  de yer alamayan tüm parametrelerin sıfıra eşit olup olmadığı test edilir.  $M'$  in uyumunu test eden bu istatistik  $G^2(M)$  ile gösterilir.  $L_S$  doymuş modelin maksimize edilmiş log-olabilirliğini gösterebilir. Örnek olarak  $M_0$  ve  $M_1$  modelleri için artıklar sırasıyla  $G^2(M_0) = -2(L_0 - L_S)$  ve  $G^2(M_1) = -2(L_1 - L_S)$  şeklindedir<sup>47</sup>.  $M_1$  gibi belirli bir model kullanılarak  $M_0$ 'ın test edilmesi

---

<sup>47</sup> Lojistik regresyonda gözlenen ve beklenen değerlerin karşılaştırılması log olabilirlik fonksiyonu ile yapılmaktadır.  $D = -2 \ln(\text{Şu andaki modelin olabilirliği} / \text{Doymuş modelin Olabilirliği})$  Şeklinde dir. Parantez içinde verilen ifade olabilirlik oranı olarak adlandırılır. Bağımsız değişkenin önemine karar vermek için denklemden bağımsız değişkenin olduğu ve olmadığı durumlardaki  $D$  değerleri karşılaştırılır. Bağımsız değişkeni kapsamasından dolayı ortaya çıkan  $D$ 'deki değişim şu şekildedir:  $G = D(\text{Değişkensiz Model için}) - D(\text{Değişkenli model için})$ . Hesaplanan bu istatistik de doğrusal regresyonda kullanılan  $F$  testindeki pay kısmı ile aynı rolü üstlenir (Bircan, 2004:185-208).  $G^2$  yi hesaplamak için farkı alınacak  $D$  değerlerinin her ikisi için de doymuş modelin olabilirlikleri ortak olduğundan  $G$  istatistiği şu şekli alır:  $G = -2 \ln(\text{Değişkensiz modelin olabilirliği} / \text{Değişkenli modelin olabilirliği})$ .

için hesaplanan olabilirlik oran istatistiği  $G^2 (M_0 | M_1)$  ile gösterilsin. Bu modellerin karşılaştırılması için bu istatistik aşağıdaki şekilde hesaplanmaktadır:

$$G^2(M_0 | M_1) = -2(L_0 - L_1) = -2(L_0 - L_S) - [-2(L_1 - L_S)] = G^2(M_0) - G^2(M_1)$$

İki modele göre Uyum İyiliği İstatistiklerindeki ( $G^2$ ) farktır. İki modelin karşılaştırılmasında olabilirlik oran istatistiği, bu modellerin sapmalarındaki farktır (Hosmer ve Lemeshow, 2000:13).  $M_0$  modeli  $M_1$  modeli ile kıyaslandığında daha az bir uyuma sahipse bu istatistik büyük bir değer almaktadır.

Bu istatistik büyük örnek ki-kare istatistiğidir ki serbestlik derecesi iki model için kullanılan artık serbestlik derecesi değerlerinin farkına eşittir. İki yönlü bir çapraz tabloda bağımsızlığın test edilmesi için kullanılan bir istatistiktir (2.4.3). Gerçekte bu istatistik gruplanmış veriler için uydurulmuş lojistik regresyon modelinde  $\beta = 0$  hipotezinin test edilmesi amacıyla hesaplanan olabilirlik oran istatistiğine  $(-2(L_0 - L_1))$  eşittir.

#### 4.3.2. Lojit Modeller İçin Artıklar

$G^2$  ve  $X^2$  şeklindeki uyum iyiliği istatistikleri; uyumun genel kalitesi konusunda özet göstergelerdir. Herhangi bir uyumsuzluğun yapısının tanımlanması amacıyla ilave teşhis analizleri gerekmektedir. Gözlemlenen ve uydurulan sayıların karşılaştırılmasını sağlayan artıklar bu amaç için oldukça kullanışlıdır (Anderson, 1990:278).  $y_i$ ;  $i$ ' ninci açıklayıcı değişken dizisinde (kümesinde)  $n_i$  denemedeki "başarı" sayısını gösterebilir.  $\hat{\pi}_i$  ise model uyumu için tahmin edilen başarı olasılığını gösterebilir.  $n_i \hat{\pi}_i$  ise uydurulmuş (beklenen) başarı sayısını ifade eder. Binom rassal bileşeni ile birlikte oluşturulan bir GLM' e göre,  $i$ . dizideki (kümedeki) uyum için Pearson artığı aşağıdaki şekildedir:

$$e_i = \frac{y_i - n_i \hat{\pi}_i}{\sqrt{[n_i \hat{\pi}_i (1 - \hat{\pi}_i)]}}$$

Her bir fark; gözlemlenen değer ve uydurulmuş değer arasındaki farkın, gözlemlenen değer için tahminlenen binom standart sapmasına bölünerek elde edilir. Model uyumunun test edilmesi için kullanılan Pearson istatistiği;  $X^2 = \sum e_i^2$  şeklindeki eşitlikle sağlanır: Her bir Pearson artığının karesi  $X^2$  nin bir bileşenidir. Binom indeksi ( $n_i$ ), büyük olduğunda Pearson artığı ( $e_i$ ) yaklaşık bir binom dağılımına sahiptir.

Model ele alındığında bu artığın yaklaşık beklenen değeri sıfırdır. Fakat standart normal bir değişkenden daha küçük bir varyansa sahiptir. Eğer model parametrelerinin sayısı, örnek lojitlerle karşılaştırıldığında küçük ise Pearson artıkları standart normal sapmalar şeklinde işleme tabi tutulur ki mutlak değerler 2' den büyüktür ve bu uyumsuzluğa işaret etmektedir (Agresti, 1996; Bircan, 2004:185-208). Uyumsuzluğun gösterilmesinde “Grafiksel Gösterimler” de oldukça fayda sağlamaktadır. Gözlemlenen ve uydurulmuş oranların, her biri karşılıklı olarak çizilerek (ya da açıklayıcı değişkenlerle karşılıklı olarak çizilerek) karşılaştırılması bunlardan biridir (Agresti, 2002). Uydurulmuş değerler (beklenen değerler) çok küçük olduğunda  $X^2$  ve  $G^2$  istatistiklerinin geçersiz olduğu daha önce belirtilmişti. Benzer şekilde artıklarda bu durumda sınırlı bilgi vermektedir.

#### 4.3.3. Etkinin Teşhis Edilmesindeki Ölçüler

Açıklayıcı değişkenlerin çoğunda ya da birinde aşırı uç bir değer söz konusu olduğunda ilgili gözlem çok büyük bir etkiye sahip olabilmektedir. Bir ya da iki gözlem silindikten sonra (bu gözlemlerle kurulan modelin uyumu yanıltıcı gözüküyorsa) uyumun raporlanması daha bilgi verici nitelikte olabilir. Birçok ölçüm etkilerin çeşitli özelliklerinin tanımlanmasını sağlamaktadır. Bu ölçümlerin çoğu veri setinden bazı gözlem/gözlemlerin atılmasının belirli karakteristikleri üzerindeki etkisi ile ilişkilidir. Bu ölçümler cebirsel anlamda bir gözlemin kaldıraç gücü ile ilişkilidir ki şapka matrisi olarak adlandırılan bir köşegen matrisi oluşturan



elemandır<sup>48</sup>(Şapka matrisi bir matris olup örnek lojitelere uygulandığında model için tahminlenen lojit değerleri vermektedir)(Agresti, 2002). Daha büyük bir kaldıraç gücüne sahip bir gözlem daha büyük bir etkiye sahiptir. Kaldıraç güçleri için formüller ve etkinin teşhis edilmesindeki ölçümler oldukça karışıktır. Bu nedenle de burada ele alınmamaktadır. Lojistik regresyon için kullanılan çoğu yazılım bu teşhisleri sağlamaktadır(Stokes, Davis ve Koch,1991).

Her bir gözlem için etki ölçümleri aşağıdaki maddeleri kapsamaktadır (Agresti, 2002):

1. Modeldeki her bir parametre için, bir gözlem silindiğinde parametre tahmini de değişir. Bu değişim (standart hataya bölünür) Dfbeta olarak adlandırılır.

2. Bir gözlemin silinmesi ile parametreler için oluşturulan ortak bir güven aralığındaki değişimin bir ölçümü söz konusudur. Bu güven aralığı yer değiştirme teşhisi “c” ile gösterilir.

3. Gözlem silindiğinde  $X^2$  veya  $G^2$  şeklindeki uyum istatistiklerinde de değişim söz konusudur.

Her bir ölçüm için daha büyük bir gözlemin etkisi sonucunda daha büyük bir değer elde edilir. Aynı zamanda, model oluşturulduğunda yaklaşık bir standart dağılıma sahip ve mutlak değeri kısmen daha büyük olan Pearson artığı ( $e_i$ ) için bir düzeltme oluşturmak amacıyla kaldıraç (leverage) değerler kullanılır. Bunlar iki yönlü tablolaradaki bağımsızlığın belirlenmesi amacıyla ikinci bölümde ele alınan ayarlanmış artıklar ile aynı amaca hizmet etmektedir.

#### **4.4. KALİTATİF TAHMİN EDİCİLER İÇİN LOJİT MODELLER**

Sıradan regresyon gibi Lojistik regresyonda, çoklu açıklayıcı değişkenler içeren modeller için genişletilir. Ayrıca bu açıklayıcı değişkenlerden bazıları ya da tümü kantitatif nitelikten daha çok kalitatif nitelikte olabilir. Bu bölümde *kalitatif*

---

<sup>48</sup> Köşegen matris: Ana köşegen öğelerinden en az biri sıfırdan farklı olup öteki öğeleri sıfır olan matris.

*tahmin edicilerin* (genellikle *faktör* olarak adlandırılmaktadırlar) dahil edilmesi için kukla değişkenlerin (dummy variable) kullanımı gösterilmektedir. Daha sonra çoklu lojistik regresyon modellerinin genel şekli ele alınacaktır.

#### 4.4.1. Lojit Modellerde Kukla Değişkenlerin Kullanımı

Varsayalım ki iki sonuçlu bir yanıt değişkeni (Y), X ve Z gibi iki sonuçlu bir tahmin ediciye sahip olsun. Her bir açıklayıcı değişken için iki seviye (0,1) şeklinde gösterilsin.  $2 \times 2 \times 2$  boyutlu bir kontenjans tablosuna göre,  $\pi$  (Y=1) olasılığı için model aşağıdaki şekildedir ve iki tahmin ediciye göre ayrı ana etkilere sahiptir:  $lojit(\pi) = \alpha + \beta_1 x + \beta_2 z$ . Bir faktör, başka bir faktörün her bir seviyesinde aynı etkiye sahip ise etkileşimin olmadığı varsayılır. Bu modeldeki x ve z değişkenleri kukla değişkenlerdir (dummy variables). Bu değişkenler tahmin edicilerin kategorilerini ifade etmektedir (Agresti, 1996; Powers ve Xie, 2000). Z' nin belirli bir z seviyesinde x=0' dan x=1' e doğru meydana gelen değişimin lojiti üzerindeki etki aşağıdaki şekildedir:

$$= [\alpha + \beta_1(1) + \beta_2 z] - [\alpha + \beta_1(0) + \beta_2 z] = \beta_1$$

İki lojit arasındaki bu fark logaritmik oddslarının farkıdır ve Z değişkeninin belirli bir seviyesinde X ve Y değişkenleri arasındaki odds oranının logaritmasına eşittir. Bundan dolayı  $\exp(\beta_1)$  ifadesi X ve Y değişkenleri arasındaki koşullu odds oranının tanımıdır. Z değişkeni kontrol altında tutularak, x=1 için başarı olasılığı (Odds'u); x=0 için bulunan başarı olasılığının (odds'unun),  $\exp(\beta_1)$  katına eşittir. Bu koşullu odds oranı Z değişkeninin her bir z seviyesinde aynıdır. Modelde bir etkileşim teriminin olmaması Z değişkeninin her iki seviyesine göre oluşturulan kısmi tablolar için ortak bir odds değerinin olduğunu göstermektedir. Model homojen bir birliktelik sağlamaktadır. Eğer  $\beta_1=0$  ise bu durumda ortak odds oranı 1' e eşittir ve Z değişkeni kontrol altına alınarak X ve Y değişkenleri arasında koşullu bir bağımsızlık vardır. Böylece üç yönlü tablo için aşağıdaki biçimde ifade edilen daha basit bir model uygulanır:  $lojit(\pi) = \alpha + \beta_2 z$ . İlk olarak, iki modelin

karşılaştırılmasında olabilirlik oran istatistiği ya da Wald istatistiği kullanılarak  $\beta_1=0$  olup olmadığı test edilmelidir.

#### 4.4.2. AZT Ve AIDS Örneği

Aşağıda yer alan Tablo-4.1' i kullanarak kalitatif (nitel) veriler kullanılarak kurulacak modeller açıklanacaktır. Bu tablo AIDS belirtilerinin gelişmesini yavaşlatmak amacıyla kullanılan AZT ilacının etkileri üzerine yapılan bir çalışmaya dayanmaktadır (New York Times, 15 Şubat 1991). Bu çalışmada, AIDS virüsü enfeksiyonu belirlendikten sonra bağışıklık sisteminde bozukluk başlayan 338 eski hasta rassal olarak belirlenmiştir. Tablo 4.1 hemen AZT kullanıp kullanmadıklarına, AIDS belirtilerinin üç yıl boyunca ilerleyip ilerlemediğine ve hasta ırklarına göre oluşturulan 2 x 2 x 2 boyutlu bir çapraz sınıflandırmayı göstermektedir.  $lojit(\pi) = \alpha + \beta_1x + \beta_2z$  şeklindeki bir modelde, X değişkeni AZT tedavisini ( $x=1$ ; hemen AZT alanları,  $x=0$  diğerlerini göstermektedir), Z değişkeni ırkları ( $z=1$  beyazlar için,  $z=0$  siyahlar için) göstermektedir. AIDS belirtilerinin ilerleme olasılığının tahmin edilmesi amaçlanmıştır. AZT etkisinin ML tahmini şu şekildedir:  $\beta_1= -0.720$  (ASE=0.279). AIDS belirtilerinin gelişimi ve hemen AZT kullanımı arasında tahmin edilen odds oranı;  $\exp(-0.720)=0.49$  olarak bulunur. Her bir ırk türüne göre, belirtilerin ilerlemesinin tahminlenen olasılığı (odds'u) derhal AZT kullanımına başlayanlar için bulunandan yarım kat daha yüksektir.

**Tablo 4.1** AZT kullanımına ve Irka Göre AIDS Belirtilerinin Gelişimi

İrk	AZT Kullanımı	Belirtiler	
		Var	Yok
Beyaz	Evet	14	93
	Hayır	32	81
Siyah	Evet	11	52
	Hayır	12	43

Kaynak: <http://www.ats.ucla.edu/stat/stata/examples/icda/azt>

İrk deęişkeni kontrol altına alınarak, AIDS belirtilerin gelişimi ve AZT tedavisinin koşullu bağımsızlığının ifade edildięi hipotez  $H_0 : \beta_1 = 0$  şeklindedir. Modellerin<sup>49</sup> karşılaştırılmasına dayanan olabilirlik oran istatistięi  $-2(L_0-L_1)$ , 1 serbestlik derecesi ile 6.9 olarak bulunur ve birliktelięin bir kanıtını gösterir (P=0.01). Wald istatistięi de benzer sonuçlar sağlamaktadır(P=0.01):  $(\hat{\beta}_1 / ASE)^2 = (-0.720 / 0.279)^2 = 6.6$ . Sonra birinci modelin uyum iyilięi analiz edilir. Bu modelin uyumu için derhal AZT kullanımına bařlayan beyaz ırktaki hastaların, inceleme boyunca AIDS belirtilerin gelişiminin tahminlenen olasılıęı 0.150 olarak bulunur. 107 tane beyaz ırka sahip hasta AZT kullandıęından dolayı, semptomların ilerlemesinin uydurulmuş sayısı;  $107(0.150)=16.0$  ve belirtilerin ilerleme göstermedięi durumlar için uydurulmuş sayı;  $107(0.850)=91.0$  olarak bulunur. Benzer şekilde Tablo 5.1' teki sekiz hücrenin tümü için uydurulmuş deęerler elde edilir. Bu deęerler ve hücre sayıları genel uyum iyilięi istatistiklerinde yerine koyulur ve  $G^2=1.4$  ile  $X^2=1.4$  deęerleri elde edilir.

Model dört tane örnek lojit deęerine sahiptir ve ilk olarak AZT kullanımı ve ırk deęişkenlerinin dört farklı kombinasyonunda her bir binom yanıt daęılımı için hesaplanır. Model üç parametreye sahip olduęundan dolayı artık serbestlik derecesi  $4-3=1$  olarak bulunur.  $G^2$  ve  $X^2$  istatistikleri için küçük deęerler model uyumunun yeterince iyi olduęunu göstermektedir (P>0.2). Daha ileri düzeyde bir analiz yapılırsa belki de daha basit bir modelin uygun olabileceęi ileri sürülebilir. Çünkü ırk deęişkeninin etkisi anlamlı deęildir.

#### 4.4.3. Faktörlerin ANOVA Tipi Gösterimi

İki seviyeli bir faktör sadece bir kukla deęişken kullanımını gerektirir. I seviyeli bir faktör ise I-1 tane kukla deęişken kullanımını gerektirmektedir. Lojistik regresyon modellerindeki faktörlerin alternatif bir gösterimi ise bu faktörlerin ifade edildięi sıradan ANOVA yöntemine benzemektedir (Agresti, 1996). Modelin

<sup>49</sup>

$$lojit(\pi) = \alpha + \beta_1 x + \beta_2 z, \quad lojit(\pi) = \alpha + \beta_2 z$$

formülü şu şekildedir :  $lojit(\pi) = \alpha + \beta_i^X + \beta_k^Z \{\beta_i^X\}$  parametreleri aracılığı ile X değişkenin etkileri ve  $\{\beta_k^Z\}$  parametreleri aracılığı ile Z değişkeninin etkileri gösterilmektedir (X ve Z üst simgeleri etiketlerdir ve kuvvetleri temsil etmezler). Her bir faktör sahip olduğu seviye kadar parametre sayısına sahiptir, fakat biri fazladır. Örnek olarak, eğer X değişkeni I seviyeli ise I-1 tane gereksiz olmayan parametreye sahiptir. Sabit bir Z değişkenine göre X ve Y değişkenleri arasındaki koşullu bağımsızlık şu şekilde ifade edilmektedir:  $\beta_1^X = \beta_2^X = \dots = \beta_I^X$ .

Son kategori sifıra eşitlenerek oluşturulan parametre dizisi aracılığı ile yukarıda sözü geçen modelde yer alan parametrelerdeki fazlalıklar hesaplanır. X ve Z değişkenleri iki kategorili olduğu zaman (Tablo 4.1) modeldeki parametreler aşağıdaki modele uymaktadır:

$$lojit(\pi) = \alpha + \beta_1 x + \beta_2 z$$

$\beta_1^X = \beta_1$  ,  $\beta_2^X = 0$  ve  $\beta_1^Z = \beta_2$  ,  $\beta_2^Z = 0$  olduğundan dolayı bu gerçekleşmektedir. Tablo 4.1' e göre bu modelin uydurulması için, Tablo 4.2' de, parametrelerin tanımlanmasında kullanılan üç ayrı yöntemle göre parametre tahminleri gösterilmektedir: (1) Sadece son parametrenin (yani  $\beta_2^X$  ) sifıra eşit olması ile oluşturulan bir yaklaşım. (2) İlk parametrenin sifıra eşitlendiği benzer bir yaklaşım. (3) Faktör parametrelerinden her birinin toplamının sifır olmasına dayanan bir yaklaşım.

İkinci yaklaşıma göre yukarıda yer alan model;  $lojit(\pi) = \alpha + \beta_i^X + \beta_k^Z$  şeklindeki modele benzemektedir ve AZT kullanımının birinci kategorisi için x=0 ve ikinci kategorisi için x=1 şeklindeki kodlanan kukla değişken kullanılmaktadır.

Üçüncü yaklaşıma göre, bir faktör iki seviyeli olduğunda bir tahmin diğerinin negatif işaretlisidir (yani  $\hat{\beta}_1^X = -\hat{\beta}_2^X$  ). Bu durum bir kukla değişkeni için birinci

kategoride  $x=1$  ve ikinci kategoride  $x= -1$  şeklinde “etki kodlaması” yapılmasından kaynaklanmaktadır (Agresti, 1996).

**Tablo 4.2** Tablo 4.1’ e göre uydurulan Lojit model için  
Parametre Tahminleri

Parametre	Parametrelerin tanımlanması		
	Son=0	İlk=0	Toplam=0
Sabit Terim	-1.074	-1.7383	-1.406
AZT Kullanımı-Evet	-0.720	0.000	-0.360
AZT Kullanımı-Hayır	0.000	0.720	0.360
İrk- Siyah	0.055	0.000	0.028
İrk-Beyaz	0.000	-0.055	-0.028

Üç kodlama düzeninden her hangi biri için  $\beta_1^X - \beta_2^X$  ve  $\beta_1^Z - \beta_2^Z$  şeklindeki farklar özdeştir ve yanıt ile birlikte X ve Z değişkenlerinin koşullu logaritmik odds oranlarını ifade etmektedir. Örnek olarak,  $\exp(\hat{\beta}_1^X - \hat{\beta}_2^X) = \exp(-0.720) = 0.49$  değeri, her bir ırka göre belirtilerin gelişimi ile derhal AZT kullanımına başlanması arasında tahmin edilen ortak odds oranına dayanmaktadır.

Bir faktörün tek bir kategorisine göre bir parametre tahmini yapmak yersizdir: Parametre fazlalıklarının belirlenmesinde farklı yöntemler söz konusudur ve bunlar ilgili tahmine göre farklı değerler oluşturmaktadır. Bir tahmin sadece başka bir kategori için bulunan tahmin ile karşılaştırılarak yorumlanır. İki kategori için yapılan tahminler arasındaki farkın üssel biçimi ile bir kategorideki sınıflamanın etkisine ilişkin odds oranı belirlenmektedir. Benzer şekilde farklı kodlama düzenleri de benzer tahminlenen olasılıklar sağlamaktadır. Belirli faktör seviyeleri için yapılan tahminler ile kesen (sabit terim) tahminin toplamı her bir düzen için aynıdır.

Örnek olarak, tablo-4.2'ye göre her bir düzene göre (sütuna göre) sabit terim, derhal AZT kullanımına başlanmasına göre yapılan tahmin ve beyaz olanlar için yapılan tahminler toplandığında hepsinde -1.738 değeri elde edilir ki bu değer, derhal AZT kullanımına başlayan hastalarda AIDS belirtilerinde iyileşme görülmesinin hesaplanan ve tahmin edilen olasılığının bulunmasında etkilidir ve bu olasılık şu şekilde bulunur:  $\exp(-1.738)/[1+\exp(-1.738)]=1.5$

#### 4.5. ÇOKLU LOJİSTİK REGRESYON

Lojistik regresyon modeli ve diğer GLM' ler de, birkaç tane açıklayıcı değişkenin var olması durumunda normal veriler için oluşturulan sıradan regresyon modelleri gibi genelleştirilebilir. İki sonuçlu Y yanıt değişkeni için,  $X_1, X_2, \dots, X_k$  değişkenleri, k tane tahmin edicinin bir kümesini gösterebilir.  $\pi$  olasılık lojiline göre en başta verilen model<sup>50</sup>, Y=1 için aşağıdaki şekilde genelleştirilir:

$$\text{lojit}(\pi) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$$

$\beta_i$  parametresi;  $X_i$ ' nin diğer X' ler kontrol altında iken Y=1' deki logaritmik olasılığı (odds'u) üzerindeki etkisini göstermektedir. Örnek olarak  $\exp(\beta_i)$ , diğer X' lerin belirli seviyelerinde  $X_i$ ' deki 1 birimlik artışın oddsu üzerindeki çarpımsal etkisidir. Çoklu lojistik regresyon modellerindeki parametreler konusunda oluşturulan hipotezlerin test edilmesi için olabilirlik oran metodu kullanılır. Örnek olarak, bir faktörün modele önemli bir katkı yapıp yapmadığının test edilmesi için  $H_0 : \beta_1 = \beta_2 = \beta_3 = 0$  hipotezi test edilir. Bu hipotez şu durumu ifade etmektedir: Bir değişken kontrol altına alındığında, Y değişkeninin gerçekleşme olasılığı ilgili faktörden bağımsızdır. Tam modele göre maksimize edilmiş log-olabilirliği ( $L_1$ ) ile daha basit bir modele (bu parametre değerleri sıfır olan model) göre maksimize

---

<sup>50</sup>  $\text{lojit} [\pi(x)] = \log \left( \frac{\pi(x)}{1 - \pi(x)} \right) = \alpha + \beta x$

edilmiş log-olabilirliği ( $L_0$ ),  $-2(L_0 - L_1)$  şeklindeki test istatistiği kullanılır. Ki-kare serbestlik derecesi iki modeldeki parametre sayıları arasındaki farktır.

Daha genel olarak, herhangi iki model için maksimize edilmiş log-olabilirlikleri karşılaştırılır ki buradaki iki modelden biri diğerinin özel bir durumu konumundadır. Bu şekildeki bir karşılaştırma ile modelin etkileşim terimlerini içerip içermediği kontrol edilir.

#### **4.5.1. Birden Fazla Tahmin Ediciye Sahip Model Seçimi**

Birçok model seçme yöntemi bulunmaktadır ve içlerinden hiç birisi en iyi değildir. Çünkü bu yöntemler her hangi bir genelleştirilmiş lineer model için eldeki normal verilerin sıradan regresyon modellemesine uygun niteliktedirler. Örnek olarak birden fazla tahmin ediciye sahip bir model *çoklu bağlantı* açısından potansiyele sahiptir: Tüm tahmin ediciler modelde yer aldığında, tahmin ediciler arasındaki güçlü korelasyon ilgili modeldeki hiçbir değişkenin anlamlı olmamasına neden olmaktadır. Bir değişken az bir etkiye sahip gibi gözüküebilmektedir. Bunun nedeni de modeldeki diğer tahmin edicilerle çakışmasıdır (overlap).

Olabilirlik oran testi, tahmin edicilerin  $Y$ 'nin gerçekleşme olasılığından ortak biçimde bağımsız olmasıdır ki eş zamanlı olarak bu tahmin edicilerin parametre değerlerinin tümünün sıfıra eşitliği (sıfır hipotezi) test edilir. Olabilirlik oran istatistiği; sadece sabit terime (kesene) sahip sıfır model ile ana etkilerin yer aldığı modelin karşılaştırılmasına dayanmaktadır. Tüm test yüksek derecede anlamlı olsa bile diğer sonuçlar bu anlamlılığı desteklemeyebilir. Kalitatif tahmin ediciler için yapılan tahminler temel olarak modele alınan son kategori ile her seviyede karşılaştırılır. Genel test için çok küçük P-değeri çoklu bağlantı açısından önemli bir uyarı oluşturmaktadır. Diğer değişkenler kontrol altına alındığında bir değişkenin kısmi etkisi konusunda çok az bir kanıt söz konusu olabilir.



#### 4.5.2. Tahmin Edicilerin Geriye Doğru Elenmesi

Bir modelin sapması, doygun model ile ilgili modelin karşılaştırılmasına dayanan  $G^2$  uyum iyiliği testidir. İki model arasında, sapmaların farkı bu modellerin karşılaştırılması için kullanılan olabilirlik oran istatistiğidir ( $-2(L_0 - L_1)$ ). Bir modeli seçmek için *geriye doğru eleme yöntemi* (Backward elimination procedure) kullanılmaktadır ve karışık bir modelle başlayarak ardaşık biçimde terimler çıkarılmaktadır (Agresti, 1996; Pampel, 2000). Her bir aşamada, model parametrelerinin sıfıra eşitliğinin test edilmesinde elde edilen P-değerlerinden en büyük değere sahip terim silinir. Her bir değişken için en yüksek dereceli terimler test edilir. Örnek olarak, ana etki terimini silmek eğer ilgili terimin dahil olduğu yüksek dereceli etkileşimler söz konusu ise uygun olmaz.

Uygunlukları karşılaştırmak amacıyla daha karışık bir modelin kurulmasına karşı daha basit bir modelin kurulması test edilir. İki modelin karşılaştırılmasını sağlayan olabilirlik oran istatistiği sapmalardaki farka eşittir.  $P > 0.05$  durumunda üç-faktörlü terimin gerekli olduğu ileri sürülemez. Bu nedenle de basitleştirme sürecine devam edilir.

Sonraki aşamada iki-faktörlü bir etkileşimi silmek amacıyla iki faktörlü etkileşimlerin tümünün yer aldığı model dikkate alınır. Sonra iki-faktör etkileşimlerinden birini silmeye odaklanılır. Bilgisayarda uygulanan değişken seçim yöntemleri tedbirli bir biçimde kullanılmalıdır (Agresti, 1996). İlk olarak bir model içinde potansiyel olarak çok sayıda terim dikkate alındığında, bunlardan bir ya da iki tanesi gerçekten önemli olmasa bile olasılıktan dolayı etkili görülebilir. Örnek olarak gerçek etkilerin tümü zayıf olduğunda büyük örneklem etkisi nedeniyle gerçek etki aşırı olarak tahmin edilebilmektedir. Ek olarak bir modelde özel olarak ilgilenilen belirli değişkenleri dahil etmek genellikle mantıklıdır ve bazı seviyelerde istatistiksel olarak anlamlı olmasa bile tahminlenen etkileri raporlanmaktadır (Le,1998; Simonoff,2003).

### 4.5.3. Tahmin Edicilik Gücünün Korelasyon Özeti

Bir veri seti için uydurulmuş çeşitli GLM'lerin, tahmin edicilik gücü açısından karşılaştırılması bilgilendirici olur (Agresti, 1996). Gözlemlenen yanıtlar  $\{Y_i\}$  ve modelin uydurulmuş değerleri  $\{\hat{\mu}_i\}$  arasındaki korelasyon (R) bunu tanımlamaktadır. En küçük kareler regresyonuna göre ( yani normal rassal bir bileşene sahip bir GLM ) R katsayısı yanıt değişkeni ile tahmin ediciler arasındaki *çoklu korelasyonu* ifade etmektedir. Lojistik regresyona göre R katsayısı yanıt üzerindeki  $Y=(0,1)$  biçiminde iki sonuçlu gözlemler ile  $\hat{\pi}$  biçimindeki tahmin edilmiş olasılıklar arasındaki korelasyondur (Agresti, 1996; Pampel,2000; Arabacı,2002). Bu şekildeki modeller için R, tahmin edicilik gücünün kaba bir göstergesidir ve normal GLM'lerde olduğu gibi iyi özelliklere sahip değildir. Örnek olarak R değerinin model karmaşık hale geldikçe azalmamasının garantisi yoktur (Agresti,1996).

Ayrıca her hangi bir korelasyon ölçümü gibi R değeri de, açıklayıcı değişkenlerin gözlemlenen değerlerinden oluşan bir aralığa bağlıdır. Bununla birlikte aynı veri seti için farklı modellerin uyumlarının karşılaştırılması bakımından kullanışlıdır. Çoklu lojistik regresyon modellerinde farklı birimlere sahip kantitatif tahmin edicilerin etkilerini karşılaştırmak için *standartlaştırılmış katsayıları* (standardized coefficients) bulmak yardımcı olmaktadır. İlk olarak tahmin edicilerin standartlaştırılmış biçimleri ile model uydurularak bu yapılır ki tahmin edici üzerindeki her bir ölçüm z-skoru [(ölçüm değeri- ortalama)/standart hata] ile değiştirilir (Agresti, 1996). Bu durumda bulunan her bir regresyon katsayısı, diğer değişkenler kontrol altında iken bir tahmin edicide meydana gelen bir standart hatalık değişimin etkisini temsil eder.

## 4.6. LOJİSTİK REGRESYON İÇİN KESİN YORUMLAMA

Modeldeki parametre sayısı ile karşılaştırıldığında örnek hacmi büyük ise model parametrelerinin ML tahmin edicileri en iyi şekilde çalışmaktadır. Örnek hacmi küçük olduğunda ya da örnek hacmine kıyasla çok sayıda parametre söz konusu olduğunda, koşullu maksimum olabilirlik metodu kullanılarak yorumlama iyileştirilir.

### 4.6.1. Koşullu Maksimum Olabilirlik Yorumlaması

Koşullu ML metodunda, ilgilenilen önemli parametrelerin yorumlanması bir koşullu olabilirlik fonksiyonuna dayanmaktadır ve diğer parametreler elenir. Teknik olarak, diğer parametreler için “etkin istatistik” değerleri belirli ve sabit olan veri setlerine göre belirlenen koşullu olasılık dağılımı kullanılır (Agresti, 1996). Bu dağılım gözlemlenen örnekte yer alan diğer parametreler konusunda aynı bilginin elde edildiği potansiyel örnekler için tanımlanmıştır.

Dağılım ve ilgili koşullu olabilirlik fonksiyonu sadece ilgilenilen parametrelere bağlıdır. Koşullu ML metodu kanonik bağıntının kullanıldığı her hangi bir GLM için uygulanabilmektedir (Bölüm-3). Örnek olarak binom verileri ile kullanılan lojit bağıntı ve Poisson verileri ile kullanılan logaritmik bağıntı verilebilir. Özellikle lojistik regresyon modelleri çok fazla sayıda sıkıcı parametre (nuisance parameter) içerdiğinde, iki sonuçlu veriler için koşullu olabilirlik metotları oldukça kullanışlıdır. Örnek olarak, eşleştirilmiş olgu-kontrol araştırmalarında yer alan eşleştirilmiş çiftler için kurulan modeller, her bir eşleştirilmiş çift için ayrı bir parametre içermektedir. Bu nedenle de örnek hacmi ile birlikte parametre sayısı da artmaktadır. Olabilirlik fonksiyonundan sıkıcı parametrelerin silinmesi ile ilgilenilen en önemli parametreler için daha doğru yorumlamalar meydana getirilmektedir. Koşullu olabilirlik metotları aynı zamanda küçük örnekler için de kullanışlıdır. İlk olarak diğer tüm parametreler yok edilerek elde edilen koşullu olabilirlik fonksiyonu kullanılarak bir parametre için *kesin* (exact) olasılık yorumlaması oluşturulur.

#### 4.6.2. Olumsuzluk Tabloları İçin Kesin Yorumlama

İlk olarak tek bir açıklayıcı değişkenin kullanıldığı lojistik regresyon modelini göz önüne alalım:  $lojit[(\pi(x))] = \alpha + \beta x$ . X sadece iki değer aldığı anda model; X değişkeninin seviyelerini temsil eden iki satır ve Y değişkeninin seviyelerini temsil eden iki sütundan meydana gelen  $2 \times 2$  boyutlu bir tabloya uygulanır. Olağan örnekleme modelinde, Y üzerindeki yanıtlar bağımsız binom değişkenleri şeklinde ayrı x değerlerine göre işleme sokulur. Satır toplamları (binom değişkenleri için deneme sayısıdır) doğal olarak sabittir (belirlenmiştir). Bu model için bağımsızlık hipotezi  $H_0 : \beta = 0$  şeklinde oluşturulur. Bilinmeyen  $\alpha$  parametresi iki tipteki yanıt sonuçlarının göreceli sayısına dayanmaktadır.  $\alpha$  için bulunan “etkin istatistikler” , verilerin  $\alpha$ ’ yı sağlaması konusunda bilgileri içermektedir. Marjinal sütun toplamlarının hepsinin sabitlenmesi (belirlenmesi) ile hipergeometrik olasılıklar elde edilir ki bu olasılıklar her hangi bilinmeyen bir parametreye bağlı değildir.

$2 \times 2$  boyutlu tablolarda  $\beta = 0$  hipotezinin test edilmesinde kullanılan kesin koşullu test açıkça Fisher Kesin Testidir (Bölüm-2). X değişkeni I tane sıralı seviyeye sahip olduğunda, satırlara skorlar atanarak elde edilen  $I \times 2$  boyutlu bir tablo için bu lojit model kullanılır (Agresti, 1996). Yine sütun toplamları üzerine koşullandırma yapılarak  $\alpha$ ’ dan bağımsız bir koşullu olabilirlik elde edilir ve  $\beta$  konusunda kesin yorumlama yapmak amacıyla kullanılır.  $\beta = 0$  hipotezinin kesin olasılık testi büyük örnek eğilim testinin (Lojit model için “etkin skor” testinin) ya da  $\beta$  için Wald ya da Olabilirlik Oran testlerinin bir alternatifidir. Şimdi modelin, K tane seviyeden oluşan Z gibi açıklayıcı bir ikinci faktörü de kapsadığını varsayalım. Eğer Z kalitatif nitelikte ise, konuya ilişkin model şu biçimde olur:

$$lojit[(\pi(x))] = \alpha + \beta x + \beta_k^Z$$

X deęişkeni iki sonuçlu olduęunda bu model,  $2 \times 2 \times K$  boyutlu olumsuzluk tabloları için gösterilen modeldir. Kesin Olasılık Testi ile her bir kısmi tablodaki marjinal toplamlar üzerinde kořullandırmalar yapılarak dięer parametreler elimine edilir. Z kontrol altına alınarak, X ve Y deęişkenleri arasındaki kořullu baęımsızlıęın bu kesin olasılık testi daha önce açıklanmıřtı. Lojistik regresyon modellerinde kesin yorumlama için yapılacak hesaplamalar oldukça karmařıktır. Fakat yazılımlarda bu hesaplamalar mevcuttur. Aslında LogXact yazılımındaki kılavuz bu yaklařım açısından iyi bir kaynaktır (Agresti, 2002). Örnek hacmi küçük olduęunda kesin testler ve bu testlere iliřkin güven aralıkları, sıradan büyük örnek-ML yorumlamasından daha güvenilirdir. Kesin olasılık metotları aynı zamanda, sıradan ML metotları ile sınırsız bir parametre tahmini söz konusu olduęunda da kullanıřlıdır.

## **BEŞİNCİ BÖLÜM**

### **OLUMSALLIK TABLOLARI İÇİN LOG-LİNEER MODELLER**

Daha önce Poisson dağılımlı veriler için genelleştirilmiş lineer modeller (GLM) şeklinde log lineer modeller ortaya konmuştu. Bu modeller içinde en yaygın biçimde kullanılan olumsuzluk (kontenjans) tablolarında yer alan hücre sayılarının modellenmesidir. Modeller hücre sayısının büyüklüğünün ilgili hücreye göre kategorik değişkenlerin seviyelerine nasıl bağlı olduğunu belirler. Bu belirlemenin doğası değişkenler arasındaki etkileşim ve birliktelik yapısı ile ilgilidir.

Log lineer modeller kategorik değişkenlerden oluşan bir küme içerisinde birliktelik ve etkileşim biçimlerini tanımlamaktadır. Diğer bir ifade ile log lineer modeller çok-yönlü çapraz tablolardaki karmaşık ilişkileri ortaya çıkarmaktadır. Bunun yanı sıra, beklenen ve gözlenen hücre frekansları arasındaki farkı azaltmaya çalışırken, beklenen hücre frekanslarını belirlemek için etkilerin bir modelini oluşturmaktır (Garson, 1999).

Bir değişken iki sonuçlu bir yanıt değişkeni ve diğer değişkenler açıklayıcı değişken olduğunda ilgili yanıt değişkeni için lojit modeller belirli log lineer modellerin karşılığıdır. Bir olumsuzluk tablosunda en azından iki değişken yanıt değişkeni olduğunda temel olarak log lineer modellerin kullanımına bakılacaktır. Tek bir yanıt olduğunda lojit modelleri kullanmak daha doğal ve daha basittir (Agresti, 1996). Log-linear modeller iki veya daha fazla kategorik değişkenin koşullu ilişkisini analiz etmek için geliştirilmiştir (Garson, 1999). Bununla birlikte, log-linear modeller sayesinde, değişkenlerin oluşturduğu bileşik dağılımı, iki veya daha fazla değişkenin birbirine bağımlı olup olmadığını ve iki veya daha fazla değişken arasındaki ilişkiyi neden-sonuç ilişkisine dayandırmaksızın test etmek mümkündür (Uygun, 1990).

## 5.1. İKİ YÖNLÜ TABLOLAR İÇİN LOG LİNEER MODELLER

$n$  tane deneğin iki kategorik yanıt üzerinde çapraz biçimde sınıflandırıldığı  $I \times J$  boyutlu bir olumsallık tablosunu dikkate alalım. Bu değişkenler istatistiksel olarak bağımsız olduğunda, ilgili tablodaki hücreler için ortak olasılıklar ( $\{\pi_{ij}\}$ ) satır ve sütun marjinal toplamları ile belirlenmektedir:

$$\pi_{ij} = \pi_{i+} \pi_{+j}, \quad i=1, \dots, I \text{ ve } j=1, \dots, J.$$

$\{\mu_{ij} = n\pi_{ij}\}$  şeklindeki beklenen frekanslar için tüm  $i$  ve tüm  $j$ ' lere göre ilgili ifade şu şekildedir:  $\mu_{ij} = n\pi_{i+} \pi_{+j}$ . Bu formüller,  $\{\pi_{ij}\}$  şeklinde gösterilen hücre olasılıklarına dayanmaktadır ki bu olasılıklar binom ve çoklu dağılımlar için kullanılan parametrelerdir. Log lineer modellerinin formüllerinde  $\{\pi_{ij}\}$  hücre olasılıklarından öte  $\{\mu_{ij}\}$  kullanılmaktadır. Bundan dolayı bu formüller,  $\{\mu_{ij}\}$  şeklindeki beklenen değerlerle hücre sayıları için Poisson örneklemesine de uygulanmaktadır.

### 5.1.1. Bağımsız Model

$X$  satır değişkenini  $Y$  ise sütun değişkenini gösterebiliriz.  $X$  ve  $Y$  arasındaki bağımsızlık koşulu ( $\mu_{ij} = n\pi_{i+} \pi_{+j}$ ) çarpımsal niteliktedir. Eşitliğin her iki yanının logaritması alındığında toplamsal bir ilişki elde edilir.

Şöyle ki  $\log \mu_{ij}$  ifadesi örnek hacmine dayanan bir terime ( $i$ ' ninci satırdaki olasılığa dayanan bir terim ile  $j$ ' ninci sütundaki olasılığa dayanan bir terim) toplamsal bir şekilde bağlıdır. Böylece bağımsızlık aşağıdaki şekle sahip olur:

$$\log \mu_{ij} = \lambda + \lambda_i^X + \lambda_j^Y \quad (5.1.1)$$

Logaritmik beklenen frekans vasıtasıyla,  $\lambda_j^Y$  şeklindeki bir sütun etkisinin ve  $\lambda_i^X$  şeklindeki bir satır etkisinin toplanabilir bir fonksiyonu elde edilir ( X ve Y şeklindeki ifadeler birer etikettir, bir üssün kuvvetini ifade etmemektedir). Bu model iki yönlü olumsuzluk tabloları için *bağımsızlığın log lineer modeli* olarak adlandırılır (Agresti, 1996; Heijden, Mooijaart ve Takane, 1994:80).  $\lambda_i^X$  parametresi X değişkeninin i' ninci satırındaki sınıflamanın etkisini temsil eder. Bu parametre büyük değerler aldıkça tablonun i' ninci satırında daha büyük beklenen frekans değerleri söz konusu olur.

$\lambda_h^X = \lambda_i^X$  olduğunda h' ninci satırdaki her bir beklenen frekans i' ninci satırdaki ilgili beklenen frekansa eşit olur. Benzer şekilde  $\lambda_j^Y$  parametresi de Y değişkeni için j' ninci sütundaki sınıflamanın etkisini temsil eder. İki kategorik değişken arasındaki bağımsızlığın ifade edildiği sıfır hipotezi açıkça log lineer modelde kurulan hipotezdir. Modeli sağlayan uydurulmuş değerler;  $\{\hat{\mu}_{ij} = n_{i+}n_{+j} / n\}$  şeklindedir ve ki-kare bağımsızlık testleri için tahminlenen beklenen frekanslardır.  $X^2$  ve  $G^2$  istatistiklerinin kullanıldığı ki-kare testleri aynı zamanda log lineer modellerin uyum iyiliği testleridir (Agresti, 1996;Acar, 2000).

Tablo-5.1' de bağımsız log lineer modeline (5.1.1) göre hücre sayıları ve bunlara karşılık gelen uydurulmuş değerlerden ( $\{\hat{\mu}_{ij}\}$ ) oluşan  $2 \times 2$  boyutlu bir tablo yer almaktadır. Hücre sayıları tablo-2.1' den alınmıştır<sup>51</sup>. X=cinsiyet ve Y=Ölümünden sonraki hayata inanma değişkenlerinin çapraz sınıflandırılmasıdır. Uydurulmuş değerler bağımsızlığı sağladığından dolayı, odds oranları 1' dir.

Uydurulmuş değerler gözlemlenen sayılara yakındır ve bağımsız log lineer modelinin test edilmesi için iyi uyum istatistikleri 1 serbestlik derecesi ile şu şekildedir:  $X^2=0.2$  ve  $G^2=0.2$ . Bu istatistikler açıkça iki yönlü tablolarda bağımsızlık



hipotezinin test edilmesinde kullanılan test istatistikleridir. Bağımsız model bu veriler için uygundur.

**Tablo-5.1** Ölümünden Sonraki Hayata İnanmanın (Sütunlar) Cinsiyet (Satırlar) İle Olan İlişkisinin Yer Aldığı Tablo 2.1 İçin Uydurulan Bağımsız Modele Göre Parametre Tahminleri

Gözlemlenen Frekanslar		Uydurulmuş Değerler		Log. Uydurulmuş Değerler	
435	147	432.10	149.90	6.069	5.010
375	134	377.90	131.10	5.935	4.876
Parametreler		Set 1	Set 2	Set 3	
$\lambda$		4.876	6.069	5.472	
$\lambda_1^X$		0.134	0	0.067	
$\lambda_2^X$		0	-0.134	-0.067	
$\lambda_1^Y$		1.059	0	0.529	
$\lambda_2^Y$		0	-1.059	-0.529	

### 5.1.2. Parametrelerin Yorumlanması

Olumsuzluk tabloları için kullanılan log lineer modeller genelleştirilmiş lineer modellerin örnekleridir.  $I \times J$  boyutlu tablolar için bu GLM; Poisson rassal bileşenin N tane bağımsız gözlemi olarak  $N=IJ$  hücre sayılarını işleme tabi tutar. Log lineer GLM' ler için veriler, n tane deneğin her bir sınıflandırılmasından öte N tane hücre sayısı şeklindedir (Agresti, 1996). Hücre sayılarının beklenen değerleri ( $\{\mu_{ij}\}$ ) logaritmik bağıntının kullanıldığı açıklayıcı terimler ile birleştirilmiştir.

Formül 5.1.1' de gösterildiği gibi olumsuzluk tabloları için uygulanan log lineer modeller değişkenleri açıklayıcı ve yanıt sınıfı şeklinde ayırmaz. Bu modeller tüm değişkenleri ortak bir şekilde yanıt değişkenleri olarak işleme tabi tutar. Değişkenlerin tüm seviyelerine göre oluşturulan tüm kombinasyonlar için hücre sayısını modeller. Fakat çoğu araştırmada yanıt değişkenleri ve açıklayıcı faktörler gibi bazı sınıflandırmalar belirlenir. Bu durum parametrelerin yorumlanmasında ve model seçiminde etkili olmaktadır (Agresti, 1996). En basit parametre yorumlaması iki sonuçlu yanıtlar için yapılmaktadır. Örnek olarak, Y yanıtının seviyeleri iki sütunla gösterilen  $I \times 2$  boyutlu tablolar için bağımsız modeli dikkate alalım.  $i$ ' ninci satırda  $\pi$  olasılığının (Y=1 için) lojiti aşağıdaki gibidir:

$$\begin{aligned} \log \left( \frac{\pi}{1 - \pi} \right) &= \log \left( \frac{\mu_{i1}}{\mu_{i2}} \right) = \log \mu_{i1} - \log \mu_{i2} \\ &= (\lambda + \lambda_i^X + \lambda_1^Y) - (\lambda + \lambda_i^X + \lambda_2^Y) \\ &= \lambda_1^Y - \lambda_2^Y. \end{aligned}$$

Son terim bir sabittir ve  $i$ ' ye bağlı değildir. Bundan dolayı Y için lojit X' in seviyesine bağlı değildir. Bağımsız Log Lineer Model,  $\text{lojit}(\pi)=\alpha$  formuna sahip olan basit bir modelin karşılığıdır ki birinci sütundaki yanıt olasılığı  $\pi$  için elde edilen lojit her bir satırda aynı değeri alır. Her bir satırda, birinci sütundaki yanıtın odds' u;  $\exp(\alpha) = \exp(\lambda_1^Y - \lambda_2^Y)$  ifadesine eşittir. Belirli bir sütunda sınıflandırma ihtimali tüm satırlarda aynıdır. (5.1.1)' deki log lineer modelde, belirli bir değişken için iki parametre arasındaki farklılıklar, ilgili değişken üzerinde diğer değişkenlerle ilgili olarak bir yanıt oluşturmanın log odds'u ile ilgilidir (Agresti, 1996; Acar, 2000; Uygun, 1990). Elbette iki sonuçlu tek bir yanıt değişkeni söz konusu olduğunda ilk olarak lojit modeller doğrudan uygulanır, log lineer modeller gereksizdir. Esasında log lineer modeller, iki ya da daha fazla kategorik yanıt değişkeni arasında var olan ilişkileri modelleme açısından kullanışlıdır.

### 5.1.3. Parametre Kısıtları

Tablo-5.1 aynı zamanda bağımsız modele göre uydurulmuş değerlerin logaritmalarını da içermektedir.  $\{\log \hat{\mu}_{ij}\}$  değerleri,  $\hat{\lambda} = 4.876$ ,  $\hat{\lambda}_1^X = 0.134$ ,  $\hat{\lambda}_2^X = 0$ ,  $\hat{\lambda}_1^Y = 1.069$  ve  $\hat{\lambda}_2^Y = 0$  şeklindeki parametre tahminleri ile birlikte (5.1.1) formülünü sağlamaktadır. Örnek olarak  $\hat{\mu}_{12} = 149.90$  ve (5.1.1)'e göre de;  $\log \hat{\mu}_{12} = \hat{\lambda} + \hat{\lambda}_1^X + \hat{\lambda}_2^Y = 4.876 + 0.134 + 0 = 5.010 = \log(149.90)$  şeklinde bulunur.

Model aynı zamanda tahmin edilen parametreler için Tablo-5.1' in altındaki kısımda diğer olası değerleri de sağlamaktadır. Parametrelerin belirlenmesinde ya da tahminlenmesinde tek bir yol yoktur ve Tablo-5.1' de tahminlerin olası üç kümesi gösterilmektedir. Bağımsız modele göre, bir  $\{\lambda_i^X\}$  parametresi ile bir  $\{\lambda_j^Y\}$  parametresi fazlalık oluşturmaktadır. Bu ANOVA ile faktör seviyelerinin sayısına kıyasla daha az sayıda kukla değişkenin gerektiği ve faktörlere göre oluşturulan çoklu lojistik regresyon modellerine benzemektedir (Agresti, 1996). Daha önce geriye doğru eleme yöntemi gibi fazlalıkların elimine edilmesine ilişkin yöntemler gösterilmişti. Tablo-5.1' de tahminlerin ilk sütununda (set1) olduğu gibi ilk olarak sifıra eşit her bir faktörün son seviyesine göre parametre tespit edilir. Diğer bir yaklaşımda ise her bir faktör için toplamı sıfır olan parametreler alınır; örnek olarak  $2 \times 2$  boyutlu bir tablo için şu şekildedir (Tablonun üçüncü sütununda yer almaktadır):  $\lambda_1^X + \lambda_2^X = 0$  ve  $\lambda_1^Y + \lambda_2^Y = 0$ . Sınırlamaların seçimi keyfidir.

Log lineer modeller için kullanılan bazı yazılımlar (SAS-GENMOD ve GLIM şeklinde) sifıra eşit bir parametreyi tespit eder. Diğer yazılımlar ise SAS-CATMOD gibi toplamları sıfır olan parametreleri belirler (Stokes, Davis ve Koch, 1991). Belirli bir türdeki ana etkiye sahip iki parametre arasındaki fark tektir.

Örnek olarak Tablo-5.1' de her bir tahmin kümesi (Set 1, Set 2, Set 3) için  $\lambda_1^Y - \lambda_2^Y = 1.059$  şeklindedir. Her bir cinsiyet durumu için, ikinci sütun yerine birinci sütundaki yanıtın tahminlenen log odds'u 1.059 olur. Bu nedenle ölümden sonraki

hayata inanmanın Tahminlenen odds' u  $\exp(1.059)=2.9$  değerini alır. Ölümden sonraki hayata inanmanın tahminlenen olasılığı ise  $\exp(1.059)/[1+\exp(1.059)]=0.74$  olarak bulunur.

#### 5.1.4. Doygun Model

İstatistiksel olarak bağımsız olmaktan çok bağımlı değişkenler, aşağıda yer alan daha karmaşık bir log lineer modeli sağlamaktadır:

$$\log \mu_{ij} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY} \quad (5.1.2)$$

$\{\lambda_{ij}^{XY}\}$  parametreleri X ve Y değişkenlerinin bağımsızlıktan sapmalarını yansıtan terimlerle ilişkilidir. Parametreler, bir değişkenin başka bir değişkenin seviyesine bağlı olan beklenen hücre sayısı üzerindeki etkisi aracılığı ile X ve Y değişkenleri arasındaki etkileşimleri temsil etmektedir. Bu model beklenen frekansların her hangi bir kümesini mükemmel şekilde tanımlamaktadır ki “Doygun Log Lineer Model” olarak adlandırılmaktadır ve iki yönlü olumsuzluk tabloları için en genel modeldir (Agresti, 1996). (5.1.1)’ de gösterilen bağımsız model, tüm  $\lambda_{ij}^{XY}$  değerlerinin sıfıra eşit olduğu özel bir durumdur. Dolaysız ilişkiler,  $\{\lambda_{ij}^{XY}\}$  şeklindeki birliktelik parametreleri ve logaritmik odds oranları arasında söz konusudur.

Örnek olarak;  $2 \times 2$  boyutlu tablolar için doyun modelin logaritmik odds oranı aşağıdaki şekildedir:

$$\begin{aligned} \log \theta &= \log \left( \frac{\mu_{11}\mu_{22}}{\mu_{12}\mu_{21}} \right) = \log \mu_{11} + \log \mu_{22} - \log \mu_{12} - \log \mu_{21} \\ &= (\lambda + \lambda_1^X + \lambda_1^Y + \lambda_{11}^{XY}) + (\lambda + \lambda_2^X + \lambda_2^Y + \lambda_{22}^{XY}) \\ &\quad - (\lambda + \lambda_1^X + \lambda_2^Y + \lambda_{12}^{XY}) - (\lambda + \lambda_2^X + \lambda_1^Y + \lambda_{21}^{XY}) \\ &= \lambda_{11}^{XY} + \lambda_{22}^{XY} - \lambda_{12}^{XY} - \lambda_{21}^{XY} \end{aligned} \quad (5.1.3)$$

Bundan dolayı  $\{\lambda_{ij}^{XY}\}$  değerleri logaritmik odds oranını belirlemektedir. Bu parametreler sıfıra eşit olduğu zaman logaritmik odds oranı da (sıfıra eşit olmaktadır. Odds oranı 1' e eşit ise X ve Y değişkenleri bağımsızdır. Tablo-5.1' de gözlemlenen sayılar  $(435 \times 134)/(147 \times 375) = 1.057$  şeklinde bir odds oranına (ve  $\log(1.057)=0.056$  şeklinde bir logaritmik odds oranına sahiptir. Tablo-5.2' de olası birliktelik parametre tahminleri gösterilmektedir ki bunlar aşağıda gösterilen bir logaritmik odds oranına sahiptir:

$$\hat{\lambda}_{11}^{XY} + \hat{\lambda}_{22}^{XY} - \hat{\lambda}_{12}^{XY} - \hat{\lambda}_{21}^{XY} = 0.056. \quad (5.1.4)$$

İlk olarak birliktelik parametreleri belirlenir. Buna göre son satır ve son sütundaki ilk değerler sıfırdır ya da ilk satır ve ilk sütundaki ilk değerler sıfırdır. Ya da ilk olarak her bir satır ve her bir sütun içinde sıfıra eşit parametrelerin toplamı alınır. Bu kombinasyon (5.1.4) ile yapılan her bir potansiyel tanım için özdeştir. 1 x J boyutlu tablolarda sadece (I-1)(J-1) tane birliktelik parametresi fazlalık teşkil etmez.

Doğun modeldeki bu “etkileşim” parametreleri, Y için belirlenen (J-1) tane kukla değişken ile X için belirlenen (I-1) tane kukla değişkenin çapraz çarpımlarının katsayılarıdır (Agresti, 1996). Bağımsızlık testleri ile bu (I-1)(J-1) tane parametrenin sıfıra eşit olup olmadığı analiz edildiğinden dolayı bunların artık serbestlik derecesi (I-1)(J-1) şeklindedir. I=J=2 olduğunda (Tablo-5.1' deki gibi), tek bir gereksiz olmayan parametre ile odds oranı belirlenir. (5.1.2)' deki doğun model;  $\lambda$  şeklinde tek bir sabit parametreye,  $\lambda_i^X$  formunda (I-1) tane gereksiz olmayan parametreye,  $\lambda_j^Y$  formunda (J-1) tane gereksiz olmayan parametreye ve  $\lambda_{ij}^{XY}$  formunda (I-1)(J-1) tane gereksiz olmayan parametreye sahiptir. Gereksiz olmayan parametrelerin toplam sayısı,  $1+(I-1)+(J-1)+(I-1)(J-1)=IJ$  şeklinde bulunur. Bu model olası maksimum parametre sayısına sahiptir. Bu nedenden dolayı da doğun olarak nitelendirilir.

Uygulamada doymamış modeller kullanılarak veriler analiz edilmeye çalışılmaktadır. Çünkü bu modellerin uyumu örnek verileri düzeltmektedir ve daha basit yorumlamalar sağlamaktadır. Üç boyutlu ve daha yüksek boyuttaki tablolar için

doymamış modeller, birliktelik terimlerini de kapsamaktadır. Log lineer modeller odds' ların tanımlanmasından (Tek faktörlü terimler aracılığı ile) ziyade birlikteliklerin tanımlanması (İki faktörlü terimler aracılığı ile) için daha yaygın biçimde kullanılmaktadır (Agresti, 1996). Log lineer modellerin uydurulmasından sonra iki faktörlü parametre tahminleri, değişken çiftleri arasındaki koşullu odds oranlarının tahminlerine dönüştürülür.

**Tablo-5.2** Doygun Log Lineer Model İçin Eşdeğer Birliktelik Parametreleri

Birliktelik Parametresi	Set 1	Set 2	Set 3
$\lambda_{11}^{XY}$	0.056	0.014	0.056
$\lambda_{12}^{XY}$	0	-0.014	0.0
$\lambda_{21}^{XY}$	0	-0.014	0.0
$\lambda_{22}^{XY}$	0	0.014	0.056

Bu çalışmada kullanılan model (5.1.2) ve diğerleri “*Hiyerarşik Modeller*” olarak adlandırılmaktadır (Agresti,1996; Uygun, 1990). Bunun anlamı ise şudur: Model, modeldeki en yüksek dereceli terimdeki değişkenlerden en düşük dereceli terimlere kadar hepsini kapsamaktadır. Üç yönlü veya çok yönlü tablolar için en çok kullanılan yöntem hiyerarşik logaritmik dorusal modellerdir. Örnek olarak model  $\lambda_{ij}^{XY}$  terimini içeriyorsa aynı zamanda  $\lambda_i^X$  ve  $\lambda_j^Y$  terimlerini de kapsar. Benzer şekilde sıradan regresyon ya da lojistik regresyon modeli iki tahmin edici için bir etkileşim terimini ( $\beta_{x_1x_2}$  şeklinde) ya da kuadratik bir terimi ( $\beta_{x_1^2}$ ) içeriyorsa aynı zamanda ana etki terimlerini de ( $\beta_1x_1$  ve  $\beta_2x_2$ ) içermektedir. Daha düşük derecedeki terimleri içermesi söz konusu olsa dahi daha yüksek derecedeki bir terimin istatistiksel anlamlılığı ve pratik yorumlaması değişkenlerin nasıl kodlandığına bağlıdır. Bu istenilmeyen sakıncalı bir durumdur ve hiyerarşik

modellerle birlikte benzer sonuçlar değişkenlerin nasıl kodlandığı sorunu olmadan elde edilir. Bir model iki-faktörlü terimlere sahipse, daha düşük dereceli tek faktör terimlerinin yorumlanmasında dikkatli olunmalıdır.

İki yönlü ANOVA ile benzeşim aracılığıyla, iki faktör etkileşimi söz konusu ise ana etkilerin raporlanması yanıltıcı olabilir. Ana etki terimlerinin tahminleri daha yüksek dereceli etkiler için kullanılan kodlama düzenine bağlıdır. Aynı zamanda yorumlama da ilgili düzene bağlıdır. Normal olarak, bu kaynakta bir değişken için en yüksek dereceli terimlere dikkat çekilerek sınırlama getirilmiştir.

## 5.2. ÜÇ YÖNLÜ TABLOLAR İÇİN LOG LİNEER MODELLER

Bu bölümde üç yönlü olumsuzluk tabloları için log lineer modeller ele alınmaktadır. Farklı modeller farklı bağımsızlık ve birliktelik şekillerine sahiptir. İyi uyumlu bir model bu şekillerin tanımlanmasında yardımcı olmaktadır ve ikili odds oranlarının gücünü de tahmin etmektedir.

### 5.2.1. Log Linear Modeller ve Bağımsızlık Yapısı

Olumsuzluk tablosundaki beklenen hücre frekansları  $\{\mu_{ijk}\}$  ile gösterilsin.  $\{\mu_{ijk}\}$  için log lineer modellerdeki tek faktörlü terimler marjinal dağılımları temsil etmektedir. Örnek olarak, bir modelin  $\lambda_i^X$  terimini içermesi gözlemlenen değerlerde olduğu gibi  $\{\hat{\mu}_{ijk}\}$  şeklindeki uydurulmuş değerleri, X değişkeninin çeşitli seviyelerinde aynı toplamlara sahip olmaya zorlamaktadır (Agresti, 1996). Log lineer modellerde başlıca kullanılan bilgi yüksek dereceli parametrelerle ilgilidir. Değişkenler arasındaki kısmi birliktelikler iki faktör terimlerinin karşılığıdır (Yılmaz ve Şıklar, 2002). Örnek olarak aşağıda yer alan log lineer modelini dikkate alalım:

$$\log \mu_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ} \quad (5.2.1)$$

Bu model X-Z iki faktör terimini içerdiğinden dolayı Y değişkeninin kontrol altına alınması ile X ve Z arasındaki birliktelik elde edilebilmektedir. Bu model aynı zamanda X değişkeni kontrol altına alınarak bir Y-Z birlikteliğinin elde edilmesine de olanak tanır. X-Y için iki faktör terimi içermemektedir. Bu log lineer model ile Z değişkeni kontrol altına alınarak X ve Y değişkenleri arasındaki koşullu bağımsızlık belirlenebilmektedir. X-Y koşullu bağımsızlığının sağlandığı log lineer model, (XZ, YZ) ile sembolize edilmektedir. Bu sembol her bir değişken için modeldeki en yüksek dereceli terimleri listeler. Örnek olarak; eğer Z değişkeni kontrol altına alındığında X ve Y arasında bir birliktelik söz konusu ise bu model kabul edilir.  $2 \times 2 \times K$  boyutlu tablolar için bu model daha önce ele alınan Cochran- Mantel-Haenszel istatistiği kullanılarak hipotez testinin yapılmasına benzemektedir.

(5.2.1)' de yer alan modelden daha basit modellerde iki-faktör birliktelik terimleri silinmiştir ve çoğu veri setine uydurmak açısından oldukça basittir. Örnek olarak, sadece tek-faktör terimlerini içeren model (X,Y,Z), “*karşılıklı bağımsızlık modeli*” (mutual independence model) olarak adlandırılır (Agresti, 1996). Değişken ikililerinden her birini bağımsız olarak işleme tabi tutar. Bir araştırmada değişkenler mantıksal olarak seçildiğinde model nadiren uygun ve yeterli olur. Üç tane değişken ikilisinin tümüne koşullu biçimde bağımlı olma imkânı veren model aşağıdaki biçimdedir:

$$\log \mu_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ} \quad (5.2.2)$$

Bu modele göre daha sonra her hangi iki değişken arasındaki koşullu odds oranlarının üçüncü bir değişkenin her bir seviyesinde özdeş olduğu gösterilmektedir. Bu homojen birliktelik özelliğidir. Homojen birliktelik modeli şeklinde bu log lineer modele başvurulmaktadır ve (XY, XZ, YZ) aracılığı ile sembolize edilmektedir.



### 5.2.2. Model Parametrelerinin Yorumlanması

Üç yönlü tablolar için kullanılan genel log-lineer modellerin çoğu aşağıdaki biçimdedir:

$$\log \mu_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ} + \lambda_{ijk}^{XYZ} \quad (5.2.3)$$

Üç-faktörlü terim, üç-faktör etkileşimi ile ilgilidir. (X,Y,Z) ile gösterilen bu model her hangi iki değişken arasındaki odds oranının (odds ratio) üçüncü bir değişkenin seviyelerine karşı değişim gösterebilmesine imkân vermektedir. Bu üç yönlü bir tablo için doygun modeldir ve mükemmel bir uyum sağlamaktadır. (5.2.2)'deki homojen birliktelik modeli ( üç-faktörlü terimin silindiği) genellikle üç faktör etkileşiminin olmadığı model olarak adlandırılmaktadır. (5.2.3)'ün sağ tarafı, üç yönlü Faktöryel ANOVA için kullanılan modeller ile aynı yapıya sahiptir(Agresti, 1996). Tek faktörlü terimler ana etkilerdir ve iki faktörlü etkileşimler kısmi birliktelikleri temsil etmektedir. Log lineer modellerin yorumlanmasında en yüksek dereceli terimlere başvurulmaktadır. Örnek olarak, homojen birliktelik modeli için yorumlamalarda birliktelikleri tanımlamak amacıyla iki faktörlü terimler kullanılır. Bu modellerde ana etki tahminleri konu dışıdır ve yüksek dereceli terimler için parametre sınırlamalarına bağlı olarak çarpıcı biçimde farklı değerler alırlar. İki faktörlü parametreler doğrudan, üç yönlü tablolar için koşullu odds oranları ile ilgilidir.  $2 \times 2 \times K$  boyutlu tablolar için durumu açıklayalım. (XY, XZ, YZ) modelini ya da X-Y terimini içeren daha basit her hangi bir modeli dikkate alalım. X-Y koşullu odds oranı,  $\theta_{XY(k)}$ , k' nıncı kısmi tabloda X ve Y arasındaki birlikteliği tanımlamaktadır:

$$\log \theta_{XY(k)} = \log \left( \frac{\mu_{11k} \mu_{22k}}{\mu_{12k} \mu_{21k}} \right) = \lambda_{11}^{XY} + \lambda_{22}^{XY} + \lambda_{12}^{XY} + \lambda_{21}^{XY} \quad (5.2.4)$$

Bu ifadenin sađ tarafı k' ya bađlı deđildir. Bu nedenle odds oranı Z deđiřkeninin her bir seviyesi iin aynıdır. Benzer řekilde (XY, XZ, YZ) modeli Y deđiřkeninin farklı seviyelerinde eřit X-Z odds oranlarına ve X deđiřkeninin farklı seviyelerinde eřit Y-Z odds oranlarına sahiptir. Ü faktör etkileřim terimine sahip olmayan her hangi bir model homojen birliktelik yapısına sahiptir.

### 5.2.3. Alkol, Sigara Ve Marihuana Kullanımına İliřkin Örneđ

Tablo-5.3 Ohio, Dayton' da bulunan Birleřmiř Sađlık Servisleri ve Wright Eyalet Üniversitesi Tıp Okulu tarafından 1992' de yürütölen bir arařtırmadan alınmıřtır. Ohio, Dayton yakınlarında kentsel olmayan bir alanda lise son sınıfta okuyan öđrencilere, alkol, sigara ve ya marihuana kullanıp kullanmadıkları sorulmuřtur (Agresti, 1996).  $2 \times 2 \times 2$  boyutlu bir tabloda deđiřkenler sırasıyla, Marihuana kullanımı iin M, sigara kullanımı iin C ve alkol kullanımı iin A ile gösterilmektedir. Geniř biimde kullanılan yazılımlar log lineer modelleri uydurmaktadır. Burada hesaplamalardan daha ok yorumlayıcı özelliklere vurgu yapılmaktadır.

**Tablo-5.3** Lise Son Sınıf Öđrencileri İin Alkol(A), Sigara(C) ve Marihuana (M) Kullanımı

Alkol Kullanımı	Sigara Kullanımı	Marihuana Kullanımı	
		Evet	Hayır
Evet	Evet	911	538
	Hayır	44	456
Hayır	Evet	3	43
	Hayır	2	279

Kaynak: Agresti, 1996

Tablo-5.4 ise birkaç log lineer model için uydurulmuş değerler gösterilmektedir. (AC, AM, CM) modelinin uyumu gözlemlenen verilere yakındır ki (ACM) modeli için uydurulmuş değerlerdir. Diğer modeller ise zayıf bir uyum göstermektedir.

Tablo-5.5' te ise bu modellerdeki koşullu ve marjinal birliktelikler için tahmin edilen odds oranlarının gösterilmesi aracılığıyla bu modellerdeki birliktelik biçimleri açıklanmaktadır. Örnek olarak; (AM, CM) modeli için A-C koşullu birlikteliği 1 şeklinde kayıt edilmiştir ki bu M' nin iki seviyesindeki A-C uydurulmuş odds oranlarının ortak değeridir ve aşağıdaki şekilde hesaplanmıştır:

$$1.0 = \frac{909.24 \times 0.24}{45.76 \times 4.76} = \frac{438.84 \times 179.84}{555.16 \times 142.16}$$

$$2.7 = \frac{(909.24 + 438.84)(0.24 + 179.84)}{(45.76 + 555.16)(4.76 + 142.16)}$$

**Tablo-5.4** Tablo-5.3' teki Verilere Uygulanan Log Lineer Modeller İçin Beklenen Değerler

Alkol Kullanımı	Sigara Kullanımı	Marihuana Kullanımı	Log-Lineer Model				
			(A,C,M)	(AC, M)	(AM,C M)	(AC, AM, CM)	(AC M)
Evet	Evet	Evet	540	611.	909.24	910.4	911
		Hayır	740.2	837.	438.84	538.6	538
	Hayır	Evet	282.1	210.	45.76	44.6	44
		Hayır	386.7	289.	555.16	455.4	456
Hayır	Evet	Evet	90.6	19.4	4.76	3.6	3
		Hayır	124.2	26.6	143.16	42.4	43
	Hayır	Evet	47.3	228.	0.24	1.4	2
		Hayır	64.9	162.	179.84	279.6	279

Bu model, marihuana kullanımı kontrol altında iken sigara ve alkol kullanımı arasındaki koşullu bağımsızlığı yansıtmaktadır. Ayrıca A-C koşullu birlikteliği için 1.0 değerindeki olasılık oranlarını sağlamaktadır. Bu modele göre A-C marjinal birlikteliği için 2.7 şeklinde bir değer kayıtlanmıştır ki bu değer marjinal A-C uydurulmuş değer gösterildiği tabloya göre uydurulmuş odds oranıdır: Gözlemlenen veriler için odds oranları, (ACM) doymun modeli için rapor edilmiştir ki bu model mükemmel bir uyum sağlamaktadır. (AC, AM, CM) modeli tüm ikili birlikteliklerin bulunmasını sağlamaktadır, fakat iki değişken arasında üçüncü bir değişkenin her bir seviyesine göre homojen olasılık oranlarını da korumaktadır. Bu modele göre A-C uydurulmuş koşullu olasılık oranları 7.8 değerine eşittir ki M' in her bir seviyesi için geçerlidir. Buna göre şu şekilde bir yorum yapılır: Sigara kullanan öğrencilerin alkol kullanmasına ilişkin tahmin edilen olasılık (odds) sigara kullanmayan öğrenciler için tahmin edilen olasılığın (odds'un) 7.8 katıdır.

İlk olarak M' in her iki seviyesinde, modelin uydurulmuş değerleri kullanılarak bu odds oranı hesaplanır ya da ((5.2.4)' den) parametre tahminleri kullanılarak aşağıdaki ifade kurulur:

$$\exp(\hat{\lambda}_{11}^{AC} + \hat{\lambda}_{22}^{AC} - \hat{\lambda}_{12}^{AC} - \hat{\lambda}_{21}^{AC})$$

17.7 değerindeki tahmin edilmiş A-C marjinal odds oranının hesaplanmasında üçüncü faktör M ihmal edilir. Oysa 7.8 değerindeki koşullu odds oranı hesaplanırken bu değişken kontrol altına alınır.

**Tablo-5.5** Tablo-5.4’ te Yer Alan Log Lineer Modeller İçin Tahmin Edilmiş Odds Oranları

Model	Koşullu Birliktelik			Marjinal Birliktelik		
	-C	-M	C-M	A-C	A-M	C-M
(A,C, M)	.0	.0	.0	1.0	1.0	1.0
(AC, M)	7.7	.0	.0	17.7	1.0	1.0
(AM, CM)	.0	1.9	25.1	2.7	61.9	25.1
(AC,AM,CM)	.8	9.8	7.3	17.7	61.9	25.1
(ACM) Seviye 1	3.8	4.3	7.5	17.7	61.9	25.1
(ACM) Seviye 2	7	3.5	9.7			

Tablo-5.5’ te (AM, CM) modelindeki A-C birlikteliğinde olduğu gibi tahmin edilen koşullu odds oranları bir modelde görünmeyen her bir ikili terim için 1.0 değerini alır. İlgili modele göre tahmin edilen A-C marjinal odds oranı 1.0 değerinden farklıdır. Daha önce koşullu bağımsızlığın marjinal bağımsızlığı yansıtmadığına işaret edilmişti. Bazı modeller, marjinal birlikteliklerle aynı olan koşullu birlikteliklere sahiptir. Bu eşitlik, tüm ikili birliktelikleri içeren log lineer modeller için gerçekleşmemektedir. Tablo-5.5’ te yüksek derecede modele bağlı marjinal ve koşullu odds oranlarının tahminleri gösterilmektedir. Bu iyi model seçiminin önemini vurgulamaktadır. Bu tablodan elde edilen bir tahmin sadece ilgili modelin iyi uyum göstermesi konusunda bilgi verici niteliktedir. Bir sonraki bölümde modelin iyi uyum gösterip göstermediğinin nasıl kontrol edileceği gösterilmektedir.

### 5.3. LOG LİNEER MODELLER İÇİN YORUMLAMA

İyi uyumlu bir log lineer model; bir dizi kategorik yanıt değişkeninin aralarındaki gerçek birliktelik konusunda yorumlamalar yapmak ve bunları tanımlamak açısından bir kaynak sağlamaktadır (Agresti, 1996). Burada log lineer modellerin uyum iyiliklerinin kontrol edilmesine ve model parametrelerinin yorumlanmasına ilişkin yöntemler ele alınmaktadır.

#### 5.3.1. Log Lineer Modellerin Uygunluğu

Bazı log lineer modeller,  $\{n_{ijk}\}$  şeklindeki örnek hücre sayıları açısından  $\{\hat{\mu}_{ijk}\}$  şeklindeki uydurulmuş (beklenen) değerlerin bulunması için açık formüllere sahiptir. Böylece tahminler direkt olarak yapılabilir. Örnek olarak, X-Y koşullu bağımsızlığın sağlandığı (XZ, YZ) modeli aşağıdaki ifadeye sahiptir:

$$\hat{\mu}_{ijk} = n_{i+k}n_{+jk} / n_{++k}$$

k tane kısmi tabloda bir hücre için uydurulmuş değer; ilgili kısmi tablodaki satır toplamının ( $n_{i+k}$ ) sütun toplamı ( $n_{+jk}$ ) ile çarpılıp örnek hacmine ( $n_{++k}$ ) bölünmesi ile bulunur. Bu sadece ilgili kısmi tabloya uygulanan bağımsızlık modeli için uydurulmuş değere eşittir. Birçok log lineer model direkt olarak tahminlemeye sahip değildir. Bu durumda ML tahmini, Newton-Raphson Algoritması gibi tekrarlı metotları gerekli kılmaktadır. Uygulamada hangi modellerin direkt tahminlemeye sahip olduğunu bilmek önemli değildir. Direkt tahminlere sahip olmayan modeller için kullanılan yinelemeli metotlar aynı zamanda direkt tahminlere sahip modellere de uygulanmaktadır. Çoğu yazılım paketi tüm durumlar için yinelemeli programları kullanmaktadır (Agresti, 1996).

### 5.3.2. Ki-Kare İyi Uyum Testleri

Üç yönlü bir tabloya göre beklenen frekansların belirli bir log lineer modeli sağlaması şeklinde kurulan bir sıfır hipotezini dikkate alalım. Her zamanki gibi büyük örnek ki-kare istatistikleri ile hücre için uydurulan değerlerle gözlemlenen sayıların karşılaştırılması vasıtasıyla modelin uyum iyiliği değerlendirilir (Agresti, 1996; Acar, 2000). Bu büyük örnek ki-kare istatistikleri Olabilirlik Oran ve Pearson istatistikleridir ve aşağıdaki şekilde hesaplanır:

$$G^2 = 2 \sum n_{ijk} \log \left( \frac{n_{ijk}}{\hat{\mu}_{ijk}} \right) \text{ ve } X^2 = \sum \frac{(n_{ijk} - \hat{\mu}_{ijk})^2}{\hat{\mu}_{ijk}}$$

Serbestlik dereceleri ise hücre sayılarının miktarından modelde gereksiz olmayan parametre sayısının çıkarılması ile elde edilir. Model karmaşık hale geldikçe serbestlik derecesinin değeri azalır.

Doygun modelin serbestlik derecesi sıfırdır. Tablo-5.3' te yer alan araştırmaya göre Tablo-5.6' da birkaç tane log lineer model için uyum testlerinin sonuçları gösterilmektedir. Belirli bir serbestlik derecesine göre, daha büyük  $G^2$  ve  $X^2$  değerleri daha küçük P-değerlerine sahiptir ve daha zayıf uyumlar olduğuna işaret etmektedir.

**Tablo 5.6** Alkol (A), Sigara (C) ve Marihuana (M) Kullanımına İlişkin Oluşturulan Log Lineer Modeller İçin Uyum İyiliği Testleri

Model	G <sup>2</sup>	X <sup>2</sup>	s.d.	P-Değeri*
(A, C, M)	1286.0	1411.4	4	<0.001
(A,CM)	534.2	505.6	3	<0.001
(C, AM)	939.6	824.2	3	<0.001
(AC, AM)	843.8	704.9	3	<0.001
(M,AC)	497.4	443.8	2	<0.001
(AC, CM)	92.0	80.9	2	<0.001
(AM, CM)	187.8	177.6	2	<0.001
(AC, AM, CM)	0.4	0.4	1	0.54
(ACM)	0.0	0.0	0	----

\*G<sup>2</sup> istatistiği için P-değeri

İki yönlü birliktelik terimlerinden herhangi birine sahip olmayan modeller zayıf biçimde uyum göstermektedir ve P-değerleri 0.001 değerinin altındadır. Tüm ikili birlikteliklere olanak tanıyan, fakat üç-faktör etkileşiminin olmadığı varsayımına dayanan (AC, AM, CM) modeli iyi şekilde uyum göstermektedir (P=0.54).

### 5.3.3. Log Lineer Artıkları

Hücre artıkları incelenerek uyum kalitesi kontrol edilir. Bunlar bir modelin neden zayıf bir uyuma sahip olduğunu göstermektedir ve belki de alternatif bir model ortaya çıkarabilirler ya da genel olarak iyi uyumlu bir modelde uyumsuzluk gösteren hücreleri belirlerler (Agresti, 1996). Bir tabloda çok sayıda hücre olduğunda bazı artıklar olasılık nedeniyle büyük olabilir. Daha önce bağımsızlık modeli için ayarlanmış artıklar açıklanmıştı. Ayrıca Poisson GLM' leri için bu artıklar tekrar



tartışılmıştır. Bu artıklar, gözlemlenen ve uydurulan değerler arasındaki farkların standart hataya bölünmesi ile elde edilmektedir. Model kurulduğunda ayarlanmış artıklar yaklaşık olarak normal bir dağılıma sahiptir. Bu nedenle, az sayıda hücre olduğunda ayarlanmış artıkların mutlak değerleri yaklaşık olarak 2' den büyükse ya da bir çok hücre söz konusu iken 3' den büyükse uyumsuzluk konusunda bir gösterge oluşturmaktadır (Agresti, 1996).

Tablo-5.7, Tablo-5.3' te yer alan A-C koşullu bağımsızlığının sağlandığı (AM, CM) modelinin uyumu açısından düzeltilmiş artıkları göstermektedir. Bu model M' nin her bir seviyesinde A ve C arasındaki bağımsızlığı gösterdiğinden dolayı M' in her seviyesine göre her bir kısmi tablo için daha önceki bölümlerde iki yönlü tablolar için ele alınan formül uygulanarak artıklar hesaplanır<sup>52</sup>. Bu model, uyumun test edilmesinde 2 serbestlik derecesine sahiptir ve iki gereksiz olmayan artık, M' in her bir seviyesinde A-C bağımsızlığının kontrol edilmesinde kullanılır. Büyük artıklar genel anlamda zayıf bir uyum olduğuna işaret etmektedir (Gerçekte  $X^2$  istatistiği iki gereksiz olmayan artıkla ilişkilidir ( $X^2 = (3.70)^2 + (12.80)^2 = 177.6$ ). Aşırı derecede büyük artık değerleri hiç marihuana içmeyen öğrenciler için söz konusudur. Bu öğrenciler için A ve C değişkenlerine verilen cevapların her ikisi de “evet” ya da “hayır” olduğunda pozitif artık değerleri elde edilir.

Eğer kullanımları birbirinden bağımsız olsaydı birini kullanmaları beklenirdi. Ancak burada öğrencilerin çoğu ya ikisini de kullanmaktadır ya da hiç birini kullanmamaktadır. Aynı durum marihuana kullanan öğrenciler için de geçerlidir. Fakat gözlemlenen ve uydurulan sayılar arasındaki farklar çarpıcı değildir.

---

$$^{52} \frac{n_{ij} - \hat{\mu}_{ij}}{\sqrt{\hat{\mu}_{ij}(1 - p_{i+})(1 - p_{+j})}}$$

**Tablo 5.7** İki Log Lineer Model İçin Ayarlanmış Artıklar

Uyusturucu Madde Kullanımı			Gözlemlenen Sayılar	Model (AM, CM)		Model (AC, AM, CM)	
A	C	M		Uydurulmuş Sayılar	Ayarlanmış Artıklar	Uydurulmuş Sayılar	Ayarlanmış
Evet	Evet	Evet	911	909.2	3.70	910.4	0.63
		Hayır	538	438.8	12.80	538.6	0.63
	Hayır	Evet	44	45.8	-3.70	44.6	-0.63
		Hayır	456	555.2	12.80	455.4	0.63
Hayır	Evet	Evet	3	4.8	-3.70	3.6	-0.63
		Hayır	43	142.2	-12.80	42.4	0.63
	Hayır	Evet	2	0.2	3.70	1.4	0.63
		Hayır	279	179.8	12.80	279.6	-0.63

Tablo-5.7 aynı zamanda (AC, AM, CM) modeli için de ayarlanmış artıkları göstermektedir. Bu model için serbestlik derecesi 1 olduğundan sadece bir artık gereksiz değildir.  $G^2$  ve  $X^2$  istatistiklerinin her ikisi de küçük olduğundan dolayı bu artıkların iyi bir uyum göstermesi şaşırtıcı olmaz (Aslında s.d=1 olduğunda  $X^2$  istatistiği açıkça her bir ayarlanmış artığın karesidir).

Pearson artığı kullanıldığında da benzer sonuçlar elde edilir. Poisson Log Lineer Modellerine göre bu artıklar aşağıdaki forma sahiptir:

$$e_{ijk} = \frac{(n_{ijk} - \hat{\mu}_{ijk})}{\sqrt{\hat{\mu}_{ijk}}}$$

#### 5.3.4. Kısmi Birliktelik Testleri

İlk olarak farklı log lineer modellerin karşılaştırılması aracılığıyla kısmi birliktelik testleri oluşturulur. Örnek olarak (AC, AM, CM) modeli için, sigara içme durumu ile alkol kullanımı arasında kısmi birliktelik olmadığı ifade edildiği sıfır hipotezi ( $\lambda^{AC}$  terimi sıfıra eşittir) test edilir. Test ile (AC, AM, CM) alternatif modeline karşı A-C koşullu bağımsızlığının ifade edildiği (AM, CM) şeklindeki daha basit bir modelin kabul edilip edilmeyeceği analiz edilmektedir. Bir log lineer terimin sıfıra eşit olduğu hipotezinin olabirlik oran testi basittir. Olabirlik oran istatistiği ( $-2(L_0 - L_1)$ ), ilgili terimin olmadığı model ile ilgili terimin yer aldığı model için ayrı ayrı hesaplanan  $G^2$  uyum iyiliği istatistikleri arasındaki fark ile özdeşdir. Serbestlik derecesi ilgili serbestlik dereceleri arasındaki farka eşittir. Lojistik regresyon modellemesine göre bu yöntem ele alınmıştır. (AC, AM, CM) modelinde  $\lambda^{AC} = 0$  olup olmadığının test edilmesi için kullanılan test istatistiği aşağıdaki şekilde ifade edilen farka eşittir:

$$G^2[(AC, CM)|(AC, AM, CM)] = G^2(AM, CM) - G^2(AC, AM, CM)$$

Buna göre (AC, AM, CM) modeli için bulunan  $G^2 = 0.4$  (s.d=1) ile (AM, CM) modeli için bulunan  $G^2 = 187.8$  (s.d=2) değerleri arasındaki farktır. Serbestlik derecesi  $(2-1)=1$  e bağlı olarak bu fark 187.4 olarak bulunur ( $P < 0.001$ ).

Küçük P-değeri, A-C kısmi birlikteliğinin lehinde ve sıfır hipotezine karşı güçlü bir kanıt sağlar. (AC, AM, CM) modeli ile (AC, CM) ve (AC, AM) modellerinin karşılaştırılmasında bulunan istatistikler de aynı zamanda A-M ve C-M kısmi birlikteliklerine ilişkin güçlü kanıtlar sağlamaktadır. Tablo-5.3 için daha ileri düzeydeki analizlerde herhangi bir basit modelden daha çok (AC, AM, CM) modeli kullanılmalıdır.

Örnek hacmi, herhangi bir yorumlama yönteminin sonuçları üzerinde güçlü bir şekilde etki göstermektedir. Örneklem hacmi arttıkça belirli bir hacmin etkisini

bulmak daha olası hale gelmektedir. Bu durumda uyarıcı bir dikkat gerekmektedir. Küçük örnek hacimleri için gerçeklik, bir uyum iyiliği testini sağlayan en basit model ile belirlenenden daha karmaşık olabilir. Bu dikkat çekici durum hipotezin test edilmesindeki sınırlamaları yansıtmaktadır. Daha çok ilgilenilen konu ise gerçek parametrelerle sıfır hipotezinde ortaya konulan değerler arasındaki farkın önem teşkil edecek kadar büyük olup olmadığıdır. Tahmin, özellikle güven aralıkları, etkilerin öneminin değerlendirilmesinde kullanılan anlamlılık testlerinden daha fazla yarar sağlamaktadır.

### 5.3.5. Odds Oranları İçin Güven Aralıkları

Log lineer parametrelerin ML tahmin edicileri büyük örnek normal dağılımı göstermektedir. Log lineer modeller için kullanılan yazılımlar tahminleri ve bu tahminlerin standart hatalarını bildirmektedir (Stokes, Davis ve Koch. 1991). Modellerde en yüksek dereceli terimler iki faktör birliktelikleridir. Tahminler koşullu odds oranlarına dayanmaktadır. Gerçek log odds oranları için güven aralıklarının oluşturulması amacıyla ilk olarak tahminler ve bu tahminlerin standart hataları kullanılır. Daha sonra bunların üstel değerleri alınarak odds oranları için aralıklar şekillendirilir. Açıklamak amacıyla Tablo-5.3' te yer alan sigara ve alkol kullanımı arasındaki koşullu odds oranını (odds ratio) (AC, AM, CM) modeli varsayımı altında tahmin edelim.

Yazılım ile ( SAS' taki PROC GENMOD gibi) ASE=0.174 ile birlikte  $\hat{\lambda}_{11}^{AC} = 2.054$  değerleri elde edilir (Agresti, 2002) ki burada son satır ve sütunda yer alan gereksiz birliktelik parametreleri sıfıra eşittir. İlgili yaklaşıma göre tek sıfırdan farklı olan terim tahmin edilmiş log odds oranına eşittir. Her bir indeksin seviyelerinde karşılıklı parametre toplamalarının sıfır olması için yazılımda (SAS' ta CATMOD) şu değerler raporlanır:  $\hat{\lambda}_{11}^{AC} = \hat{\lambda}_{22}^{AC} = 0.514$ ,  $\hat{\lambda}_{12}^{AC} = \hat{\lambda}_{21}^{AC} = -0.514$  ve

dört terimin her biri için ASE=0.044 şeklindedir. Bu yaklaşıma göre, tahmin edilen koşullu log odds oranı aşağıdaki şekilde bulunur<sup>53</sup>:

$$\hat{\lambda}_{11}^{AC} + \hat{\lambda}_{22}^{AC} - \hat{\lambda}_{12}^{AC} - \hat{\lambda}_{21}^{AC} = 2.054 = 4\hat{\lambda}_{11}^{AC}$$

Benzer biçimde ilgili koşullu log odds oranı tahmininin ASE değeri  $\hat{\lambda}_{11}^{AC}$  için bulunan ASE değerinin dört katıdır ya da 0.174 (0.044\*4). Bu nedenle parametre sınırlamaları dikkate alınmaksızın tahmin edilen A-C koşullu odds oranı 2.054 olarak bulunur ve ASE değeri de 0.174 olur. Gerçek koşullu log odds oranı için %95 güven aralığı ise  $2.054 \pm 1.96(0.174)$  ya da (1.71, 2.39) şeklindedir. Buradan odds oranı için  $(e^{1.71}, e^{2.39}) = (5.5, 11.0)$  değerleri elde edilir. Buna göre marihuana kullanan ve kullanmayanların hepsi açısından alkol kullanımı ve sigara kullanımı arasında pozitif yönde güçlü bir birliktelik söz konusudur. (AC, AM, CM) modeline göre, A-M koşullu odds oranı için %95 güven aralığı; (8.0, 49.2) ve C-M koşullu odds oranı için %95 güven aralığı (12.5, 23.8) olarak elde edilir. Aralıklar geniştir fakat bu birliktelikler aynı zamanda güçlüdür.

Özet olarak bu log lineer modelin kullanımı ile yapılan tahminler, uyuşturucu maddelerin her bir çifti için güçlü koşullu birliktelikler göstermektedir. Bir uyuşturucu madde kullanıcılarında ikinci bir maddeyi kullanma konusunda güçlü bir eğilim söz konusudur. Ayrıca üçüncü bir uyuşturucu madde kullananların ve kullanmayanların hepsi için bu durum geçerlidir. Bu modele göre tahmin edilen marjinal birliktelikler Tablo-5.5' de kontrol edildiğinde bu değerlerin daha da güçlü olduğu görülmektedir ki her hangi bir madde üzerinde elde edilen sonuçların kontrol altına alınması diğer iki madde arasındaki birlikteliği az çok azaltmaktadır. Bu bölümdeki analizler birliktelik yapısı ile ilişkilidir. Farklı türdeki bir analiz marjinal dağılımların karşılaştırılması ile ilgilidir ki örnek olarak, bir maddenin diğerlerine

---

<sup>53</sup>  $\log \theta_{XY(k)} = \log \left( \frac{\mu_{11k} \mu_{22k}}{\mu_{12k} \mu_{21k}} \right) = \lambda_{11}^{XY} + \lambda_{22}^{XY} - \lambda_{12}^{XY} - \lambda_{21}^{XY}$

göre daha fazla kullanımının söz konusu olup olmadığının belirlenmesi verilebilir. Bu çalışmada bu konuya değinilmeyecektir.

#### 5.4. DAHA YÜKSEK BOYUTLAR İÇİN LOG LİNEER MODELLER

Üç yönlü tablolar için log lineer modellerin kullanımı iki yönlü tablolardan daha karmaşık bir yapıya sahiptir. Bunun nedeni ise potansiyel birliktelik modellerinin çeşitliliğidir (Agresti, 1996). Fakat üç yönlü tablolara uygulanan log lineer modellerde kullanılan temel kavramlar çok yönlü tablolar için kolayca genişletilebilir. İstatistiksel olarak anlamlı bir etki uygulamada önemli değildir. Çok büyük örneklerle birlikte hipotez testlerinden daha çok tahminlemeye odaklanmak çok önemlidir. Daha basit modeller özetleme açısından daha fazla kolaylık sağlamaktadır. Final modelin seçiminde iyi uyum istatistikleri tek başına kullanılmamalıdır.  $\{n_i = np_i\}$  şeklindeki hücre sayıları ve  $\{\hat{\mu}_i = n\hat{\pi}_i\}$  şeklindeki uydurulmuş değerler ile birlikte keyfi bir boyuttaki bir tablo için; ilk olarak *benzersizlik indeksi* (dissimilarity index) aracılığı ile bir modelin örnek verilerine yakınlığı özetlenir(Agresti, 1996). Benzersizlik indeksi aşağıdaki şekildedir:

$$D = \sum |n_i - \hat{\mu}_i| / 2n = \sum |p_i - \hat{\pi}_i| / 2$$

Bu indeksin alabileceği değerler 0 ile 1 arasında değerler almaktadır ve daha küçük değerler daha iyi bir uyumu temsil etmektedir. Örnek durumlarının oranını temsil etmektedir ki mükemmel bir uyuma ulaşmak amacıyla modele göre sırasıyla farklı hücreler için hesaplanır. D, benzersizlik indeksi ile  $\Delta$  şeklinde bir uyum indeksi tahmin edilir ki örneklem alınan popülasyona uygun olmayan modeli belirler. Model mükemmel olarak oluşturulduğunda  $\Delta$ ' nın değeri 0 olur. Örneklem varyasyonundan dolayı D ile  $\Delta$  değeri aşırı tahminlenir (küçük örnekler için bu önemlidir). Uygulamada  $\Delta > 0$  olduğunda modeller mükemmel şekilde kabul edilir(Agresti, 1996). Yeterince büyük n için model kabul edilmediğinde,  $G^2$  ve  $X^2$  şeklindeki uyum iyiliği istatistikleri büyük olacaktır ve uyumsuzluğa işaret edecektir. D tahmin edicisi ise bu istatistikler tarafından ileri sürülen uyumsuzluğun uygulama açısından önemli olup olmadığını gösterir.

## 5.5. LOG LINEER – LOJİT İLİŞKİSİ

Olumsuzluk tabloları için kullanılan log lineer modeller yanıt ve açıklayıcı değişkenler arasında ayırım yapmaz. Esas olarak, bu modeller tüm değişkenleri yanıt değişkenler olarak işleme sokar. Diğer yandan lojit modeller iki sonuçlu bir yanıtın bir açıklayıcı değişken kümesine nasıl bağlı olduğunu tanımlamaktadır. Model türleri farklı olsa bile bu modeller arasında güçlü ilişkiler söz konusudur. Bir log lineer modele göre modeli yorumlama konusunda yardım etmesi amacıyla bir yanıt değişken için lojitler oluşturulur. Ayrıca kategorik açıklayıcı değişkenler için oluşturulan lojit modeller log lineer modeller ile eşittir (Agresti, 1996).

### 5.5.1. Log Linear Modelleri Yorumlamak İçin Lojit Modellerin Kullanımı

Bir log lineer formülünü anlamak için, bir değişkeni yanıt diğerlerini açıklayıcı değişken olarak işleme tabi tutan bir lojit oluşturmak yardımcı olur. Üç yönlü tablolarda homojen birlikteliğe sahip log lineer model ile durumu açıklayalım. Homojen birlikteliğin log lineer modeli aşağıdaki şekildedir:

$$\log \mu_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ}$$

Varsayalım ki Y iki sonuçlu bir değişkendir ve X ile Z değişkenleri açıklayıcı değişken ve Y değişkeni de yanıt değişkeni olarak işleme sokulsun.  $\pi$ ,  $Y=1$  şeklindeki olasılığı gösterir ki X ve Z değişkenlerinin seviyelerine bağlı bir değerdir. Y için lojit aşağıdaki biçimdedir (Agresti, 1996):

$$\begin{aligned}
\text{logit}(\pi) &= \log\left(\frac{\pi}{1-\pi}\right) = \log\left(\frac{P(Y=1 | X=i, Z=k)}{P(Y=2 | X=i, Z=k)}\right) \\
&= \log\left(\frac{\mu_{i1k}}{\mu_{i2k}}\right) = \log(\mu_{i1k}) - \log(\mu_{i2k}) \\
&= (\lambda + \lambda_i^X + \lambda_i^Y + \lambda_k^Z + \lambda_{i1}^{XY} + \lambda_{ik}^{XZ} + \lambda_{1k}^{YZ}) \\
&\quad - (\lambda + \lambda_i^X + \lambda_2^Y + \lambda_k^Z + \lambda_{i2}^{XY} + \lambda_{ik}^{XZ} + \lambda_{2k}^{YZ}) \\
&= (\lambda_1^Y - \lambda_2^Y) + (\lambda_{i1}^{XY} - \lambda_{i2}^{XY}) + (\lambda_{1k}^{YZ} - \lambda_{2k}^{YZ}).
\end{aligned}$$

İlk parantezdeki terim bir sabittir ki i ya da k değerlerine bağlı değildir. İkinci parantez içi terim X' in i tane seviyesine bağlıdır. Üçüncü parantez içi terim ise Z' nin k tane seviyesine bağlıdır. Lojit aşağıdaki şekilde toplanabilir bir şekle sahiptir:

$$\text{logit}(\pi) = \alpha + \beta_i^X + \beta_k^Z \quad (5.5.2)$$

Daha önce ele alınan bu model (lojit) toplanabilirlik özelliğine göre Z ve X' in seviyelerine bağlıdır. X değişkeninin lojit üzerindeki etkisi Z değişkeninin her bir seviyesinde aynıdır. Z değişkeninin etkisi de ise X değişkeninin her bir seviyesinde aynıdır. Lojit ölçekteki toplanabilirlik, kategorik değişkenler arasında hiçbir etkileşimin olmadığı yani etkileşimsizliğin standart tanımıdır (Agresti, 1996). Y değişkeni iki sonuçlu olduğunda homojen birlikteliğin log lineer modeli ( üç faktörlü etkileşimin olmadığı) bu lojit modele eşittir. X değişkeni de iki sonuçlu olduğunda yukarıda ele alınan lojit model ve (XY, XZ, YZ) şeklindeki log lineer model; Z değişkeninin K tane seviyesinin her birinde X ve Y arasında eşit odds oranları aracılığı ile nitelendirilir. Bu daha önce ele alınan Breslow- Day testinin kullanılması için gerekli koşuldur. Bu modeller için  $G^2$  veya  $X^2$  uyum istatistikleri bir ortak odds oranına göre test yapmak için alternatif yöntemler sağlamaktadır. Örnek hacmi, K' ya göre yeteri kadar büyük olduğunda bu istatistikler s.d=K-1 serbestlik derecesi ile yaklaşık ki kare dağılımlarına sahiptir.



### 5.5.2. Log Lineer Modeller ve Lojit Modeller Arasındaki Benzerlik

(XY, XZ, YZ) şeklindeki log lineer modelden türetilen (6.5.2) şeklindeki lojit modele dönelim. (XY, XZ, YZ) modelindeki  $\lambda_{ik}^{XZ}$  terimi lojit oluşturulduğu zaman iptal edilir. Bu terimin ihmal edildiği (XY, YZ) şeklindeki model gibi görülebilir ki aynı zamanda ilgili lojit modele eşittir. Gerçekte (XY, YZ) şeklindeki log lineer modele göre Y üzerinde lojit oluşturulursa, aynı forma sahip bir lojit model elde edilir. Fakat lojit model ile aynı uyuma sahip log lineer model, açıklayıcı değişkenler arasındaki ilişki açısından sadece genel bir etkileşim terimini kapsamaktadır. Lojit model açıklayıcı değişkenler arasındaki ilişkileri tanımlayamaz ve bu nedenle de bunlara ilişkin keyfi bir etkileşim biçimine fırsat verir (Agresti, 1996). Tablo-5.11, Y değişkeni iki sonuçlu bir yanıt değişkeni olduğunda üç yönlü tablolar için eşdeğer lojit ve log lineer modelleri özetlemektedir. (Y, XZ) şeklindeki basit log lineer modelde, Y değişkeni X ve Z değişkenlerinin her ikisinden de ortaklaşa biçimde bağımsızdır. Bu model,  $\{\beta_i^X\}$  ve  $\{\beta_k^Z\}$  terimlerinin sıfıra eşit olduğu, (5.5.2) şeklindeki lojit modelin özel bir durumu ile eşdeğerdir.

Tablo-5.11’ de yer alan karşılıklı modellerin her birinde log lineer model, lojit modellerdeki açıklayıcı değişkenlere ilişkin X-Z birliktelik terimini de kapsamaktadır. (XYZ) şeklindeki doymuş log lineer model üç faktörlü etkileşim terimini kapsamaktadır. Y iki sonuçlu bir yanıt olduğunda bu model, X ve Z tahmin edicileri arasında bir etkileşime sahip lojit model ile eşdeğerdir. Örnek olarak, X değişkeninin Y değişkeni üzerindeki etkisi Z değişkeninin seviyesine bağlıdır ki bu durum, X-Y odds oranının Z değişkeninin seviyeleri karşısında değişim gösterdiği anlamına gelmektedir. İlgili lojit model de aynı zamanda doymuş modeldir. Daha yüksek boyuttaki tablolar içinde paralel benzerlikler elde edilir. Bir lojit modelde yanıt değişkeni ile her bir açıklayıcı değişken arasındaki birliktelikler ile açıklayıcı değişkenler arasındaki etkileşim teriminin hepsini kapsayan log lineer model ile benzeşebilir.

### 5.5.3. Model Seçimindeki Stratejiler

Bu bölüm, bir log lineer modelin seçimi için bazı stratejilerin tanımlanmasıyla son bulmaktadır. Çoğu uygulamada potansiyel olarak yararlı ve kullanışlı modeller tüm log lineer modellerin küçük bir alt kümesini oluşturmaktadır. Örnek olarak, birçok incelemede açıklayıcı ve yanıt değişken ayrımı yapılmaktadır. Bu nedenle, açıklayıcı değişkenler arasındaki ilişkilerin modellenmesinden daha çok açıklayıcı değişkenlerin yanıt değişkenler üzerindeki etkilerinin modellenmesine ve yanıt değişkenler arasındaki birlikteliklerin modellenmesine odaklanılır. Tek bir yanıt değişken söz konusu olduğunda bu yanıt değişken iki sonuçlu ise ilgili log lineer modeller yanıtı göre oluşturulan lojit modellere uymaktadır. Yanıt değişken iki kategoriden fazla kategoriye sahip olduğunda ilgili log lineer modeller çok kategorili lojit modellere uymaktadır. Elbette bu şekildeki durumlarda ilk olarak log lineer modellerle probleme yaklaşmaktan çok doğrudan doğruya lojit modeller uydurulur.

**Tablo-5.8** İki Sonuçlu Y Yanıt Değişkeni İle Üç Yönlü Bir Tablo İçin Eşdeğer Log Lineer ve Lojit Modeller

Loglinear	Logit
$(Y, XZ)$	$\alpha$
$(XY, XZ)$	$\alpha + \beta_i^X$
$(YZ, XZ)$	$\alpha + \beta_k^Z$
$(XY, YZ, XZ)$	$\alpha + \beta_i^X + \beta_k^Z$
$(XYZ)$	$\alpha + \beta_i^X + \beta_k^Z + \beta_{ik}^{XZ}$

Log lineer yaklaşımın bir avantajı genelleştirilebilir olmasıdır. Birden fazla yanıt değişken var olduğunda bu genelleştirmeler uygulanır. Örnek olarak uyuşturucu madde kullanımına ilişkin örnekte üç yanıt değişken arasındaki birliktelik biçimlerinin incelenmesi için log lineer modeller kullanılmıştır. Özet olarak, en

azından iki deęişken yanıt deęişkeni olduęunda log lineer modellerin kullanılması en doęal olanıdır. Sadece bir deęişken yanıt deęişkeni olduęunda doęrudan doęruya lojit modellerin kullanılması daha mantıklıdır (Agresti, 1996). Her iki model türü için seçim süreci, deęişken sayısı arttıkça daha zor bir hale gelmektedir. Çünkü olası birlikteliklerde ve etkileşimlerde artma meydana gelmektedir. Tüm olası modellerin uydurulması genellikle pratik bir durum deęildir. İlk olarak iki karşıt hedef arasında denge kurulmalıdır: Model iyi bir uyum sağlayacak kadar kompleks olmalıdır, fakat yorumlama için de basit olmalıdır (yani verilere aşırı uyum göstermek yerine verilerin düzeltilmesi).

Sıklıkla bir araştırma ile ortaya atılan ana sorular, bir modelde yer alması gereken potansiyel terimlerin seçimine yön vermek konusunda yardımcı olmaktadır ve sınırlı sayıda modeller ortaya koyulabilmektedir. Örnek olarak, birlikteliklere ilişkin bir araştırmadaki hipotez, belirli modellerin uyumlarının karşılaştırılması ile test edilebilmektedir.

Yapısal araştırma şeklindeki çalışmalar için potansiyel modeller arasında bir tarama yapılması ile bağımlılık yapısının önemli yön ve görünümüleri konusunda bilgi sağlanabilir ve ayrıca keşfedici analizler ile deęişkenler arasındaki birliktelik bağlantılarının karmaşıklığı çözümlenmeye çalışılır (Agresti, 1996). İlk olarak ele alınan yaklaşımda tek faktörlü terimlere sahip log lineer model kurulur. Daha sonra sadece iki faktörlü ve tek faktörlü terimleri içeren log lineer model, sonra üç faktörlü ve üç faktörden düşük faktör derecesine sahip terimleri içeren modeller sırasıyla incelenir (faktör sayısına göre incelenecek model sayısı da deęişmektedir). Bu şekildeki modellerin uygunluęuna bakıldığında genellikle sınırlı sayıda iyi uyumlu model elde edilir. Aynı zamanda bu amaçla deęişken eleme yöntemleri de kullanılır. **Geriye doęru eleme yöntemi**<sup>54</sup> doęgun model gibi karışık bir modelle başlar ve her defasında bir terim atılır. Atılacak terimler çalışan modelin uyumunu en az bozan terime göre belirlenir. Stratejiye bakılmaksızın örnekleme tasarımı ile ya da

---

<sup>54</sup> Önce tüm deęişkenler modelde bulunur, en düşük bölümsel korelasyonu gösteren deęişken modelden çıkarılarak yeniden inceleme yapılır. Eğer yeni modelin kestirimi eskiye göre anlamlı olarak arttı ise sırası gelen güçsüz deęişken elenerek işlem sürdürülür.

açıklayıcı-yanıt deęişken ayırımı yapılarak belirli marjinal toplamlar sabitlenir. Model bu toplamların uydurulmuş toplamlar ile eş deęer olmasını saęlayan terimleri içermelidir. Bu noktada modellerde yanıt ve açıklayıcı deęişkenler arasında bir ayırım yapılması gerekmektedir. Modelleme sürecinde açıklayıcı deęişkenler ile yanıt deęişkenleri baęlayan terimlere ve yanıt deęişkenleri birbirine baęlayan terimlere odaklanılmalıdır. Boyut sayısı arttıkça faktörlerin karmaşıklığı ortaya çıkmaktadır. Örneklem hacmi çok büyük olmadıkça birçok sayının sıfır olması söz konusu olmaktadır. Bu durum sınırsız parametre tahminlerine ve uyum iyilięi istatistiklerine göre zayıf ki-kare yaklaşımlarına neden olmaktadır. Ek olarak bazı deęişkenler ordinal nitelikte olabilmektedir ve modeller bunu yansıtmalıdır(Agresti, 1996).

## ALTINCI BÖLÜM

### KATEGORİK VERİ ANALİZİ KONUSUNDA YIRMİNCİ YÜZYILDAKİ GELİŞMELER

Kategorik veri analizi (CDA) için kullanılan metotların evrimine ilişkin tarihsel bir inceleme yapıldığında konuya ilişkin esasların genellikle tartışmalar içerisinde gömülü kaldığı görülmektedir. İstatistik biliminin gelişiminde önemli biçimde çığır açan katkılar sağlanmasına rağmen bu katkıları sağlayan istatistikçiler arasında genellikle hararetli biçimde anlaşmazlıklar söz konusu olmuştur.

#### 6.1. PEARSON-YULE BİRLİKTELİK TARTIŞMASI

Kategorik veri analizinde kullanılan metotlardaki ilk gelişmelerin çoğu yirminci yüzyılın başında İngiltere’ de gerçekleşmiştir. 1900’ lü yılların başında kategorik veri analizi konusunda literatürün çoğunu *birlikteliklerin tanımlanmasında kullanılan özet göstergeler* konusundaki tartışmalar oluşturmuştur. 1900 yılında Karl Pearson ki-kare istatistiğini ( $X^2$ ) ve G istatistiğini ortaya koymuştur. Udny Yule’ da odds oranını ve birlikteliğe ilişkin ölçüleri göstermiştir. 1900 yılına kadar Karl Pearson (1857–1936) zaten istatistik çevresi tarafında oldukça iyi tanınmaktaydı. Bundan önceki on yıl boyunca çalışmaları; çarpık olasılık dağılımlarının bir ailesinin geliştirilmesi (Pearson Eğrileri olarak adlandırılmaktadır), korelasyon katsayısının çarpım momenti tahminin ve bu tahminin standart hatasının elde edilmesi ve doğrusal regresyon üzerinde Francis Galton tarafından yapılan çalışmanın ilerletilmesi konularını kapsamaktadır. Pearson’ un ki-kare testini geliştirirken hareket noktası Monte Carlo’ da bir rulet masasında oluşan sonuçların rasgele olup olmadığını test etmek olmuştur. Pearson eğrilerini oluşturan çeşitli veri setlerine uygunluğunu kontrol ederek iki yönlü olumsuzluk tablolarında istatistiksel bağımsızlığı test etmiştir.

Pearson’ un yaklaşımında, çapraz sınıflandırılmış tabloların temelini oluşturanın sürekli nitelikte iki değişkenli dağılımların olduğu varsayılmıştır. Pearson yaklaşık bir ölçü aracılığı ile (korelasyon gibi) birlikteliğin tanımlanabileceğini öne

sürmüştür. 1904 yılında Pearson, bağımsızlık olasılığından, sınıfın toplam sapmasının bir ölçüsü olarak *kontenjans* (*olumsallık*) terimini ortaya koymuş ve bunu tanımlamak amacıyla çeşitli ölçüler oluşturmuştur. Dört düzeyli korelasyon,  $2 \times 2$  boyutlu tablolarda yer alan sayılar için normal dağılım varsayımı yapılarak kullanılan maksimum olabilirlik (ML) tahminidir (korelasyon için). Ortalama kare kontenjansı ile kontenjans katsayısı<sup>55</sup> ise  $X^2$  istatistiğinin (0,1) ölçeğine göre normalleştirilmeleridir.

George Udny Yule (1871–1951) (Pearson ile aynı döneme ait bir İngiliz) alternatif bir yaklaşım ortaya koymuştur ve çoklu regresyon modelleri ile ilgili çoklu ve kısmi korelasyon katsayılarının geliştirilmesine öncülük etmiştir. Yule 1900–1912 yılları arasında, kontenjans tablolarındaki birlikteliklerin incelenmesine yönelmiştir. Kategorik değişkenlerin çoğunun yapısal olarak kesikli olduğuna inanan Yule hücre sayılarına ilişkin göstergeleri sürekliliği temel almaksızın tanımlamıştır. Yule’ a ait ölçüler, odds oranını ( $\theta$ ) ve odds oranının  $[-1,+1]$  ölçeğine dönüştürülmesini ( $Q = [\theta - 1]/[\theta + 1]$  şeklinde olup günümüzde Yule-Q olarak adlandırılmaktadır.) kapsamaktadır. Pearson’ un ölçülerinde normallik varsayımı söz konusudur. Yule ise; en iyi normal katsayının ancak varsayımsal değişkenler arasında hipotetik (kuramsal) bir korelasyon şeklinde elde edilebileceğini öne sürmüştür (Agresti, 1996). Yule aynı zamanda olumsuzluk tablolarında marjinal ve koşullu birliktelikler arasındaki potansiyel tutarsızlıkları (farklılıkları) göstermiştir. Daha sonra bu konu 1951 yılında E. H. Simpson tarafından ele alınmıştır ve Simpson paradoksu olarak literatüre geçmiştir.

Yirminci yüzyılın ilk çeyreğinde Karl Pearson, İngiltere’ de istatistik biliminde lider konumdaydı. Pearson, Yule’ un fikirlerine negatif biçimde tepki gösterdi. Özellikle Yule’ un elde etmiş olduğu katsayının elverişsiz olduğunu ileri

---

<sup>55</sup> Pearson Kontenjans katsayısı şu şekilde hesaplanır:  
$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}}$$
: Asla negatif bir değer alamaz. 0-1 değerleri arasında değişir.  
Maksimum değeri “N” şeklindeki tablo hacmine bağlıdır (Janes,2001).

sürmüştür. Örnek olarak,  $I \times J$  boyutlu tablolarda yer alan farklı  $2 \times 2$  tablolarında oldukça farklı ölçüm değerleri elde edildiğinden bu katsayının değerlerinin kararsız özellik gösterdiğini ileri sürmüştür.

1913 yılında Pearson ve D. Heron, Yule' un eleştirisine sert bir cevap olarak Biometrika alanında 150 sayfalık bir dergi hazırlamışlar ve Yule' un “iyi oluşmamış kavramları” ile “yanıltıcı sonuçlarına” karşı gelmiştir. Yule' un en çok bilinen “İstatistik Teorisine Giriş” adlı kitabına da aynı zamanda eleştiride bulunmuşlardır. Geçmişe bakıldığında Pearson ve Yule geçerli ve doğru birçok noktaya sahiptir. Çoğu nominal değişkenlerden oluşan bazı sınıflamalar sürekli dağılıma uymamaktadır. Diğer yandan birçok uygulama doğal olarak bir süreklilik temeline bağlıdır ki bu durum doğrudan model kurulmasında ve yorumlanmasında kullanışlı olmaktadır. Ordinal modeller Yule ve Pearson arasında bir bakıma uzlaşma sağlamaktadır. Çünkü Yule tarafından kullanılan odds oranı, yaklaşık olarak normallik gösteren dağılımlar söz konusu olduğunda da iyi oluşturulmuş modelleri tanımlayabilmektedir (Agresti, 1996).

## 6.2. R.A. FISHER' IN KATKILARI

Pearson' un Yule ile uyuşmazlığı daha sonra Ronald A. Fisher (1890–1962) ile yaşadıklarına kıyasla önemsiz kalmıştır (Stigler, 1986). 1922 yılında Fisher geometrik bir gösterim kullanarak “serbestlik derecesi” kavramını ortaya koymuştur. Serbestlik derecesi (s.d) göstergesi ki-kare dağılımlarının oluşturduğu dağılım ailesini tanımlamaktadır. Fisher, serbestlik derecesinin  $I \times J$  boyutlu tablolarda bağımsızlıkların test edilmesi için kullanılabileceğini ileri sürmüştür ki  $X^2$  istatistiğinin serbestlik derecesi  $s.d=(I-1)(J-1)$  şeklindedir.

Bunun aksine 1900 yılında Pearson, kendi istatistiğinin kullanıldığı herhangi bir uygulamada serbestlik derecesinin toplam hücre sayısından 1 çıkarılarak elde edildiğini ya da iki yönlü tablolarda  $(IJ-1)$  şeklinde olduğunu iddia etmiştir. Fakat Fisher; hücre olasılıklarına ilişkin hipotezlerin tahminlenmesinde, tahmin edilmiş sıra ve sütun olasılıklarının kullanılmasının uydurulmuş değerler üzerinde  $(I-1)+(J-1)$  kadar ilave sınırlamalar meydana geleceğini ve bu nedenle  $X^2$  dağılımının

etkileneceğini belirtmiştir (Stigler, 1986). Pearson' un Fisher' in önerdiği bu serbestlik derecesi formülüne tepki göstermesi şaşırtıcı olmamıştır. Pearson, bilinmeyen olasılıkların tahminlenmesinde satır ve sütun örnek oranlarının kullanılmasının büyük örnek dağılımları üzerinde önemsenmeyen bir etkiye sahip olduğunu ileri sürmüştür. Fisher İstatistik Birliği tarafından hatalı olduğunu gösteren yayıma itiraz edememiş ve üyelikten sonunda ayrılmıştır. Ancak çok geçmeden Fisher' in haklı olduğu fark edilmiş ve Pearson ile diğer tartışmaları ve bu tartışma daha da şiddetlenmiştir. Fisher' in bugün istatistikçiler arasındaki ünü, onun teorik çalışmalarından (ML tahmincilerinin optimal özellikleri, yeterliliği şeklindeki kavramları tanıtmıştır) ve varyans analizi ile deney tasarımı gibi alanlara yaptığı metodolojik katkılardan meydana gelmektedir (Stigler, 1986). Kategorik veri analizi konusunda çok iyi bilinen bir çalışması olmamasına rağmen bu konuda dikkate değer katkıları olmuştur ve uygulamalı çalışmalarında metotların kullanımını iyi bir şekilde gerçekleştirmiştir (Agresti, 1996).

Fisher, laboratuvar çalışması için büyük örneklem istatistik metotlarının sınırlarını ortaya koymuştur ve özelleştirilmiş küçük örneklem yöntemlerini savunarak ön plana çıkarmıştır (Agresti, 2002). “t-dağılımı” konusunda W.S. Gosset (“Student” takma isimli) tarafından yapılan çalışmanın ilerlemesinde de katkıları olmuştur (Stigler, 1986). Ayrıca klasikleşen kitabının (Edinburgh: Oliver ve Boyd (1934)) beşinci basımında  $2 \times 2$  boyutlu tablolar için Fisher- Kesin (Exact) Olasılık testini tanıtmıştır.

Son olarak 1930' lu yılların ortalarında CDA için model kurulumuna ilişkin bazı çalışmalar görülmektedir. Örnek olarak, Chester Bliss, iki sonuçlu yanıtlara sahip zehir bilim alanındaki uygulamalar için probit modelin kullanımını popüler hale getirmiştir (Stigler, 1986). Bliss' in 1935 yılındaki makalelerinden birinin ekinde probit modellerde parametrelerin ML tahminlerinin elde edilmesine ilişkin Fisher tarafından ortaya konan bir algoritma yer almaktadır. İlgili algoritma Newton-Raphson türünde bir metot olup bugün genel olarak Fisher Puanlaması (Fisher's Scoring) olarak adlandırılmaktadır (Agresti, 2002). Olumsuzluk tablolarında homojen birlikteliğe ilişkin tanım (üç faktörlü etkileşimin olmadığı,  $2 \times 2 \times 2$  boyutlu tablolar için) 1935 yılında bir İngiliz İstatistikçisi olan Maurice Bartlett tarafından yapılmıştır



ve Bartlett üçüncü değişkenin her seviyesinde diğer iki değişken arasındaki odds oranlarının eşitliğini sağlayan hücre olasılıklarına ilişkin ML tahminlerinin nasıl hesaplanacağını göstermiş, bu fikir için Fisher' a atıfta bulunmuştur (Stigler, 1986).

1940 yılında Fisher, olumsuzluk tabloları için “Kanonik Korelasyon Metotları” konusunu geliştirmiş ve korelasyonu maksimize etmek amacıyla bir olumsuzluk tablosunun satır ve sütunlarına skorların nasıl atanacağını göstermiştir ve bu çalışma özellikle Fransa’ da Uygunluk Analizi (Correspondence Analysis) metotlarının gelişimine de katkı sağlamış; buna bağlı olarak modern istatistik bilimindeki uygulamalarda çok büyük etkilere sahip olmuştur ve Fienberg tarafından CDA konusuna yaptığı katkılar özetlenmiştir (Agresti, 2002).

Paik (1985), daire diyagramlarının üç yönlü tabloların özetlenmesi amacıyla kullanılmasını önermiş, Friendly (2000) kategorik verilerin grafiksel gösterimini ele almış ve Simpson (1951) yılında Simpson Paradoksunu geliştirmiştir (Agresti, 2002).

### **6.3. LOJİSTİK REGRESYON**

Bartlett (1937),  $y$  gözlem değerlerinin dönüşümü için (sürekli oranlara) ANOVA ve Regresyon’ da  $\log\left[\frac{y}{1-y}\right]$  ifadesini kullanmıştır (Agresti, 2002). 1938 yılında basılan, istatistik tablolarından oluşan bir kitapta R.A. Fisher ve Frank Yates, iki sonuçlu verilerin (binary data) analiz edilmesi için bir binom parametresinin olası birkaç dönüşümünden biri olarak  $\log\left[\frac{\pi}{1-\pi}\right]$  ifadesinin kullanılmasını önermişlerdir (Stigler, 1986). 1944 yılında, hem doktor hem de istatistikçi olan Joseph Berkson bu dönüşüm için “lojit” ifadesini tanımlamış ve lojistik regresyon modelinin probit model ile benzer şekilde kurulduğunu göstermiştir (Agresti, 1996).

1951 yılında medikal alt yapıya sahip başka bir istatistikçi olan Jerome Cornfield, olgu-kontrol çalışmalarında göreceli risk (relative risk) oranlarının hesaplanmasında odds oranını kullanmıştır (Agresti, 2002). Kalitatif tahmin

edicilerle kurulan modellerde lojiti ilk olarak 1952 yılında Dyke ve Patterson kullanmıştır (Agresti, 1996). Ayrıca David R. Cox (Halen Oxford Üniversitesinde bulunmaktadır) lojistik regresyonun tanıtılmasında ve ön plana çıkmasında dikkate değer bir etkiye sahip olmuştur. Hem 1958 yılında yayımladığı makalesinde hem de 1970 yılında basılan kitabında (İki Sonuçlu Verilerin Analizi) bunu görmekteyiz (Agresti, 2002). Aynı dönemlerde Danimarkalı istatistikçi ve matematikçi Georg Rasch tarafından yayımlanan ve madde-yanıt modellerine ilişkin büyük bir yazın oluşmasını teşvik eden makalesi ile de lojistik regresyon önem kazanmıştır. En önemli katkısı, denek ve madde parametreleri ile kurulan lojit model olup şu an da “Rasch Model”<sup>56</sup> olarak bilinmektedir.

Bu çalışma Kuzey Avrupa da Psikometri Birlikleri arasında büyük bir etkiye sahip olmuştur (Stigler, 1986). 1970 yılından önce *çok kategorili yanıtlar* için lojistik regresyonun genişletilmesi konusu ara sıra dikkate alınmıştır (Örnek olarak Mantel 1966). Fakat daha sonra önemli çalışmalar yapılmıştır. *Nominal yanıtlar* açısından ilk çalışmalar başlıca ekonometri yazınında görülmektedir: Bock (1970), McFadden (1974), Nerlove ve Press (1973) ve Theil (1969,1970) şeklinde çalışmalar örnek olarak verilebilir (Stigler, 1986).

Daniel Mcfadden, kesikli seçim (discrete-choice) modeli konusunda 1970 ve 1980 yıllarında yaptığı çalışmalar ile 2000 yılında Ekonomi alanında Nobel Ödülü kazanmıştır (Agresti, 2002). *Ordinal yanıtlar* için kümülatif lojit modellerin uygulanması konusunda da ilk olarak şu çalışmalar görülmektedir (Agresti, 1996): Bock ve Jones (1968), Simon (1974), Snell (1964), Waker ve Duncan (1967) ve Williams ve Grizzle (1972). *Normal yanıtlara* bağlı olan kümülatif probit konusu daha eski bir geçmişe sahip olup örnek olarak Aitchison ve Silvey (1957) ile Bock ve

---

<sup>56</sup> Rasch analizi herhangi bir kişinin herhangi bir maddedeki bir kategoriye seçme olasılığının kişi yetenek düzeyi ile madde zorluk düzeyi arasındaki farkın bir lojistik fonksiyonu olduğunu varsayar. Sağlık alanında kullanılan anket ve ölçeklerin birçoğu sıralı ölçeğe sahiptir. Bu nedenle, maddelere verilen doğru cevapların toplanmasıyla elde edilen ham puanları kullanarak anket ya da ölçek değerlendirilmeye çalışıldığı zaman birtakım sorunlarla karşılaşılır. Rasch analizi bu sorunların üstesinden gelmek için kullanılan yöntemlerden biridir. Özellikle fiziksel tıp ve rehabilitasyon dalında sıralı ölçekli ölçümleri aralık ölçekli hale dönüştürmek ve sonuçta sağlık bakımının etkinliğini belirlemek amacıyla oldukça yaygın bir şekilde kullanılmaktadır. (Elhan ve Atakurt,, 2005)

Jones (1968) çalışmaları gösterilmektedir. McCullagh (1980) tarafından yapılan ve tüm kümülatif bağıntı modellerinin ML (maximum likelihood) değerlerinin uydurulmasında bir Fisher skorlama algoritması sağlayan çalışma ile kümülatif lojit ve probit modeller daha fazla dikkat çekmiştir (Agresti, 2002). Lojistik regresyona ilişkin daha sonraki dönemlerde meydana gelen en önemli ilerlemeler; olgu-kontrol (case-control) çalışmalarına uygulanması (Örnek olarak Breslow 1996; Mantel 1973) ve çok sayıdaki sıkıcı parametreye dayanan çalışmalarla olgu kontrol çalışmalarında modelin uydurulmasında koşullu ML yaklaşımının kullanılması (Örnek olarak Breslow ve arkadaşları 1978–1980) şeklinde gerçekleşmiştir (Agresti, 2002).

Nathan Mantel, kategorik veri analizine ilginç katkılarda bulunmuştur ki *odds oranı* ile 1959 yılında ortaya koyduğu *Mantel-Haenszel testi* en bilinenleridir (Agresti,1996)).

Aynı zamanda şu konularda da çalışmalar yaparak katkı sağlamıştır: Eğilim (Trend) testleri (1963), çok terimli lojit ve log-lineer modelleme (1966), olgu-kontrol çalışmalarında elde edilen verilere lojistik regresyonun uygulanması (1973), sabit marjnlere sahip olumsuzluk (kontenjans) tablolarının analizi (Gail ve Mantel 1977), kare biçimindeki olumsuzluk tablolarının analizi (Mantel ve Byar 1978) ve minimum ki-kare ile Wald testlerine ilişkin problemler (Stigler, 1986)). Daha yakın tarihlerde ise kümelenmiş (gruplandırılmış) verilerde ilişkili (korelasyonlu) yanıtlara lojistik modellerin uygulanmasına yoğunlaşmıştır. Bu durum ilk olarak boyutsal verilerin marjinal modellenmesi şeklinde ortaya çıkmıştır (Agresti, 2002). Bu yazının birçoğunda genelleştirilmiş tahmin denklemleri (GTD) şeklindeki yarı olabilirlik metotlarına odaklanılmıştır. Bir diğer hareket noktası ise Genelleştirilmiş Doğrusal Karma Modellerdir ve örnek olarak Breslow ve Clayton, 1993 verilebilir(Agresti, 2002).

Belki geçen yarım yüzyılın en geniş kapsamlı ve etkili katkısı, İngiliz istatistikçileri olan John Nelder ve R. W. M. Wedderburn tarafından 1972 yılında “Genelleştirilmiş Lineer Modeller (GLM)” kavramının ortaya konması ile sağlanmıştır (Agresti, 1996). Bu kavram üç ayrı veri tipi için kullanılan modelleri

bütünleştirmiştir (benzerleştirmiştir): *Binom verileri* için kullanılan lojistik ve probit regresyon modelleri, *poisson verileri* için kullanılan logaritmik-doğrusal modeller ile *normal yanıt verileri* için kullanılan klasik regresyon ve ANOVA modelleri.

Genelleştirilmiş doğrusal modellerin uydurulması için kullanılan algoritma; R.A. Fisher tarafından 1935 yılında probit modellerde ML uygunluğu için ortaya konulan Fisher puanlamasıdır (Fisher scoring). McCulloch (2000) probit modellerden genelleştirilmiş modellere kadar hatta bu modellere ilişkin yarı-olabilirlik gibi daha sonradan eklenen genelleştirmelerde dahil bir inceleme yapmıştır (Agresti, 2002).

#### **6.4. ÇOK YÖNLÜ OLUMSALLIK TABLOLARI VE LOG LİNEER MODELLER**

İkinci dünya savaşının sonunu takiben çeyrek yüzyılda olumsuzluk tabloları için kullanılan modeller konusunda teorik bir temelin geliştiği görülmektedir. H.Cramer (1946) parametre tahmin edicilerinin büyük örneklem dağılımları için genel ifadeler sağlamış ve C.R. Rao (1957,1963) bu konuya ilişkin bir çalışma yürütmüştür (Stigler, 1986). 1949 yılında istatistikçi Jerzy Neyman (E.S. Pearson ile birlikte aralık tahmini ve hipotez testleri metotları üzerine temel bir çalışma oluşturmuştur) “En İyi Asimptotik Normal Tahmin Edicilerin (EİANTE)” bir ailesini tanımlamıştır (Stigler, 1986). Bu EİANTE ailesi; gözlemlenen oranlar ile model aracılığı ile tahmin edilen oranların karşılaştırılmasına dayanan ki-kare tipi ölçülerin minimize edilmesiyle elde edilen tahmin edicileri kapsamaktadır (Agresti, 1996). Bu tip tahmin ediciler kendiliğinden birçok ağırlıklandırılmış en küçük kareler (AEK) tahmincilerini içermektedir. Modern hesaplamalarda ilerlemeler sağlanmadan önce ML tahmincilerine kıyasla hesaplamalarındaki basitlik önem kazanmalarını sağlamıştır.

Neyman’ a göre Fisher ML dışındaki tahmin edicilerin EİANTE olduğunu fark etmemiştir ve şunu ifade etmiştir (Agresti, 2002): “Maksimum olabilirlik (ML) eşitliği gerçekte frekanslardaki doğrusallık ve tüm  $\theta$  değerleri için etkinliğin varlığı

gibi koşullardan türemiş olabilir.” 1950’ li yılların başlarında William Cochran kategorik veri analizi konusunda çeşitli ve önemli başlıklara değinen çalışmasını yayımlamıştır. 1940 yılında varyans durağanlaştırma dönüşümleri ile iki sonuçlu (binomial) ve Poisson yanıtları modellemiş, 1943 yılında aşırı yayılımın üstesinden gelmenin yöntemlerini ele almış ve 1950 yılında bir kaç eşleştirilmiş örnekleme oranların karşılaştırılmasına ilişkin McNemar testinin bir genelleştirmesini (Cochran Q) ortaya koymuştur (Agresti, 1996).

1954 yılında yayımlanan klasik makalesi ise uygulamacılar için yeni bir yöntem karması ve önerisi şeklindedir ki  $X^2$  istatistiğinin iyi işlemesi için ki-kare yakınsamalarının örnek hacminin kılavuz olduğunu ileri sürmektedir. Aynı zamanda yorumlamaların ki-kare istatistiğinin bileşenlere parçalanımı yönünde yapılabileceğinin önemini de vurgulamaktadır. Buna bir örnek, birkaç  $2 \times 2$  boyutlu tabloda koşullu bağımsızlık testini önermesidir ki bu Mantel ve Haenszel testi (1959) ile aynı türdendir (Agresti, 1996). Bartlett’ in  $2 \times 2 \times 2$  boyutlu olumsuzluk tablolarındaki etkileşim yapısı konusundaki çalışması görece olarak az bir etkiye sahip olmuştur. Fakat 1950’ li yılların ortalarında ve 1960’ ların başlarında Bartlett’ in çalışması çok yönlü tablolar için birçok açıdan genişletilmiştir. Örnek olarak Darroch (1962), Good (1963) ve Goodman (1964) verilebilir ki bu makaleler 1965–1975 yılları arasında Loglineer Modeller üzerine yapılan araştırma çalışmalarının temeli olmuştur (Agresti, 1996).

Birch’ in çalışması Glasgow Üniversitesinde asla bildirmediği doktora tezinin bir parçasıdır ki çeşitli koşullar altında üç yönlü tablolarda hücre olasılıklarının ML tahminlerinin nasıl elde edildiğini göstermektedir (Agresti, 2002). Birch ve Watson, 1959 yılında Cramer ve Rao’ nun büyük örneklem dağılımları konusunda sağladıkları teorik sonuçları olumsuzluk tablolarının modellenmesi açısından genişletmişlerdir ve Mantel 1966 yılında bu sonuçları ele alarak Loglineer modelin formülünü açık bir biçimde ortaya koymuştur (Agresti, 1996).

Sonraki on yılda log-lineer ve lojit modelleme konusunda çalışmaların çoğu üç Amerikan Üniversitesinde yapılmıştır: Kuzey Carolina Üniversitesi, Harvard Üniversitesi ve Chicago Üniversitesi ve Chicago Üniversitesinde Leo Goodman çığır

açan bir dizi makale yazmıştır (Agresti, 1996). Bu makalelerin başlıkları şu şekildedir: Ki-kare parçalanımları (bölümlemeleri), kare tablolar için modeller (yarı-bağımsızlık gibi), aşamalı lojit ve log-lineer model oluşturma yöntemleri, log-lineer parametrelerinin ML tahminlerinden asimptotik varyansların türetilmesi, gizli sınıf modelleri, birliktelik modelleri, korelasyon modelleri ve uygunluk analizi.

Goodman'ın ilk çalışmasının araştırmaları için 1968 yılındaki yayımlarına daha sonraki çalışmaları için 1985, 1996 ve 2000 yıllarındaki yayımlarına bakılmalıdır. Goodman aynı zamanda sosyal bilim dergileri için de birçok makale yazmıştır ki; uygulamalarda log-lineer ve lojit metotların popüler olmasında önemli etkilere sahip olmuştur (Örnek olarak, Goodman 1969) (Agresti, 2002). Geçmiş elli yıl boyunca kategorik veri analizinin gelişiminde en verimli katkıları Goodman sağlamıştır ve ayrıca Goodman'ın Chicago'daki öğrencilerinden bazılarının da esaslı katkıları olmuştur. 1970 yılında Shelby Haberman, log-lineer modellemeye önemli teorik katkılar sağlayan bir doktora tezi hazırlamıştır (Agresti, 1996). İlgili tezde dikkate aldığı başlıklar arasında *artık analizleri, ML tahminlerinin varlığı, ordinal değişkenler için Loglineer Modeller ve Rasch Model gibi parametre sayısı örnek büyüklüğü ile artan modeller için teorik sonuçlar* yer almaktadır. Clifford Clogg ise birliktelik modelleri, demografi, nüfus sayımı, oranlar için modeller ve değişik başlıklar üzerine yaptığı çalışmalarla istatistik ve sosyal bilimler alanında etki yaratarak Goodman'ın izinden gitmiştir (Agresti, 2002).

Goodman'ın çalışmalarıyla eş zamanlı olarak Frederick Mosteller ve William Cochran'ın öğrencileri (Stephen Fienberg gibi) tarafından Harvard'da Loglineer-Lojit Modeller için ML Metotları konusunda araştırmalar gerçekleştirilmiştir ki bu araştırmaların çoğunda çok değişkenli ve büyük veri setlerinin analiz edilmesi sırasında ortaya çıkan problemler ilham vermiştir (Agresti, 1996). Mosteller 1968 yılında kesikli ve çok boyutlu veri setlerinin problemsiz hale getirilmesi için log lineer modellerin faydalarını tanımlanmıştır. Fienberg ve öğrencileri ise bu çalışmayı daha da ileri götürmüştür. Fienberg'in 1975 yılında yayımlanan dönüm noktası olan "Kesikli Çok Değişkenli Analiz" adlı kitabı istatistik topluluğuna log lineer modelleri tanıtmak amacıyla yazılmıştır ve mükemmel bir

başvuru kaynağı olarak yerini almıştır (Agresti, 2002). Kuzey Carolina’ da Gary Koch ve bazı öğrenciler tarafından yapılan araştırma biyomedikal bilim alanında kategorik verilerin modellenmesinde ağırlıklandırılmış regresyon metotlarının kullanımında gelişme sağlayarak oldukça etkili olmuştur (Agresti, 1996).

Loglineer modeller konusunda literatürde önceleri tüm kategoriler (sınıflamalar) nominal olarak dikkate alınmıştır. Haberman (1974) ve Simon (1974), log lineer modellerde sınıfların sıralarının nasıl kullanılacağını gösteren bir çalışma oluşturmuşlar ve Leo Goodman tarafından yazılan birçok makalede (1979,1981, 1986) bu çalışma genişletilmiştir(Agresti, 1996). Bu genişletmeler log lineer modellerde yer alan sıralanmış skorların yerine parametreler tarafından oluşturulan birliktelik modellerini de kapsamaktadır. Koşullu bağımsızlık yapısıyla birlikte oluşturulan bazı log lineer modeller olumsuzluk tabloları için grafiksel modeller sağlamaktadır (Bunlar birliktelik grafikleri ile ilişkilidir.).

## 6.5. SON (VE GELECEKTEKİ) GELİŞMELER

Kategorik veri analizi konusunda 90’lı yılların sonlarına kadar en aktif araştırma alanı boylamsal çalışmalarda ve tekrarlı ölçümlerin diğer formlarında olduğu gibi kümelenmiş verilerin modellenmesi olmuştur. Yöntemlerden biri aynı küme içerisinde yer alan yanıtlar arasındaki korelasyonun açıklanmasında kullanılan modelledir (Agresti, 2001). ML tahmini bu şekildeki modeller bakımından oldukça zordur. Örnek olarak *genelleştirilmiş doğrusal karma modellerin* karmaşık biçimleri için varyans bileşenlerini ve regresyon parametrelerini iyi bir şekilde tahmin etmek güçtür. Olabilirlik fonksiyonunu elde etmek için rassal etkiyi birleştirmek nümerik birleşme (integrasyon) şeklinde bir yaklaşımı gerektirmektedir.

Gelecek yıllarda “kümelenmiş korelasyonlu verilerin modellenmesi” aktif bir araştırma alanı olacaktır ve genelleştirilmiş doğrusal karma modellerin bu biçiminde daha fazla genelleştirme ve önemli çalışmalar görülmesi kesindir (Agresti, 2001). Bunlardan biri *genelleştirilmiş toplanabilir karma* modellerdir.

Kategorik yanıtlar için zaman serisi modelleri şimdiye kadar göreceli olarak daha az dikkat çekmiştir. Korelasyonlu yanıtlarla kurulan bu şekildeki bütün modeller için model tanıları ciddi bir öneme sahiptir ve geliştirilmesi gerekmektedir (Agresti, 2002). Son dönemlerdeki bir diğer ilerleme, kesin (exact) küçük örneklem metotları için etkin (yeterli) algoritmaların geliştirilmesidir. Bu şekildeki metotlar ile birlikte ilk olarak, bir testin büyüklüğünün önceden aşağı yukarı belirlenmiş bir seviyeden daha büyük olmaması ve bir güven aralığı için kapsama olasılığının en azından nominal seviyede olması garanti altına alınır (Agresti, 2002).

Kesinlik (exactness) sadece olasılık dağılımına göre yapılan yorumlama ile ilgilidir. Bilinmeyen parametrelere bağlı olarak yapılmaz. Bunu yapmanın da tek bir yolu yoktur. Literatürün çoğunda koşullu yaklaşımlar (sıkıcı parametreler, etkin istatistikleri üzerine koşullar konularak elenir) söz konusudur. Bu nedenle temel fikir Fisher'ın kesin olasılık testi üzerine kurulmuştur. Koşullu metotlar çok yönlüdür ki üssel aileden meydana gelen doğrusal modellere de (Poisson yanıtları için logaritmik doğrusal modeller ve binom yanıtları için lojit modeller gibi) uygulanır ve kanonik bağıntı fonksiyonu kullanılır. Kesin (exact) koşullu yaklaşıma ilişkin hesaplamalarda ilerlemelerin çoğu Cyrus Mehta ve Nitin Patel tarafından ağ algoritmaları kullanılarak meydana getirilmiştir (Agresti, 1996).

Ayrıca Alan Agresti (1992), Mehta (1994) ve Mehta ve Patel (1995) tarafından yapılan araştırmalar ile StatXact ve LogXact programlarının kılavuzlarında da yer almaktadır (Agresti, 2002). “Kesin (exact) Olasılık” metotlarının gelişimi önemli bir süreç gibi görülse de belirli analizler hala yapılabilir değildir ve örnek büyüklüğü ya da tablo boyutu arttıkça üssel olarak hesaplamadaki artış dolayısı ile bazen kullanılması söz konusu olabilir. Son olarak kategorik veri analizi konusunda Bayesyen yaklaşımların gelişimi artan oranda etkin bir alan olmuştur. Parametre çeşitliliği Bayesyen modellemeyi zorlaştırmaktadır.

Bayesyen olasılık tahminlerinin ilk kullanımları Good (1965) ve Lindley (1964) makalelerinde görülmektedir (Stigler, 1986). Kategorik veri analizi için Bayesyen yaklaşımın gelişimi Alan Agresti (2002) tarafından kategorik veri



analizine ilişkin yazılan kitapta ve makalelerde ele alınmaktadır. Örnek olarak, 2005 yılında İtalya İstatistik Birliği dergisine, David B. Hitchcock ile yazdığı “Bayesian Inference for Categorical Data Analysis: A Survey” adlı çalışmaya bakılarak bu konu da hem tarihi hem teknik gelişim konusunda bilgi edinilebilir.

Gelecek konusunda tahminde bulunmak her zaman tehlikelidir. Fakat gelecek araştırmaların hesaplama bakımından genelleştirilmiş doğrusal karma modeller gibi daha yoğun metotlara odaklanmasının olası olduğu görülmektedir. Bir diğer başlık ise geleneksel modelleme alanı dışında, çok fazla sayıda değişkeni içeren çok büyük veri setleri için algoritmali metotların geliştirilmesidir. Bu şekildeki metotlarda (genellikle veri madenciliğine dayanarak) karmaşık veri yapılarının işleme sokulması ile ilgilenilmektedir.

Son sekiz senedir Avrupa’ da ve ülkemizde bir sürü uygulama alanına konu olan kategorik veri analizi konusunda bundan sonra meydana gelebilecek en önemli katkıların neler olacağını tahmin edemeyiz. Ancak sadece uygulamalarda daha önceden oldukça yarar sağlayan bazı çalışmalardan bahsedebiliriz. Özellikle son zamanlarda birçok doktora tezinde bu konu ayrıntıları ile alınmıştır. Örnek olarak, 2002 yılında Deborah Susan Hosler tarafından yazılan “ Models and Graphics in the Analysis of Categorical Variables: The Case of the Youth Tobacco Survey” başlıklı doktora tezi ile yine aynı yılda Chakarida Nukoolkit tarafından yazılan “ Knowledge Discovery in Databases with Categorical Data” adlı doktora tezi verilebilir. Özet olarak kronolojik bir sırayla bu çalışmaların en önemlileri şunlardır (Agresti, 2002):

\* Ordinal verilerin modellenmesi: Leo Goodman tarafından 1979 yılında logaritmik doğrusal modeller konusunda yazılan makale ile Peter Mcgullagh tarafından 1980 yılında oldukça etkili olan oransal odds yaklaşımı konusunda yazılan makalede yer almaktadır.

\* Kesin küçük örneklem metotlarının oluşturulmasında etkin algoritmaların kullanımı: StatXact ve LogXact yazılımlarında Cyrus Mehta, Nitin Patel ve Harvard’ daki meslektaşları tarafından oluşturulmuştur.

\* Çok yönlü olumsuzluk (kontenjans) tabloları için geliştirilen grafiksel modeller (Birliktelik grafikleri ile ilgilidir): John Darroch, Steffen Lauritzen ve Terry Speed tarafından 1980 yılında yazılan makale ile teşvik edilmiştir.

\* Norman Breslow, Ross Prentice ve Washington Üniversitesindeki meslektaşları tarafından olgu-kontrol çalışmalarında odds oranlarının modellenmesi için koşullu olabilirlik metotları ortaya konmuştur.

\* Çok değişkenli kategorik yanıtlar ile boylamsal yanıtlar için Kung-Yee Liang ve Scott Zeger ile Harvard Üniversitesindeki biyoistatistikçiler tarafından geliştirilen ve genel tahminleme denklemlerinin kullanıldığı metodoloji.

Burada değinilen tarihsel gelişim turu için yararlanılan kaynaklar; Stigler (1986) tarafından yazılan “İstatistiğin ve Olasılığın Tarihine İlişkin İncelemeler” adlı kaynağı ile Alan Agresti’ nin üç önemli kitabı (1984, 1996, 2002) ve bunlara dayanarak yıllar boyunca diğer istatistikçilerle (Erling Andersen, R.L. Anderson, William Cochran, Sir David Cox, Leo Goodman, Frederick Mosteller, Stephen Stigler, John Nelder, C.R. Rao ve Gary Koch) paylaştığı kişisel sohbetleri de kapsamaktadır. Kategorik veri analizinin tarihsel gelişiminin daha fazla anlaşılmasını sağlamak için aşağıda kronolojik sıraya göre listelenmiş kaynaklar incelenebilir. Bu kaynaklar, metodolojinin nasıl geliştirildiği konusunda bir fikir sağlamaktadır (Agresti, 2002).

Alternatif olarak, bu konuda daha önceleri yazılmış kitaplara da bakılabilir ki bunlar (Stigler, 1986)): A. E. Maxwell tarafından yazılan “ Nitel Verilerin Analizi” (New York:Methuen, 1961), R.L. Plackett tarafından yazılan “Kategorik Verilerin Analizi” (Londra:Griffin,1974).

Kronolojik kaynak listesi (Agresti, 2002): “Person (1900), Yule (1912), Fisher (1922), Bartlett (1935), Berkson (1944), Neyman (1949), Cochran (1954), Goodman ve Kruskal (1954), Roy ve Mitra (1956), Cox (1958), Mantel ve Haenszel (1959), Birch (1963, 1964b), Caussinus (1966), Goodman (1968),Grizzle ve arkadaşları (1969), Goodman (1970), Haberman (1974), Nelder ve Wedderburn (1972), McFadden (1974), Goodman (1979), McGullagh (1980), Liang ve Zeger (1986), Breslow ve Clayton (1993), Alan Agresti (1996).

## **YEDİNCİ BÖLÜM**

### **MAĞAZACILIK SEKTÖRÜNDE BİR KATEGORİK VERİ**

### **ANALİZİ UYGULAMASI**

Kategorik veri analiz yöntemlerinin incelenmesi amacıyla yapılan çalışmanın uygulama bölümünde uygulama alanı olarak “departman mağazacılığında rekabet olanaklarının geliştirilmesi” için bu alanda hizmet veren bir mağaza ve müşterileri seçilmiştir. Mağazaya ait temel süreçlere ilişkin kriterlerden yola çıkarak cazibe değişkenleri belirlenmiş ve bir anket hazırlanmıştır. Anket üzerinden müşteriler ile yapılan karşılıklı görüşmeler sonucu ilgili değişkenlere ait veriler toplanmıştır. Anket sonuçlarının değerlendirilmesi ve analizi için kategorik veri analizi yöntemlerinden yararlanmıştır.

Uygulamanın yer aldığı bu bölümde sırasıyla uygulanan anketlerin geçerliliği ve güvenilirliği, araştırma evren ve örneklem sayıları, araştırmanın varsayımları ve hipotezleri ile araştırmada kullanılan analiz yöntemlerine ilişkin gerekli açıklamalar yapılmıştır.

Araştırma için hazırlanan anket sorularında “5’li likert ölçeği” kullanılmıştır. Likert Ölçeği, derecelendirmenin yapılması amacıyla hazırlanan anketlerde en sık kullanılan ve güvenilirliği yüksek bir ölçektir. Hazırlanan anket sorularında; ankete cinsiyet ve çocuk sayılarının sorulduğu 2 soru haricinde, 15 soruluk ikinci bölümde müşterilerin “Rekabet faktörleri üzerinden mağaza seçiminde değerlendirme kriterleri” tespit edilmeye çalışılmıştır. Hazırlanan ankette “kapalı uçlu” soru tipi kullanılmıştır.

## 7.1. ANKETLERİN GEÇERLİLİĞİ VE GÜVENİLİRLİĞİ

Geçerlilik ve güvenilirlik tespitlerinde SPSS güvenirlik analizleri kullanılmıştır. Programdan elde edilen ekran görüntüleri ve değerlendirmeleri aşağıda belirtilmiştir.

### 7.1.1 Soruların Zorluk Derecesi Testi

Hotelling's T-Squared Test				
Hotelling's T-Squared	F	df1	df2	Sig
4377,045	275,013	14	95	,000

Yapılan test sonucunda soru ortalamalarının farklı olduğu ( $T^2=275,013$ ;  $p=0,000<0,05$ ) görülmüştür. Soru ortalamalarının birbirine eşit olmaması; soruların denekler tarafından aynı yaklaşımla algılanmadığını, soruların zorluk derecelerinin birbirine eşit olmadığını göstermektedir.

### 7.1.2. Anket Güvenirlik Analizi:

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,853	,865	15

Anket soruları içerisinde yapılan analiz neticesinde elde edilen güvenirlik değeri alpha katsayısı (0,865) 1' e çok yakın çıkmıştır. Bu durum; 109 katılımcı üzerinde yapılan 15 soruluk 5'li Likert ölçekli anketin “yüksek derecede güvenilir” olduğunu ifade etmektedir.

### 7.1.3. Soru-Bütün Korelasyon Testi:

Korelasyon değerlerinin tüm sorular için pozitif çıkması ölçeğin “*toplanabilirlik özelliğinin olduğu*” anlamını taşımaktadır. Tabloya bakıldığında çıkarılması gereken soru olmadığı görülmektedir.

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Fiyat Aralığı	48,7523	21,929	,579	,533	,838
Ziyaret Sıklığı	48,9908	19,731	,597	,660	,841
Değişik Urun Bulabilme	48,8716	21,909	,597	,666	,837
Ödeme Koşulları	48,2477	23,892	,359	,362	,850
Kaliteli Marka Bulabilme	48,4404	24,619	,311	,558	,852
Görsellik	49,4128	21,467	,493	,419	,845
Ulaşım Kolaylığı	48,0459	24,433	,254	,230	,854
Cevre Entegrasyonu (Cafe, Sinema)	50,8899	23,543	,423	,281	,847
Hizmet Performansı	48,4587	23,177	,655	,568	,839
Urun Çeşitliliği	48,9817	23,111	,363	,578	,851
Moda Takibi	48,4312	23,933	,489	,722	,846
Müşteri Odaklılık	48,5963	22,336	,739	,711	,833
Kişiyeye Özel ürün Seçenekleri Sunmak	48,6697	22,094	,780	,722	,831
Beden Kalıp Performansı	48,6422	23,713	,415	,610	,847
Özel Bedenlerin Bulunması	49,0642	21,190	,545	,668	,841

## 7.2. ARAŞTIRMANIN EVRENİ VE ÖRNEKLEM

Uygulamanın evreni olarak Türkiye genelinde departman mağazacılığı sektöründe hizmet veren mağazalar seçilmiştir. Araştırmanın örneklem sınırları ise, bu alanda İzmir sınırları içinde hizmet veren bir mağaza ve müşterileri seçilmiştir.

### **7.3. PROBLEM CÜMLESİ**

Uygulamada araştırılan problem cümlesi “Mağaza müşterilerinin rekabet faktörleri üzerinden mağaza seçiminde değerlendirme kriterleri nelerdir?” şeklinde belirlenmiştir.

### **7.4. VERİLERİ TOPLAMA VE ANALİZ YÖNTEMİ**

Verilerin düzenli bir şekilde toplanması amacıyla, hazırlanan anketin mağaza içinde uygulanabilmesi için mağaza yönetiminden gerekli izinler alınmıştır. Hazırlanan anketler 1 ay süreyle mağazaya gelen müşterilere birebir uygulanmıştır. Toplam 109 kişi üzerinde uygulanan anketlerin tamamı analizlerde kullanılmıştır.

Araştırmada, verilerin analizi bilgisayarda “SPSS 15.0”, “STATA 10.0” ve “MINITAB 15.0” programları kullanılarak gerçekleştirilmiştir. Bu programlar veri toplama aracı ile elde edilen verilerin istatistiksel olarak çözümlerinin yapılması ve anlamlandırılabilmesi amacıyla, araştırma probleminin gerektirdiği istatistik araçları (ortalama, standart sapma, ilişki analizleri, bağımlılık testleri, dağılım grafikleri, çözümlene teknikleri vb.) içermektedir. Bu programların yanında veri toplama aracından elde edilen verilerin sağlıklı şekilde girişlerinin sağlanması ve tasniflerinin yapılması maksadıyla “Microsoft Office Excel” paket programından yararlanılmıştır.

Çalışmada öncelikle; araştırmaya denek olarak katılan bireyler hakkında bilgi sağlanması amacıyla anketin birinci bölümünde yer alan sosyodemografik sorulara ait frekans tabloları ve dağılım grafikleri verilmiştir. Daha sonraki bölümde probleme uygun olarak belirlenen araştırma hipotezleri sırasıyla uygun istatistiksel araçlar kullanılarak test edilmiştir.

Analizlerde kullanılan test araçları; “Reliability Tests” – Güvenilirlik Testleri, “Factor Analysis” – Faktör Analizi, “Pareto Analyze” – Pareto Analizleri, “Residual Analyze Plots” – Model Artık Analizleri, “Bar Chart Graphs” – Yığılımlı Gösterim Grafikleri, “Crosstabs-Olumsuzluk (Kontenjans) Tabloları”, “Chi-Square Tests- Ki-kare Testleri”, “Ordinal Logit Analysis” – Sıralı Lojit Analiz Yöntemi ve “Multinomial Logistic Regression” – Çoklu Lojistik Regresyon Analiz Yöntemi olarak verilebilir.

#### **7.5. ARAŞTIRMA SINIRLILIKLARI**

1. Araştırma mağazaya 1 ay süre içerisinde gelen müşteriler ile sınırlandırılmıştır.
2. Ölçek 15 soru ile sınırlandırılmıştır.
3. Araştırma 1 aylık süre ile sınırlandırılmıştır.

#### **7.6. ARAŞTIRMA VARSAYIMLARI**

Araştırma süresince yapılan analizlerin anlamlılık kazanabilmesi için, analizler aşağıdaki varsayımlara dayandırılmıştır.

1. Araştırmada denek olarak kullanılan bireyler anket sorularına içtenlikle ve doğru olarak cevap vermişlerdir.
2. Ankette amaca yönelik olarak sorularda Likert Ölçeği'nin kullanılması (diğer benzer uygulamalar baz alınarak) uygundur.
3. Tüm veriler ölçülebilir ve tekrarlanabilir verilerdir.

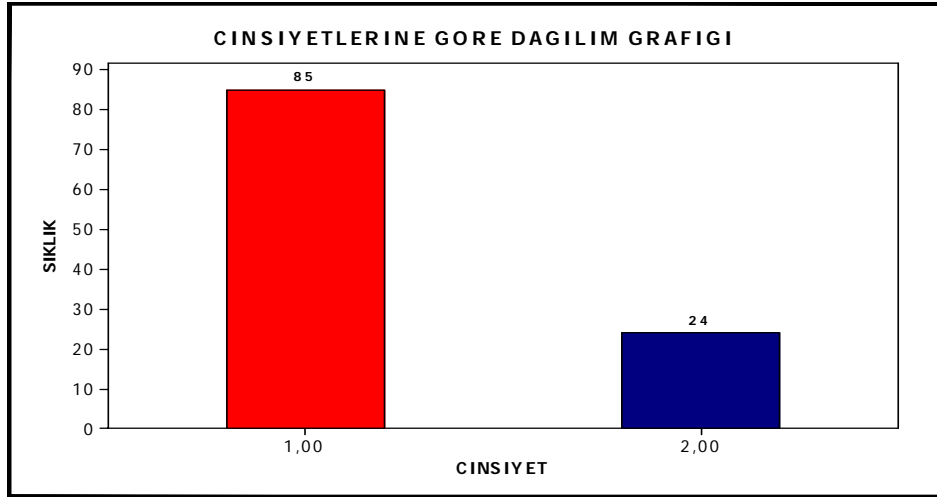
#### **7.7. VERİLERİN ANALİZİ VE BULGULAR**

Araştırmada toplam 109 mağaza müşterisinin sorulan Sosyodemografik özelliklerine ilişkin bilgiler aşağıdaki tablolarda ve şekillerde gösterilmiştir.

Araştırmaya katılanların; 85'inin (% 78.0) "Kadın", 24'ünün (% 22.0) "Erkek" olduğu tespit edilmiştir. Tablo 7.1 ve Şekil 7.1'de araştırmaya katılanların cinsiyetlerine göre dağılımları görülmektedir.

**Tablo 7.1** Cinsiyetlerine Göre Frekans ve Yüzde Değerleri

Cinsiyet	FREKANS	%	% GEÇERLİ	% YİĞİLMALİ
KADIN	85	78,0	78,0	78,0
ERKEK	24	22,0	22,0	100,0
Toplam	109	100,0	100,0	



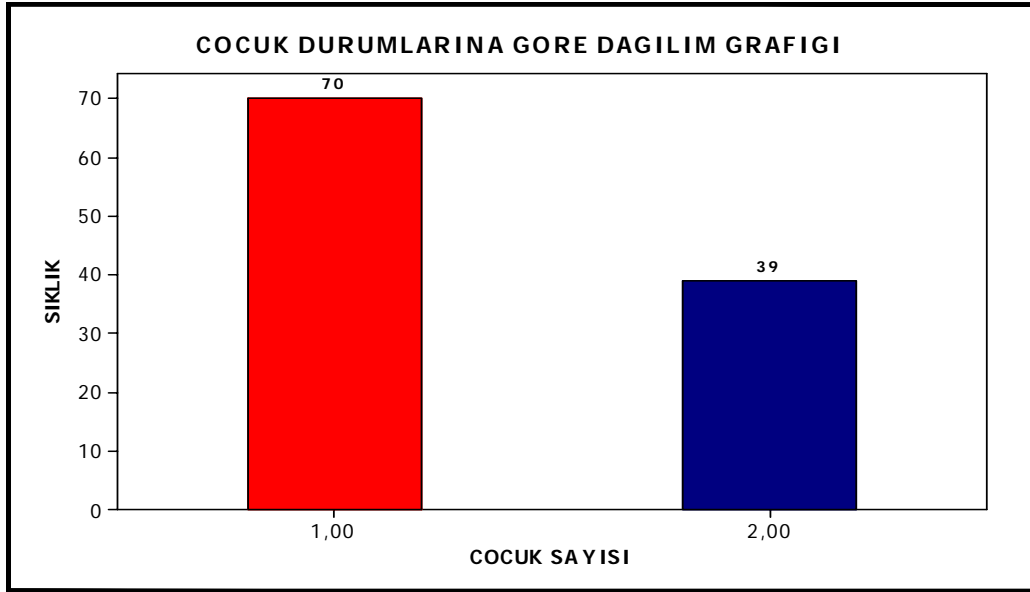
**Şekil 7.1** Cinsiyetlerine Göre Dağılım Grafiği

Araştırmaya katılanların; 70'inin (% 64.2) çocuğunun olmadığı, 39'unun (% 35.8) çocuğunun olduğu tespit edilmiştir. Tablo 7.2 ve Şekil 7.2'de (1: Yok, 2: Var olarak gösterilmiştir) araştırmaya katılanların çocuk durumlarına göre dağılımları görülmektedir.



**Tablo 7.2** Çocuk Durumlarına Göre Frekans ve Yüzde Değerleri

Çocuk Durumu	FREKANS	%	% GEÇERLİ	% YİĞİLMALİ
YOK	70	64,2	64,2	64,2
VAR	39	35,8	35,8	100,0
Toplam	109	100,0	100,0	



**Şekil 7.2** Çocuk Durumlarına Göre Dağılım Grafiği

Analizlere geçmeden bir önceki basamakta, anket soruları neticesinde elde edilen değişkenlerin faktör analizi ile gruplanması bulunmaktadır. Bu şekilde elde edilecek analizler kuvvetlendirilecek, kurulacak ilişkiler daha olumlu sonuçlar doğurabilecektir. Faktör analizi spss ekran çıktıları ve tabloları aşağıda belirtilmiştir.

**Tablo 7.3** KMO and Bartlett's Test Çıktısı

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,719
Bartlett's Test of Sphericity	Approx. Chi-Square	645,377
	<b>Df</b>	<b>91,000</b>
	<b>Sig.</b>	<b>,000</b>

Faktör analizinde Eigen değeri 1 olarak alınmış. Kaiser-Meyer Olkin örneklem doğruluk ölçütü; 0,719 olarak saptanmıştır. Kaiser'in ölçütüne göre bu orta derece bir örneklem büyüklüğüdür. Barlett's sferisite testi anlamlıdır ( $p < 0.001$ ). Bu açılardan veriler faktör analizi uygulamaya uygundur.

**Tablo 7.4** Açıklanan Toplam Varyans Çıktısı

Component	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4,878	34,841	34,841	2,814	20,101	20,101
2	1,765	12,604	47,445	2,235	15,966	36,067
3	1,281	9,152	56,597	2,120	15,143	51,210
<b>4</b>	<b>1,159</b>	<b>8,277</b>	<b>64,874</b>	<b>1,913</b>	<b>13,663</b>	<b>64,874</b>
5	,999	7,139	72,013			
6	,828	5,918	77,930			
7	,653	4,666	82,597			
8	,637	4,552	87,148			
9	,516	3,688	90,836			
10	,367	2,624	93,460			
11	,344	2,454	95,914			
12	,239	1,708	97,622			
13	,196	1,398	99,020			
14	,137	,980	100,000			

**Tablo 7.5** Çevrilmiş Faktör Matrisi (Rotated Component Matrix)

	Faktörler			
	1	2	3	4
Ziyaret Sıklığı	,748			
Hizmet Performansı	,709			
Ürün Çeşitliliği	,708			
Değişik Urun Bulabilme	,646			
Kişiyeye Özel ürün Seçenekleri Sunmak	,573			
Fiyat Aralığı	,536			
Kaliteli Marka Bulabilme		,860		
Moda Takibi		,788		
Beden Kalıp Performansı			,834	
Özel Bedenlerin Bulunması			,833	
Görsellik			,480	
Ulaşım Kolaylığı				,693
Ödeme Koşulları				,689
Çevre Entegrasyonu (Cafe, Sinema)				,383

Yukarıda SPSS ekran çıktıları verilen analizlerde görüldüğü üzere; faktör analizi ile 15 soruluk likert bölümü 4 faktöre indirgenmiştir. İndirgenmiş faktörler üzerinden toplam verilerin yaklaşık **%64'ü** (Bkz. Tablo 7.4) ifade edilebilmektedir. Analizler ile belirlenen faktör gruplarının adları ve ortalamaları Tablo 7.6'da ifade edilmiştir.

**Tablo 7.6** Faktör Grupları ve Ortalama Değerleri Tablosu

LİKERT SORULARI	FAKTÖRLER	ORTALAMA
Ziyaret Sıklığı	ÇEŞİTLİLİK_HİZMET (Faktör 1)	4,24
Hizmet Performansı		
Ürün Çeşitliliği		
Değişik Urun Bulabilme		
Kişiyeye Özel urun Seçenekleri Sunmak		
Fiyat Aralığı		
Kaliteli Marka Bulabilme	KALİTE (Faktör 2)	3,89
Moda Takibi		
Beden Kalıp Performansı	BEDEN (Faktör 3)	3,28
Özel Bedenlerin Bulunması		
Görsellik		
Ulaşım Kolaylığı	KONUM (Faktör 4)	3,26
Ödeme Koşulları		
Çevre Entegrasyonu (Kafe, Sinema)		

Yapılacak diğer analizlerde belirlenen bu faktör grupları bağımsız değişken olarak kullanılacaktır. Sırasıyla yapılan analizler ve hipotezleri müteakip paragraflarda belirtilmiştir. Analizlere geçmeden önce faktör gruplarını oluşturan sorular üzerinden SPSS ile her bir faktör grubu için ayrı ayrı ikinci bir güvenilirlik analizi uygulanacaktır.

➤ ÇEŞİTLİLİK\_HİZMET (Faktör 1) Güvenilirlik analiz sonucu:

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,802	,833	6

Faktör grubu içerisinde yer alan altı soru üzerinden yapılan güvenilirlik analizinde elde edilen güvenilirlik değeri alpha katsayısı (0,833) 1' e çok yakın çıkmıştır. Bu durum; 6 soruluk faktör grubunun “yüksek derecede güvenilir” olduğunu ifade etmektedir.

Hotelling's T-Squared	F	df1	df2	Sig
137,936	26,566	5	104	,000

Yapılan test sonucunda soru ortalamalarının farklı olduğu ( $T^2=137,936$ ;  $p=0,000<0,05$ ) görülmüştür.

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Fiyat Aralığı	17,6330	5,679	,542	,470	,775
Ziyaret Sıklığı	17,8716	4,335	,611	,454	,781
Değişik Urun Bulabilme	17,7523	5,558	,605	,576	,761
Hizmet Performansı	17,3394	6,263	,674	,495	,768
Urun Çeşitliliği	17,8624	5,861	,461	,515	,793
Kişiyeye Özel Ürün Seçenekleri Sunmak	17,5505	5,898	,700	,520	,753

Korelasyon deęerleri tüm sorular için pozitif çıktığından, çıkarılması gereken soru olmadığı görölmektedir.

➤ KALİTE (Faktör 2) Güvenilirlik analiz sonucu:

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,676	<b>,714</b>	2

Faktör grubu içerisinde yer alan iki soru üzerinden yapılan güvenilirlik analizinde elde edilen güvenilirlik deęeri alpha katsayısı (0,714) 1' e yakın çıkmıştır. Bu durum; 2 soruluk faktör grubunun “*güvenilir*” olduğunu ifade etmektedir.

Hotelling's T-Squared	F	df1	df2	Sig
9,329	9,329	1	108	<b>,003</b>

Yapılan test sonucunda soru ortalamalarının farklı olduğu ( $T^2=9,329$ ;  $p=0,003<0,05$ ) görölmüştür.

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Kaliteli Marka Bulabilme	3,5688	,414	,555	,308	.....
Moda Takibi	3,3303	,945	,555	,308	.....

Analiz neticesinde tüm sorular için korelasyon deęerleri pozitif çıkmasına karşın, soruların silinmesi halinde güvenilirliğe ne kadar etki edecekleri istatistiksel olarak elde edilememiştir. Ancak, analizden çıkarılması gereken soru olmadığı görölmektedir.

➤ BEDEN (Faktör 3) Güvenilirlik analiz sonucu:

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,660	<b>,690</b>	3

Faktör grubu içerisinde yer alan üç soru üzerinden yapılan güvenilirlik analizinde elde edilen güvenilirlik değeri alpha katsayısı (0,690) 1' e yakın çıkmıştır. Bu durum; 3 soruluk faktör grubunun “*güvenilir*” olduğunu ifade etmektedir.

Hotelling's T-Squared	F	df1	df2	Sig
109,567	54,276	2	107	<b>,000</b>

Yapılan test sonucunda soru ortalamalarının farklı olduğu ( $T^2=109,567$ ;  $p=0,000<0,05$ ) görülmüştür.

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Görsellik	6,9358	1,320	,391	,180	,687
Beden Kalıp Performansı	6,1651	1,843	,506	,383	,594
Özel Bedenlerin Bulunması	6,5872	1,059	,615	,463	,342

Korelasyon değerleri tüm sorular için pozitif çıktığından, çıkarılması gereken soru olmadığı görülmektedir.

➤ KONUM (Faktör 4) Güvenilirlik analiz sonucu:

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,453	,454	3

Faktör grubu içerisinde yer alan üç soru üzerinden yapılan güvenilirlik analizinde elde edilen güvenilirlik değeri alpha katsayısı (0,454) 0,5' e yakın çıkmıştır. Bu durum; 3 soruluk faktör grubunun “*kabul edilebilir*” olduğunu ifade etmektedir.

Hotelling's T-Squared	F	Df1	df2	Sig
2583,691	1279,884	2	107	,000

Yapılan test sonucunda soru ortalamalarının farklı olduğu ( $T^2=2835,691$ ;  $p=0,000<0,05$ ) görülmüştür.

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Ödeme Koşulları	5,7064	,524	,352	,132	,217
Ulaşım Kolaylığı	5,5046	,586	,280	,104	,351
Cevre Entegrasyonu (Cafe, Sinema)	5,8486	,537	,367	,149	,482

Korelasyon değerleri tüm sorular için pozitif çıktığından, çıkarılması gereken soru olmadığı görülmektedir.



Yukarıdaki paragraflarda ifade edilen güvenilirlik analizleri neticesinde anketten çıkarılması gereken herhangi bir sorunun varlığına rastlanmamıştır.

Faktör grupları ile elde edilen veriler kategorik veri olmaları sebebiyle, uygulanacak analiz yöntemleri de kategorik veri analiz yöntemleri olmalıdır. Kategorik verilere uygulanabilecek iki farklı model temelli analiz yöntemi ile ilişkilerin gösteriminde kullanılacak olumsuzluk tabloları çalışmanın teorik bölümünde ifade edilmiştir. Analizlerde her iki yöntem ile modeller oluşturulacak ve analiz yöntemlerinin uygunluğu araştırılacaktır. Uygulanacak analiz yöntemleri, Ordinal Logit Analiz Yöntemi ile Multinomial Logistic Regresyon Analiz Yöntemi olarak belirlenmiştir.

Bahse konu analizlerde bağımsız değişken olarak faktör grupları olarak gözükmekte; ankette yer bulan mağazanın “Müşteri Odaklılık” seviyesini ölçmek amacıyla hazırlanan soru bağımlı değişkenimiz olarak tespit edilmiştir. Oluşturulacak modeller bu değişkenlerin etkileşimleri üzerinden ifade edilecektir.

Öncelikle belirlenen faktör grupları arasındaki ilişkinin boyutu ve şiddetinin incelenmesi amacıyla, değişkenlerimiz arasında korelasyon analizi yapılacaktır. Analiz neticesinde elde edilen program çıktıları ve yorumları aşağıdaki paragraflarda ifade edilmiştir.

		ÇEŞİTLİLİK_ HİZMET	KALİTE	BEDEN	KONUM
ÇEŞİTLİLİK_ HİZMET	Pearson Correlation	1	,407(**)	,444(**)	,201(*)
	Sig. (2-tailed)		,000	,000	,036
	N	109	109	109	109
KALİTE	Pearson Correlation	,407(**)	1	,325(**)	,224(*)
	Sig. (2-tailed)	,000		,001	,019
	N	109	109	109	109
BEDEN	Pearson Correlation	,444(**)	,325(**)	1	,279(**)
	Sig. (2-tailed)	,000	,001		,003
	N	109	109	109	109
KONUM	Pearson Correlation	,201(*)	,224(*)	,279(**)	1
	Sig. (2-tailed)	,036	,019	,003	
	N	109	109	109	109

Korelasyon analizi neticesinde; elde edilen tüm korelasyon değerleri (sign.2-tailed <0,05 olduğundan) istatistiksel olarak anlamlıdır. Anlamlılık seviyeleri tablo içinde (\*\*) ve (\*) şeklinde ifade edilmiştir. (\*\*) işareti değişkenler arası ilişkinin yüksek seviyede (%99 güven aralığında), (\*) işareti ise değişkenler arası ilişkinin normal seviyede (%95 güven aralığında) anlamlı olduğunu göstermektedir. Anlamlı olduğu görülen faktörler arası ilişki şiddetlerini ifade eder tablo aşağıda verilmiştir.

	ÇEŞİTLİLİK_ HİZMET	KALİTE	BEDEN	KONUM
ÇEŞİTLİLİK_ HİZMET	1			
KALİTE	OLDUKÇA ŞİDDETLİ	1		
BEDEN	OLDUKÇA ŞİDDETLİ	OLDUKÇA ŞİDDETLİ	1	
KONUM	ŞİDDETLİ	OLDUKÇA ŞİDDETLİ	OLDUKÇA ŞİDDETLİ	1

### 7.7.1. Kontenjans (Olumsuzluk) Tabloları ve Ki-Kare Analizleri

Bu kısımda yapılan analizlerde faktör grupları ile müşteri odaklılık değişkeni ve demografik özellikler ile müşteri odaklılık arasındaki ilişkileri tespit etmeye yönelik olumsuzluk tabloları oluşturulacak ve ki-kare analizleri yapılacaktır. Sırasıyla hipotezler, tablolar ve yorumları aşağıdaki paragraflarda ifade edilmiştir.

Aşağıda farklı değişkenler arasında yapılan olumsuzluk tablolarında ortak nokta, ankete katılanların büyük çoğunluğunun (%72,8) mağazanın genel müşteri odaklılık politikasının “İyi” seviyede olduğunu düşünmeleridir.

#### ► Çeşitlilik\_Hizmet ile Müşteri Odaklılık İçin Olumsuzluk Tablosu ve Ki-Kare Analizleri

Ankete katılanlar arasında Çeşitlilik\_Hizmet seviyesini; “Vasat” olarak nitelendiren katılımcıların %92,3’ünün mağazanın Müşteri Odaklılık politikasını da “Vasat” olarak değerlendirdiği, “İyi” olarak nitelendiren katılımcıların ise %82,2’lik kısmının mağazanın Müşteri Odaklılık politikasının “İyi” seviyede gördüğü tespit edilmiştir.

**Tablo 7.7** Çeşitlilik-Hizmet ile Müşteri Odaklılık İçin Olumsuzluk Tablosu

			Müşteri Odaklılık		
			VASAT	İYİ	Toplam
ÇEŞİTLİLİK_ HİZMET	VASAT	Sıklık	12	1	13
		% ÇEŞİTLİLİK_HİZMET	<b>92,3%</b>	7,7%	100,0%
	İYİ	Sıklık	16	74	90
		% ÇEŞİTLİLİK_HİZMET	17,8%	<b>82,2%</b>	100,0%
	Toplam	Sıklık	28	75	103
		% ÇEŞİTLİLİK_HİZMET	27,2%	72,8%	100,0%

Ki-kare analizi neticesinde elde edilen sonuçlarda; **Pearson Ki-Kare** değerine ( $\chi^2=31,876$ ;  $p=0,000<0,05$ ) bakılarak Çeşitlilik\_Hizmet ve Müşteri Odaklılık değişkenlerinin **birbirinden bağımsız olmadığı, olabilirlik oran istatistiğine** bakıldığında (LR=29,235;  $p=0,000$ ) oluşturulacak log-linear modelde etken olduğu, **doğrusala doğrusal birliktelik istatistiğine** bakıldığında ise değişkenlerin doğrusal bir ilişki seviyesine sahip olduğu (31,567;  $p=0,000$ ) kararlarına varılmaktadır.

Simetrik ölçüler tablosunda ifade edilen istatistiki değerlere bakıldığında; verilerin ordinal olmaları veya nominal olmalarına göre farklılık gösteren istatistiklerin tamamında aynı sonuç çıkmıştır. Phi ve Cramer's değerleri ile olasılık katsayılarına bakılarak değişkenler arasındaki ilişkinin şiddetinin kuvvetli seviyede olduğu ( $p=0,000<0,05$ ); ordinal veriler olduğu kabul edilerek yapılan analizlerde ise aynı şekilde kuvvetli ilişki düzeyi yakalanmıştır ( $p=0,000<0,05$ ).

**Tablo 7.8** Çeşitlilik-Hizmet Faktörü ile Müşteri Odaklılık İçin Ki-Kare Analiz Tablosu

	Değer	Serbestlik Derecesi	Asimptotik Değer (2-yanlı)
Pearson Ki-Kare	31,876(a)	1	,000
<b>Olabilirlik Oran İstatistiği</b>	29,235	1	,000
Doğrusala Doğrusal Birliktelik	31,567	1	,000
Geçerli Durumların Sayısı	103		
a. 1 hücre (25,0%) için beklenen değer 5'ten azdır. Beklenen en az değer 3,53'dür.			

**Tablo 7.9** Çeşitlilik-Hizmet ile Müşteri Odaklılık Arasındaki Simetrik Ölçüler Tablosu

		Değer	Tahmini Standart Hata	Tahmini T	Tahmini Anlamlılık Değeri
Nominal-Nominal	Phi	,556			,000
	Cramer's V	,556			,000
	Olasılık Katsayısı	,486			,000
Ordinal-Ordinal	Kendall's tau-b	,556	,086	4,018	,000
	Kendall's tau-c	,329	,082	4,018	,000
	Gamma	,965	,037	4,018	,000
Geçerli Durum Sayısı		109	103		

► **Kalite ve Müşteri Odaklılık İçin Olumsuzluk Tabloları ve Ki-Kare Analizleri**

Ankete katılanlar arasında Kalite seviyesini; “Vasat” olarak nitelendiren katılımcıların %68,4’ünün mağazanın Müşteri Odaklılık politikasını da “Vasat” olarak değerlendirdiği, “İyi” olarak nitelendiren katılımcıların ise %82,1’lik kısmının mağazanın Müşteri Odaklılık politikasının “İyi” seviyede gördüğü tespit edilmiştir.

**Tablo 7.10** Kalite-Müşteri Odaklılık İçin Olumsuzluk Tablosu

			Müşteri Odaklılık		
			VASAT	İYİ	Toplam
KALİTE	VASAT	Sıklık	13	6	19
		% KALİTE	<b>68,4%</b>	31,6%	100,0%
	İYİ	Sıklık	15	69	84
		% KALİTE	17,9%	<b>82,1%</b>	100,0%
	Toplam	Sıklık	28	75	103
		% KALİTE	27,2%	<b>72,8%</b>	100,0%

Ki-kare analizi neticesinde elde edilen sonuçlarda; **Pearson Ki-Kare** değerine ( $\chi^2=20,014$ ;  $p=0,000<0,05$ ) bakılarak kalite ile müşteri odaklılık değişkenlerinin **birbirinden bağımsız olmadığı, olabilirlik oran istatistiğine** bakıldığında (LR=18;  $p=0,000$ ) oluşturulacak log-linear modelde etken olduğu, **doğrusala doğrusal birliktelik istatistiğine** bakıldığında ise değişkenlerin doğrusal bir ilişki seviyesine sahip olduğu ( $19,820$ ;  $p=0,000$ ) kararlarına varılmaktadır.

**Tablo 7.11** Kalite ve Müşteri Odaklılık Ki-Kare Analiz Tablosu

	Değer	Serbestlik Derecesi	Asimptotik Değer (2-yanlı)
Pearson Ki-Kare	20,014(a)	1	,000
Olabilirlik Oran İstatistiği	18,000	1	,000
Doğrusala Doğrusal Birliktelik	19,820	1	,000
Geçerli Durumların Sayısı	103		
a. 0 hücre (0%) için beklenen değer 5'ten azdır. Beklenen en az değer 5,17'dir.			

Simetrik ölçüler tablosunda ifade edilen istatistiki değerlere bakıldığında; verilerin ordinal olmaları veya nominal olmalarına göre farklılık gösteren istatistiklerin tamamında aynı sonuç çıkmıştır. Phi ve Cramer's değerleri ile olasılık katsayılarına bakılarak değişkenler arasındaki ilişkinin şiddetinin kuvvetli seviyede olduğu ( $p=0,000<0,05$ ); ordinal veriler olduğu kabul edilerek yapılan analizlerde ise aynı şekilde kuvvetli ilişki düzeyi tespit edildiği ( $p=0,004<0,05$  görülmektedir).

**Tablo 7.12** Kalite-Müşteri Odaklılık Simetrik Ölçüler Tablosu

		Değer	Tahmini Standart Hata	Tahmini T	Tahmini Anlamlılık Değeri
Nominal-	Phi	441			,000
Nominal	Cramer's V	441			,000
	Kontenjans Katsayısı	,403			,000
Ordinal-	Kendall's tau-b	,441	,102	3,604	,000
Ordinal	Kendall's tau-c	,304	,084	3,604	,000
	Gamma	,818	,094	3,604	,000
Geçerli Durum Sayısı		109	103		

► **Beden-Müşteri Odaklılık Olumsuzluk Tablosu ve Ki-Kare Analizleri**

Ankete katılanlar arasında Beden seviyesini; “Vasat” olarak nitelendiren katılımcıların %58,6’sının, “İyi” olarak nitelendiren katılımcıların ise %91,1’lik kısmının mağazanın Müşteri Odaklılık politikasının “İyi” seviyede gördüğü tespit edilmiştir.



**Tablo 7.13** Beden-Müşteri Odaklılık Olumsuzluk Tablosu

			Müşteri Odaklılık		
			VASAT	İYİ	Toplam
BEDEN	VASAT	Sıklık	24	34	58
		% BEDEN	41,4%	<b>58,6%</b>	100,0%
	İYİ	Sıklık	4	41	45
		% BEDEN	8,9%	<b>91,1%</b>	100,0%
	Toplam	Sıklık	28	75	103
		% BEDEN	27,2%	<b>72,8%</b>	100,0%

Ki-kare analizi neticesinde elde edilen sonuçlarda; **Pearson Ki-Kare** değerine ( $\chi^2=13,514$ ;  $p=0,000<0,05$ ) bakılarak beden ile müşteri odaklılık değişkenlerinin **birbirinden bağımsız olmadığı, olabilirlik oran istatistiğine** bakıldığında (LR=14,859;  $p=0,000$ ) oluşturulacak log-lineer modelde etken olduğu, **doğrusala doğrusal birliktelik istatistiğine** bakıldığında ise değişkenlerin doğrusal bir ilişki seviyesine sahip olduğu (13,382;  $p=0,000$ ) kararlarına varılmaktadır.

**Tablo 7.14** Beden-Müşteri Odaklılık Ki-Kare Analiz Tablosu

	Değer	Serbestlik Derecesi	Asimptotik Değer (2-yanlı)
Pearson Ki-Kare	13,514(a)	1	,000
Olabilirlik Oran İstatistiği	14,859	1	,000
Doğrusala Doğrusal Birliktelik	13,382	1	,000
Geçerli Durumların Sayısı	103		

a. Beklenen değeri 5'ten az olan hücre söz konusu değildir. Minimum beklenen değer 12,23'tür.

Simetrik ölçüler tablosunda ifade edilen istatistiki değerlere bakıldığında; verilerin ordinal olmaları veya nominal olmalarına göre farklılık gösteren istatistiklerin tamamında aynı sonuç çıkmıştır. Phi ve Cramer's değerleri ile olasılık katsayılarına bakılarak değişkenler arasındaki ilişkinin şiddetinin kuvvetli seviyede olduğu ( $p=0,000<0,05$ ); ordinal veriler olduğu kabul edilerek yapılan analizlerde ise aynı şekilde kuvvetli ilişki düzeyi tespit edildiği ( $p=0,000<0,05$ ) görülmektedir.

**Tablo 7.15** Beden-Müşteri Odaklılık Simetrik Ölçüler Tablosu

		Değer	Tahmini Standart Hata	Tahmini T	Tahmini Anlamlılık Değeri
Nominal-Nominal	Phi	,362			,000
	Cramer's V	,362			,000
	Contingency Coefficient	,341			,000
Ordinal-Ordinal	Kendall's tau-b	,362	,080	,178	,000
	Kendall's tau-c	,320	,077	,178	,000
	Gamma	,757	,125	,178	,000
Geçerli Durum Sayısı			103		

► **Konum-Müşteri Odaklılık İçin Olumsuzluk Tablosu ve Ki-Kare Analizleri**

Ankete katılanlar arasında Konum seviyesini; “Vasat” olarak nitelendiren katılımcıların %66,3’inin, “İyi” olarak nitelendiren katılımcıların ise tamamının mağazanın Müşteri Odaklılık politikasının “İyi” seviyede gördüğü tespit edilmiştir.

**Tablo 7.16** Konum-Müşteri Odaklılık Olumsuzluk Tablosu

		Müşteri Odaklılık			
		VASAT	İYİ	Toplam	
KONUM	VASAT	Sıklık	28	55	83
		% KONUM	33,7%	<b>66,3%</b>	100,0%
	İYİ	Sıklık	0	20	20
		% KONUM	0,0%	<b>100,0%</b>	100,0%
	Toplam	Sıklık	28	75	103
		% KONUM	27,2%	<b>72,8%</b>	100,0%

Ki-kare analizi neticesinde elde edilen sonuçlarda; **Pearson Ki-Kare** değerine ( $\chi^2=9,266$ ;  $p=0,002<0,05$ ) bakılarak konum ile müşteri odaklılık değişkenlerinin **birbirinden bağımsız olmadığı, olabilirlik oran istatistiğine** bakıldığında (LR=14,410;  $p=0,000$ ) oluşturulacak log-linear modelde etken olduğu, **doğrusala doğrusal birliktelik istatistiğine** bakıldığında ise değişkenlerin doğrusal bir ilişki seviyesine sahip olduğu (9,176;  $p=0,002$ ) kararlarına varılmaktadır.

**Tablo 7.17** Konum-Müşteri Odaklılık Ki-Kare Analiz Tablosu

	Değer	Serbestlik Derecesi	Asimptotik Değer (2-yanlı)
Pearson Ki-Kare	9,266(a)	1	,002
Olabilirlik Oran İstatistiği	14,410	1	,000
Doğrusala Doğrusal Birliktelik	9,176	1	,002
Geçerli Durumların Sayısı	103		

a. 5' ten küçük beklenen değere sahip hücre bulunmamaktadır. Minimum beklenen değer 5,44'tür.

Simetrik ölçüler tablosunda ifade edilen istatistiki değerlere bakıldığında; verilerin ordinal olmaları veya nominal olmalarına göre farklılık gösteren istatistiklerin tamamında aynı sonuç çıkmıştır. Phi ve Cramer's değerleri ile olasılık katsayılarına bakılarak değişkenler arasındaki ilişkinin şiddetinin kuvvetli seviyede olduğu ( $p=0,002<0,05$ ); ordinal veriler olduğu kabul edilerek yapılan analizlerde ise aynı şekilde kuvvetli ilişki düzeyi tespit edildiği ( $p=0,000<0,05$ ) görülmektedir.

**Tablo 7.18** Konum-Müşteri Odaklılık Simetrik Ölçüler Tablosu

		Değer	Tahmini Standart Hata	Tahmini T	Tahmini Anlamlılık Değeri
Nominal-Nominal	Phi	,300			,002
	Cramer's V	,300			,002
	Contingency Coefficient	,287			,002
Ordinal-Ordinal	Kendall's tau-b	,300	,042	4,619	,000
	Kendall's tau-c	,211	,046	4,619	,000
	Gamma	1,000	,000	4,619	,000
Geçerli Durum Sayısı			103		

➤ **Cinsiyet-Müşteri Odaklılık Olumsuzluk Tablosu ve Ki-Kare Analizleri**

Ankete katılanlar arasında çoğunluğu oluşturan “Kadın” katılımcıların %71,3’ünün, “Erkek” katılımcıların ise %78,3’lük kısmının mağazanın müşteri odaklılık politikasının “İyi” seviyede olduğu görüşünde olduğu görülmektedir.

**Tablo 7.19** Cinsiyet-Müşteri Odaklılık Olumsuzluk Tablosu

			Müşteri Odaklılık		
			VASAT	İYİ	Toplam
CİNSİYET	KADIN	Sıklık	23	57	80
		% CİNSİYET	28,8%	<b>71,3%</b>	100,0%
	ERKEK	Sıklık	5	18	23
		% CİNSİYET	21,7%	<b>78,3%</b>	100,0%
	Toplam	Sıklık	28	75	103
		% CİNSİYET	27,2%	<b>72,8%</b>	100,0%

Ki-kare analizi neticesinde elde edilen sonuçlarda; **Pearson Ki-Kare** değerine ( $p=0,505>0,05$ ) bakılarak cinsiyet ile müşteri odaklılık değişkenlerinin **birbirinden bağımsız olduğu, olabilirlik oran istatistiğine** bakıldığında ( $LR=0,459$ ;  $p=0,498$ ) oluşturulacak log-linear modelde etken olmadığı, **doğrusala doğrusal birliktelik istatistiğine** bakıldığında ise değişkenlerin doğrusal bir ilişkidenden bahsedilemeyeceği ( $0,439$ ;  $p=0,507$ ) kararlarına varılmaktadır.

**Tablo 7.20** Cinsiyet-Müşteri Odaklılık Ki-Kare Analiz Tablosu

	Değer	Serbestlik Derecesi	Asimptotik Değer (2-yanlı)
Pearson Ki-Kare	,444(a)	1	,505
Olabilirlik Oran İstatistiği	,459	1	,498
Doğrusala Doğrusal Birliktelik	,439	1	,507
Geçerli Durumların Sayısı	103		

a. 5' ten küçük beklenen değere sahip hücre söz konusu değildir. Minimum beklenen değer 6,25' tir.

Simetrik ölçüler tablosunda ifade edilen istatistiki değerlere bakıldığında; verilerin ordinal olmaları veya nominal olmalarına göre farklılık gösteren istatistiklerin tamamında aynı sonuç çıkmıştır. Phi ve Cramer's değerleri ile olasılık katsayılarına bakılarak değişkenler arasındaki ilişkinin şiddetinin çok düşük seviyede olduğu ( $p=0,505>0,05$ ); ordinal veriler olduğu kabul edilerek yapılan analizlerde ise aynı şekilde çok düşük ilişki düzeyi tespit edildiği ( $p=0,484<0,05$ ); uyuşma ölçüsü olarak kappa değeri hesaplanamadığından örtüşme durumlarından bahsedilemeyeceği görülmektedir (Kappa istatistiği simetrik bir tablo olmadığından hesaplanamamıştır.).

**Tablo 7.21** Cinsiyet-Müşteri Odaklılık Simetrik Ölçüler Tablosu

		Değer	Tahmini Standart Hata	Tahmini T	Tahmini Anlamlılık Değeri
Nominal-Nominal	Phi	,066			,505
	Cramer's V	,066			,505
	Contingency Coefficient	,065			,505
Ordinal-Ordinal	Kendall's tau-b	,066	,093	,700	,484
	Kendall's tau-c	,049	,070	,700	,484
	Gamma	,185	,272	,700	,484
Geçerli Durum Sayısı			103		

► **Çocuk Durumu-Müşteri Odaklılık Kontenjans Tabloları ve Ki-Kare Analizleri**

Çoğunluğu oluşturan çocuğu olmayan katılımcıların %69,2'sinin, Çocuğu olan katılımcıların ise %78,9'luk kısmının müşteri odaklılık politikasının "İyi" seviyede olduğu görüşünde olduğu görülmektedir.

**Tablo 7.22** Çocuk Durumu -Müşteri Odaklılık Olumsuzluk Tablosu

			Müşteri Odaklılık		
			VASAT	İYİ	Toplam
ÇOCUK DURUMU	YOK	Sıklık	20	5	5
		% ÇOCUK DURUMU	30,8%	<b>69,2%</b>	100,0%
	VAR	Sıklık	8	30	38
		% ÇOCUK DURUMU	21,1%	<b>78,9%</b>	100,0%
	Toplam	Sıklık	28	75	103
		% ÇOCUK DURUMU	27,2%	<b>72,8%</b>	100,0%

Ki-kare analizi neticesinde elde edilen sonuçlarda; **Pearson Ki-Kare** değerine ( $p=0,285>0,05$ ) bakılarak çocuk durumu ile müşteri odaklılık değişkenlerinin birbirinden **bağımsız olduğu, olabilirlik oran istatistiğine** bakıldığında ( $LR=1,172$ ;  $p=0,279$ ) oluşturulacak log-linear modelde etken olmadığı, **doğrusala doğrusal birliktelik istatistiğine** bakıldığında ise değişkenlerin doğrusal bir ilişkiden bahsedilemeyeceği ( $1,133$ ;  $p=0,287$ ) kararlarına varılmaktadır.

**Tablo 7.23** Çocuk Durumu-Müşteri Odaklılık Ki-Kare Analiz Tablosu

	Değer	Serbestlik Derecesi	Asimptotik Değer (2-yanlı)
Pearson Ki-Kare	1,144(a)	1	,285
Olabilirlik Oran İstatistiği	1,172	1	,279
Doğrusala Doğrusal Birliktelik	1,133	1	,287
Geçerli Durumların Sayısı	103		
a. 5' ten küçük beklenen değer söz konusu değildir. Minimum beklenen değer 10,33' tür.			

Simetrik ölçüler tablosunda ifade edilen istatistiki değerlere bakıldığında; verilerin ordinal olmaları veya nominal olmalarına göre farklılık gösteren istatistiklerin tamamında aynı sonuç çıkmıştır. Phi ve Cramer's değerleri ile olasılık katsayılarına bakılarak değişkenler arasındaki ilişkinin şiddetinin düşük seviyede olduğu ( $p=0,285>0,05$ ); ordinal veriler olduğu kabul edilerek yapılan analizlerde ise aynı şekilde düşük ilişki düzeyi tespit edildiği ( $p=0,267>0,05$ ); uyuşma ölçüsü olarak kappa değeri simetrik bir tablo olmadığından dolayı hesaplanamamıştır ki bu durumda örtüşme durumlarından bahsedilemeyeceği görülmektedir.



**Tablo 7.24** Çocuk Durumu-Müşteri Odaklılık Simetrik Ölçüler Tablosu

		Değer	Tahmini Standart Hata	Tahmini T	Tahmini Anlamlılık Değeri
Nominal-Nominal	Phi	,105			,285
	Cramer's V	,105			,285
	Kontenjans Tablosu	,105			,285
Ordinal-Ordinal	Kendall's tau-b	,105	,094	1,109	,267
	Kendall's tau-c	,090	,082	1,109	,267
	Gamma	,250	,225	1,109	,267
Geçerli Durum Sayısı			103		

### 7.7.2.. Sıralı Lojit Analiz Yöntemi Uygulaması

Öncelikle, Ordinal Logit Analiz Yöntemi'ne göre bir model oluşturulacaktır. Analiz neticesinde elde edilen STATA ekran çıktıları aşağıda ifade edilmiştir. Analizde her bir faktör için kullanılacak hipotez:

$H_0$ : Faktörün müşteri odaklılığı üzerinde etkisi yoktur.

Sıralı Lojit Analiz Yöntemi ile yapılan birinci denemede Faktör 1 ve Faktör 4'ün sıfır hipotezimizi gerçekleştirmediği yani model içerisinde tutulabileceği, ancak Faktör 2 ( $p=0.627>0.05$ ) ve Faktör 3'ün ( $p=0.223>0.05$ ) sıfır hipotezimizi gerçekleştirdiği yani modelden çıkarılmasının ve ikinci bir denemenin yapılmasının gerektiği görülmüştür.

İkinci denemede Faktör 1 ( $p=0.000<0.05$ ) ve Faktör 4 ( $p=0.043<0.05$ ) ile oluşturulacak modelin uygun olduğu tespit edilmiştir. Oluşturulan model ile verilerin % 36,33'ü temsil edilebilmektedir.

**Tablo 7.25** Sıralı Lojit Model Tablosu (1.Deneme)

Ordered logistic regression	Number of obs	=	102				
	LR chi2(4)	=	50.76				
	Prob > chi2	=	0.0000				
Log likelihood = -41.679232	Pseudo R2	=	0.3785				
+-----+-----							
Müş.Odaklılık	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]		
+-----+-----							
<b>fak_1</b>	<b>2.363208</b>	<b>.6347702</b>	<b>3.72</b>	<b>0.000</b>	<b>1.119081</b>	<b>3.607335</b>	
fak_2	.3341011	.6879711	0.49	0.627	-1.014297	1.6825	
fak_3	.5593022	.459031	1.22	0.223	-.3403819	1.458986	
<b>fak_4</b>	<b>1.978477</b>	<b>1.113746</b>	<b>1.78</b>	<b>0.076</b>	<b>-.2044247</b>	<b>4.161378</b>	
+-----+-----							
/cut1	18.20832	4.39683			9.590691	26.82595	
/cut2	26.0148	5.316712			15.59424	36.43537	

**Tablo 7.26** Sıralı Lojit Model Tablosu (2.Deneme)

Ordered logistic regression		Number of obs	=	102			
		LR chi2(2)	=	48.73			
		Prob > chi2	=	0.0000			
Log likelihood = -42.694792		Pseudo R2	=	<b>0.3633</b>			
-----+-----							
Müş.Odaklılık		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
-----+-----							
fak_1		<b>2.620357</b>	.6054986	4.33	<b>0.000</b>	1.433602	3.807113
fak_4		<b>2.213342</b>	1.095467	2.02	<b>0.043</b>	.066267	4.360417
-----+-----							
/cut1		16.93399	4.118727			8.861432	25.00654
/cut2		24.72571	5.101771			14.72642	34.725

### 7.7.3. Multinomial Logistic Regresyon Analiz Yöntemi Uygulaması

Bir önceki kısımda sıralı lojit yöntemi ile yapılan analizin aynısı Multinomial Logistic Regresyon Analiz Yöntemi ile yapılacaktır. Analizde yine her bir faktör için kullanılacak hipotez:

$H_0$ : Faktörün müşteri odaklılığı üzerinde etkisi yoktur.

Sıralı Lojit Analiz Yöntemi'nde olduğu gibi Multinomial Lojistik Regresyon Analiz Yöntemi'nde de yapılan birinci denemede Faktör 1 ve Faktör 4'ün sıfır hipotezimizi gerçekleştirmediği yani model içerisinde tutulabileceği, ancak Faktör 2 ( $p=0.322>0.05$ ) ve Faktör 3'ün ( $p=0.162>0.05$ ) sıfır hipotezimizi gerçeklediği yani modelden çıkarılmasının ve ikinci bir denemenin yapılmasının gerektiği görülmüştür.

İkinci denemede Faktör 1 ( $p=0.000<0.05$ ) ve Faktör 4 ( $p=0.000<0.05$ ) ile oluşturulacak modelin uygun olduğu tespit edilmiştir. Oluşturulan model ile verilerin %39,76'sı kapsamaktadır.

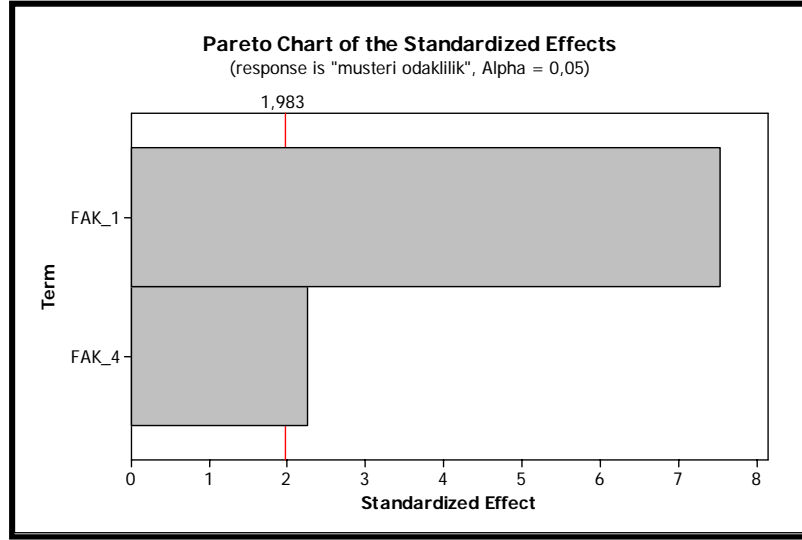
**Tablo 7.27** Multinomial Lojistik Regresyon Model Tablosu (1.Deneme)

Multinomial logistic regression		Number of obs	=	102			
		LR chi2(8)	=	65.93			
		Prob > chi2	=	0.0000			
Log likelihood = -34.098135		Pseudo R2	=	0.4915			
-----+-----							
Müş.Odaklılık		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
-----+-----							
3							
	fak_1	-2.21497	.6364007	-3.48	0.001	-3.462292	-.9676476
	<b>fak_2</b>	<b>-.6964761</b>	<b>.7027708</b>	<b>-0.99</b>	<b>0.322</b>	<b>-2.073881</b>	<b>.6809293</b>
	<b>fak_3</b>	<b>-.6784247</b>	<b>.4855914</b>	<b>-1.40</b>	<b>0.162</b>	<b>-1.630166</b>	<b>.2733171</b>
	fak_4	-17.9401	1.114414	-16.10	0.000	-20.12431	-15.75589
	_cons	67.28553					
-----+-----							
5							
	fak_1	68.86677	3.52e+07	0.00	1.000	-6.91e+07	6.91e+07
	fak_2	-77.80079	9.35e+08	-0.00	1.000	-1.83e+09	1.83e+09
	fak_3	.2737372	3.90e+07	0.00	1.000	-7.65e+07	7.65e+07
	fak_4	4.709989	9.45e+08	0.00	1.000	-1.85e+09	1.85e+09
	cons	-92.98005	4.46e+08	-0.00	1.000	-8.74e+08	8.74e+08

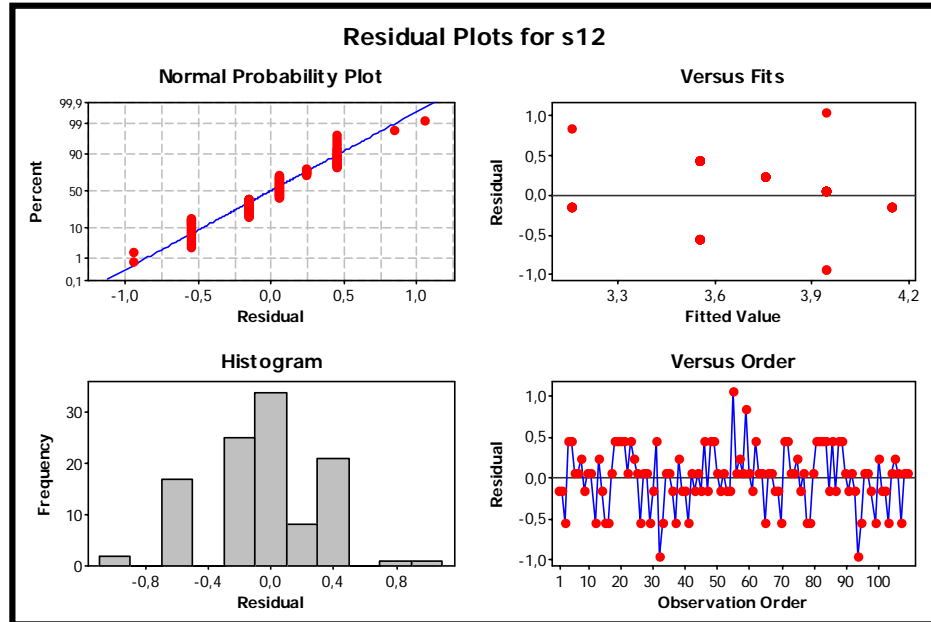
**Tablo 7.28** Multinomial Lojistik Regresyon Model Tablosu (2.Deneme)

Multinomial logistic regression		Number of obs	=	102			
		LR chi2(4)	=	53.32			
		Prob > chi2	=	0.0000			
Log likelihood = -40.39959		Pseudo R2	=	<b>0.3976</b>			
-----+-----							
Müş.Odaklılık		Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
-----+-----							
3							
fak_1		<b>-2.586516</b>	.608354	-4.25	<b>0.000</b>	-3.778868	-1.394164
fak_4		<b>-18.2795</b>	.8351025	-21.89	<b>0.000</b>	-19.91627	-16.64273
_cons		65.0605					
-----+-----							
5							
fak_1		17.54863	.2031575	86.38	0.000	17.15045	17.94681
fak_4		-19.10125					
_cons		-33.87153					

Oluşturulan modellerin uygunluğu ve grafikleri MINITAB ile yapılmış olup, Şekil 7.3'te gösterilmiştir. Modellerde görüldüğü üzere; analizler neticesinde etkin olduğu tespit edilen iki faktör grubu “Design of Experiment” (Deney Tasarımı) Pareto Grafiği ile yakalanmıştır.



Şekil 7.3 Sıralı Lojit Modeli



Şekil 7.4 Sıralı Lojit Model Dörtlü Uygunluk Test Grafiği



Analizin birinci basamağında % 80 güven aralığıyla ana etkilerin etkileri baz alınarak çalışılmış; faktör 2 ve faktör 3 değişkenlerinin model üzerinde hiçbir etkisinin olmadığı görülmüştür. Diğer basamağında ise iki değişkene indirgenen model üzerinde yine ana etkilerinin etkisi analiz edilmiş ve her iki faktörün de modelde etkin olduğu görülmüştür.

Modelin uygunluk testi yine MINITAB ile dörtlü analiz yöntemi ile model artıkları üzerinden analiz edilmiştir (Şekil 7.4). Modelin artıkları üzerinde yapılan analiz neticesinde modelin uygun olduğu şekil üzerinden de görülmektedir.

## SONUÇ VE ÖNERİLER

Kategorik veri analiz yöntemlerinin incelenmesi amacıyla yapılan çalışma kapsamında farklı alt başlıklarda bu alandaki analiz yöntemlerinin tamamına yer verilmiştir. Bölümlerde analiz yöntemlerinin kullanım alanları, kullanım şekilleri, değerlendirme kriterleri ifade edilmiştir.

Çalışmanın uygulama bölümünde ise; kategorik veri analiz yöntemleri içerisinde en fazla kullanım yeri bulan başlıca iki veri analiz yöntemini içeren bir anket uygulamasına yer verilmiştir. Ankette katılımcılara, anketin uygulama alanına göre belirlenmiş olan beşli likert ölçekli ve kapalı uçlu soru tipindeki sorular yöneltilmiştir. Anket sorularına verilen cevaplar neticesinde elde edilen veriler olumsuzluk tabloları, sıralı lojit analiz yöntemi ve multinomial regresyon analiz yöntemi ile analiz edilmiştir. Analizlerde her iki analiz yöntemiyle de aynı model oluşturulmuştur. Uygulanan modellerin Anova analiz yönteminin kategorik modeller için kullanılan şekli olduğu bilinmektedir. Modellerde etki eden faktörler aynı olmasına rağmen, değişken katsayıları ve verileri ifade edebilme yüzdelerinin ( $R^2$ ) farklı olduğu görülmüştür. Ancak Ki-Kare analizleri neticesinde modellerde yakalanan ilişkiler net olarak yakalanamamıştır. Bu durum ise Ki-Kare analizlerine kıyasla uygulanan iki model analizinin daha etkin olduğunu göstermektedir.

Bahse konu iki analiz yönteminin birbirlerine göre üstünlüklerinden bahsedebilmenin çok zor olduğu aşıkardır. Ancak oluşan modellerde; birinci modelde sıralı lojit ile verilerin yaklaşık % 36'sı ifade edilebilirken, ikinci modelde multinomial regresyon analiz yöntemi ile yaklaşık % 40'ı ifade edilebilmektedir. Modellere katsayılarına göre bakıldığında ise; sıralı lojit modelinin pozitif gradyene sahip oluşu, multinomial regresyon modelinin ise negatif gradyene sahip olduğu görülmektedir.

Sonuç olarak; kategorik veri analiz yöntemleri detaylı olarak gösterilmiştir. Ayrıca; uygulama bölümü sonucunda, kategorik verinin ordinal veya nominal olmasına bağlı olarak seçim yapılan analiz yöntemleri arasında kayda değer bir fark tespit edilememiş, ordinal veriler için uygulanan sıralı lojit analiz yöntemi yerine nominal verilerin analizinde kullanılan multinominal regresyon analiz yönteminin de kullanılabileceği gösterilmiştir.

## KAYNAKÇA

Acar, F. (2000). Çapraz Tabloların Çözümlemesinde Logaritmik Doğrusal Modellerin Kullanımı. IV. Ulusal Ekonometri ve İstatistik Sempozyumu Kitabı (ss.783-799). Antalya.

Agresti, A. (2002). *Categorical Data Analysis*. New York: John Wiley & Sons

Agresti, A. (1996). *An introduction to categorical data analysis*. New York: John Wiley & Sons.

Anderson, E. B. (1990). *The Statistical Analysis of Categorical Data*. Berlin: Springer Verlag.

Akçil, M. ve Karağaoğlu, E. (2001). Meta-Analizinde İstatistiksel Yöntemler. *Hacettepe Tıp Dergisi*, 32 (4): 362-369.

Arabacı, Ö. (2002). *Lojistik Regresyon Analizi ve Bir Uygulama Denemesi*. Yayımlanmamış Yüksek Lisans Tezi. Bursa: Uludağ Üniversitesi Sosyal Bilimler Enstitüsü.

Bircan, H., Coşkun, A., Coşkun, S. Ve Kartal, M. (2004). Lojistik Regresyon Analizinin İncelenmesi ve Diş Hekimliğinde Bir Uygulaması. *Cumhuriyet Üniversitesi Diş Hekimliği Fakültesi Dergisi*, 7 (1):41-49

Bircan, H. (2004). Lojistik Regresyon Analizi: Tıp Verileri Üzerine Bir Uygulama. *Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 2 : 185-208 .

Breslow, N. ve Day, N. E. (1980). *Statistical Methods in Cancer Research, Vol. 1: The Analysis of Case- Control Studies*, Lyon:IARC.

Czado, C. (2003). Overdispersion in Logistic Regression. . [http://odin.stochastik.math.uni-goettingen.de/zfs/past\\_courses/GLMcourse/lec5.pdf](http://odin.stochastik.math.uni-goettingen.de/zfs/past_courses/GLMcourse/lec5.pdf) (13.04.2007)

Eyduran, E., Özdemir, T. ve Küçük, M. (2005). Hayvancılıkta Kategorik Verilerde Ki-Kare ve G Testi. *Yüzüncü Yıl Üniversitesi Veterinerlik Fakültesi Dergisi*. 16(1):1-3.

Elhan, A. H. ve Atakurt, Y. (2005). Ölçeklerin değerlendirilmesinde niçin Rasch analizi kullanılmalıdır?. Ankara Üniversitesi Tıp Fakültesi Mecmuası, 58(1):47-50.

Everit, B.S. (1992). *The Analysis of Contingency Tables*. Second Edition. London:: Chapman&Hall

Garson, D. (1999). Log-Linear, Logit and Probit Models. <http://www2.chass.ncsu.edu/garson/pa765/logit.htm> . (13 Nisan 2007).

Gujarati, D. N. (1995). *Basic Econometrics*. Third Edition. NY: McGraw – Hill Inc.

Heijden, P.G.M., Mooijaart, A., ve Takane, Y. (1994). Correspondence analysis and contingency table models. New York: Academic Press.

Hosmer, D. W. ve Lemeshow, S. (2000). Applied Logistic Regression. NY: John Wiley and Sons.

Işığışok, E. (2003). Bebeklerin Doğum Ağırlıklarını ve Boylarını Etkileyen Faktörlerin Lojistik Regresyon Analizi ile Araştırılması. *IV. Ulusal Ekonometri ve İstatistik Sempozyumu*, Gazi Üniversitesi, Ankara.

Lawles, J. F. (1987). Negative binomial and mixed poisson regression. *The Canadian Journal of Statistics*, 15(3): 209-225.

Le, C. T. (1998). *Applied Categorical Data Analysis*. New York: John Wiley & Sons.

Oktay, E. (2003). Kontenjans Tablolarından elde edilen ilişki ölçüleri. Ankara: Aktif yayın evi.

Liu, Z.(1992). Smoking ve Lung Cancer in China. *Intern. J. Epidemiol.*,21(2):197

Özdamar K. (1999). Paket Programlarla İstatistiksel Veri Analizi 1. Eskişehir: Kaan Kitapevi.

- Pampel, F. C. (2000). Logistic Regression. CA: Sage Publications.
- Powers, D. A. ve Xie, Y.(2000). Statistical Methods for Categorical Data Analysis. ABD: Academic Press,
- Simonoff,, J. S. (2003). *Analysis Categorical Data*. New York: Springer-Verlag.
- Stigler, S. (1986). The History of Statistics: the measurement of uncertainty before 1900. Cambridge, MA: Harvard Univ. Pres.
- Stokes, M. E, Davis, C.S ve Koch, G.G. (1991). *Categorical Data Analysis: Using the SAS System*. Cary, NC: SAS Institute Inc.
- Therneau, T.M. ve Grambsch, P.M. (2000). *Modeling Survival Data: Extending The Cox Model*. New York: Springer-Verlag.
- Uğurman, F. ve Berktaş, M. B. (2003). Tıpta Tanı Tesleri ve Risk Değerlendirmelerinde Kullanılan Temel Kavramlar. *Atatürk Göğüs Hastalıkları ve Göğüs Cerrahisi Eğitim ve Araştırma Hastanesi Dergisi*, 4: 235-240
- Uygun, H. (1990). Çapraz Tabloların Çözümlemesi ve Log-Linear Modeller. *H.Ü.İ.İ.B.F. Dergisi*, 8(1):299-308.
- Yay, M., Çoker, E. ve Uysal, Ö. (2007).Yaşam Analizinde Cox Regresyon Modeli ve Artıkların İncelenmesi. *İstanbul Üniversitesi Cerrahpaşa Tıp Fakültesi Dergisi*, 38:139-145.
- Yeşilova, A. (2003). Poisson Regresyonunda Meydana Gelen Aşırı Yayılımın İncelenmesi. Yayınlanmış Doktora Tezi, VAN: Yüzüncü Yıl Üniversitesi, Ziraat Fakültesi, Zootekni Bölümü.
- Yılmaz, V. ve Şıklar, E. (2002). İntiharlarla ilgili çok değişkenli kategorik verilerin analizinde logaritmik doğrusal modellerin kullanılması. *Anadolu Üniversitesi Bilim ve Teknoloji Dergisi*, 3 (2) :271-280.
- Zelderman, D. (1999). Models for Discreate Data. Clarependon Press. Oxford.

**EK**

**REKABET ŞARTLARINI DEĞERLENDİRME ANKETİ**

<b>Cazibe Değişkenleri</b>	<b>Durum</b>				
	<b>Çok Kötü</b>	<b>Kötü</b>	<b>Vasat</b>	<b>İyi</b>	<b>Çok İyi</b>
S.1.Fiyat Aralığı					
S.2.Ziyaret Sıklığı					
S.3.Değişik Ürün Bulabilme					
S.4.Ödeme Koşulları					
S.5.Kaliteli Marka Bulabilme					
S.6. Görsellik					
S.7.Ulaşım Kolaylığı					
S.8.Çevre Entegrasyonu (Cafe, Sinema)					
S.9.Hizmet Performansı					
S.10.Ürün Çeşitliliği					
S.11.Moda Takibi					
S.12.Müşteri Odaklılık					
S.13.Kişiyeye Özel Ürün Seçenekleri Sunmak					
S.14.Beden Kalıp Performansı					
S.15.Özel Bedenlerin Bulunması					

Cinsiyetiniz ?

KADIN

ERKEK

Çocuğunuz var mı?

YOK

VAR