DOKUZ EYLÜL UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

ESTIMATION OF PARAMETER WITH KNOWN COEFFICIENT OF VARIATION

by

Ayşe ÜNSAL

July, 2010

İZMİR

ESTIMATION OF PARAMETER WITH KNOWN COEFFICIENT OF VARIATION

A Thesis Submitted to the

Graduate School of Natural and Applied Sciences of Dokuz Eylül University

In Partial Fulfillment of the Requirements for the Degree of Master of Science

In Statistics Program

by

Ayşe ÜNSAL

July, 2010

İZMİR

M.Sc THESIS EXAMINATION RESULT FORM

We have read the thesis entitled **'ESTIMATION OF PARAMETER WITH KNOWN COEFFICIENT OF VARIATION'** completed by **AYŞE ÜNSAL** under the supervision of **ASSIST. PROF. DR. ÖZLEM EGE ORUÇ** with the contribution of **PROF. DR. GÖTZ TRENKLER** as the co-supervisor. We certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

………………………………………

Assist. Prof. Dr. ÖZLEM EGE ORUÇ

─────────────────────────────

Supervisor

────────────────────────   ─────────────────────────────

(Jury Member)   (Jury Member)

───────────────────────────────

Prof. Dr. Mustafa SABUNCU

Director

Graduate School of Natural and Applied Sciences

ACKNOWLEDGEMENTS

ESTIMATION OF PARAMETER WITH KNOWN COEFFICIENT OF VARIATION

ABSTRACT

Improved estimation is an interesting and also important concept in statistical inference. This concept is defined by explaining all the factors that have an effect on the improvement to be gained in this thesis. With the help of former studies such as Searls (1964), Bibby and Toutenburg (1972), Thompson (1968) and Hirano (1973), improved estimation techniques were clarified.

In this study, using coefficient of variation as prior information improved estimators for mean, variance and ratio parameters are obtained. These biased estimators were proposed as alternatives to the usual unbiased estimators. In order to make the comparisons between usual unbiased estimators and improved estimators mean square error (MSE) is used as the criterion function. Relative efficiency values are presented in tables to prove that the suggested improved estimators are more efficient than the unbiased estimators.

Key words: improved estimation, mean square error, efficiency, coefficient of variation, bias.

DEĞİŞİM KATSAYISININ BİLİNMESİ DURUMUNDA PARAMETRE TAHMİNİ

ÖZ

İyileştirilmiş tahminleme istatistiksel çıkarsamada ilginç ve bir o kadar da önemli bir kavramdır. Bu çalışmada, iyileştirme üzerinde etkisi bulunan tüm faktörlerin açıklanmasıyla iyileştirilmiş tahminleme tanıtılmıştır. Searls (1964), Bibby ve Toutenburg (1972), Thompson (1968) ve Hirano (1973) gibi daha önce yapılmış çalışmaların da yardımıyla iyileştirilmiş tahminleme teknikleri açıklanmıştır.

Önsel bilgi olarak değişim katsayısının kullanılmasıyla ortalama, varyans ve oran parametreleri için iyileştirilmiş tahmin ediciler elde edilmiştir. Bu yanlı tahmin ediciler, söz konusu parametreler için bilinen yansız tahmin edicilere alternatif olarak önerilmiştir. Yansız tahmin ediciler ile alternatifi olan iyileştirilmiş tahmin ediciler arasında karşılaştırma yapmak için hata kareler ortalaması (MSE) ölçüt alınmıştır. Önerilen tahmin edicilerin yansız tahmin edicilerden daha etkin olduğunu gösteren göreli etkinlik değerleri tablolar halinde verilmiştir.

Anahtar sözcükler: iyileştirilmiş tahminleme, hata kareler ortalaması, etkinlik, değişim katsayısı, yanlılık.

CONTENTS

CHAPTER ONE

INTRODUCTION

To know about future, the time to come has always been interesting and popular independently from time. The requirement of prediction or estimation of unknown entities could occur in several areas of science and daily life quite frequently in scientific and/or unscientific ways.

Although it is indispensable to predict the unknown, it is not exactly an easy task to do. At this point statistical inference provides a scientific solution, which is not flawless but could be improved. An improved method is the main subject of this study. But before explaining what should be done to gain an improvement, statistical inference and especially one of its parts; estimation should be defined properly.

Statistical inference covers two main parts; estimation and tests of hypotheses. Estimation is to obtain a value or an interval for an unknown parameter θ, which belongs to the parameter space of population, utilizing the observed sample values. Casella & Berger (2002) define an estimator briefly as any function $f(X_1,...,X_n)$ of a sample. If it is desired to determine a statistic for the unknown parameter, then this statistic is called a point estimator. To determine two statistics as the lower and upper bounds for the unknown parameter is called interval estimation.

Point estimation deals with two matters. The first problem is to obtain a statistic that could be used as an estimator and the second problem is the selection of the criterion to compare estimators and determine best estimator among all possible choices. This thesis focuses on a special case of the point estimation, which is called improved estimation. The goal of this estimation will be given in the following chapter including its method. Throughout the whole study, mean square error (MSE) is held as the criterion function

to make the comparisons. And the relative efficiency values are presented as the proof of that the improved estimators are more efficient alternatives to usual unbiased estimators.

A satisfactory definition of the improved estimation concept is given in the following section. This section also includes the method and the goal of gaining improved estimators.

Chapter 3 contains examples for improved estimators for mean, variance and ratio parameters for various distributions.

The last chapter is the conclusion of this thesis. A brief summary of the whole study and the obtained results are also given in this chapter.

CHAPTER TWO

IMPROVED ESTIMATION

As noted in the previous section, a detailed definition for improved estimation will be given in this section divided in two parts. The method to obtain an improvement and the best improved estimator will be explained in these parts, respectively.

It is possible to improve distribution parameters, to gain better estimators, by utilization of biased estimators. An important point besides the method to obtain new estimators is the choice of the criterion in order to compare estimators. Mean square error (MSE), which provides to observe the bias and variance simultaneously, will be the criterion to decide which estimator is better.

2.1 Improvement Region

Let $\hat{\theta}$ be an unbiased estimator of the distribution parameter θ with zero bias and known variance. Using these known properties another estimator, say $a\hat{\theta}$, is an improvement on $\hat{\theta}$ if $MSE(a\hat{\theta}) < MSE(\hat{\theta})$, where $a \in (0,1)$. Bibby & Toutenburg (1977) defined the 'improvement region', an interval, using the inequality $MSE(a\hat{\theta})/MSE(\hat{\theta}) < 1$. The improvement region is dependent on the estimator $\hat{\theta}$, the parameter θ, the function adapting $\hat{\theta}$ to $a\hat{\theta}$ and lastly the criterion function to make a comparison. In this thesis, the estimator $\hat{\theta}$ is used as an unbiased estimator of the parameters mean, variance and ratio which are $\bar{X}, S^2, \hat{p}$ respectively.

Suppose that $\hat{\theta}$ is unbiased for θ, and thus its variance coincides with its MSE. $MSE(a\hat{\theta})$ is calculated using these properties of $\hat{\theta}$ as follows

$$MSE(a\hat{\theta}) = Var(a\hat{\theta}) + [E(a\hat{\theta}) - E(\hat{\theta})]^2 = a^2\sigma^2 + (a\theta - \theta)^2. \tag{1}$$

If $a\hat{\theta}$ is an improvement over $\hat{\theta}$ in terms of MSE criterion, then the proportion of $\dfrac{MSE(\hat{\theta})}{MSE(a\hat{\theta})}$ should be greater than 1.

$$\frac{MSE(\hat{\theta})}{MSE(a\hat{\theta})} = \frac{\sigma^2}{a^2\sigma^2 + \theta^2(a-1)^2} > 1. \tag{2}$$

Solving the inequality (2) and substituting $v = \sigma/\theta$, the improvement region is obtained based on the coefficient of variation as follows

$$v^2\theta^2 > a^2v^2\theta^2 + \theta^2(a-1)^2$$

$$v^2 - a^2v^2 > (a-1)^2$$

$$v^2(1-a)(1+a) > (a-1)^2$$

Case I: $1-a > 0 \quad \Rightarrow \quad a(v^2+1) > (1-v^2)$

$$a > \frac{1-v^2}{1+v^2}. \tag{3}$$

Case II: $1-a < 0 \quad \Rightarrow \quad a > 1$. \hfill (4)

Inequality (4) is in contradiction with the assumption $a \in (0,1)$, this is why it is ignored. As a result inequality (3) provides a lower bound for a based on the coefficient of variation v. Now, the improvement region is obtained as follows

$$\frac{1-v^2}{1+v^2} < a < 1. \tag{5}$$

Inequality (5) means that, when the coefficient of variation of the population is known, any value between $\dfrac{1-v^2}{1+v^2}$ and 1 provides an improvement; a better estimator with smaller MSE than the unbiased estimator, the case where a is equal to 1.

Considering the length of the improvement region, a long interval would include more possible values to yield smaller MSE values compared with the original estimator.

On the other hand, a short interval is more beneficial to find the exact value, which provides the minimum MSE. (G. Trenkler, private correspondence, 16 March 2010)

Next, we will present a table that consists of the lower bound of a for different values of the coefficient of variation v.

Table 2.1 Lower bound of the improvement region for various values of v

v	0	0.50	1.00	1.50	2.00	2.50	5.00	7.50	10.00
$\dfrac{1-v^2}{1+v^2}$	1	0.60	0.00	-0.38	-0.60	-0.72	-0.92	-0.96	-0.98

Lower bound values decrease as the coefficient of variation gets larger. As v approaches to zero, which means that either the population variance approaches zero or the population mean approaches infinity the improvement region is expected to vanish.

2.2 Best Improved Estimation

A point in the improvement region provides the minimum mean square error, so that it is also possible to find the value of a which minimizes the mean square error of $a\hat{\theta}$. The estimator with minimum MSE was called by Bibby & Toutenburg (1977) as minimum mean square error estimator (MIMSEE). To determine this point, we will simply derive the mean square error function of the estimator $a\hat{\theta}$ and find the value of a for which the derivative vanishes.

$$\frac{\partial MSE(a\hat{\theta})}{\partial a} = 2a\sigma^2 + 2(a-1)\theta^2. \qquad (6)$$

The value of a with the minimum MSE will be denoted by a^* throughout this study. Substitution of the coefficient of variation in equation (6) gives

$$a^* = \frac{1}{1+v^2}. \qquad (7)$$

Note that, a^* is the midpoint of the improvement region (5).Using the value given in (7) $MSE(a\hat{\theta})$ can be calculated. The variance of the improved estimator is

$$Var(a^*\hat{\theta}) = \frac{1}{(1+v^2)^2}Var(\hat{\theta}) = \sigma^2/(1+v^2)^2 .$$

(8)

The squared bias of the estimator $a^*\hat{\theta}$ is given by

$$[E(a^*\hat{\theta}) - E(\hat{\theta})]^2 = (\frac{1}{1+v^2}\theta - \theta)^2 = \theta^2 v^4(1+v^2)^{-2} = \sigma^2 v^2(1+v^2)^{-2} .$$

(9)

since $\hat{\theta}$ is unbiased for θ, where $v = \sigma/\theta$. The sum of variance and squared bias is given by equations (8) and (9) respectively gives the total MSE for the improved estimator.

$$MSE(a^*\hat{\theta}) = \frac{\sigma^2}{(1+v^2)^2}(1+v^2) = \sigma^2/(1+v^2) .$$

(10)

Clearly, MSE of the improved estimator is $a^* = \frac{1}{1+v^2}$ times smaller than the MSE of the original unbiased estimator $\hat{\theta}$. To prove the improvement gained with the proposed estimator, relative efficiency value is calculated as

$$RE = \frac{MSE(\hat{\theta})}{MSE(a^*\hat{\theta})} = 1+v^2 > 1 .$$

(11)

The suggested estimator dominates the unbiased one with respect to MSE, because relative efficiency is greater than1. It should be noted that this improvement is gained through a known coefficient of variation as prior information.

Although it could seem like running away from the main subject of this study, which is the utilization of known coefficient of variation in the estimation procedure, a few important points should be also mentioned for the case that there is no certain information of the coefficient of variation. If v is unknown, this improvement method becomes inapplicable. But there are some circumstances; under those still an improvement can be achieved

1. If the original estimator, $\hat{\theta}$ in this case, has a specific distribution.

2. If there is a given interval including the coefficient of variation.

3. If there is prior information on the coefficient of variation.

Since this study is focused on the situation that the coefficient of variation is known, three circumstances are not investigated deeply (Bibby & Toutenburg, 1977, chap. 2.3, p. 26).

2.2.1 *Best Improved Estimator for Mean*

It is well known that, the sample mean is the unbiased estimator of the population mean. Searls (1964) proved that a biased estimator with smaller MSE for the mean can be obtained when coefficient of variation is known. Determining the optimum weight for the sum of observations by minimizing MSE is the method used in this study. In other words, the approach is to find an alternative to $1/n$ for $\sum_{i=1}^{n} x_i$ a factor with minimum MSE for a random sample of n observations $x_1, x_2, ..., x_n$.

$$\bar{x}' = a \sum_{i=1}^{n} x_i . \tag{12}$$

The goal is to evaluate a by minimizing the mean square error of estimator (12), which is $E(\bar{x}' - \mu)^2$ where μ and σ^2 are population mean and variance, respectively.

$$MSE(\bar{x}') = a^2 n \sigma^2 + \mu^2 (an-1)^2 . \tag{13}$$

Differentiating equation (13) with respect to a and setting equal to zero will give us the minimal value of the MSE function.

$$\frac{\partial MSE(\bar{x}')}{\partial a} = 2an\sigma^2 + 2\mu^2 n(an-1) . \tag{14}$$

$$\frac{\partial^2 MSE(\bar{x}')}{\partial a^2} = 2n\sigma^2 + 2n^2 \mu^2 . \tag{15}$$

The value of a which puts equation (14) to zero is the minimum point of the MSE function since the second derivative given by equation (15) is always positive. Substitution of the coefficient of variation $v = \sigma/\mu$ gives the optimum weight for the sum of observations by

$$a^* = \frac{1}{n + v^2}. \tag{16}$$

Note that, the statistic to be improved causes the differences between values of factors (7) and (16). Using the optimum weight, the alternative estimator and its MSE are obtained as

$$\bar{x}' = \frac{1}{n + v^2} \sum_{i=1}^{n} x_i, \tag{17}$$

$$MSE(\bar{x}') = \sigma^2 / (n + v^2). \tag{18}$$

Clearly, the MSE of the unbiased estimator is equal to its variance since it is the sum of variance and squared bias. Although it is obvious that the expression for the MSE in (18) is always smaller than the MSE of the unbiased estimator, the relative efficiency is calculated to prove that the suggested biased estimator is more efficient than the unbiased one.

$$RE = \frac{\sigma^2 / n}{\sigma^2 / (n + v^2)} = \frac{n + v^2}{n} = 1 + v^2 / n. \tag{19}$$

Since improved estimator dominates $\bar{x}$, the relative efficiency with respect to MSE is greater than 1. Below, we present relative efficiencies for various sample size and coefficient of variation values.

Table 2.2 Relative efficiencies of improved estimator $\bar{x}'$ for various values of n and v.

RE	Sample Size					
v	5	20	100	250	500	1000
1	1.20	1.05	1.01	1.00	1.00	1.00
2	1.80	1.20	1.04	1.02	1.00	1.00
3	2.80	1.45	1.09	1.04	1.01	1.01
4	4.20	1.80	1.16	1.06	1.03	1.02

Note that, relative efficiency increases as sample size decreases and coefficient of variation increases.

Another improved estimator by Thompson (1968) will be defined in Chapter 3. And we will analyze the relation between improved estimators proposed by Searls (1964) and Thompson (1968).

2.2.2 *Best Improved Estimator for Variance*

Using the value of kurtosis as prior information provides a better estimator, a biased but more efficient estimator for variance. Although we focus on utilizing coefficient of variation as prior information in the estimation procedure, in this section coefficient of kurtosis will be used. Additionally, in the next chapter a more efficient estimator for variance will be given using the coefficient of variation following another way.

Using the same approach described in the previous subsection, a more efficient estimator for variance is defined as an alternative to the usual unbiased estimator s^2 (Singh, Pandey & Hirano, 1973). For a random sample size of n, the estimator to be improved is defined as

$$r^2 = a\sum_{i=1}^{n}(X_i - \bar{X})^2 ,\qquad(20)$$

where $\bar{X}$ is the sample mean. Population mean and variance are μ and σ^2, respectively.

The goal is to obtain the optimum value for a by minimizing MSE of the proposed estimator r^2. Before that, in order to calculate MSE, variance and bias-square of r^2 are calculated and given respectively.

$$Var(r^2) = a^2(n-1)^2 Var(s^2) = \frac{a^2(n-1)^2}{n}[\mu_4 + \frac{(3-n)\sigma^4}{(n-1)}], \tag{21}$$

$$(Biasr^2)^2 = [a(n-1)\sigma^2 - \sigma^2]^2 = \sigma^4[a(n-1)-1]^2, \tag{22}$$

since $E(s^2) = \sigma^2$. The sum of variance (21) and bias-square (22) gives the MSE of the suggested estimator r^2.

$$MSE(r^2) = \frac{a^2(n-1)^2}{n}[\mu_4 + \frac{(3-n)\sigma^4}{(n-1)}] + \sigma^4[a(n-1)-1]^2. \tag{23}$$

The value of a that makes MSE minimal is given by

$$a^* = \frac{n}{n^2 - 2n + 3 + \beta_2(n-1)}. \tag{24}$$

β_2 is the coefficient of kurtosis and equals to $\dfrac{\mu_4}{\sigma^4}$. Finally the improved estimator by Singh et al (1973) is

$$r^2 = \frac{n}{n^2 - 2n + 3 + \beta_2(n-1)}\sum_{i=1}^{n}(X_i - \bar{X})^2. \tag{25}$$

Next, it will be proved that the new estimator is more efficient than the unbiased estimator by computing relative efficiency. Since s^2 is unbiased, its variance and MSE are equal.

$$MSE(s^2) = 1/n(\mu_4 + \frac{3-n}{n-1}\sigma^4). \tag{26}$$

Inserting the factor (24) in equation (23) yields the relative efficiency with respect to MSE.

$$RE(r^2) = \frac{n^2 - 2n + 3}{n(n-1)} + \frac{\beta_2}{n}. \tag{27}$$

In the subsequent table, we present relative efficiencies for various values of sample size and coefficient of kurtosis.

Table 2.3 Relative efficiencies of improved estimator r^2 for various values of n and β_2.

RE	Sample Size					
β_2	5	20	100	250	500	1000
1	1.10	1.01	1.00	1.00	1.00	1.00
2	1.30	1.06	1.01	1.00	1.00	1.00
3	1.50	1.11	1.02	1.01	1.00	1.00
4	1.70	1.16	1.03	1.01	1.01	1.00

The same situation in connection with the best improved estimator for mean arises also for variance. It is observed that the relative efficiency increases as sample size decreases and the coefficient of kurtosis increases.

Let $X_1, X_{2,}..., X_n$ be a random sample of size n from a population having a normal distribution with unknown mean μ and variance σ^2. We get the best improved estimator for variance of the normal distribution by substitution of normal distributions kurtosis in estimator (25).

$$r^2 = \frac{1}{n+1}\sum_{,=1}^{n}(X_i - \bar{X})^2 .$$
(28)

The estimator (28) can also be defined in terms of s^2.

$$r^2 = \frac{n-1}{n+1}s^2 .$$
(29)

Since variance and expected value of s^2 are $2\sigma^4/(n-1)$ and σ^2 respectively, $MSE(r^2)$ of a normal distribution is

$$MSE(r^2) = \frac{n-1}{(n+1)^2}2\sigma^4 + \left[\frac{n-1}{n+1}\sigma^2 - \sigma^2\right]^2 .$$
(30)

The relative efficiency of r^2 with respect to unbiased estimator is obtained as

$$RE(r^2) = \frac{MSE(s^2)}{MSE(r^2)} = \frac{2\sigma^4/(n-1)}{2\sigma^4/(n+1)} = \frac{n+1}{n-1} .$$
(31)

The improvement is demonstrated by relative efficiency values for various values of sample size and tabulated.

Table 2.4 Relative efficiencies of improved estimator r^2 for a normal distribution.

	Sample Size					
	5	20	100	250	500	1000
RE	1.50	1.11	1.02	1.01	1.00	1.00

As n increases, relative efficiency of r^2 is reduced. Again the suggested estimator is more efficient than s^2 when the sample sizes are small.

In Chapter 3, we will describe another way to derive the best improved estimator for variance of gamma family distributions such as exponential, laplace, etc.

2.2.3 *Best Improved Estimator for Proportion*

The proportion of a characteristic is defined as the number of individuals with this property divided by the total number of individuals.

$$\hat{p} = \frac{x}{n} \ . \tag{32}$$

Individuals with the characteristic are denoted by x and the total number of individuals is denoted by n. Suppose that a variable is created which is equal to 1 if the subject has the characteristic and 0 if not. The proportion of individuals with the characteristic is the mean of this variable because the sum of these 0's and 1's is the number of individuals with the characteristic.

MIMSEE for proportion which means to determine another scalar for x alternatively to $1/n$ is obtained proceeding in the same way as the former parameters mean and variance.

The improving estimator of proportion is defined as

$$\overset{\wedge}{p'} = ax, \tag{33}$$

for a random sample size of n from a population with unknown mean and variance μ and σ^2 respectively.

Next, to find the optimum weight for x in other words the value of a that minimizes MSE of estimator (33) variance and bias-square is calculated and given as follows

$$Var(\overset{\wedge}{p'}) = a^2\sigma^2, \tag{34}$$

$$(Bias\ \overset{\wedge}{p'})^2 = (a\mu - \mu)^2 = \mu^2(a-1)^2. \tag{35}$$

Clearly, summarizing equations (34) and (35) we get MSE of estimator $\overset{\wedge}{p'}$.

$$MSE(\overset{\wedge}{p'}) = a^2\sigma^2 + \mu^2(a-1)^2. \tag{36}$$

The value of a which makes $MSE(\overset{\wedge}{p'})$ minimal is derived by differentiating the MSE function dependent on a and equalizing the derivative to zero. As a result, the optimum weight for individuals with the characteristic in the sample and MIMSEE using this weight is

$$a^* = \frac{1}{1+v^2}, \tag{37}$$

$$\overset{\wedge}{p'} = \frac{1}{1+v^2}x, \tag{38}$$

respectively, where $v = \sigma/\mu$.

Let the distribution of the population be binomial. In this case, the estimator (32) is unbiased and has MSE equal to its variance, say σ^2. The MSE of the suggested biased estimator is obtained substituting equation (38) in equation (36). To make a comparison between the estimators $\hat{p}$ and $\overset{\wedge}{p'}$, the relative efficiency is calculated.

$$RE(\hat{p}') = \frac{\sigma^2}{\sigma^2/(1+v^2)} = 1+v^2.$$ (39)

Clearly, the estimator given in (38) is more efficient than the unbiased estimator for a binomial distribution since the relative efficiency (39) always exceeds 1.

In the following chapter, we will give some examples for the usage of best improved estimator for proportion in geometric and negative binomial distributions.

CHAPTER THREE
ALTERNATIVE IMPROVED ESTIMATORS

In this chapter, we give some examples about improved estimator for parameters of mean, variance and proportion.

Firstly, the alternative improved estimator by Thompson is explained and the relationship between improved estimators by Thompson (1968) and Searls (1964) is constructed. Thus a different way to improve estimators for parameters of mean is presented.

Next, a biased but more efficient estimator for the variance of gamma family distributions, gamma, exponential, laplace and chi-square distributions, is obtained.

Finally, efficient estimators of proportion are investigated for geometric and negative binomial distributions.

3.1 Alternative Improved Estimator for Mean

An improved estimator is suggested by Thompson (1968) by determining a factor for mean in order to gain shrinkage towards a natural origin, say μ_0 and reduce mean square error. Unlikely former studies, Thompson defined the estimator to be improved as

$$\hat{\mu} = c(\bar{x} - \mu_0) + \mu_0. \tag{40}$$

It is clear that the estimator (40) is equal to the mean, the usual estimator of the location parameter, where $c = 1$. But to yield a reduction in MSE the shrinkage factor c is given as

$$c = \frac{(\bar{x} - \mu_0)^2}{(\bar{x} - \mu_0)^2 + s^2/n}. \tag{41}$$

The statistics $\bar{x}$ and s^2 are unbiased estimators of mean and variance, respectively. Substitution of the shrinkage factor (41) in the proposed estimator $\hat{\mu}$ yields a general form for the shrunken estimator.

$$\hat{\mu} = \frac{(\bar{x} - \mu_0)^2}{(\bar{x} - \mu_0)^2 + s^2/n}(\bar{x} - \mu_0) + \mu_0. \tag{42}$$

Rewriting the proposed estimator (17) defined in Chapter 2 in dependence of the sample mean, it is clear to see that improved estimators are identical where the natural origin μ_0 of estimator (42) is equal to zero. Estimator (17) could also be written as

$$\bar{x}' = \frac{n}{n + v^2}\bar{x} \tag{43}$$

or equivalently

$$\bar{x}' = \frac{1}{1 + v^2/n}\bar{x}. \tag{44}$$

Thompson uses observed values to obtain coefficient of variation and forms the shrinkage factor for mean.

In this part, we will show that the same shrunken (improved) estimator (42) can be obtained by using a suitable pivotal quantity.

The utilization of a proper pivotal quantity as a shrinkage factor instead of the coefficient of variation leads to an improvement. The method is explained for normal and binomial distributions in the following.

3.1.1 Improved Estimation of Mean for Normal Distribution

Let $X_i's$ be normally distributed with mean μ_0 and variance σ^2. Rewriting estimator (42) in another form shows that it is possible to obtain an improved estimator using a suitable pivotal quantity instead of the coefficient of variation. We have

$$\hat{\mu} = c(\bar{x} - \mu_0) + \mu_0 = \frac{(\bar{x} - \mu_0)^2}{(\bar{x} - \mu_0)^2 + s^2/n}(\bar{x} - \mu_0) + \mu_0$$

$$= \frac{1}{1 + \dfrac{s^2/n}{(\bar{x} - \mu_0)^2}}(\bar{x} - \mu_0) + \mu_0 \tag{45}$$

$$= \frac{1}{1 + (t^{-1})^2}(\bar{x} - \mu_0) + \mu_0,$$

where $t = \sqrt{n}(\bar{x} - \mu_0)/s$ is the pivotal quantity. Substituting the inverse of the pivotal quantity in the shrinking factor instead of cv in equation (44), improved estimator of mean suggested by Searls (1964), yields the shrunken estimator suggested by Thompson (1968).

If the natural origin μ_0 equals zero, then the pivotal quantity is given by

$$t = \frac{(\bar{x} - 0)}{s/\sqrt{n}} = \frac{\sqrt{n}\,\bar{x}}{s}. \tag{46}$$

The shrunken estimator (42) becomes

$$\hat{\mu} = \frac{\bar{x}^2}{\bar{x}^2 + s^2/n}\bar{x} = \frac{1}{1 + \dfrac{s^2/n}{\bar{x}^2}} = \frac{1}{1 + (t^{-1})^2}\bar{x}. \tag{47}$$

Clearly, shrunken estimators can be derived using the inverse of the pivotal quantity instead of the coefficient of variation in the shrinkage factor for mean.

3.1.2 *Improved Estimation of Proportion for Binomial Distribution*

Let X_i's have a binomial distribution with parameters n and p. An unbiased estimator of p is denoted by $\hat{p}$. The mean and standard deviation of sampling distribution of $\hat{p}$ is $\mu_p = p_o$ and $\sigma_p = \sqrt{\dfrac{p_0(1-p_0)}{n}}$, respectively. Proceeding in the same way with mean, the estimator of proportion is defined as

$$\hat{p} = c(\bar{x} - p_0) + p_0. \tag{48}$$

The suggested estimator is constructed by shrinking the mean towards p_0. Clearly the estimator (48) equals to sample mean when $c = 1$. In order to achieve that the shrinkage factor c is given by

$$c = \frac{(\bar{x} - p_0)^2}{(\bar{x} - p_0)^2 + p_0(1-p_0)/n}. \tag{49}$$

Consequently, the shrunken estimator $\hat{p}$ is defined as follows

$$\hat{p} = \frac{(\bar{x} - p_0)^2}{(\bar{x} - p_0)^2 + p_o(1-p_0)/n}(\bar{x} - p_0) + p_0. \tag{50}$$

The same estimator can also be constructed by using the pivotal quantity for proportion, that is

$$\begin{aligned}
\hat{p} &= \frac{(\bar{x} - p_0)^2}{(\bar{x} - p_0)^2 + p_0(1-p_0)/n}(\bar{x} - p_0) + p_0 \\
&= \frac{1}{1 + \dfrac{p_0(1-p_0)}{n(\bar{x} - p_0)^2}}(\bar{x} - p_0) + p_0 \\
&= \frac{1}{1 + (t^{-1})^2}(\bar{x} - p_0) + p_0,
\end{aligned} \tag{51}$$

where c, the shrinking factor, equals $\dfrac{1}{1+(t^{-1})^2}$.

As a result, it is also easy here to see that using the inverse of the pivot instead of the cv in the shrinking factor gives a shrunken estimator.

A relation between the shrunken estimators suggested by Searls (1964) and Thompson (1968) is constructed using a suitable pivotal quantity. We observe that any pivotal quantity for location parameter can be used instead of the coefficient of variation to gain an improvement, a reduction of MSE.

3.2 Efficient Estimator for Variance of Gamma Family Distributions

Let $X_1, X_2, ..., X_n$ be a random sample of size n from a population having a gamma distribution with unknown parameters α and λ .

$$X_i \sim Gamma(\alpha, \lambda). \tag{52}$$

Since S^2 is unbiased for variance, its expected value equals to variance of the gamma distribution.

$$E(S^2) = \alpha\lambda^2. \tag{53}$$

In order to calculate the variance of the unbiased estimator, the relationship between the fourth central moment and kurtosis is used. The general form of the variance for S^2 is

$$Var(S^2) = 1/n(\mu_4 + \frac{3-n}{n-1}\sigma^4), \tag{54}$$

where μ_4 is the fourth central moment and σ^4 is the square of the variance (Mood, Graybill & Boes 1974). Substituting the values μ_4 and σ^4 in equation (54) yields

$$Var(S^2) = \frac{\lambda^4\alpha}{n}\left(\frac{2n\alpha + 6n - 6}{n-1}\right). \tag{55}$$

Since S^2 is unbiased for σ^2, formula (55) gives also the variance of S^2. To find the MIMSEE for the gamma distribution the coefficient of variation v is needed.

$$v = \frac{\sqrt{Var(S^2)}}{E(S^2)}. \tag{56}$$

The coefficient of variation is obtained in a form of the shape parameter α and sample size n by substituting $E(S^2)$ and $Var(S^2)$ in equation (56), i.e.

$$v = \sqrt{\frac{2}{n-1} + \frac{6}{n\alpha}}. \tag{57}$$

Using equation (57), we obtain a^*, which will lead us to MIMSEE, an alternative to the unbiased estimator of the variance of the gamma distribution. The alternative estimator to S^2 is

$$
\begin{aligned}
a^* S^2 &= \frac{n\alpha(n-1)}{n\alpha(n+1) + 6(n-1)} S^2 = \frac{n\alpha}{n\alpha(n+1) + 6(n-1)} \sum_{i=1}^{n} (X_i - \bar{X})^2 \\
&= \frac{1}{(n+1) + 6(n-1)/n\alpha} \sum_{i=1}^{n} (X_i - \bar{X})^2
\end{aligned}
\tag{58}
$$

This estimator does not only depend on the observations and sample size but also on the shape parameter α.

The estimator $a^* S^2$ is a special case of the minimum variance biased estimator of variance derived by J. Singh et al (1972) which was noted in Chapter 2 as the best improved estimator for variance. The authors found the optimum weight for the total sum of squares about the mean as a function of the coefficient of kurtosis β_2 and sample size n. Substituting the coefficient of kurtosis of the gamma distribution in equation (25) provides the estimator $a^* S^2$, just because both estimators were aimed at the minimum mean square error.

Although a comparison between MIMSEE for variance and its unbiased estimator was made in the previous chapter, we will also compare estimator $a^* S^2$ and S^2 to interpret the results not based on the kurtosis β_2 but on the shape parameter of the gamma distribution. This will also provide a perspective for the other distributions related to gamma.

Again, the biased and unbiased estimators will be compared by using their MSE's. The equation (55) is also MSE of S^2. The mean square error for a^*S^2 is given by

$$MSE(a^*S^2) = Var(a^*S^2) + \left[E(a^*S^2) - E(S^2)\right]^2 = \frac{\alpha^2\lambda^4(2n\alpha + 6n - 6)}{n\alpha(n+1) + 6(n-1)}. \tag{59}$$

Accordingly, the relative efficiency based on MSE is

$$RE(a^*S^2) = \frac{MSE(S^2)}{MSE(a^*S^2)} = \frac{n+1}{n-1} + \frac{6}{n\alpha}. \tag{60}$$

If the suggested estimator a^*S^2 dominates the usual unbiased estimator of the variance respect to the MSE, then clearly the relative efficiency should exceed 1. Solving the inequality, we received an interesting result. The expression of relative efficiency (60) exceeds 1, if only the shape parameter α is greater than $3/n-3$, which means that the suggested improved estimator a^*S^2 is always a better estimator than the unbiased estimator of the variance.

We present the relative efficiencies in the table 1 for various sample sizes and shape parameter in order to observe the changes in RE.

Table 3.1 Relative efficiencies $RE(a^*S^2)$ for different values of α and n

RE	Sample size n				
α	5	10	20	50	100
1/2	3.90	2.42	1.71	1.28	1.14
1	2.70	1.82	1.41	1.16	1.08
2	2.10	1.52	1.25	1.10	1.05
3	1.90	1.42	1.20	1.08	1.04
10	1.62	1.28	1.14	1.05	1.03
20	1.56	1.25	1.12	1.04	1.02

The largest RE values are obtained with small sample sizes. As a second result related to parameter α; we can say that as α increases, RE decreases. High relative efficiency values are observed with small sample sizes and small values of shape parameter for a gamma distribution.

3.2.1   Efficient Estimator for Variance of Exponential Distribution

Let $X_1, X_2, ..., X_n$ be a random sample of size n from a population having exponential distribution with unknown parameter λ.

$$X_i \sim Exp(\lambda). \tag{61}$$

Proceeding in the same way as described for the gamma distribution using the variance and the coefficient of kurtosis for exponential distribution, leads to the improved estimator for the variance of an exponential distribution. The same estimator can also be obtained by substituting $\alpha = 1$ in the equation of a^*S^2 provided by the relationship between gamma and exponential distributions. MIMSEE for the variance of an exponential distribution is given by

$$a^*S^2 = \frac{n}{n^2 + 7n - 6} \sum_{i=1}^{n} (X_i - \bar{X})^2. \tag{62}$$

Since the relative efficiency for the gamma distribution is calculated and given in the previous subsection, it is not needed to calculate it all over again for exponential distribution. The gamma distribution becomes an exponential distribution when the shape parameter equals 1. Hence second row of Table 3.2.1 ($\alpha = 1$) shows the relative efficiency values of estimator (62) with respect to unbiased estimator of variance. The largest values of $RE(a^*S^2)$ are obtained with small sample sizes.

3.2.2 Efficient Estimator for Variance of Laplace Distribution (Double exponential distribution)

Let $X_1, X_2, ..., X_n$ be a random sample of size n from a population has Laplace distribution with unknown parameter λ. The alternative estimator to S^2 for Laplace distribution is defined as

$$a^* S^2 = \frac{n}{n^2 + 4n - 3} \sum_{i=1}^{n} (X_i - \bar{X})^2 \,. \tag{63}$$

Due to the relationship between gamma and laplace distributions the third column of Table 3.2.1 (where $\alpha = 2$) yields the relative efficiencies for the improved estimator of the Laplace distribution for the variance with respect to the usual unbiased estimator. Clearly, the suggested estimator (63) is more efficient than the unbiased alternative.

3.2.3 Efficient Estimator for Variance of Chi-square Distribution

MIMSEE for variance of chi-square distribution can be found either by substituting $\alpha/2$ in equation of $a^* S^2$, which is the best improved estimator for the gamma distribution, or by proceeding in the same way described in section 3.2. It should be noted that this time kurtosis of chi-square distribution will be used.

The improved estimator of the variance for a sample size of n from a population with chi-square distribution and its relative efficiency with respect to unbiased estimator are given, respectively, as

$$a^* S^2 = \frac{1}{(n+1) + 12(n-1)/n\alpha} \sum_{i=1}^{n} (X_i - \bar{X})^2 \,, \tag{64}$$

$$RE(a^* S^2) = \frac{n+1}{n-1} + \frac{12}{n\alpha} \,. \tag{65}$$

In Table 3.2.3.1, we present the relative efficiency of the estimator $a^* S^2$ with respect to S^2 for different values of n and α.

Table 3.2. Relative efficiencies $RE(a^* S^2)$ for various values of n and α

RE	Sample size n				
α	5	10	20	50	100
1	3.90	2.42	1.71	1.28	1.14
2	2.70	1.82	1.41	1.16	1.08
10	1.74	1.28	1.17	1.06	1.03
20	1.62	1.25	1.14	1.05	1.02

The suggested estimator (64) is more efficient than the unbiased estimator for the variance of a chi-square distribution like the former three distributions. Clearly, the largest values are gained for small sample sizes and high values of α. The same RE values are obtained when the shape parameter of the gamma distribution is one half of the α values from the chi-square distribution.

3.3 Efficient Estimator of Proportion for Geometric and Negative Binomial Distributions

An alternative estimator to the unbiased estimator of proportion is obtained for the binomial distribution with minimum mean square error using coefficient of variation as prior information in previous chapter. The original unbiased estimator of binomial distribution is

$$\hat{p} = \frac{x}{n} . \tag{66}$$

The improving estimator is defined as

$$\hat{p}' = ax . \tag{67}$$

Minimizing its MSE function based on a, the best improved estimator of proportion (MIMSEE) for the binomial distribution is obtained for

$$a^* = \frac{1}{1 + v^2} . \tag{68}$$

The alternative efficient estimator of proportion is defined as

$$\hat{p}' = \frac{1}{1+v^2}x. \tag{69}$$

But the estimator (69) does not provide the minimum point of MSE for geometric or negative binomial distributions since $\hat{p}$ is biased for these distributions. Thus, minimum mean square error estimator for geometric and negative binomial distributions is investigated individually.

   The MSE of $\hat{p}'$ is calculated as follows by using the properties of expected value and variance of a geometric distribution.

$$MSE(\hat{p}') = Var(\hat{p}') + [E(\hat{p}') - p]^2 = a^2 Var(x) + [aE(x) - p]^2, \tag{70}$$

where $E(x) = 1/p$ and $Var(x) = \dfrac{1-p}{p^2}$. The value of a, which makes $MSE(ax)$ minimal, is found by equalizing first derivative to zero.

$$\frac{\partial MSE(ax)}{\partial a} = \frac{2a(1-p)}{p^2} + 2\left(\frac{a}{p} - p\right)\frac{1}{p}$$
$$\frac{\partial^2 MSE(ax)}{(\partial a)^2} = \frac{2(1-p)}{p^2} + \frac{2}{p^2} > 0 \tag{71}$$

Since the second derivative is positive for any value of $p \in (0,1)$, first derivative provides the minimum mean square error estimator and this value is given by

$$a^* = \frac{p^2}{1+q}, \tag{72}$$

where $q = 1 - p$. Hence the MIMSEE of proportion for a geometric distribution is given by

$$\hat{p}' = \frac{p^2}{1+q}x. \tag{73}$$

Clearly this form of estimator is not applicable, thus substitution of the coefficient of variation v is needed. The coefficient of variation for geometric distribution and suggested estimator (73) in a form of coefficient of variation are given, respectively as

$$v = \frac{\sqrt{Var(x)}}{E(X)} = \frac{\sqrt{(1-p)/p^2}}{1/p} = \sqrt{q} \, , \tag{74}$$

$$\hat{p}' = \frac{(1-v^2)^2}{1+v^2} x \, . \tag{75}$$

The MSE of $\hat{p}'$ is calculated this time for the negative binomial distribution where the degrees of freedom k are held fixed.

$$X \sim NB(k, p) \, . \tag{76}$$

It is known that the expected value and variance of a negative binomial distribution with parameters k and p are $E(X) = k/p$, $Var(X) = \dfrac{k(1-p)}{p^2}$ respectively.

$$MSE(\hat{p}') = \frac{a^2 k(1-p)}{p^2} + \left(\frac{ak}{p} - p \right)^2 \, . \tag{77}$$

The constant a is determined to minimize MSE of $\hat{p}'$. We have

$$\frac{\partial MSE(\hat{p}')}{\partial a} = \frac{2ak(1-p)}{p^2} + \frac{2ak^2}{p^2} - 2k \, , \tag{78}$$

$$a^* = \frac{p^2}{k+q} \, . \tag{79}$$

The factor (79) can also be written in a form of coefficient of variation and parameter k since $v^2 = q/k$.

In contrast to the improved estimators for parameters of mean and variance, a general solution, a more efficient alternative to usual estimator of proportion, cannot be provided independently from distributions. Considering both examples for proportion, improved estimation techniques does not seem to provide practicable solutions for proportion just because the original unbiased estimator for proportion differs from one distribution to another.

CHAPTER FOUR
CONCLUSIONS

Utilization of prior information in the estimation procedure, which is the distinctive feature of Bayesian approach in statistical inference, allows the scientists to incorporate knowledge from former studies into current researches. As mentioned, there are a lot of studies in those coefficient of variation or kurtosis are used as prior information, such as Khan (1968), Hirano (1973), Arnholt and Hebert (1995).

A simple prior information such as the coefficient of variation or kurtosis is important in many biological and industrial studies since it is accessible in designing experiments, estimating sample size, etc. Usage of a known coefficient of variation in estimation procedure is investigated throughout this thesis. Following results gained in this study, are given as a brief summary to contribute studies in application of estimation theory.

Best improved estimators for mean and variance (MIMSEE's) are introduced. In both cases we observed the relative efficiencies between usual unbiased estimators and proposed improved estimators to make a comparison. The improved estimators are more efficient than the usual minimum variance unbiased estimators when sample sizes are small. It is observed that relative efficiencies decrease as sample size increases.

As an alternative to the improved estimator of mean, another estimator is given by shrinking a natural origin towards usual minimum variance unbiased estimator. We have analyzed the relationship between MIMSEE and alternative shrunken estimator using pivotal quantities. We came to the conclusion that improved estimators for mean can be developed using a suitable pivotal quantity.

Proceeding in another way, a general estimator for the variance of gamma family distributions is obtained, which was a special case of MIMSEE for variance dependent

to shape parameter and sample size. Relative efficiency values are presented for gamma, exponential, laplace and chi-square distributions. And it is observed that, the largest gains are obtained with both small sample sizes and shape parameters.

Improved estimation is investigated also for the parameter of proportion. Unlike to the improved estimators of mean and variance, no general form of minimum mean square error estimator for proportion is obtained, since the unbiased estimator of proportion depends on the distribution. This is why three different estimators are obtained for each of binomial, geometric and negative binomial distributions.

In this study, different approaches to estimation procedure are applied to estimate some parameters. The results may contribute to studies in statistical inference and turn out to be helpful for anybody working in this field.

REFERENCES

Arnholt, A.T., & Hebert, J.L. (1995) Estimating the Mean With Known Coefficient of Variation. *The American Statistician, 49,* 367-369.

Bibby, J., Toutenburg, H. (1977) *Prediction and Improved Estimation in Linear Models* (Second Edition). Chichester: Wiley.

Casella, G., Berger, R.L. (2002) *Statistical Inference* (Second Edition). Pacific Grove: Duxbury.

Hirano, K. (1973). Biased Efficient Estimator Utilizing Some A Priori Information. *Journal of the Japan Statistical Society,* (4), 11-13.

Khan, R.A. (1968). A Note on Estimating the Mean of a Normal Distribution with Known Coefficient of Variation. *Journal of the American Statistical Association,* (63), 1039-1041.

Mood, A.M., Graybill, F.A., Boes, D.C. (1974). *Introduction to Theory of Statistics* (Third Edition). New York: McGraw-Hill Publishing Company.

Searls, D.T. (1964). The Utilization of a known coefficient of variation in the estimation procedure. *Journal of the American Statistical Association,* (59), 1225-1226.

Singh, J., Pandey, N., & Hirano, K. (1973). On the Utilization of a Known Coefficient of Kurtosis in the Estimation Procedure of Variance. *Annals of the Institute of Statistical Mathematics,* (25), 51-55.

Thompson, J.R. (1968). Some Shrinkage Techniques for Estimating the Mean. *Journal of the American Statistical Association,* (63), 113-122.