Revista Colombiana de Estadística
July 2016, Volume 39, Issue 2, pp. 229 to 245
DOI: http://dx.doi.org/10.15446/rce.v39n2.51586

Improved Linear Combination of Two Estimators for a Function of Interested Parameter

Estimación eficiente de una función de un parámetro a través de una combinación lineal de dos estimadores

Afshin Fallah[a], Hamid Khoshtarkib[b]

Department of Statistics, Faculty of Basic Sciences, Imam Khomeini International University, Qazvin, Iran

Abstract

In this paper, we consider the problem of improving the efficiency of a linear combination of two estimators when the population coefficient of variation is known. We generalized the discussion from the case of a parameter to a function of are interested parameter. We show that two estimators obtained from a improved linear combination of two estimators and a linear combination of two improved estimators are equivalent in terms of efficiency. We also show how a doubly-improved linear combination of two estimators can be constructed when the population coefficient of variation is known.

***Key words*:** Coefficient of variation, Mean squared error, Efficiency, Linear combination.

Resumen

En este artículo, se considera el problema de mejorar la eficiencia de una combinación lineal de dos estimadores cuando el coeficiente de variación poblacional es conocido. Se generaliza el caso de un solo parámetro al de una función del parámetro. Se muestra que hay equivalencia, en términos de eficiencia, entre usar combinaciones lineales mejoradas y combinaciones lineales de estimadores mejorados. También se muestra como construir una combinación lineal doblemente mejorada cuando el coeficiente de variación poblacional es conocido.

***Palabras clave*:** coeficiente de variación, combinación lineal, eficiencia, error cuadrado medio.

[a]Professor. E-mail: a.fallah@sci.ikiu.ac.ir
[b]Researcher. E-mail: khoshtarkibhamid@yahoo.com

1. Introduction

In many practical inferential studies some prior information such as coefficient of variation (CV), kurtosis or skewness of population are available in advance. Using prior information to improve the efficiency of a given estimator has been considered in literature, repeatedly Searls (1964) and Arnholt & Hebert (2001) proposed an improved estimators for the population mean given the population CV. Wencheko & Wijekoon (2007) improved their results and obtained a shrunken estimator for the mean of one parameter exponential families. Also, given the population CV, Khan (1968) constructed a convex combination of two uncorrelated and unbiased estimators of the population mean with minimum mean square error (MSE). Improved estimators for the population variance that utilize the population kurtosis have been discussed by many authors notably Searls (1964), Kleffe (1985), Searls & Intarapanich (1990), Kanefuji & Iwase (1998), Wencheko & Chipoyera (2005) and Subhash & Cem (2013). In this regard, Laheetharan & Wijekoon (2010) proposed an improved estimator for the population variance and compared it with other estimators based on the scaled squared error loss function. The problem of finding improved estimators given an additional information has also been considered, for situations in which the dimension of sufficient statistics is grater than the dimension of the interested parameter. Gleser & Healy (1976) considered the problem of minimizing the MSE of a non-convex combination of two uncorrelated and unbiased estimators given a known population coefficient of variation. Samuel-Cahn (1994) expand their solution to a more general case for two correlated and unbiased estimators. Also, Arnholt & Hebert (1995) discussed non-convex combination of two correlated and biased estimators for an unknown parameter when the CV of both two estimators are known. It should be noted that the process of finding improved estimator usually leads to a biased estimator; therefore, the MSE criterion plays a main role in all results due to its emphasis on both variance and biasness of estimators. Some important results related to improving biased estimators are given by Bibby (1972), Bibby & Toutenburg (1977) and Bibby & Toutenburg (1978). The following theorems provide some of the most important results related to the problem of finding improved estimators in the presence of some prior information.

Theorem 1 (Arnholt & Hebert, 2001 and Laheetharan & Wijekoon, 2011)**.** *Let* $\boldsymbol{X} = (X_1, \ldots, X_n)'$ *be a random sample from a population with distribution* $f(x|\theta)$ *and* $T_1(\boldsymbol{X})$ *and* $T_2(\boldsymbol{X})$ *be estimators of* θ*, possibly correlated with* $E(T_i(\boldsymbol{X})) = k_i\theta$*,* $i = 1, 2$*. Suppose that the ratios* $\nu_i = \frac{Var(T_i)}{\theta^2}, i = 1, 2$ *are free from* θ *and* $Var(T_1(\boldsymbol{X})) < Var(T_2(\boldsymbol{X}))$*. Under these conditions, the estimator* $T^*(\boldsymbol{X}) = \alpha_1^* T_1(\boldsymbol{X}) + \alpha_2^* T_2(\boldsymbol{X})$ *uniformly has the minimum MSE among all estimators that are linear in* $T_1(\boldsymbol{X})$ *and* $T_2(\boldsymbol{X})$*, where*

$$\alpha_1^* = \frac{1 - \rho\lambda}{k_1(1 - 2\rho\lambda + \lambda^2 + (1 - \rho^2)(\nu_1/k_1^2))},$$
$$\alpha_2^* = \frac{\lambda(\lambda - \rho)}{k_2(1 - 2\rho\lambda + \lambda^2 + (1 - \rho^2)(\nu_1/k_1^2))}.$$

Furthermore, $\lambda^2 = \frac{k_2^2\nu_1}{k_1^2\nu_2}$ *and* $\rho = \frac{Cov(T_1(\boldsymbol{X}),T_2(\boldsymbol{X}))}{\sqrt{Var(T_1(\boldsymbol{X}))Var(T_2(\boldsymbol{X}))}}$ *are known and free from* θ.

Theorem 2 (Laheetharan & Wijekoon, 2010)**.** *Let* $\boldsymbol{X} = (X_1, \ldots, X_n)'$ *be a random sample from a population with distribution* $f(x|\theta)$ *and* $g(\theta)$ *be a real-valued function on the parameter space* Θ. *Let* $T_1(\boldsymbol{X})$ *and* $T_2(\boldsymbol{X})$ *be point estimators of* $g(\theta)$ *with* $E(T_i(\boldsymbol{X})) = k_i g(\theta)$, *where* $k_i \in \Re$. *Then, the estimators* $T_i^*(\boldsymbol{X}) = \alpha_i^* T_i(\boldsymbol{X}), i = 1, 2$ *uniformly have the minimum MSE among all estimators in class of* $C_{T_i}(\alpha_i) = \{\alpha_i T_i(\boldsymbol{X}) \mid 0 < \alpha_i < \infty\}$, *where* $\alpha_i^* = \frac{k_i}{k_i^2+\tau_i^2}$ *and* $\tau_i^2 = \frac{Var(T_i(\boldsymbol{X}))}{[g(\theta)]^2}, i = 1, 2$ *are free from* θ. *Furthermore, if* $k_2 < k_1(\tau_2/\tau_1)$ *then* $T_1^*(\boldsymbol{X})$ *has smaller MSE and if* $k_2 > k_1(\tau_2/\tau_1)$, *then* $T_2^*(\boldsymbol{X})$ *has smaller MSE.*

In this paper, we consider the problem of improving the efficiency of a linear combination of two estimators, when the population CV is known. The rest of paper is organized as follows: in Section 2, we briefly review the main results related to the improved linear combination of estimators. In Section 3 we generalized the discussion from the case of a parameter to a function of an interested parameter by expanding the results of Gleser & Healy (1976) and and Arnholt & Hebert (2001). In section 4, we show that two estimators obtained from a improved linear combination of two estimators and a linear combination of two improved estimators are the same in terms of efficiency. In section 5, we show that how a doubly-improved linear combination of two estimators can be construct when the population CV is known. In section 6, we provide some illustrative examples.

2. Improved Linear Combination of Two Estimators for a Parameter

In this section, we briefly review the main results related to an improved linear combination of estimators, when some additional information is available.

Using some prior information may reduce the dimension of parameter space. For example, when the coefficient of variation $\nu = \frac{\sigma}{\mu}$ is known, the distribution of $N(\mu, \sigma^2), \mu \neq 0$ can be written as $N(\mu, \nu^2\mu^2)$ due to the equation $\sigma^2 = \nu^2\mu^2$. It can be seen that the dimension of sufficient statistics, $(\bar{X}, S^2)$, is more than the dimension of the parameter of interest, μ. In this situation using only a part of the sufficient statistics leads to a loss of some information about the parameter of interest. Therefore, the simultaneous use of two or more estimator is necessary to achieve more possible information about the parameter of interest. One can use a combination of estimators to construct an efficient estimator. Khan (1968) proposed the optimal combinations of two independent and unbiased estimators of the population mean when the sampling distribution is normal and the population coefficient of variation, ν, is known. Consider $T_1(\mathbf{X}) = \bar{X}_n$, $T_2(\mathbf{X}) = c_n S$, $c_n = (n^{1/2}\Gamma((n-1)/2))/((2a)^{1/2}\Gamma(n/2))$, as two unbiased and independent estimators for μ, where S is the sample standard deviation and $a = \sqrt{\nu}$. Then, the shrinkage estimator

$$T(\mathbf{X}) = \alpha^*\bar{X} + (1-\alpha^*)c_n S,$$

is the optimal combination of estimators $\bar{X}$ and c_nS, where $\alpha^* = d_n/(d_n + n^{-1}a)$ and $d_n = [n^{-1}(n-1)ac_n{}^2 - 1]$. Of course, it is not necessary to restrict these combinations to be convex. Gleser & Healy (1976) considered a more general case with $T = \alpha_1 T_1 + \alpha_2 T_2$, where T_i are any independent and unbiased estimators of θ and $\alpha_1 + \alpha_2$ is not necessarily equal to 1. The only restriction is that the ratios $\nu_i^2 = \theta^2 Var(T_i), i = 1, 2$ are free from θ, where ν_i denotes the CV of estimator T_i. This restriction holds, for example, when the T_i, $i = 1, 2$ are unbiased and ν is known. Since the estimator T is not necessarily convex, it is not necessarily an unbiased estimator for θ. The authors showed that the optimal weights in this case are given by

$$\alpha_1^* = \frac{\nu_2}{\nu_1 + \nu_2 + \nu_1\nu_2} \quad , \quad \alpha_2^* = \frac{\nu_1}{\nu_1 + \nu_2 + \nu_1\nu_2}.$$

Samuel-Cahn (1994) studied another generalized case of optimizing a convex combination of two unbiased, dependent estimators with a known correlation coefficient ρ. They derived the optimal weight as $\alpha^* = (1 - \rho\lambda)/(1 - 2\rho\lambda + \lambda^2)$, where $\lambda^2 = Var(T_1)/Var(T_2)$. The authors assumed that λ^2 is known and free from θ. It should be noted that when the estimators CV are known and free from θ for both estimators, this restriction is held.

3. Improved Linear Combination of Two Estimators for a Function of a Parameter

In a population with distribution $f(x|\theta)$ there are different interested parameters such as mean, variance, etc. these appear as different functions of θ, hence it is interesting to look for improved estimators for a function of a parameter. In recent years some authors, notably Laheetharan & Wijekoon (2010), have considered the problem of finding improved estimators for a function of an interested parameter, say $g(\theta)$. In this section, we derived an optimal shrinkage estimator for a function of a parameter with assumption of known population CV. The following lemma, which is left without proof, provides a preliminary necessary fact for the next theorem.

Lemma 1. *Let $T(\boldsymbol{X})$ be an estimator of parameter θ and $g(\cdot)$ be a real valued function, where $E(T(\boldsymbol{X})) = kg(\theta)$. If the population CV is known, then the ratio $\tau^2 = \frac{Var(T(\boldsymbol{X}))}{[g(\theta)]^2}$ is free from θ.*

Using the Lemma 1, we improved the Gleser & Healy (1976) and Arnholt & Hebert (2001) results to estimate a function of parameter, $g(\theta)$, in the next theorem.

Theorem 3. *Let $\boldsymbol{X} = (X_1, \ldots, X_n)'$ be a random sample from a population with distribution $f(x|\theta)$ and let $T_1(\boldsymbol{X})$ and $T_2(\boldsymbol{X})$ be estimators for $g(\theta)$, possibly correlated with $E(T_i(\boldsymbol{X})) = k_i g(\theta)$, and $i = 1, 2$. Under these conditions, $T^*_{LC}(\boldsymbol{X}) = \alpha_1^* T_1(\boldsymbol{X}) + \alpha_2^* T_2(\boldsymbol{X})$ uniformly has the minimum MSE among all estimators in the class $C_{T_1,T_2}(\alpha_1, \alpha_2) = \{\alpha_1 T_1(\boldsymbol{X}) + \alpha_2 T_2(\boldsymbol{X}) \mid 0 < \alpha_1, \alpha_2 < \infty\}$,* *where*

$$\alpha_1^* = \frac{1-\rho\lambda}{k_1(1-2\rho\lambda+\lambda^2+(1-\rho^2)(\tau_1^2/k_1^2))}, \qquad \alpha_2^* = \frac{\lambda(\lambda-\rho)}{k_2(1-2\rho\lambda+\lambda^2+(1-\rho^2)(\tau_1^2/k_1^2))}. \tag{1}$$

In addition, $\tau_i^2 = \frac{Var(T_i(\boldsymbol{X}))}{[g(\theta)]^2}$, $\lambda^2 = \frac{k_2^2\tau_1^2}{k_1^2\tau_2^2}$ *and* $\rho = \frac{Cov(T_1(\boldsymbol{X}),T_2(\boldsymbol{X}))}{\sqrt{Var(T_1(\boldsymbol{X}))Var(T_2(\boldsymbol{X}))}}$ *are known and free from* θ*.*

Proof. Let $T_{LC}(\mathbf{X}) = \alpha_1 T_1(\mathbf{X}) + \alpha_2 T_2(\mathbf{X}); 0 < \alpha_1, \alpha_2 < \infty$. Without loss of generality, we assume that $k_1, k_2 > 0$, then

$$\begin{aligned} MSE(T_{LC}(\mathbf{X})) &= Var(T_{LC}(\mathbf{X})) + bias^2(T_{LC}(\mathbf{X})) \\ &= Var(\alpha_1 T_1(\mathbf{X}) + \alpha_2 T_2(\mathbf{X})) \\ &+ [E(\alpha_1 T_1(\mathbf{X}) + \alpha_2 T_2(\mathbf{X})) - g(\theta)]^2 \\ &= \alpha_1^2 Var(T_1(\mathbf{X})) + \alpha_2^2 Var(T_2(\mathbf{X})) \\ &+ 2\alpha_1\alpha_2\rho\sqrt{Var(T_1(\mathbf{X}))Var(T_2(\mathbf{X}))} \\ &+ (\alpha_1 k_1 + \alpha_2 k_2 - 1)^2 [g(\theta)]^2. \end{aligned} \tag{2}$$

Differentiating (2) with respect to α_1 and α_2 and equating it to zero leads to the following system of equations:

$$\begin{cases} \frac{\partial MSE(T_{LC}(\mathbf{X}))}{\partial \alpha_1} = & 2\alpha_1 Var(T_1(\mathbf{X}))) + 2\alpha_2\rho\sqrt{Var(T_1(\mathbf{X}))Var(T_2(\mathbf{X}))} \\ & +2k_1[g(\theta)]^2(\alpha_1 k_1 + \alpha_2 k_2 - 1) = 0, \\ \frac{\partial MSE(T_{LC}(\mathbf{X}))}{\partial \alpha_2} = & 2\alpha_2 Var(T_2(\mathbf{X}))) + 2\alpha_1\rho\sqrt{Var(T_1(\mathbf{X}))Var(T_2(\mathbf{X}))} \\ & +2k_2[g(\theta)]^2(\alpha_1 k_1 + \alpha_2 k_2 - 1) = 0. \end{cases} \tag{3}$$

The solutions of equations (3) are given by

$$\begin{aligned} \alpha_1^* &= \frac{(1-\alpha_2^* k_2)k_1 g(\theta)^2 - \alpha_2^*\rho\sqrt{Var(T_1(\mathbf{X}))Var(T_2(\mathbf{X}))}}{Var(T_1(\mathbf{X})) + k_1^2 g(\theta)^2} \\ & \frac{k_1 - \alpha_2^*(k_1k_2 + \rho\tau_1\tau_2)}{k_1^2 + \tau_1^2} \end{aligned} \tag{4}$$

$$\begin{aligned} \alpha_2^* &= \frac{(1-\alpha_1^* k_1)k_2 g(\theta)^2 - \alpha_1^*\rho\sqrt{Var(T_1(\mathbf{X}))Var(T_2(\mathbf{X}))}}{Var(T_2(\mathbf{X})) + k_2^2 g(\theta)^2} \\ &= \frac{k_2 - \alpha_1^*(k_1k_2 + \rho\tau_1\tau_2)}{k_2^2 + \tau_2^2}. \end{aligned} \tag{5}$$

Substituting (4) in (5), we have

$$\begin{aligned} \alpha_2^* &= \frac{k_2 - \frac{k_1 - \alpha_2^*(k_1k_2+\rho\tau_1\tau_2)}{k_1^2+\tau_1^2}(k_1k_2 + \rho\tau_1\tau_2)}{k_2^2 + \tau_2^2} \\ &= \frac{k_2\tau_1^2 - \rho k_1\tau_1\tau_2 + \alpha_2^* k_1^2 k_2^2 + 2\alpha_2^*\rho k_1 k_2\tau_1\tau_2 + \alpha_2^*\rho^2\tau_1^2\tau_2^2}{(k_1^2+\tau_1^2)(k_2^2+\tau_2^2)}. \end{aligned}$$

Therefore

$$\begin{aligned}\alpha_2^* &= \frac{k_2\tau_1^2 - \rho k_1\tau_1\tau_2}{(k_1^2+\tau_1^2)(k_2^2+\tau_2^2) - k_1^2k_2^2 - 2\rho k_1k_2\tau_1\tau_2 - \rho^2\tau_1^2\tau_2^2}\\ &= \frac{\lambda(\lambda-\rho)}{k_2(1+\lambda^2-2\rho\lambda-(1-\rho^2)\frac{\tau_1^2}{k_1^2})}.\end{aligned}$$

Similarly, we have

$$\begin{aligned}\alpha_1^* &= \frac{k_1 - \frac{\lambda(\lambda-\rho)}{k_2(1+\lambda^2-2\rho\lambda-(1-\rho^2)\frac{\tau_1^2}{k_1^2})}(k_1k_2+\rho\tau_1\tau_2)}{k_1^2+\tau_1^2}\\ &= \frac{1-\rho\lambda}{k_1(1+\lambda^2-2\rho\lambda-(1-\rho^2)\frac{\tau_1^2}{k_1^2})}.\end{aligned}$$

The second order partial derivations of (2) with respect to α_1 and α_2 given by

$$\begin{cases} \frac{\partial^2 MSE(T_{LC}(\mathbf{X}))}{\partial\alpha_1^2} = 2Var(T_1(\mathbf{X})) + 2k_1^2[g(\theta)]^2 \\ \frac{\partial^2 MSE(T_{LC}(\mathbf{X}))}{\partial\alpha_2^2} = 2Var(T_2(\mathbf{X})) + 2k_2^2[g(\theta)]^2, \end{cases} \tag{6}$$

which are both positive, therefore α_1^* and α_2^* minimize the value of $MSE(T_{LC}(\mathbf{X}))$, and the estimator $T_{LC}^*(\mathbf{X}) = \alpha_1^*T_1(\mathbf{X}) + \alpha_2^*T_2(\mathbf{X})$ uniformly has the minimum MSE among all estimators in the class $C_{T_1,T_2}(\alpha_1,\alpha_2) = \{\alpha_1T_1(\mathbf{X}) + \alpha_2T_2(\mathbf{X}) \mid 0 < \alpha_1, \alpha_2 < \infty\}$. □

Obviously, Theorem 3 is assumptions are culmination of the required assumptions for Theorems 1 and 2, which are provided in the section. The next corollary is an immediate consequence of Theorem 3.

Corollary 1. *The estimators proposed by Arnholt & Hebert (2001) can be obtained as special cases, in Theorem 3, for $g(\theta) = \theta$ and $\tau_i^2 = \nu_i$,. Also, for $g(\theta) = \theta$, $k_i = 1$, $\rho = 0$ and $\tau_i^2 = \nu_i$, we obtained Gleser & Healy (1976) results.*

4. Linear Combination of Two Improved Estimators

One may expected, intuitively, that using two estimators with improved efficiency to construct an optimal linear combination, leads to a more efficient estimator. In the other words, it may be expected that improving the two estimators $T_1(\mathbf{X})$ and $T_2(\mathbf{X})$ by using Theorem 2 and then constructing an optimal combination of these improved estimators by Theorem 3 leads to a more efficient linear combination. The following theorem shows that this intuitive expectation is not true. In fact, it shows that two estimators obtained from an improved linear combination of two estimators and a linear combination of two improved estimators are equivalent, in terms of efficiency.

Theorem 4. *Suppose based on Theorem 2, $T_1^*(\boldsymbol{X})$ and $T_2^*(\boldsymbol{X})$ are improved versions of estimators $T_1\boldsymbol{X})$ and $T_2(\boldsymbol{X})$, respectively. Let $T_{LC}^*(\boldsymbol{X}) = \alpha_1^* T_1(\boldsymbol{X}) + \alpha_2^* T_2(\boldsymbol{X})$ is the optimal Linear Combination (LC) of $T_1(\boldsymbol{X})$ and $T_2(\boldsymbol{X})$, based on Theorem 3. If $T_{LCI}^*(\boldsymbol{X}) = \alpha_1^{**}T_1^*(\boldsymbol{X}) + \alpha_2^{**}T_2^*(\boldsymbol{X})$ be the optimal Linear Combination of Improved (LCI) estimators $T_1^*(\boldsymbol{X})$ and $T_2^*(\boldsymbol{X})$, respectively, then $T_{LCI}^*(\boldsymbol{X}) = T_{LC}^*(\boldsymbol{X})$.*

***Proof*.** Let $T_{LCI}(\mathbf{X}) = \alpha_1 T_1^*(\mathbf{X}) + \alpha_2 T_2^*(\mathbf{X}) = \alpha_1 b_1 T_1(\mathbf{X}) + \alpha_2 b_2 T_2(\mathbf{X}); 0 < \alpha_1, \alpha_2 < \infty$.
Without loss of generality, we assume that $k_1, k_2 > 0$. Then,

$$\begin{aligned}
MSE(T_{LCI}(\mathbf{X})) &= Var(T_{LCI}(\mathbf{X})) + bias^2(T_{LCI}(\mathbf{X})) \\
&= Var(\alpha_1 b_1 T_1(\mathbf{X}) + \alpha_2 b_2 T_2(\mathbf{X})) \\
&+ (E(\alpha_1 b_1 T_1(\mathbf{X}) + \alpha_2 b_2 T_2(\mathbf{X})) - g(\theta))^2, \\
&= \alpha_1^2 b_1^2 Var(T_1(\mathbf{X})) + \alpha_2^2 b_2^2 Var(T_2(\mathbf{X})) \\
&+ 2\alpha_1\alpha_2 b_1 b_2 \rho\sqrt{Var(T_1(\mathbf{X}))Var(T_2(\mathbf{X}))} \\
&+ (\alpha_1 b_1 k_1 + \alpha_2 b_2 k_2 - 1)^2 (g(\theta))^2.
\end{aligned} \tag{7}$$

Differentiating (7) with respect to α_1 and α_2 and equating to zero leads to the following system of equations:

$$\begin{cases}
\frac{\partial MSE(T_{LCI}(\mathbf{X}))}{\partial\alpha_1} = & 2\alpha_1 b_1^2 Var(T_1(\mathbf{X}))) + 2\alpha_2 b_1 b_2 \rho\sqrt{Var(T_1(\mathbf{X}))Var(T_2(\mathbf{X}))} \\
& +2k_1[g(\theta)]^2(\alpha_1 k_1 + \alpha_2 k_2 - 1) = 0, \\
\frac{\partial MSE(T_{LCI}(\mathbf{X}))}{\partial\alpha_2} = & 2\alpha_2 Var(T_2(\mathbf{X}))) + 2\alpha_1 \rho\sqrt{Var(T_1(\mathbf{X}))Var(T_2(\mathbf{X}))} \\
& +2k_2[g(\theta)]^2(\alpha_1 k_1 + \alpha_2 k_2 - 1) = 0.
\end{cases} \tag{8}$$

The solutions of equation (8) are given by

$$\begin{aligned}
\alpha_1^{**} &= \frac{(1-\alpha_2^* b_2 k_2) b_1 k_1 g(\theta)^2 - \alpha_2^* b_1 b_2 \rho\sqrt{Var(T_1(\mathbf{X}))Var(T_2(\mathbf{X}))}}{b_1^2[Var(T_1(\mathbf{X})) + k_1^2 g(\theta)^2]} \\
&= \frac{k_1 - \alpha_2^* b_2(k_1 k_2 + \rho\tau_1\tau_2)}{b_1(k_1^2 + \tau_1^2)},
\end{aligned} \tag{9}$$

$$\begin{aligned}
\alpha_2^{**} &= \frac{(1-\alpha_1^* b_1 k_1) b_2 k_2 g(\theta)^2 - \alpha_1^* b_1 b_2 \rho\sqrt{Var(T_1(\mathbf{X}))Var(T_2(\mathbf{X}))}}{b_2^2[Var(T_2(\mathbf{X})) + k_2^2 g(\theta)^2]} \\
&= \frac{k_2 - \alpha_1^* b_1(k_1 k_2 + \rho\tau_1\tau_2)}{b_2(k_2^2 + \tau_2^2)}.
\end{aligned} \tag{10}$$

Substituting (9) in (10) we have the following,

$$\begin{aligned}
\alpha_1^{**} &= \frac{1}{b_1}\frac{1-\rho\lambda}{k_1(1+\lambda^2 - 2\rho\lambda - (1-\rho^2)\frac{\tau_1^2}{k_1^2})} = \frac{1}{b_1}\alpha_1^*, \\
\alpha_2^{**} &= \frac{1}{b_2}\frac{\lambda(\lambda-\rho)}{k_2(1+\lambda^2 - 2\rho\lambda - (1-\rho^2)\frac{\tau_1^2}{k_1^2})} = \frac{1}{b_2}\alpha_2^*,
\end{aligned}$$

where α_1^* and α_2^* presented in equation (1). Then

$$\begin{aligned} T^*_{LCI}(\mathbf{X}) &= \alpha_1^{**} T_1^*(\mathbf{X}) + \alpha_2^{**} T_2^*(\mathbf{X}) \\ &= \frac{\alpha_1^*}{b_1} T_1^*(\mathbf{X}) + \frac{\alpha_2^{**}}{b_2} T_2^*(\mathbf{X}) \\ &= \alpha_1^* T_1(\mathbf{X}) + \alpha_2^* T_2(\mathbf{X}) \\ &= T^*_{LC}(\mathbf{X}). \end{aligned}$$

□

5. A Doubly-Improved Linear Combination of Two Estimators

In this section, we show how a Doubly-Improved (DI) linear combination of two estimators of a parameter can be construct when the population CV is known. In the next theorem, we try to further improve the improved linear combination estimator that resulted from Theorem 3 by applying Theorem 2.

Theorem 5. *Consider the assumptions of Theorem 3. Suppose* $T^*_{LC}(\boldsymbol{X}) = \alpha_1^* T_1(\boldsymbol{X}) + \alpha_2^* T_2(\boldsymbol{X})$ *are the optimal linear combination of estimators* $T_1(\boldsymbol{X})$ *and* $T_2(\boldsymbol{X})$ *where* α_1^* *and* α_2^* *are given in equation* (1)*. Then,*

a) The doubly-improved estimator $T^*_{DI}(\boldsymbol{X}) = \alpha^* T^*_{LC}(\boldsymbol{X})$ *uniformly has the minimum MSE among all estimators of* $g(\theta)$ *that are in the class* $C_{T^*_{LC}}(\alpha) = \{\alpha T^*_{LC}(\boldsymbol{X}) \mid \alpha \in (0,\infty)\}$ *where* $\alpha^* = \frac{k}{k^2+\tau^2}$.

b) The minimum value of MSE($T^*_{DI}(\boldsymbol{X})$*) is given by* $\frac{\tau^2}{k^2+\tau^2}[g(\theta)]^2$

Proof. a) Since

$$\begin{aligned} T^*_{DI}(\mathbf{X}) &= \alpha T^*_{LC}(\mathbf{X}); \quad 0<\alpha<\infty, \\ E(T^*_{LC}(\mathbf{X})) &= (\alpha_1^* k_1 + \alpha_2^* k_2) g(\theta) = k g(\theta), \\ k &= \alpha_1^* k_1 + \alpha_2^* k_2, \\ \tau^2 &= \frac{Var(T^*_{LC}(\mathbf{X}))}{[g(\theta)]^2} = {\alpha_1^*}^2 \tau_1^2 + {\alpha_2^*}^2 \tau_2^2 + 2\alpha_1^*\alpha_2^* \rho \tau_1 \tau_2, \end{aligned}$$

hence,

$$\begin{aligned} E(T^*_{DI}(\mathbf{X})) &= \alpha E(T^*_{LC}(\mathbf{X})) \\ &= \alpha(\alpha_1^* k_1 g(\theta) + \alpha_2^* k_2 g(\theta)) \\ &= \alpha(\alpha_1^* k_1 + \alpha_2^* k_2) g(\theta) \\ &= \alpha k g(\theta). \end{aligned}$$

Also, $Var(T^*_{DI}(\mathbf{X})) = Var(\alpha T^*_{LC}(\mathbf{X})) = \alpha^2 Var(T^*_{LC}(\mathbf{X})))$. Therefore the MSE of $T^*_{DI}(\mathbf{X}) \in C_{T^*_s}(\alpha)$ that is obtained is:

$$\begin{aligned} MSE(T^*_{DI}(\mathbf{X})) &= Var(T^*_{DI}(\mathbf{X})) + bias^2(T^*_{DI}(\mathbf{X})) \\ &= Var(\alpha T^*_{LC}(\mathbf{X})) + [E(\alpha T^*_{LC}(\mathbf{X})) - g(\theta)]^2 \\ &= \alpha^2 Var(T^*_{LC}(\mathbf{X})) + (\alpha k - 1)^2 [g(\theta)]^2. \end{aligned}$$

Due to following system of equations

$$\begin{cases} \frac{\partial MSE(T^*_{DI}(\mathbf{X}))}{\partial \alpha} = & 2\alpha Var(T^*_{LC}(\mathbf{X})) + 2k(\alpha k - 1)[g(\theta)]^2, \\ \frac{\partial^2 MSE(T^*_{DI}(\mathbf{X}))}{\partial \alpha^2} = & 2Var(T^*_{LC}(\mathbf{X})) + 2k^2[g(\theta)]^2 > 0, \end{cases} \tag{11}$$

it can be easily shown that the estimator $T^*_{DI}(\mathbf{X}) = k(k^2+\tau^2)^{-1} T^*_{LC}(\mathbf{X})$ has the minimum MSE in the class $C_{T^*_{LC}}(\alpha)$.
b) We have

$$\begin{aligned} \min_{\alpha^*} MSE(T^*_{DI}(\mathbf{X})) &= \min_{\alpha^*} MSE(\alpha^* T^*_{LC}(\mathbf{X})) \\ &= \min_{\alpha^*} \{ {\alpha^*}^2 Var(T^*_{LC}(\mathbf{X})) + [E(\alpha^* T^*_{LC}(\mathbf{X})) - g(\theta)]^2 \} \\ &= \frac{k^2}{(k^2+\tau^2)^2} Var(T^*_{LC}(\mathbf{X})) + \frac{\tau^4}{(k^2+\tau^2)^2}[g(\theta)]^2 \\ &= [\frac{g(\theta)}{k^2+\tau^2}]^2 (k^2 \frac{Var(T^*_{LC}(\mathbf{X}))}{[g(\theta)]^2} + \tau^4) \\ &= [\frac{g(\theta)}{k^2+\tau^2}]^2 [k^2\tau^2 + \tau^4] \\ &= \frac{\tau^2}{k^2+\tau^2}[g(\theta)]^2. \end{aligned}$$

□

6. Illustrative Examples

Using Theorem, 2 it is possible to obtain optimal shrunken estimators for both the population mean, say $T^*_\mu(X)$, and the population variance, say $T^*_{\sigma^2}(X)$. Note that if the population CV, ν, is known, then one can easily use the mean based estimator $T^*_{\sigma_\nu{}^2}(X) = \nu^2[T^*_\mu(X)]^2$ as another estimator for the population variance. Laheetharan & Wijekoon (2010) compared the MSE of estimators $T^*_{\sigma^2}(X)$ and $T^*_{\sigma_\nu{}^2}(X)$.

Suppose $E(T_\mu(\mathbf{X})) = k_1\mu$ and $E(T_{\sigma^2}(\mathbf{X})) = k_2\sigma^2$ are estimators of the population mean and variance, respectively. Since the population CV is known, then the estimator $T_{\sigma_\nu{}^2}(\mathbf{X}) = \nu^2[T_\mu(\mathbf{X})]^2$ can be considered as another estimator for the population variance. Hence, if $Var(T_\mu(\mathbf{X})) = c\sigma^2$, we have the following

$$\begin{aligned}
E(T_{\sigma^2}(\mathbf{X})) &= k_2\sigma^2, \\
E(T_{\sigma_\nu{}^2}(\mathbf{X})) &= E(\nu^2[T_\mu(\mathbf{X})]^2) \\
&= \nu^2[Var(T_\mu(\mathbf{X})) + E^2(T_\mu(\mathbf{X}))] \\
&= c\nu^2\sigma^2 + k_1^2\nu^2\mu^2 \\
&= (c\nu^2 + k_1^2)\sigma^2 \\
&= k_{\sigma_\nu{}^2}\sigma^2,
\end{aligned}$$

where $k_1,\ k_2,\ k_{\sigma^2{}_\nu}$ and c are known constants. Using the above information, and based on Theorems 4 and 5, we have the following theorem to estimate the population variance.

Theorem 6. *Let $\boldsymbol{X} = (X_1, \dots, X_n)'$ be a random sample from a population with distribution $f(x|\theta)$, and let $T_{\sigma_\nu{}^2}(\boldsymbol{X})$ and $T_{\sigma^2}(\boldsymbol{X})$ be estimators of σ^2, possibly correlated with $E(T_{\sigma_\nu{}^2}(\boldsymbol{X})) = k_{\sigma_\nu{}^2}\sigma^2$ and $E(T_{\sigma^2}(\boldsymbol{X})) = k_2\sigma^2$. Then,*

i) Based on Theorem 4, the linear combination $T^*_{LC}(\boldsymbol{X}) = \alpha_{\sigma_\nu{}^2}{}^*T_{\sigma_\nu{}^2}(\boldsymbol{X}) + \alpha_{\sigma^2}{}^*T_{\sigma^2}(\boldsymbol{X})$ uniformly has the minimum MSE of all estimators in the class $C_{T_{\sigma_\nu{}^2},T_{\sigma^2}}(\alpha_{\sigma_\nu{}^2}, \alpha_{\sigma^2}) = \{\alpha_{\sigma_\nu{}^2}T_{\sigma_\nu{}^2}(\boldsymbol{X}) + \alpha_{\sigma^2}T_{\sigma^2}(\boldsymbol{X}) \mid 0 < \alpha_{\sigma_\nu{}^2}, \alpha_{\sigma^2} < \infty\}$, where

$$\begin{aligned}
\alpha^*_{\sigma_\nu{}^2} &= \frac{1 - \rho\lambda}{k_{\sigma_\nu{}^2}(1 - 2\rho\lambda + \lambda^2 + (1 - \rho^2)(\tau^2_{\sigma_\nu{}^2}/k^2_{\sigma_\nu{}^2}))}, \\
\alpha^*_{\sigma^2} &= \frac{\lambda(\lambda - \rho)}{k_2(1 - 2\rho\lambda + \lambda^2 + (1 - \rho^2)(\tau^2_{\sigma_\nu{}^2}/k^2_{\sigma_\nu{}^2}))}.
\end{aligned} \tag{12}$$

Also, $\tau^2_{\sigma_\nu{}^2} = \frac{Var(T_{\sigma_\nu{}^2})}{[\sigma^2]^2}$,$\tau^2_{\sigma^2} = \frac{Var(T_{\sigma^2})}{[\sigma^2]^2}$, $\lambda^2 = \frac{k_2^2\tau^2_{\sigma_\nu{}^2}}{k^2_{\sigma_\nu{}^2}\tau^2_{\sigma^2}}$ and

$$\rho = \frac{Cov(T_{\sigma_\nu{}^2}(\boldsymbol{X}), T_{\sigma^2}(\boldsymbol{X}))}{\sqrt{Var(T_{\sigma_\nu{}^2}(\boldsymbol{X}))Var(T_{\sigma^2}(\boldsymbol{X}))}}$$

are known and free from σ^2.

ii) Since

$$E(T^*_{LC}(\boldsymbol{X})) = (\alpha_{\sigma_\nu{}^2}{}^* + \alpha_{\sigma^2}{}^*k_2)\sigma^2 = k\sigma^2,$$

and

$$\begin{aligned}
\tau^2 &= \frac{Var(T^*_{LC}(\boldsymbol{X}))}{[\sigma^2]^2} \\
&= [\alpha_{\sigma_\nu{}^2}{}^*]^2\tau^2_{\sigma_\nu{}^2} + [\alpha_{\sigma^2}{}^*]^2\tau^2_{\sigma^2} + 2\alpha_{\sigma_\nu{}^2}{}^*\alpha_{\sigma^2}{}^*\rho\tau_{\sigma_\nu{}^2}\tau_{\sigma^2}
\end{aligned}$$

is free from σ^2 and known, based on Theorem 5, the doubly-improved estimator $T^*_{DI}(\boldsymbol{X}) = \alpha^*T^*_{LC}(\boldsymbol{X})$ uniformly has the minimum MSE of all σ^2 estimators that are in the class $C_{T^*_{LC}}(\alpha) = \{\alpha T^*_{LC}(\boldsymbol{X}) \mid \alpha \in (0, \infty)\}$, where $\alpha^* = \frac{k}{k^2+\tau^2}$.

Example 1. Let $X = (X_1, \dots, X_n)'$ be a random sample from a population with a location-scale exponential distribution $\mathrm{E}(\theta, \theta)$, given by

$$f(x) = \frac{1}{\theta}\exp(\frac{x-\theta}{\theta})I_{(\theta,+\infty)}(x).$$

Since the estimators $T_1(X) = X_{(1)}$ and $T_2(X) = \sum_{i=1}^{n}(X_i - X_{(1)})$ are jointly sufficient statistics for $g(\theta) = \theta$, our motivation is to use a combination of these two estimators. We can to estimate an interested parameter. It is easy to shaw that the mean and variance of T_1 and T_2 are given by

$$\begin{aligned} E(T_1(X)) &= \frac{n+1}{n}\theta, \\ Var(T_1(X)) &= \frac{1}{n^2}\theta^2, \end{aligned} \tag{13}$$

and

$$\begin{aligned} E(T_2(X)) &= (n-1)\theta, \\ Var(T_2(X)) &= (n-1)\theta^2, \end{aligned} \tag{14}$$

respectively. Hence, based on the notation of Theorem 3.1, we have $k_1 = \frac{n+1}{n}$, $k_2 = n-1$, $\tau_1^2 = \frac{1}{n^2}$, $\tau_2^2 = n-1$, $\lambda^2 = \frac{1}{(n+1)^2}$, $\rho_{T_1,T_2} = -1$. Therefore, according to equation (1), the improved linear combination of two estimators T_1 and T_2 is given by $T^*_{LC}(\mathbf{X}) = \alpha_1^* T_1(\mathbf{X}) + \alpha_2^* T_2(\mathbf{X})$, where

$$\begin{aligned} \alpha_1^* &= \frac{n}{n+2}, \\ \alpha_2^* &= \frac{1}{(n-1)(n+2)}. \end{aligned} \tag{15}$$

This improved estimator uniformly has the minimum MSE among all estimators in the class $C_{T_1,T_2}(\alpha_1, \alpha_2) = \{\alpha_1 T_1(\mathbf{X}) + \alpha_2 T_2(\mathbf{X}) \mid 0 < \alpha_1, \alpha_2 < \infty\}$. The value of MSE for an improved estimator has been computed for different sample sizes and plotted in Figure 1. Decreasing the value of MSE by increasing the sample size, indicates that the improved shrinkage estimator will become more consistent.

Example 2. Let $X = (X_1, \dots, X_n)$ be a random sample from a population with normal distribution $N(\theta, \theta^2)$. This is a curved exponential family with a two-dimensional sufficient statistic. The joint minimal sufficient statistic for $g(\theta) = \theta^2$ is $(T_1(X), T_2(X)) = (\bar{X}^2, \sum_{i=1}^{n}(X_i - \bar{X})^2)$ and the following equations hold for these estimators:

$$\begin{aligned} E(T_1(X)) &= \frac{n+1}{n}\theta^2, \\ Var(T_1(X)) &= \frac{12n^2+2}{n^2}\theta^4, \\ E(T_2(X)) &= (n-1)\theta^2, \\ Var(T_2(X)) &= 2(n-1)\theta^4. \end{aligned} \tag{16}$$

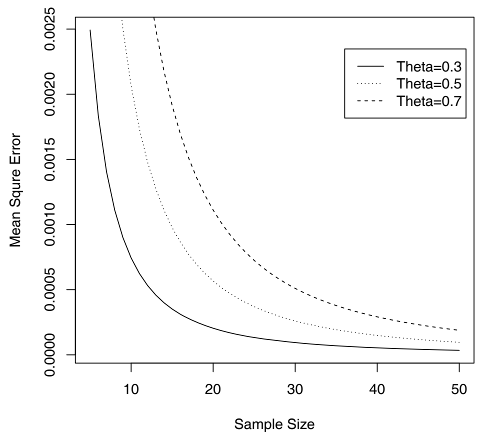

FIGURE 1: MSE of the improved shrinkage estimator for estimation of $g(\theta) = \theta^2$ versus sample size in $E(\theta,\theta)$ distribution.

Hence, based on Theorem 3.1 is notation, we obtain $k_1 = \frac{n+1}{n}$, $k_2 = n - 1$, $\tau_1^2 = \frac{12n^2+2}{n^2}$, $\tau_2^2 = 2(n-1)$, $\lambda^2 = \frac{(6n^2+1)(n-1)}{(n+1)^2}$ and $\rho_{T_1,T_2} = 0$ due to the independence of $\bar{X}$ and S^2 in normal distribution. Therefore, according to Equation (1), improved linear combination of two estimators T_1 and T_2 is given by $T_{LC}^*(\mathbf{X}) = \alpha_1^* T_1(\mathbf{X}) + \alpha_2^* T_2(\mathbf{X})$, where

$$\alpha_1^* = \frac{n(n+1)^3}{(n+1)^4 + 2(n-1)(6n^2+1)^2},$$
$$\alpha_2^* = \frac{(n+1)^2(6n^2+1)}{(n+1)^4 + 2(n-1)(6n^2+1)^2}.$$

Again, this improved estimator uniformly has the minimum MSE among all estimators in the class $C_{T_1,T_2}(\alpha_1,\alpha_2) = \{\alpha_1 T_1(\mathbf{X}) + \alpha_2 T_2(\mathbf{X}) \mid 0 < \alpha_1, \alpha_2 < \infty\}$. The value for the improved estimator has been computed for different sample sizes and plotted in Figure 2. Decreasing the value of MSE by increasing the sample size indicates that the improved shrinkage estimator becoming more consistent.

Example 3. Let $X = (X_1, \dots, X_n)'$ be a random sample from a population with Inverse Gaussian distribution $IG(\theta, \theta)$. Let

$$T_1(X) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{X_i},$$
$$T_2(X) = \frac{1}{\bar{X}}, \tag{17}$$

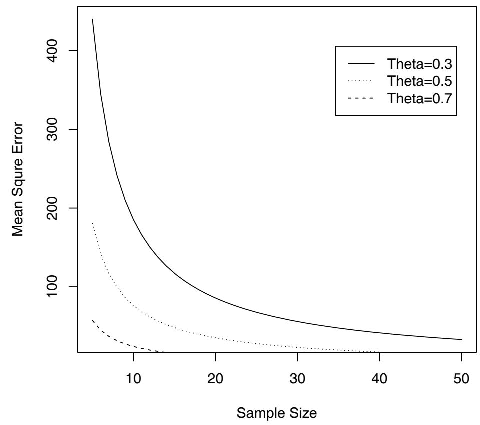

FIGURE 2: MSE of the improve shrinkage estimator for estimation of $g(\theta) = \theta$ versus sample size in N(θ,θ^2) distribution.

and $g(\theta) = \frac{1}{\theta}$. It is easy to see that the mean and variance of T_1 and T_2 are given by

$$\begin{aligned} E(T_1(X)) &= \frac{2}{\theta}, \\ Var(T_1(X)) &= \frac{3}{n\theta^2}, \end{aligned} \tag{18}$$

and

$$\begin{aligned} E(T_2(X)) &= \frac{n+1}{n\theta}, \\ Var(T_2(X)) &= \frac{n+2}{n^2\theta^2}, \end{aligned} \tag{19}$$

respectively. To compute the coefficient of correlation between T_1 and T_2, let $V = \sum_{i=1}^{n}(\frac{1}{X_i} - \frac{1}{\bar{X}})$, Then

$$\begin{aligned} Var(V) &= Var(\sum_{i=1}^{n}(\frac{1}{X_i} - \frac{1}{\bar{X}})) \\ &= \sum_{i=1}^{n} Var(\frac{1}{X_i} - \frac{1}{\bar{X}}) \\ &= n(Var(\frac{1}{X_i}) + Var(\frac{1}{\bar{X}}) - 2Cov(\frac{1}{X_i}, \frac{1}{\bar{X}})) \\ &= n(\frac{3}{\theta^2} + \frac{n+2}{n^2\theta^2} - 2Cov(\frac{1}{X_i}, \frac{1}{\bar{X}})). \end{aligned}$$

Therefore,

$$Cov(\frac{1}{X_i}, \frac{1}{\bar{X}}) = \frac{3}{2\theta^2} + \frac{n+2}{2n^2\theta^2} - \frac{1}{2n} Var(V). \tag{20}$$

From Inverse Gaussian distribution we know that $E(V^p) = \frac{2^p}{\theta^p}\frac{\Gamma(\frac{n+2p-1}{2})}{\Gamma(\frac{n-1}{2})}$, (see for example, Singh & Pandit 2008). Therefore,

$$\begin{aligned} Var(V) &= E(V^2) - E(V)^2 \\ &= \frac{4}{\theta^2}c(n), \end{aligned} \tag{21}$$

where $c(n) = \left[\frac{\Gamma(\frac{n+3}{2})}{\Gamma(\frac{n-1}{2})} - (\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n-1}{2})})^2\right]$. Substituting (21) in (20), we have

$$\begin{aligned} Cov(\frac{1}{X_i}, \frac{1}{\bar{X}}) &= \frac{3}{2\theta^2} + \frac{n+2}{2n^2\theta^2} - \frac{2}{n\theta^2}c(n) \\ &= k(n)\frac{1}{\theta^2}, \end{aligned}$$

where $k(n) = \left[\frac{3}{2} + \frac{n+2}{2n^2} - \frac{2c(n)}{n}\right]$. Therefore, the coefficient of correlation between T_1 and T_2 obtained is:

$$\begin{aligned} \rho_{T_1,T_2} &= \frac{Cov(T_1, T_2)}{\sqrt{Var(T_1)Var(T_2)}} \\ &= \frac{Cov(\frac{1}{n}\sum_{i=1}^n \frac{1}{X_i}, \frac{1}{\bar{X}})}{\sqrt{Var(T_1)Var(T_2)}} \\ &= \frac{\frac{1}{n}\sum_{i=1}^n Cov(\frac{1}{X_i}, \frac{1}{\bar{X}})}{\sqrt{Var(T_1)Var(T_2)}} \\ &= \frac{Cov(\frac{1}{X_i}, \frac{1}{\bar{X}})}{\sqrt{Var(T_1)Var(T_2)}} \\ &= h(n), \end{aligned} \tag{22}$$

where $h(n) = \frac{k(n)}{\sqrt{\frac{3n+6}{n^3}}}$ is a free from θ quantity. Considering the equations (18), (19) and (22), and based on the notation of Theorem 3.1, we have $k_1 = 2$, $k_2 = \frac{n+1}{n}$, $\tau_1^2 = \frac{3}{n}$, $\tau_2^2 = \frac{n+2}{n^2}$, $\lambda^2 = \frac{3n(n+1)^2}{4(n+2)^2}$ and $\rho_{T_1,T_2} = h(n)$. Therefore, according to equation (1), the improved linear combination of two estimators T_1 and T_2 is given by $T^*_{LC}(\mathbf{X}) = \alpha_1^* T_1(\mathbf{X}) + \alpha_2^* T_2(\mathbf{X})$ where,

$$\alpha_1^* = \frac{1 - h(n)\frac{(n+1)\sqrt{3n}}{2(n+2)}}{2\left(1 - \frac{h(n)(n+1)\sqrt{3n}}{n+2} + \frac{3n(n+1)^2}{4(n+2)^2} + (1-h(n)^2)(\frac{9}{4n^2})\right)},$$

$$\alpha_2^* = \frac{\frac{(n+1)\sqrt{3n}}{2(n+2)}\left(\frac{(n+1)\sqrt{3n}}{2(n+2)} - h(n)\right)}{\frac{n+1}{n}\left(1 - \frac{h(n)(n+1)\sqrt{3n}}{n+2} + \frac{3n(n+1)^2}{4(n+2)^2} + (1-h(n)^2)(\frac{9}{4n^2})\right)}.$$

This improved estimator, uniformly has the minimum MSE among all estimators in class $C_{T_1,T_2}(\alpha_1, \alpha_1) = \{\alpha_1 T_1 + \alpha_2 T_1\}$. The value of MSE for improved estimators has been computed for different sample sizes and plotted in Figure 3. Decreasing the value of MSE by increasing the sample size indicates that the improved shrinkage estimator becomes more consistent.

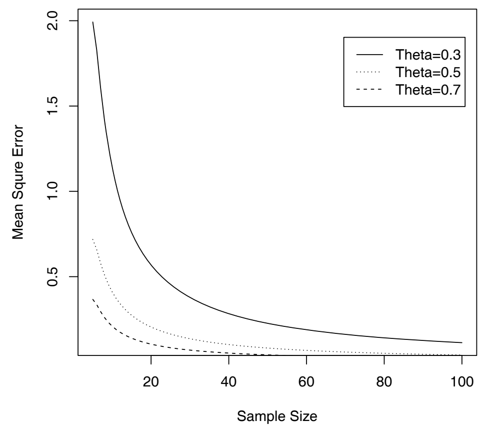

FIGURE 3: MSE of improved shrinkage estimator for the estimation of $g(\theta) = 1/\theta$ versus sample size in IG(θ,θ) distribution.

7. Discussion and Results

Sometimes the complete information about the interested parameter is distributed in two or more different estimators. In these situations, using only one of given estimators leads to loss of information of other estimators. Therefore, a combination of estimators must be employed to achieve a more efficient estimator. Moreover, it is interesting to look for improved estimators for a general form function of interested parameters, say $g(\theta)$. In recent years, some authors, notably Laheetharan & Wijekoon (2010), have been considered the problem in term of finding improved estimators for a function of an interested parameter. In this context, we have presented an optimal shrinkage estimator for a general form function of an interested parameter with an assumption of a known population coefficient of variation. We have also showed that two estimators obtained from the improved linear combination of two estimators and the linear combination of two improved estimators are equivalent, in terms of efficiency.

We think that using other coefficients of distributions, as additional information to be able to achieve a more efficient linear combination of two or more estimators, is an interesting field of research. Future studies will need to address this problem. Of course, it is our opinion that using the coefficient of variation in this direction, as an informative coefficient of distribution, will remain forever interesting. In fact, whenever prior information about the size of coefficient of variations is available, the shrinkage procedure could be useful. The possible results for some distributions with particular properties may be more interesting. For example, considering one-parameter exponential family of distributions is quite interesting. In some members of this distributions family such as normal, Poisson, gamma, binomial and negative binomial, it is known that variance is at most a quadratic function of the mean. Therefore, identifying the pertinent coefficients in the quadratic function is equivalent to determining the coefficient of the variations. As is obvious from theorems' assumptions, one can use any correlated or uncorrelated pair of estimators to construct an optimal linear combination to estimate any parametric function of an interested parameter. The results show that the efficiency of the proposed improved shrinkage estimator increase when the sample size increases.

[Received: June 2015 — Accepted: April 2016]

References

Arnholt, A. T. & Hebert, J. L. (1995), 'Estimating the mean with known coefficient of variation', *Journal of the American Statistical Associatio* **49**, 367–369.

Arnholt, A. T. & Hebert, J. L. (2001), 'Combinations of pairs of estimator', *Statistics on the Internet* pp. 1–8.

Bibby, J. (1972), 'Minimum means square error estimation, ridge regression and some unanswered questions', *Progress in Statistics* **1**, 107–121.

Bibby, J. & Toutenburg, H. (1977), *Prediction and Improved Estimation in Linear Models*, 1 edn, Wiley, New York.

Bibby, J. & Toutenburg, H. (1978), 'Improved estimation and prediction', *Zeitschriftur Angewandte Mathematik and Mechanik* **58**, 5–49.

Gleser, L. J. & Healy, J. D. (1976), 'Estimating the mean of a normal distribution with known coefficient of variation', *Journal of the American Statistical Association* **71**, 977–981.

Kanefuji, K. & Iwase, K. (1998), 'Estimation for a scale parameter with known coefficient of variation', *Statistical Papers* **39**, 377–388.

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.

Kleffe, J. (1985), Some remarks on improving unbiased estimators by multiplication with a constant, *in* T. Alinski & W. Klonecki, eds, 'Linear Statistical Inference', Cambridge.

Laheetharan, A. & Wijekoon, P. (2010), 'Improved estimation of the population parameters when some additional information is available', *Statistical Papers* **51**, 889–914.

Laheetharan, A. & Wijekoon, P. (2011), 'Mean square error comparison among variance estimators with known coefficient of variation', *Statistical Papers* **52**, 171–201.

Samuel-Cahn, E. (1994), 'Combining unbiased estimators', *The American Statistician* **48**, 34–36.

Searls, D. T. (1964), 'The utilization of known coefficient of variation in the estimation procedure', *Journal of the American Statistical Association* **59**, 1225–1226.

Searls, D. T. & Intarapanich, P. (1990), 'A note on an estimator for the variance that utilizes the kurtosis', *The American Statistician* **44**, 295–296.

Singh, H. P. & Pandit, S. (2008), 'Estimation of the reciprocal of the mean of the inverse gaussian distribution with prior information', *Statistica* **2**, 201–216.

Subhash, K. Y. & Cem, K. (2013), 'Improved exponential type ratio estimator of population variance', *Revista Colombiana de Estadística* **36**(1), 145–152.

Wencheko, E. & Chipoyera, H. W. (2005), 'Estimation of the variance when kurtosis is known', *Statistical Papers* **50**, 455–464.

Wencheko, E. & Wijekoon, P. (2007), 'Improved estimation of the mean in one-parameter exponential families with known coefficient of variation', *Statistical Papers* **46**, 101–115.