Sri Lankan Journal of Applied Statistics, Vol.9. 2008. p.84-92

Estimation of Variance in Inverse Gaussian Distribution with Known Kurtosis

A. LAHEETHARAN[1] AND P. WIJEKOON[2]

[1]**Department of Mathematics and Statistics, University of Jaffna, Sri Lanka**
[2]**Department of Statistics and Computer Science, University of Peradeniya, Sri Lanka**

ABSTRACT

This work concerns with the problem of estimating the population variance of inverse Gaussian distribution using the prior knowledge of kurtosis. A new estimator based on the sufficient statistic is obtained by using the method introduced by Laheetharan and Wijekoon (2007), and it is shown that this new estimator has a gain in efficiency with compared to the estimator based on the sample variance given by Searls and Intarapanich (1990) and Wencheko and Chipoyera (2007). To make the comparison, the scalar mean square error loss function is utilized, and the relative efficiency of the estimator is shown by considering different values of kurtosis with increasing sample size.

Key words: Inverse Gaussian, Kurtosis, Optimal shrunken estimator, Scaled mean square error loss

INTRODUCTION

The inverse Gaussian distribution (IG) known as the first passage time distribution of rownian motion with positive drift, is now widely used to model positively skewed data in diverse areas of applied research as cardiology, hydrology, demography, linguistics, employment service, labor disputes and finance (see, Chhikara and Folks, 1989; Seshadri, 1999). One of the problems that the researchers are involved is the estimation of mean and variance of this distribution. Korwar (1980) has derived the uniformly minimum variance unbiased (UMVU) estimator of the variance and its reciprocals, and Iwase and Setô (1983) have obtained the uniformly minimum variance unbiased (UMVU) estimator of the cumulants including mean and variance of this distribution. The estimation of mean and variance having known prior information such as coefficient of variation, kurtosis or skewness of the distribution concerned is considered recently in the literature. Searls (1964), Khan (1968) and Arnholt and Hebert (1995) utilized the known coefficient of variation on estimating the population mean, and recently Wencheko and Wijekoon (2005) further improved their results and derived shrunken estimators for of variation on estimating the population mean, and recently Wencheko and Wijekoon (2005) further improved their results and derived shrunken estimators for the mean in one parameter exponential families. Searls and Intarapanich (1990) and Wencheko and Chipoyera (2007) have developed alternative estimators for the population variance over the traditional sample variance when kurtosis is known, and showed that their estimators have minimum mean square error over the sample variance.

The goal of this paper is to derive an improved estimator over the estimator given by Wencheko and Chipoyera (2007) for the variance of the inverse Gaussian distribution (IG) when kurtosis is known. Some applications of known IG- kurtosis are presented in Kolodziejski and Betz (2000). To assess the efficiency of the new estimator, the scalar mean square error loss function is utilized, and the relative efficiency of the estimator is shown by considering different values of kurtosis by increasing the sample size.

OPTIMAL SHRUNKEN ESTIMATORS

Gleser and Healy (1976) have considered the minimization of mean square error of linear combination of two uncorrelated and unbiased estimators of θ having a known coefficient of variation. By improving their results Laheetharan and Wijekoon (2007) have developed a general method to estimate a function of θ, say $g(\theta)$, when prior information (coefficient of variation, kurtosis, skewness etc.) is available, and this can be used to estimate the mean and variance of distributions since they are functions of unknown parameters. According to their method a shrunken estimator for $g(\theta)$ can be obtained, and this estimator has uniformly minimum mean square error (MSE) in a certain class. These results are summarized in the following theorem:

Theorem 1 (Laheetharan and Wijekoon, 2007)

Let $X = (X_1,\dots,X_n)'$ 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 $T(X)$ be a point estimator of $g(\theta)$ with $E[T(X)] = kg(\theta)$ where $k \in \Re$, and without loss of generality, assume that $k > 0$. If the ratio $\tau^2 = [g(\theta)]^{-2} Var[T(X)]$ is independent of θ, then $T^*(X) = \alpha^* T(X)$ has uniformly minimum MSE (in $g(\theta)$) among all estimators that are in the class $C_T(\alpha) = \{\alpha T(X) \mid 0 < \alpha < \infty\}$, where

$\alpha^* = k/(k^2 + \tau^2)$, and the minimum MSE is given by $MSE[T^*(X)] = \tau^2 (k^2 + \tau^2)^{-1} [g(\theta)]^2$.

Laheetharan and Wijekoon (2007) have applied this method to derive class of shrunken estimators for both mean and variance of natural exponential family of distributions with known coefficient of variation. However this theorem can be applied not only to obtain shrunken estimators for the exponential family of distributions but also to other distributions which are not members of this family. When applying this theorem to find shrunken estimators for the population mean and variance, $g(\theta)$ is taken as the mean or variance of the distribution, and $T(X)$, preferably a sufficient statistic if available, is a point estimator of $g(\theta)$.

IG DISTRIBUTION AND SHRUNKUN ESTIMATORS FOR IT'S VARIANCE

Suppose the random variable X_i has an inverse Gaussian distribution with mean $\mu > 0$ and scaling parameter $\lambda > 0$, and we write $X_i \sim IG(\mu, \lambda), i = 1 \cdots n$. The probability density function of IG distribution is given by

$$f(x, \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left[\frac{-\lambda(x-\mu)^2}{2x\mu^2}\right] \quad (1)$$

with $E(X_i) = \mu$ and $Var(X_i) = \mu^3/\lambda$. Hence the joint probability density function is

$$f(x_1, \ldots, x_n \mid \mu, \lambda) = \left(\frac{\lambda}{2\pi}\right)^{n/2} \left(\prod_{i=1}^{n} x_i^{-3/2}\right) \exp\left\{\frac{-\lambda}{2}\sum_{i=1}^{n}\frac{1}{x_i}\right\} \exp\left\{\frac{-\lambda}{2\mu^2}\sum_{i=1}^{n} x_i + \frac{n\lambda}{\mu}\right\}, \quad (.2)$$

and that implies the distribution belongs to the exponential family. The statistic $T(X) = \sum_{i=1}^{n} X_i$ is sufficient, and the first four cumulants of this distribution are

$\kappa_1 = E[T(X)] = n\mu$, $\kappa_2 = Var[T(X)] = n\mu^3/\lambda$, $\kappa_3 = 3n\mu^5/\lambda^2$, and $\kappa_4 = 15n\mu^7/\lambda^3$.

Note that the kurtosis is the degree of peakedness of a distribution, and defined by $\alpha_4 = (\mu_4/\mu_2^2) - 3 = \kappa_4/\kappa_2^2$, which is the normalized form of the fourth central moment of the distribution, where $\mu_j = E[X - E(X)]^j$ is the j^{th} central moment and κ_r is the r th cumulant of the distribution.

The kurtosis of the IG-distribution is given by $\alpha_4 = \kappa_4/\kappa_2^2 = 15\mu/\lambda$, and hence $\lambda = 15\mu/\alpha_4$. Thus $X_i \sim IG(\mu, 15\alpha_4^{-1}\mu)$, and this indicates the distribution belongs to the curved exponential family when the kurtosis α_4 is known. Now using cumulants we derive the moments of $T(X)$ as follows: (Lehmann and Casella, 1998)

$$E(T(X))^2 = \kappa_2 + \kappa_1^2 = n((\mu/\lambda) + n)\mu^2 = n\,((\alpha_4/15) + n)\mu^2. \tag{3}$$

$$E(T(X))^3 = \kappa_3 + 3\kappa_1\kappa_2 + \kappa_1^3 = n(n^2 + 3n\,\mu/\lambda + 3\,\mu^2/\lambda^2)\mu^3 = n\,(n^2 + n\,\alpha_4/5 + \alpha_4^2/75)\,\mu^3 \tag{4}$$

$$\begin{aligned} E(T(X))^4 &= \kappa_4 + 3\kappa_2^2 + 4\kappa_1\kappa_3 + 6\kappa_1^2\kappa_2 + \kappa_1^4 = n(n^3 + 6n^2\,\mu/\lambda + 15n\,\mu^2/\lambda^2 + 15\,\mu^3/\lambda^3)\mu^4 \\ &= n\,(n^3 + 2n^2\,\alpha_4/5 + n\,\alpha_4^2/15 + \alpha_4^3/225)\,\mu^4 \end{aligned} \tag{5}$$

Note that according to (3) the statistic $(T(X))^2 = \left(\sum_{i=1}^{n} X_i\right)^2$ can be used as an estimator for μ^2. Since $Var(X_i) = \alpha_4\mu^2/15$ this estimator is important for estimating the variance of the distribution when kurtosis is known. Now we consider the optimal estimator for μ^2 belongs to the class $C_{T_1}(\alpha) = \{\alpha(T(X))^2 \mid 0 < \alpha < \infty\}$, where $T(X) = \sum_{i=1}^{n} X_i$.

Let $T_1(X) = [T(X)]^2 = \left(\sum_{i=1}^{n} X_i\right)^2$ and $g(\mu) = \alpha_4\mu^2/15$. Then according to (3) $E[T_1(X)] = kg(\mu)$ with $k = 15n\,((\alpha_4/15) + n)/\alpha_4$. Also using (3) and (5) we obtain

$$\begin{aligned} Var[(T_1(X))] &= Var[(T(X))^2] = E(T(X))^4 - (E(T(X))^2)^2 \\ &= n(n^3 + 2n^2\,\alpha_4/5 + n\alpha_4^2/15 + \alpha_4^3/225)\mu^4 - [n((\alpha_4/15) + n)\mu^2]^2. \end{aligned} \tag{6}$$

Now by applying theorem 1 we can show that $\alpha^* = \frac{k}{k^2+\tau^2} = \frac{15(15n+\alpha_4)}{225n^3+90n^2\alpha_4+15n\alpha_4^2+\alpha_4^3}$ with

$\tau^2 = n(n^3+2n^2\alpha_4/5+n\alpha_4^2/15+\alpha_4^3/225)-[n((\alpha_4/15)+n)]^2$ which is independent of unknown parameters. Hence the optimal shrunken estimator of $g(\mu)=Var(X_i)$ in the class $C_{T_1}(\alpha)=\{\alpha(T(X))^2 \mid 0<\alpha<\infty\}$ is

$$T_1^*(X) = \frac{15(15n+\alpha_4)}{225n^3+90n^2\alpha_4+15n\alpha_4^2+\alpha_4^3}\left(\sum_{i=1}^{n} X_i\right)^2 \tag{7}$$

Note that the estimator $T_1^*(X)$ is based on the sufficient statistic $\sum_{i=1}^{n} X_i$. Since the common estimator for the variance is the sample variance $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ now we derive another estimator based on S^2, which is a uniformly minimum mean square error estimator in the class $C_{T_2}(\alpha)=\{\alpha S^2 \mid 0<\alpha<\infty\}$. Note that $E(S^2)=\sigma^2=\mu^3/\lambda$, and

$$Var(S^2)=\frac{1}{n}\left[\kappa_4+3\kappa_2^2-\left(\frac{n-3}{n-1}\right)\kappa_2^2\right]=\frac{1}{n}\left[\frac{\kappa_4}{\kappa_2^2}+3-\left(\frac{n-3}{n-1}\right)\right]\kappa_2^2=\frac{1}{n}\left[\alpha_4+3-\left(\frac{n-3}{n-1}\right)\right]\sigma^4 \tag{8}$$

Now we apply theorem 1 again with $T_2(X)=S^2, k=1$, and $g(\sigma^2)=\sigma^2$. Then we can show that $\tau^2 = \frac{Var[T(X)]}{[g(\sigma^2)]^2} = \frac{Var(S^2)}{\sigma^4} = \frac{1}{n}\left[\alpha_4+3-\left(\frac{n-3}{n-1}\right)\right]$, which is independent of σ^2, and $\alpha^* = \frac{k}{k^2+\tau^2} = \frac{1}{1+n^{-1}[\alpha_4+3-(n-3)(n-1)^{-1}]} = \frac{1}{\alpha_4/n+(n+1)/(n-1)}$.

This gives the optimal shrunken estimator

$$T_2^*(X) = \frac{1}{(\alpha_4/n) + ((n+1)/(n-1))} S^2 \quad (9)$$

for the variance $\sigma^2 = \mu^3/\lambda$ in the class $C_{T_2}(\alpha) = \{\alpha S^2 \mid 0 < \alpha < \infty\}$. Note that this estimator is already proposed by Searls and Intarapanich (1990), and Wencheko and Chipoyera (2007), and it is interesting to notice that the same estimator can be obtained by applying theorem1.

MSE COMPARISON OF ESTIMATORS

To assess the efficiency of estimators $T_1^*(X)$ and $T_2^*(X)$, the Scalar Mean Square Error Loss (SMSEL) function will be used, since it is more appropriate for the estimation of a scale parameter and is independent of unknown parameters. The SMSEL function is defined as $MSE_L(\hat{\theta}) = E_\theta((\hat{\theta} - \theta)/\theta)^2 = MSE(\hat{\theta})/\theta^2$, where MSE($\hat{\theta}$) is the mean squared error of the estimator $\hat{\theta}$ of an unknown parameter θ. (Refer Kanfuji and Iwase, 1998). Then for the estimator $T^*(X)$ defined in theorem 1 we can easily show that $MSE_L[T^*(X)] = MSE[T^*(X)]/(g(\theta))^2 = \tau^2(k^2 + \tau^2)^{-1}$. Note that $MSE_L = 0$ when the estimated value is equal to the actual value so that the distance between $T^*(X)$ and $g(\theta)$ is around zero.

Note that the scalar mean square error loss of $T_1^*(X)$ defined on the class $C_{T_1}(\alpha)$ is

$$MSE_L[T_1^*(X)] = \tau^2(k^2 + \tau^2)^{-1} = \left\{ \frac{\alpha_4(60n^2 + 14n\alpha_4 + \alpha_4^2)}{225n^3 + 90n^2\alpha_4 + 15n\alpha_4^2 + \alpha_4^3} \right\}, \quad (10)$$

and the scalar mean square error loss of $T_2^*(X)$ defined on the class $C_{T_2}(\alpha)$ is

$$MSE_L[T_2^*(X)] = \tau^2(k^2 + \tau^2)^{-1} = \frac{\alpha_4(n-1) + 2n}{\alpha_4(n-1) + n(n+1)}. \quad (11)$$

Now we investigate the MSE_L of $T_1^*(X)$ and $T_2^*(X)$ by increasing the sample size for three different values of kurtosis, and the results are shown in the figure 1.

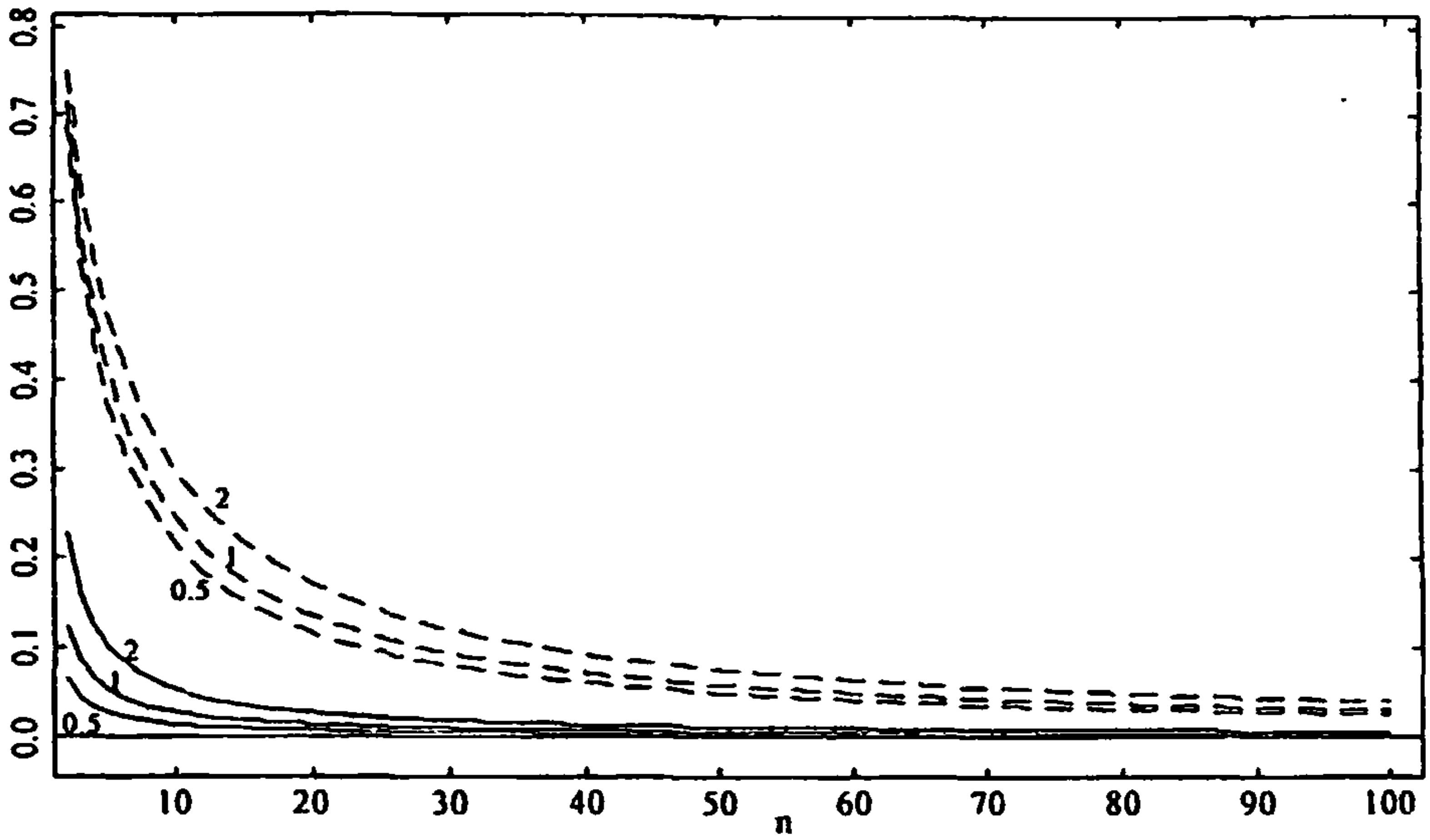

Figure 1: MSE_L versus n (>1) the variance of shrunken estimators of IG distribution for kurtosis = 0.5, 1and 2.

Figure 1 demonstrates that $T_2^*(X)$ is always less efficient than $T_1^*(X)$. Also MSE_L of both estimators depend on the kurtosis α_4, and as α_4 decreases, MSE_L decreases. When sample size increases the MSE_L of both estimators approximate to zero.

The relative efficiency (RE) of the two estimators is given by

$$\text{RE} = \frac{MSE\left[T_2^*(X)\right]}{MSE\left[T_1^*(X)\right]} = \frac{\left(\alpha_4(n-1)+2n\right)\left(225n^3+90n^2\alpha_4+15n\alpha_4^2+\alpha_4^3\right)}{\alpha_4\left(\alpha_4(n-1)+n(n+1)\right)\left(60n^2+14n\alpha_4+\alpha_4^2\right)} \qquad (12)$$

Figure 2 shows the behaviour of RE versus sample size for three different values of kurtosis.

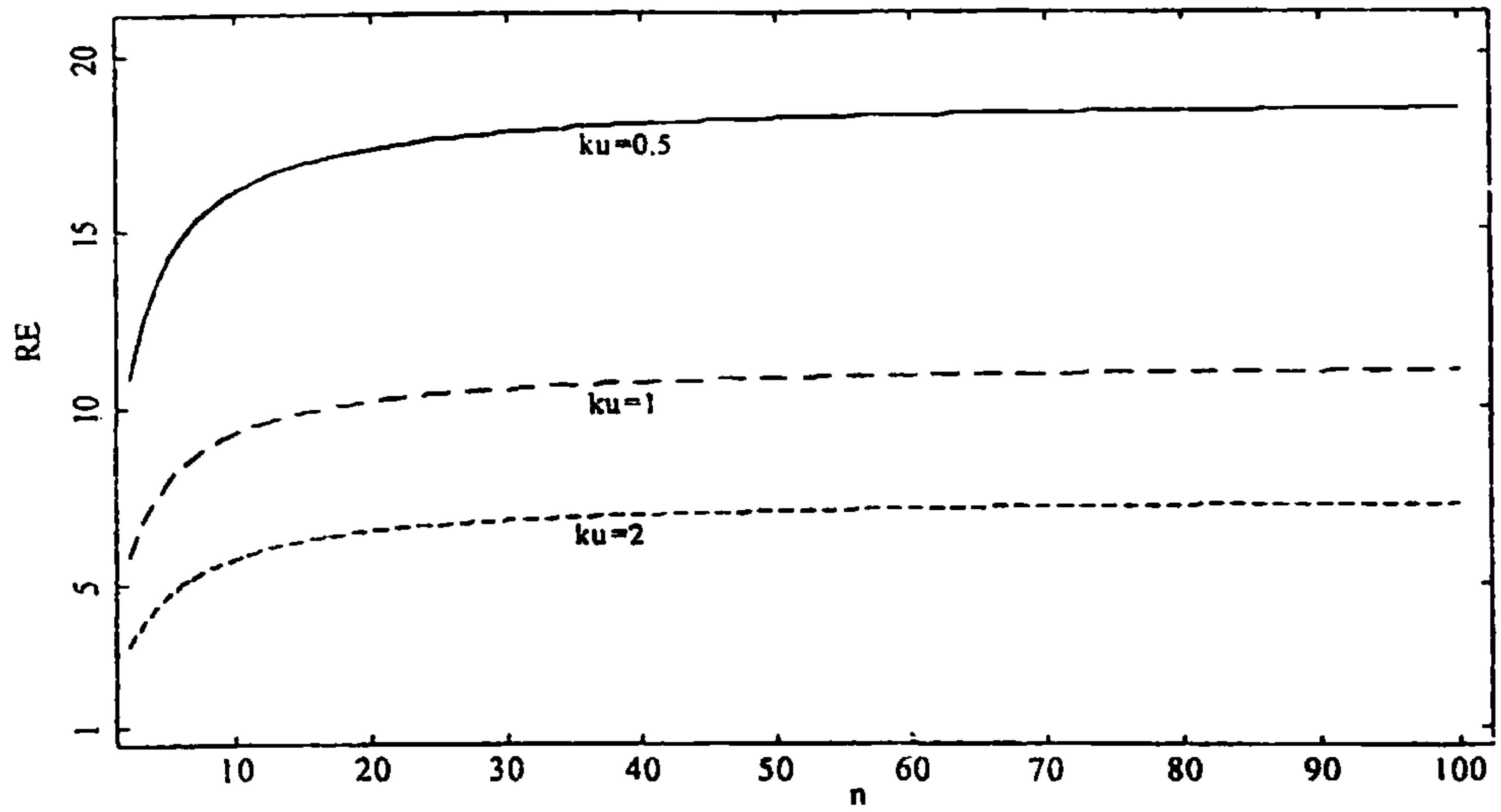

Figure 2: Relative efficiency (RE) versus sample size *n* (>1) for kurtosis (ku) = 0.5, 1, 2.

According to Figure 2, it is clear that $T_1^*(X)$ is always more efficient than $T_2^*(X)$, and efficiency increases as n increases up to a certain level and then stabilizes. However the RE decreases when kurtosis α_4 increases.

CONCLUSION

It has been seen that the estimator $T_1^*(X)$ based on the sufficient statistic has a considerable gain in efficiency with respect to the estimator $T_2^*(X)$ based on the sample variance when kurtosis is known even for a small sample. For large samples, $T_1^*(X)$ is still more efficient, and both estimators have very low Scaled Mean Square Error Loss.

REFERENCES

Arnholt, A.T. and J.L. Hebert (1995). Estimating the mean with known coefficient of variation. *The American Statistician,* 49: 367-369.

Chhikara R.S. and L. Folks (1989). The inverse Gaussian distribution: Theory, methodology and applications, CRC Press.

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

Iwase, K, and N. Setô (1983). Uniformly minimum variance unbiased estimation for the inverse Gaussian distribution. *Journal of the American Statistical Association,* 78: 660-663.

Kanefuji, K. and K. Iwase (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.

Kolodziejski, K.R. and J.W. Betz (2000). Detection of weak random signals in IID Non-Gausian noise. *IEEE Transactions on Communications*, 48: 222-230.

Korwar, R.M. (1980). On the Uniformly minimum variance unbiased estimators of the variance and its reciprocal of an inverse Gaussian distribution. *Journal of the American Statistical Association,* 75: 734-735.

Laheetharan, A. and P. Wijekoon (2007). Deriving shrunken estimators for the variance. in one-parameter natural exponential families. Proceedings of the PURSE, University of Peradeniya, 12(Part II): 281-283.

Lehmann, E.L. and G. Casella (1998). *Theory of Point Estimation, 2nd Edition.* Springer-Verlag, Berlin Heidelberg, New York.

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.

Searls, D.T. and P. Intarapanich (1990). A note on an estimator for the variance that utilizes the kurtosis. *The American Statistician,* 44: 295-296.

Seshadri, V. (1999) *The Inverse Gaussian Distribution: Statistical Theory and Applications*, Springer Verlag, New York.

Wencheko, E. and H.W. Chipoyera (2007). Estimation of the variance when kurtosis is known. *Statistical Papers*, DOI: 10.1007/s00362-007-0084-1.

Wencheko. E. and P. Wijekoon (2005). Improved estimation of the mean in one-parameter exponential families with known coefficient of variation, *Statistical Papers*, 46: 101-115.