STATISTICA, anno LXXVII, n. 1, 2017

ON SEQUENTIAL ESTIMATION OF A NORMAL DISTRIBUTION HAVING EQUAL MEAN AND VARIANCE

Saralees Nadarajah [1]
University of Manchester, Manchester M13 9PL, UK
Idika E. Okorie
University of Manchester, Manchester M13 9PL, UK

1. Introduction

The normal distribution with equal mean and variance arises in many applied areas: collective theory of risk (Ammeter (1962)); contracts and supply assurance in the UK health care market (Fenn *et al.* (1994)); estimation of individual asymmetry (Dongen (1999)); the effect of patenting on the networks and connections of academic scientists (Forti *et al.* (2007)); gene expression data biclustering stability (Badea and Tilivea (2008)); first passage percolation on random graphs with finite mean degrees (Bhamidi *et al.* (2010)); postmortem body cooling (Kaliszan (2011)); detection of ADC clipping, quantization noise, and amplifier saturation in surface electromyography (Fraser *et al.* (2012)); rainfall frequency analysis (E. S. Chung (2013)); finding discriminatory genes (Khan and Greiner (2013)); impact of pacemaker failover configuration on mean time to recovery for small cloud clusters (Benz and Bohnert (2014)); interaction networks underlying the minority game (Caridi (2014)); to mention just a few.

This paper relates to one of the latest estimation problems for the normal distribution with equal mean and variance. Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed observations from a normal distribution with both the mean and variance equal to θ. Mukhopadhyay and Cicconetti (2004) derived (among other things) the MLE and the UMVUE of θ and discussed their application to purely sequential and two-stage bounded risk estimation of θ. The MLE of θ was given by

$$\widehat{\theta}_{1n} = \sqrt{T_n + \frac{1}{4}} - \frac{1}{2}, \tag{1}$$

where

$$T_n = \frac{1}{n} \sum_{i=1}^{n} X_i^2.$$

[1] Corresponding author. E-mail: mbbsssn2@manchester.ac.uk

The UMVUE of θ was given by

$$\widehat{\theta}_{5n} = b(u,n) n^{-1/2} I(u,n), \tag{2}$$

where

$$I(u,n) = \int_{-\sqrt{u}}^{\sqrt{u}} y \exp\left(\sqrt{n} y\right) \left(u - y^2\right)^{(n-3)/2} dy, \tag{3}$$

$$b(u,n) = \frac{1}{\sqrt{\pi}\Gamma\left((n-1)/2\right) q(u,n)}, \tag{4}$$

$$q(u,n) = \sum_{k=0}^{\infty} \frac{u^{n/2+k-1} n^k}{2^{2k} \Gamma\left(n/2+k\right) k!} \tag{5}$$

and

$$u = \sum_{i=1}^{n} X_i^2.$$

Mukhopadhyay and Cicconetti (2004) used the formulas given by the MLE (1) and the UMVUE (2) for purely sequential and two-stage bounded risk estimation of θ. Mukhopadhyay and Cicconetti (2004), however, encountered difficulties in evaluating the UMVUE (2) and this limited the use of (2) as a sequential estimator for θ. Mukhopadhyay and Cicconetti (2004) stated, for instance, that the UMVUE (2) can be reduced to a "more closed form expression" and evaluated "quickly and exactly" only when n is odd. Mukhopadhyay and Cicconetti (2004) also stated that the complexity of the UMVUE (2) "increases as n increases and quickly renders computation" of the UMVUE (2) intractable. As far as we can see, much of the discussion in Mukhopadhyay and Cicconetti (2004) concerning the evaluation of the UMVUE (2) appears to be not correct, as shown in Sections 2 and 3. We feel that it is important that we point this out especially since Mukhopadhyay and Cicconetti Mukhopadhyay and Cicconetti (2004) has been cited by several other papers, see Mukhopadhyay (2006), Choi (2005), Kim *et al.* (2007), Mukhopadhyay (2008), Mukhopadhyay and Bhattacharjee (2010), Bhattacharjee and Mukhopadhyay (2011) and S. Banerjee (2016).

The aim of this paper is two folded. Firstly, we derive a much simpler expression for the UMVUE (2). This expression turns out to be a ratio of two modified Bessel functions of the first kind, see Section 2. Secondly, using the derived formula, we compare the performance of the MLE (1) versus the UMVUE (2) for purely sequential estimation of θ, see Section 3.

Various special functions are used in Sections 2 and 3. Their definitions and detailed properties can be found in

`http://functions.wolfram.com/Bessel-TypeFunctions/BesselI/`,
`http://functions.wolfram.com/Bessel-TypeFunctions/StruveL/`,
`http://functions.wolfram.com/HypergeometricFunctions/Hypergeometric0F1/`,
`http://functions.wolfram.com/HypergeometricFunctions/Hypergeometric1F1/`.

Some of the properties in these links are used for the derivations in Section 2.

The computations for Sections 2 and 3 were performed using the R and Maple software in a Windows 10 environment. The R version used was R 3.3.0 (R Development Core Team (2016)). The Maple version used was Maple 2015.

2. Simpler expression for the UMVUE (2)

Here, we would like to show that the UMVUE (2) can be reduced in terms of a well known special function for any real number $n > 1$. Note that

$$
\begin{aligned}
I(u,n) &= \left\{ \int_0^{\sqrt{u}} y \exp\left(\sqrt{n} y\right) \left(u - y^2\right)^{(n-3)/2} dy \right. \\
&\qquad \left. - \int_0^{\sqrt{u}} y \exp\left(-\sqrt{n} y\right) \left(u - y^2\right)^{(n-3)/2} dy \right\} \\
&= 2 \int_0^{\sqrt{u}} y \left(u - y^2\right)^{(n-3)/2} \sinh\left(\sqrt{n} y\right) dy \\
&= \sqrt{\pi} u^{n/4} 2^{n/2-1} n^{1/2-n/4} \Gamma\left(\frac{n-1}{2}\right) I_{n/2}\left(\sqrt{un}\right), \qquad (6)
\end{aligned}
$$

where the last step follows by equation (2.4.3.11) in Prudnikov *et al.* (1986) and $I_\nu(\cdot)$ denotes the modified Bessel function of the first kind of order ν defined by

$$
I_\nu(x) = \sum_{k=0}^{\infty} \frac{1}{k!\Gamma(k+\nu+1)} \left(\frac{x}{2}\right)^{2k+\nu}. \qquad (7)
$$

The equality in (6) holds for any real number $n > 1$. Using the definition in (7), one can express $q(u,n)$ in (5) as

$$
q(u,n) = u^{n/4-1/2} n^{1/2-n/4} 2^{n/2-1} I_{n/2-1}\left(\sqrt{un}\right). \qquad (8)
$$

Combining (2)-(4), (6) and (8), one obtains the simple form

$$
\widehat{\theta}_{5n} = \sqrt{\frac{u}{n}} \frac{I_{n/2}\left(\sqrt{un}\right)}{I_{n/2-1}\left(\sqrt{un}\right)}. \qquad (9)
$$

The closed form expression in (9) is not only mathematically simpler than (2) which involves an infinite series and an integral. The closed form expression can also lead to more efficient computation of $\widehat{\theta}_{5n}$ in terms of computational time and computational accuracy. In-built routines are usually based on efficient algorithms. The direct computation of infinite sums or integrals is usually not so efficient and can lead to round off errors.

There are many in-built routines for computing the modified Bessel function of the first kind. In the R software, it can be computed using besselI. In Maple, it can be computed using BesselI. The following code in R was used for computing the UMVUE (9).

```
f=function (u,n)
{
   t1=besselI(sqrt(u*n),nu=(n/2),expon.scaled=TRUE)
   t2=besselI(sqrt(u*n),nu=((n/2)-1),expon.scaled=TRUE)
   tt=sqrt(u/n)*t1/t2
   return(tt)
}
```

This code was able to compute the UMVUE (9) for a wide range of values of u and n without any computational problems. Figure 1 plots the UMVUE (9) versus $n = 1, 2, \ldots, 100$ for a range of different values of u. The UMVUE (9) appears to be a decreasing function of n and an increasing function of u.

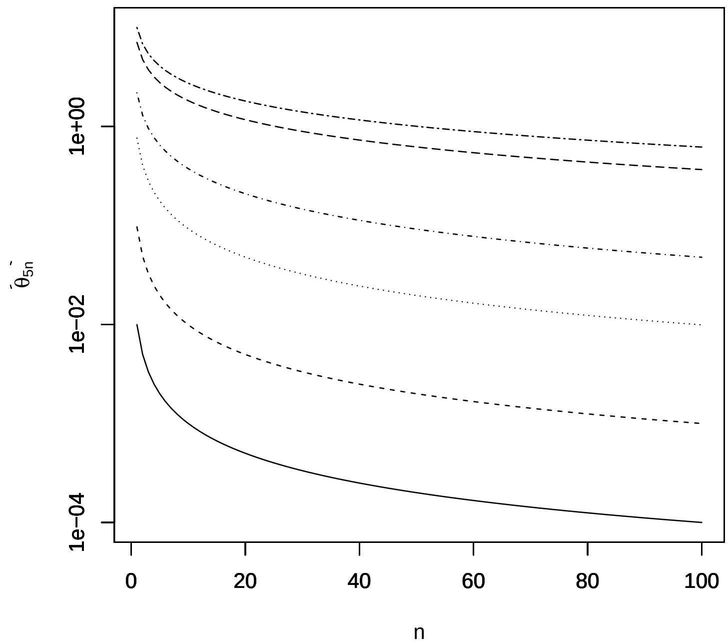

Figure 1 – $\widehat{\theta}_{5n}$ versus n for $u = 0.01$ (solid line), $u = 0.1$ (dashed line), $u = 1$ (dotted line), $u = 5$ (dotdash line), $u = 50$ (longdash line) and $u = 100$ (twodash line). The y axis is in log scale.

To check the accuracy of the R code, we also computed the UMVUE (9) using BesselI in Maple. Maple like most other algebraic manipulation packages allows for arbitrary precision, so the accuracy of the values computed using Maple was not an issue. These values were plotted on the same axes of Figure 1. There appears to be no visual distinction between the values computed using Maple and R. Hence, the values computed using the R code can be considered accurate enough.

Mukhopadhyay and Cicconetti (2004) claimed that $\widehat{\theta}_{5n}$ can be computed much more accurately for odd n and that the accuracy "also appears to be affected adversely when u is either large or small". But there is no evidence of this in Figure 1. According to this figure, $\widehat{\theta}_{5n}$ can be computed equally accurately for all $n = 1, 2, \ldots, 100$ and for a range of values of u including small u and large u.

Mukhopadhyay and Cicconetti (2004) also claimed that $\widehat{\theta}_{5n}$ "can be evaluated quickly and exactly based on observed data only when n is odd" and the computational complexity of $\widehat{\theta}_{5n}$ "increases many fold as n increases and quickly renders computation of $\widehat{\theta}_{5n}$" intractable. To verify this, we plotted the central processing unit time taken for ten thousand computations of $\widehat{\theta}_{5n}$ versus $n = 1, 2, \ldots, 100$ for a range of values of u, see Figure 2. We see that the central processing unit times increase only slightly with increasing n. There are no significant changes in the central processing unit time between n odd and n even. There are also no significant changes in the central processing unit time with respect to u.

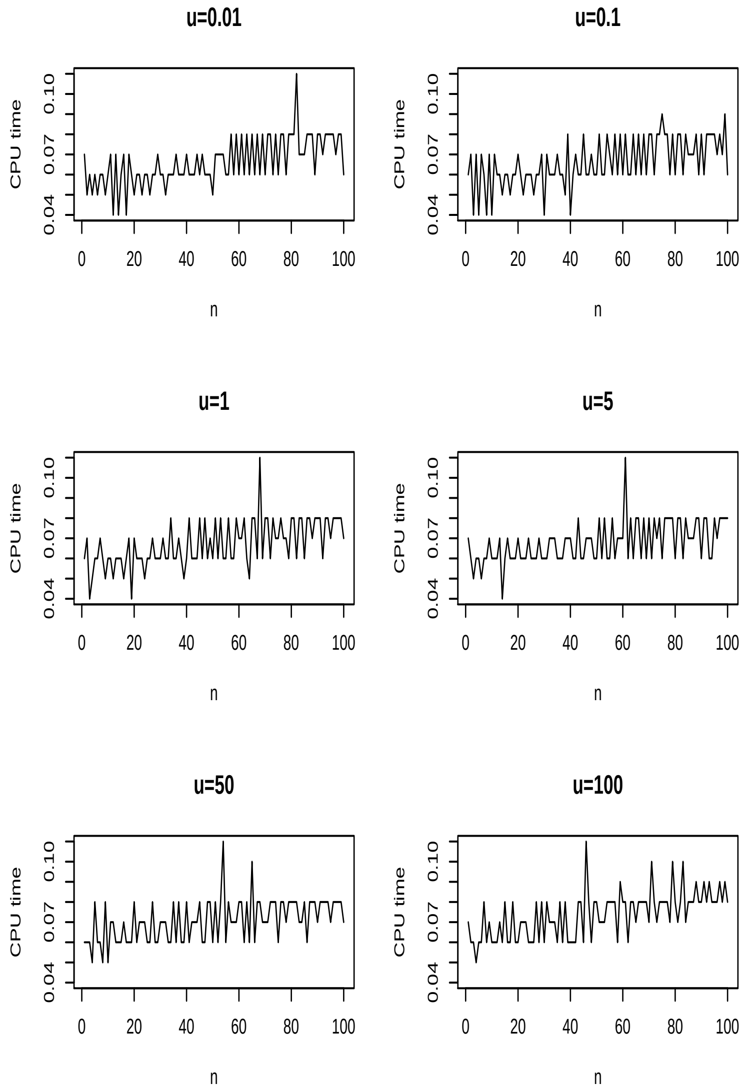

Figure 2 – Central processing unit time taken to compute $\widehat{\theta}_{5n}$ ten thousand times versus n for $u = 0.01, 0.1, 1, 5, 50, 100$.

It is well known that $I_\nu(\cdot)$ takes an elementary form if ν is a half integer. Actually, if $\nu - \frac{1}{2}$ is an integer then

$$\begin{aligned} I_\nu(z) \quad = \quad & -\sqrt{\frac{2}{\pi z}} \exp\left[\frac{\pi \mathrm{i}}{2}\left(\frac{1}{2}-\nu\right)\right] \left\{ \sinh\left[\frac{\pi \mathrm{i}}{2}\left(\frac{1}{2}-\nu\right) - z\right] \right. \\ & \cdot \sum_{k=0}^{\left[\frac{2|\nu|-1}{4}\right]} \frac{\left(\mid \nu \mid +2k - \frac{1}{2}\right)!}{(2k)!\left(\mid \nu \mid -2k-\frac{1}{2}\right)!(2z)^{2k}} \\ & + \cosh\left[\frac{\pi \mathrm{i}}{2}\left(\frac{1}{2}-\nu\right) - z\right] \\ & \left. \cdot \sum_{k=0}^{\left[\frac{2|\nu|-3}{4}\right]} \frac{\left(\mid \nu \mid +2k + \frac{1}{2}\right)!}{(2k+1)!\left(\mid \nu \mid -2k-\frac{3}{2}\right)!(2z)^{2k+1}} \right\}, \qquad (10) \end{aligned}$$

where $\mathrm{i} = \sqrt{-1}$ and $[x]$ denotes the largest integer less than or equal to x. In particular,

$$\begin{aligned} I_{-1/2}(z) &= \sqrt{\frac{2}{\pi}} \frac{\cosh(z)}{\sqrt{z}}, \\ I_{1/2}(z) &= \sqrt{\frac{2}{\pi}} \frac{\sinh(z)}{\sqrt{z}}, \\ I_{3/2}(z) &= \sqrt{\frac{2}{\pi}} \frac{z\cosh(z) - \sinh(z)}{z^{3/2}}, \\ I_{5/2}(z) &= \sqrt{\frac{2}{\pi}} \frac{\left(z^2+3\right)\sinh(z) - 3z\cosh(z)}{z^{5/2}}, \end{aligned}$$

and so on. So, the UMVUE (9) reduces to an elementary form if n is a positive odd integer.

Mukhopadhyay and Cicconetti (2004) also claimed that $\widehat{\theta}_{5n}$ reduces to a "closed-form expression solution for odd values of $n \geq 5$". Actually, $\widehat{\theta}_{5n}$ reduces to a closed-form expression for all odd values of $n \geq 1$. For $n = 1, 3$, the UMVUE (9) reduces to

$$\widehat{\theta}_{5n} = \sqrt{u}\tanh\left(\sqrt{u}\right)$$

and

$$\widehat{\theta}_{5n} = \frac{1}{\sqrt{3}}\left[\sqrt{u}\coth\left(\sqrt{u}\right) - 1\right],$$

respectively.

Equivalent representations for the UMVUE (9) in terms of other special functions are

$$\widehat{\theta}_{5n} = \sqrt{\frac{u}{n}} \frac{L_{-n/2}\left(\sqrt{un}\right)}{L_{1-n/2}\left(\sqrt{un}\right)},$$

$$\widehat{\theta}_{5n} = \frac{u}{n} \frac{{}_0F_1\left(; \frac{n}{2}+1; \frac{un}{4}\right)}{{}_0F_1\left(; \frac{n}{2}; \frac{un}{4}\right)}$$

and

$$\widehat{\theta}_{5n} = \frac{u}{n} \frac{{}_1F_1\left(\frac{n+1}{2}; n+1; 2\sqrt{un}\right)}{{}_1F_1\left(\frac{n-1}{2}; n-1; 2\sqrt{un}\right)},$$

where $L_\nu(\cdot)$ denotes the Struve L function defined by

$$L_\nu(z) = \left(\frac{z}{2}\right)^{\nu+1} \sum_{k=0}^{\infty} \frac{1}{\Gamma\left(k+\frac{3}{2}\right)\Gamma\left(k+\nu+\frac{3}{2}\right)} \left(\frac{z}{2}\right)^{2k}$$

while ${}_0F_1(\cdots)$ and ${}_1F_1(\cdots)$ denote hypergeometric functions defined by

$${}_0F_1(;a;z) = \sum_{k=0}^{\infty} \frac{1}{(a)_k} \frac{z^k}{k!}$$

and

$${}_1F_1(a;b;z) = \sum_{k=0}^{\infty} \frac{(a)_k}{(b)_k} \frac{z^k}{k!},$$

respectively, where $(e)_k = e(e+1)\cdots(e+k-1)$ denotes the ascending factorial.

3. Sequential estimation of θ

Here, we compare the performance of the UMVUE (2) versus the MLE (1) for purely sequential estimation of θ. The MLE-based and the UMVUE-based sequential estimators of θ are given by (1) and (2), respectively, with the sample size n defined by the stopping rule: the smallest integer $n \geq m$ such that $n \geq a^* w^{-1} 2\widehat{\theta}_{1n}^2 \left(2\widehat{\theta}_{1n}+1\right)^{-1}$, where a^* is a constant and w and m are determined by the equations

$$w = \frac{a^* 2\theta^2}{(2\theta+1)n^*} \tag{11}$$

and

$$m = 1 + \left[\frac{a^* 2\theta_L^2}{w\left(2\theta_L+1\right)}\right], \tag{12}$$

respectively, for given n^*, θ and θ_L.

Mukhopadhyay and Cicconetti (2004) provided extensive tabulations of the sequential estimators given by the MLE (1) and the UMVUE (2) for various combinations of n^*, θ and θ_L with $a^* = 1$. For $n^* = 25$ and $n^* = 35$, Mukhopadhyay and Cicconetti argued that the MLE (1) was the better estimator of the two. For the other values of n^* considered, Mukhopadhyay and Cicconetti were unable to compute the UMVUE (2). Here, we provide a more comprehensive investigation of the relative performance of the MLE (1) versus the UMVUE (2) by using the derived formula (9). We computed 2,000 replications of both the MLE (1) and the UMVUE (9) for each of the following combinations: $a^* = 1$, $n^* = 25, 50, 100, 200$, $\theta = 0.1, 0.2, \ldots, 1.5$ and $\theta_L = 0.01\theta, 0.02\theta, \ldots, 0.99\theta$. For each set of 2,000 replications, we computed the risks of estimation as

$$R_1 = a^* \widehat{E}\left[\left(\widehat{\theta}_{1n} - \theta\right)^2\right]$$

and

$$R_5 = a^* \widehat{E}\left[\left(\widehat{\theta}_{5n} - \theta\right)^2\right].$$

Figure 3 shows how the difference in the estimated risk, $R_1 - R_5$, varies with respect to θ_L for $n^* = 25, 50, 100, 200$ and $\theta = 0.1, 0.2, \ldots, 1.5$. The following function in R was used to compute $R_1 - R_5$ for given a^*, n^*, θ and θ_L.

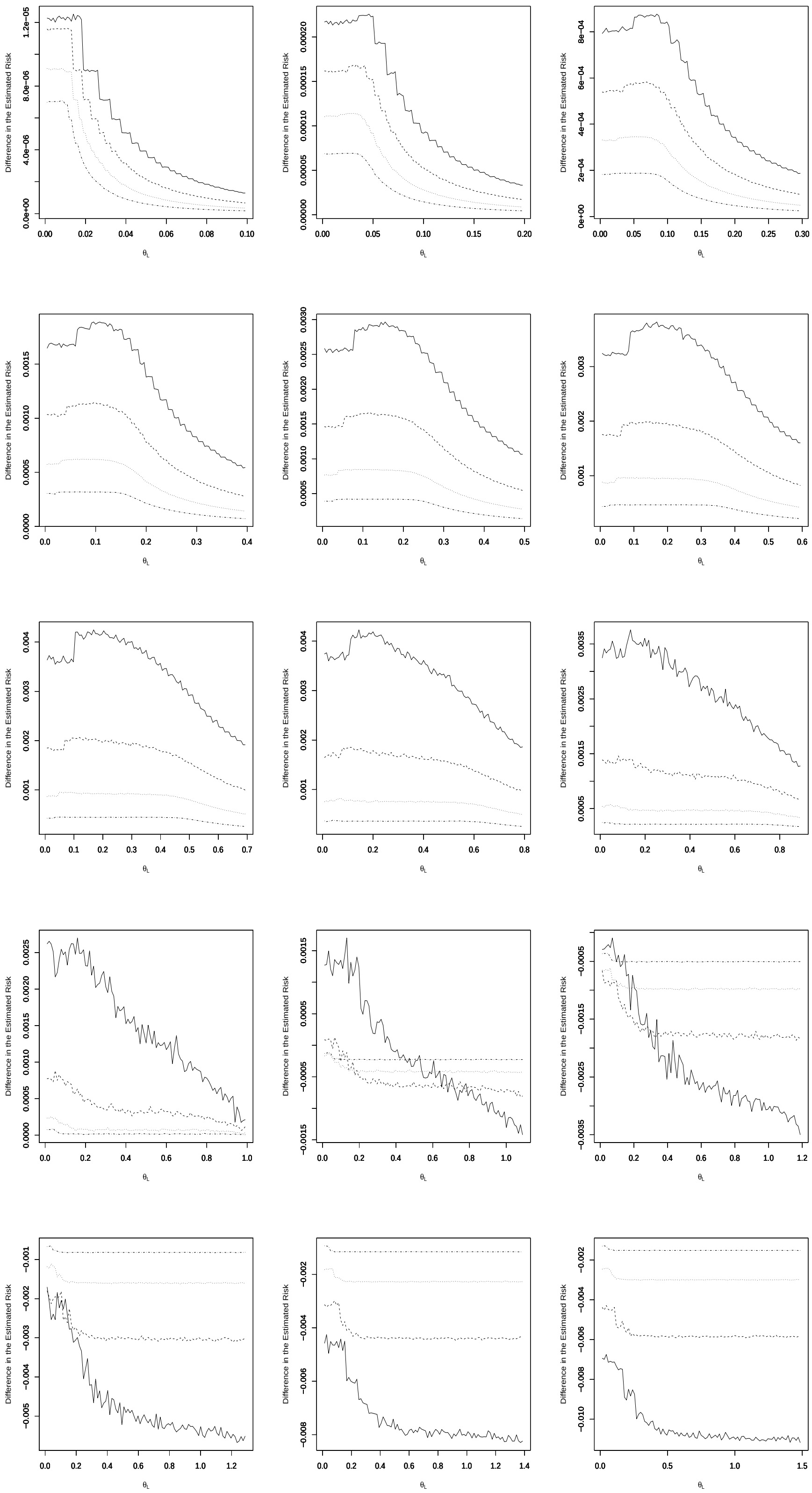

Figure 3 – The difference in the estimated risk, $R_1 - R_5$, versus θ_L for $\theta = 0.1, 0.2, \ldots, 1.5$, $n^* = 25$ (solid line), $n^* = 50$ (dashed line), $n^* = 100$ (dotted line) and $n^* = 200$ (dotdash line).

```
ff=function (astar,nstar,theta,thetaL)
{
   tt=astar*2*theta**2*(2*theta+1)**(-1)
   w=tt/nstar
   m=trunc(astar*(1/w)*2*thetaL**2*(2*thetaL+1)**(-1))+1
   for (i in 1:2000)
     {
       x=rnorm(m,mean=theta,sd=theta**2)
       tn=mean(x**2)
       theta1n=sqrt(tn+0.25)-0.5
       n=m
       repeat
          {
            tt=astar*(1/w)*2*theta1n**2*(2*theta1n+1)**(-1)
            if (n>=tt) break
            n=n+1
            x=c(x,rnorm(1,mean=theta,sd=theta**2))
            tn=mean(x**2)
            theta1n=sqrt(tn+0.25)-0.5
          }
         theta1[i]=theta1n
         u=sum(x**2)
         theta5[i]=sqrt(u/n)*besselI(sqrt(u*n),nu=n/2,expon.scaled=TRUE)
            /besselI(sqrt(u*n),nu=n/2-1,expon.scaled=TRUE)
     }
  ttt=astar*mean((theta1-theta)**2)-astar*mean((theta5-theta)**2)
  return(tt)
}
```

We can observe the following from Figure 3: the UMVUE is the better of the two estimators for $\theta \leq 1$; the UMVUE performs relatively better as θ increases to 1; the UMVUE performs relatively better as θ_L decreases; the UMVUE performs relatively better as n^* decreases; the MLE is the better of the two estimators for $\theta > 1$ and θ_L not too small; the MLE performs relatively better as $\theta > 1$ increases; the MLE performs relatively better as θ_L increases; the MLE performs relatively better as n^* decreases.

The estimates given in the tables in Mukhopadhyay and Cicconetti (2004) for $n^* = 25, 35$ do coincide with the estimates we obtained up to the decimal places given. The codes given in this paper can be used to compute the estimates to any decimal place.

4. Conclusions

Given a normal random variable with both mean and standard deviation equal to θ, we have derived a simple expression for the UMVUE of θ. This is the simplest known to date, hence it can be applied efficiently to sequential problems involving the normal random variable.

We have compared the performances of the UMVUE and MLE in terms of risk. Some of the main findings are that the UMVUE is the better for $\theta \leq 1$, the UMVUE performs relatively better as θ increases to 1, the MLE is the better for $\theta > 1$ and the MLE performs relatively better as $\theta > 1$ increases.

A future work is to see if some expressions can be derived for sequential estimation involving other distributions like the gamma, beta, Gumbel, lognormal, Weibull and inverse Gaussian distributions.

Acknowledgements

The authors would like to thank the Editor and the referee for careful reading and comments which improved the paper.

References

H. Ammeter (1962). *Experience rating a new application of the collective theory of risk*. ASTIN Bulletin, 2, pp. 261–270.

L. Badea, D. Tilivea (2008). *Nonnegative decompositions with resampling for improving gene expression data biclustering stability*. Proceedings of the 18th European Conference on Artificial Intelligence.

K. Benz, T. M. Bohnert (2014). *Impact of pacemaker failover configuration on mean time to recovery for small cloud clusters*. Proceedings of the 7th IEEE International Conference on Cloud Computing, pp. 384–391.

S. Bhamidi, R. van der Hofstad, G. Hooghiemstra (2010). *First passage percolation on random graphs with finite mean degrees*. Annals of Applied Probability, 20, pp. 1907–1965.

D. Bhattacharjee, N. Mukhopadhyay (2011). *On mp test and the mvues in a $n(\theta, c\theta)$ distribution with θ unknown: Illustrations and applications*. Journal of the Japan Statistical Society, 41, pp. 75–91.

I. Caridi (2014). *Properties of interaction networks underlying the minority game*. Physical Review E, 90.

K. Choi (2005). *Some properties of sequential point estimation of the mean*. Journal of Korean Data and Information Science Society, 16, pp. 657–663.

V. Dongen (1999). *Unbiased estimation of individual asymmetry*. Journal of Evolutionary Biology, 13, pp. 107–112.

S. U. K. E. S. Chung (2013). *Bayesian rainfall frequency analysis with extreme value using the informative prior distribution*. KSCE Journal of Civil Engineering, 17, pp. 1502–1514.

P. Fenn, N. Rickman, A. McGuire (1994). *Contracts and supply assurance in the uk health care market*. Journal of Health Economics, 13, pp. 125–144.

E. Forti, C. Franzoni, M. Sobrero (2007). *The effect of patenting on the networks and connections of academic scientists*. Working Paper.

G. D. Fraser, A. D. C. Chan, J. R. Green, D. MacIsaac (2012). *Detection of adc clipping, quantization noise, and amplifier saturation in surface electromyography*. Proceedings of the IEEE International Symposium on Medical Measurements and Applications, pp. 1–5.

M. Kaliszan (2011). *Does a draft really influence postmortem body cooling?* Journal of Forensic Sciences, 56, pp. 1310–1314.

S. Khan, R. Greiner (2013). *Finding discriminatory genes: A methodology for validating microarray studies*. Proceedings of the 13th IEEE International Conference on Data Mining Workshops, pp. 64–71.

S. K. Kim, S. L. Kim, Y. W. Lee (2007). *Sequential confidence intervals with β-protection in a normal distribution having equal mean and variance.* Journal of Applied Mathematics and Computing, 23, pp. 479–488.

N. Mukhopadhyay (2006). *Mvue for the mean with one observation.* The American Statistician, 60, pp. 71–74.

N. Mukhopadhyay, D. Bhattacharjee (2010). *A note on minimum variance unbiased estimation.* Communications in Statistics—Theory and Methods, 39, pp. 1466–1476.

N. Mukhopadhyay, G. Cicconetti (2004). *Applications of sequentially estimating the mean in a normal distribution having equal mean and variance.* Sequential Analysis, 23, pp. 625–665.

B. M. d. S. N. Mukhopadhyay (2008). *Theory and applications of a new methodology for the random sequential probability ratio test.* Statistical Methodology, 5, pp. 424–453.

A. P. Prudnikov, Y. A. Brychkov, O. I. Marichev (1986). *Integrals and Series, volume 1.* Gordon and Breach Science Publishers, Amsterdam.

R Development Core Team (2016). *R: A Language and Environment for Statistical Computing.* Vienna, Austria.

N. M. S. Banerjee (2016). *A general sequential fixed-accuracy confidence interval estimation methodology for a positive parameter: Illustrations using health and safety data.* Annals of the Institute of Statistical Mathematics, 68, pp. 541–570.

Summary

Mukhopadhyay and Cicconetti (2004) derived the Maximum Likelihood Estimator (MLE) and the Uniformly Minimum Variance Unbiased Estimator (UMVUE) of θ in $N(\theta, \theta)$ and discussed their application to purely sequential and two-stage bounded risk estimation of θ. In this paper, a much simpler expression is derived for the UMVUE of θ. Using this expression, a comprehensive investigation is provided for comparing the performances of the sequential estimators based on the MLE and the UMVUE.

Keywords: MLE; Normal distribution; Sequential estimation; UMVUE