بعض مقدرات الاختبار الأولي المقلصة لمتوسط التوزيع الطبيعي

م.م. بيداء عطية خلف*

المستخلص :

اقترح هذا البحث بعض مقدرات الاختبار الأولي المقلصة ذات المرحلة الواحدة والمرحلتين لمتوسط التوزيع الطبيعي عندما يكون التباين معلوماً وعند توافر معلومات مسبقة حول المتوسط الحسابي بشكل قيم أولية (μ_0) ومن خلال استخدام دوال تقلص موزونة $\psi(\cdot)$ بالاضافة الى استخدام مجال الاختبار الاولي R.

ولدراسة سلوك المقدرات المقترحة تم اشتقاق بعض المؤشرات الاحصائية حيث تم اشتقاق معادلة التحيز $[B(\cdot)]$، متوسط مربعات الخطأ $[MSE(\cdot)]$، الكفاءة $[Eff(\cdot)]$، حجم العينة المتوقع $[E(n \mid \sigma^2)]$ لهذه المقدرات. ثم بعد ذلك أعطيت بعض النتائج العددية والاستنتاجات الخاصة بالمؤشرات المذكورة اعلاه والخاصة بالمقدرات المقترحة من خلال اختيار بعض القيم للثوابت المتظمنة فيها. وأخيراً تمت بعض المقارنات للمقدرات المقترحة مع المقدرات الكلاسيكية وبعض البحوث المنجزة حديثاً لبيان فائدتها وافضليتها.

Abstract

This paper concerned with preliminary single and double stage shrunken estimators for estimate the mean (μ) of normal distribution when a prior estimate (μ_0) of the actual (μ) is a available using specifying shrinkage weight factors $\psi(\cdot)$ as well as pre- test region (R).

Expression for the Bias $[B(\cdot)]$, Mean Squared error $[MSE(\cdot)]$, Efficiency $[EFF(\cdot)]$ and Expected sample size $[E(n \mid \sigma^2)]$ for proposed estimators are derived. Numerical results and conclusions are drawn

---

* مدرس مساعد / جامعة بغداد/ كلية التربية – ابن الهيثم /قسم الرياضيات

مقبول للنشر بتاريخ 2010/3/1

about selection different (constant) including in these expressions. Comparisons between suggested estimators with classical and some existing estimators are given in the sense of Bias and relative efficiency.

1. Introduction

Assume that $x_1,x_2,\ldots,x_n$ be a random sample of size (n) from a normal population with unknown mean (μ) and known variance (σ^2). In conventional notation, we write $x \sim N(\mu,\sigma^2)$.

In this work we suggest the problem of estimating the mean (μ) when some prior information (μ_0) regarding the mean (μ) is available. More specifically we assume that the prior information regarding due the following reasons, (1)

1. **we belive that (μ_0) is closer to the true value of μ, or**
2. **we fear that (μ_0) may near the true value of (μ) i.e., something bad happens if (μ) is approximately equal to (μ_0) and we do not know about it in such a situation it is natural to start with the MLE($\hat{\mu}$) of μ and modify it by moving it closer to (μ_0) using shrinkage weight factor [$\psi(\cdot)$], so that the resulting estimator though perhaps biased, has a smaller mean square error [MSE] than that of ($\hat{\mu}$) in some interval around (μ_0). Preliminary test single stage shrinkage estimator is $\tilde{\mu}$ estimator of level of significance (α) for test the hypothesis $H_0:\mu=\mu_0$ against the hypothesis $H_A:\mu\neq \mu_0$, if H_0 accepted we feel comfortable in using the prior information (μ_0) with ($\hat{\mu}$) in estimating μ using shrinkage weight factor $\psi_1(\hat{\mu})$, $0 \leq \psi_1(\hat{\mu}) \leq 1$.**

i.e. $\tilde{\mu} = \psi_1(\hat{\mu})\hat{\mu} + [1-\psi_1(\hat{\mu})]\mu_0$ **…[1]**

or, $\tilde{\mu} = \psi_1(\hat{\mu})[\hat{\mu}-\mu_0] + \mu_0$

However, if H_0 is rejected, we assume shrinkage estimator using another shrinkage weight factor $\psi_2(\hat{\mu})$, $0 \leq \psi_2(\hat{\mu}) \leq 1$.

i.e. $\tilde{\mu} = \psi_2(\hat{\mu})\hat{\mu} + [1-\psi_2(\hat{\mu})]\mu_0$ **…[2]**

or, $\tilde{\mu} = \psi_2(\hat{\mu})[\hat{\mu}-\mu_0] + \mu_0$

Therefore, the general preliminary test single stage shrunken estimator (PTSSSE) for the mean will be:

$$\tilde{\mu}_{SS} = \begin{cases} \psi_1(\hat{\mu})\hat{\mu} + [1-\psi_1(\hat{\mu})]\mu_0, & \text{if } \hat{\mu} \in \mathrm{R}, \\ \psi_2(\hat{\mu})\hat{\mu} + [1-\psi_2(\hat{\mu})]\mu_0, & \text{if } \hat{\mu} \notin \mathrm{R}, \end{cases} \qquad \ldots[3]$$

where $\psi_i(\hat{\mu})$, $0 \le \psi_i(\hat{\mu}) \le 1$ [i=1,2] is shrinkage weighted factor specifying the belief in μ_0 which can be a function of $\hat{\mu}$ or a constant.

Preliminary test double stage shrunken estimator for the mean that utilize a prior estimate (μ_0) represents as following steps:

1. **Select two positive integers (n_1) and (n_2).**
2. **Obtain a random sample of size (n_1) on x [first stage sample]. Compute sample mean $\hat{\mu}_1$ by M.L.E. technique.**
3. **Choose a suitable region (R_1) around μ_0. In this work we suggest pre-test region.**
4. **If $\hat{\mu}_1 \in R_1$, suggest the shrinkage estimator which defined in [1]. However, if $\hat{\mu}_1 \notin R_1$, obtain a second stage random sample of size n_2 on x and consider the estimator of μ as a polling of two samples mean ($\hat{\mu}_p$)**

i.e. $\hat{\mu}_p = (n_1\hat{\mu}_1 + n_2\hat{\mu}_2)/n$ **...[4]**

where $\hat{\mu}_2$ is the second sample mean and $n = n_1 + n_2$. Thus, the general preliminary double stage shrunken estimator (PTDSSE) has the following form

$$\tilde{\mu}_{DS} = \begin{cases} \psi_1(\cdot)(\hat{\mu}) + [1-\psi_1(\cdot)]\mu_0 & \text{if } \hat{\mu}_1 \in R \\ \hat{\mu}_p & \text{if } \hat{\mu} \notin R \end{cases} \quad ...[5]$$

The aim of this paper is to study the (PTSSSE) and (PTDSSE) for estimate the mean of normal distribution using new shrinkage weight factors and to get higher efficiency than the classical estimator and exists studies.

Where, R is the pre-test region for testing the hypothesis H_0: $\hat{\mu} \neq \mu_0$ against H_A:$\mu \neq \mu_0$ with level of significance (α) using test statistics T($\hat{\mu}$ /σ) $= \dfrac{\sqrt{n}(\hat{\mu}-\mu_0)}{\sigma}$.

As well as R_1 is the pre-test region for testing the hypothesis H_0: $\hat{\mu}_1 \neq \mu_0$ against H_A: $\hat{\mu}_1 \neq \mu_0$ with level of significance (α) using test statistics

T($\hat{\mu}_1$ /σ) = $\dfrac{\sqrt{n_1}(\hat{\mu}_1-\mu_0)}{\sigma}$.

i.e.;

$$\left.\begin{array}{l} R=[\mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}] \\ \text{and} \\ R_1=[\mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n_1}}, \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n_1}}] \end{array}\right\} \quad ...[6]$$

Several authors have been studied (PTSSSE) and (PTDSSE) for various parameters of different distributions and to estimate the parameters of regression models, also for complete and Censoed samples, for example Thompson (1), Katti (2), Mehta and Srinivasan (3), Pandy (4), Waikar, Schuurmann and Raghunathan (5), Al-Hemyari (6), Al-Bayyati and Arnold (7), Al-Hemyari and Al-Juboori (8), Al-Juboori (9), Al-Juboori (10), Al-Bermani (11) and Al-Rabassi (12).

2. Preliminary Test Single-Stage Shrunken Estimator (PTSSSE)

In this section, we consider the (PTSSSE) defined in (3) with the following form

$$\tilde{\mu}_{SS} = \begin{cases} \mu_0 & ,\text{if } \hat{\mu} \in R \\ w_p(\hat{\mu}-\mu_0)+\mu_0 & ,\text{if } \hat{\mu} \notin R \end{cases}$$

Where R is pre-test region and $\psi_1(\hat{\mu}) = 0$, $\psi_2(\hat{\mu}) = w_p = k_1(p) \mid k_2(p)$

such that $k_i(p) = \left(\frac{2}{n-2}\right)^{ip}\left[\Gamma\left(\frac{2+2ip-2}{2}\right)\Big|\Gamma\left(\frac{n-2}{2}\right)\right]$, **i=1,2, n > 2**

and p ∈ N.

The expression for Bias [B(·)] and mean squared error [MSE(·)] of $\tilde{\mu}_{SS}$ **represented respectively as follows:**

$$\text{Bias}(\tilde{\mu}_{SS} \mid \mu, R) = E(\tilde{\mu}_{SS}) - \mu$$

$$= \int_R [-(\mu-\mu_0)] f(\hat{\mu} \mid \mu) d\hat{\mu} + \int_{\bar{R}} [w_p(\hat{\mu}-\mu_0) + (\mu_0-\mu)]\, f(\hat{\mu} \mid \mu) d\hat{\mu}$$

where $\bar{R}$ **is the complement region of R in real space,** $\hat{\mu} = \bar{x} \sim N(\mu, \frac{\sigma^2}{n})$

and

$$f(\hat{\mu}\mid\mu,\sigma^2)=\frac{\sqrt{n}}{\sigma\sqrt{2\pi}}\exp[-n(\hat{\mu}-\mu)^2/2\sigma^2],\ -\infty<\hat{\mu}<\infty,-\infty<\mu<\infty,\sigma^2\geq 0$$

The previous expression will result.

$$B(\tilde{\mu}_{SS}\mid\mu,R)=\frac{\sigma}{\sqrt{n}}[-\lambda j_0(a,b)+w_p\lambda-\lambda-w_p(j_1(a,b)+\lambda j_0(a,b))+\lambda j_0(a,b)] \quad \text{...[7]}$$

where $j_l(a,b)=\int_a^b \frac{1}{\sqrt{2\pi}} z^l \exp(-z^2/2)dz,\ l=0,1,2,\ldots$ …**[8]**

$$z=\frac{\hat{\mu}-\mu}{\sigma/\sqrt{n}},\ \lambda=\frac{\mu-\mu_0}{\sigma/\sqrt{n}}, \quad \text{...[9]}$$

$a=-\lambda-z_{\alpha/2}$ **,** $b=-\lambda+z_{\alpha/2}$**,** …**[10]**

and

$$MSE(\tilde{\mu}_{SS}\mid\mu,R)=E(\tilde{\mu}_{SS}-\mu)^2$$

$$=\int_R(\mu-\mu_0)^2 f(\hat{\mu}\mid\mu)d\hat{\mu}+\int_{\bar{R}}[w_p(\hat{\mu}-\mu_0)+(\mu_0-\mu)]^2 f(\hat{\mu}\mid\mu)d\hat{\mu}$$

We can conclude,

$$MSE(\tilde{\mu}_{SS}\mid\mu,R)=\frac{\sigma^2}{n}[2w_p\lambda[j_1(a,b)+\lambda j_0(a,b)]-w_p^2[j_2(a,b)+2\lambda j_1(a,b)+j_0(a,b)\lambda^2]+ w_p^2(1+\lambda^2)-\lambda^2(2w_p-1)] \quad \text{...[11]}$$

The Efficiency of $\tilde{\mu}_{SS}$ **relative to** $\hat{\mu}$ **is given by**

$$R.Eff(\tilde{\mu}_{SS}\mid\mu,R)=\frac{MSE(\hat{\mu}\mid\mu)}{MSE(\tilde{\mu}_{SS}\mid\mu,R)} \quad \text{...[12]}$$

3. Preliminary Test double-Stage Shrunken Estimator (PTDSSE)

In this section, we consider the (PTDSSE) defined in (5) for the mean of normal distribution as follows

$$\tilde{\mu}_{DS}(\mu\mid\hat{\mu},R_1)=\begin{cases} w_p(\hat{\mu}_1-\mu_0)+\mu_0 & ,\text{if } \hat{\mu}_1\in R \\ \hat{\mu}_p=\dfrac{\hat{\mu}_1 n_1+\hat{\mu}_2 n_2}{n} & ,\text{if } \hat{\mu}_1\notin R \end{cases}$$

where $\psi_1(\hat{u})=w_p$**, which defined in previous section and** $n=n_1+n_2$**.**

The expression for Bias [B(·)] and mean squared error [MSE(·)] of $\tilde{\mu}_{DS}$ are respectively given as follows

$$\text{Bias}(\tilde{\mu}_{DS} \mid \mu, R_1) = E(\tilde{\mu}_{DS} - \mu)$$

$$= \int_{\hat{\mu}_1 \in R_1} \int_{\hat{\mu}_2=-\infty}^{\infty} [w_p(\hat{\mu}_1 - \mu_0) + (\mu_0 - \mu)] f(\hat{\mu}_1 \mid \mu) f(\hat{\mu}_2 \mid \mu) d\hat{\mu}_2 d\hat{\mu}_1 +$$

$$\int_{\hat{\mu}_1 \in \overline{R_1}} \int_{\hat{\mu}_2=-\infty}^{\infty} [\frac{\hat{\mu}_1 n_1 + \hat{\mu}_2 n_2}{n} - \mu] f(\hat{\mu}_1 \mid \mu) f(\hat{\mu}_2 \mid \mu) d\hat{\mu}_2 d\hat{\mu}_1$$

and by simple calculations, we get

$$B(\tilde{\mu}_{DS} \mid \mu, R_1) = \frac{\sigma}{\sqrt{n_1}} w_p j_1(a_1, b_1) + \lambda j_0(a_1, b_1) - \lambda j_0(a_1, b_1) - \frac{1}{1+r} j_1(a_1, b_1)$$

where

$$r = \frac{n_2}{n_1},\ z_1 = \frac{\hat{\mu}_1 - \mu}{\sigma / \sqrt{n}},$$

$$j_l(a_1, b_1) = \frac{1}{\sqrt{2\pi}} \int_{a_1}^{b_1} z_1^l \exp(-z_1^2/2) dz_1, l = 0, 1, 2, K$$

$a_1 = -\lambda - z_{\alpha/2}$, $b_1 = -\lambda + z_{\alpha/2}$,

and,

$$\text{MSE}(\tilde{\mu}_{DS} \mid \mu, R_1) = E(\tilde{\mu}_{DS} - \mu)^2$$

$$= \int_{\hat{\mu}_1 \in R_1} \int_{\hat{\mu}_2=-\infty}^{\infty} [w_p(\hat{\mu}_1 - \mu_0) + (\mu_0 - \mu)]^2 f(\hat{\mu}_1 \mid \mu) f(\hat{\mu}_2 \mid \mu) d\hat{\mu}_1 d\hat{\mu}_2 +$$

$$\int_{\hat{\mu}_1 \in \overline{R}} \int_{\hat{\mu}_2=-\infty}^{\infty} [\frac{\hat{\mu}_1 n_1 + \hat{\mu}_2 n_2}{n} - \mu)]^2 f(\hat{\mu}_1 \mid \mu) f(\hat{\mu}_2 \mid \mu) d\hat{\mu}_1 d\hat{\mu}_2$$

We can conclude,

$$MSE(\tilde{\mu}_{DS} \mid \mu, R_1) = \frac{\sigma^2}{n}\left[w_p^2\, j_2(a_1,b_1) + 2\lambda j_1(a_1,b_1) + \lambda^2 j_0(a_1,b_1)\right] - 2w_p\left[\lambda j_1(a_1,b_1) + \lambda^2 j_0(a_1,b_1)\right] + \lambda^2\, j_0(a_1,b_1)] + \frac{1}{(1-r)^2} + \frac{r}{(1+r)^2} - \frac{1}{(1-r)^2}\, j_2(a_1,b_1) - \left(\frac{r}{1+r}\right)^2 j_0(a_1,b_1) \qquad K[13]$$

The expected sample size can be obtained as below

$E(n/\mu,R_1) = n_1\,[1 + r\,(1 - j_0(a_1,b_1)]$ **...[14]**

The Efficiency of $\tilde{\mu}_{DS}$ relative to $\hat{\mu}(\bar{x})$ is given by

$$R.Eff\,(\tilde{\mu}_{DS} \mid \mu, R) = \frac{MSE(\hat{\mu} \mid \mu)}{[MSE(\tilde{\mu}_{DS} \mid \mu, R_1) * E(n/\mu, R_1)]} \qquad ...[15]$$

And the probability of a voiding the second sample given by $p(\hat{\mu} \in R)$.

The percentage of the over all saved (P.O.S.S.) of $\tilde{\mu}_{DS}$ can be compute by:

$$P.O.S.S. = \frac{n_2}{n}\, j_0(a_1,b_1) * 100$$

4. Conclusion and Numerical Results

4.1 From the Expressions for Bias and MSE of $\tilde{\mu}_{SS}$, the following could be easily seen:

1. i. $B(\tilde{\mu}_{SS}/\mu,R)$ is an odd function of $\lambda, (\lambda = \sqrt{n}(\mu - \mu_0)/\sigma)$.

ii. MSE $((\tilde{\mu}_{SS} \mid \mu, R)$ **is an even function of λ.**

iii. The estimator $\tilde{\mu}_{SS}$ is consistent, i.e. $\lim_{n\to\infty} MSE(\tilde{\mu}_{SS} \mid \mu, R) = 0$.

iv. The estimator $\tilde{\mu}_{SS}$ dominates $(\hat{\mu})$ with large sample size (n) in the term of MSE, i.e. $\lim_{n\to\infty}[MSE(\tilde{\mu}_{SS} \mid \mu, R) - MSE(\hat{\mu})] \le 0$.

v. As we expected the estimator $\tilde{\mu}_{SS}$ is unbiased when $\mu = \mu_0$,

i.e. $\lim_{\lambda\to 0} B(\tilde{\mu} \mid \mu, R) = 0$

2. The computations of relative Efficiency [R.Eff($\tilde{\mu}_{SS}$)] and Bias ratio $B(\hat{\mu}_{SS}) = \sqrt{n}\, B(\hat{\mu}_{SS})/\sigma]$ **of consider estimators $\tilde{\mu}_{SS}$ were made for different values involved in it, some of these computations are given in**

table (1) for some samples of these constant e.g. α = 0.01,0.05, 0.1, n = 4,8,12, p = 2and λ = 0.0(0.1)1,2. The following numerical results from the mentioned table were mad:

i. R.Eff($\tilde{\mu}_{SS}$) is maximum when $\mu \approx \mu_0$ and decreases with increasing value of λ.

ii. R.Eff($\tilde{\mu}_{SS}$) has maximum value with small value of α.

iii. B($\tilde{\mu}_{SS}$) are reasonably small when $\mu \approx \mu_0$ ($\lambda \approx 0$).

iv. B($\tilde{\mu}_{SS}$) is increases when α increases.

v. The Effective interval [the values of λ that make R.Eff. greater than one] using proposed estimator $\tilde{\mu}_{SS}$ in [-1,1]

3. The consider estimator $\tilde{\mu}_{SS}$ is better than classical estimator (MLE) and also than the estimators suggested by [1], [3], [8] and [9] in the terms of higher relative Efficiancy especially at $\mu \approx \mu_0$.

4.2 From the Expressions for Bias and MSE of $\tilde{\mu}_{BS}$, we can see the following:

1. i. $\tilde{\mu}_{BS}$ is consistent of μ.

ii. $\tilde{\mu}_{BS}$ dominates $\hat{\mu}$ with large sample size (n).

iii. $\tilde{\mu}_{BS}$ is unbiased estimator when $\mu \approx \mu_0$.

2. The computations of relative Efficiency [R.Eff(·)], Bias ratio [B(·)], Expected sample size [E(n/μ,R)], expected sample size proportion, percentage of the overall sample saved and probability of avoiding the second sample were used for the estimator $\tilde{\mu}_{BS}$.

These computations were performed for n_1 = 12, p = 4,8,12, r = 0.5,1,2,4,8,12, λ = 0.0(0.1),1,2, α = 0.01,0.05,0.1.

Some of these computations are given in tables (2), …, (12).

The observations mention the tables above leads to the following results:

i. R.Eff($\tilde{\mu}_{BS}$) are adversely proportional with small value of α.

ii. R.Eff($\tilde{\mu}_{BS}$) are maximum when $\mu \approx \mu_0$ and decreasing with increasing value of λ.

iii. B($\tilde{\mu}_{BS}$) are reasonably small when $\mu \approx \mu_0$, otherwise B($\tilde{\mu}$) will be maximum.

iv. B($\tilde{\mu}_{BS}$) are reasonably small with small value of α.

v. R.Eff.($\tilde{\mu}_{BS}$) and B($\tilde{\mu}_{BS}$) are decreasing function w.r.t. first sample size (n_1).

vi. E(n | μ,R) of $\tilde{\mu}_{BS}$ is closer to (n_1) specially when $\mu \approx \mu_0$ and start faraway slowly with (λ) increases.

vii. Percentage of the over all sample saved $[\frac{n_2}{n} j_0(a_1,b_1) * 100]$ is decreasing function of λ and has maximum value when $\mu = \mu_0$.

viii. R.Eff.($\tilde{\mu}_{BS}$) is an increasing function w.r.t. r ($r = n_2/n_1$) and p.

3. Suggested estimator $\tilde{\mu}_{BS}$ is more efficient than (MLE) and than estimators introduced by (5), (6) and (12).

References:

1. **Thompson,J.R., (1968), "Some Shrinkage Techniques for Estimating the Mean", J. Amer. Statist. Assoc, 63, p.113-122.**
2. **Katti,S.K., (1962), "Use of Some a Prior Knowledge in the Estimation of Means from Double Samples", Biometrics, 18, p. 139-147.**
3. **Mehta,J.S. and Srinivasan,R., (1971), "Estimation of the Mean by Shrinkage to a Point", J. Amer. Statist. Assoc., 66, p.86-90.**
4. **Pandey,B.N., (1979), "On Shrinkage Estimation of Normal Population Variance", Commum. Statist-Theor. Meth., A8(4), p.359-365.**
5. **Waikar,V.B., Schuurmann,F.J. and Raghunathan,T.E., (1984), "On a Two-Stage Shrinkage Testimator of the Mean of a Normal Distribution", Commum. Statist-Theor. Meth., A13(15), p.1901-1913.**
6. **Al-Hemyari,Z.A., (1990), "A Two-stage, Estimator of the Mean of Normal Distribution", Proc. Int. Conf. statistic, comp. Cairo, 3, 2, p.139-153.**
7. **Al-Bayyati,H.A, and Arnold, J.C., (1972), "On Double Stage Estimation in Sample Linear Regression Using Prior Knowledge", Technometrics, 14, p.405-414.**
8. **Al-Hemyari, Z.A. and AL-Jubori A.N., (1999), "Modifical Single Stage Estimators of the Mean of Normal Population", Al-Fath J. of the College of Education for Pure sci. and Humanities, 3(5), p.45-57.**
9. **Al-Jubori, A.N., (2008), "An Efficient Shrunken Estimators for the Mean of Normal Population With Known Variance", Ibn Al-Haitham J. for Pure and Applied Sciences, 21(3), p.159-165.**

10. Al-Juboori,A.N., (2001), "On Shrinkage Techniques for Estimating the Variance of Normal Distribution Using Stein-Type Estimator", Ibn-Al-Haitham J. for Pure and Applied Sciences, 14 (4B), p.59-66.

11. AL-Bermani,M.H.,(2008), "Comparison Between Bayesian Shrinkage Estimators and Shrinkage Estimators for the Variance of Normal Distribution by Using Simulation", Ph.D. Thesis, College of Administration and Economics, University of Baghdad.

12. AL-Rabassi,A.M.,(2000), "Some Single and Double Stage Shrunken Estimators for the mean of Normal Distribution-Comparison Study", Ph.D. Thesis, College of Administration and Economics, University of Baghdad.