On Shrunken Bayesian Estimators for The Mean of Normal Distribution

Ahmed. M. Al –Robassi
Faculty of Administration Science, Taiz University, Yemen
Email:alrobasi5@yhoo.com

Abstract

This paper studies the double-stage shrunken Bayesian estimators (DSSBE) for the mean suggested. In this estimator a shrinkage factor k is taken and the region R was found by minimizing mean squared error. The numerical result shows improvement of the double-stage shrunken Bayesian estimators over the double stage Bayesian estimators in some situations

Introduction

Let x_{ji}, $j = 1,2$, $i = 1,2,\ldots,n_j$, denote two random samples independently normally distributed population with mean θ and variance σ^2.
Arnold, and Al-Bayyati [4] considered a double stage shrunken estimator of the mean θ when an a prior information about θ is available in the form of an initial estimate θ_0.
Their estimator is given by:

$$\tilde{\theta} = \begin{cases} k(\hat{\theta}_1 - \theta_0) + \theta_0 & if\ \hat{\theta}_1 \in R \\ \dfrac{n_1\hat{\theta}_1 + n_2\hat{\theta}_2}{n_1 + n_2} & if\ \hat{\theta}_1 \notin R \end{cases} \qquad \ldots(1)$$

Where $0 \le k \le 1$, is the shrinkage factor, R is the parameter, and $\hat{\theta}_1$ is MLE (maximum likelihood estimator) of θ based on n_1.

Several authors studied the estimator $\tilde{\theta}$ (see, e.g. Whiker, Shurmann, and Raghunath [5], Al-Robassi [1]).

Bayesian method assumes as before that the random sample, $x_1, x_2, \ldots x_n$ can from a population with probability density function $f(x,\theta)$,but further more the unknown parameter θ is a random variable that is there is additional information about θ.

In this paper we study a double stage shrunken Bayesian estimator (DSSBE) of the mean θ of normal distribution when the variance σ^2 is known.

Now suppose that $\hat{\theta}_j = \overline{x}_j, j = 1,2$, the sufficient statistic, is the mean of random sample of size n_j; $g(\hat{\theta}_j / \theta)$ is $N(\theta, \frac{\sigma^2}{n_j})$. Further more suppose that the prior distribution of θ is defined by:

$$h(\theta) = \frac{1}{\sqrt{2\pi}} \exp(\frac{\theta^2}{2}) \qquad \ldots(2)$$

If the square-error loss function $L(\hat{\theta}_j, \theta) = (\hat{\theta}_j - \theta)^2$ is used [3] , then the Bayes estimator of θ is given by:

$$\hat{\theta}_{Bj} = \frac{\hat{\theta}_j}{1 + \sigma^2 / n_j} \qquad \ldots(3)$$

Therefore a double stage shrunken Bayesian estimator is defined as follows:

$$\widetilde{\theta} = \begin{cases} k(\hat{\theta}_{B1} - \theta_0) + \theta_0 & if\ \hat{\theta}_{B1} \in R \\ \dfrac{n_1 \hat{\theta}_{B1} + n_2 \hat{\theta}_{B2}}{n_1 + n_2} & if\ \hat{\theta}_{B1} \notin R \end{cases} \qquad \ldots(4)$$

where, $0 \le k \le 1$, R is the suitable region in the parameter space, and $\hat{\theta}_{B1}$ is the Bayes estimator of θ based on first sample of size n_1.

Mean squared error, Expected sample size, and Efficiency.

In this section the DSSBE of the form (4) is considered, mean squared error, expected sample size and relative efficiency of the estimator $\widetilde{\theta}_B$ are derived as follows.

$$MSE(\widetilde{\theta}_B / \theta, R) = E(\widetilde{\theta}_B - \theta)^2$$

$$= MSE(\hat{\theta}_B) + (k^2 - c_1^2) g_2(\hat{\theta}_{B1} / \theta, R) + [2k(1-k)(\theta_0 - \theta)$$

$$- \frac{2n_1 n_2}{n^2} Bias(\hat{\theta}_{B2})] g_1(\hat{\theta}_{B1} / \theta, R) + [(1-k)^2 (\theta_0 - \theta)^2$$

$$- c_2^2 MSE(\hat{\theta}_{B2})] g_0(\hat{\theta}_{B1} / \theta, R) \qquad \ldots(5)$$

where $\hat{\theta}_B = \frac{n_1 \hat{\theta}_{B1} + n_2 \hat{\theta}_{B2}}{n_1 + n_2}$, $c_1^2 = \frac{n_1^2}{n^2}$, $c_2^2 = \frac{n_2^2}{n^2}$, $n = n_1 + n_2$,

and $g_i(\hat{\theta}_{B1} / \theta, R) = \int_R (\hat{\theta}_{B1} - \theta)^i f(\hat{\theta}_{B1}) d(\hat{\theta}_{B1}), i = 1,2$

If θ_0 is the true value of θ then equation (5) becomes:

$$MSE(\tilde{\theta}_B / \theta_0, R) = E(\tilde{\theta}_B - \theta_0)^2$$
$$= MSE(\hat{\theta}_B) + (k^2 - c_1^2) g_2(\hat{\theta}_{B1} / \theta_0, R)$$
$$- \frac{2n_1 n_2}{n^2} Bias(\hat{\theta}_{B2}) g_1(\hat{\theta}_{B1} / \theta_0, R)$$

$$- c_2^2 MSE(\hat{\theta}_{B2}) g_0(\hat{\theta}_{B1} / \theta_0, R) \qquad \ldots(6)$$

Let us choose the region R so that $MSE(\theta_B / \theta_0, R)$ is minimum (see [4]) we get :

$$R = \theta_0 + \frac{n_1 n_2 Bias(\hat{\theta}_{B1}) - a}{n^2 (k^2 - c_1^2)}, \theta_0 + \frac{n_1 n_2 Bias(\hat{\theta}_{B1}) + a}{n^2 (k^2 - c_1^2)} \qquad (7)$$

where $a = \sqrt{c_1^2 c_2^2 [Bias(\hat{\theta}_{B2)})]^2 + (k^2 - c_1^2) MSE(\hat{\theta}_{B2})}$.

The expected sample size is given by:

$$E(n / \theta, R) = n - (n - n_1) P_r(\hat{\theta}_{B1} \in R)$$

$$= n - (n - n_1) g_0(\hat{\theta}_{B1} / \theta, R) \qquad \ldots(8)$$

Therefore, we define the efficiency of the estimator $\tilde{\theta}_B$ with respect to the Bayesian estimator $\hat{\theta}_B$ by:

$$Eff(\tilde{\theta}_B / \theta, R) = \frac{MSE(\hat{\theta}_B, based\ on\ a\ sample\ of\ equivalent\ size)}{MSE(\tilde{\theta}_B / \theta, R)} \qquad \ldots(9)$$

where $\hat{\theta}_B = \dfrac{n_1 \hat{\theta}_{B1} + n_2 \hat{\theta}_{B2}}{n_1 + n_2}$.

Numerical results

The computation of mean squared error and efficiency of the estimator $\tilde{\theta}_B$ with respect to the Bayes estimator $\hat{\theta}_B$ considered in this section by taking

$$n = 50,\ n_1 = 5,10,15,20,\ k = 0.25,\ 0.5,\ 0.75\ and\ t = \sqrt{n_1} \frac{|\theta - \theta_0|}{\sigma}.$$

The following conclusions are based on these computations :

(1) The probability of avoiding the second sample , is decreasing function of the shrinkage factor, whereas it is increasing function of the first sample n_1 (see table (1)).

(2) Mean squared error of $\widetilde{\theta}_B$ is increasing function of the shrinkage factor, and decreasing function with the first sample n_1 (see table (2)).

(3) As expected, the efficiency of the double stage shrunken Bayesian estimator $\widetilde{\theta}_B$, is better than the efficiency of the Bayesian estimator $\hat{\theta}_B$,when θ is close to θ_0 .

(4) Table (3) indicates that the efficiency of the suggested estimator is decreasing function of the shrinkage factor k , and increasing function of n_1.

(5) All value of k and n_1 give highest efficiency only in neighborhood t=0 .

Table(1)

Probability 0f avoiding the second sample for, $n = 50,\ \theta = 10,\ \sigma = 5,\ t = \dfrac{|\theta - \theta_0|}{\sigma}\sqrt{n_1}$

	$n_1 = 5$			$n_1 = 10$			$n_1 = 15$			$n_1 = 20$		
k t	0.25	0.50	0.75	0.25	0.50	0.75	0.25	0.50	0.75	0.25	0.50	0.75
0.0	0.938	0.823	0.761	0.959	0.873	0.837	0.962	0.890	0.871	0.969	0.913	0.892
0.2	0.917	0.786	0.710	0.936	0.859	0.791	0.943	0.875	0.830	0.947	0.891	0.864
0.4	0.895	0.737	0.663	0.918	0.839	0.739	0.927	0.858	0.811	0.929	0.875	0.830
0.6	0.872	0.694	0.624	0.889	0.778	0.701	0.903	0.826	0.775	0.912	0.850	0.801
0.8	0859	0.657	0.562	0.867	0.733	0.671	0.889	0.783	0.729	0.896	0.821	0.771
1.0	0.847	0.615	0.519	0.849	0.692	0.604	0.868	0.749	0.683	0.871	0.785	0.729
1.2	0819	0.572	0.468	0.826	0.657	0.569	0.841	0.709	0.623	0.853	0.736	0.670
1.4	0.789	0.529	0.419	0.810	0.617	0.520	0.829	0.673	0.560	0.837	0.706	0.612
1.6	0.764	0.481	0.361	0.795	0.561	0.482	0.815	0.660	0.516	0.820	0.673	0.564
1.8	0.735	0.428	0.321	0.774	0.510	0.441	0.790	0.569	0.469	0.805	0.626	0.521
2.0	0.697	0.357	0.211	0.730	0.483	0.387	0.779	0.517	0.413	0.789	0.580	0.460

Table(2)

MSE($\widetilde{\theta}_B / \theta, R$) for, $\theta = 10,\ \sigma = 5,\ and\ t = \dfrac{|\theta - \theta_0|}{\sigma}\sqrt{n_1}$

$n_1 = 5$ $\quad$ $n_1 = 10$ $\quad$ $n_1 = 15$ $\quad$ $n_1 = 20$

k	0.25	0.50	0.75	0.25	0.50	0.75	0.25	0.50	0.75	0.25	0.50	0.75
0.0	0.099	0.158	0.171	0.061	0.097	0.117	0.046	0.89	0.108	0,043	0.087	0.103
0.2	0.118	0.169	0.186	0.073	0.118	0.128	0.057	0.115	0.126	0.051	0.113	0.123
0.4	0.141	0.182	0.203	0.097	0.133	0.143	0.071	0.128	0.139	0.078	0.126	0.137
0.6	0.185	0.196	0.221	0.113	0.155	0.164	0.103	0.147	0.157	0.095	0.146	0.155
0.8	0.230	0.208	0.237	0.129	0.178	0.181	0.121	0.173	0.176	0.117	0.170	0.175
1.0	0.279	0.225	0.251	0.141	0.195	0.204	0.149	0.192	0.198	0.141	0.189	0.187
1.2	0.331	0.291	0.304	0.169	0.237	0.233	0.169	0.219	0.223	0.163	0.214	0.221
1.4	0.398	0.312	0.327	0.182	0.261	0.274	0.187	0.247	0.257	0.184	0.239	0.243
1.6	0.461	0.326	0.339	0.227	0.315	0.301	0.219	0.293	0.294	0.215	0.270	0.280
1.8	0.365	0.355	0.348	0.251	0.341	0.356	0.249	0.330	0.330	0.247	0.312	0.327
2.0	0.571	0.458	0.427	0.289	0.407	0.419	0.287	0.387	0.379	0.278	0.358	0.368

Table(3)

$$Eff\,(\tilde{\theta}_B/\theta,R)\ \text{for,}\ \theta=10,\ \ \sigma=5,\ and\ t=\frac{|\theta-\theta_0|}{\sigma}\sqrt{n_1}$$

$n_1=5$ $n_1=10$ $n_1=15$ $n_1=20$

k t	0.25	0.50	0.75	0.25	0.50	0.75	0.25	0.50	0.75	0.25	0.50	0.75
0.0	6.526	4.011	2.953	7.821	4.189	3.413	8.201	5.203	3.932	8.900	5.813	4.310
0.2	5.703	3.315	2.341	6.395	3.507	2.971	7.053	4.047	3.045	8.015	4.965	3.591
0.4	4.250	2.712	1.932	4.395	2.915	2.576	5.917	3.341	2.610	7.217	4.072	2.843
0.6	3.801	1.903	1.057	3.317	2.014	1.789	4.731	2.865	1.969	6.044	3.001	2.169
0.8	2.753	1.142	0.980	2.650	1.581	1.316	3.021	2.054	1.497	4.351	2.379	1.629
1.0	1.830	0.873	0.789	1.813	0.899	0.983	2.450	1.631	1.000	3.190	1.832	1.110
1.2	1.012	0.705	0.703	1.259	0.758	0.837	1.534	0.971	0.872	2.418	1.006	0.901
1.4	0.831	0.571	0.639	0.985	0.696	0.721	1.069	0.823	0.763	1.505	0.879	0.829
1.6	0.679	0.485	0.561	0.739	0.587	0.649	0.869	0.765	0.703	0.892	0.781	0.740
1.8	0.485	0.445	0.483	0.590	0.569	0.573	0.753	0.682	0.649	.782	0.718	0.683
2.0	0.401	0.406	0.415	0.543	0.550	0.525	0.614	0.603	0.583	0.631	0.662	0.616

References

[1] Al---Robassi , A. M. , (2003) , Some Double Stage Shrinkage Estimator for the Mean of Normal Distribution . The Annual Conference on Statistics, Computer sciences and Operation Research, part II, p 86 .

[2] Arnold ,J.C. and AL—Bayyati , H.A. , (1970) ,On Double Stage Estimation of the Mean Using Prior Knowledge . Biometrics,Vol. (26),No.4,P.787.

[3] Hogg , R.V. and, Craig, A.T., (1995) , Introduction to Mathematical Statistics , Prentic-Hall International , Inc.

[4] Katte ,S.K., (1962) , Use Of Some Prior Knowledge In Estimation Means From Double Samples , Biometrics , Vol.,(18) ,P 139.

[5] Waikar,V.R., Schuurman,F.J. ,and Raghunathann ,(1984),On a Two Stage Shrinkage Testimator of The Mean of a Normal Distribution. Communication in Statistics.-Theory and Methods., 13(15),p, 1901-1913

مقدرات التقلص البيزية
لتقدير الوسط الحسابي للتوزيع الطبيعي

أحمد محمد غالب الرياصي
قسم الإحصاء، كلية العلوم الإدارية
جامعة تعز، اليمن

الملخص

بفرض أن x متغيراً عشوائياً يتبع التوزيع الطبيعي بوسط حسابي مجهول وتباين معلوم في حالة توافر معلومات مسبقة حول المعلمة المجهولة على شكل قيمة ابتدائية يكون من المفيد استخدام مقدرات التقلص ذات المرحلتين لتقدير المعلمة المجهولة إنظر (2) .

تناولت هذه الدراسة مقدرات التقلص البيزية ذات المرحلتين لتقدير الوسط الحسابي للتوزيع الطبيعي لقيم مختلفة لعامل التقلص K ، كما تم اختيار المجال R بواسطة تصغير متوسط مربعات الخطأ ، وقد بينت النتائج العددية أن المقدر المقترح يمتلك متوسط مربعات خطأ أقل من مقدرات البيزية وبالتالي كفاءة نسبية أعلى .