Linear Bayes Estimator for the Parameters of the Uniform Distribution $U(\theta_1, \theta_2)$

王立春
北京交通大学

2015.12.28

Outline

1. Introduction(介绍)
2. Linear Bayes estimator(线性贝叶斯估计)
3. Numerical comparison(数值比较)
4. Linear empirical Bayes estimator(线性经验贝叶斯估计)
5. Conclusions(结论)

1.介绍

Perhaps one of the most important distributions is the uniform distribution for continuous random variables. One reason is that the uniform $(0, 1)$ distribution is used as the basis for simulating most random variables.
A random variable that is uniformly distributed over the interval (θ_1, θ_2), which is often abbreviated $U(\theta_1, \theta_2)$, follows the probability density function given by

$$f(x; \theta_1, \theta_2) = (\theta_2 - \theta_1)^{-1} I(\theta_1 < x < \theta_2),$$

where $I(C)$ denotes the indicator function of the set C.

1.介绍

Let $X_1, X_2, \cdots, X_n$ be independently drawn from the uniform distribution $U(\theta_1, \theta_2)$. Note that $X_{(1)}$ and $X_{(n)}$ are sufficient and complete statistics, we obtain the classic uniformly minimum variance unbiased estimators (UMVUE) for the parameters θ_1 and θ_2 in the sense of minimizing mean squared error as follows:

$$\hat{\theta}_{1,U} = \frac{n}{n-1}X_{(1)} - \frac{1}{n-1}X_{(n)}, \quad \hat{\theta}_{2,U} = -\frac{1}{n-1}X_{(1)} + \frac{n}{n-1}X_{(n)}, \tag{1}$$

where $X_{(1)} = \min_{1\leq i\leq n} X_i$ and $X_{(n)} = \max_{1\leq i\leq n} X_i$.

1.介绍

Denote $\theta = (\theta_1, \theta_2)'$ and assume that the prior distribution is $G(\theta)$. Then the posterior distribution of θ given $X_{(1)}$ and $X_{(n)}$ would be

$$dH(\theta|x_{(1)}, x_{(n)}) \propto (\theta_2 - \theta_1)^{-n} I(\theta_1 < x_{(1)} < x_{(n)} < \theta_2) dG(\theta). \tag{2}$$

Thus, under the quadratic loss

$$L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)' D(\hat{\theta} - \theta), \tag{3}$$

where D is a positive definite matrix, the Bayes estimator of the parameter θ is the posterior expectation of $H(\theta|x_{(1)}, x_{(n)})$.

However, it is not easy to handle the relevant integration even if the prior $G(\theta)$ is common and ordinary, which may result in $H(\theta|x_{(1)}, x_{(n)})$ to be complicated or non-standard. Normally, the Bayes estimators of θ_1 and θ_2 are hardly expressed in explicit forms even for the single parameter uniform distribution $U(0, \theta_2)$, and approximate Bayes estimators are raised and computed using the idea of MCMC such as Gibbs sampling procedure and Metropolis method. Hence, in this situation the Bayes estimators are somewhat complicated and inconvenient to use.

2.线性贝叶斯估计

Assume that the prior $G(\theta)$ belongs to the following prior family

$$\mathcal{G} = \{G(\theta) : E[\theta_1^4 + \theta_2^4] < \infty\}. \tag{4}$$

Put $T = (X_{(1)}, X_{(n)}, X_{(1)}X_{(n)})'$ and define the LBE of θ, say $\hat{\theta}_{LB}$, be of the form $\hat{\theta} = BT + b$ satisfying

$$R(\hat{\theta}_{LB}, \theta) = \min_{B,b} E_{(T,\theta)} L(\hat{\theta}, \theta), \tag{5}$$

$$\left[E_{(T,\theta)}(\hat{\theta}_{LB} - \theta) = 0 \right]$$

where B and b are unknown matrices and $E_{(T,\theta)}$ denotes the expectation w.r.t the joint distribution of T and θ and the loss $L(\hat{\theta}, \theta)$ is given by (3).

Thus, we have the following two main results.
Theorem 2.1. Let $\hat{\theta}_{LB}$ be defined by (5). Then,

$$\hat{\theta}_{LB} = Cov(\theta, \tilde{\theta})A'[W + ACov(\tilde{\theta}, \tilde{\theta})A']^{-1}[T - AE(\tilde{\theta})] + E\theta,$$

where $\tilde{\theta} = (\theta_1, \theta_2, \theta_1^2, \theta_2^2, \theta_1\theta_2)'$ and the matrix

$$A = \begin{pmatrix} \frac{n}{n+1} & \frac{1}{n+1} & 0 & 0 & 0 \\ \frac{1}{n+1} & \frac{n}{n+1} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{n+2} & \frac{1}{n+2} & \frac{n}{n+2} \end{pmatrix}$$

and $W = E[Cov(T|\theta)]$.

Theorem 2.2. Let $\hat{\theta}_U = GT$ denote the UMVUE of θ with

$$G = \begin{pmatrix} \frac{n}{n-1} & \frac{-1}{n-1} & 0 \\ \frac{-1}{n-1} & \frac{n}{n-1} & 0 \end{pmatrix},$$

and $\hat{\theta}_{LB}$ be given by Theorem 2.1, then for $n \geq 2$,

$\mathrm{MSEM}(\hat{\theta}_{LB}) \leq \mathrm{MSEM}(\hat{\theta}_U)$.

Moreover, note that the MLE of θ, denoted by $\hat{\theta}_M = (\hat{\theta}_{1,M}, \hat{\theta}_{2,M})'$, equals

to FT with

$$F = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

If the sample size $n \geq 2$, then
$\mathrm{MSEM}(\hat{\theta}_{LB}) \leq \mathrm{MSEM}(\hat{\theta}_M)$.

The following are the measurements made on the tear strengths of 16 sample sheets of a silicone rubber used in a high voltage transformer (used by Tamhane and Dunlop (2000)):

33.74	34.40	32.62	32.57	34.69	33.78	36.76	34.31
37.61	33.78	35.43	33.22	33.53	33.68	33.24	32.98

Assume the above sample comes from the uniform distribution $U(\theta_1, \theta_2)$ and the parameter vector $\theta = (\theta_1, \theta_2)'$ have three classes of priors, and each class has the same mean $E(\theta) = (33, 37)'$ but has different covariance matrixes $Cov(\theta)$, i.e.,

$$I := \begin{pmatrix} 36 & 24 \\ 24 & 25 \end{pmatrix}, \quad II := \begin{pmatrix} 20 & 16 \\ 16 & 20 \end{pmatrix}, \quad III := \begin{pmatrix} 25 & 8 \\ 8 & 4 \end{pmatrix}.$$

We define the percentages of improvement of $\hat{\theta}_{LB}$ over $\hat{\theta}_U$ and $\hat{\theta}_M$, respectively, by

$$POI_U = \frac{tr(MSEM(\hat{\theta}_U) - MSEM(\hat{\theta}_{LB}))}{tr(MSEM(\hat{\theta}_U))} \tag{6}$$

and

$$POI_M = \frac{tr(MSEM(\hat{\theta}_M) - MSEM(\hat{\theta}_{LB}))}{tr(MSEM(\hat{\theta}_M))}. \tag{7}$$

Table 1—Estimation under different prior classes

prior	$\hat{\theta}_{LB}$	$\hat{\theta}_U$	$\hat{\theta}_M$	POI_U	POI_M	$tr(C$
class I	$\begin{pmatrix} 32.2478 \\ 37.9291 \end{pmatrix}$	$\begin{pmatrix} 32.2340 \\ 37.9460 \end{pmatrix}$	$\begin{pmatrix} 32.5700 \\ 37.6100 \end{pmatrix}$	0.0103	0.4722	6
class II	$\begin{pmatrix} 32.2524 \\ 37.9272 \end{pmatrix}$	$\begin{pmatrix} 32.2340 \\ 37.9460 \end{pmatrix}$	$\begin{pmatrix} 32.5700 \\ 37.6100 \end{pmatrix}$	0.0126	0.4734	4
class III	$\begin{pmatrix} 32.2665 \\ 37.8669 \end{pmatrix}$	$\begin{pmatrix} 32.2340 \\ 37.9460 \end{pmatrix}$	$\begin{pmatrix} 32.5700 \\ 37.6100 \end{pmatrix}$	0.0405	0.4883	2

From Table 1 we see that $\hat{\theta}_{LB}$ changes with the prior, while $\hat{\theta}_U$ and $\hat{\theta}_M$ remain unchanged. However, as stated in Theorem 2.2, since all three classes of priors belong to the prior family (2.1), both MSEM $(\hat{\theta}_U)-$ MSEM$(\hat{\theta}_{LB})$ and MSEM $(\hat{\theta}_M)-$ MSEM$(\hat{\theta}_{LB})$ are always nonnegative definite. Moreover, the percentages of improvement of $\hat{\theta}_{LB}$ over $\hat{\theta}_U$ and $\hat{\theta}_M$ tend to increase as the variation of the prior gets smaller. Note also that POI_M is always larger than POI_U, which is consistent with the fact that MSEM$(\hat{\theta}_U) \leq$ MSEM$(\hat{\theta}_M)$.

In the following we generate n random numbers from the uniform distribution $U(\theta_1, \theta_2)$, and the parameter vector $\theta = (\theta_1, \theta_2)'$ is assumed to follow a 2-dimensional Normal distribution $N_2(E(\theta), Cov(\theta))$, where the mean $E(\theta) = (33, 37)'$ but the covariance matrix $Cov(\theta)$ has a number of alternative values, i.e.,

$$(i) := \begin{pmatrix} 36 & 24 \\ 24 & 25 \end{pmatrix}, (ii) := \begin{pmatrix} 2\sqrt{5} & 4 \\ 4 & 5\sqrt{5}/2 \end{pmatrix},$$

$$(iii) := \begin{pmatrix} 1.2 & 0.8 \\ 0.8 & 5/6 \end{pmatrix}, (iv) := \begin{pmatrix} 0.8 & 0.64 \\ 0.64 & 0.8 \end{pmatrix}$$

Table 2—Estimations under the Normal prior (i): $tr(Cov(\theta)) = 61$

n	$\hat{\theta}_{LB}$	$\hat{\theta}_U$	$\hat{\theta}_M$	POI_U	POI_M
30	$\begin{pmatrix} 21.2365 \\ 49.0177 \end{pmatrix}$	$\begin{pmatrix} 21.1799 \\ 49.0863 \end{pmatrix}$	$\begin{pmatrix} 22.0801 \\ 48.1861 \end{pmatrix}$	0.0030	0.4843
60	$\begin{pmatrix} 21.6514 \\ 48.6117 \end{pmatrix}$	$\begin{pmatrix} 21.6376 \\ 48.6285 \end{pmatrix}$	$\begin{pmatrix} 22.0801 \\ 48.1861 \end{pmatrix}$	7.5258e-004	0.4919
100	$\begin{pmatrix} 21.7999 \\ 50.6357 \end{pmatrix}$	$\begin{pmatrix} 21.7945 \\ 50.6424 \end{pmatrix}$	$\begin{pmatrix} 22.0801 \\ 50.3568 \end{pmatrix}$	2.7179e-004	0.4951

Table 3—Estimations under the Normal prior (ii): $tr(Cov(\theta)) = 9\sqrt{5}/2$

n	$\hat{\theta}_{LB}$	$\hat{\theta}_U$	$\hat{\theta}_M$	POI_U	POI_M
30	$\begin{pmatrix} 30.1129 \\ 42.5919 \end{pmatrix}$	$\begin{pmatrix} 30.0298 \\ 42.6711 \end{pmatrix}$	$\begin{pmatrix} 30.4376 \\ 42.2633 \end{pmatrix}$	0.0109	0.4884
60	$\begin{pmatrix} 28.1733 \\ 42.4744 \end{pmatrix}$	$\begin{pmatrix} 28.1506 \\ 42.4985 \end{pmatrix}$	$\begin{pmatrix} 28.1773 \\ 42.2633 \end{pmatrix}$	0.0028	0.4929
100	$\begin{pmatrix} 28.2551 \\ 42.3949 \end{pmatrix}$	$\begin{pmatrix} 28.2456 \\ 42.4035 \end{pmatrix}$	$\begin{pmatrix} 28.3858 \\ 42.2633 \end{pmatrix}$	0.0010	0.4955

Table 4—Estimations under the Normal prior (iii): $tr(Cov(\theta)) = 61/30$

n	$\hat{\theta}_{LB}$	$\hat{\theta}_U$	$\hat{\theta}_M$	POI_U	POI_M
30	$\begin{pmatrix} 31.1658 \\ 38.4252 \end{pmatrix}$	$\begin{pmatrix} 31.0337 \\ 38.5807 \end{pmatrix}$	$\begin{pmatrix} 31.2771 \\ 38.3372 \end{pmatrix}$	0.0471	0.5071
60	$\begin{pmatrix} 31.1882 \\ 38.7678 \end{pmatrix}$	$\begin{pmatrix} 31.1515 \\ 38.8122 \end{pmatrix}$	$\begin{pmatrix} 31.2771 \\ 38.6867 \end{pmatrix}$	0.0125	0.4979
100	$\begin{pmatrix} 31.2153 \\ 38.7456 \end{pmatrix}$	$\begin{pmatrix} 31.2023 \\ 38.7615 \end{pmatrix}$	$\begin{pmatrix} 31.2771 \\ 38.6867 \end{pmatrix}$	0.0046	0.4973

Table 5—Estimations under the Normal prior (iv): $tr(Cov(\theta)) = 1.6$

n	$\hat{\theta}_{LB}$	$\hat{\theta}_U$	$\hat{\theta}_M$	POI_U	POI_M
30	$\begin{pmatrix} 31.5071 \\ 38.5584 \end{pmatrix}$	$\begin{pmatrix} 31.3397 \\ 38.7265 \end{pmatrix}$	$\begin{pmatrix} 31.5780 \\ 38.4882 \end{pmatrix}$	0.0566	0.5120
60	$\begin{pmatrix} 31.5034 \\ 38.5626 \end{pmatrix}$	$\begin{pmatrix} 31.4609 \\ 38.6053 \end{pmatrix}$	$\begin{pmatrix} 31.5780 \\ 38.4882 \end{pmatrix}$	0.0153	0.4993
100	$\begin{pmatrix} 31.5223 \\ 38.7452 \end{pmatrix}$	$\begin{pmatrix} 31.5062 \\ 38.7617 \end{pmatrix}$	$\begin{pmatrix} 31.5780 \\ 38.6898 \end{pmatrix}$	0.0056	0.4978

First, when the sample size n is fixed, as expected, both POI_U and POI_M increase as the variation of the prior gets smaller (i.e, as $tr(Cov(\theta))$ tends to be smaller); secondly, for four different priors, POI_U uniformly decreases as the sample size n grows larger, the reason is that $\hat{\theta}_U$ gets closer to $\hat{\theta}_{LB}$ as n gets larger;

thirdly, in Tables 2 and 3, when the variation of the prior is larger, POI_M changes a little as n grows larger, whereas, Tables 4 and 5 imply that when the variation of the prior is smaller POI_M decreases as n grows larger; finally, as seen in Table 1, POI_M is always larger than POI_U since $\text{MSEM}(\hat{\theta}_U) \leq \text{MSEM}(\hat{\theta}_M)$.

We simulate a case to see what are the performances of estimators like as the prior variances tend to infinity.

$Var(\theta_1) = 100$, $Var(\theta_2) = 400$ and $Cov(\theta_1, \theta_2) = 160$

n	$\hat{\theta}_{LB}$	$\hat{\theta}_U$	$\hat{\theta}_M$	POI_U	POI_M
30	$\begin{pmatrix} 0.1089 \\ 85.2534 \end{pmatrix}$	$\begin{pmatrix} -0.1922 \\ 85.4065 \end{pmatrix}$	$\begin{pmatrix} 2.5690 \\ 82.6452 \end{pmatrix}$	0.0037	0.4847
60	$\begin{pmatrix} -1.0385 \\ 84.0028 \end{pmatrix}$	$\begin{pmatrix} -1.1165 \\ 84.0413 \end{pmatrix}$	$\begin{pmatrix} 0.2795 \\ 82.6452 \end{pmatrix}$	9.3322e-004	0.4920
100	$\begin{pmatrix} -0.5245 \\ 83.4636 \end{pmatrix}$	$\begin{pmatrix} -0.5525 \\ 83.4772 \end{pmatrix}$	$\begin{pmatrix} 0.2795 \\ 82.6452 \end{pmatrix}$	3.3774e-004	0.4951

When the (norm of the) covariance matrix reaches beyond a certain level, if both the sample size and the prior correlation coefficient are the same then the percentages of improvement are almost unchanged.

Let us assume the parameters θ_1 and θ_2 have independent prior distributions, i.e., $\theta_1 \sim U(a, b)$ and $\theta_2 \sim U(b, c)$. Together with (2) we know that the UBE $\hat{\theta}_{UB} = (\hat{\theta}_{1,UB}, \hat{\theta}_{2,UB})'$ is given by

$$\begin{aligned} \hat{\theta}_{1,UB} &= \int\int \theta_1 f(\theta_1, \theta_2 | x_{(1)}, x_{(n)}) d\theta_1 d\theta_2, \quad (8) \\ \hat{\theta}_{2,UB} &= \int\int \theta_2 f(\theta_1, \theta_2 | x_{(1)}, x_{(n)}) d\theta_1 d\theta_2, \quad (9) \end{aligned}$$

where $f(\theta_1, \theta_2 | x_{(1)}, x_{(n)})$ denotes the posterior density of $\theta = (\theta_1, \theta_2)'$.

Note that
(i) If $b < x_{(1)} < x_{(n)}$, then the full conditional distributions θ_1 and θ_2 would be

$$\begin{aligned} f(\theta_1|\theta_2, x_{(1)}, x_{(n)}) &= \left[\frac{(\theta_2 - b)^{-n+1}}{n-1} - \frac{(\theta_2 - a)^{-n+1}}{n-1}\right]^{-1} (\theta_2 - \theta_1)^{-n} \\ &\quad \times I(a < \theta_1 < b), \\ f(\theta_2|\theta_1, x_{(1)}, x_{(n)}) &= \left[\frac{(c - \theta_1)^{-n+1}}{1-n} - \frac{(x_{(n)} - \theta_1)^{-n+1}}{1-n}\right]^{-1} (\theta_2 - \theta_1)^{-n} \\ &\quad \times I(x_{(n)} < \theta_2 < c). \end{aligned}$$

(ii) If $x_{(1)} < b < x_{(n)}$, then the full conditional distributions θ_1 and θ_2 would be

$$\begin{aligned} f(\theta_1|\theta_2, x_{(1)}, x_{(n)}) &= \left[\frac{(\theta_2 - x_{(1)})^{-n+1}}{n-1} - \frac{(\theta_2 - a)^{-n+1}}{n-1}\right]^{-1} (\theta_2 - \theta_1)^{-n} \\ &\quad \times I(a < \theta_1 < x_{(1)}), \\ f(\theta_2|\theta_1, x_{(1)}, x_{(n)}) &= \left[\frac{(c - \theta_1)^{-n+1}}{1-n} - \frac{(x_{(n)} - \theta_1)^{-n+1}}{1-n}\right]^{-1} (\theta_2 - \theta_1)^{-n} \\ &\quad \times I(x_{(n)} < \theta_2 < c). \end{aligned}$$

3.数值比较

(iii) If $x_{(1)} < x_{(n)} < b$, then the full conditional distributions θ_1 and θ_2 would be

$$
\begin{aligned}
f(\theta_1|\theta_2, x_{(1)}, x_{(n)}) &= \left[\frac{(\theta_2 - x_{(1)})^{-n+1}}{n-1} - \frac{(\theta_2 - a)^{-n+1}}{n-1}\right]^{-1} (\theta_2 - \theta_1)^{-n} \\
&\quad \times I(a < \theta_1 < x_{(1)}), \\
f(\theta_2|\theta_1, x_{(1)}, x_{(n)}) &= \left[\frac{(c - \theta_1)^{-n+1}}{1-n} - \frac{(b - \theta_1)^{-n+1}}{1-n}\right]^{-1} (\theta_2 - \theta_1)^{-n} \\
&\quad \times I(b < \theta_2 < c).
\end{aligned}
$$

3.数值比较

Employing the MCMC procedure we first use the idea of Devroye (1984) to generate samples from the above (i)-(iii) and then combine Geman and Geman (1984) with the following scheme to compute $\hat{\theta}_{1,UB}$ and $\hat{\theta}_{2,UB}$.

Table 7—Case I: $N = 10000, m_0 = 1000, \theta_1 \sim U(a, b), \theta_2 \sim U(b, c)$

n	a	b	c	$\hat{\theta}_{LB} = (\hat{\theta}_{1,LB}, \hat{\theta}_{2,LB})'$	$\hat{\theta}_{UB} = (\hat{\theta}_{1,UB}, \hat{\theta}_{2,UB})'$	$\|\|\hat{\theta}_{LB} - \hat{\theta}$
20	-8	4	11	$(-1.9586, 7.6018)'$	$(-1.9712, 7.6168)'$	0.1235
	-2	4	11	$(0.8702, 7.3304)'$	$(0.8674, 7.3334)'$	0.0832
	3	4	11	$(3.4957, 7.4288)'$	$(3.4825, 7.4325)'$	0.0855
50	-8	4	11	$(-2.5905, 7.3553)'$	$(-2.5924, 7.3508)'$	0.0299
	-2	4	11	$(1.1362, 7.4760)'$	$(1.1329, 7.4745)'$	0.0223
	3	4	11	$(3.5049, 7.3122)'$	$(3.5089, 7.3127)'$	0.0192
100	-8	4	11	$(-2.1455, 7.2938)'$	$(-2.1454, 7.2953)'$	0.0084
	-2	4	11	$(1.1751, 7.0134)'$	$(1.1746, 7.0139)'$	0.0040
	3	4	11	$(3.5287, 7.3993)'$	$(3.5302, 7.3999)'$	0.0053

3.数值比较

Table 8—Case I: $N = 10000, m_0 = 1000, \theta_1 \sim U(a, b), \theta_2 \sim U(b, c)$

n	a	b	c	$\hat{\theta}_{LB} = (\hat{\theta}_{1,LB}, \hat{\theta}_{2,LB})'$	$\hat{\theta}_{UB} = (\hat{\theta}_{1,UB}, \hat{\theta}_{2,UB})'$	$\|\|\hat{\theta}_{LB} - \theta$
20	-8	4	11	$(-1.9586, 7.6018)'$	$(-1.9712, 7.6168)'$	0.123
	-2	4	7.5	$(1.1080, 5.8169)'$	$(1.1152, 5.8054)'$	0.070
	3	4	5	$(3.5693, 4.5260)'$	$(3.5697, 4.5270)'$	0.011
50	-8	4	11	$(-2.5905, 7.3553)'$	$(-2.5924, 7.3508)'$	0.029
	-2	4	7.5	$(0.8998, 5.7803)'$	$(0.9038, 5.7765)'$	0.021
	3	4	5	$(3.5014, 4.5627)'$	$(3.5007, 4.5628)'$	0.003
100	-8	4	11	$(-2.1455, 7.2938)'$	$(-2.1454, 7.2953)'$	0.008
	-2	4	7.5	$(1.2067, 5.6953)'$	$(1.2074, 5.6957)'$	0.003
	3	4	5	$(3.4812, 4.5152)'$	$(3.4810, 4.5152)'$	7.1417e-

Case II. Let us assume $f(\theta_1, \theta_2) = \frac{2}{k^2} I(0 < \theta_1 < \theta_2 < k)$.

Table 9—Case II: $N = 10000$, $m_0 = 1000$

n	k	$\hat{\theta}_{LB} = (\hat{\theta}_{1,LB}, \hat{\theta}_{2,LB})'$	$\hat{\theta}_{UB} = (\hat{\theta}_{1,UB}, \hat{\theta}_{2,UB})'$	$\|\|\hat{\theta}_{LB} - \hat{\theta}_{UB}\|\|$
20	100	$(26.2690, 47.2641)'$	$(26.1913, 47.2133)'$	0.0928
	10	$(5.0995, 7.6376)'$	$(5.0957, 7.6506)'$	0.0135
	5	$(3.0665, 4.4994)'$	$(3.0651, 4.5102)'$	0.0110
	2	$(0.8912, 1.5219)'$	$(0.8889, 1.5263)'$	0.0050
50	100	$(41.6557, 65.3176)'$	$(41.6499, 65.3277)'$	0.0117
	10	$(6.2990, 6.7385)'$	$(6.3032, 6.7366)'$	0.0046
	5	$(2.4443, 3.4494)'$	$(2.4445, 3.4498)'$	5.8255e-4
	2	$(1.2458, 1.7987)'$	$(1.2460, 1.7992)'$	4.2735e-4
100	100	$(50.0809, 76.9095)'$	$(50.0791, 76.9165)'$	0.0073
	10	$(6.5906, 9.9839)'$	$(6.5912, 9.9698)'$	4.5475e-4
	5	$(2.2015, 4.1888)'$	$(2.2011, 4.1895)'$	3.2280e-4
	2	$(1.1354, 1.9578)'$	$(1.1353, 1.9579)'$	6.2221e-5

For the single parameter uniform distribution $U(0, \theta_2)$, we assume the prior $\pi(\theta_2)$ has finite second-order moment and mimic the above discussions, then the LBE of the parameter θ_2 is given by

$$\hat{\theta}_{2,LB} = a_0 X_{(n)} + b_0$$

with $a_0 = \frac{(n+1)(n+2)Var(\theta_2)}{(n+1)^2 E\theta_2^2 - n(n+2)(E\theta_2)^2}$ and $b_0 = \frac{[(1-a_0)n+1]E\theta_2}{n+1}$.
Specifically, we take the inverse Gamma distribution as the prior, i.e., $\pi(\theta_2) = \frac{\beta^\alpha}{\Gamma(\alpha)}(\frac{1}{\theta_2})^{\alpha+1} \exp(-\frac{\beta}{\theta_2}) I(\theta_2 > 0)$.

3.数值比较

Thus, note that $f(x_{(n)}|\theta_2) = \frac{n x_{(n)}^{n-1}}{\theta_2^n} I(0 < x_{(n)} < \theta_2)$, hence under the squared loss the UBE $\hat{\theta}_{2,UB}$ is

$$E(\theta_2|x_{(n)}) = \frac{\int_{x_{(n)}}^{\infty} (\frac{1}{\theta_2})^{\alpha+n} \exp(-\frac{\beta}{\theta_2}) d\theta_2}{\int_{x_{(n)}}^{\infty} (\frac{1}{\theta_2})^{\alpha+n+1} \exp(-\frac{\beta}{\theta_2}) d\theta_2}$$

$$= \frac{\beta}{\alpha+n-1} \frac{P(\chi^2(2(\alpha+n-1)) \leq 2\beta/x_{(n)})}{P(\chi^2(2(\alpha+n)) \leq 2\beta/x_{(n)})},$$

For instance, let $n = 5$, $x_{(n)} = 2$ and $\alpha = 3$ and $\beta = 8$, simple computations show that $a_0 = 1.1351$, $b_0 = 0.2163$ and $P(\chi^2(14) \leq 8) = 0.1107$ and $P(\chi^2(16) \leq 8) = 0.0511$.
Hence, we have $\hat{\theta}_{2,LB} = 2.4865$ and $\hat{\theta}_{2,UB} = 2.4758$, which show that the linear Bayes estimator is very close to the usual Bayes estimator.

4.线性经验贝叶斯估计

Denote $Y_i = \frac{n}{n-1}X_{(1)}^{(i)} - \frac{1}{n-1}X_{(n)}^{(i)}$ and $S_i = -\frac{1}{n-1}X_{(1)}^{(i)} + \frac{n}{n-1}X_{(n)}^{(i)}$. Following the idea of Samaniego and Vestrup (1999), we define the LEB estimator for θ as follows

$$\hat{\theta}_{LEB} = (\hat{\theta}_{1,LEB}, \hat{\theta}_{2,LEB})' = (\sum_{i=1}^{m+1} c_i Y_i, \sum_{i=1}^{m+1} d_i S_i)', \tag{10}$$

where $\sum_{i=1}^{m+1} c_i = 1$ and $\sum_{i=1}^{m+1} d_i = 1$ and $c_i \geq 0, \ d_i \geq 0$ for $i = 1, 2, \cdots, m+1$.

Rewrite

$$\begin{aligned}\hat{\theta}_{LEB} &= \begin{pmatrix} c_1 & 0 \ldots & c_{m+1} & 0 \\ 0 & d_1 \ldots & 0 & d_{m+1} \end{pmatrix} (Y_1, S_1, \ldots, Y_{m+1}, S_{m+1})' \\ &\triangleq C_d L, \end{aligned} \tag{11}$$

where $L = (Y_1, S_1, \ldots, Y_{m+1}, S_{m+1})'$. Setting $Q_i = (Y_i, S_i)', i = 1, 2, \cdots, m+1$ and following from the fact that

$$Cov(Q_i|\theta) = \begin{pmatrix} \frac{n(\theta_1-\theta_2)^2}{(n-1)(n+1)(n+2)} & \frac{-(\theta_1-\theta_2)^2}{(n-1)(n+1)(n+2)} \\ \frac{-(\theta_1-\theta_2)^2}{(n-1)(n+1)(n+2)} & \frac{n(\theta_1-\theta_2)^2}{(n-1)(n+1)(n+2)} \end{pmatrix} \triangleq Q, \tag{12}$$

we have

$$\begin{aligned}&Cov(L|\theta)\\&= \text{diag}\left(Q, \cdots, Q\right). \end{aligned} \tag{13}$$

Thus

$$\text{MSEM}(\hat{\theta}_{LEB}) = E_{(T_1,\cdots,T_{m+1},\theta)}(\hat{\theta}_{LEB} - \theta)(\hat{\theta}_{LEB} - \theta)'$$

$$= \begin{pmatrix} \frac{nE(\theta_1-\theta_2)^2}{(n-1)(n+1)(n+2)} \sum_{i=1}^{m+1} c_i^2 & -\frac{E(\theta_1-\theta_2)^2}{(n-1)(n+1)(n+2)} \sum_{i=1}^{m+1} c_i d_i \\ -\frac{E(\theta_1-\theta_2)^2}{(n-1)(n+1)(n+2)} \sum_{i=1}^{m+1} c_i d_i & \frac{nE(\theta_1-\theta_2)^2}{(n-1)(n+1)(n+2)} \sum_{i=1}^{m+1} d_i^2 \end{pmatrix}. \tag{14}$$

Theorem 4.1. If $n \geq 2$ and $[1-\sum_{i=1}^{m+1} c_i^2][1-\sum_{i=1}^{m+1} d_i^2]n^2 \geq [1-\sum_{i=1}^{m+1} c_i d_i]^2$, then $\hat{\theta}_{LEB}$ is superior to $\hat{\theta}_U$ in terms of MSEM criterion.

Moreover, under the conditions that $n \geq 2$ and $[2-\frac{n}{n-1}\sum_{i=1}^{m+1} c_i^2][2-\frac{n}{n-1}\sum_{i=1}^{m+1} d_i^2] \geq [1-\frac{1}{n-1}\sum_{i=1}^{m+1} c_i d_i]^2$, $\hat{\theta}_{LEB}$ is superior to $\hat{\theta}_M$ in terms of MSEM criterion,too.

Remark 4.1. For example, let $c_i = d_i = (m+1)^{-1}(i = 1, 2, \cdots, m+1)$, then it is easy to see that the conditions of Theorems 4.1 are satisfied.

The overall Bayes risk of an estimator $\hat{\theta}$ under the loss (3) and the prior (4) is defined by

$$R(\hat{\theta}, G(\theta)) = E_{(T_1,\cdots,T_m,(T,\theta))}(\hat{\theta}-\theta)' D(\hat{\theta}-\theta).$$

Theorem 4.2. If the conditions of Theorem 4.1 are satisfied, then

$$R(\hat{\theta}_{LEB}, G(\theta)) \le R(\hat{\theta}_U, G(\theta));$$

$$R(\hat{\theta}_{LEB}, G(\theta)) \le R(\hat{\theta}_M, G(\theta)).$$

5.结论

The paper employs the linear Bayes method to simultaneously estimate the parameters θ_1 and θ_2 of the uniform distribution $U(\theta_1, \theta_2)$ and proves their superiorities over the classical estimators UMVUE and MLE in terms of mean squared error matrix (MSEM) criterion as well.

simple and easy to calculate

has good approximation performances

The method can be extended easily to Weibull, log-normal and two-parameter Inverse Gaussian distribution, etc.

谢谢大家的聆听

部分参考文献

- Devroye, L., 1984. A simple algorithm for generating random variates with a log-concave density. Computing 33, 247-257.
- Geman, S., Geman, D., 1984. Stochastic relaxation,Gibbs distributions and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence 6, 721-741.
- Hartigan, J. A., 1969. Linear Bayesian methods. J. Roy. Statist. Soc. Ser. B 31, 440-454.
- Mao, S. S., Tang, Y. C., 2012. Bayesian statistics. Second Edition, China Statistics Press, Beijing.
- Samaniego, F. J., Vestrup, E., 1999. On improving standard estimators via linear empirical Bayes methods. Statistics & Probability Letters 44, 309-318.
- Zhang, W. P., Wei. L. S., Chen. Y., 2011. The superiorities of Bayes linear unbiased estimation in partitioned linear model. Journal of System Science and Complexity 5, 945-954.