

การประมาณค่าเฉลี่ยของการแจกแจงแบบปกติเมื่อทราบสัมประสิทธิ์ความแปรผัน Estimating the Mean of a Normal Distribution with Known Coefficient of Variation

นพคุณ ทองมวล¹และวุฒิชัย ศรีโสคาพล²
Noppakun Tongmol¹ and Wuttichai Srisodaphol²

บทคัดย่อ

วัตถุประสงค์ของงานวิจัยนี้เพื่อหาตัวประมาณของค่าเฉลี่ยประชากร θ ที่มีความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด เมื่อประชากรมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น θ และความแปรปรวน $a\theta^2$ โดยที่ $a > 0$ และ $\theta > 0$ เมื่อทราบสัมประสิทธิ์ความแปรผัน ตัวประมาณที่ผู้วิจัยนำเสนอ คือ $\hat{\theta}_{pr1}$ และ $\hat{\theta}_{pr2}$ ผลการวิจัยพบว่า ตัวประมาณที่ผู้วิจัยนำเสนอมีความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำกว่าตัวประมาณที่เสนอโดย Khan [1] Arnholt และ Hebert [2] Laheetharan และ Wijekoon [3]

คำสำคัญ : สัมประสิทธิ์ความแปรผัน ค่าเฉลี่ยประชากร

Abstracts

The objective of this research is to find the estimators of the population mean θ in the sense of minimum mean squared error (MSE) for a normal distribution with mean θ and variance $a\theta^2$, $a > 0$, $\theta > 0$ when known coefficient of variation. The two estimators, $\hat{\theta}_{pr1}$ and $\hat{\theta}_{pr2}$, are proposed. The results show that the proposed estimators are less MSE than the estimators which are proposed by Khan [1] Arnholt and Hebert [2] Laheetharan and Wijekoon [3]

Keywords : Coefficient of Variation, Population Mean

¹ นักศึกษาปริญญาโท สาขาสถิติประยุกต์ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น ขอนแก่น 40000

² อ.ดร., ภาควิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น ขอนแก่น 40000

* Corresponding author: E-mail : nop_stat@hotmail.com Tel. 087-4490248

จากงานประชุมวิชาการระดับชาติครั้งที่ 22 ปี 2555

บทนำ

ในการประมาณค่าเฉลี่ยประชากร θ จากประชากรที่มีการแจกแจงแบบปกติมีค่าเฉลี่ย θ และความแปรปรวนเป็นที่ทราบกันว่าค่าเฉลี่ยตัวอย่าง $\bar{X}$ เป็นตัวประมาณที่ไม่เอนเอียงและมีความแปรปรวนต่ำสุดแต่หากว่าตัวอย่างที่ได้มาจากประชากรที่มีการแจกแจงแบบปกติและทราบสัมประสิทธิ์ความแปรผัน (β) นั่นคือ $\sigma^2 = a\theta^2$ ซึ่ง $a = \beta^2 > 0$ และ $\theta > 0$ ในสถานการณ์นี้ Khan [1] ได้นำเสนอตัวประมาณผลรวมเชิงเส้นระหว่างค่าเฉลี่ยตัวอย่าง ($\bar{X}$) และส่วนเบี่ยงเบนมาตรฐานตัวอย่าง (s_n) ซึ่งเป็นตัวประมาณที่ไม่เอนเอียงและมีความแปรปรวนเท่ากับค่าขอบเขตล่างครามเมอร์และราว (Cramér – Rao Lower Bound) เมื่อ $n \rightarrow \infty$ แสดงให้เห็นว่า $\bar{X}$ มิใช่ตัวประมาณที่มีความแปรปรวนต่ำสุดของค่าเฉลี่ยประชากร θ โดยตัวประมาณที่ Khan นำเสนอ คือ

$$d^* = \omega d_2 + (1 - \omega) d_1$$

โดยที่ $\omega = \frac{2a}{(1+2a)}$

เมื่อตัวประมาณ d_1 คือ

$$d_1 = \bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

และตัวประมาณ d_2 คือ

$$d_2 = c_n \sqrt{n} s_n$$

โดยที่ $c_n = \frac{1}{\sqrt{2a}} \Gamma\left(\frac{n-1}{2}\right) / \Gamma\left(\frac{n}{2}\right)$

ซึ่งตัวประมาณ d_1 และตัวประมาณ d_2 เป็นตัวประมาณที่ไม่เอนเอียงของ θ

ดังนั้น

$$E(d_1) = E(d_2) = \theta$$

จะได้

$$\text{Bias}(d^*) = E(d^*) - \theta = 0$$

และ

$$\text{Var}(d^*) = \frac{a\theta^2}{n(1+2a)}$$

ความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ d^* คือ

$$\text{MSE}(d^*) = \text{Var}(d^*) + [\text{Bias}(d^*)]^2 = \frac{a\theta^2}{n(1+2a)}$$

จากงานวิจัยที่เกี่ยวข้องกับการประมาณค่าเฉลี่ยประชากร θ สำหรับประชากรที่มีการแจกแจงแบบปกติ $N(\theta, a\theta^2)$ มีผู้วิจัยหลายท่านได้เสนอแนวคิดในการลดความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ โดย Arnholt และ Hebert [2] ได้นำเสนอตัวประมาณ $\delta_k^* = k\bar{T}$ โดยที่ $\bar{T} = \bar{X}$ ไม่เอนเอียง

ของ θ ซึ่งมีความแปรปรวน $\sigma_T^2 = \frac{\sigma_X^2}{n}$ โดยที่มีความคลาดเคลื่อนกำลังสองเฉลี่ยของ δ_k คือ

$$MSE(\delta_k) = k^2 \sigma_T^2 + (k-1)^2 \theta^2$$

ได้ค่าคงที่ k ที่ทำให้ความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ δ_k ต่ำสุด คือ

$$k = \left(\frac{a}{n} + 1 \right)^{-1}$$

ดังนั้นตัวประมาณของ Arnholt และ Hebert นำเสนอ คือ

$$\delta_k^* = \left(\frac{a}{n} + 1 \right)^{-1} \tilde{T}$$

จะได้

$$Bias(\delta_k^*) = E(\delta_k^*) - \theta = \theta(k-1)$$

และ

$$Var(\delta_k^*) = k^2 \sigma_T^2$$

ความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ δ_k^* คือ

$$MSE(\delta_k^*) = Var(\delta_k^*) + [Bias(\delta_k^*)]^2 = k^2 \sigma_T^2 + (k-1)^2 \theta^2 = \frac{a\theta^2}{a+n}$$

Laheetharan และ Wijekoon [3] ได้นำเสนอการปรับปรุงตัวประมาณโดยให้ $T = k\bar{X}$ เป็นตัว

ประมาณที่เอนเอียงของ θ ซึ่ง $E(T) = k\theta$ โดยที่ $k = \left(\frac{a}{n} + 1 \right)^{-1}$ โดยที่นำ α คูณกับตัวประมาณ T จะได้ตัวประมาณ คือ

$$T^* = \alpha T$$

ได้ค่าคงที่ α ที่ทำให้ความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ T^* ต่ำสุด คือ

$$\alpha = \frac{k\theta^2}{[Var(T) + k^2\theta^2]}$$

กำหนดสัดส่วน $\tau^2 = \frac{Var(T)}{\theta^2}$ จะได้ $\alpha = \frac{k}{k^2 + \tau^2}$

จะได้

$$Bias(T^*) = E(T^*) - \theta = \theta(\alpha k - 1)$$

และ

$$Var(T^*) = \theta^2 \alpha^2 \tau^2$$

ความคลาดเคลื่อนกำลังสองเฉลี่ยของ T^* คือ

$$MSE(T^*) = Var(T^*) + [Bias(T^*)]^2 = \tau^2 \theta^2 (k^2 + \tau^2)^{-1} = \frac{a\theta^2}{a+n}$$

จากงานประชุมวิชาการระดับชาติครั้งที่ 22 ปี 2555

จากที่กล่าวมาแล้ว ผู้วิจัยจึงได้ปรับปรุงตัวประมาณค่าเฉลี่ย θ โดยนำตัวประมาณของ Khan มาปรับปรุง ซึ่งอาศัยแนวคิดจากงานวิจัยของ Arnholt และ Hebert หลังจากนั้นนำตัวประมาณที่ปรับปรุงด้วยแนวคิดของ Arnholt และ Hebert มาปรับปรุงอีกครั้งโดยใช้แนวคิดของ Laheetharan และ Wijekoon เพื่อหาตัวประมาณของ θ ที่มีความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด

วิธีการวิจัย

วิธีการดำเนินงานมีขั้นตอนดังต่อไปนี้

ตัวประมาณ $\hat{\theta}_{pr1}$ เป็นตัวประมาณที่ผู้วิจัยเสนอโดยที่เป็นการปรับปรุงตัวประมาณของ Khan ใช้แนวคิดของ Arnholt และ Hebert โดยให้

$$E(kd^*) = k\theta \quad \text{ซึ่ง } kd^* \text{ เป็นตัวประมาณที่เอนเอียงของ } \theta$$

ความเอนเอียงของ kd^* คือ

$$\text{Bias}(kd^*) = E(kd^*) - \theta = \theta(k-1)$$

ซึ่งความแปรปรวนของ kd^* คือ

$$\text{Var}(kd^*) = k^2 \text{Var}(d^*)$$

ดังนั้นความคลาดเคลื่อนกำลังสองเฉลี่ยของ kd^* คือ

$$\text{MSE}(kd^*) = \text{Var}(kd^*) + [\text{Bias}(kd^*)]^2 = k^2 \text{Var}(d^*) + \theta^2 (k-1)^2$$

หาค่าของ k ที่ทำให้ความคลาดเคลื่อนกำลังสองเฉลี่ยของ kd^* ต่ำสุด โดยหาอนุพันธ์อันดับที่ 1 ของความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ kd^* เทียบกับ k นั่นคือ

$$\frac{d\text{MSE}(kd^*)}{dk} = 2k\text{Var}(d^*) + 2k\theta^2 - 2\theta^2$$

กำหนดให้ $\frac{d\text{MSE}(kd^*)}{dk}$ เท่ากับศูนย์และแก้สมการหาค่า k ได้

$$k = \frac{n(1+2a)}{a+n(1+2a)}$$

และ $\frac{d^2\text{MSE}(kd^*)}{dk^2} = 2\text{Var}(d^*) + 2\theta^2 > 0$ แสดงว่า $k = \frac{n(1+2a)}{a+n(1+2a)}$ ทำให้ความคลาดเคลื่อนกำลังสอง

เฉลี่ยของ kd^* ต่ำสุด

ดังนั้นตัวประมาณ $\hat{\theta}_{pr1}$ คือ

$$\hat{\theta}_{pr1} = kd^* = \frac{n}{a + n(1 + 2a)} (2ad_2 + d_1)$$

ได้

$$Bias(\hat{\theta}_{pr1}) = E(\hat{\theta}_{pr1}) - \theta = \theta(k - 1)$$

และ

$$Var(\hat{\theta}_{pr1}) = k^2 Var(d^*)$$

ดังนั้นความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ $\hat{\theta}_{pr1}$ คือ

$$MSE(\hat{\theta}_{pr1}) = Var(\hat{\theta}_{pr1}) + [Bias(\hat{\theta}_{pr1})]^2 = \frac{a\theta^2}{a + n + 2an}$$

ตัวประมาณ $\hat{\theta}_{pr2}$ เป็นตัวประมาณที่ผู้วิจัยเสนอซึ่งนำตัวประมาณ $\hat{\theta}_{pr1}$ มาปรับปรุง เป็นการใช้แนวคิดของ Laheetharan และ Wijekoon โดยนำค่าคงที่ α คูณกับตัวประมาณ $\hat{\theta}_{pr1}$ และกำหนดสัดส่วนส่วน

$\tau^2 = \frac{Var(\hat{\theta}_{pr1})}{\theta^2}$ แล้วหลังจากนั้นหาค่าคงที่ α ที่ทำให้ ความคลาดเคลื่อนกำลังสองเฉลี่ยของ $\hat{\theta}_{pr2}$ ต่ำสุด

ให้

$$E(\alpha kd^*) = \alpha k\theta$$

ความเอนเอียงของ αkd^* คือ

$$Bias(\alpha kd^*) = E(\alpha kd^*) - \theta = \theta^2 (\alpha k - 1)$$

และความแปรปรวนของ αkd^* คือ

$$Var(\alpha kd^*) = \alpha^2 k^2 Var(d^*)$$

ความคลาดเคลื่อนกำลังสองเฉลี่ยของ αkd^* คือ

$$MSE(\alpha kd^*) = \alpha^2 k^2 Var(d^*) + \theta^2 (\alpha k - 1)^2$$

หาค่าคงที่ α ที่ทำให้ความคลาดเคลื่อนกำลังสองเฉลี่ยของ αkd^* ต่ำสุด โดยหาอนุพันธ์อันดับที่ 1 ของความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ αkd^* เทียบกับ α นั่นคือ

$$\frac{dMSE(\alpha kd^*)}{d\alpha} = 2\alpha k^2 Var(d^*) + 2\alpha k^2 \theta^2 - 2k\theta^2$$

กำหนดให้ $\frac{dMSE(\alpha kd^*)}{d\alpha}$ เท่ากับศูนย์และแก้สมการหาค่า α ได้

$$\alpha = \frac{k^2 \theta^2}{Var(\hat{\theta}_{pr1}) + k^2 \theta^2}$$

จากงานประชุมวิชาการระดับชาติครั้งที่ 22 ปี 2555

$$\text{และ } \frac{d^2 MSE(\alpha k d^*)}{d\alpha^2} = 2k^2 Var(d^*) + 2k^2 \theta^2 > 0 \text{ แสดงว่า } \alpha = \frac{k^2 \theta^2}{Var(\hat{\theta}_{pr1}) + k^2 \theta^2} \text{ ทำให้ความ}$$

คลาดเคลื่อนกำลังสองเฉลี่ยของ $\alpha k d^*$ ต่ำสุด

ดังนั้นตัวประมาณ $\hat{\theta}_{pr2}$ คือ

$$\hat{\theta}_{pr2} = \alpha k d^* = \frac{n}{a + n(1 + 2a)} (2ad_2 + d_1)$$

$$\text{กำหนดสัดส่วน } \tau^2 = \frac{Var(\hat{\theta}_{pr1})}{\theta^2} \text{ จะได้ } \alpha = \frac{k}{k^2 + \tau^2}$$

$$\text{จะได้ } Bias(\hat{\theta}_{pr2}) = E(\hat{\theta}_{pr2}) - \theta = \theta(\alpha k - 1)$$

$$\text{และ } Var(\hat{\theta}_{pr2}) = \theta^2 \alpha^2 \tau^2$$

ดังนั้นความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ $\hat{\theta}_{pr2}$ คือ

$$MSE(\hat{\theta}_{pr2}) = Var(\hat{\theta}_{pr2}) + [Bias(\hat{\theta}_{pr2})]^2 = \tau^2 \theta^2 (k^2 + \tau^2)^{-1} = \frac{a\theta^2}{a + n + 2an}$$

หลังจากนั้นเปรียบเทียบความคลาดเคลื่อนกำลังสองเฉลี่ยระหว่างตัวประมาณของ Khan(d^*) Arnholt

และ Hebert (δ_k^*) Laheetharan และ Wijekoon (T^*) กับ $\hat{\theta}_{pr1}$ และ $\hat{\theta}_{pr2}$

ผลการวิจัยและอภิปรายผลการวิจัย

ในการเปรียบเทียบประสิทธิภาพของตัวประมาณระหว่างตัวประมาณ d^* δ_k^* T^* $\hat{\theta}_{pr1}$ และ $\hat{\theta}_{pr2}$

โดยพิจารณาจากประสิทธิภาพสัมพัทธ์ (Relative efficiency : RE)

ประสิทธิภาพสัมพัทธ์ของตัวประมาณ $\hat{\theta}_{pr1}$ เทียบกับ ตัวประมาณ d^* ของค่าเฉลี่ย θ คือ

$$RE(\hat{\theta}_{pr1} | d^*) = \frac{MSE(d^*)}{MSE(\hat{\theta}_{pr1})} = \frac{a + n + 2an}{n + 2an} > 1$$

ดังนั้น $MSE(\hat{\theta}_{pr1}) < MSE(d^*)$

ประสิทธิภาพสัมพัทธ์ของตัวประมาณ $\hat{\theta}_{pr1}$ เทียบกับ ตัวประมาณ δ_k^* ของค่าเฉลี่ย θ คือ

$$RE(\hat{\theta}_{pr1} | \delta_k^*) = \frac{MSE(\delta_k^*)}{MSE(\hat{\theta}_{pr1})} = \frac{a + n + 2an}{a + n} > 1$$

ดังนั้น $MSE(\hat{\theta}_{pr1}) < MSE(\delta_k^*)$

ประสิทธิภาพสัมพัทธ์ของตัวประมาณ d^* เทียบกับ ตัวประมาณ δ_k^* ของค่าเฉลี่ย θ คือ

$$RE(d^* | \delta_k^*) = \frac{MSE(\delta_k^*)}{MSE(d^*)} = \frac{n + 2an}{a + n} > 1$$

ดังนั้น $MSE(d^*) < MSE(\delta_k^*)$

ความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ $\hat{\theta}_{pr1}$ เท่ากับความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ $\hat{\theta}_{pr2}$ นั่นคือ

$$MSE(\hat{\theta}_{pr1}) = MSE(\hat{\theta}_{pr2}) = \frac{a\theta^2}{a + n + 2an}$$

ความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ T^* เท่ากับความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ δ_k^* นั่นคือ

$$MSE(T^*) = MSE(\delta_k^*) = \frac{a\theta^2}{a + n}$$

จากการพิจารณาความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ พบว่า

$$MSE(\hat{\theta}_{pr1}) = MSE(\hat{\theta}_{pr2}) < MSE(d^*) < MSE(\delta_k^*) = MSE(T^*)$$

ผลการคำนวณค่าความคลาดเคลื่อนกำลังสองเฉลี่ยระหว่างตัวประมาณของ Khan (d^*) Arnholt และ Hebert (δ_k^*) Laheetharan และ Wijekoon (T^*) และตัวประมาณที่ผู้วิจัยนำเสนอ คือ $\hat{\theta}_{pr1}$ และ $\hat{\theta}_{pr2}$ โดยที่กำหนดค่า a θ และ n ดังตารางที่ 1-3

ตารางที่ 1 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ $\hat{\theta}_{pr1}$ $\hat{\theta}_{pr2}$ d^* δ_k^* และ T^* เมื่อ $n = 10$

θ	a	MSE		
		$\hat{\theta}_{pr1}, \hat{\theta}_{pr2}$	d^*	δ_k^*, T^*
10	0.01	0.097943	0.098039	0.099900
	0.25	1.639344	1.666667	2.439024
	0.81	2.998889	3.091603	7.493062
20	0.01	0.391773	0.392157	0.399600
	0.25	6.557377	6.666667	9.756098
	0.81	11.995557	12.366412	29.972248
30	0.01	0.881489	0.882353	0.899101
	0.25	14.754098	15.000000	21.951220
	0.81	26.990004	27.824427	67.437558

จากงานประชุมวิชาการระดับชาติครั้งที่ 22 ปี 2555

ตารางที่ 2 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ $\hat{\theta}_{pr1}$ $\hat{\theta}_{pr2}$ d^* δ_k^* และ T^* เมื่อ $n = 30$

θ	a	MSE		
		$\hat{\theta}_{pr1}, \hat{\theta}_{pr2}$	d^*	δ_k^*, T^*
10	0.01	0.032669	0.032680	0.033322
	0.25	0.552486	0.555556	0.826446
	0.81	1.020023	1.030534	2.629017
20	0.01	0.130676	0.130719	0.133289
	0.25	2.209945	2.222222	3.305785
	0.81	4.080091	4.122137	10.516066
30	0.01	0.294022	0.294118	0.299900
	0.25	4.972376	5.000000	7.438017
	0.81	9.180204	9.274809	23.661149

ตารางที่ 3 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยของตัวประมาณ $\hat{\theta}_{pr1}$ $\hat{\theta}_{pr2}$ d^* δ_k^* และ T^* เมื่อ $n = 50$

θ	a	MSE		
		$\hat{\theta}_{pr1}, \hat{\theta}_{pr2}$	d^*	δ_k^*, T^*
10	0.01	0.019604	0.019608	0.019996
	0.25	0.332226	0.333333	0.497512
	0.81	0.614521	0.618321	1.594174
20	0.01	0.078416	0.078431	0.079984
	0.25	1.328904	1.333333	1.990050
	0.81	2.458084	2.473282	6.376698
30	0.01	0.176436	0.176471	0.179964
	0.25	2.990033	3.000000	4.477612
	0.81	5.530688	5.564885	14.347569

จากตารางที่ 1-3 พบว่า ตัวประมาณ $\hat{\theta}_{pr1}$ และ $\hat{\theta}_{pr2}$ มีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำกว่าตัวประมาณ d^* δ_k^* และ T^*

สรุปผลการวิจัย

ในการประมาณค่าเฉลี่ยประชากร θ สำหรับประชากรที่มีการแจกแจงแบบปกติมีค่าเฉลี่ยเป็น θ และความแปรปรวน $a\theta^2$ โดยที่ $a > 0$ และ $\theta > 0$ เมื่อทราบสัมประสิทธิ์ความแปรผัน พบว่าตัวประมาณที่ผู้วิจัยนำเสนอ คือ ตัวประมาณ $\hat{\theta}_{pr1}$ และ $\hat{\theta}_{pr2}$ มีความแม่นยำกว่าตัวประมาณของ Khan (d^*) Arnholt และ Hebert (δ_k^*) Laheetharan และ Wijekoon (T^*) นั่นคือ ตัวประมาณ $\hat{\theta}_{pr1}$ และ $\hat{\theta}_{pr2}$ มีความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำกว่า ตัวประมาณของ Khan (d^*) Arnholt และ Hebert (δ_k^*) Laheetharan และ Wijekoon (T^*)

เอกสารอ้างอิง

- [1] 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), 1039-104.
- [2] Arnholt, A. T., and Hebert, J. E. (1995). Estimating the Mean with known Coefficient of Variation. **The American Statistician**, (49), 367-369.
- [3] Laheetharan, A. and Wijekoon, P. (2008). Improved estimation of the population parameters when some additional information is available. **Statistical Papers**, (51) , 889-914
- [4] Gleser, L. J., and Healy, J. D. (1976). Estimating the Mean of a Normal Distribution with Known Coefficient of variation. **Journal of the American Statistical Association**, (71), 977-981.