INTERVAL KEPERCAYAAN UNTUK PERBEDAAN KOEFISIEN VARIASI DARI DISTRIBUSI LOGNORMAL

I. Pebriyani[1*], Bustami[2], S. Sugiarto[2]

[1]Mahasiswa Program S1 Matematika
[2]Dosen Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Riau
Kampus Bina Widya Pekanbaru, 28293, Indonesia.

*irma_pebriyani@yahoo.com

ABSTRACT

This article discusses confidence interval for the difference of variation coefficients for lognormal distribution using the pivotal quantity, where pivotal quantity for the variation coefficients is unavailable, but there is a parameter having pivotal quantity in coefficients of variation *i.e* σ^2. Therefore, σ^2 is constructed for pivotal quantity coefficients of variation using the *Generalized Pivotal Approach* (*GPA*). This confidence interval is shown in the term of coverage probabilities by implementing simulation studies with Matlab 7.6.0.

Keywords: *coefficient of variation, confidence interval, generalized pivotal approach, lognormal distribution.*

ABSTRAK

Artikel ini membahas interval kepercayaan untuk perbedaan koefisien variasi dari distribusi Lognormal dengan menggunakan kuantitas pivot, dimana kuantitas pivot untuk koefisien variasi tidak tersedia tetapi didalam koefisien variasi terdapat parameter yang mempunyai kuantitas pivot yaitu σ^2. Oleh karena itu, σ^2 dikonstruksi untuk kuantitas pivot koefisien variasi dengan menggunakan *Generalized Pivotal Approach (GPA)*. Bentuk interval ini diperlihatkan dalam bentuk peluang cangkupan dengan menggunakan studi simulasi melalui program Matlab versi 7.6.0.

Kata Kunci: *koefisien variasi, interval kepercayaan, generalized pivotal approach, distribusi lognormal.*

1. PENDAHULUAN

Salah satu aspek yang penting dalam statistika inferensi adalah menaksir nilai parameter dari suatu populasi melalui analisa data sampel yang telah diperoleh dari populasi tersebut. Penaksiran parameter ini dapat dilakukan dengan dua cara yaitu penaksiran titik dan penaksiran interval. Penulisan ini membahas tentang taksiran interval. Secara umum, taksiran interval diperoleh dengan menggunakan metode kuantitas pivot. Dalam hal ini kuantitas pivot untuk koefisien variasi belum tersedia, tetapi berdasarkan Buntao & Niwipong [2] diketahui bahwa koefisien variasi hanya bergantung pada σ^2. Oleh karena itu, dikonstruksi kuantitas pivot dari σ^2 untuk koefisien variasi dari distribusi lognormal melalui kajian distribusi normal, metode ini disebut dengan *GPA*.

Penulis mendetailkan taksiran interval untuk perbedaan koefisien variasi dari distribusi lognormal berdasarkan Buntao & Niwipong [2].

2. DISTRIBUSI LOGNORMAL SERTA TAKSIRAN TITIK UNTUK DISTRIBUSI LOGNORMAL

Distribusi lognormal adalah suatu distribusi yang terkait dengan distribusi normal, tetapi diasumsikan nilai variabel random hanya yang bernilai positif [1:h. 199]. Misalkan Y berdistribusi $N(\mu,\sigma^2)$, maka variabel random $X = e^Y$ disebut variabel random lognormal dimana X berdistribusi $LN(\mu,\sigma^2)$.

Fungsi densitas peluang untuk distribusi lognormal adalah

$$f(x; ln(\mu,\sigma^2)) = \begin{cases} \dfrac{1}{x\sigma\sqrt{2\pi}} exp\left(-\dfrac{(ln(x)-\mu)^2}{2\sigma^2}\right) & ; \quad 0 \le x, -\infty < \mu < \infty, \sigma < \infty \\ 0 & ; \quad x \le 0. \end{cases}$$

Populasi ekspektasi dari X dinotasikan dengan $E(X) = e^{\mu+\frac{\sigma^2}{2}}$ dan populasi variansi dari X dinotasikan dengan $Var(X) = e^{2\mu+\sigma^2}\left(e^{\sigma^2}-1\right)$.

Dalam [2:h. 6692] koefisien variasi dari distribusi lognormal bisa dimodifikasi sebagai berikut

$$KV(\eta) = \frac{\sqrt{Var(X)}}{E(X)}$$

$$KV(\eta) = \sqrt{e^{\sigma^2}-1}.$$

3. INTERVAL KEPERCAYAAN UNTUK KOEFISIEN VARIASI DARI DISTRIBUSI LOGNORMAL

Perhitungan interval kepercayaan untuk koefisien variasi memuat tiga langkah. Langkah pertama karena koefisien variasi hanya bergantung pada σ^2 maka dibentuk interval kepercayaan untuk σ^2 dari distribusi normal, langkah kedua membentuk interval kepercayaan untuk koefisien variasi dari distribusi lognormal melalui variansi dari distribusi normal dan langkah ketiga adalah mengkonstruksi interval kepercayaan untuk koefisien variasi berdasarkan langkah kedua.

Teorema 1 [3:h. 220] $Y_1,...,Y_n$ suatu sampel random dimana $Y \sim N(\mu,\sigma^2)$. Jika $\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ dan $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2$, maka

1. $\overline{Y}$ dan S^2 independen.

2. $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$.

Bukti teorema ini dapat dilihat pada buku [3:h.221].

Berdasarkan Teorema (1) bisa dibentuk interval kepercayaan untuk variansi dari distribusi normal menjadi

$$\left(\frac{(n-1)S^2}{\chi^2_{(n-1),(\alpha/2)}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{(n-1),(1-\alpha/2)}} \right).$$

Interval kepercayaan untuk koefisien variasi dari distribusi lognormal adalah

$$\sqrt{\exp\left(\frac{(n-1)S^2}{\chi^2_{(n-1),(\alpha/2)}} \right) - 1} < \sqrt{\exp(\sigma^2) - 1} < \sqrt{\exp\left(\frac{(n-1)S^2}{\chi^2_{(n-1),(1-\alpha/2)}} \right)}$$

$$\sqrt{\exp\left(\frac{(n-1)S^2}{\chi^2_{(n-1),(\alpha/2)}} \right) - 1} < \eta < \sqrt{\exp\left(\frac{(n-1)S^2}{\chi^2_{(n-1),(1-\alpha/2)}} \right)}.$$

Kemudian, dapat dikonstruksi interval kepercayaan dua sisi untuk $100(1-\alpha)\%$ untuk koefisien variasi dari distribusi Lognormal sebagai

$$CI_{LN} = [L,U] = \left[\sqrt{\exp\left(\frac{(n-1)S^2}{\chi^2_{(n-1),(\alpha/2)}} \right) - 1}, \sqrt{\exp\left(\frac{(n-1)S^2}{\chi^2_{(n-1),(1-\alpha/2)}} \right)} \right].$$

CI_{LN} merupakan interval kepercayaan untuk distribusi lognormal, L dan U masing-masing adalah batas atas dan batas bawah untuk interval tersebut.

4. INTERVAL KEPERCAYAAN UNTUK PERBEDAAN KOEFISIEN VARIASI DARI DISTRIBUSI LOGNORMAL

GPA adalah metode yang digunakan untuk mengkonstriksi interval kepercayaan apabila kuantitas pivot belum tersedia, tetapi dapat menggunakan kuantitas pivot lain yang bersesuaian. Ide untuk mengkonstruksi interval kepercayaan untuk perbedaan koefisien variasi dari distribusi lognormal adalah dengan menggunakan kuantitas pivot, secara umum yang didasarkan pada nilai X .

Misalkan ada dua populasi yang bebas untuk $i = 1,2$ dari $X_i = (X_{i1}, X_{i2}, ..., X_{in})$ dengan $Y_i = \ln(X_i)$ dan $Y_i \sim N(\mu_i, \sigma_i^2)$. Misalkan X_i berdistribusi lognormal, $LN(\mu_i, \sigma_i^2)$, koefisien variasi dari berbagai kelompok adalah $\eta_i = \sqrt{\exp(\sigma_i^2) - 1}$.

Perbedaan dari dua koefisien variasi dari distribusi lognormal adalah

$$\psi = \eta_1 - \eta_2$$

$$\psi = \left(\sqrt{\exp(\sigma_1^2) - 1}\right) - \left(\sqrt{\exp(\sigma_2^2) - 1}\right)$$

Perbedaan koefisien variasi dari distribusi lognormal dapat dikonstruksi dengan menggunakan metode *GPA*.

Definisi 1 [4] Misalkan $R = r(Z; z, \nu)$ adalah fungsi dari Z, z, ν, dimana $\nu = (\theta, \delta)$. Jika R memenuhi kedua sifat berikut ini maka R disebut dengan *generalized pivotal quantity*. Sifat A : R berdistribusi peluang bebas dari parameter yang tidak diketahui. Sifat B : pivot pengamatan, didefinisikan sebagai $r_{obs} = r(z; z, \nu)$, tidak bergantung pada parameter pengganggu δ.

Berdasarkan Definisi (1) metode umum didefinisikan sebagai sebuah statistik yang berdistribusi bebas dengan parameter yang tidak diketahui dan tidak bergantung pada parameter pengganggu. Koefisien variasi hanya bergantung pada pada parameter σ^2.

Misalkan S_i^2 disimbolkan dengan variansi sampel untuk transformasi log data $Y_i = \ln(X_i)$ dan s^2 disimbolkan dengan sampel variansi dari populasi $ke - i$. Kuantitas pivot secara umum untuk σ_i^2 adalah $R_{\sigma_i^2} = (n_i - 1)s_i^2 / U_i$ dimana U_i adalah distribusi chi-kuadarat dengan derajat kebebasan $n_i - 1$, $U_i = (n_i - 1)S_i^2 / \sigma_i^2$.

Kuantitas pivot secara umum untuk dua populasi $R_{\sigma_1^2}$ dan $R_{\sigma_2^2}$ adalah bebas. Kuantitas pivot secara umum untuk ψ adalah

$$R_\psi = \left(\sqrt{\exp(R_{\sigma_1^2}) - 1}\right) - \left(\sqrt{\exp(R_{\sigma_2^2}) - 1}\right)$$

Catatan untuk s_i^2, adalah sebagai berikut :

(i) Distribusi dari R_ψ adalah bebas untuk semua parameter yang tidak diketahui.

(ii) Pivot pengamatan tidak bergantung pada parameter pengganggu dan nilai R_ψ adalah sama dengan ψ sebagai $s_i^2 = S_i^2$. Oleh karena itu, R_ψ adalah kuantitas pivot secara umum untuk mengkonstruksi interval kepercayaan untuk ψ dan kuantil ini mungkin digunakan untuk mengkonstruksi R_ψ. Jika $R_\psi(1-\alpha)$ adalah dinotasikan dengan persentil ke $100(1-\alpha)$ dari distribusi dari R_ψ, kemudian $R_\psi(1-\alpha)$ adalah $100(1-\alpha)\%$ interval kepercayaan bagian atas untuk ψ. Kemudian, $CI = [\psi_L, \psi_U] = \left[R_\psi\left(\alpha/2\right), R_\psi\left(1 - \alpha/2\right)\right]$ adalah $100(1-\alpha)\%$ interval kepercayaan umum dua sisi untuk perbedaan koefisien variasi. Kemudian Peluang cangkupan dari interval kepercayaan umum dapat dihitung dengan simulasi menggunakan program Matlab versi 7.6.0.

5. SIMULASI STUDI DAN PEMBAHASAN

Dalam simulasi studi ini akan ditunjukkan peluang cangkupan untuk interval kepercayaan dari perbedaan koefisien variasi dari distribusi lognormal. Dengan menggunakan 2 populasi diambil sampel yang berukuran n_1 dan n_2, dimana $n_1 = (10,50,100,200,500)$ dan n_2 merupakan kombinasi dari n_1. Selanjutnya, parameter yang digunakan adalah μ dan σ^2, nilai dari masing-masing parameter adalah 0,6 dan 0,5. Simulasi ini dilakukan dengan 2 kali pengulangan, pengulangan petama disimbolkan dengan m_1 dan pengulangan kedua disimbolkan dengan m_2 masing-masing pengulangan ini adalah 10 dan 100. Setelah itu, untuk memperoleh nilai perkiraan peluang cangkupan dari $100(1-\alpha)\%$ interval kepercayaan umum dua sisi untuk perbedaan koefisien variasi digunakan berbagai nilai α yaitu (0.1,0.4,0.8). Sehingga, nilai perkiraan peluang cangkupan tersebut dapat dilihat pada Tabel 1.

Tabel 1. Nilai peluang cangkupan interval kepercayaan umum

n_1	n_2	$\alpha_1 = 0.1$	$\alpha_2 = 0.4$	$\alpha_3 = 0.8$
10	10	0.48	0.55	0.49
	50	0.52	0.41	0.50
	100	0.45	0.50	0.51
	200	0.51	0.42	0.40
	500	0.53	0.52	0.40
50	10	0.56	0.53	0.48
	50	0.39	0.50	0.57
	100	0.43	0.48	0.56
	200	0.45	0.50	0.49
	500	0.50	0.49	0.56
100	10	0.46	0.52	0.59
	50	0.55	0.51	0.60
	100	0.48	0.41	0.55
	200	0.56	0.54	0.49
	500	0.39	0.49	0.48
200	10	0.52	0.51	0.51
	50	0.59	0.52	0.51
	100	0.43	0.52	0.51
	200	0.52	0.50	0.45
	500	0.38	0.46	0.51
500	10	0.55	0.58	0.56
	50	0.50	0.54	0.51
	100	0.51	0.47	0.44
	200	0.59	0.50	0.62
	500	0.52	0.51	0.52

Pada Tabel 1 kolom pertama dan kedua merupakan ukuran sampel yang diambil, kolom ketiga, keempat dan kelima merupakan nilai α yang digunakan.

Pada Tabel 1 dapat dilihat bahwa nilai perkiraan peluang cangkupan dari $100(1-\alpha)\%$ interval kepercayaan umum untuk dua sisi untuk perbedaan koefisien variasi, semakin menurun, semakin besar α nilai perkiraan peluang cangkupan dari $100(1-\alpha)\%$ interval kepercayaan umum untuk dua sisi untuk perbedaan koefisien variasi semakin kecil.

DAFTAR PUSTAKA

[1] Bain, L. J & M. Engelhardt. 1993. *Introduction to Probability and Mathematical Statistics. Second Edition.* Duxbury Press, California.

[2] Buntao, N & S. Niwipong. 2012. Confidence Intervals for the Difference of Coeffisients of Variation for Lognormal Distributions and Delta-Lognormal Distribution. *Applied Mathematical Sciences*, **134**: 6691–6704.

[3] Casella, G & R. L. Berger. 1990. *Statistical Inference.* Duxbury Press, California.

[4] Weerahandi, S. 1993. Generalized Confidence Intervals. *Journal of the American Statistical Association,* **88**: 899-905.