

Pengembangan Bagan Kendali 3 Sigma Berbasis Simulasi Untuk Pengawasan Variabilitas Proses Multivariat

Rahmat Sagara¹⁾, Khoirul Umam²⁾

Program Studi Matematika, Fakultas Bisnis Institut Teknologi dan Bisnis Kalbis
Jalan Pulomas Selatan kav.22, Jakarta Timur 13210

¹⁾Email: rahmat.sagara@kalbis.ac.id

²⁾Email: khoirul.umam@kalbis.ac.id

Abstract: This article describes a method in developing control charts for generalized variance as a quality statistics in terms of process variability through simulation. Mathematical equation that maps sample size n and number of quality variables p onto constant multiplier of standard deviation K is obtained thorough least square method using simulated data. The constant K for a certain n and p is used for control charts with the upper control limits of $UCL = \mu + K\sigma$ where μ and σ are the mean and the standard deviation of the generalized variance, respectively. The simulation of finding the constant K is used with the constraint of 3 sigma paradigm.

Keywords: control chart, generalized variance, simulation, 3 sigma paradigm

Abstrak: Artikel ini memaparkan sebuah metode pengembangan bagan kendali untuk generalized variance sebagai sebuah statistik mutu yang berkaitan dengan variabilitas proses menggunakan simulasi. Persamaan matematika yang mengaitkan ukuran sampel n dan banyaknya variabel mutu p terhadap konstanta pengali standar deviasi K didapatkan melalui metode kuadrat terkecil menggunakan data simulasi. Konstanta K yang didapatkan untuk suatu nilai n dan p kemudian digunakan untuk bagan kendali dengan batas kendali atas $UCL = \mu + K\sigma$ dimana μ dan σ masing-masing adalah rata-rata dan standar deviasi dari generalized variance. Simulasi pencarian K ini dilakukan dengan kendala memenuhi paradigma 3 sigma.

Kata kunci: bagan kendali, generalized variance, paradigma 3 sigma, simulasi

I. PENDAHULUAN

Bentuk umum dari bagan kendali adalah

$$UCL = \mu + K\sigma$$

$$CL = \mu$$

$$LCL = \mu - K\sigma$$

dimana μ dan σ masing-masing adalah rata-rata dan standar deviasi dari statistik mutu yang digunakan. Dengan mengadopsi paradigma 3 sigma bahwa dari sejuta produk yang dihasilkan hanya 2700 produk yang cacat [1], konstanta K bernilai 3 jika distribusi dari statistik mutu tersebut adalah normal berdasarkan fakta bahwa jika variabel acak $X \sim N(\mu, \sigma^2)$ maka

$$P(X < \mu - 3\sigma \text{ atau } X > \mu + 3\sigma) \cong 0.002700$$

atau

$$P(X < \mu - 3\sigma) = P(X > \mu + 3\sigma) \cong 0.00135$$

Konstanta K tidak sama dengan 3 jika statistik

mutu tidak berdistribusi normal, namun dekat dengan 3 jika statistik mutu berdistribusi normal asimtotis – berdistribusi normal jika ukuran sampel menuju tak hingga. Sebagai contoh untuk statistik mutu generalized variance dan vector variance, [2] memperlihatkan bahwa ketika $K = 3$ maka nilai PFA (probability of a false alarm) jauh dari angka 0.002700 meskipun untuk $n = 100$.

Untuk mencari nilai K yang lebih baik, [2] melakukan simulasi untuk ukuran sampel $n = 5, 20, \text{ dan } 100$ dan banyaknya variabel $p = 3$. Dengan menggunakan metode simulasi serupa, pada artikel ini akan dibahas pencarian konstanta K_p yang diekspresikan sebagai sebuah fungsi $K_p = f(n)$ untuk statistik mutu generalized variance.

II. METODE PENELITIAN

Misalkan $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n$ sebuah sampel acak berukuran n dari populasi berdistribusi normal

multivariat $N_p(\bar{\mathbf{r}}, \Sigma)$ dimana $\bar{\mathbf{r}}$ adalah vektor rata-rata dan Σ adalah matriks kovariansi populasinya. Misalkan pula S adalah matriks kovariansi sampelnya. Maka (lihat [3]) distribusi eksak dari generalized variance $|S|$ sama dengan distribusi dari statistik

$$\frac{|\hat{Q}|}{(n-1)^p} \sum_{k=1}^p Z_k$$

dimana variabel acak Z_1, Z_2, \dots, Z_p saling bebas dan masing-masing berdistribusi χ^2 dengan derajat kebebasan $n-1, n-2, \dots, n-p$. Dengan demikian, (lihat [4]), didapatkan bahwa

$$E[|S|] = b_1|\Sigma| \text{ dan } \text{Var}[|S|] = b_2|\Sigma|^2$$

dimana

$$b_1 = \frac{1}{(n-1)^p} \prod_{k=1}^p (n-k)$$

$$b_2 = \frac{1}{(n-1)^p} \prod_{k=1}^p (n-k) \times \left(\prod_{j=1}^p (n-j+2) - \prod_{i=1}^p (n-i) \right)$$

Sedangkan distribusi asimtotis dari generalized variance (lihat [4]) adalah

$$\frac{1}{\sqrt{b_2}} (|S| - b_1|\hat{Q}|) \rightarrow N(0, |\hat{Q}|^2)$$

Karena $E[|S|] = b_1|\Sigma|$ dan $\text{Var}[|S|] = b_2|\Sigma|^2$ maka bagan kendali untuk generalized variance $|S|$ adalah

$$UCL = b_1|\Sigma| + K_p \sqrt{b_2|\Sigma|^2}$$

$$CL = b_1|\Sigma|$$

dimana K_p merupakan konstanta pengali yang berkaitan dengan p dan merupakan fungsi dari ukuran sampel n serta memenuhi paradigma 3 sigma, yaitu bahwa:

$$P(|S| > b_1|\hat{Q}| + K_p \sqrt{b_2|\hat{Q}|^2}) \cong 0.00135$$

A. Simulasi

Tanpa mengurangi keumuman, simulasi dilakukan menggunakan data hasil pembangkitan dari populasi normal multivariat dengan vektor rata-rata vektor nol dan matriks kovariansi matriks identitas. Algoritma untuk menentukan nilai K_p ini adalah sebagai berikut: Untuk suatu nilai p , bangkitkan 100000 (seratus ribu) data berukuran n dari populasi berdistribusi $N_p(\bar{\mathbf{0}}, \mathbf{I})$; Untuk masing-masing matriks data berukuran $n \times p$ itu, hitung nilai dari *generalized variance* yaitu determinan dari matriks kovariansi sampel; Urutkan ke 100000 nilai *generalized variance* itu dari kecil ke besar; Dapatkan Q , yaitu nilai *generalized variance* urutan ke 99865; dan Dapatkan konstanta K_p dengan formula $K_p = \frac{Q - b_1}{\sqrt{b_2}}$

B. Hasil simulasi

Hasil simulasi konstanta pengali untuk $p = 2, 3, 5, 10, 15, 20$ dan $n = 25, 50, 100, 250, 500, 1000$ disajikan pada Tabel 1.

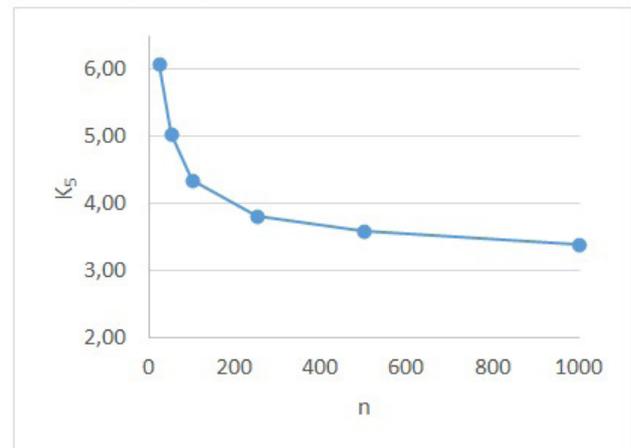
Tabel 1 Hasil Simulasi Konstanta Pengali

p	2	3	5	10	15	20	
n	25	4.44183	5.08007	6.08987	8.08241	9.8437	9.84422
	50	4.02038	4.38765	5.04556	6.54118	7.38947	8.21068
	100	3.70155	3.98221	4.35005	5.18595	5.99643	6.47568
	250	3.47854	3.64598	3.81005	4.27615	4.71758	4.99328
	500	3.2566	3.42571	3.59104	3.89642	4.10858	4.24538
	1000	3.20954	3.32293	3.37985	3.58576	3.67614	3.99297

C. Perumusan Model

Untuk suatu nilai p , hasil simulasi pasangan terurut (n, K_p) untuk $n = 25, 50, 100, 250, 500$, dan 1000 digambarkan diagram plotnya dan ditentukan model persamaan K_p sebagai fungsi dari n .

Sebagai contoh, untuk banyaknya variabel $p = 5$ dan ukuran sampel $n = 25, 50, 100, 250, 500$, dan 1000 didapatkan plot pasangan terurut (n, K_5) sebagai mana disajikan pada Gambar 1.



Gambar 1 Plot hubungan n dan K_5

Gambar 1 memperlihatkan bahwa K_5 berbanding terbalik dengan n , dan karena generalized variance berdistribusi normal asimtotis maka $K_5 \rightarrow 3$ jika $n \rightarrow \infty$. Oleh karena itu model hubungan n dan K_5 bisa dirumuskan melalui persamaan

$$K_5 = 3 + An^{-B}$$

dengan A dan B adalah dua konstanta positif.

III. HASIL DAN PEMBAHASAN

Pandang kembali model hubungan n dan K_5 dalam persamaan

$$K_5 = 3 + An^{-B} \tag{1}$$

dimana A dan B adalah dua konstanta positif. Pada bagian ini, akan dicari taksiran dari A dan B menggunakan data konstanta pengali K_5 dari hasil simulasi pada Tabel 1.

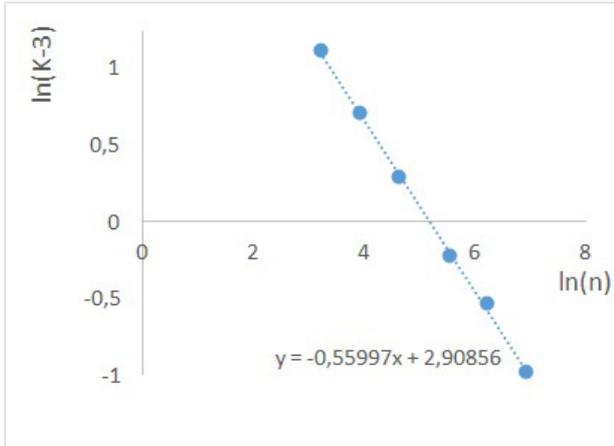
Persamaan 1 ekuivalen dengan persamaan

$$K_5 - 3 = An^{-B} \tag{2}$$

Penerapan fungsi logaritma natural pada kedua sisi Persamaan 2 mendapatkan

$$\ln(K - 3) = \ln A - B \ln n \tag{3}$$

sehingga didapatkan hubungan linier antara $h(K_5 - 3)$ dan $h(n)$.



Gambar 2 Hubungan linier $\ln(K_5-3)$ dan $\ln(n)$

Dengan menggunakan metode kuadrat terkecil didapatkan taksiran $h(A) = 2.908563$ atau $A = 18.33043$, dan $B = 0.55997$. Dengan demikian, untuk $p = 5$, konstansa pengali K_5 sebagai fungsi dari n dituliskan sebagai berikut

$$K_5(n) = 3 + 18.33043n^{-0.55997}$$

Menggunakan prosedur yang sama, untuk $p = 2, 3, 5, 10, 15, 17,$ dan 20 didapat taksiran untuk A dan B seperti disajikan pada Tabel 2.

Tabel 2 Taksiran nilai untuk A dan B

p	A	B
2	8.36591	0.53896
3	10.25519	0.50514
5	18.33043	0.55997
10	33.99401	0.58872
15	50.61895	0.61861
20	42.73245	0.55325

IV. SIMPULAN

Pada paper ini sudah dipaparkan mengenai pengembangan bagan kendali 3 sigma berbasis simulasi untuk pengawasan variabilitas proses multivariat menggunakan general variance dengan hasil sebagai berikut.

Bagan kendali berupa

$$UCL = b_1|\Sigma| + K_p \sqrt{b_2|\Sigma|^2}$$

$$CL = b_1|\Sigma|$$

dimana

Ó adalah matriks definit positif berukuran $p \times p$ yang ditargetkan sebagai matriks kovariansi populasi,

$$b_1 = \frac{1}{(n-1)^p} \prod_{k=1}^p (n-k)$$

$$b_2 = \frac{1}{(n-1)^{2p}} \prod_{k=1}^p (n-k) \times \left(\prod_{j=1}^p (n-j+2) - \prod_{i=1}^p (n-i) \right)$$

K_p adalah konstanta pengali yang berkaitan dengan p dan merupakan fungsi dari ukuran sampel n yang dirumuskan sebagai $\bar{K}_p = 3 + An^{-B}$ dengan nilai A dan B kedua-duanya bilangan positif.

Konstanta pengali K_p nilainya dipengaruhi oleh p selain dipengaruhi oleh n . Dengan demikian konstanta pengali merupakan sebuah fungsi dari dua variabel p dan n . Akan tetapi pada penelitian ini belum kepada pencarian bentuk eksplisit $K = f(p,n)$. Diharapkan, pada penelitian selanjutnya didapat model hubungan eksplisit konstanta pengali K sebagai fungsi dua variabel dari p dan n tersebut.

V. DAFTAR RUJUKAN

- [1] D. C. Montgomery, Introduction to Statistical Quality Control, New York: John Wiley & Sons, 2009, p. 29.
- [2] M. A. Djauhari, R. M. Salleh, Z. Zolkeply dan L. S. Li, "On the reliability of Shewhart-type control charts for multivariate process variability," AIP Conference Proceedings, vol. 1842, no. 1, pp. 030028-6, 2017.
- [3] M. A. Djauhari, "Improved monitoring of multivariate process variability," Journal of Quality Technology, vol. 37, no. 1, p. 33, 2005.
- [4] M. A. Djauhari, Advanced Monitoring Techniques for Complex Process Variability in Manufacturing Industry, Serdang: Universiti Putra Malaysia Press, 2016, pp. 13-15, 21.