

Evaluasi Kinerja Bagan Kendali \tilde{F}_m Pada Proses *Short-Run*

Darmanto

Jurusan Statistika, FMIPA, Universitas Brawijaya, Jln. Veteran 1 – Malang, 65145, Indonesia

DOI: 10.20961/performa.17.1.18740

Abstrak

Proses produksi manufaktur yang tren saat ini adalah *short-run*. Proses *short-run* bersifat *job shop* (spesifikasi berbeda tiap proses) dan *just in-time* (jumlah produk sesuai pesanan). Hal tersebut menyebabkan parameter proses tidak diketahui karena tidak tersedianya data dan umumnya jumlah produk sedikit. Bagan kendali \tilde{F}_m adalah salah satu bagan kendali yang dirancang untuk proses *short-run*. Cara kerja bagan kendali \tilde{F}_m mengikuti konsep *successive difference* dan di bawah asumsi distribusi Normal multivariat. Tingkat sensitifitas suatu bagan kendali dievaluasi berdasarkan nilai *average run length* (ARL). Pada penelitian ini, nilai ARL dihitung berdasar simulasi pergeseran terhadap vektor rata-rata dengan mencatat titik ke- m pertama yang keluar dari batas kendali. Simulasi pergeseran vektor rata-rata dari target (μ_0) dilakukan secara serempak dengan sifat pergeseran positif ($\mu_s = \mu_0 + \delta$). Variasi ukuran data dan banyak variabel pada penelitian ini berturut-turut adalah $m = 20, 50$ and $p = 2, 4, 8$. Tiap skema kombinasi dari δ, m and p diiterasi masing-masing sebanyak 250.000 kali. Hasil simulasi menunjukkan bahwa untuk semua skema ketika kedua parameter diketahui $ARL_0 \approx 370$. Adapun nilai ARL_1 , seiring bertambah jauhnya vektor rata-rata digeser menjauh dari target, nilainya semakin kecil. Artinya, sensitifitas bagan kendali \tilde{F}_m meningkat. Kesimpulan ini juga berlaku untuk bertambahnya banyaknya variabel p , sedangkan bertambahnya ukuran n akan mengurangi sensitifitas bagan kendali \tilde{F}_m .

Kata kunci: *average run length* (ARL), *job shop*, *just in-time*, *short-run*, *successive difference*

Abstract

The manufacturing production process that is currently trend is *short-run*. *Short-run* process is a *job shop* and a *just in-time*. These causes the process parameters to be unknown due to unavailability of data and generally a small amount of product. The \tilde{F}_m control chart is one of the control charts which designed for the short run. The procedure of the \tilde{F}_m control chart follows the concept of *successive difference* and under the assumption of the multivariate Normal distribution. The sensitivity level of a control chart is evaluated based on the *average run length* (ARL) value. In this study, the ARL value was calculated based on the shift simulation of the average vector by recording the first m -point out of the control limits. The average vector shift simulation of the target (μ_0) is performed simultaneously with the properties of a positive shift ($\mu_s = \mu_0 + \delta$). Variations of data size and many variables in this study were $m = 20, 50$ and $p = 2, 4, 8$, respectively. Each scheme (a combination of δ, m and p) is iterated 250,000 times. The simulation results show that for all schemes when both parameters are known $ARL_0 \approx 370$. But, when parameters are unknown, ARL_1 turn to smaller. This conclusion also implied when the number of p and n are increased, it reduce the sensitivity of the \tilde{F}_m control chart.

Keywords: *average run length* (ARL), \tilde{F}_m , *job shop*, *just in-time*, *short-run*, *successive difference*

1. Pendahuluan

Saat ini, proses produksi yang sedang tren adalah proses produksi *short-run*, yakni suatu proses produksi dengan kondisi: (1) Proses *job shop* yaitu karakteristik kualitas beserta standar (spesifikasi) yang dimonitor berbeda-beda tiap produk, (2) Sistem *Just-in-Time* (JIT) yaitu jumlah produk yang diproduksi sedikit (*low volume*) sehingga proses produksi berjalan lebih pendek dari yang konvensional, (3) Parameter proses tidak tersedia (Khoo, dkk., 2005; Marques, dkk., 2015). Hal ini dikarenakan tidak cukup atau tidak tersedia data produksi sebelumnya (*no historical data*). Dinyatakan pula bahwa secara umum proses produksi *short-run* memproduksi produk dengan kuantitas kurang dari 50 (20-50) (Khoo dan Quah, 2002; Elam dan Case, 2005; Fonseca, dkk., 2007; Montgomery, 2009; Jaupi, dkk., 2013). Dikarenakan jumlah produk yang sedikit, maka amatan yang paling banyak diterapkan pada kondisi *short-run* adalah amatan individu.

Secara konvensional, bagan kendali multivariat amatan individu didasarkan pada bagan kendali klasik T^2 -Hotelling. Untuk menghitung statistik T^2 bagan kendali T^2 -Hotelling, diperlukan statistik vektor rata-rata

dan matriks kovarians yang diperoleh berdasarkan data produksi sebelumnya (*historical data*). Dikarenakan pada kondisi *short-run* proses produksi berjalan cepat dan karakteristik kualitas dan spesifikasi berbeda-beda tiap proses produksi, maka data historis relatif sulit diperoleh. Oleh karenanya, hal ini menjadi kendala untuk memonitor proses produksi dengan menggunakan bagan kendali T^2 -Hotelling konvensional.

Untuk mengatasi kondisi (1) dan (2) pada proses produksi *short-run*, para peneliti telah mengusulkan beberapa bagan kendali multivariat *short-run*. Di antaranya adalah Scholz-Tosch (1994) yang mengusulkan bagan kendali multivariat X untuk sampel berukuran kecil. Pada tahun 2001, Quesenberry mengusulkan bagan kendali multivariat *snapshot* Q yang merupakan perluasan dari bagan kendali univariat Q oleh Quesenberry (1993). Selanjutnya, Khoo-Quah (2002) mengusulkan bagan kendali multivariat V_m . Statistik V_m merupakan modifikasi dan perluasan dari statistik T^2 dengan mentransformasi ke dalam distribusi multivariat normal standar, $N_p(\mathbf{0}, \mathbf{I}_p)$.

Kemudian, pada tahun 2005, Khoo-Gan mengusulkan bagan kendali multivariat CUSUM untuk amatan individu menggunakan statistik V_m . Selanjutnya, pada tahun 2012, Zou, dkk. mengusulkan bagan kendali multivariat *short-run* EWMA untuk data yang tidak diketahui distribusinya dengan basis *rank-spatial*. Pada tahun 2013, Jaupi dkk. juga mengusulkan bagan kendali multivariat *short-run* untuk mengendalikan rata-rata dan variabilitas dengan melibatkan fungsi pengaruh (*influence function*). Dari beberapa bagan kendali tersebut terdapat dua bagan kendali yang mempunyai persamaan yaitu bagan kendali yang diusulkan oleh Scholz-Tosch (1994) dan bagan kendali yang diusulkan oleh Khoo-Quah (2002). Persamaan dari kedua bagan kendali tersebut adalah keduanya dikembangkan dari konsep *successive difference* (SD) untuk menghitung estimator parameter dan mengasumsikan bahwa data yang diamati berdistribusi normal multivariat. Pada tulisan ini, peneliti mengkaji dan mengevaluasi kinerja bagan kendali \tilde{F}_m yang diusulkan oleh Scholz dan Tosch (1994) berupa tingkat sensitifitas dalam mendeteksi ada tidaknya titik yang tidak terkendali.

2. Metode Penelitian

Evaluasi kinerja bagan kendali dapat diketahui dengan melihat nilai ARL_1 yang diperoleh dari hasil simulasi. Pada kondisi kedua parameter diketahui, langkah awal evaluasi kinerja bagan kendali multivariat *short-run* \tilde{F}_m adalah inisiasi awal kedua parameter yang diketahui yaitu $\boldsymbol{\mu}_0 = \mathbf{0}$ dan $\boldsymbol{\Sigma}_0 = \mathbf{I}_p$. Selanjutnya adalah membangkitkan data secara random sebanyak $N = 10.000$ yang kemudian dinyatakan sebagai \mathbf{x}_f (*future observation*), $\mathbf{x}_f \sim N_p(\boldsymbol{\mu}_s, \mathbf{I}_p)$ untuk $p = 2, 4$ dan 8 . p adalah banyaknya karakteristik kualitas yang diamati dan $\boldsymbol{\mu}_s$ adalah vektor rata-rata yang telah mengalami pergeseran proses sebesar δ yaitu $\boldsymbol{\mu}_s = \boldsymbol{\mu}_0 + \delta$, $\delta = 0; 0,5; 1; 1,5; 2; 2,5; 3; 4; 5$.

Selanjutnya adalah menghitung statistik \tilde{F}_m sebagaimana pada Persamaan (1). Kemudian membandingkan statistik \tilde{F}_m dengan $\chi_{\alpha,p}^2$ untuk mendeteksi ada tidaknya titik yang keluar pertama kali (RL_1), $\tilde{F} < \chi_{\alpha,p}^2$. Jika $RL_1 = 0$, ulangi langkah-langkah sebelumnya. Namun, jika $RL_1 \neq 0$, maka catat sebagai nilai RL_1 . Seluruh proses tersebut diiterasi sebanyak 250.000 kali. Langkah terakhir adalah mencari rata-rata RL_1 (ARL_1). Nilai ARL_1 didapatkan dengan cara menjumlahkan seluruh RL_1 yang diperoleh dari tiap-tiap iterasi dibagi dengan banyaknya iterasi.

Sedangkan untuk kondisi kedua parameter tidak diketahui, proses simulasi terbagi menjadi dua fase yakni Fase I dan Fase II. Fase I adalah fase estimasi parameter vektor rata-rata dan matriks kovarians berdasarkan m titik yang terkendali (*in-control*). Fase II adalah fase menghitung nilai ARL_1 berdasarkan data bangkitan yang telah mengalami pergeseran terhadap statistik vektor rata-rata dan matriks kovarians yang telah diperoleh pada Fase I.

Fase I dimulai dengan membangkitkan data secara random sebanyak $N = 10.000$ yang berdistribusi normal multivariat $N_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ dengan variasi $p = 2, 4$ dan 8 . Kemudian, mengambil secara random dari $N = 10.000$ untuk tiap-tiap p sebanyak $m = 20$ dan 50 . Selanjutnya, menghitung statistik $\bar{\mathbf{x}}_m$ dan $\tilde{\mathbf{S}}_m$ pada Persamaan (3). Langkah berikutnya adalah menghitung statistik \tilde{F}_m sebagaimana pada Persamaan (2). Jika ada 1 titik yang tidak terkendali ($\tilde{F}_m > F_{\alpha,p,d-p+1}$), maka tahapan awal pada fase ini diulangi hingga semua

titik m terkendali ($\tilde{F}_m < F_{\alpha,p,d-p+1}$). Jika semua titik m terkendali, maka statistik \bar{x}_m dan \tilde{S}_m dari m titik terkendali ini disimpan untuk perhitungan pada Fase II.

Pada Fase II, langkah awalnya adalah membangkitkan data secara random sebanyak $N = 10.000$ yang kemudian dinyatakan sebagai \mathbf{x}_f (*future observation*), $\mathbf{x}_f \sim N_p(\boldsymbol{\mu}_s, \mathbf{I}_p)$ untuk $p = 2, 4$ dan 8 . Kemudian, menghitung statistik \tilde{F}_m sebagaimana pada Persamaan (2) dengan memanfaatkan statistik \bar{x}_m dan \tilde{S}_m yang diperoleh pada Fase I. Langkah selanjutnya adalah membandingkan antara statistik \tilde{F}_m dengan $F_{\alpha,p,d-p+1}$ dan menyatakan tidak terkendali secara statistik jika dan hanya jika $\tilde{F}_m > F_{\alpha,p,d-p+1}$. Kemudian, menghitung nilai ARL_I dengan cara seperti sebelumnya.

Setelah memperoleh nilai ARL_I masing-masing untuk bagan kendali, maka selanjutnya adalah membandingkan nilai ARL_I kedua bagan kendali. Semakin kecil nilai ARL_I , maka dapat dinyatakan bahwa semakin sensitif suatu bagan kendali mendeteksi adanya pergeseran vektor rata-rata proses dari vektor rata-rata yang sebenarnya (*target*). Selanjutnya, menerapkan kedua bagan kendali dengan data numerik.

3. Hasil dan Pembahasan

Adapun analisis evaluasi kinerja dari kedua bagan kendali tersebut adalah dengan melihat nilai ARL_I hasil simulasi. Semakin kecil nilai ARL_I , maka akan semakin sensitif suatu bagan kendali dalam mendeteksi adanya titik yang tidak terkendali (*out of control*). Secara praktik, nilai ARL_I dihitung dengan mencatat titik ke- m pertama yang keluar dari batas bagan kendali.

Pada penelitian ini, simulasi pergeseran terhadap vektor rata-rata dilakukan secara serempak dengan sifat pergeseran positif (+). Misal: untuk $p = 3$, maka

$$\begin{aligned} \boldsymbol{\mu}_s &= \boldsymbol{\mu}_0 + \boldsymbol{\delta} \\ &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \boldsymbol{\delta} = \begin{pmatrix} \delta \\ \delta \\ \delta \end{pmatrix} \end{aligned}$$

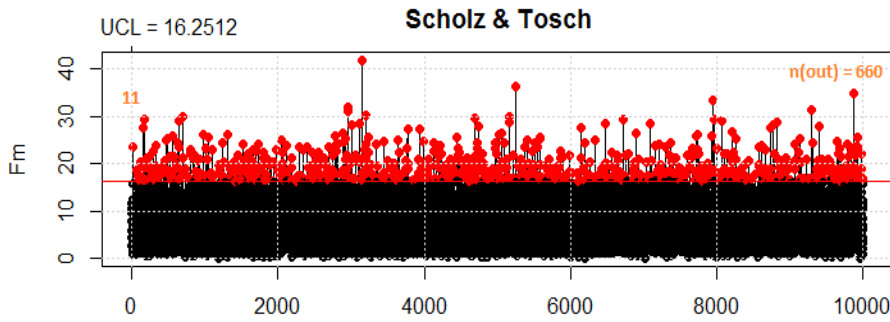
Tabel 1 merupakan tabel yang berisi nilai ARL bagan kendali multivariat *short-run* \tilde{F}_m untuk kondisi kedua parameter diketahui. Nilai δ dimulai dari 0 yang artinya tidak ada pergeseran vektor rata-rata proses, 0,5 artinya vektor rata-rata proses digeser menjauh secara serempak sebesar 0,5 dari vektor rata-rata yang sebenarnya ($\boldsymbol{\mu}_0 = \mathbf{0}$), 1 artinya vektor rata-rata proses digeser menjauh secara serempak sebesar 1 dari vektor rata-rata yang sebenarnya, dan demikian seterusnya untuk nilai $\delta = 1,5; 1; 1,5; 2; 2,5; 3; 4; 5$.

Berdasarkan Tabel 1, untuk besar $\delta = 0$ atau tidak ada pergeseran vektor rata-rata, bagan kendali \tilde{F}_m berlaku untuk setiap p , diketahui nilai ARL_0 berada di sekitar angka 370. Nilai $ARL_0 = 370$ artinya, titik amatan ke-370 adalah titik yang pertama kali keluar yang dideteksi oleh bagan kendali. Atau dengan kata lain, bagan kendali membutuhkan amatan sebanyak 370 untuk menyatakan bahwa proses mengalami pergeseran vektor rata-rata.

Tabel 1. Nilai ARL Kondisi Kedua Parameter Diketahui dengan $m = 20$

δ	$p = 2$	$p = 4$	$p = 8$
	\tilde{F}_m	\tilde{F}_m	\tilde{F}_m
0	370.52	370.29	370.19
0.5	129.56	101.33	71.86
1	27.74	15.16	7.26
1.5	7.71	3.64	1.77
2	3.07	1.57	1.08
2.5	1.68	1.11	1.00
3	1.21	1.01	1.00
4	1.01	1.00	1.00
5	1.00	1.00	1.00

Berdasarkan Tabel 1 pula, jika besar pergeseran vektor rata-rata adalah $\delta = 0,5$ dengan banyak karakteristik kualitas, p , misal: $p = 4$, diketahui nilai ARL_I bagan kendali \tilde{F}_m adalah 101. Artinya, ketika vektor rata-rata digeser sebesar $\delta = 0,5$ dari vektor rata-rata spesifikasi, bagan kendali \tilde{F}_m memberikan sinyal bahwa telah terjadi pergeseran vektor rata-rata untuk yang pertama kali adalah pada amatan ke-101.



Gambar 1. Plot Simulasi untuk $p = 4, m = 20$ dan $\delta = 1$ Pada Kondisi Kedua Parameter Diketahui

Gambar 1 adalah plot satu kali perulangan simulasi. Data pada Gambar 1 dibangkitkan secara random dengan $\mathbf{x}_f \square N_4(\boldsymbol{\mu}_s = \boldsymbol{\mu}_0 + \delta, \mathbf{I}_4)$. Besar pergeseran dari Gambar 1 adalah $\delta = 1$, sehingga $\boldsymbol{\mu}_s = \boldsymbol{\mu}_0 + 1 = \mathbf{1}$. Batas kendali atas dari bagan kendali \tilde{F}_m adalah $UCL = \chi^2_{0,0027,4} = 16,2512$. Berdasarkan Gambar 1, dapat diketahui bahwa nilai RL_1 untu bagan kendali \tilde{F}_m adalah 11. Artinya, titik pertama kali yang terdeteksi ketika telah terjadi pergeseran vektor rata-rata sebesar $\mathbf{1}$ adalah titik ke-11.

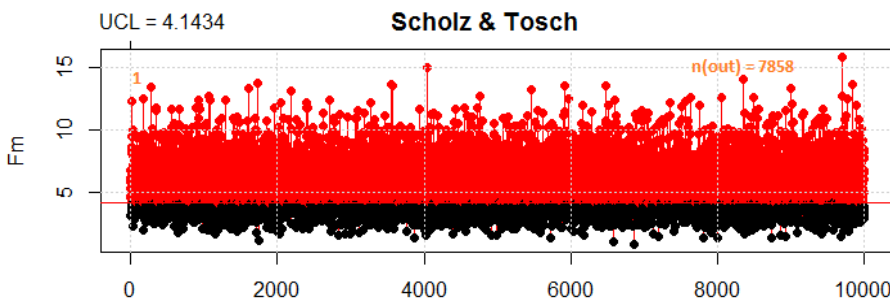
Tabel 2 merupakan tabel yang berisi nilai ARL bagan kendali \tilde{F}_m untuk kondisi kedua parameter tidak diketahui. Berdasarkan Tabel 2, dapat diketahui bahwa untuk $p = 2$, tanpa ada pergeseran vektor rata-rata, nilai ARL_0 untuk bagan kendali \tilde{F}_m adalah 1593. Artinya, ketika vektor rata-rata tidak bergeser dari vektor rata-rata target, bagan kendali ini mendeteksi adanya titik yang keluar batas kendali untuk yang pertama kali (RL_1) adalah pada amatan ke-1593.

Tabel 2. Nilai ARL Kondisi Kedua Parameter Diketahui dengan $m = 20$

δ	$p = 2$	$p = 4$	$p = 8$
	\tilde{F}_m	\tilde{F}_m	\tilde{F}_m
0	1593.59	1577.66	1419.89
0.5	499.78	412.78	256.30
1	82.42	56.60	40.44
1.5	17.76	14.58	10.82
2	5.50	3.45	2.90
2.5	2.41	1.39	1.15
3	1.46	1.25	1.10
4	1.03	1.02	1.00
5	1.00	1.00	1.00

Dibandingkan dengan kondisi parameter diketahui, ketika parameter tidak diketahui sensitifitas bagan kendali \tilde{F}_m menurun. Hal ini ditunjukkan dengan nilai ARL pada parameter tidak diketahui yang jauh lebih besar dibandingkan dengan nilai ARL pada parameter diketahui. Sensitifitas bagan kendali akan sama ketika pergeseran vektor rata-rata pada nilai $\delta > 2$.

Gambar 2 merupakan plot bagan kendali hasil simulasi. Data pada Gambar 2 dibangkitkan secara random dengan $\mathbf{x}_f \square N_8(\boldsymbol{\mu}_s = \boldsymbol{\mu}_0 + \delta, \mathbf{I}_8)$. Besar pergeseran dari Gambar 2 adalah $\delta = 2$, sehingga $\boldsymbol{\mu}_s = \boldsymbol{\mu}_0 + 2 = \mathbf{2}$. Batas kendali atas dari bagan kendali \tilde{F}_m adalah $UCL = F_{0,0027,8,33} = 4,1434$.

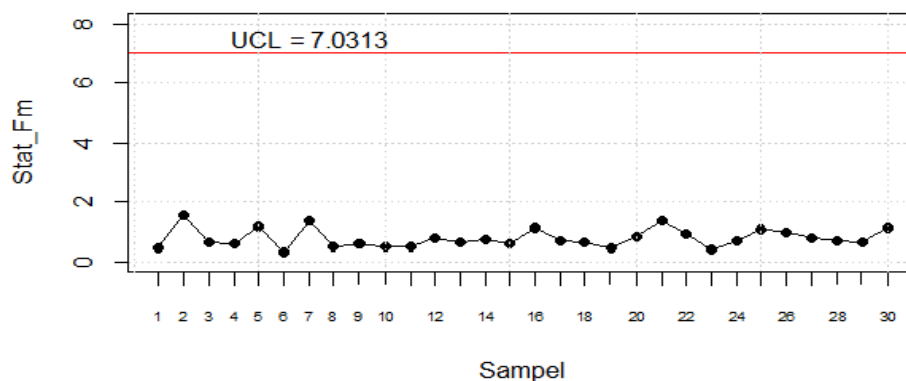


Gambar 2. Plot Simulasi untuk $p = 8, m = 50$ dan $\delta = 2$ Pada Kondisi Kedua Parameter Tidak Diketahui

Berdasarkan Gambar 2 dapat diketahui bahwa nilai RL_1 dari kedua bagan kendali adalah sama yakni 1. Artinya titik pertama kali yang keluar dari batas kendali dari kedua bagan kendali adalah titik amatan ke-1.

Dapat diartikan juga bahwa ketika vektor rata-rata digeser menjauh sebesar $\delta = 2$, kedua bagan kendali tersebut dapat dengan segera pada titik pertama mendeteksi bahwa proses memang benar telah terjadi pergeseran. Gambar 2 juga menunjukkan bahwa sekitar 80% titik amatan yang keluar dari batas kendali. Artinya, semakin besar pergeseran vektor rata-rata, maka semakin banyak titik amatan yang keluar dari batas kendali.

Sebagai terapan, berikut adalah hasil analisis data Quesenberry (2001) [9]. Data terdiri atas 11 karakteristik kualitas ($p = 11$) dengan banyak amatan $m = 30$. Berdasarkan uji Mardia, dinyatakan bahwa data tersebut berdistribusi normal multivariat. Kemudian, dengan menggunakan Persamaan (2) diperoleh nilai statistik untuk bagan kendali dan diplotkan sebagaimana pada Gambar 3.



Gambar 3. Bagan Kendali \tilde{F}_m Untuk Data Terapan Quesenberry (2001)

4. Simpulan

Ada beberapa kesimpulan yang dapat diambil berdasarkan hasil analisis yang telah dilakukan yaitu pada kondisi kedua parameter diketahui, konsep SD pada bagan kendali \tilde{F}_m hanya digunakan untuk menghitung \tilde{S}_m , \bar{x}_m melibatkan semua titik amatan. Akibat dari penghitungan \bar{x}_m yang melibatkan semua titik amatan, maka bagan kendali \tilde{F}_m tidak dapat segera terbentuk. Adapun pada kondisi kedua parameter diketahui, bagan kendali \tilde{F}_m mempunyai sensitifitas lebih tinggi dibandingkan dengan kondisi kedua parameter tidak diketahui. Selain itu, berdasarkan evaluasi kinerja kedua bagan kendali, semakin bertambahnya karakteristik kualitas (p) yang diamati, maka semakin kecil nilai ARL_1 , namun semakin banyak titik amatan yang diamati, maka nilai ARL_1 semakin besar. Kesimpulan lain yang dapat diambil adalah semakin jauh pergeseran vektor rata-rata proses, maka kedua bagan kendali ini semakin cepat mendeteksi.

Adapun saran yang dapat diberikan adalah perlu dikembangkan rumusan baru dengan menggunakan konsep SD untuk estimasi kedua parameter pada bagan kendali \tilde{F}_m , sehingga bagan kendali dapat segera terbentuk setelah didapat beberapa data amatan. Pola pergeseran vektor rata-rata yang disimulasikan pada penelitian ini bersifat simultan, agar lebih komprehensif, perlu dilakukan pola pergeseran vektor rata-rata secara parsial. Pada penelitian ini yang dikaji adalah hanya bagan kendali pergeseran rata-rata proses, sehingga perlu dikaji lebih lanjut bagan kendali untuk dispersi.

Daftar Pustaka

- Elam, M. E., dan Case, K. E. 2005. Two-Stage Short-Run Control Charts. *Quality Engineering*, 17, hal. 95-107.
- Fonseca, D. J., M. E. Elam, dan L. Tibbs. 2007. Fuzzy Short-Run Control Charts. *Mathware & Soft Computing*, 14, hal. 81-101.
- Jaupi, L., D. E. Herwindiati, Ph. Durand dan D. Ghorbanzadeh. 2013. Short-Run Multivariate Control Charts for Process Mean and Variability. *Proceeding of the World Congress on Engineering*, London.
- Khoo, M. B. C., dan S. H. Quah. 2002. Proposed Short Run Multivariate Control Chart for The Proses Mean. *Quality Engineering*, 14, hal. 603-621.
- Khoo, M. B. C., S. H. Quah, H. C. Low, dan C. K. Ch'Ng. 2005. Short Runs Multivariate Control Chart for Process Dispersion. *International Journal of Reliability, Quality and Safety Engineering*, 12, hal. 127-147.

Marques, P. A., Carlos B. C., Paula P., Sousa R., dan Helena G. 2015. Selection of The Most Suitable Statistical Process Control Approach for Short Production Runs: A Decision-Model. *International Journal of Information and Education Technology*, 5, hal. 303-310.

Montgomery, D. C. 2009. *Introduction to Statistical Quality Control*. Edisi ke-5. John Wiley & Sons: New York.

Quesenberry, C. P. 2001. The Multivariate Short-Run Snapshot Q Chart. *Quality Engineering*, 13, hal. 679-683.

Scholz, F. W. dan T. J. Tosch. 1994. Small Sample Uni- and Multivariate Control Charts for Means. *Proceedings of The American Statistical Association, Quality and Production Section*.

Apendiks

KEDUA PARAMETER DIKETAHUI

Dimisalkan bahwa $\{x_1, x_2, \dots, x_m\}$; $x_i \square N_p(\mu, \Sigma)$. Jika $x_i \square N_p(\mu, \Sigma)$, maka $(x_i - \mu) \square N_p(0, \Sigma)$.

Definisi 1.

Jika x_1, x_2, \dots, x_m berdistribusi normal p-variati, identik dan independen, $N_p(0, \Sigma)$; maka $w = \sum_{i=1}^m x_i x_i'$ akan berdistribusi Wishart dengan derajat bebas m . (Seber, 2004)

Berdasarkan Definisi 1, $\sum_{i=1}^m (x_i - \mu_0)(x_i - \mu_0)' \square W_p(m, \Sigma); (: \mu = \mu_0)$.

Jika $cc' = \Sigma$, di mana c adalah matriks non-singular berukuran $p \times p$, maka $U = c^{-1}(x_i - \mu_0) \square N_p(0, I_p)$ dan $U'U = (x_i - \mu_0)' \Sigma^{-1} (x_i - \mu_0)$. Perkalian matriks $U'U$ merupakan jumlah kuadrat dari p variabel saling bebas yang masing-masing berdistribusi $N(0,1)$, sehingga $U'U \square \chi_p^2$. Jika $\Sigma = \Sigma_0$, maka benar bahwa $\tilde{F}_i = (x_i - \mu_0)' \Sigma_0^{-1} (x_i - \mu_0) \square \chi_p^2$.

KEDUA PARAMETER TIDAK DIKETAHUI

Dimisalkan bahwa $\{x_1, x_2, \dots, x_m\}$; $x_i \square N_p(\mu, \Sigma)$. Dinyatakan pula bahwa $y_i = x_{i+1} - x_i$; $i = 1, \dots, m-1$ disebut vektor selisih lokal (*local difference vector*).

$$\begin{pmatrix} y'_1 \\ y'_2 \\ y'_3 \\ \vdots \\ y'_{m-1} \end{pmatrix} = \begin{pmatrix} x'_2 - x'_1 \\ x'_3 - x'_2 \\ x'_4 - x'_3 \\ \vdots \\ x'_m - x'_{m-1} \end{pmatrix}$$

$$\begin{pmatrix} y'_1 \\ y'_2 \\ y'_3 \\ \vdots \\ y'_{m-1} \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & -1 & 1 & \dots & 0 & 0 \\ \dots & & & & \dots & & \\ 0 & 0 & 0 & 0 & \dots & -1 & 1 \end{pmatrix} \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ \vdots \\ x'_m \end{pmatrix}$$

Y = DX.

di mana D disebut sebagai matriks pembeda (*differencing matrix*). Diketahui bahwa $s_i = \frac{1}{2} y_i y_i'$ dan estimator tak bias gabungan (\tilde{S}_m) untuk Σ adalah

$$\tilde{S}_m = \frac{1}{(m-1)} \sum_{i=1}^{m-1} S_i = \frac{1}{2(m-1)} \sum_{i=1}^{m-1} y_i y_i'$$

$$= \frac{1}{2(m-1)} Y'Y.$$

Jika $A = \frac{1}{2} D'D$, maka

$$\tilde{S}_m = \frac{1}{2(m-1)} Y'Y = \frac{1}{m-1} X'AX.$$

Teorema 1.

Diketahui bahwa $X' = (x_1, x_2, \dots, x_m)$ di mana $x_i \square N_p(0, \Sigma)$ dan $v = X\xi$, $\xi (\neq 0)$ adalah vektor konstan berukuran $p \times 1$. Jika A adalah suatu matriks berukuran $m \times m$ dengan rank r , maka $X'AX \square W_p(r, \Sigma)$ jika dan hanya jika $v'Av \square \sigma_\xi^2 \chi_r^2$ untuk setiap ξ , di mana $\sigma_\xi^2 = \xi' \Sigma \xi$. (Seber, 2004)

Teorema 2.

Jika $w \square W_p(m, \Sigma)$ dan c adalah matriks berukuran $r \times p$ dengan rank r , maka $c'wc \square W_p(m, c\Sigma c')$.

Akibat Teorema 2.

Jika ξ adalah sembarang vektor konstan berukuran $p \times 1$, ($\xi \neq \mathbf{0}$), maka $\xi' \mathbf{W} \xi \sim \sigma_{\xi}^2 \chi_m^2$ di mana $\sigma_{\xi}^2 = \xi' \Sigma \xi > 0$ karena $\Sigma > \mathbf{O}$.

Berdasarkan Teorema 1, diketahui bahwa perkalian beberapa $\tilde{\mathbf{S}}_m$ mempunyai distribusi yang dapat didekati dengan distribusi Wishart, $W_p(d, \Sigma)$.

Jika diasumsikan $\mu = \mathbf{0}$, maka $v_i = \mathbf{x}_i' \xi \sim N(0, \sigma_{\xi}^2)$ dan $\mathbf{v}' = (v_1, v_2, \dots, v_m)$. Dengan persamaan pendekatan seperti berikut:

$\frac{1}{\eta} \mathbf{v}' \mathbf{A} \mathbf{v} \approx \sigma_{\xi}^2 \chi_d^2$. Maka ekspektasi dan varians dari kedua sisi adalah

$$E\left(\frac{1}{\eta} \mathbf{v}' \mathbf{A} \mathbf{v}\right) = \frac{m-1}{\eta} \sigma_{\xi}^2 \quad \text{dan} \quad \text{var}\left(\frac{1}{\eta} \mathbf{v}' \mathbf{A} \mathbf{v}\right) = \frac{\sigma_{\xi}^4}{\eta^2} (3m-4).$$

$$E(\sigma_{\xi}^2 \chi_d^2) = d \sigma_{\xi}^2 \quad \text{dan} \quad \text{var}(\sigma_{\xi}^2 \chi_d^2) = 2d \sigma_{\xi}^4.$$

Selanjutnya menyamakan hasil ekspektasi dan varians untuk masing-masing sisi.

$\frac{m-1}{\eta} \sigma_{\xi}^2 = d \sigma_{\xi}^2$ dan $\frac{\sigma_{\xi}^4}{\eta^2} (3m-4) = 2d \sigma_{\xi}^4$, sehingga diperoleh $\eta = \frac{3m-4}{2(m-1)}$ dan $d = \frac{2(m-1)^2}{3m-4}$. Dikarenakan η dan d tidak

tergantung pada ξ , maka dapat dinyatakan bahwa $\frac{(m-1)\tilde{\mathbf{S}}_m}{\eta} = d\tilde{\mathbf{S}}_m \approx W_p(d, \Sigma)$.

Dinyatakan bahwa $\bar{\mathbf{x}}_m = \frac{1}{m} \sum_{i=1}^m \mathbf{x}_i \sim N_p\left(\mu, \frac{\Sigma}{n}\right)$ independen terhadap $\tilde{\mathbf{S}}_m$. Lebih lanjut, jika $\mathbf{x}_f \sim N_p(\mu, \Sigma)$ adalah

amatan selanjutnya (*future observation*), maka $(\mathbf{x}_f - \bar{\mathbf{x}}_m) \sim N_p\left(\mathbf{0}, \Sigma\left(1 + \frac{1}{m}\right)\right)$ juga independen terhadap $\tilde{\mathbf{S}}_m$. Oleh

karena itu, dapat dinyatakan benar bahwa $\tilde{F}_m = \frac{d-p+1}{d \cdot p} \frac{m}{m+1} (\mathbf{x}_f - \bar{\mathbf{x}}_m)' \tilde{\mathbf{S}}_m^{-1} (\mathbf{x}_f - \bar{\mathbf{x}}_m)$ mengikuti distribusi F dengan

numerator p dan denominator $d-p+1$. Dengan demikian dapat ditulis:

$$\tilde{F}_m = \frac{d-p+1}{d \cdot p} \frac{m}{m+1} (\mathbf{x}_f - \bar{\mathbf{x}}_m)' \tilde{\mathbf{S}}_m^{-1} (\mathbf{x}_f - \bar{\mathbf{x}}_m) \sim F_{p, d-p+1}.$$