UNIVERSITAS INDONESIA

BAGAN KENDALI BERDASARKAN MEDIAN DAN *MEDIAN ABSOLUTE DEVIATION*

SKRIPSI

PUTRI MARLINA
0706261833

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
PROGRAM STUDI SARJANA MATEMATIKA
DEPOK
JUNI 2012

UNIVERSITAS INDONESIA

BAGAN KENDALI BERDASARKAN MEDIAN DAN *MEDIAN ABSOLUTE DEVIATION*

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana sains

PUTRI MARLINA

0706261833

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

PROGRAM STUDI SARJANA MATEMATIKA

DEPOK

JUNI 2012

HALAMAN PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya saya sendiri, dan semua
sumber yang dikutip maupun dirujuk telah saya
nyatakan dengan benar.

Nama	: Putri Marlina
NPM	: 0706261833
Tanda Tangan	:
Tanggal	: 15 Juni 2012

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh

Nama	: Putri Marlina
NPM	: 0706261833
Program Studi	: Matematika
Judul Skripsi	: Bagan Kendali Berdasarkan Median dan *Median Absolute Deviation*

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Indonesia.

DEWAN PENGUJI

Pembimbing	: Dra. Saskya Mary S., M.Si	( )
Penguji	: Dra. Rianti Setiadi M.Si	( )
Penguji	: Dra. Ida Fithriani M.Si	( )
Penguji	: Mila Novita S.Si., M.Si	( )

Ditetapkan di : Depok

Tanggal : 15 Juni 2012

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah Swt., karena atas rahmat dan hidayah-Nya, penulis dapat menyelesaikan skripsi ini. Penulisan skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Sains Program Studi Matematika pada Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Indonesia.

Penulis menyadari bahwa tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyususnan skripsi ini, sangatlah sulit bagi penulis untuk menyelesaikan skripsi ini. Oleh karena itu, penulis mengucapkan terima kasih kepada berbagai pihak antara lain:

1. Pembimbing tugas akhir penulis, Dra. Saskya Mary S., M.Si, yang telah menyediakan waktu, tenaga dan kasih sayangnya untuk mengarahkan, membimbing dan mengajarkan penulis berbagai hal dalam penyusunan skripsi ini.
2. Dra. Ida Fithriani, M.S., selaku pembimbing akademis penulis yang telah memberikan dukungan, arahan dan motivasi untuk penulis selama menjalani masa perkuliahan.
3. Dr. Yudi Satria dan Rahmi Rusin, M.ScTech, selaku ketua dan sekretaris Departemen Matematika. Terima kasih atas segala bantuan dan dukungan yang telah diberikan.
4. Seluruh staf Tata Usaha, staf Perpustakaan, serta karyawan Departemen Matematika, terima kasih atas bantuannya.
5. Kedua orang tua penulis, Bpk. Mardiman (Alm.) dan Ibunda tersayang Murniati, yang senantiasa memberikan kasih sayang yang begitu tulus, bantuan materi dan doa serta motivasi yang tiada henti sehingga penulis mampu menyelesaikan pendidikan S1 ini.
6. Adik-adik penulis, De’ Febi dan De’ Dani, terima kasih atas semangat dan doanya ya. Semoga kelak, kalian bisa meraihnya juga, amiin.

7. Teman-teman mahasiswa Matematika UI angkatan 2007 : Winda, Safa, Misda, Sica, Gamar, Anjar, Isna, Stefi, Wiwi, Mita, Siti, Afni, Nene, Shafir, Widya, Amanda, Hikmah, Widi, Dita, Lois, Farah, Nora, Siska, Anis, Arip, Ashari, Adi, Zul, Bowo, Adit, Andi, Fauzan, Ferdi, Syahrul, Iki, Toto, Yos, Putu, Anggun Danar dan Hanif. Terima kasih ya.. telah berbagi kisah dan memotivasi penulis selama kuliah dan menyelesaikan skripsi ini.

8. Teman-teman mahasiswa Matematika UI lainnya, seperti K'Ajat dan K'Nola (2004), K'Dia, K'Desti, K'Nurma, K'Amri dan K'Sally (2005), K'Alfa, K'Dian, K'Farah, K'Indah, K'Lena, K'Rizqiyatul, K'Nurgi, K'Opie, K'Stefani, K'Rafli, K'Tika, K'Tino, K'Yuko, K'Rifza, K'Rita Prihatiningsih (2006), Lutfa, Dea, Janu, Uci, Vika, Shita, Nora, Mei, Ica, Fani (2008), Eja, Azki, Emil, Sofi (2009) dan lainnya. Terima kasih atas doa dan semangat yang diberikan kepada penulis.

9. Teman-teman se-MIPA, se-kosan Ar-Rizal dan se-UI, seperti Yuli (Fisika'07) dan Muhtar (Kimia'07) : Semangat! Insya Allah kita pasti bisa! Juga teman-teman BPDU MII dan RUMBEL UI. Terimakasih atas semangat dan pengalaman-pengalaman berharga yang telah diberikan kepada penulis.

10. Teman-teman alumni IPA 5 SMAN 4 Bekasi (Imam, Osha, Deni, Dito, Ucup, Jihan, Kiki: terimakasih untuk S2Club-nya sebelum SPMB dulu ☺), BTA Depok (K'Rahmat, K'Nike, K'Pipit, dll: senangnya bisa berbagi ilmu, dapet duit jajan lagi, hhe.. ☺) dan juga tempat saya biasa ngeprint di kampus (Mba Dede dan Mba Farida: terimakasih atas bantuannya ya mba.. ☺).

Akhir kata, penulis berharap semoga Allah Swt. berkenan membalas semua kebaikan dari semua pihak yang telah membantu. Semoga skripsi ini membawa manfaat bagi perkembangan ilmu pengetahuan.

Penulis

2012

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika Universitas Indonesia, penulis yang bertanda tangan di bawah ini:

Nama	: Putri Marlina
NPM	: 0706261833
Program Studi	: Sarjana
Departemen	: Matematika
Fakultas	: Matematika dan Ilmu Pengetahuan Alam
Jenis karya	: Skripsi

demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Indonesia Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah penulis yang berjudul:

Bagan Kendali Berdasarkan Median dan *Median Absolute Deviation*

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Indonesia berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat dan memublikasikan tugas akhir penulis selama tetap mencantumkan nama penulis sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini penulis buat dengan sebenarnya.

Dibuat di: Depok
Pada tanggal: 15 Juni 2012
Yang menyatakan

(Putri Marlina)

ABSTRAK

Nama : Putri Marlina
Program Studi : Matematika
Judul : Bagan Kendali Berdasarkan Median dan *Median Absolute Deviation*

Bagan kendali berdasarkan median dan *median absolute deviation* merupakan modifikasi dari bagan kendali Shewhart, ketika ada sedikit penyimpangan asumsi kenormalan pada sampel. Modifikasi ini dilakukan dengan menggunakan penaksir-penaksir *robust*, median sampel dan *median absolute deviation* sampel, dalam pembuatan batas-batas kendali pada bagan kendali yang akan digunakan. Median sampel digunakan sebagai penaksir untuk mean proses, dan *median absolute deviation* sampel digunakan sebagai penaksir untuk standar deviasi proses. Bagan kendali berdasarkan median dan *median absolute deviation* ini terdiri dari dua bagan, yaitu bagan kendali $\bar{X}_{MDMAD}$ untuk mengawasi mean proses dan bagan kendali S_{MDMAD} untuk mengawasi standar deviasi proses. Sebagai contoh penerapan, studi ini menggunakan data sampel berdistribusi normal standar. Dari contoh penerapan tersebut diperoleh hasil bahwa bagan kendali berdasarkan median dan *median absolute deviation* dapat digunakan sebagai alternatif dari bagan kendali Shewhart, ketika sampel yang digunakan telah terkontaminasi *outlier* yang bukan merupakan kesalahan atau dengan kata lain terdapat penyimpangan asumsi kenormalan pada sampel.

Kata Kunci : bagan kendali, median, *median absolute deviation*, *robust*, *influence function*, *M-estimator*
xiv+63 halaman ; 6 gambar; 5 tabel
Daftar Pustaka : 13 (1978-2008)

ABSTRACT

Name : Putri Marlina
Study Program: Mathematics
Title : Control Chart Based On Median and Median Absolute Deviation

Control chart based on the median and median absolute deviation is a modification of the Shewhart control chart, when there are slight deviations assuming normality in the sample. This modification is done by using robust estimators, the sample median and median absolute deviation of samples, in the constructing of the control limits on control charts to be used. Sample median is used as an estimator for the process mean and sample median absolute deviation is used as an estimator for the standard deviation of the process. Control chart based on median and median absolute deviation consists of two charts, the $\bar{X}_{MDMAD}$control chart to monitor the process mean and the S_{MDMAD} control chart to monitor standard deviation of the process. As an example of the application, this study uses the standard normal distribution of sample data. It showed that the control chart based on the median and median absolute deviation can be used as an alternative to the Shewhart control chart, when the sample used was contaminated with outliers that are not a mistake or in other words, there are deviations assuming normality in the sample.

Keywords : control chart, median, median absolute deviation, robust, influence function, M-estimator
xiv+63 pages ; 6 pictures; 5 tables
Bibliography : 13 (1978-2008)

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

DAFTAR LAMPIRAN

BAB 1

PENDAHULUAN

1.1 Latar Belakang Masalah

Perbaikan mutu didefinisikan sebagai pengurangan variabilitas dalam proses produksi dan produk yang dihasilkan (Montgomery: 2005). Variabilitas yang besar dalam suatu proses produksi sering menghasilkan produk yang tidak sesuai dengan spesifikasi, sehingga dapat menimbulkan kerugian bagi perusahaan baik dari segi biaya, waktu maupun tenaga ekstra sebagai konsekuensi dari perbaikan ulang produk. Oleh karena itu, perlu adanya pengendalian mutu di dalam sebuah proses produksi.

Bagan kendali adalah salah satu teknik pengendalian mutu yang sering digunakan. Bagan ini memplot rata-rata pengukuran suatu karakteristik mutu sampel yang diambil dari proses produksi berdasarkan urutan proses produksi. Sampel yang diambil berdasarkan urutan proses produksi tersebut harus rasional mewakili proses produksi, yang selanjutnya disebut dengan subgrup. Dari plot rata-rata karakteristik mutu tersebut variabilitas mutu produksi dapat dikendalikan.

Bagan kendali mempunyai sebuah *center line* (CL), *upper control limit* (UCL) dan *lower control limit* (LCL). *Center line* merepresentasikan nilai dimana seharusnya rata-rata karakteristik mutu proses berada, jika tidak terdapat sumber variabilitas yang tidak biasa. Dengan demikian, bagan kendali merupakan sebuah teknik pemantauan rata-rata proses yang sangat berguna karena keberadaan sumber variabilitas dapat terdeteksi langsung dengan adanya sinyal berupa rata-rata karakterisitik mutu sampel yang terplot berada di luar batas-batas kendali.

Walter A. Shewhart telah memperkenalkan sebuah konsep bagan kendali statistika yang dapat membuat proses pengendalian mutu berjalan efektif, yang dikenal dengan bagan kendali Shewhart. Bagan ini dibuat dengan asumsi bahwa sampel yang diambil dari sebuah proses produksi mempunyai distribusi normal. Jika asumsi kenormalan ini terpenuhi, maka mean sampel $(\bar{X})$ dapat digunakan sebagai penaksir untuk mean proses (μ) dan standar deviasi sampel (S) digunakan sebagai penaksir untuk standar deviasi proses (σ).

Namun, jika terdapat *outlier* pada sampel dimana *outlier* tersebut bukan merupakan kesalahan sehingga akan mengakibatkan terjadinya penyimpangan pada asumsi distribusi normal, maka mean dan standar deviasi sampel tidak lagi menjadi penaksir yang baik untuk mean dan standar deviasi proses. Oleh karena itu, jika asumsi kenormalan sedikit menyimpang, maka diperlukan bagan kendali alternatif dengan penaksir-penaksir yang lebih *robust*.

Pada tugas akhir ini, akan dibahas tentang bagan kendali yang dapat digunakan sebagai alternatif untuk bagan kendali Shewhart ketika asumsi kenormalan pada sampel menyimpang, yaitu bagan kendali median dan *median absolute deviation* (MDMAD). Bagan kendali MDMAD adalah modifikasi dari bagan kendali Shewhart yang dapat digunakan untuk mengawasi mean dan standar deviasi proses. Bagan kendali MDMAD menggunakan median sampel sebagai penaksir yang *robust* untuk mean proses dan menggunakan *median absolute deviation* sampel sebagai penaksir yang *robust* untuk standar deviasi proses. Dengan menggunakan kedua penaksir *robust* ini, akan dibuat suatu bagan kendali alternatif untuk bagan kendali Shewhart, yaitu bagan kendali $\bar{X}_{MDMAD}$ untuk mengawasi mean proses dan bagan kendali S_{MDMAD} untuk mengawasi standar deviasi proses.

1.2 Perumusan Masalah dan Ruang Lingkup Masalah

Perumusan masalah yang diajukan pada tugas akhir ini adalah:

1. Bagaimana cara menentukan penaksir *robust* untuk mean proses dan standar deviasi proses?
2. Bagaimana cara membuat bagan kendali $\bar{X}_{MDMAD}$ untuk mengawasi mean proses?
3. Bagaimana cara membuat bagan kendali S_{MDMAD} untuk mengawasi standar deviasi proses?

Ruang lingkup masalah yang digunakan dalam tugas akhir ini dibatasi pada karakteristik mutu yang digunakan hanya terdiri dari satu peubah.

1.3 Jenis Penelitian dan Metode yang Digunakan

Jenis penelitian yang digunakan dalam pembuatan tugas akhir ini adalah studi literatur. Sedangkan metode yang digunakan untuk menaksir mean dan standar deviasi proses pada bagan kendali median dan *median absolute deviation* (MDMAD) adalah metode *robust*.

1.4 Tujuan Penulisan

Tujuan penulisan tugas akhir ini adalah:

1. Menjelaskan prosedur untuk mendapatkan penaksir *robust* untuk mean proses dan standar deviasi proses
2. Menjelaskan prosedur pembuatan bagan kendali $\bar{X}_{MDMAD}$ untuk mengawasi mean proses
3. Menjelaskan prosedur pembuatan bagan kendali S_{MDMAD} untuk mengawasi standar deviasi proses

BAB 2
LANDASAN TEORI

2.1 Percobaan Acak dan Ruang Sampel

Misalkan terdapat suatu percobaan, hasilnya tidak dapat diprediksi dengan pasti. Percobaan acak adalah suatu percobaan yang dilakukan berulang kali dan di dalam kondisi yang sama. Ruang sampel adalah koleksi dari semua hasil yang mungkin dari suatu percobaan acak.

2.2 Fungsi Himpunan Probabilitas

Misalkan $\mathcal{C}$ menyatakan himpunan dari semua hasil yang mungkin dari suatu percobaan acak, atau disebut ruang sampel. Berikut akan didefinisikan fungsi himpunan P sedemikian sehingga jika $C \subset \mathcal{C}$, maka $P(C)$ menyatakan probabilitas bahwa hasil dari suatu percobaan acak berada di C. Dengan demikian maka fungsi himpunan P adalah suatu fungsi yang didefinisikan dari power set $\mathcal{C}$ (notasi $P^{\mathcal{C}}$) ke interval [0,1] atau dapat ditulis

$$P: P^{\mathcal{C}} \rightarrow [0,1]$$
$$C \mapsto P(C).$$

Apabila fungsi P di atas memenuhi sifat-sifat berikut ini:

a) $P(C) \geq 0$

b) $P(C_1 \cup C_2 \cup C_3 \cup \ldots) = P(C_1) + P(C_2) + P(C_3) + \cdots$, dimana himpunan-himpunan C_i, $i=1, 2, 3, \ldots$, adalah saling lepas, yaitu $C_i \cap C_j = \emptyset, i \neq j$

c) $P(\mathcal{C}) = 1$,

maka P disebut fungsi himpunan probabilitas.

Untuk setiap subhimpunan C dari $\mathcal{C}$, bilangan $P(C)$ disebut probabilitas bahwa hasil dari percobaan acak adalah elemen dari himpunan C, atau probabilitas kejadian C.

Suatu fungsi himpunan probabilitas memberitahukan bagaimana probabilitas didistribusikan terhadap berbagai subhimpunan C dari suatu ruang sampel $\mathcal{C}$. Dalam hal ini disebut distribusi probabilitas.

Beberapa sifat dari suatu fungsi himpunan probabilitas adalah:

1. Untuk setiap $C \subset \mathcal{C}$, $P(C) = 1 - P(C^*)$, $C^* \cup C = \mathcal{C}$
2. Probabilitas dari himpunan kosong adalah nol, yaitu $P(\emptyset) = 0$
3. Jika C_1 dan C_2 adalah subhimpunan-subhimpunan dari $\mathcal{C}$ sedemikian sehingga $C_1 \subset C_2$, maka $P(C_1) \leq P(C_2)$
4. Untuk setiap $C \subset \mathcal{C}$, $0 \leq P(C) \leq 1$
5. Jika C_1 dan C_2 adalah subhimpunan-subhimpunan dari $\mathcal{C}$, maka

$$P(C_1 \cup C_2) = P(C_1) + P(C_2) - P(C_1 \cap C_2)$$

2.3 Probabilitas Bersyarat

Pada beberapa percobaan acak, misalkan ingin diselidiki hasil-hasil yang merupakan elemen-elemen dari suatu subset C_1, dimana C_1 subset dari $\mathcal{C}$. Ini berarti bahwa ruang sampel yang efektif adalah C_1. Untuk selanjutnya, akan didefinisikan suatu fungsi himpunan probabilitas dengan C_1 sebagai ruang sampel baru.

Misalkan fungsi himpunan probabilitas $P(C)$ didefinisikan pada ruang sampel $\mathcal{C}$ dan misalkan C_1 adalah subset dari $\mathcal{C}$ sedemikian sehingga $P(C_1) > 0$. Karena yang ingin diperhatikan selanjutnya adalah hasil percobaan acak yang merupakan elemen dari C_1, berarti disini C_1 adalah ruang sampel baru. Ambil C_2 adalah subset lain dari $\mathcal{C}$. Relatif terhadap ruang sampel baru, akan didefinisikan probabilitas dari kejadian C_2. Probabilitas ini disebut probabilitas bersyarat dari C_2 relatif terhadap kejadian C_1, atau probabilitas bersyarat dari C_2 diberikan C_1, yang dinotasikan sebagai $P(C_2|C_1)$. Karena C_1 sekarang adalah ruang sampel yang baru, maka elemen-elemen C_2 yang berhubungan dengan ini hanya elemen-elemen C_2 yang juga merupakan elemen-elemen dari C_1, yaitu elemen-elemen dari $C_1 \cap C_2$.

$P(C_2|C_1)$ didefinisikan sedemikian sehingga $P(C_1|C_1) = 1$ dan $P(C_2|C_1) = P(C_1 \cap C_2|C_1)$. Dalam hal ini

$$\frac{P(C_1 \cap C_2|C_1)}{P(C_1|C_1)} = \frac{P(C_1 \cap C_2)}{P(C_1)}$$

Berarti

$$P(C_2|C_1) = \frac{P(C_1 \cap C_2)}{P(C_1)}$$

Ini merupakan definisi dari probabilitas bersyarat kejadian C_2 diberikan C_1, dengan syarat $P(C_1) > 0$.

Beberapa sifat dari probabilitas bersyarat adalah:

1. $P(C_2|C_1) \geq 0$
2. $P(C_2 \cup C_3 \cup \ldots |C_1) = P(C_2|C_1) + P(C_3|C_1) + \cdots$,
   dimana C_2, C_3,… merupakan himpunan-himpunan yang saling lepas
3. $P(C_1|C_1) = 1$.

Teorema Bayes

Misalkan kejadian-kejadian $C_1, C_2, \ldots, C_k$ adalah subset-subset dari ruang sampel $\mathcal{C}$. Apabila $C_1, C_2, \ldots, C_k$ tidak saling beririsan, maka $C_1, C_2, \ldots, C_k$ disebut kejadian saling lepas atau *mutually exclusive events*. Jika berlaku

$P(C_1) + P(C_2) + \ldots + P(C_k) = 1$, maka $C_1, C_2, \ldots, C_k$ disebut *mutually exclusive events and exhaustive*.

Misalkan terdapat suatu kejadian C di ruang sampel $\mathcal{C}$, sedemikian sehingga

$C = (C \cap C_1) \cup (C \cap C_2) \cup \ldots \cup (C \cap C_k)$ dimana $(C \cap C_1), (C \cap C_2), \ldots, (C \cap C_k)$ saling lepas. Artinya, berlaku bahwa

$$P(C) = P(C \cap C_1) + P(C \cap C_2) + \cdots + P(C \cap C_k).$$

Dari penjelasan sebelumnya, telah diketahui bahwa $P(C \cap C_i) = P(C_i)P(C|C_i)$, $i = 1,2, \ldots, k$. Dengan demikian, maka $P(C)$ dapat ditulis menjadi

$$\begin{aligned} P(C) &= P(C_1)P(C|C_1) + P(C_2)P(C|C_2) + \cdots + P(C_k)P(C|C_k) \\ &= \sum_{i=1}^{k} P(C_i)P(C|C_i) \end{aligned}$$

Persamaan di atas disebut hukum probabilitas total.

Selanjutnya, misalkan $P(C) > 0$. Dari definisi probabilitas bersyarat dan dengan menggunakan hukum probabilitas total diperoleh

$$P(C_j|C) = \frac{P(C \cap C_j)}{P(C)} = \frac{P(C_j)P(C|C_j)}{\sum_{i=1}^{k} P(C_i)P(C|C_i)}$$

Persamaan di atas disebut dengan Teorema Bayes.

2.4 Peubah Acak

Berikut ini diberikan definisi dari peubah acak.

Definisi 2.2

Perhatikan suatu percobaan acak dengan ruang sampel $\mathcal{C}$. Suatu fungsi X, yang memetakan setiap elemen $c \in \mathcal{C}$ satu dan hanya satu bilangan riil $X(c) = x$, disebut peubah acak.

(Hogg and Craig: 1995)

Ruang nilai dari peubah acak X adalah himpunan dari bilangan-bilangan riil $\mathcal{A} = \{ x | x = X(c), c \in \mathcal{C}\}$.

2.5 Fungsi Kepadatan Probabilitas

Misalkan X menyatakan suatu peubah acak dengan ruang nilai satu dimensi $\mathcal{A}$. Misalkan $\mathcal{A}$ berisi nilai-nilai bilangan yang terhitung, maka ruang $\mathcal{A}$ disebut himpunan diskrit dari nilai-nilai. Tetapi sebaliknya, jika $\mathcal{A}$ berisi nilai-nilai bilangan yang tidak terhitung, maka ruang $\mathcal{A}$ disebut himpunan kontinu dari nilai-nilai.

Untuk kasus diskrit, misalkan X menyatakan peubah acak dengan ruang satu dimensi $\mathcal{A}$, yang memuat titik-titik bilangan yang terhitung. Misalkan $f(x)$ adalah suatu fungsi sedemikian sehingga $f(x) \geq 0$, $x \in \mathbb{R}$, dan $\Sigma_{\mathcal{A}} f(x) = 1$.

Ketika suatu fungsi himpunan probabilitas $P(A), A \subset \mathcal{A}$, dapat dinyatakan dalam bentuk

$$P(A) = Pr(X \in A) = \Sigma_A f(x), \tag{2.1}$$

maka X disebut sebagai peubah acak tipe diskrit dan $f(x)$ disebut sebagai fungsi kepadatan probabilitas (*probability density function (pdf)*) dari X.

Untuk kasus kontinu, misalkan X menyatakan peubah acak dengan ruang satu dimensi $\mathcal{A}$, yang memuat suatu interval atau gabungan dari interval-interval. Misalkan $f(x)$ adalah suatu fungsi sedemikian sehingga $f(x) \geq 0$, $x \in \mathbb{R}$, dan $\int_{\mathcal{A}} f(x)dx = 1$.

Ketika suatu fungsi himpunan probabilitas $P(A), A \subset \mathcal{A}$, dapat dinyatakan dalam bentuk

$$P(A) = Pr(X \in A) = \int_A f(x)dx, \qquad (2.2)$$

maka X disebut peubah acak tipe kontinu dan $f(x)$ disebut fungsi kepadatan probabilitas (*probability density function (pdf)*) dari X.

2.6 Fungsi Distribusi

Misalkan peubah acak X mempunyai suatu fungsi himpunan probabilitas $P(A)$, dimana A adalah himpunan satu dimensi.

Ambil bilangan riil x dan perhatikan himpunan A yang merupakan himpunan yang tidak terbatas dari $-\infty$ sampai x, termasuk titik x itu sendiri. Untuk setiap himpunan A yang demikian, diperoleh $P(A) = Pr(X \in A) = Pr(X \leq x)$. Probabilitas ini bergantung pada nilai x, yaitu, probabilitas ini adalah fungsi dari x. Fungsi nilai ini dinyatakan dengan

$$F(x) = Pr(X \leq x). \qquad (2.3)$$

Fungsi $F(x)$ dikenal dengan sebutan fungsi distribusi, atau fungsi distribusi kumulatif (*cumulative distribution function (cdf)*) dari peubah acak X.

Misalkan *pdf* dari X adalah $f(x)$ maka fungsi distribusi dari X adalah untuk X peubah acak diskrit:

$$F(x) = Pr(X \leq x) = \sum_{w \leq x} f(w). \qquad (2.4)$$

untuk X peubah acak kontinu:

$$F(x) = Pr(X \leq x) = \int_{-\infty}^{x} f(w)dw, \text{ sehingga } F'(x) = f(x). \qquad (2.5)$$

Berikut diberikan sifat-sifat dari suatu fungsi distribusi:

1. $0 \leq F(x) \leq 1$
2. *F(x)* merupakan fungsi tidak turun
3. $F(\infty) = 1$ dan $F(-\infty) = 0$
4. *F(x)* kontinu kanan.

2.7 Ekspektasi dari Peubah Acak

Misalkan *X* adalah suatu peubah acak yang mempunyai *pdf f(x)* sedemikian sehingga dimiliki kekonvergenan absolut, dalam kasus diskrit,

$\sum_x |x| f(x)$ konvergen ke suatu batas berhingga,

atau, dalam kasus kontinu,

$\int_{-\infty}^{\infty} |x| f(x) dx$ konvergen ke suatu batas berhingga.

Ekspektasi dari suatu peubah acak adalah

$E(X) = \sum_x x f(x)$, dalam kasus diskrit, atau (2.6)

$E(X) = \int_{-\infty}^{\infty} x f(x) dx$, dalam kasus kontinu. (2.7)

Ekspektasi *E(X)* disebut juga sebagai ekspektasi matematika dari *X* atau nilai harapan dari *X*.

Perhatikan suatu peubah acak *X* dengan ruang nilai $\mathscr{A}$. Misalkan fungsi dari peubah acak *X* adalah $Y = u(X)$. Misalkan *X* merupakan peubah acak yang bertipe kontinu dan $y = u(x)$ merupakan fungsi kontinu naik dari *X* dengan invers fungsinya $x = w(y)$, yang juga merupakan fungsi naik. Jadi, *Y* adalah suatu peubah acak dan fungsi distribusinya adalah

$$G(y) = Pr(Y \leq y) = Pr[u(X) \leq y] = Pr[X \leq w(y)] = \int_{-\infty}^{w(y)} f(x) dx$$

dimana *f(x)* adalah *pdf* dari *X*.

Dengan salah satu bentuk dari Teorema Dasar Kalkulus,

$$\begin{aligned} g(y) = G'(y) &= f[w(y)]w'(y), \quad y \in \mathscr{B} \qquad (2.8) \\ &= 0 \qquad\qquad\qquad\qquad , \text{lainnya}, \end{aligned}$$

dimana $\mathscr{B} = \{y \mid y = u(x), x \in \mathscr{A}\}$.

Dengan definisi, nilai harapan nilai harapan dari Y adalah

$$E(Y) = \int_{-\infty}^{\infty} yg(y)dy. \tag{2.9}$$

Dengan menggunakan teknik perubahan variabel dari integrasi melalui *y=u(x)* atau, secara ekivalen, *x=w(y)*. Karena

$$\frac{dx}{dy} = w'(y) > 0 \tag{2.10}$$

maka diperoleh

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} yg(y)dy \\ &= \int_{-\infty}^{\infty} u(x)g[u(x)]\frac{1}{w'[u(x)]}dx \\ &= \int_{-\infty}^{\infty} u(x)f(x)dx. \end{aligned} \tag{2.11}$$

Hal ini benar secara umum dan juga tidak ada perbedaan apakah *X* peubah acak bertipe diskrit atau kontinu dan *Y=u(X)* tidak harus merupakan fungsi naik dari *X*. Jadi, *Y = u(X)* mempunyai ekspektasi, dapat diperoleh dari (2.11) bahwa

$E[u(X)] = \int_{-\infty}^{\infty} u(x)f(x)dx,$ untuk kasus kontinu, dan (2.12)

$E[u(X)] = \sum_x u(x)f(x),$ untuk kasus diskrit. (2.13)

Berikut ini diberikan beberapa sifat dari ekspektasi matematika:

1. Jika k adalah sebuah konstanta, maka $E(k)=k$
2. Jika k adalah sebuah konstanta dan V adalah suatu peubah acak, maka $E(kV)=kE(V)$
3. Jika $k_1, k_2, \ldots, k_m$ adalah konstanta-konstanta dan $V_1, V_2, \ldots, V_m$ adalah peubah-peubah acak, maka $E(k_1V_1 + k_2V_2 + \ldots + k_mV_m) = k_1E(V_1) + k_2E(V_2) + \ldots + k_mE(V_m)$.

2.8 Variansi

Misalkan X adalah suatu peubah acak yang mempunyai *pdf* $f(x)$. Variansi dari suatu peubah acak X adalah suatu ekspektasi matematika dari $(X-\mu)^2$, dengan $\mu=E(X)$.

$$\begin{aligned}
Var(X) &= E[(X-\mu)^2] \\
&= E(X^2 - 2\mu X + \mu^2) \\
&= E(X^2) - E(2\mu X) + E(\mu^2) \\
&= E(X^2) - 2\mu E(X) + E(\mu^2) \\
&= E(X^2) - \mu^2 \\
&= E(X^2) - [E(X)]^2 \qquad (2.14)
\end{aligned}$$

2.9 Distribusi Bivariat

Pada bagian ini akan diijelaskan beberapa bagian, yaitu mengenai distribusi-distribusi dari dua peubah acak dan distribusi bersyarat.

2.9.1 Fungsi Kepadatan Probabilitas

Definisi 2.3

Diberikan suatu percobaan acak dengan ruang sampel $\mathcal{C}$. Perhatikan dua peubah acak X_1 dan X_2, yang memetakan setiap elemen c dari $\mathcal{C}$ ke satu dan hanya satu pasangan terurut dari bilangan-bilangan $X_1(c)=x_1$, $X_2(c)=x_2$. Ruang dari X_1 dan X_2 merupakan himpunan dari pasangan-pasangan terurut $\mathcal{A} = \{ (x_1,x_2) \mid x_1=X_1(c), x_2=X_2(c), c \in \mathcal{C} \}$.

Misalkan $\mathcal{A}$ merupakan ruang yang dihubungkan dengan dua peubah acak X_1 dan X_2, dan misalkan A adalah subhimpunan dari $\mathcal{A}$. Probabilitas dari kejadian A dinyatakan dengan $\Pr[(X_1, X_2) \in A]$. Ambil $C=\{c \mid c\in\mathcal{C}$ dan $\Pr[X_1(c), X_2(c)\in A]\}$, dimana $\mathcal{C}$ adalah ruang sampel. Kemudian didefinisikan $\Pr[(X_1, X_2) \in A] = P(C)$, dimana P adalah fungsi himpunan probabilitas yang didefinisikan untuk subhimpunan-subhimpunan C dari $\mathcal{C}$. $\Pr[(X_1, X_2) \in A]$ dapat dinyatakan dengan fungsi himpunan probabilitas $P_{X_1,X_2}(A)$, atau yang lebih dikenal, ditulis

$$P(A) = \Pr[(X_1, X_2) \in A]. \qquad (2.15)$$

Notasi dari *pdf* dari peubah acak X dapat diperluas untuk notasi dari *pdf* dari peubah-peubah acak bivariat. Di bawah batasan-batasan tertentu pada ruang $\mathscr{A}$ dan fungsi pada $\mathscr{A}$, dua peubah acak X dan Y disebut berdistribusi probabilitas tipe diskrit atau kontinu, dan mempunyai distribusi sesuai tipenya, berdasarkan fungsi himpunan probabilitas $P(A), A \subset \mathscr{A}$, dapat dinyatakan sebagai

$$P(A) = \Pr[(X_1, X_2) \in A] = \Sigma\Sigma_A f(x,y), \quad (2.16)$$

atau sebagai

$$P(A) = \Pr[(X_1, X_2) \in A] = \iint_A f(x,y)dxdy. \quad (2.17)$$

Dalam kedua kasus, $f(x,y)$ disebut *pdf* dari peubah-peubah acak X dan Y. Untuk setiap kasus, $P(\mathscr{A}) = 1$.

Definisi dari suatu *pdf* $f(x,y)$ dapat diperluas pada keseluruhan bidang-xy dengan menggunakan "nol untuk yang lainnya". Setelah ini dilakukan, gantikan

$$\iint_{\mathscr{A}} f(x,y)dxdy \qquad \text{dengan} \qquad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy, \quad (2.18)$$

untuk kasus kontinu, dan untuk kasus diskrit adalah

$$\Sigma\Sigma_{\mathscr{A}}\, f(x,y) \qquad \text{dengan} \qquad \Sigma_y\, \Sigma_x\, f(x,y). \quad (2.19)$$

Fungsi f untuk dua peubah acak ini, pada dasarnya memenuhi kondisi-kondisi untuk menjadi suatu *pdf* jika:

(a) $f(x,y) \geq 0, \quad x, y \in \mathbb{R}$

(b) untuk kasus kontinu $\iint_{\mathscr{A}} f(x,y)dxdy=1$, dan untuk kasus diskret $\Sigma\Sigma_{\mathscr{A}}\, f(x,y)=1$.

2.9.2 Fungsi Distribusi

Misalkan peubah-peubah acak X dan$Y$ mempunyai fungsi himpunan probabilitas $P(A)$, dimana A adalah himpunan $\{(u,v) \mid u \leq x, v \leq y\}$ yang tidak terbatas, dimana x dan y adalah bilangan-bilangan riil, maka

$$P(A) = \Pr[(X, Y) \in A] = \Pr(X \leq x, Y \leq y). \quad (2.20)$$

Fungsi pada titik (x,y) ini disebut fungsi distribusi dari peubah acak X dan Y, dan dinyatakan oleh

$$F(x,y) = (X \leq x, Y \leq y). \quad (2.21)$$

Jika X dan Y adalah peubah-peubah acak yang berdistribusi probabilitas kontinu yang mempunyai *pdf* $f(x,y)$, maka

$$F(x,y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(u,v)dudv, \tag{2.22}$$

Sehingga, pada titik-titik kontinuitas dari $f(x, y)$,

$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y) \tag{2.23}$$

Dalam setiap kasus,

$$Pr(a < X \leq b, c < Y \leq d) = F(b,d) - F(b,c) - F(a,d) + F(a,c), \tag{2.24}$$

untuk setiap konstanta riil $a < b$, $c < d$.

2.9.3 Fungsi Kepadatan Probabilitas Marjinal

Misalkan $f(x_1, x_2)$ adalah *pdf* dari peubah-peubah acak X_1 dan X_2. Untuk selanjutnya, dalam rangka penekanan dan penjelasan, *pdf* atau fungsi distribusi dari peubah acak yang lebih dari satu akan disebut *pdf* bersama atau disebut fungsi distribusi bersama. Jadi, $f(x_1, x_2)$ adalah *pdf* bersama dari peubah-peubah acak X_1 dan X_2.

Perhatikan kejadian $a < X_1 < b$, $a < b$. Kejadian ini dapat terjadi jika dan hanya jika kejadian $a<X_1<b$, $-\infty<X_2<\infty$ terjadi, yaitu, kedua kejadian ekivalen, sehingga probabilitasnya sama. Probabilitas dari $a<X_1<b$, $-\infty<X_2<\infty$ didefinisikan oleh

$$Pr(a < X_1 < b, -\infty < X_2 < \infty) = \int_{a}^{b} \int_{-\infty}^{\infty} f(x_1,x_2)dx_2\, dx_1 \tag{2.25}$$

untuk kasus kontinu, dan

$$Pr(a < X_1 < b, -\infty < X_2 < \infty) = \sum_{a<x_1<b} \sum_{x_2} f(x_1,x_2) \tag{2.26}$$

untuk kasus diskrit. Masing-masing dari

$$\int_{-\infty}^{\infty} f(x_1,x_2)dx_2 \qquad \text{dan} \qquad \sum_{x_2} f(x_1,x_2) \tag{2.27}$$

adalah fungsi dari x_1 saja, sebut $f_1(x_1)$.

Jadi, untuk setiap $a < b$, diperoleh

$$Pr(a < X_1 < b) = \int_{a}^{b} f_1(x_1)dx_1, \quad \text{untuk kasus kontinu,} \tag{2.28}$$

$$Pr(a < X_1 < b) = \sum_{a<x_1<b} f_1(x_1), \quad \text{untuk kasus diskrit.} \tag{2.29}$$

Sehingga $f_1(x_1)$ adalah *pdf* dari X_1 saja. Fungsi $f_1(x_1)$ ini disebut dengan *pdf* marjinal dari X_1.

Begitu juga untuk X_2,

$$f_2(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2)dx_1, \quad \text{untuk kasus kontinu,} \tag{2.30}$$

$$f_2(x_2) = \sum_{x_1} f(x_1, x_2), \quad \text{untuk kasus diskrit,} \tag{2.31}$$

adalah *pdf* marjinal dari X_2.

2.9.4 Distribusi Bersyarat

Misalkan X_1 dan X_2 adalah peubah acak tipe diskrit yang mempunyai *pdf* bersama $f(x_1, x_2)$, dimana $f(x_1, x_2) > 0$ untuk $(x_1. x_2) \in \mathscr{A}$ dan bernilai nol untuk lainnya. Misalkan $f_1(x_1)$ adalah *pdf* marjinal dari X_1 dan $f_2(x_2)$ adalah *pdf* marjinal dari X_2. Ambil himpunan $A_1 = \{(x_1, x_2)|x_1 = x_1', -\infty < x_2 < \infty\}$, dimana x_1' adalah suatu nilai dari $X_1 \ni P(A_1) = Pr(X_1 = x_1') = f_1(x_1') > 0$. Ambil himpunan $A_2 = \{(x_1, x_2)| -\infty < x_1 < \infty, x_2 = x_2'\}$. Maka probabilitas bersyarat A_2 diberikan A_1 adalah

$$P(A_2|A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)} = \frac{Pr(X_1 = x_1', X_2 = x_2')}{Pr(X_1 = x_1')} = \frac{f(x_1', x_2')}{f_1(x_1')},$$

dimana $f_1(x_1') > 0$.

Sehingga, jika (x_1, x_2) merupakan suatu titik dimana $f_1(x_1) > 0$, maka probabilitas bersyarat bahwa $X_2 = x_2$, diberikan $X_1 = x_1$ adalah

$$f(x_2|x_1) = \frac{f(x_1, x_2)}{f_1(x_1)}.$$

Dengan x_1 tetap, dan dengan $f_1(x_1) > 0$, maka fungsi dari x_1 ini memenuhi syarat-syarat untuk menjadi suatu *pdf* dari suatu peubah acak X_2 jenis diskrit, karena:

1. $f(x_2|x_1) = \frac{f(x_1,x_2)}{f_1(x_1)} \geq 0$
2. $\sum_{x_2} \frac{f(x_1,x_2)}{f_1(x_1)} = \frac{1}{f_1(x_1)} \sum_{x_2} f(x_1, x_2) = \frac{f_1(x_1)}{f_1(x_1)} = 1.$

Setelah didefinisikan notasi untuk $f(x_2|x_1)$ dengan hubungan:

$f(x_2|x_1) = \frac{f(x_1,x_2)}{f_1(x_1)}$ untuk $f_1(x_1) > 0$, yang disebut dengan *pdf* bersyarat dari peubah acak tipe diskrit X_2 diberikan peubah acak tipe diskrit $X_1 = x_1$. Dengan cara yang sama, notasi $f(x_1|x_2)$ didefinisikan dengan hubungan: $f(x_1|x_2) = \frac{f(x_1,x_2)}{f_2(x_2)}$, $f_2(x_2) > 0$ dan disebut *pdf* bersyarat dari peubah acak tipe diskrit X_1 diberikan peubah acak tipe diskrit $X_2 = x_2$.

Selanjutnya, misalkan X_1 dan X_2 menyatakan peubah acak-peubah acak tipe kontinu yang mempunyai *pdf* bersama $f(x_1, x_2)$, dan *pdf* marjinal masing-masing $f_1(x_1)$ dan $f_2(x_2)$. Fungsi kepadatan probabilitas (*pdf*) bersyarat untuk peubah acak tipe kontinu X_2 diberikan peubah acak kontinu $X_1 = x_1$ didefinisikan dengan hubungan

$$f(x_2|x_1) = \frac{f(x_1, x_2)}{f_1(x_1)}.$$

dalam hubungan ini, x_1 dianggap mempunyai nilai tertentu dimana $f_1(x_1) > 0$.

$f(x_2|x_1)$ mempunyai sifat-sifat *pdf* tipe kontinu dengan satu peubah acak dan disebut *pdf* bersyarat tipe kontinu dari peubah acak X_2 diberikan peubah acak tipe kontinu $X_1 = x_1$, karena:

1. $f(x_2|x_1) \geq 0$
2. $f(x_2|x_1) = \frac{f(x_1,x_2)}{f_1(x_1)}, \quad f_1(x_1) > 0, \quad f(x_1, x_2) \geq 0$
3. $\int_{-\infty}^{\infty} f(x_2|x_1)dx_2 = \int_{-\infty}^{\infty} \frac{f(x_1,x_2)}{f_1(x_1)}dx_2 = \frac{1}{f_1(x_1)}\int_{-\infty}^{\infty} f(x_1, x_2)dx_2 = \frac{f_1(x_1)}{f_1(x_1)} = 1.$

Jika $f_2(x_2) > 0$, *pdf* bersyarat dari peubah acak tipe kontinu X_1 diberikan peubah acak tipe kontinu $X_2 = x_2$, didefinisikan sebagai

$$f(x_1|x_2) = \frac{f(x_1, x_2)}{f_2(x_2)}, \qquad f_2(x_2) > 0.$$

$f(x_1|x_2)$ mempunyai sifat-sifat *pdf* tipe kontinu dengan satu peubah acak dan disebut *pdf* bersyarat tipe kontinu dari peubah acak X_1 diberikan peubah acak tipe kontinu $X_2 = x_2$, karena:

1. $f(x_1|x_2) \geq 0$
2. $f(x_1|x_2) = \frac{f(x_1,x_2)}{f_2(x_2)}, \quad f_2(x_2) > 0, \quad f(x_1,x_2) \geq 0$
3. $\int_{-\infty}^{\infty} f(x_1|x_2)dx_1 = \int_{-\infty}^{\infty} \frac{f(x_1,x_2)}{f_2(x_2)} dx_1 = \frac{1}{f_2(x_2)} \int_{-\infty}^{\infty} f(x_1,x_2)dx_1 = \frac{f_2(x_2)}{f_2(x_2)} = 1.$

Karena $f(x_1|x_2)$ dan $f(x_2|x_1)$ masing-masing merupakan suatu *pdf* dari satu peubah acak (baik tipe diskrit maupun kontinu), maka masing-masing mempunyai semua sifat-sifat dari suatu *pdf* . Dengan demikian, dapat dihitung probabilitas dan ekspektasi matematikanya sebagai berikut:

Untuk peubah acak tipe kontinu:

$$Pr(a < X_2 < b|X_1 = x_1) = \int_a^b f(x_2|x_1)dx_2$$

disebut probabilitas bersyarat bahwa $a < X_2 < b$ diberikan $X_1 = x_1$.

Dengan cara yang sama, probabilitas bersyarat bahwa $c < X_1 < d$ diberikan $X_2 = x_2$, adalah

$$Pr(c < X_1 < d|X_2 = x_2) = \int_c^d f(x_1|x_2)dx_1.$$

Jika $u(X_2)$ adalah suatu fungsi dari X_2, maka

$$E(u(X_2)|x_1) = \int_{-\infty}^{\infty} u(x_2)f(x_2|x_1)dx_2$$

disebut ekspektasi bersyarat dari $u(X_2)$ diberikan $X_1 = x_1$.

Ekspektasi khusus:

- $E(X_2|x_1)$ adalah mean dari *pdf* bersyarat dari X_2 diberikan $X_1 = x_1$
- $E\{[X_2 - E(X_2|x_1)]^2|x_1\}$ adalah variansi dari *pdf* bersyarat dari X_2 diberikan $X_1 = x_1$, dan dapat dinotasikan dengan $Var(X_2|\, x_1)$

Untuk selanjutnya, $E(X_2|x_1)$ disebut mean bersyarat dari X_2 diberikan $X_1 = x_1$. Sedangkan $Var(X_2|\, x_1)$ disebut variansi bersyarat dari X_2 diberikan $X_1 = x_1$ dan dapat ditulis sebagai berikut

$$Var(X_2|\, x_1) = E(X_2^2|x_1) - [E(X_2|x_1)]^2.$$

Dengan cara yang sama, ekspektasi bersyarat dari $u(X_1)$ diberikan $X_2 = x_2$ adalah

$$E(u(X_1)|X_2 = x_2) = \int_{-\infty}^{\infty} u(x_1) f(x_1|x_2) dx_1.$$

Untuk peubah acak tipe diskret, probabilitas bersyarat dan ekspektasi bersyaratnya dihitung dengan cara analog dengan peubah acak tipe kontinu, namun menggunakan sumasi sebagai pengganti integral.

(Hogg and Craig: 1995)

2.10 Distribusi Normal

Definisi 2.4

Sebuah peubah acak kontinu X dikatakan mempunyai suatu distribusi normal jika *pdf*-nya adalah

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \qquad -\infty < x < \infty, \tag{2.32}$$

dimana μ dan σ^2 adalah mean dan variansi dari X. Selanjutnya, X dikatakan mempunyai suatu distribusi $N(\mu, \sigma^2)$.

(Hogg, McKean and Craig: 2005)

Mean μ dari $N(\mu, \sigma^2)$ disebut parameter lokasi karena perubahan nilainya mengakibatkan perubahan lokasi dari posisi tengah (*middle*) dari pdf normal, maksudnya, bentuk grafik dari pdf normal tersebut tidak berubah melainkan hanya ada suatu pergeseran pada lokasinya. Standar deviasi σ dari $N(\mu, \sigma^2)$ disebut parameter skala karena perubahan nilainya dapat mengubah penyebaran distribusinya. Maksudnya, sebuah nilai yang kecil dari σ mengakibatkan bentuk grafik dari pdf normal menjadi tinggi dan curam, sedangkan nilai yang besar dari σ akan membuat bentuk grafiknya lebih menyebar dan landai. Berapapun nilai dari μ dan σ, bentuk grafik dari pdf normal akan menyerupai lonceng atau *bell-shape*.

2.11 Distribusi Campuran (*Mixture Distribution*)

Teorema 2.5

Misalkan X adalah peubah acak yang bergantung pada parameter λ yang merupakan nilai dari suatu peubah acak Λ dengan pdf $f_{\Lambda}(\lambda)$. Fungsi probabilitas bersyarat dari X diberikan $\Lambda = \lambda$ adalah $f_{X|\Lambda}(x|\lambda)$. Maka distribusi campuran (*mixture distribution*) didefinisikan sebagai berikut:

$$f_{X}(x) = \int_{\lambda\in\Lambda} f_{X|\Lambda}(x|\lambda) f_{\Lambda}(\lambda) d\lambda, \qquad (2.33)$$

dengan distribusi dari Λ disebut sebagai distribusi pencampur (*mixing distribution*).

Fungsi distribusi campuran dapat dibentuk dari

$$F_X(x) = \int_{-\infty}^{x} \int_{\lambda\in\Lambda} f_{X|\Lambda}(y|\lambda) f_{\Lambda}(\lambda) d\lambda\, dy$$

$$= \int_{\lambda\in\Lambda} \left[\int_{-\infty}^{x} f_{X|\Lambda}(y|\lambda) dy\right] f_{\Lambda}(\lambda) d\lambda$$

$$= \int_{\lambda\in\Lambda} F_{X|\Lambda}(x|\lambda) f_{\Lambda}(\lambda) d\lambda.$$

Moment ke-*k* dari distribusi campuran ini adalah

$$E(X^k) = E[E(X^k|\Lambda)]$$

dan, secara khusus, bentuk variansinya adalah

$$Var(X) = E[Var(X|\Lambda)] + Var[E(X|\Lambda)].$$

(A. Klugman, Harry H. Panjer and Gordon E. Willmot: 2004)

Bukti:

Karena X mempunyai pdf $f_{X|\Lambda}(x|\lambda)$, dimana λ adalah parameter dari X dan karena Λ mempunyai pdf $f_{\Lambda}(\lambda)$. Maka pdf marginal dari X adalah

- jika Λ peubah acak kontinu

$$f_{X}(x) = \int_{\lambda\in\Lambda} f_{X|\Lambda}(x|\lambda) f_{\Lambda}(\lambda) d\lambda$$

- jika Λ peubah acak diskrit

$$f_X(x) = \sum_{\lambda \in \Lambda} f_{X|\Lambda}(x|\lambda) f_\Lambda(\lambda)$$

Dengan kata lain, $f_X(x)$ dapat dinyatakan dalam bentuk ekspektasi sebagai berikut

$$f_X(x) = E_\Lambda\left(f_{X|\Lambda}(x|\Lambda)\right)$$

dimana subskrip Λ pada E mengindikasikan bahwa peubah acaknya adalah Λ.

Dengan demikian, *moment* ke-*k* $\left(E(X^k)\right)$ adalah

- jika Λ peubah acak kontinu

$$E(X^k) = \int_{-\infty}^{\infty} \int_{\lambda \in \Lambda} x^k f_{X|\Lambda}(x|\lambda) f_\Lambda(\lambda)\, d\lambda\ dx$$

$$E(X^k) = \int_{\lambda \in \Lambda} \left[\int_{-\infty}^{\infty} x^k f_{X|\Lambda}(x|\lambda) dx\right] f_\Lambda(\lambda) d\lambda$$

$$E(X^k) = \int_{\lambda \in \Lambda} E(X^k|\lambda) f_\Lambda(\lambda) d\lambda$$

$$E(X^k) = E[E(X^k|\Lambda)]. \tag{2.34}$$

- jika Λ peubah acak diskrit

$$E(X^k) = \sum_{x} \sum_{\lambda \in \Lambda} x^k f_{X|\Lambda}(x|\lambda) f_\Lambda(\lambda)$$

$$E(X^k) = \sum_{x} \left[\sum_{\lambda \in \Lambda} x^k f_{X|\Lambda}(x|\lambda)\right] f_\Lambda(\lambda)$$

$$E(X^k) = \sum_{x} E(X^k|\lambda) f_\Lambda(\lambda)$$

$$E(X^k) = E[E(X^k|\Lambda)]. \tag{2.35}$$

Sehingga, nilai variansi $Var(X)$ adalah

$$\begin{aligned} Var(X) &= E(X^2) - \{E(X)\}^2 \\ &= E[E(X^2|\Lambda)] - \{E[E(X|\Lambda)]\}^2 \\ &= E\{Var(X|\Lambda) + [E(X|\Lambda)]^2\} - \{E[E(X|\Lambda)]\}^2 \\ &= E[Var(X|\Lambda)] + Var[E(X|\Lambda)] \end{aligned}$$

Definisi 2.6

Sebuah peubah acak Y adalah suatu campuran k-titik (*k-point mixture)* dari peubah acak X_1, X_2, …, X_k jika fungsi distribusinya berbentuk

$$F_Y(y) = a_1 F_{X_1}(y) + a_2 F_{X_2}(y) + \cdots + a_k F_{X_k}(y) \qquad (2.34)$$

dimana semua $a_j > 0$ dan $a_1 + a_2 + \cdots + a_k = 1$.

(A. Klugman, Harry H. Panjer and Gordon E. Willmot: 2004)

Selanjutnya, akan dibahas salah satu bentuk khusus dari distribusi campuran, yaitu distribusi normal terkontaminasi (*contaminated normal distribution*). Pembahasannya dimulai dengan sebuah peubah acak yang terstandardisasi.

2.12 Distribusi Normal Terkontaminasi (*Contaminated Normal Distribution*)

Sebuah peubah acak dikatakan mempunyai distribusi normal terkontaminasi jika distribusinya merupakan gabungan dari dua distribusi normal dengan mean sama tetapi variansinya berbeda, sehingga hanya terdapat sebagian kecil nilai observasi yang berasal dari distribusi normal dengan variansi yang lebih besar. Jika proporsi banyaknya nilai observasi yang berasal dari distribusi normal dengan variansi yang lebih besar itu cukup kecil, maka distribusi normal terkontaminasi akan terlihat seperti distribusi normal dengan *outlier*.

Misalkan terdapat sebuah peubah acak W yang sebagian besar observasinya mengikuti suatu distribusi normal standar $N(0,1)$ namun terkadang terlihat seolah mengikuti suatu distribusi normal dengan variansi yang lebih besar, yaitu $N(0,\sigma_c^2)$ dimana $\sigma_c^2 > 1$. Pada aplikasinya, mungkin dapat dikatakan bahwa sebagian besar observasi dari data tersebut "baik" namun terdapat outlier di dalamnya. Oleh karena itu, untuk mendapatkan distribusi yang lebih tepat dari W, lakukan langkah-langkah berikut ini:

- misalkan peubah acak Z mempunyai distribusi $N(0,1)$
- misalkan $I_{1-\epsilon}$ adalah sebuah peubah acak diskrit yang didefinisikan sebagai

$$I_{1-\epsilon} = \begin{cases} 1 & dengan\ probabilitas\ 1-\epsilon \\ 0 & dengan\ probabilitas\ \epsilon. \end{cases}$$

- misalkan $W = ZI_{1-\epsilon} + \sigma_c Z(1 - I_{1-\epsilon})$, dan asumsikan bahwa Z dan $I_{1-\epsilon}$ independen.

Bila peubah acak $W = ZI_{1-\epsilon} + \sigma_c Z(1 - I_{1-\epsilon})$ diperhatikan kembali, maka akan terlihat bahwa peubah acak $I_{1-\epsilon}$ digunakan sebagai penanda (indikator) terjadi atau tidaknya kontaminasi pada sebuah titik observasi. Misalnya, jika sebuah titik *w* telah terkontaminasi berarti $I_{1-\epsilon} = 0$, sehingga sebenarnya titik observasi tersebut adalah $w = \sigma_c z$. Sebaliknya, jika *w* bukanlah titik yang terkontaminasi, maka $I_{1-\epsilon} = 1$, artinya titik tersebut adalah $w = z$. Oleh karena itu, probabilitas dari W merupakan probabilitas bersyarat dari W diberikan $I_{1-\epsilon}$.

Selanjutnya akan dicari fungsi distribusi dari W. Karena probabilitas dari W merupakan probabilitas bersyarat dari W diberikan $I_{1-\epsilon}$, dan karena independensi antara Z dan $I_{1-\epsilon}$, maka bentuk fungsi distribusi dari W adalah

$$\begin{aligned} F_W(w) &= P[W \leq w, I_{1-\epsilon} = 1] + P[W \leq w, I_{1-\epsilon} = 0] \\ &= P[W \leq w |\ I_{1-\epsilon} = 1]P[I_{1-\epsilon} = 1]\ + P[W \leq w |\ I_{1-\epsilon} = 0]P[I_{1-\epsilon} = 0] \\ &= P[Z \leq w](1-\epsilon) + P[Z \leq w/\sigma_c]\epsilon \\ &= \Phi(w)(1-\epsilon) + \Phi(w/\sigma_c)\epsilon \qquad (2.35) \end{aligned}$$

dimana Φ adalah *cdf* normal standar.

Persamaan (2.35) menunjukkan bahwa distribusi dari W merupakan gabungan dari dua distribusi normal, dengan demikian dapat ditarik kesimpulan bahwa distribusi dari W adalah distribusi normal terkontaminasi.

Jika formula (2.35) diturunkan terhadap *w*, maka akan didapat pdf dari W yaitu

$$f_W(w) = \phi(w)(1-\epsilon) + \phi(w/\sigma_c)\frac{\epsilon}{\sigma_c}, \qquad (2.36)$$

dimana ϕ adalah suatu *pdf* normal standar.

Kemudian, karena $W = ZI_{1-\epsilon} + \sigma_c Z(1 - I_{1-\epsilon})$, maka dapat dibuktikan bahwa ekspektasi dan variansi dari W berturut-turut adalah

$$E(W) = 0 \text{ dan } Var(W) = 1 + \epsilon(\sigma_c^2 - 1). \qquad (2.37)$$

Bukti:

a) Akan dibuktikan bahwa $E(W) = 0$.

Karena Z berdistribusi $N(0,1)$, maka $E(Z) = 0$ dan $Var(Z) = 1$. Kemudian, karena W merupakan peubah acak yang mempunyai distribusi campuran, maka berdasarkan teorema 2.5, ekspektasi W merupakan ekspektasi bersyarat sebagai berikut

$E(W) = E[E(W|I_{1-\epsilon})]$

$= E(W|I_{1-\epsilon} = 0)P(I_{1-\epsilon} = 0) + E(W|I_{1-\epsilon} = 1)P(I_{1-\epsilon} = 1)$

$= E(\sigma_c Z)P(I_{1-\epsilon} = 0) + E(Z)P(I_{1-\epsilon} = 1)$

$= \sigma_c E(Z)\epsilon + E(Z)(1-\epsilon)$

$E(W) = 0.$

b) Akan dibuktikan bahwa $Var(W) = 1 + \epsilon(\sigma_c^2 - 1)$.

Karena Z berdistribusi $N(0,1)$, maka $E(Z) = 0$ dan $Var(Z) = 1$.

$Var(Z) = E(Z^2) - [E(Z)]^2$

$1 = E(Z^2) - [0]^2$

$E(Z^2) = 1.$

Kemudian, karena W berdistribusi campuran, maka berdasarkan teorema 2.5, momen ke-2 dari distribusi ini adalah

$E(W^2) = E(W^2|I_{1-\epsilon})$

$= E(W^2|I_{1-\epsilon} = 0)P(I_{1-\epsilon} = 0) + E(W^2|I_{1-\epsilon} = 1)P(I_{1-\epsilon} = 1)$

$= E(\sigma_c^2 Z^2)P(I_{1-\epsilon} = 0) + E(Z^2)P(I_{1-\epsilon} = 1)$

$= \sigma_c^2 E(Z^2)\epsilon + E(Z^2)(1-\epsilon)$

$= \sigma_c^2\epsilon + (1-\epsilon)$

$= 1 + \epsilon(\sigma_c^2 - 1).$

Dengan demikian, variansi dari *W* adalah

$Var(W) = E(W^2) - [E(W)]^2$

$= [1 + \epsilon(\sigma_c^2 - 1)] - [0]^2$

$= 1 + \epsilon(\sigma_c^2 - 1).$

∴ Terbukti.

Misalkan, terdapat suatu peubah acak baru $X = a + bW$, dimana *a* dan *b* konstanta, dengan *b* > 0. Dengan mensubtitusi W ke persamaan (2.35), maka diperoleh bahwa fungsi distribusi dari *X* adalah

$$F_X(x) = \Phi\left(\frac{x-a}{b}\right)(1-\epsilon) + \Phi\left(\frac{x-a}{b\sigma_c}\right)\epsilon, \tag{2.38}$$

dimana distribusi ini juga merupakan suatu distribusi normal terkontaminasi atau distribusi campuran normal (*mixture of normals*). Sedangkan pdf dari *X* adalah

$$f_X(x) = \phi\left(\frac{x-a}{b}\right)\frac{(1-\epsilon)}{b} + \phi\left(\frac{x-a}{b\sigma_c}\right)\frac{\epsilon}{b\sigma_c}. \tag{2.39}$$

Berdasarkan (2.37), mean dan variansi dari X adalah

$$E(X) = a \text{ dan } Var(X) = b^2\big(1 + \epsilon(\sigma_c^2 - 1)\big). \quad (2.40)$$

(Hogg, McKean and Craig: 2005)

2.13 Metode Maksimum *Likelihood*

Misalkan $X_1, X_2, \ldots, X_n$ adalah peubah acak – peubah acak yang tidak diketahui apakah saling independen dan berdistribusi identik atau tidak. Misalkan *pdf* bersama $g(x, y, \ldots, z; \theta_1, \theta_2, \ldots, \theta_m), (\theta_1, \theta_2, \ldots, \theta_m) \in \boldsymbol{\Omega}$ bergantung pada m parameter yang tidak diketahui dan $\boldsymbol{\Omega}$ adalah ruang parameter. Untuk selanjutnya, *pdf* bersama dari $X_1, X_2, \ldots, X_n$ disebut fungsi *likelihood*, dan dinotasikan dengan $L[\theta_1, \theta_2, \ldots, \theta_m; x_1, x_2, \ldots, x_n]$ atau $L[\boldsymbol{\theta}; x_1, x_2, \ldots, x_n]$. Misalkan terdapat fungsi $u_1(X_1, X_2, \ldots, X_n),\ u_2(X_1, X_2, \ldots, X_n), \ldots, u_m(X_1, X_2, \ldots, X_n)$ yang memaksimumkan fungsi *likelihood* $L[\boldsymbol{\theta}; x_1, x_2, \ldots, x_n]$. Taksiran maksimum *likelihood* diperoleh dengan menentukan

$\hat{\theta}_1 = u_1(X_1, X_2, \ldots, X_n),\ \hat{\theta}_2 = u_2(X_1, X_2, \ldots, X_n),\ \ldots, \hat{\theta}_m = u_m(X_1, X_2, \ldots, X_n)$

dengan cara menyelesaikan persamaan berikut:

$$\frac{\partial \ln L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 0\,.$$

(Hogg and Craig: 1995)

2.14 *M-Estimator*

Untuk beberapa keluarga distribusi parametric, penaksir maksimum *likelihood es* (MLE) adalah sebuah penaksir parameter yang konsisten dan mempunyai variansi minimum. Namun, terkadang ditemukan kasus dimana MLE diperoleh, tetapi tidak dalam bentuk yang sederhana, sehingga perlu proses iterasi untuk mendapatkannya. Selain itu, masih terdapat permasalahan saat MLE diperoleh, yaitu seberapa baik MLE tersebut dapat digunakan sebagai penaksir ketika distribusi sebenarnya menyimpang dari distribusi yang diasumsikan.

Dari permasalahan tersebut, munculah ide tentang metode *M-estimate* yang diperkenalkan oleh P.J Huber pada tahun 1964. *M-estimate* digunakan untuk mendapatkan penaksir yang tidak sensitif terhadap sedikit penyimpangan asumsi distribusi. Oleh karena itu, penaksir yang diperoleh dengan menggunakan metode ini disebut dengan *robust M-estimator*. Selanjutnya, akan dijelaskan *M-estimator* untuk parameter lokasi dan skala.

2.14.1 *M-estimator* Untuk Parameter Lokasi

Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dari distribusi yang mempunyai fungsi probabilitas $f(x-\theta)$, dimana θ adalah suatu parameter lokasi sedemikian sehingga

$$\ln L(\theta) = \sum_{i=1}^{n} \ln f(x_i - \theta) = -\sum_{i=1}^{n} \rho(x_i - \theta),$$

dimana $\rho(x) = -\ln f(x)$ dan

$$\frac{d \ln L(\theta)}{d\theta} = -\sum_{i=1}^{n} \frac{f'(x_i-\theta)}{f(x_i-\theta)} = \sum_{i=1}^{n} \psi(x_i - \theta),$$

dimana $\rho'(x) = \psi(x)$.

(Hogg and Craig: 1995)

Penaksir yang diperoleh dengan mencari solusi dari persamaan

$$\sum_{i=1}^{n} \psi(x_i - \theta) = 0$$

disebut *robust M-estimator*. Oleh sebab itu, untuk mendapatkan sebuah *M-estimator* terlebih dahulu harus dipilih suatu fungsi ψ yang dapat menghasilkan penaksir yang baik untuk setiap distribusi sampel yang terambil.

Sekarang, perhatikan suatu fungsi $\psi(x) = sgn(x)$ yang disarankan oleh Huber untuk dapat menghasilkan *M-estimator* yang baik untuk menaksir parameter lokasi.

M-estimator yang dinotasikan dengan $\hat{\theta}$ ini akan didapat dengan menyelesaikan persamaan $\sum_{i=1}^{n} \psi(X_i - \theta) = 0$ sebagai berikut:

$$\sum_{i=1}^{n} \psi(X_i - \theta) = 0$$

$$\sum_{i=1}^{n} sgn(X_i - \theta) = 0 \tag{2.41}$$

Misalkan didefinisikan $I(x>0)=\begin{cases}1 & jika\ x>0\\ 0 & jika\ x\le 0\end{cases}$ dan

$I(x<0)=\begin{cases}1 & jika\ x<0\\ 0 & jika\ x\ge 0\end{cases}$, maka fungsi $sgn(x)$ dapat dinyatakan sebagai

$sgn(x)=I(x>0)-I(x<0)$

Sehingga persamaan (2.41) dapat dinyatakan sebagai

$$\sum_{i=1}^{n}\big(I(X_i-\theta>0)-I(X_i-\theta<0)\big)=0$$

$$\sum_{i=1}^{n}I(X_i-\theta>0)-\sum_{i=1}^{n}I(X_i-\theta<0)=0$$

$$\#(X_i>\theta)-\#(X_i<\theta)=0$$

dimana $\#(x)$=banyaknya *x*, sehingga diperoleh

$$\#(X_i>\theta)=\#(X_i<\theta). \qquad (2.42)$$

Dari persamaan (2.42) diperoleh bahwa banyaknya *x* dengan nilai yang lebih besar dari θ sama dengan banyaknya *x* dengan nilai yang lebih kecil dari θ, maka penaksir M-robust yang diperoleh untuk θ adalah $\hat{\theta}=median(X_1,X_2,\dots,X_n)$. Untuk selanjutnya, $median(X_1,X_2,\dots,X_n)$ dinotasikan dengan *MD* dan pada bab 3 akan dijelaskan bahwa *MD* merupakan penaksir yang *robust* untuk mean proses.

2.14.2 Penaksir-M untuk parameter skala

Dengan adanya fungsi Huber (ψ), permasalahan baru muncul, yaitu tidak semua solusi dapat ditemukan saat menyelesaikan persamaan $\sum_{i=1}^{n}\psi(x_i-\theta)=0$. Oleh karena itu, untuk menyelesaikannya, dilakukan modifikasi sebagai berikut:

$$\sum_{i=1}^{n}\psi\left(\frac{x_i-\theta}{d}\right)=0,$$

dimana θ adalah parameter lokasi dan *d* adalah parameter skala.

Penaksir *d* yang sering digunakan adalah

$$d = \frac{median|x_i - median(x_i)|}{0.6745},$$

dimana pembilang dari *d* merupakan penaksir skala yang disebut *median absolute deviation* (MAD). Sedangkan penyebut dari *d* yaitu 0,6745 digunakan agar *d* menjadi penaksir yang dekat dengan σ, ketika sampel yang diambil berasal dari suatu distribusi normal, seperti yang akan dijelaskan berikut ini.

Definisikan

$$\begin{aligned} MAD(x) &= MAD(x_1, x_2, \ldots, x_n) \\ &= median\{|x_i - median(x_i)|\} \end{aligned}$$

Jika distribusi dari *x* simetris, maka $median(x_i) = \mu$, sehingga $MAD(x) = median\{|x_i - \mu|\}$ dan berlaku

$$Pr\big(|x_i - \mu| \leq MAD(x)\big) = \frac{1}{2}$$

$$Pr\big(-MAD(x) \leq x_i - \mu \leq MAD(x)\big) = \frac{1}{2}$$

$$Pr\left(-\frac{MAD(x)}{\sigma} \leq \frac{x_i - \mu}{\sigma} \leq \frac{MAD(x)}{\sigma}\right) = \frac{1}{2}$$

$$Pr\left(-\frac{MAD(x)}{\sigma} \leq Z \leq \frac{MAD(x)}{\sigma}\right) = \frac{1}{2}$$

Jika *Z* berdistribusi *N(0,1)*, maka didapat $\frac{MAD(x)}{\sigma} = 0.6745$. Sehingga, $\hat{\sigma} = \frac{MAD(x)}{0.6745} = 1.4826\ MAD(x)$.

Untuk selanjutnya $d = 1.4826\ MAD(x) = \frac{median|x_i - median(x_i)|}{0.6745}$ dinotasikan dengan *MAD*, dan pada bab 3 akan dijelaskan bahwa *MAD* merupakan penaksir yang *robust* untuk standar deviasi proses.

2.15 Bagan Kendali Shewhart

Dalam suatu proses produksi, walaupun telah dirancang atau dipelihara dengan baik, pasti terdapat variabilitas natural di dalamnya. Variabilitas natural ini adalah efek kumulatif yang tidak dapat dihindari, dan sering disebut dengan "sebab acak dari sistem stabil". Sebuah proses produksi yang beroperasi dengan hanya dipengaruhi oleh variasi sebab acak dapat dikatakan sebagai proses yang terkendali secara statistik. Dengan kata lain, sebab acak sudah menjadi bagian lazim dari suatu proses produksi.

Pada output sebuah proses, terkadang juga dipengaruhi oleh variabilitas jenis lain. Variabilitas jenis ini biasanya berasal dari tiga sumber yaitu pengaturan mesin yang salah atau tidak sesuai, kesalahan operator, atau kecacatan bahan material. Variabilitas jenis ini pada umumnya mempunyai andil yang lebih besar dalam mempengaruhi tingkat penolakan dari performa proses produksi dibandingkan dengan variabilitas natural. Oleh karena itu, keberadaan variabilitas ini harus segera diusut atau diselidiki penyebabnya agar dapat segera dihilangkan. Variabilitas ini sering disebut dengan "sebab terusut". Sebuah proses yang beroperasi dibawah keberadaan sebab terusut dapat dikatakan sebagai proses yang tidak terkendali.

Pengendalian proses statistik atau *Statistical Process Control* (SPC) adalah sekumpulan perangkat statistik yang berguna untuk mengendalikan proses produksi agar dapat berjalan dengan stabil dan meningkatkan kemampuan produksi melalui pengurangan variabilitas pada proses produksi. Salah satu perangkat statistik dalam prosedur SPC adalah bagan kendali (*control chart*). Bagan kendali yang paling umum dikenal adalah bagan kendali Shewhart.

Bagan kendali Shewhart diperkenalkan pertama kali oleh Walter A. Shewhart pada tahun 1924. Ide dasar dari bagan kendali Shewhart adalah melakukan analisis dengan mengambil sampel dari proses produksi secara periodik. Sampel dari proses produksi yang diambil secara periodik disebut dengan istilah subgrup. Kemudian, hasil pengukuran pada setiap subgrup digunakan untuk menghitung batas-batas kendali pada bagan yaitu garis tengah (*CL*), batas kendali atas (*UCL*), dan batas kendali bawah (*LCL*). Setelah itu, batas-batas kendali dan hasil pengukuran dari setiap subgrup diplot bersama-sama untuk diperiksa apakah terdapat sinyal *out of control process* yang muncul sebagai indikasi bahwa proses tidak terkendali secara statistik.

Western Electric Handbook (1956) menyarankan sekumpulan aturan keputusan untuk mendeteksi pola nonrandom pada bagan kendali. Secara khusus, dia menyimpulkan bahwa sebuah proses dapat dikatakan tidak terkendali jika terdapat:

1. Satu titik terplot di luar batas kendali 3σ,
2. Dua dari tiga titik berurutan terplot di luar batas-batas kendali 2σ,
3. Empat dari lima titik terplot pada sebuah jarak 1σ atau dekat dengan *CL*, atau
4. Delapan titik berurutan terplot pada satu sisi dari *CL*.

Aturan-aturan ini diterapkan pada satu sisi dari *CL* pada satu waktu. Oleh sebab itu, sebuah titik yang berada di atas *UCL* yang diikuti dengan sebuah titik di bawah *LCL* tidak mengindikasikan sebuah sinyal bahwa proses tidak terkendali. Selain itu, terdapat juga beberapa kriteria tambahan yang terkadang digunakan untuk meningkatkan sensitifitas dari bagan kendali Shewhart terhadap pergeseran proses yang kecil. Beberapa aturan sensitifitas yang telah digunakan secara luas pada bagan kendali Shewhart terangkum dalam tabel berikut ini.

Tabel 2.2 Beberapa aturan sensitifitas pada bagan kendali Shewhart

Standard action signal	1. Satu atau lebih titik terletak di luar batas-batas kendali (UCL dan LCL)	Western Electric Rules
	2. Dua dari tiga titik berurutan pada sisi yang sama, terletak di luar batas 2σ, tetapi masih di dalam batas-batas kendali (UCL dan LCL)	
	3. Empat dari lima titik berurutan pada sisi yang sama, terletak di luar batas 1σ	
	4. Delapan titik berurutan terletak pada satu sisi yang sama dari CL	
	5. Enam titik menaik atau menurun terletak pada sisi dan zona yang sama dari CL	
	6. Lima belas titik dalam sebuah baris terletak pada zona 1-σ dari CL	
	7. Empat belas titik dalam sebuah baris bergantian naik dan turun	
	8. Delapan titik dalam sebuah baris terletak di sisi atas dan bawah CL dengan tidak ada titik yang terletak pada zona 1-σ dari CL	
	9. Sebuah pola tidak biasa atau non random pattern pada data	
	10. Satu atau lebih titik terletak di dekat batas kendali (UCL dan LCL)	

Bagan kendali Shewhart yang biasa digunakan adalah bagan $\bar{X}_{Shewhart}$ untuk mengawasi mean proses dan bagan $S_{Shewhart}$ untuk mengawasi standar deviasi proses. Kedua bagan ini memberikan informasi mengenai mean dan standar deviasi suatu proses.

Rumus yang digunakan untuk menghitung batas kendali pada bagan $\bar{X}$ dan bagan S adalah

a. Jika μ dan σ diketahui

- Batas-batas kendali untuk bagan $\bar{X}$

  $CL = E(\bar{X}) = \mu$

  $$LCL = \mu - 3\sigma_{\bar{X}} = \mu - 3\frac{\sigma}{\sqrt{n}} = \mu - \frac{3}{\sqrt{n}}\sigma$$

  $$UCL = \mu + 3\sigma_{\bar{X}} = \mu + 3\frac{\sigma}{\sqrt{n}} = \mu + \frac{3}{\sqrt{n}}\sigma$$

  Dengan demikian, batas-batas kendali untuk bagan $\bar{X}$ dapat ditulis sebagai berikut

  $CL = \mu$

  $LCL = \mu - A\sigma$

  $UCL = \mu + A\sigma$

  dimana $A = \frac{3}{\sqrt{n}}$

- Batas-batas kendali untuk bagan S

  Karena μ dan σ diketahui, maka dapat dibuktikan bahwa $E(S^2) = \sigma^2$ dan $E(S) = c_4\sigma.$ (lihat lampiran1)

  Sehingga

  $$\begin{aligned} \sigma_S^2 &= E(S^2) - [E(S)]^2 \\ &= \sigma^2 - [c_4\sigma]^2 \\ &= \sigma^2 - {c_4}^2\sigma^2 \\ &= (1 - {c_4}^2)\sigma^2. \end{aligned}$$

Nilai konstanta c_4 bergantung pada ukuran subgroup yang tertera pada tabel di lampiran 2.

Karena $\sigma_S^2 = (1 - c_4{}^2)\sigma^2$, maka $\sigma_S = \sigma\sqrt{(1 - c_4{}^2)}$.

Oleh karena itu,

$CL = E(S) = c_4\sigma$

$LCL = c_4\sigma - 3\sigma_S = c_4\sigma - 3c_5\sigma = (c_4 - 3c_5)\sigma$

$UCL = c_4\sigma + 3\sigma_S = c_4\sigma + 3c_5\sigma = (c_4 + 3c_5)\sigma$

Dengan demikian, batas-batas kendali untuk bagan S dapat ditulis sebagai berikut

$CL = c_4\sigma$

$LCL = B_5\sigma$

$UCL = B_6\sigma$

dimana $B_5 = (c_4 - 3c_5)$ dan $B_6 = (c_4 + 3c_5)$

b. Jika μ dan σ tidak diketahui

Jika μ dan σ tidak diketahui, maka perlu dicari penaksir untuk parameter lokasi dan parameter skala. Penaksir *unbiased* untuk μ dalam konstruksi bagan kendali Shewhart adalah $\bar{X}$. Sedangkan penaksir *unbiased* untuk σ adalah $\frac{\bar{S}}{c_4}$.

- Batas-batas kendali untuk bagan $\bar{X}$

$CL = \hat{\mu} = \bar{X}$

$$LCL = \bar{X} - 3\frac{\hat{\sigma}}{\sqrt{n}} = \bar{X} - \frac{3}{\sqrt{n}}\frac{\bar{S}}{c_4} = \bar{X} - \frac{3}{c_4\sqrt{n}}\bar{S}$$

$$UCL = \bar{X} + 3\frac{\hat{\sigma}}{\sqrt{n}} = \bar{X} + \frac{3}{\sqrt{n}}\frac{\bar{S}}{c_4} = \bar{X} + \frac{3}{c_4\sqrt{n}}\bar{S}$$

Dengan demikian, batas-batas kendali untuk bagan $\bar{X}$ dapat ditulis sebagai berikut

$CL = \bar{X}$

$LCL = \bar{X} - A_3\bar{S}$

$UCL = \bar{X} + A_3\bar{S}$

dimana $A_3 = \frac{3}{c_4\sqrt{n}}$.

- Batas-batas kendali untuk bagan S

  $CL = c_4\hat{\sigma} = c_4\ \frac{\bar{S}}{c_4} = \bar{S}$

  $LCL = \bar{S} - 3\sigma_S = \bar{S} - 3c_5\hat{\sigma} = \bar{S} - 3c_5\frac{\bar{S}}{c_4} = \left(1 - 3\frac{c_5}{c_4}\right)\bar{S} = \left(\frac{c_4 - 3c_5}{c_4}\right)\bar{S}$

  $= \frac{B_5}{c_4}\bar{S}$

  $UCL = \bar{S} + 3\sigma_S = \bar{S} + 3c_5\hat{\sigma} = \bar{S} + 3c_5\frac{\bar{S}}{c_4} = \left(1 + 3\frac{c_5}{c_4}\right)\bar{S} = \left(\frac{c_4 + 3c_5}{c_4}\right)\bar{S}$

  $= \frac{B_6}{c_4}\bar{S}$

  Dengan demikian, batas-batas kendali untuk bagan S dapat ditulis sebagai berikut

  $CL = \bar{S}$

  $LCL = B_3\bar{S}$

  $UCL = B_4\bar{S}$

  dimana $B_3 = \frac{B_5}{c_4}$ dan $B_4 = \frac{B_6}{c_4}$.

(Montgomery: 2001)

BAB 3
BAGAN KENDALI BERDASARKAN MEDIAN DAN *MEDIAN ABSOLUTE DEVIATION*

Ketika terdapat *outlier* pada data yang diasumsikan berdistribusi normal yang diambil dari sebuah proses produksi, dimana *outlier* tersebut bukan merupakan kesalahan, maka bagan kendali Shewhart yang dibuat dari data tersebut tidak lagi baik untuk digunakan sebagai alat pengendali mutu secara statistika. Hal ini terjadi karena *outlier* tersebut akan mempengaruhi mean sampel $(\bar{X})$ dan standar deviasi sampel (S) yang biasa digunakan sebagai penaksir untuk mean proses (μ) dan standar deviasi proses (σ) pada bagan kendali Shewhart, akibatnya batas-batas kendali pada bagan tersebut menjadi tidak valid. Oleh sebab itu, diperlukan bagan kendali alternatif dengan penaksir-penaksir yang *robust* atau penaksir yang tahan terhadap keberadaan *outlier*.

Salah satu bagan kendali alternatif yang dapat digunakan adalah bagan kendali berdasarkan median dan *median absolute deviation*, yang selanjutnya disebut dengan bagan kendali MDMAD. Sesuai dengan namanya, bagan kendali MDMAD ini menggunakan dua penaksir *robust* yaitu median sampel (MD) yang digunakan untuk menaksir mean proses (μ) dan *median absolute deviation* sampel (MAD) yang digunakan untuk menaksir standar deviasi proses (σ).

3.1 Penaksir *Robust*

Penaksir *robust* adalah penaksir yang tahan terhadap keberadaan *outlier*, sehingga taksirannya tidak berubah oleh sedikit penyimpangan asumsi distribusinya. (Moustafa Omar: 2008)

Penyimpangan asumsi distribusi diantaranya dapat disebabkan oleh adanya outlier yang bukan merupakan kesalahan atau observasi yang terkontaminasi. Efek yang ditimbulkan oleh sebuah observasi yang terkontaminasi pada saat menaksir suatu parameter dapat diukur dengan menggunakan *influence function* dari suatu penaksir.

Influence Function

Ide tentang *influence function* diperkenalkan oleh Hampel pada tahun 1974, sebagai berikut:

Misalkan $X_1, X_2, \dots, X_n$ adalah peubah acak independen, dengan fungsi distribusi masing-masing F_X. Kemudian didefinisikan sebuah *functional* yaitu suatu fungsi dari F_X yang dinotasikan dengan $\theta(F_X)$. Penaksir untuk parameter dari distribusi F_X dapat dinyatakan dalam bentuk *functional*. Misalnya, mean (μ) dari fungsi distribusi F_X dinyatakan dalam *functional* sebagai

$$\theta_\mu(F_X) = \int_{-\infty}^{\infty} x \, dF_X \,.$$

Setelah mengetahui konsep *functional*, hal lain yang perlu diperhatikan dalam mencari *influence function* dari sebuah penaksir adalah efek kontaminasi (ϵ) pada beberapa titik *x* yang dapat menyebabkan *x* seolah-olah mempunyai distribusi yang berbeda dengan distribusi yang diasumsikan sebelumnya. Distribusi dari *x* yang dimaksud adalah distribusi campuran yaitu $F_{X,\epsilon}$. Efek kontaminan (ϵ) pada *x* dari suatu penaksir $\hat{\theta}$ diukur dengan *influence function*.

Definisi 3.1

Influence function dari penaksir $\hat{\theta}$ pada *x* diberikan oleh

$$IF(x;\hat{\theta}) = \lim_{\varepsilon \to 0} \frac{\theta(F_{X,\epsilon}) - \theta(F_X)}{\epsilon}$$

dimana $\theta(F_{X,\epsilon})$ adalah notasi *functional* untuk penaksir dari suatu parameter θ dibawah distribusi campuran $F_{X,\epsilon}$, sedangkan $\theta(F_X)$ adalah notasi *functional* untuk penaksir dari suatu parameter θ dibawah distribusi yang diasumsikan F_X. (Dudewicz dan Mishra: 1988)

Penaksir yang diharapkan adalah penaksir yang taksirannya untuk setiap *x*, pada saat data terkontaminasi tetap dekat dengan taksiran yang dihasilkan di bawah distribusi yang diasumsikan. Dengan demikian, *influence function* dari penaksir *robust* harus merupakan fungsi yang terbatas dari *x*.

Definisi 3.2

Penaksir $\hat{\theta}$ disebut penaksir robust jika $\left|IF\left(x;\hat{\theta}\right)\right|$ terbatas untuk setiap *x*. (Hogg, Mc.Kean, and Craig: 2005)

Definisi 3.3

Sebuah fungsi $f: A \to \mathbb{R}$ dikatakan terbatas pada A jika terdapat sebuah konstanta M > 0 sedemikian sehingga $|f(x)| \leq M,\ \forall x \in A.$
(Bartle and Sherbert: 2000)

Selanjutnya akan dicari penaksir *robust* untuk mean proses dan standar deviasi proses. Penaksir yang dapat digunakan untuk menaksir mean proses adalah mean sampel dan median sampel, sedangkan penaksir yang digunakan untuk menaksir standar deviasi proses adalah standar deviasi sampel dan *meadian absolute deviation* sampel. Oleh karena itu, pencarian penaksir *robust* untuk mean proses dan standar deviasi proses dilakukan dengan membandingkan *influence function* dari empat penaksir tersebut seperti yang akan dijelaskan di bawah ini.

***Influence function* dari mean sampel**

Misalkan *X* adalah sebuah peubah acak kontinu yang mempunyai fungsi distribusi F_X dengan mean μ dan variansi σ^2. Misalkan *W* peubah acak diskrit yang mempunyai fungsi distribusi $\Delta_x(w) = \begin{cases} 0, & w < x \\ 1, & w \geq x \end{cases}$, dan misalkan $I_{1-\epsilon}$ adalah sebuah peubah acak indikator yang didefinisikan sebagai berikut:

$$I_{1-\epsilon} = \begin{cases} 1 & dengan\ probabilitas\ 1-\epsilon \\ 0 & dengan\ probabilitas\ \epsilon. \end{cases}$$

Kemudian, misalkan terdapat sebuah peubah acak $Y = I_{1-\epsilon}X + [1 - I_{1-\epsilon}]W$, dimana *X*, *W* dan $I_{1-\epsilon}$ independen. Dengan demikian, dapat dikatakan bahwa *Y* adalah peubah acak yang bergantung pada $I_{1-\epsilon}$, karena jika $I_{1-\epsilon} = 1$, maka *Y* diambil dari *X*, dan jika $I_{1-\epsilon} = 0$, maka *Y* diambil dari *W*.

Selanjutnya akan dibuktikan bahwa Y adalah sebuah peubah acak yang mempunyai distribusi campuran.

Fungsi distribusi dari Y adalah

$$
\begin{aligned}
F_{X,\epsilon}(y) &= Pr(Y \leq y) \\
&= Pr(Y \leq y, I_{1-\epsilon} = 1) + Pr(Y \leq y, I_{1-\epsilon} = 0) \\
&= Pr(X \leq y)Pr(I_{1-\epsilon} = 1) + Pr(W \leq y)Pr(I_{1-\epsilon} = 0) \\
&= (1-\epsilon)F_X + \epsilon\Delta_X. \qquad (3.1)
\end{aligned}
$$

Berdasarkan teorema 2.5, mean dari distribusi $F_{X,\epsilon}$ adalah

$$
\begin{aligned}
E(Y) &= E[E(Y|I_{1-\epsilon})] \\
&= E(Y|I_{1-\epsilon} = 0)P(I_{1-\epsilon} = 0) + E(Y|I_{1-\epsilon} = 1)P(I_{1-\epsilon} = 1) \\
&= E(W)P(I_{1-\epsilon} = 0) + E(X)P(I_{1-\epsilon} = 1) \\
&= \epsilon x + (1-\epsilon)\mu. \qquad (3.2)
\end{aligned}
$$

Kemudian, karena E(Y) merupakan fungsi dari $F_{X,\epsilon}$, maka E(Y) dapat dinotasikan dalam bentuk *functional* $T(F_{X,\epsilon})$ dan karena μ merupakan fungsi dari F_X, maka μ dapat dinotasikan dalam bentuk *functional* $T(F_X)$. Dengan demikian, persamaan (3.2) dapat ditulis menjadi

$$T(F_{X,\epsilon}) = \epsilon x + (1-\epsilon)T(F_X)$$

$$T(F_{X,\epsilon}) - T(F_X) = \epsilon x - \epsilon T(F_X)$$

Sehingga, *influence function* dari mean sampel adalah

$$
\begin{aligned}
IF(x;\bar{X}) &= \lim_{\epsilon\to 0} \frac{T(F_{X,\epsilon}) - T(F_X)}{\epsilon} \\
&= \lim_{\epsilon\to 0} \frac{\epsilon x - \epsilon T_\mu(F_X)}{\epsilon} \\
&= x - \mu. \qquad (3.3)
\end{aligned}
$$

Dari persamaan (3.3) diperoleh *influence function* dari mean sampel adalah $IF(x;\bar{X}) = x - \mu$, yang merupakan fungsi linier dari *x*. Selanjutnya akan diperlihatkan bahwa *influence function* dari mean sampel adalah fungsi yang tidak terbatas.

Misalkan $M > 0$ adalah sebuah nilai $IF(x;\bar{X})$ di titik $x = x_c$. Misalkan terdapat sebuah titik x_e dimana $x_e = x_c + 1$, maka

$$
\begin{aligned}
|IF(x_e;\bar{X})| &= |x_e - \mu| \\
&= |x_c + 1 - \mu| \\
&= |x_c - \mu + 1| \\
&= |M + 1| \\
|IF(x_e;\bar{X})| &= M + 1 > M. \qquad (3.4)
\end{aligned}
$$

Sehingga, berdasarkan definisi 3.3 diketahui bahwa *influence function* dari mean sampel adalah fungsi yang tidak terbatas. Karena $|IF(x;\bar{X})|$ merupakan fungsi yang tidak terbatas, maka berdasarkan definisi 3.2, mean sampel bukan penaksir *robust* untuk mean proses.

Telah diketahui bahwa mean sampel bukan merupakan penaksir *robust*. Selanjutnya, akan dilihat apakah median sampel pada bagan kendali MDMAD merupakan penaksir *robust* untuk mean proses.

***Influence function* dari median sampel**

Misalkan $M(F_X)$ adalah notasi *functional* untuk median (MD) dari distribusi yang diasumsikan F_X, yang didefinisikan oleh

$$\int_{-\infty}^{MD} dF_X = \frac{1}{2}.$$

Misalkan $M(F_{X,\epsilon})$ adalah notasi *functional* untuk median dari distribusi campuran $F_{X,\epsilon}$. Maka nilai median setelah ada titik yang terkontaminasi, adalah $MD + \delta$, dimana δ dapat bernilai positif atau negatif tergantung pada nilai titik yang terkontaminasi tersebut, apakah lebih besar atau lebih kecil dari MD. Selanjutnya, asumsikan bahwa F_X kontinu, dengan *pdf* f_X. Maka untuk $\delta < 0$ dapat diperoleh

$$\frac{1}{2} = (1-\epsilon)F_X(MD+\delta) + \epsilon$$

$$\frac{1}{2} = (1-\epsilon)\left[\frac{1}{2} + \delta f_X(MD)\right] + \epsilon$$

dengan demikian,

$$\frac{1}{2} = \frac{1}{2} + \frac{1}{2}\epsilon + \delta f_X(MD)$$

$$\delta = \frac{-\epsilon}{2f_X(MD)}.$$

Sebaliknya, jika $\delta > 0$, maka $\delta = \frac{\epsilon}{2f_X(MD)}$.

Oleh karena itu, *influence curve* dari median sampel adalah

$$IF(x;MD) = \lim_{\epsilon\to 0}\frac{M\left(F_{X,\epsilon}\right) - M(F_X)}{\epsilon}$$

$$= \lim_{\epsilon\to 0}\frac{\delta}{\epsilon}$$

$$= \begin{cases} \frac{1}{2f_X(MD)}\ ,\ x < MD \\ \frac{-1}{2f_X(MD)}\ ,\ x > MD \end{cases}$$

$$= \frac{sgn\ (x-MD)}{2f_X(MD)} \tag{3.5}$$

Dari persamaan (3.5), diperoleh bahwa *influence function* dari median sampel merupakan fungsi signum. Selanjutnya, akan diperlihatkan bahwa *influence function* dari median sampel merupakan fungsi yang terbatas.

Karena telah diketahui bahwa nilai terbesar dari $IF(x;MD)$ adalah sebuah konstanta $M = \frac{1}{2f_X(MD)} > 0$, maka untuk setiap *x* diperoleh

$$|IF(x;MD)| = \left|\frac{sgn\ (x-MD)}{2f_X(MD)}\right| = \frac{1}{2f_X(MD)} = M.$$

Sehingga, berdasarkan definisi 3.3, diketahui bahwa *influence function* dari median sampel merupakan fungsi yang terbatas dari *x*. Jadi, berdasarkan definisi 3.2, median sampel merupakan penaksir *robust* untuk mean proses.

***Influence function* dari standar deviasi sampel**

Misalkan *X* adalah sebuah peubah acak kontinu yang mempunyai fungsi distribusi F_X dengan mean μ dan variansi σ^2. Misalkan W peubah acak diskrit yang mempunyai fungsi distribusi $\Delta_x(w) = \begin{cases} 0, & w < x \\ 1, & w \geq x \end{cases}$, dan misalkan $I_{1-\epsilon}$ adalah sebuah peubah acak indikator yang didefinisikan sebagai berikut:

$$I_{1-\epsilon} = \begin{cases} 1 & dengan\ probabilitas\ 1-\epsilon \\ 0 & dengan\ probabilitas\ \epsilon. \end{cases}$$

Kemudian, misalkan terdapat sebuah peubah acak $Y = I_{1-\epsilon}X + [1 - I_{1-\epsilon}]W$, dimana *X*, *W* dan $I_{1-\epsilon}$ independen, maka dari persamaan (3.1) terbukti bahwa Y mempunyai distribusi campuran dengan fungsi distribusi

$$F_{X,\epsilon} = (1-\epsilon)F_X + \epsilon\Delta_X.$$

Kemudian, karena diketahui X adalah sebuah peubah acak dengan mean μ $(E(X) = \mu)$ dan variansi σ^2 $(Var(X) = \sigma^2)$, maka

$$E(X^2) = Var(X) + [E(X)]^2 = \sigma^2 + \mu^2.$$

Karena W mempunyai fungsi distribusi $\Delta_X(w)$, dimana

$\Delta_x(w) = \begin{cases} 0, & w < x \\ 1, & w \geq x \end{cases}$, maka fungsi probabilitas dari W adalah

$\zeta(w) = \begin{cases} 1, & w = x \\ 0, & lainnya \end{cases}$, sehingga

$$E(W^2) = \sum w^2 \zeta(w) = x^2 \zeta(x) = x^2.$$

Dengan demikian, variansi dari Y adalah

$$\begin{aligned} Var(Y) &= E(Y^2) - [E(Y)]^2 \\ &= [E(Y^2|I_{1-\epsilon} = 0)P(I_{1-\epsilon} = 0) + E(Y^2|I_{1-\epsilon} = 1)P(I_{1-\epsilon} = 1)] \\ &\qquad - [\epsilon x + (1-\epsilon)\mu]^2 \\ &= E(W^2)\epsilon + E(X^2)(1-\epsilon) - [\epsilon^2 x^2 + (1-\epsilon)^2\mu^2 + 2\epsilon x\mu - 2\epsilon^2 x\mu] \\ &= \epsilon x^2 + (1-\epsilon)(\sigma^2 + \mu^2) - \epsilon^2 x^2 - \mu^2 - \epsilon^2\mu^2 + 2\epsilon\mu^2 - 2\epsilon x\mu \\ &\qquad + 2\epsilon^2 x\mu \\ &= \epsilon x^2 + (1-\epsilon)\sigma^2 + \epsilon\mu^2 - \epsilon^2 x^2 - \epsilon^2\mu^2 - 2\epsilon x\mu + 2\epsilon^2 x\mu \end{aligned} \tag{3.6}$$

Kemudian, karena Var(Y) merupakan fungsi dari $F_{x,\epsilon}$, maka Var(Y) dapat dinotasikan dalam bentuk *functional* $V(F_{X,\epsilon})$. Begitu pula dengan σ^2, karena σ^2 merupakan fungsi dari F_X, maka σ^2 dapat dinotasikan dalam bentuk *functional* $V(F_X)$, sehingga Var(Y) pada persamaan (3.6) dapat ditulis menjadi

$$V(F_{x,\epsilon}) = \epsilon x^2 + (1-\epsilon)V(F_X) + \epsilon\mu^2 - \epsilon^2 x^2 - \epsilon^2\mu^2 - 2\epsilon x\mu + 2\epsilon^2 x\mu$$

$$V(F_{x,\epsilon}) - V(F_X) = \epsilon x^2 - \epsilon V(F_X) + \epsilon\mu^2 - \epsilon^2 x^2 - \epsilon^2\mu^2 - 2\epsilon x\mu + 2\epsilon^2 x\mu$$

Jadi, *influence function* dari variansi sampel adalah

$$\begin{aligned} IF(x; S^2) &= \lim_{\epsilon\to 0} \frac{V(F_{x,\epsilon}) - V(F_X)}{\epsilon} \\ &= \lim_{\epsilon\to 0} \frac{\epsilon x^2 - \epsilon V(F_X) + \epsilon\mu^2 - \epsilon^2 x^2 - \epsilon^2\mu^2 - 2\epsilon x\mu + 2\epsilon^2 x\mu}{\epsilon} \\ &= \lim_{\epsilon\to 0} \frac{\epsilon(x^2 - 2x\mu + \mu^2) - \epsilon V(F_X) - \epsilon^2(x^2 - 2\epsilon^2 x\mu + \mu^2)}{\epsilon} \\ &= \lim_{\epsilon\to 0} \frac{\epsilon(x-\mu)^2 - \epsilon\sigma^2 - \epsilon^2(x^2 - 2x\mu + \mu^2)}{\epsilon} \\ &= (x-\mu)^2 - \sigma^2. \end{aligned} \tag{3.7}$$

Dari persamaan (3.7), diketahui bahwa *influence function* dari variansi sampel merupakan fungsi kuadrat dari *x*. Selanjutnya, akan ditunjukkan bahwa *influence function* dari variansi sampel merupakan fungsi yang tidak terbatas dari x.

Misalkan $M = (x_c - \mu)^2 - \sigma^2 > 0$ adalah sebuah nilai $IF(x; S^2)$ di titik $x = x_c$. Misalkan terdapat sebuah titik x_e dimana $x_e = x_c + 1$, maka

$$\begin{aligned}
|IF(x_e; S^2)| &= |(x_e - \mu)^2 - \sigma^2| \\
&= |(x_c + 1 - \mu)^2 - \sigma^2| \\
&= \left|\left((x_c - \mu) + 1\right)^2 - \sigma^2\right| \\
&= |(x_c - \mu)^2 + 2(x_c - \mu) + 1 - \sigma^2| \\
&= |(x_c - \mu)^2 - \sigma^2 + 2(x_c - \mu) + 1| \\
&= |M + 2(x_c - \mu) + 1| \\
&= M + 2|x_c - \mu| + 1 > M
\end{aligned} \tag{3.8}$$

Sehingga, berdasarkan definisi 3.3 diketahui bahwa *influence function* dari variansi sampel adalah fungsi yang tidak terbatas. Karena $|IF(x; S^2)|$ merupakan fungsi yang tidak terbatas, maka berdasarkan definisi 3.2, variansi sampel bukan penaksir *robust*. Dengan demikian, standar deviasi sampel bukan merupakan penaksir *robust* untuk standar deviasi proses.

Telah diketahui bahwa standar deviasi sampel bukan merupakan penaksir *robust.* Selanjutnya, akan dilihat apakah *median absolute deviation* sampel pada bagan kendali MDMAD merupakan penaksir *robust* untuk standar deviasi proses.

***Influence function* dari *median absolute deviation* sampel**

Dari subbab 2.14.2 diketahui bahwa *median absolute deviation* (MAD) merupakan penaksir-M yang diperoleh dengan mencari solusi dari persamaan

$$\sum_{i=1}^{n} \psi\left(\frac{x_i - \theta}{d}\right) = 0 \tag{3.9}$$

dimana θ adalah parameter lokasi dan d adalah parameter skala.

Dari Huber (1981) halaman 136, diketahui bahwa *influence function* dari *median absolute deviation* sampel diperoleh dengan mensubtitusikan $F_{X,\epsilon} = (1-\epsilon)F_X + \epsilon\Delta_X$ ke dalam persamaan berikut:

$$\int \psi\left(\frac{x_i - T(F)}{S(F)}\right) F dx = 0 \tag{3.10}$$

dimana, $T(F)$ adalah penaksir untuk parameter lokasi θ yang dinotasikan dalam bentuk *functional*, dan $S(F)$ adalah penaksir untuk parameter lokasi d yang dinotasikan dalam bentuk *functional*, sedangkan ψ adalah fungsi Huber untuk menaksir parameter skala, seperti yang telah dijelaskan pada subbab 2.14.2.

Menurut Huber (1981) halaman 137, turunan persamaan (3.10) terhadap ϵ adalah

$$IF\big(x; T(F)\big) \int \psi'\left(\frac{x_i - T(F)}{S(F)}\right) F(dx) + IF\big(x; S(F)\big) \int \psi'\left(\frac{x_i - T(F)}{S(F)}\right)\left(\frac{x_i - T(F)}{S(F)}\right) F(dx) = \psi\left(\frac{x_i - T(F)}{S(F)}\right) S(F) \tag{3.11}$$

Dari persamaan (3.11), *influence function* dari penaksir parameter skala *S(F)* adalah

$$IF\big(x; S(F)\big) = \frac{\psi\left(\frac{x}{S(F)}\right) S(F)}{\int \psi'\left(\frac{x}{S(F)}\right)\left(\frac{x}{S(F)}\right) F(dx)}. \tag{3.12}$$

Karena fungsi Huber untuk MAD adalah $\psi(x) = sgn(|x| - 1)$, maka dari persamaan (3.12) diperoleh

$$IF(x; MAD) = \frac{\psi\left(\frac{x}{MAD}\right) MAD}{\int \psi'\left(\frac{x}{MAD}\right)\left(\frac{x}{MAD}\right) F(dx)}$$

$$= \frac{sgn\left(\left|\frac{x}{MAD}\right| - 1\right) MAD}{\int \psi'\left(\frac{x}{MAD}\right)\left(\frac{x}{MAD}\right) F(dx)}$$

Sehingga, *influence function* dari MAD adalah

$$IF(x; MAD) = \frac{sgn(|x| - MAD)}{4f(MAD)}. \tag{3.13}$$

(Huber, halaman 137-138)

Dari persamaan (3.13) diperoleh bahwa *influence function* dari *median absolute deviation* sampel adalah fungsi signum. Sehingga $|IF(x; MAD)|$ merupakan fungsi yang terbatas dari *x*. Jadi, berdasarkan definisi 3.2, *median absolute deviation* sampel merupakan penaksir *robust* untuk standar deviasi proses.

Setelah diketahui bahwa median sampel dan *median absolute deviation* sampel merupakan penaksir yang *robust* untuk menaksir mean proses dan standar deviasi proses, maka selanjutnya akan dibahas prosedur pembuatan bagan kendali berdasarkan median dan *median absolute deviation*.

3.2 Bagan Kendali $\bar{X}_{MDMAD}$

Bagan kendali $\bar{X}_{MDMAD}$ adalah bagan kendali yang digunakan untuk mengawasi mean proses. Bagan kendali $\bar{X}_{MDMAD}$ ini dibuat dengan menggunakan penaksir-penaksir robust, yaitu median sampel (*MD*) dan *median absolute deviation* (*MAD*). Median sampel (*MD*) adalah penaksir alternatif dari mean sampel ($\bar{X}$) yang digunakan untuk menaksir mean proses (μ), sedangkan *median absolute deviation* (*MAD*) merupakan penaksir alternatif dari standar deviasi sampel (*S*) yang digunakan untuk menaksir standar deviasi proses (σ).

Misalkan $x_{i1}, x_{i2}, \ldots, x_{in}$ dimana *i* = 1, 2, ..., *m* adalah sampel yang berasal dari *m* buah subgrup berukuran *n*, maka median subgrup ke-*i* (MD_i) didefinisikan sebagai berikut:

$$MD_i = \begin{cases} X_{\left(\frac{n+1}{2}\right)} & , jika\ n\ genap \\ \dfrac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2} & , jika\ n\ ganjil \end{cases} \tag{3.14}$$

sedangkan *median absolute deviation* subgrup ke-*i* (MAD_i) didefinisikan sebagai median dari penyimpangan mutlak dari MD_i sebagai berikut:

$MAD_i = median\left|x_{ij} - MD_i\right|$, dimana *i* = 1, 2, ..., *m* dan *j* = 1, 2, ..., *n*.

Kemudian, kedua penaksir ini (*MD_i* dan *MAD_i*) akan digunakan untuk membuat bagan kendali $\bar{X}_{MDMAD}$ dengan mengikutsertakannya dalam perhitungan batas-batas kendali dan garis tengah dari bagan kendali $\bar{X}_{MDMAD}$ sebagai berikut:

1. Mean proses μ ditaksir dengan rata-rata dari *m* subgrup median, yaitu

$$\hat{\mu} = \overline{MD} = \sum_{i=1}^{m} MD_i/m$$

2. Standar deviasi proses ditaksir dengan rata-rata dari *m* subgrup *median absolute deviation*, yaitu

$$\hat{\sigma} = b_n\overline{MAD} = b_n \sum_{i=1}^{m} MAD_i/m$$

3. Batas kendali bawah (*LCL*), bagan kendali atas (*UCL*) dan garis tengah (*CL*) dihitung dengan menggunakan *m* buah subgrup berukuran *n* yang didapat dari proses yang terkendali (*in control process*) sebagai berikut:

$$LCL = \hat{\mu} - Z_{\frac{\alpha}{2}}\hat{\sigma}_{MD} = \overline{MD} - 3\hat{\sigma}_{MD} = \overline{MD} - 3\frac{1.253\hat{\sigma}}{\sqrt{n}}$$

$$= \overline{MD} - 3\frac{1.253b_n\overline{MAD}}{\sqrt{n}} = \overline{MD} - R_1\overline{MAD}$$

$$CL = \hat{\mu} = \overline{MD}$$

$$UCL = \hat{\mu} + Z_{\frac{\alpha}{2}}\hat{\sigma}_{MD} = \overline{MD} + 3\hat{\sigma}_{MD} = \overline{MD} + 3\frac{1.253\hat{\sigma}}{\sqrt{n}}$$

$$= \overline{MD} + 3\frac{1.253b_n\overline{MAD}}{\sqrt{n}} = \overline{MD} + R_1\overline{MAD}$$

dimana $R_1 = 3.759b_n/\sqrt{n}$. Nilai-nilai dari faktor koreksi b_n dan faktor batas kendali R_i ada pada tabel 3.1.

Tabel 3.1 Nilai-nilai untuk faktor batas kendali R_1

n	b_n	R_1	n	b_n	R_1
2	1.196	3.17899	14	1.061	1.06592
3	1.495	3.24454	15	1.056	1.02492
4	1.363	2.56176	16	1.053	0.98956
5	1.206	2.02738	17	1.049	0.95636
6	1.200	1.84153	18	1.047	0.92765
7	1.140	1.61968	19	1.044	0.90032
8	1.129	1.50045	20	1.042	0.87584
9	1.107	1.38707	21	1.040	0.85309
10	1.087	1.29212	22	1.038	0.83188
11	1.078	1.22178	23	1.036	0.81202
12	1.071	1.16217	24	1.034	0.79339
13	1.066	1.11137	25	1.033	0.77661

Sementara itu, untuk $n > 25$, nilai faktor koreksi b_n diperoleh dari formula $b_n = \frac{n}{n-0.8}$, sedangkan nilai faktor batas kendali R_1 diperoleh dengan menggunakan formula R_1, yaitu $R_1 = 3.759 b_n/\sqrt{n}$.

4. Selanjutnya, nilai mean dari setiap subgrup $(\bar{X}_i)$ dimana $i = 1, 2, \ldots, m$ diplot pada bagan $\bar{X}_{MDMAD}$. Oleh karena itu, bagan kendali $\bar{X}_{MDMAD}$ dapat digunakan untuk mengawasi mean proses karena mean sampel diplot pada bagan ini sesuai dengan pernyataan Rocke (1989), bahwa "prosedur terbaik adalah untuk memplot mean sampel tetapi dengan menggunakan batas-batas kendali yang dihitung berdasarkan penaksir *robust*". Mean sampel dari setiap subgrup baik digunakan untuk mendeteksi keberadaan outlier atau adanya pergeseran dari mean proses.
5. Jika ada satu nilai $\bar{X}_i$, $i = 1, 2, \ldots, m$ yang terletak di luar batas-batas kendali, maka proses produksi yang diamati berada di luar kendali (*out of control process*).

3.3 Bagan Kendali S_{MDMAD}

Bagan kendali S_{MDMAD} adalah bagan kendali yang digunakan untuk mengawasi standar deviasi proses. Bagan kendali S_{MDMAD} ini tetap memplot standar deviasi sampel S_i , namun menggunakan penaksir *MAD* untuk mendapatkan batas-batas kendali dan garis tengahnya. Oleh karena itu, batas-batas kendali (*LCL* dan *UCL*) dan garis tengah (*CL*) dari bagan ini diperoleh sebagai berikut:

1. Batas kendali bawah (*LCL*), batas kendali atas (*UCL)* dan garis tengah (*CL*) dihitung dengan menggunakan *MAD* dan faktor koreksi b_n , serta faktor batas kendali c_4, B_5 dan B_6 sebagai berikut:

$$LCL = c_4\hat{\sigma} - 3\hat{\sigma}\sqrt{1 - c_4^2} = c_4 b_n\overline{MAD} - 3b_n\overline{MAD}\sqrt{1 - c_4^2}$$

$$= B_5 b_n\overline{MAD} = B_5^*\overline{MAD}$$

$$CL \quad = c_4\hat{\sigma} = c_4 b_n\overline{MAD} = c_4^*\overline{MAD}$$

$$UCL = c_4\hat{\sigma} + 3\hat{\sigma}\sqrt{1 - c_4^2} = c_4 b_n\overline{MAD} + 3b_n\overline{MAD}\sqrt{1 - c_4^2}$$

$$= B_5 b_n\overline{MAD} = B_5^*\overline{MAD}$$

dimana

$$c_4 = \frac{E(\bar{S})}{\sigma} \approx \frac{4(n-1)}{4n-3} \; , c_4^* = b_n c_4$$

$$B_5^* = b_n B_5 = b_n\left(c_4 - 3\sqrt{1 - c_4^2}\right)$$

$$B_6^* = b_n B_6 = b_n\left(c_4 + 3\sqrt{1 - c_4^2}\right)$$

Nilai-nilai dari faktor-faktor batas kendali c_4^*, B_5^* dan B_6^* dihitung dan diberikan pada tabel 3.2.

Tabel 3.2 Faktor-faktor batas kendali untuk bagan kendali *robust* dengan peubah univariat

n	b_n	c_4*	B_5*	B_6*	n	b_n	c_4*	B_5*	B_6*
2	1.196	0.954	0.000	3.117	14	1.061	1.041	0.423	1.658
3	1.495	1.325	0.000	3.403	15	1.056	1.037	0.445	1.630
4	1.363	1.256	0.000	2.846	16	1.053	1.036	0.463	1.607
5	1.206	1.134	0.000	2.369	17	1.049	1.033	0.480	1.585
6	1.200	1.142	0.035	2.249	18	1.047	1.032	0.497	1.566
7	1.140	1.094	0.129	2.059	19	1.044	1.030	0.512	1.548
8	1.129	1.089	0.202	1.977	20	1.042	1.028	0.525	1.532
9	1.107	1.073	0.257	1.890	21	1.040	1.027	0.537	1.517
10	1.087	1.057	0.300	1.814	22	1.038	1.026	0.548	1.503
11	1.078	1.051	0.337	1.765	23	1.036	1.024	0.558	1.490
12	1.071	1.047	0.371	1.724	24	1.034	1.023	0.568	1.478
13	1.066	1.044	0.399	1.690	25	1.033	1.022	0.577	1.467

2. Selanjutnya, nilai-nilai standar deviasi dari setiap subgroup S_i , i = 1,2,...,m diplot pada bagan S_{MDMAD}. Oleh karena itu, bagan kendali S_{MDMAD} ini dapat digunakan untuk mengawasi variabilitas proses.
3. Jika terdapat S_i yang jatuh di luar batas-batas kendali, maka proses produksi yang diamati dapat dikatakan tidak terkendali (*out of control*).

BAB 4

CONTOH PENERAPAN

Dalam bab ini, akan dibahas penerapan bagan kendali MDMAD pada data. Kemudian akan dijelaskan interpretasi dari bagan kendali yang diperoleh. Selain itu, akan dilihat juga bagan kendali Shewhart yang dibentuk dengan menggunakan data yang sama, dan hasil dari bagan kendali MDMAD akan dibandingkan dengan hasil dari bagan kendali Shewhart.

4.1 Data

Data yang digunakan dalam skripsi ini adalah data simulasi yang berasal dari jurnal A robust Control Chart Based on Median Absolute Deviation yang ditulis oleh Moustafa Omar pada tahun 2008. Data tersebut berupa 300 observasi yang berasal dari distribusi Normal Standar ($N(0,1)$), yang terdiri dari m = 30 subgrup berukuran n = 10. Data ini terlampir pada tabel 4.1.

Dari data tersebut, akan dibuat bagan kendali Shewhart untuk mengawasi mean proses (bagan kendali $\bar{X}_{Shewhart}$) dan bagan kendali Shewhart untuk mengawasi variabilitas proses (bagan kendali $S_{Shewhart}$).

4.2 Pembuatan Bagan Kendali Shewhart

Karena mean proses (μ) dan standar deviasi proses (σ) tidak diketahui, maka penaksir untuk μ dalam konstruksi bagan kendali Shewhart adalah $\bar{X}$, dan penaksir *unbiased* untuk σ adalah $\frac{\bar{S}}{c_4}$.

Dari data pada tabel 4.1, dapat diperoleh mean dan standar deviasi dari masing-masing subgrup seperti yang diperlihatkan pada tabel 4.2.

Tabel 4.1 Data simulasi dari distribusi *N*(0,1) dimana *m*=30 dan *n*=1

i	X_{i1}	X_{i2}	X_{i3}	X_{i4}	X_{i5}	X_{i6}	X_{i7}	X_{i8}	X_{i9}	X_{i10}
1	1.4835	0.0376	0.0722	-1.8837	-0.7156	0.3126	0.1178	0.3896	-1.0478	1.4713
2	-2.2635	0.8896	-0.5676	0.2314	-0.3133	-1.0031	-0.7193	0.3692	0.207	-0.6791
3	-0.7842	-0.1933	0.3397	0.6804	1.8069	0.3704	-0.6364	-1.6107	-0.3319	-0.4416
4	1.8807	1.9536	-0.2151	-1.989	1.1574	0.1117	-0.618	-0.0344	0.1037	1.2879
5	-0.2782	0.0261	-1.2148	-0.3072	0.8678	-0.5908	0.9473	-1.0318	1.0879	0.9132
6	-1.7016	0.7459	0.1955	0.3813	1.6286	-1.4217	-0.1743	0.0492	-0.1788	-1.0088
7	0.5299	-2.6118	-2.4908	1.0802	1.7977	-0.1522	-2.277	2.2526	0.689	-1.4098
8	-0.2373	-0.8532	-0.4249	0.468	-0.5686	0.375	1.6013	0.0354	1.0108	-1.3431
9	1.0986	0.3344	-0.565	-0.1682	0.4486	-0.5084	0.4823	0.9617	-1.3246	0.4549
10	0.9639	0.3666	0.3424	-1.3559	-1.0383	-1.0383	-1.5773	0.4897	0.3031	1.9253
11	0.8036	-0.2394	0.9071	1.305	0.8026	0.0705	0.8722	0.4805	2.0662	-1.8403
12	0.2451	0.8715	-0.7216	0.3562	0.3337	-0.0436	-1.0722	2.0567	0.8635	-0.6404
13	-1.1704	-0.5599	0.2078	0.1277	1.3502	0.3061	1.7752	0.1702	1.287	-0.1969
14	-0.041	-0.9462	0.3148	0.1625	-0.8268	0.7635	-1.3772	1.1971	-0.4499	0.9468
15	0.4757	-0.1437	-1.7909	0.7318	-1.2359	-1.5585	-0.0528	0.7145	-0.467	1.4883
16	-0.0088	-0.5943	0.0364	1.2101	0.0954	-1.6292	-0.9782	0.0266	0.4573	-1.0764
17	-0.3347	0.3659	-0.0693	-0.3473	0.1825	-0.7915	-0.5884	0.4206	1.345	-0.5754
18	-1.1306	-0.8849	1.4665	0.6482	0.4037	0.5172	-0.0659	0.3637	-0.0072	0.3601
19	-1.1953	0.5811	-0.6265	1.6277	-0.2339	1.8788	-0.0306	-0.7576	-0.6517	-1.302
20	-0.3743	0.3014	0.3535	0.033	-1.3977	0.8398	1.046	0.2259	-0.0296	-0.1785
21	-2.443	0.0219	-0.6683	0.0772	-0.1099	-0.9951	0.1513	-0.3407	1.111	-1.0888
22	-0.3248	-0.1301	1.2401	0.0311	-0.7363	-0.3991	0.8195	-0.817	-0.0088	1.1887
23	0.6001	-0.6047	-0.5429	-1.0755	0.3088	1.1351	0.1287	-1.5995	1.4185	-0.2448
24	-2.247	-0.2824	0.4995	0.3882	-1.4335	-0.0054	-0.3867	-1.2682	0.7874	-0.944
25	-1.3835	1.3497	-0.2675	0.0872	-0.2292	-0.671	-1.1851	-0.9307	-0.8666	-0.0157
26	-0.1835	0.9426	0.59	0.0604	-0.2276	1.264	1.8589	-0.6672	1.4359	-0.2702
27	0.6427	-0.7793	-2.3813	0.9137	-0.9337	-1.4447	0.2132	-0.235	0.1888	0.9655
28	0.9965	0.1248	0.7862	-0.5655	1.5301	1.109	0.5112	-0.0067	0.8158	0.0268
29	-0.8081	0.834	-0.8316	-0.5117	0.8855	0.7094	-0.4049	-1.5813	-0.6667	0.0083
30	1.8602	-0.1939	-1.0404	-2.4239	1.1495	-1.2498	-0.0945	2.3545	1.7392	-0.3095

Tabel 4.2 Mean dan standar deviasi dari masing-masing subgrup

Subgrup (i)	Mean Subgrup ($\bar{X}_i$)	Standar deviasi subgrup ke-i (S_i)
1	0.02375	1.04287
2	-0.38487	0.8864
3	-0.08007	0.93531
4	0.36385	1.22187
5	0.04195	0.8636
6	-0.14847	1.00946
7	-0.25922	1.81895
8	0.00634	0.88433
9	0.12143	0.75259
10	-0.06188	1.13839
11	0.5228	1.04309
12	0.22489	0.91954
13	0.3297	0.9097
14	-0.02564	0.86352
15	-0.18385	1.08293
16	-0.24611	0.83013
17	-0.03926	0.63935
18	0.16708	0.74906
19	-0.071	1.10742
20	0.08195	0.67713
21	-0.42844	0.95142
22	0.08633	0.74835
23	-0.04762	0.95367
24	-0.48921	0.96952
25	-0.41124	0.79221
26	0.48033	0.8597
27	-0.28501	1.09245
28	0.53282	0.633
29	-0.23671	0.82584
30	0.17914	1.55246
	$\bar{\bar{X}}$= -0.0079	$\bar{S}$= 0.9585

4.2.1 Bagan Kendali $\bar{X}_{Shewhart}$

Sebelum menghitung batas-batas kendali untuk bagan kendali $\bar{X}_{Shewhart}$, perlu diketahui faktor A_3 untuk bagan kendali ini. Karena jumlah observasi dalam setiap subgrup adalah n=10, maka didapat nilai faktor A_3=0.98.

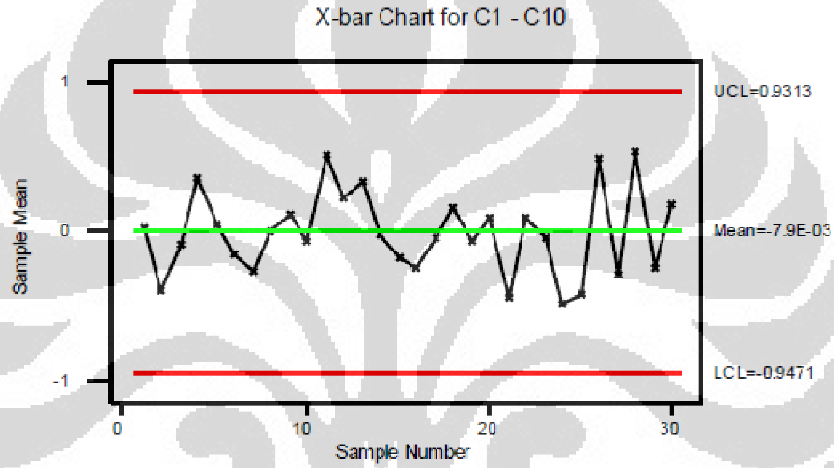
a) Menentukan batas-batas kendali

$CL = \bar{X} = -0.0079$

$LCL = \bar{X} - A_3\bar{S} = -0.0079 - (0.98 * 0.9585) = -0.94718$

$UCL = \bar{X} + A_3\bar{S} = -0.0079 + (0.98 * 0.9585) = 0.93143$

b) Memplot mean dari setiap subgrup pada bagan kendali $\bar{X}_{Shewhart}$



Gambar 4.1 Bagan kendali $\bar{X}_{Shewhart}$

4.2.2 Bagan Kendali $S_{Shewhart}$

Sebelum menghitung batas-batas kendali untuk bagan kendali $S_{Shewhart}$, perlu diketahui faktor B_3 dan B_4 untuk bagan kendali ini. Karena jumlah observasi dalam setiap subgrup adalah n=10, maka berdasarkan tabel didapat nilai faktor B_3=0.28 dan B_4=1.72.

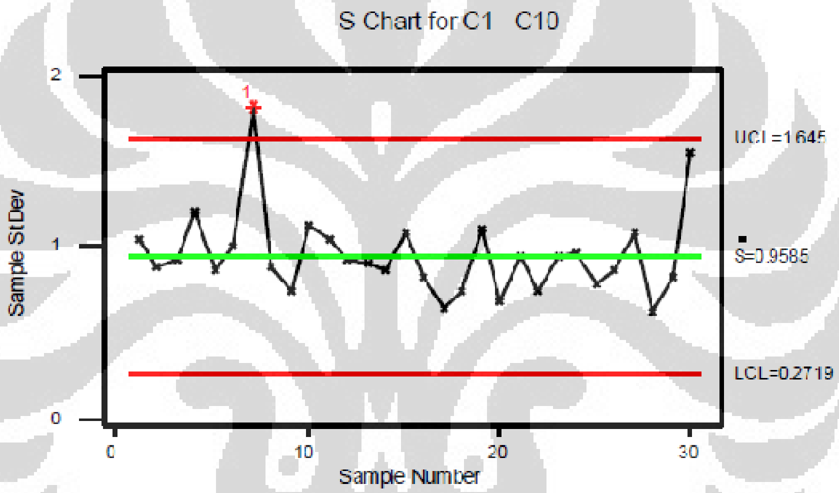
a) Menentukan batas-batas kendali

$CL = \bar{S} = 0.9585$

$LCL = B_3\bar{S} = (0.28 * 0.9585) = 0.2719$

$UCL = B_4\bar{S} = (1.72 * 0.9585) = 1.645$

b) Memplot standar deviasi dari setiap subgrup pada bagan kendali $S_{Shewhart}$



Gambar 4.2 Bagan kendali $S_{Shewhart}$

4.3 Pembuatan Bagan Kendali MDMAD

Karena mean proses (μ) dan standar deviasi proses (σ) tidak diketahui, maka

a) Mean proses (μ) ditaksir dengan rata-rata dari *m* subgrup median (MD_i).

b) Standar deviasi proses (σ) ditaksir dengan rata-rata dari *m* subgrup MAD (MAD_i).

Tabel 4.3 Median dan median absolute deviation dari masing-masing subgrup

Subgrup (i)	Median subgroup ke-i (MD_i)	MAD subrup ke-i (MAD_i)
1	0.095	0.819285
2	-0.4405	0.897047
3	-0.2626	0.833147
4	0.1077	1.316104
5	-0.1261	1.408173
6	-0.0626	0.92833
7	0.18885	2.37772
8	-0.101	0.774584
9	0.3915	0.837595
10	0.32275	1.484231
11	0.8031	0.611202
12	0.2894	0.857091
13	0.189	0.841227
14	0.06075	1.177777
15	-0.0983	1.217808
16	0.0089	0.779551
17	-0.202	0.571468
18	0.3619	0.485848
19	-0.4302	0.863392
20	0.12945	0.394372
21	-0.2253	0.607569
22	-0.0695	0.738705
23	-0.0581	0.893118
24	-0.3346	1.154056
25	-0.4693	0.67829
26	0.3252	0.899049
27	-0.0231	1.235599
28	0.6487	0.729587
29	-0.4583	0.622618
30	-0.1442	1.778601
	$\overline{MD}$= 0.013898	$\overline{MAD}$= 0.960438

Setelah penaksir untuk mean proses dan standar deviasi proses didapatkan, selanjutnya akan dihitung batas-batas kendali untuk bagan kendali MDMAD.

4.3.1 Bagan Kendali $\bar{X}_{MDMAD}$

Sebelum menghitung batas-batas kendali pada bagan kendali $\bar{X}_{MDMAD}$, perlu diketahui terlebih dahulu faktor batas kendali R_1-nya. Oleh karena data yang digunakan merupakan subgrup berukuran n=10, maka dari tabel 3.1 didapat faktor batas kendalinya adalah $R_1 = 1.29212$.

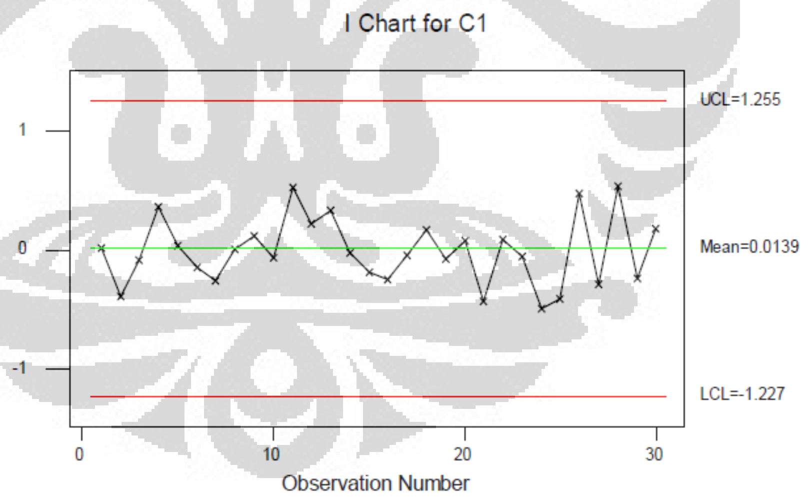
a) Menentukan batas-batas kendali bagan $\bar{X}_{MDMAD}$

$$CL = \overline{MD} = 0.013898$$

$$LCL = \overline{MD} - R_1\overline{MAD} = 0.013898 - (1.29212 * 0.960438)$$
$$= -1.22710$$

$$UCL = \overline{MD} + R_1\overline{MAD} = 0.013898 + (1.29212 * 0.960438)$$
$$= 1.25489.$$

b) Memplot mean dari setiap subgrup pada bagan kendali $\bar{X}_{MDMAD}$



Gambar 4.3 Bagan kendali $\bar{X}_{MDMAD}$

4.3.2 Bagan Kendali S_{MDMAD}

Sebelum menghitung batas-batas kendali pada bagan kendali S_{MDMAD}, perlu diketahui terlebih dahulu faktor-faktor batas kendali c_4^*, B_5^*, dan B_6^*-nya. Karena data yang digunakan merupakan subgrup berukuran 10 (n=10), maka dari tabel 3.2 didapat factor-faktor batas kendalinya adalah $c_4^* = 1.057, B_5^* = 0.300$, dan $B_6^* = 1.814$.

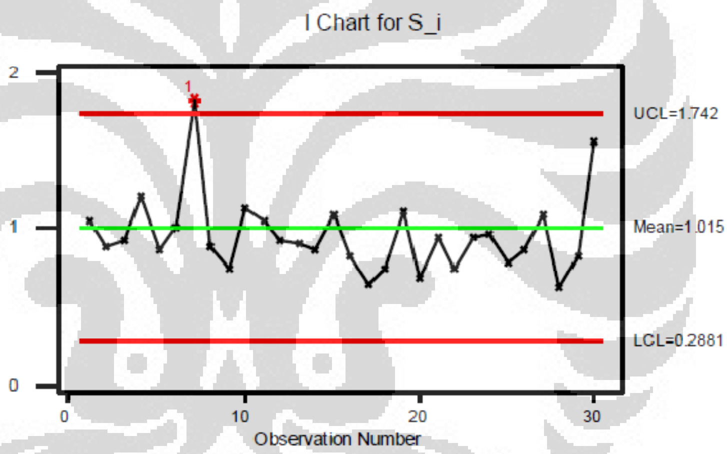
a) Menentukan batas-batas kendali bagan S_{MDMAD}

$CL = c_4^*\overline{MAD} = 1.057 * 0.960438 = 1.015183$

$LCL = B_5^*\overline{MAD} = 0.300 * 0.960438 = 0.288131$

$UCL = B_6^*\overline{MAD} = 1.814 * 0.960438 = 1.742235.$

b) Memplot standar deviasi dari setiap subgrup pada bagan kendali S_{MDMAD}



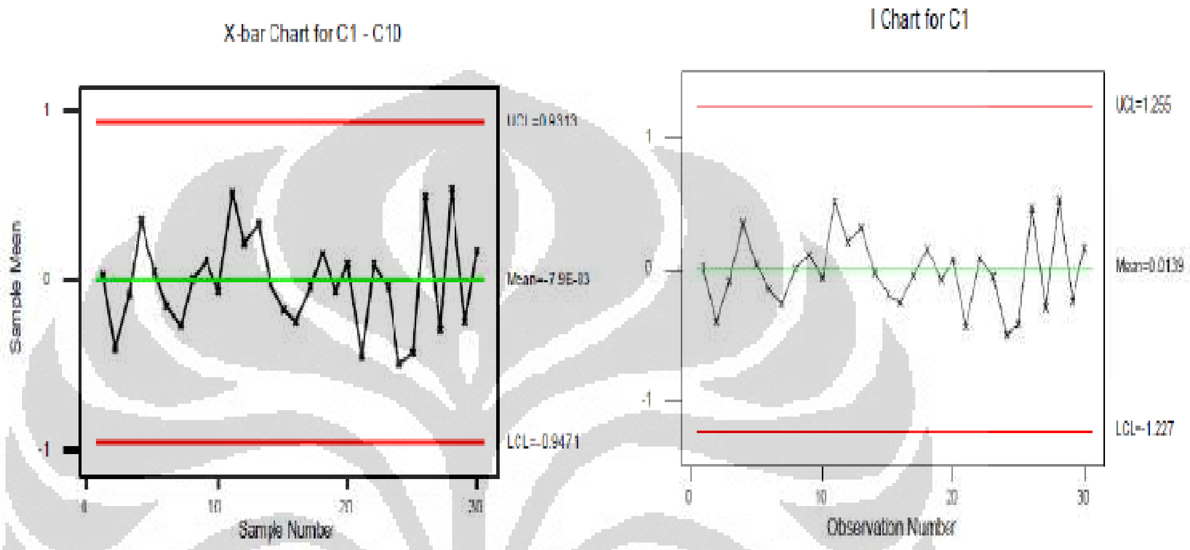
Gambar 4.4 Bagan kendali S_{MDMAD}

4.4 Analisa

Pada subbab ini, akan dibandingkan hasil dari bagan kendali Shewhart dan bagan kendali MDMAD dalam mengawasi mean proses dan variabilitas proses.

4.4.1 Bagan Kendali Untuk Mengawasi Mean Proses

Bagan kendali yang dapat digunakan untuk mengawasi mean proses adalah bagan kendali $\bar{X}_{Shewhart}$ dan $\bar{X}_{MDMAD}$.



(a) Bagan kendali $\bar{X}_{Shewhart}$ (b)Bagan kendali $\bar{X}_{MDMAD}$

Gambar 4.5 Bagan kendali untuk mengawasi mean proses

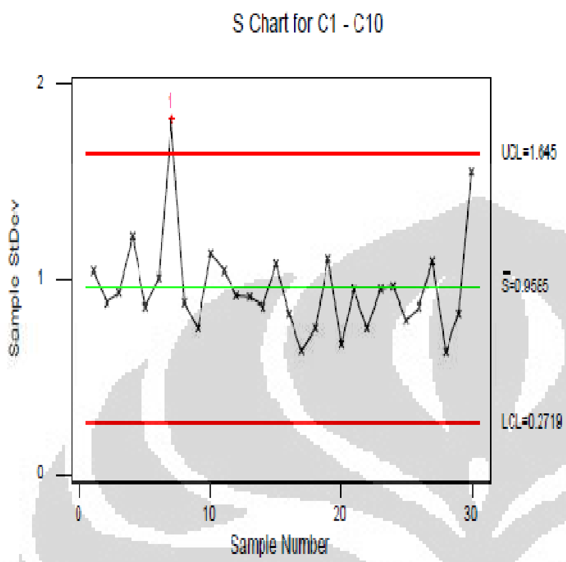
Analisa:

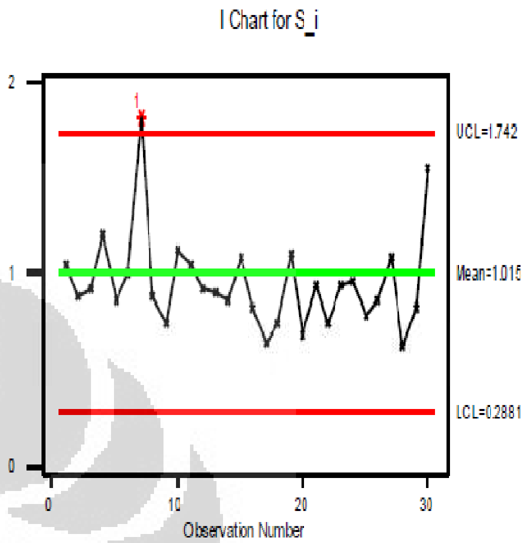
Dari gambar (a) bagan kendali $\bar{X}_{Shewhart}$ di atas, terlihat bahwa tidak terdapat sinyal *out of control process* yang muncul. Dengan demikian, kesimpulan yang dapat diambil dari bagan kendali $\bar{X}_{Shewhart}$ ini adalah mean proses terkendali secara statistik.

Salah satu bagan kendali kendali robust yang dapat digunakan untuk mengawasi mean proses adalah bagan kendali $\bar{X}_{MDMAD}$. Sinyal *out of control process* pada bagan kendali $\bar{X}_{MDMAD}$ hanya ditandai dengan adanya satu atau lebih titik yang diplot di luar batas kendali (UCL dan LCL). Dari gambar (b) bagan kendali $\bar{X}_{MDMAD}$ di atas, terlihat bahwa tidak ada titik mean subgrup yang terletak di luar bagan kendali. Sehingga kesimpulan yang dapat diambil adalah mean proses terkendali secara statistik.

4.4.2 Bagan kendali untuk mengawasi variabilitas proses

Bagan kendali yang dapat digunakan untuk mengawasi variabilitas proses adalah bagan kendali $S_{Shewhart}$ dan S_{MDMAD}.





(a) Bagan Kendali $S_{Shewhart}$ (b) Bagan Kendali S_{MDMAD}

Gambar 4.6 Bagan kendali untuk mengawasi standar deviasi proses

Analisa:

Dari gambar (a) bagan kendali $S_{Shewhart}$ di atas, terlihat bahwa terdapat satu titik standar deviasi subgrup yang terletak di luar batas kendali, yaitu S_7. Selain itu, terdapat juga sinyal *out of control process* berupa satu titik yang terletak di dekat batas kendali, yaitu S_{30}. Sehingga, dapat ditarik kesimpulan bahwa variabilitas proses berada di luar kendali secara statistik.

Kesimpulan yang sama pun dihasilkan dengan mengamati gambar (b) bagan kendali S_{MDMAD}. Karena dari gambar tersebut terlihat bahwa terdapat satu titik standar deviasi subgrup yang terplot di luar batas kendali, yaitu S_7, maka hal ini mengindikasikan bahwa variabilitas proses berada di luar kendali secara statistik.

Jika dilihat lebih teliti, lebar bagan kendali Shewhart lebih ketat bila dibandingkan dengan bagan kendali MDMAD. Hal ini wajar, karena bagan kendali Shewhart dibuat dengan menggunakan penaksir-penaksir yang tidak *robust* terhadap *outlier,* sehingga batas-batas kendali dari bagan tersebut menjadi tidak valid. Akibatnya, jika ada titik yang terplot di luar bagan kendali Shewhart (mengindikasikan tidak terkendali secara statistik), maka belum tentu titik tersebut merupakan sinyal *out of control process* pada bagan kendali MDMAD.

BAB 5

PENUTUP

5.1 Kesimpulan

Berdasarkan pembahasan pada bab-bab sebelumnya, dapat ditarik kesimpulan bahwa:

1. Jika terdapat *outlier* pada rata-rata karakteristik mutu yang diplot pada bagan kendali Shewhart, dimana *outlier* tersebut bukan merupakan kesalahan, maka dapat dicurigai bahwa data sampel yang diambil dari proses produksi tersebut telah terkontaminasi. Dengan demikian dapat dikatakan bahwa telah terjadi sedikit penyimpangan asumsi kenormalan pada data yang digunakan.
2. Ketika asumsi distribusi normal pada data untuk bagan kendali Shewhart menyimpang, maka hasil taksiran yang diberikan oleh mean sampel dan standar deviasi sampel menjadi tidak valid untuk digunakan sebagai penaksir dari mean proses dan standar deviasi proses. Hal ini terjadi karena mean sampel dan standar deviasi sampel bukan merupakan penaksir *robust*.
3. Kekokohan suatu penaksir dapat diamati melalui *influence function*-nya. Penaksir *robust* mempunyai *influence function* yang terbatas. Pada bab 3 skripsi ini, diperoleh kesimpulan bahwa median sampel merupakan penaksir *robust* untuk mean proses, sedangkan penaksir *robust* untuk standar deviasi proses adalah *median absolute deviation* sampel.
4. Karena kekokohannya terhadap *outlier,* median sampel dan *median absolute deviation* sampel dapat digunakan untuk membuat bagan kendali alternatif yang lebih *robust* daripada bagan kendali Shewhart, yaitu bagan kendali median dan *median absolute deviation* (MDMAD). Bagan kendali ini terdiri dari bagan kendali $\bar{X}_{MDMAD}$ untuk mengawasi mean proses dan bagan kendali S_{MDMAD} untuk mengawasi standar deviasi proses.

5. Dari contoh penerapan, dapat ditarik kesimpulan bahwa bagan kendali MDMAD memberikan hasil yang lebih valid dan dapat menjelaskan variabilitas mean proses dan standar deviasi proses dengan lebih baik.

5.2 Saran

Saran untuk pengembangan skripsi ini adalah

1. Kelebihan dari median sampel dan *median absolute deviation* sampel sebagai penaksir *robust* untuk mean proses dan standar deviasi proses juga dapat dilihat melalui kriteria-kriteria penaksir *robust* seperti efisiensi penaksir yang lebih tinggi, *breakdown point* penaksir yang besar dan *high error sensitivity* yang kecil bila dibandingkan dengan mean sampel dan standar deviasi sampel.
2. Kelebihan bagan kendali MDMAD dibandingkan bagan kendali Shewhart dapat dilihat juga melalui perbandingan *average run length* (ARL) dari tiap bagan.
3. Pembuatan bagan kendali alternatif untuk bagan kendali Shewhart dapat menggunakan metode *L-estimator* atau *R-estimator*.

DAFTAR PUSTAKA

Abu-Shawiesh, Moustafa Omar Ahmed. 2008. *A Control Chart Based on Robust Estimators for Monitoring The Process Mean of A Quality Characteristics.http://www.jstor.org/*

Abu-Shawiesh, Moustafa Omar Ahmed. 2008. *A Simple Robust Control Chart Based on MAD. http://www.jstor.org/*

A. Klugman, Harry H. Panjer and Gordon E. Willmot. 2004. *Loss Models From Data to Decisions*. New Jersey: John Wiley & Sons

Dudewicz, Edward J. and Satya N. Mishra. 1988. *Modern Mathematical Statistics*. Singapore: John Wiley & Sons

Hampel, F.R. 1974. *The Influence Curve and Its Role in Robust Estimation. http://www.jstor.org/*

Huber, Peter J. 1981. *Robust Statistics*. USA: John Wiley & Sons

Lax, D.A. 1985.*Robust Estimators of Scale: Finite Sample Performance in Long-Tailed Symmetric Distributions. http://www.jstor.org/*

Montgomery, D.C. 2005. *Introduction to Statistical Quality Control, 5th ed*. USA: John Wiley & Sons

Robert V. Hogg and Allen T. Craig. 1995. *Introduction to Mathematical Statistics, 5thed*. New Jersey: Prentice-Hall

Robert V. Hogg, Joseph W. McKean and Allen T. Craig. 2005. *Introduction to Mathematical Statistics, 6thed*. USA: Prentice-Hall

Maronna, Ricardo A., R. Douglas Martin and Victor J. Yohai. 2006. *Robust Statistics: Theory and Methods*. USA: John Wiley & Sons

Barnett, Vic and Toby Lewis. 1978. *Outliers in Statistical Data*. USA: John Wiley & Sons

Bartle, Robert G. and Donald R. Sherbert. 2000. *Introduction to Real Analysis, 3rd ed.*, USA: John Wiley & Sons, Inc.

Lampiran 1

Bukti bahwa $E(\bar{X}) = \mu$, $E(S^2) = \sigma^2$ dan $E(S) = c_4\sigma$

Misalkan peubah acak *X* mempunyai mean μ dan variansi σ^2. Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak berukuran *n* dari populasi. Maka

$$\begin{aligned} E(\bar{X}) &= E\left(\tfrac{1}{n}\textstyle\sum_{i=1}^{n} X_i\right) \\ &= \tfrac{1}{n}\textstyle\sum_{i=1}^{n} E(X_i) \\ &= \tfrac{1}{n}\textstyle\sum_{i=1}^{n} \mu \\ &= \mu \end{aligned}$$

Sekarang perhatikan

$$\begin{aligned} E(S^2) &= E\left(\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}\right) \\ &= \frac{1}{n-1} E\left(\sum_{i=1}^{n}(X_i - \bar{X})^2\right) \\ &= \frac{1}{n-1} E\left(\sum_{i=1}^{n}(X_i^2 - 2X_i\bar{X} + \bar{X}^2)\right) \\ &= \frac{1}{n-1} E\left(\sum_{i=1}^{n} X_i^2 - 2\bar{X}\sum_{i=1}^{n} X_i + \sum_{i=1}^{n}\bar{X}^2\right) \\ &= \frac{1}{n-1} E\left(\sum_{i=1}^{n} X_i^2 - 2n\bar{X}^2 + n\bar{X}^2\right) \\ &= \frac{1}{n-1} E\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right) \end{aligned}$$

Dengan demikian,

$$\frac{1}{n-1} E\left(\sum_{i=1}^{n}(X_i - \bar{X})^2\right) = \frac{1}{n-1} E\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right)$$

$$E\left(\sum_{i=1}^{n}(X_i - \bar{X})^2\right) = \sum_{i=1}^{n} E(X_i^2) - E(n\bar{X}^2).$$

Karena

$\sum_{i=1}^{n} E(X_i^2) = \mu^2 + \sigma^2$ dan $E(\bar{X}^2) = \mu^2 + {\sigma^2}/{n}$, maka

$$E(S^2) = \frac{1}{n-1}\left[\sum_{i=1}^{n}(\mu^2 + \sigma^2) - n\left(\mu^2 + {\sigma^2}/{n}\right)\right]$$

$$= \frac{1}{n-1}[n\mu^2 + n\sigma^2 - n\mu^2 - \sigma^2]$$

$$= \frac{(n-1)\sigma^2}{n-1}$$

$$= \sigma^2$$

Meskipun $E(S^2) = \sigma^2$, standar deviasi sampel bukan penaksir yang unbiased untuk standar deviasi populasi. Buktinya sebagai berikut:

Misalkan peubah acak X berdistribusi normal dengan mean μ dan variansi σ^2. Misalkan $X_1, X_2, \ldots, X_n$ sampel acak berukuran n. Dapat dibuktikan bahwa

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

Sehingga distribusi dari variansi sampel adalah

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

Jadi, ketika sampel berasal dari distribusi normal, nilai ekspektasi dari S^2 adalah

$$E(S^2) = E\left(\frac{\sigma^2}{(n-1)}\chi^2_{n-1}\right)$$

$$= \frac{\sigma^2}{(n-1)}E(\chi^2_{n-1})$$

$$= \frac{\sigma^2}{(n-1)}(n-1)$$

$$= \sigma^2.$$

Kemudian, dapat dibuktikan juga bahwa

$$\frac{\sqrt{n-1}S}{\sigma} \sim \chi_{n-1}$$

Sehingga, nilai ekspektasi dari S dapat ditulis sebagai

$$E(S) = E\left(\frac{\sigma}{\sqrt{n-1}}\chi_{n-1}\right)$$
$$= \frac{\sigma}{\sqrt{n-1}}E(\chi_{n-1})$$

Mean dari distribusi Chi dengan derajat bebas $n-1$ adalah

$$E(\chi_{n-1}) = \sqrt{2}\frac{\Gamma(\mathrm{n}/2)}{\Gamma[(\mathrm{n}-1)/2]}$$

Dimana fungsi gamma $\Gamma(r) = \int_0^\infty y^{r-1}e^{-y}dy$. Maka

$$E(S) = \sqrt{\frac{2}{n-1}}\frac{\Gamma(\mathrm{n}/2)}{\Gamma[(\mathrm{n}-1)/2]}\sigma$$
$$= c_4\sigma$$

Lampiran 2

Tabel Konstanta c_4

Ukuran subgrup (n)	Nilai c_4
2	0,7979
3	0,8862
4	0,9213
5	0,9400
6	0,9515
7	0,9594
8	0,9650
9	0,9693
10	0,9727
11	0,9754
12	0,9776
13	0,9794
14	0,9810
15	0,9823
16	0,9835
17	0,9845
18	0,9854
19	0,9862
20	0,9869
25	0,9896
30	0,9914
35	0,9927
40	0,9936
45	0,9943
50	0,9949

.