

PERBANDINGAN BAGAN KENDALI T^2 HOTELLING KLASIK DENGAN T^2 HOTELLING PENDEKATAN *BOOTSTRAP* PADA DATA BERDISTRIBUSI NON-NORMAL MULTIVARIAT

KHAULAH BINTI AFRINALDI, MAIYASTRI, YUDIANTRI ASDI

*Program Studi Matematika,
Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Andalas,
Kampus UNAND Limau Manis Padang, Indonesia,
email : khaulahbintiafrinaldi@ymail.com*

Abstrak. Bagan kendali T^2 Hotelling merupakan bagan yang berguna untuk memonitor rata-rata pergeseran proses dengan asumsi distribusi normal harus dipenuhi. Pada penelitian ini akan digunakan data berdistribusi non-normal multivariat untuk melihat kinerja dari bagan kendali T^2 Hotelling dengan pendekatan *Bootstrap* yang efisien memantau proses ketika distribusi yang diamati adalah tidak normal atau tidak diketahui. Data yang digunakan pada penelitian ini adalah data sekunder yaitu data IPK dan Lama studi lulusan matematika FMIPA Unand tahun 2015. Tujuan penelitian ini adalah membandingkan kepekaan bagan kendali T^2 Hotelling klasik dengan T^2 Hotelling pendekatan *bootstrap* dalam mendeteksi titik-titik yang berada diluar batas kendali. Pada bagan kendali T^2 Hotelling klasik terdapat tujuh titik yang berada diluar batas kendali, sedangkan bagan kendali T^2 Hotelling pendekatan *bootstrap* mendeteksi sembilan titik yang berada diluar batas kendali.

Kata Kunci: Bagan kendali T^2 Hotelling, Transformasi Johnson, Bootstrap

1. Pendahuluan

Statistika pengendalian proses merupakan salah satu metode statistik yang mengendalikan kualitas produk atau jasa agar hasilnya tetap memberikan *output* yang sesuai dengan spesifikasi. Banyak alat yang digunakan untuk mengendalikan kualitas, salah satunya adalah dengan menggunakan bagan kendali. Bila terdapat dua karakteristik kualitas atau lebih yang harus dikendalikan secara bersamaan, maka bagan kendali yang digunakan adalah bagan kendali multivariat. Salah satu bagan kendali multivariat yang dapat digunakan adalah bagan kendali T^2 Hotelling. Bagan kendali T^2 Hotelling digunakan untuk mengendalikan dua atau lebih karakteristik kualitas yang mempunyai korelasi signifikan dan data yang dianalisis memenuhi asumsi distribusi normal.

Apabila asumsi distribusi normal multivariat tidak dipenuhi, maka sebelum membentuk bagan kendali T^2 Hotelling harus dilakukan transformasi pada data terlebih dahulu. Transformasi ini dilakukan untuk mengubah skala data sehingga data memenuhi asumsi distribusi normal dan pembentukan bagan kendali T^2 Hotelling dapat dilanjutkan. Namun, masih ada alternatif lain yang bisa dilakukan apabila data yang dianalisis tidak memenuhi asumsi distribusi normal, yaitu dengan

melakukan pendekatan nonparametrik. Ada beberapa metode nonparametrik yang bisa digunakan, antara lain dengan pendekatan fungsi densitas kernel, Linkage, *bootstrap*, dan lain-lain.

2. Tinjauan Pustaka

2.1. Bagan Kendali T^2 Hotelling

Bagan kendali T^2 Hotelling merupakan salah satu bagan kendali multivariat yang dapat mendeteksi pergeseran proses dengan menggunakan vektor rata-rata dan matriks ragam peragam sampel. Asumsi yang harus dipenuhi untuk menggunakan bagan kendali T^2 Hotelling adalah semua vektor pengamatan harus mengikuti distribusi normal multivariat.

Bagan Kendali T^2 Hotelling untuk Pengamatan Individu

Bagan kendali T^2 Hotelling untuk pengamatan individu digunakan apabila ukuran subgroup sampel $n = 1$. Misalkan \mathbf{X} adalah matriks berukuran $m \times p$ yang merupakan matriks data dari m -sampel dengan p -karakteristik kualitas sebagai berikut:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mp} \end{bmatrix}$$

Setelah diperoleh vektor rata-rata dan matriks ragam peragamnya, maka nilai T^2 Hotelling untuk pengamatan individu dapat dihitung menggunakan rumus:

$$T_k^2 = (\mathbf{x}_k - \bar{\mathbf{x}})' \mathbf{S}^{-1} (\mathbf{x}_k - \bar{\mathbf{x}}), \quad k = 1, 2, \dots, m \quad (2.1)$$

dimana : T_k^2 : nilai T^2 untuk setiap pengamatan ke- k
 \mathbf{x}_k : nilai setiap karakteristik kualitas
 $\bar{\mathbf{x}}$: vektor rata-rata karakteristik kualitas
 \mathbf{S}^{-1} : invers matriks ragam peragam

Matriks ragam peragam dihitung menggunakan rumus:

$$\mathbf{S} = \frac{1}{m-1} \sum_{k=1}^m (\mathbf{x}_k - \bar{\mathbf{x}})(\mathbf{x}_k - \bar{\mathbf{x}})' \quad (2.2)$$

dimana : \mathbf{S} : matriks ragam peragam dari karakteristik kualitas
 m : banyak sampel
 \mathbf{x}_k : nilai setiap karakteristik kualitas
 $\bar{\mathbf{x}}$: vektor rata-rata karakteristik kualitas

Batas kendali yang digunakan adalah

$$\begin{aligned} \text{BKA} &= \frac{(m-1)^2}{m} \beta_{\alpha, \frac{p}{2}, \frac{(m-p-1)}{2}} \\ \text{BKB} &= 0 \end{aligned} \quad (2.3)$$

dimana m merupakan jumlah sampel, p adalah jumlah karakteristik kualitas, sedangkan $\beta_{\alpha, \frac{p}{2}, \frac{(m-p-1)}{2}}$ adalah sebaran β dengan taraf nyata α serta parameter $\frac{p}{2}$ dan $\frac{(m-p-1)}{2}$.

2.2. Analisis Korelasi

Analisis korelasi merupakan analisis yang dilakukan untuk mengukur ke-kuatan hubungan antara dua peubah, sebut saja X dan Y melalui sebuah bilangan yang disebut koefisien korelasi. Koefisien korelasi didefinisikan sebagai ukuran ke-eratan hubungan linier antara dua peubah acak X dan Y yang dilambangkan dengan r untuk sampel, dan ρ untuk populasi.

Ukuran korelasi linier antara dua peubah yang paling banyak digunakan adalah Koefisien korelasi *Pearson's product-moment* yang dapat dihitung dengan menggunakan rumus berikut [6]:

$$r_{XY} = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{\sqrt{[n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2][n \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2]}} \quad (2.4)$$

dimana : n : banyaknya pengamatan
 x_i : nilai x ke- i , dengan $i = 1, 2, \dots, n$
 y_i : nilai y ke- i , dengan $i = 1, 2, \dots, n$

2.3. Uji Normal Multivariat

Mardia memperkenalkan uji normal multivariat berdasarkan perhitungan multivariat *skewness* ($b_{1,p}$) dan *kurtosis* ($b_{2,p}$) yang masing-masing dihitung menggunakan rumus [4]:

$$b_{1,p} = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n m_{ij}^3 \quad (2.5)$$

dan

$$b_{2,p} = \frac{1}{n} \sum_{i=1}^n m_{ii}^2 \quad (2.6)$$

dimana $m_{ij} = (\mathbf{x}_i - \bar{\mathbf{x}})' \mathbf{S}^{-1} (\mathbf{x}_j - \bar{\mathbf{x}})$ adalah jarak Mahalanobis, p adalah banyaknya peubah, dan n adalah banyaknya sampel.

2.4. Metode Bootstrap

Bootstrap adalah metode berbasis komputer yang dikembangkan untuk mengestimasi berbagai perhitungan statistik. Metode *bootstrap* pertama kali diperkenalkan oleh Bradley Efron pada tahun 1979 dan kemudian dikembangkan oleh Kotz dan Johnson pada tahun 1992. Metode *bootstrap* merupakan salah satu metode *resampling* dimana dari sekelompok data dilakukan pengambilan kembali sejumlah sampel secara berulang kali (replikasi) dengan pengembalian untuk memperoleh suatu kesimpulan. Metode *bootstrap* dapat digunakan untuk mengatasi permasalahan dalam statistika baik masalah data yang sedikit, data yang tidak memenuhi asumsi distribusi normal maupun data yang tidak diketahui distribusinya.

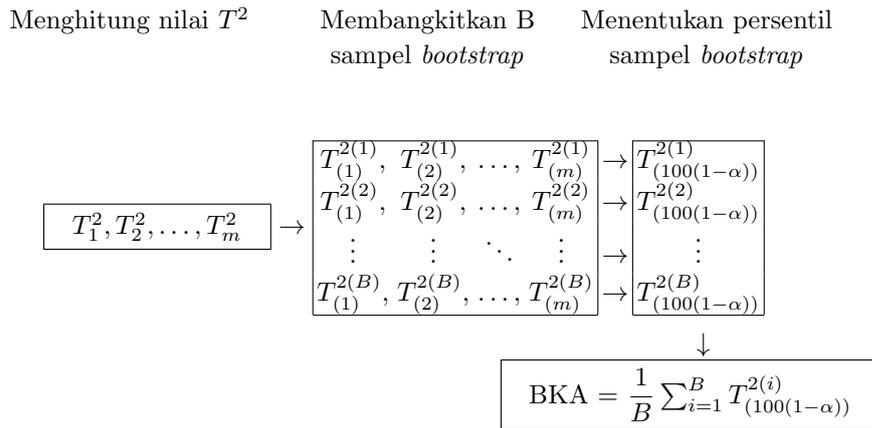
2.5. Jumlah Replikasi Bootstrap

Ada beberapa hal yang perlu diperhatikan mengenai jumlah replikasi *bootstrap*, yaitu [1] :

- (1) Meskipun jumlah replikasi *bootstrap* kecil, misalnya $B = 25$, biasanya sudah cukup informatif tetapi dengan $B = 50$, sudah sangat cukup untuk memberikan estimasi standar error yang akurat.
- (2) Jumlah replikasi *bootstrap* yang besar, misalnya $B = 200$, jarang dilakukan untuk mengestimasi standar error. Biasanya jumlah replikasi yang besar diperlukan untuk pendugaan selang kepercayaan *bootstrap*.

2.6. Bagan Kendali T^2 Hotelling Pendekatan Bootstrap

Prosedur dalam menghitung batas kendali untuk bagan kendali T^2 Hotelling dengan pendekatan *bootstrap* dapat dilihat pada gambar berikut.



Gambar 1. Prosedur *bootstrap* dalam menghitung batas kendali

Langkah pertama dalam membuat bagan kendali T^2 Hotelling pendekatan *bootstrap* yaitu dengan menghitung nilai T^2 dari data yang tidak berdistribusi normal tanpa melakukan transformasi pada data aslinya. Setelah menghitung nilai T^2 , maka dilakukan *resampling* dengan pengembalian sebanyak B . Setelah didapatkan nilai T^2 dengan ulangan sebanyak B , selanjutnya dalam setiap sampel *bootstrap* B dihitung nilai persentil ke-100 $(1 - \alpha)$, kemudian ditentukan batas kendali dengan mengambil rata-rata nilai persentil dari B .

3. Metode Penelitian

3.1. Data

Data yang digunakan pada tugas akhir ini adalah data sekunder yaitu data IPK dan lama studi lulusan matematika tahun 2015 sebanyak 73 sampel. Untuk setiap sampel dicatat data berupa IPK dan lama studi sebagai karakteristik kualitas.

3.2. Metode Penelitian

Terdapat beberapa tahapan dalam melakukan analisis data, dimulai dengan pembentukan bagan kendali T^2 Hotelling klasik, kemudian dilanjutkan dengan pembentukan bagan kendali T^2 Hotelling pendekatan *bootstrap*, serta membandingkan kinerja kedua bagan kendali tersebut. Dalam melakukan perhitungan dan pembuatan bagan kendali digunakan *software* R.3.3.1.

Berikut ini akan dijelaskan lebih rinci langkah-langkah dalam menganalisis data.

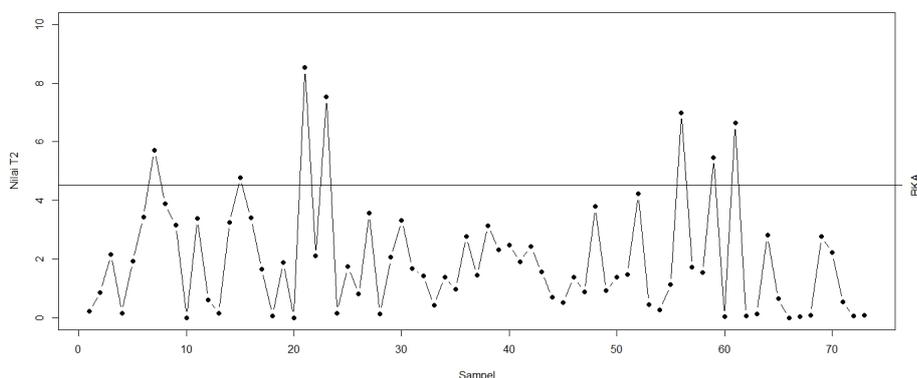
- (1) Membentuk Bagan Kendali T^2 Hotelling Klasik.
Langkah-langkah dalam membuat bagan kendali T^2 Hotelling klasik adalah:
 - (a) Menghitung nilai korelasi dua karakteristik kualitas.
 - (b) Menguji asumsi distribusi normal multivariat.
 - (c) Mengubah skala data dengan menggunakan transformasi Johnson pada data yang tidak memenuhi asumsi distribusi normal multivariat.
 - (d) Menguji kembali asumsi distribusi normal multivariat pada data yang telah ditransformasi.
 - (e) Menghitung nilai T^2 untuk setiap sampel.
 - (f) Menghitung nilai batas kendali atas.
 - (g) Membentuk bagan kendali.
 - (h) Mengidentifikasi titik-titik yang berada diluar batas kendali.
- (2) Membentuk Bagan Kendali T^2 Hotelling dengan Pendekatan Bootstrap.
Langkah-langkah dalam membuat bagan kendali T^2 Hotelling dengan pendekatan *bootstrap* adalah :
 - (a) Menghitung nilai T^2 untuk setiap sampel.
 - (b) Menjadikan $T_1^{2(i)}, T_2^{2(i)}, \dots, T_n^{2(i)}$ sebagai nilai T^2 dari sampel *bootstrap* ke- i , $i = 1, \dots, B$ yang diambil secara acak dari nilai T^2 .
 - (c) Menentukan nilai persentil dengan $\alpha = 0.1$ dalam setiap sampel *bootstrap* $B = 200$.
 - (d) Menentukan batas kendali atas dengan mengambil rata-rata nilai persentil dari B .
 - (e) Membentuk bagan kendali.
 - (f) Mengidentifikasi titik-titik sampel yang berada diluar batas kendali.
- (3) Membandingkan Kinerja Bagan Kendali.
Langkah terakhir adalah membandingkan kinerja kedua bagan kendali, yaitu bagan kendali T^2 Hotelling klasik dengan bagan kendali T^2 Hotelling pendekatan *bootstrap* dalam mendeteksi titik-titik yang berada diluar batas kendali.

4. Hasil dan Pembahasan

4.1. Membentuk Bagan Kendali T^2 Hotelling Klasik

Pada uji korelasi diperoleh hasil bahwa nilai *p-value* kurang dari α , maka terdapat korelasi yang signifikan antara IPK dan lama studi. Dari nilai koefisien korelasi juga dapat disimpulkan bahwa korelasi antara kedua karakteristik kualitas tersebut cukup erat dengan nilai koefisien korelasi sebesar -0.8 .

Pada uji distribusi normal multivariat, diperoleh hasil bahwa data IPK dan lama studi lulusan matematika tahun 2015 tidak berdistribusi normal multivariat, maka dilakukan transformasi pada data tersebut sehingga data tersebut berdistribusi normal multivariat. Setelah dilakukan transformasi, maka nilai T^2 untuk bagan kendali T^2 Hotelling klasik dapat dihitung.



Gambar 2. Bagan Kendali T^2 Hotelling Klasik

Berdasarkan bagan kendali pada Gambar 2 dapat dilihat bahwa pada bagan kendali T^2 Hotelling klasik terdapat tujuh titik sampel yang berada di luar batas kendali, yaitu sampel ke 7, 15, 21, 23, 56, 59 dan 61.

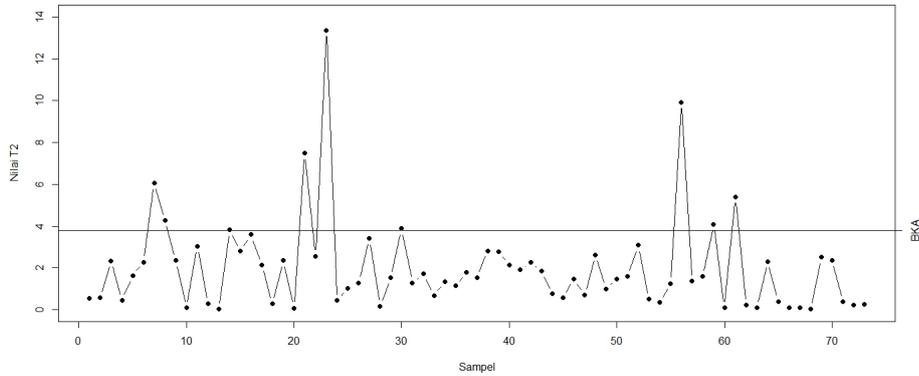
4.2. Pembentukan Bagan Kendali T^2 Hotelling dengan Pendekatan Bootstrap

Untuk membuat bagan kendali T^2 Hotelling yang tidak memenuhi asumsi distribusi normal, maka digunakan bagan kendali T^2 Hotelling dengan pendekatan *bootstrap* sehingga data asli tidak perlu ditransformasi terlebih dahulu. Perbedaan antara bagan kendali T^2 Hotelling klasik dan bagan kendali T^2 Hotelling dengan pendekatan *bootstrap* terletak pada batas kendalinya.

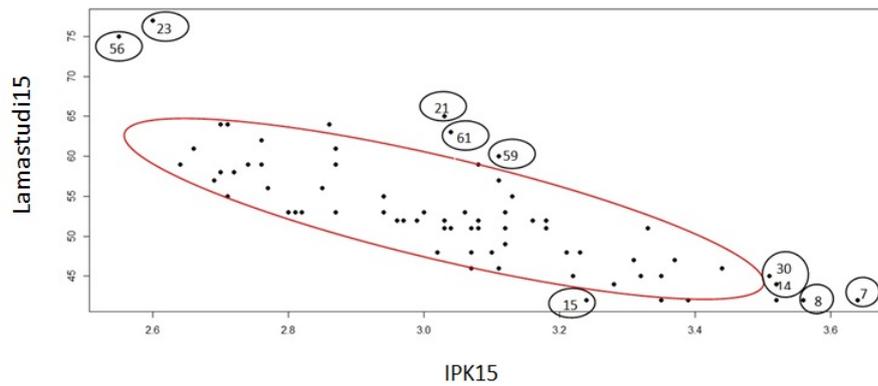
Berdasarkan bagan kendali pada Gambar 3 dapat dilihat bahwa pada bagan kendali T^2 Hotelling pendekatan *bootstrap* terdapat sembilan titik sampel yang berada diluar batas kendali, yaitu sampel ke 7, 8, 14, 21, 23, 30, 56, 59 dan 61.

Berdasarkan bagan kendali yang terbentuk, dapat dilihat bahwa pada bagan kendali T^2 Hotelling klasik hanya tujuh titik yang keluar dari batas kendali atas, sedangkan pada bagan kendali T^2 Hotelling pendekatan *bootstrap* terdapat sembilan titik yang keluar dari batas kendali atas, sehingga dapat dikatakan bahwa bagan kendali T^2 Hotelling pendekatan *bootstrap* lebih peka dalam mendeteksi titik-titik yang berada diluar batas kendali.

Berikut ini adalah plot IPK dan lama studi lulusan matematika FMIPA UNAND tahun 2015.



Gambar 3. Bagan Kendali T^2 Hotelling Pendekatan Bootstrap



Gambar 4. Plot IPK dan Lama studi lulusan matematika

Secara umum, terdapat kecenderungan mahasiswa yang masa studinya singkat, IPK-nya cenderung tinggi, sementara mahasiswa yang masa studinya lama, IPK-nya cenderung rendah. Akan tetapi, terdapat beberapa titik yang tidak mengikuti pola di atas, dan titik inilah yang menjadi titik yang berada di luar batas kendali.

5. Kesimpulan

Berdasarkan uraian dari pembahasan sebelumnya, dapat diambil kesimpulan sebagai berikut:

- (1) Baik bagan kendali T^2 Hotelling klasik maupun bagan kendali T^2 Hotelling pendekatan bootstrap sama-sama memberikan sinyal *out of control*. Bagan kendali T^2 Hotelling klasik hanya mendeteksi 7 titik yang keluar dari batas kendali atas dengan nilai $BKA = 4.521499003$, sedangkan bagan kendali T^2 Hotelling pendekatan bootstrap mendeteksi 9 titik yang keluar dari batas kendali atas

dengan nilai $BKA = 3.80156$. Mengacu pada banyaknya titik yang keluar dari batas kendali atas, maka dapat diartikan bahwa bagan kendali T^2 *Hotelling* pendekatan *bootstrap* lebih sensitif dalam mendeteksi nilai-nilai yang berada diluar batas kendali dibandingkan bagan kendali T^2 *Hotelling* klasik dalam mendeteksi pergeseran proses.

- (2) Secara umum, terdapat kecenderungan mahasiswa yang masa studinya sebentar, IPK nya cenderung tinggi, sementara mahasiswa yang masa studinya lama, IPK-nya cenderung rendah. Akan tetapi, terdapat beberapa titik yang tidak mengikuti pola di atas, dan titik inilah yang menjadi titik yang berada di luar batas kendali.
- (3) Metode *bootstrap* dapat dijadikan sebagai metode alternatif ketika asumsi distribusi normal multivariat tidak dipenuhi dalam pembentukan bagan kendali, karena metode *bootstrap* tidak memerlukan syarat asumsi.

Daftar Pustaka

- [1] Efron, B dan Tibshirani, R. 1993. *An Introduction to the Bootstrap*. Chapman and Hall, New York, London.
- [2] George, F. 2007. *Johnson's System of Distributions and Microarray Data Analysis*. Dissertations. Department of Mathematics, University of South Florida, USA.
- [3] Kim S.B. 2011. *Bootstrap-Based T^2 Multivariate Control Charts*. School of Industrial Management Engineering, Korea University, Seoul, Korea.
- [4] Korkmaz, S., Dincer Goksuluk, dan Gokmen Zararsiz. 2015. *An R Package for Assessing Multivariate Normality*. Department of Biostatistics, Hacettepe University, Ankara, Turki.
- [5] Montgomery, D. C. 2002. *Introduction To Statistical Quality Control*. John Wiley & Sons, Inc. New York.
- [6] Walpole, RE. 1993. *Pengantar Statistika*. Edisi ke-3. Jakarta: PT Gramedia Pustaka Utama.
- [7] Yudistira, I.G.A.A. 2015. *Penerapan Metode Resampling untuk Pendugaan Indeks Kemampuan Proses*. Jurnal Widya Eksakta, Vol 1 No 1, Maret 2015: 28 – 33.