

**BAGAN KENDALI NONPARAMETRIK DENGAN
ESTIMASI FUNGSI KEPEKATAN KERNEL
(Studi Kasus: Indeks Prestasi Mahasiswa Jurusan Matematika Angkatan 2014-2016
FMIPA Untan pada Semester Genap 2016/2017)**

Sani, Setyo Wira Rizki, Nurfitri Imro'ah

INTISARI

Indeks Prestasi (IP) merupakan salah satu tolok ukur prestasi belajar pada proses perkuliahan. Demi menjaga konsistensi dan kestabilan IP agar tetap baik, maka perlu dilakukan pengendalian kualitas. Pengendalian kualitas dapat dilakukan dengan bagan kendali. Bagan kendali yang digunakan adalah bagan kendali nonparametrik dengan pendekatan Kernel. Langkah yang dilakukan adalah menentukan fungsi Kernel yang digunakan. Setelah itu mengestimasi fungsi kepekatan dengan masing-masing fungsi Kernel tersebut. Berdasarkan fungsi estimasi kepekatan Kernel, selanjutnya dicari nilai bandwidth yang optimal dan ditentukan nilai-nilai batas kendali. Batas-batas kendali yang diperoleh digunakan untuk membentuk bagan kendali. Data yang digunakan adalah data IP mahasiswa Jurusan Matematika FMIPA Untan angkatan 2014-2016 semester genap 2016/2017. Hasil penelitian menunjukkan bahwa estimasi fungsi kepekatan Kernel Triangular adalah fungsi yang paling baik jika dibandingkan dengan fungsi Kernel Epanechnikov, Biweight, dan Gaussian pada kasus ini.

Kata kunci: *Bagan kendali, Kepekatan Kernel, Nonparametrik*

PENDAHULUAN

Perguruan tinggi merupakan sebuah institusi pendidikan lanjutan setelah lulus dari Sekolah Menengah Atas (SMA)/ sederajat yang mempunyai peran sangat luas dalam pengembangan sumber daya manusia. Kualitas dari perguruan tinggi dapat dilihat dari proses perkuliahan yang berlangsung. Salah satu tolok ukur prestasi belajar mahasiswa selama proses perkuliahan tiap semester di perguruan tinggi dapat ditunjukkan melalui nilai yang diberikan oleh dosen atau biasa yang disebut dengan Indeks Prestasi (IP).

Demi menjaga konsistensi dan kestabilan nilai Indeks Prestasi agar tetap baik, maka perlu dilakukan pengendalian kualitas. Pengendalian kualitas merupakan langkah-langkah penyelesaian masalah yang digunakan untuk memonitor, menganalisis, mengendalikan, dan mampu mendeteksi penyimpangan-penyimpangan yang terjadi sehingga dapat dilakukan suatu tindakan perbaikan terhadap suatu proses.

Penemu bagan kendali Shewhart adalah Walter Andrew Shewhart. Dalam bagan kendali Shewhart harus memenuhi ketentuan bahwa data yang dianalisis harus berdistribusi normal. Namun, pada saat di lapangan data sampel yang didapat tidak selalu berdistribusi normal. Seringkali data yang dimiliki tidak diketahui bentuk sebarannya, sehingga bagan kendali Shewhart tidak cocok digunakan. Oleh karena itu dikembangkan alternatif bagan kendali dengan pendekatan nonparametrik karena metode nonparametrik tidak membutuhkan asumsi distribusi normal [1]. Metode untuk membangun bagan kendali nonparametrik dengan menggunakan pendekatan Kernel seperti yang telah diperkenalkan oleh [2]. Pada pendekatan ini, bagan kendali dibangun dengan terlebih dahulu mengestimasi bentuk fungsi kepekatan peluang dari data dengan metode estimasi fungsi kepekatan Kernel, yang dalam pengestimasiannya menggunakan fungsi Kernel. Fungsi Kernel yang digunakan pada penelitian ini diantaranya Kernel Epanechnikov, Biweight, Triangular, dan Gaussian.

Data yang digunakan dalam penelitian ini merupakan data sekunder yang didapat dari Unit Pelaksana Teknis Teknologi, Informasi dan Komunikasi (UPT TIK) Untan. Langkah pertama dalam penelitian ini melakukan pengumpulan data Indeks Prestasi (IP) mahasiswa Jurusan Matematika angkatan 2014-2016.

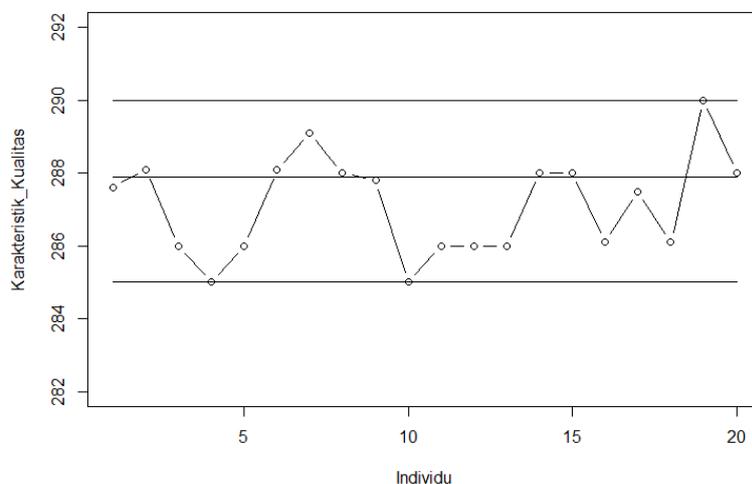
Setelah data terkumpul selanjutnya mendefinisikan masing-masing fungsi Kernel dan rumus estimasi fungsi kepekatan masing-masing Kernel dalam bahasa pemrograman R.3.2.4. Selanjutnya untuk mencari nilai *bandwidth* (h) optimal pada estimasi fungsi kepekatan masing-masing Kernel dengan metode *Cross Validation*. Menghitung estimasi fungsi kepekatan masing-masing Kernel dengan program R. Selanjutnya menggambarkan grafik estimasi fungsi kepekatan Kernel dengan *bandwidth* yang optimal dengan program R, setelah itu menentukan nilai batas kendali atas (UCL), garis tengah (CL) dan batas kendali bawah (LCL) berdasarkan grafik estimasi fungsi kepekatan Kernel. Selanjutnya menggambarkan bagan kendali berdasarkan estimasi fungsi kepekatan Kernel dengan batas-batas kendali yang telah diperoleh. Langkah terakhir yaitu membandingkan bagan kendali dengan estimasi fungsi kepekatan masing-masing Kernel yang digunakan.

STATISTICAL PROCESS CONTROL (SPC)

Proses merupakan suatu transformasi atau perubahan dari sekumpulan *input* yang dapat berupa materi, tindakan, metode, dan operasi, sehingga dihasilkan *output* dalam bentuk produk, informasi, dan pelayanan atau *service* [3]. Pada tahap perubahan *input* menjadi *output* pasti terdapat gangguan yang membuat *output* menjadi beragam atau terjadi variasi dalam pengukuran kualitas dari proses yang satu dengan proses yang lainnya. Upaya yang diperlukan untuk meminimalisir keragaman dengan melakukan suatu pengendalian. Pengendalian terhadap proses ini dapat dilakukan dengan menggunakan alat statistik, yang disebut *Statistical Process Control* (SPC). SPC merupakan teknik penyelesaian masalah yang digunakan untuk menganalisis data dalam mengontrol produksi sehingga sesuai dengan yang diinginkan. Salah satu alat SPC adalah bagan kendali (*Control Chart*) [4].

BAGAN KENDALI (CONTROL CHART)

Ide pembuatan bagan kendali pertama kali dikembangkan oleh Dr. Walter Andrew Shewhart tahun 1924 yang bertujuan untuk memonitor karakteristik kualitas dari suatu proses, apakah proses tersebut sudah berjalan sesuai prosedur yang diharapkan. Hal utama yang dilakukan untuk membuat sebuah bagan kendali adalah melakukan pengambilan data sampel berukuran n . Pada masing-masing data sampel diukur atau diambil karakteristik kualitas yang menjadi perhatian. Data sampel kemudian yang menjadi *input* untuk membuat suatu bagan kendali. Secara umum bagan kendali dapat digambarkan sebagai berikut:



Gambar 1 Bagan kendali secara umum

Berdasarkan Gambar 1 dapat dilihat bentuk umum dari suatu bagan kendali, dimana sumbu horizontal menunjukkan individu sedangkan sumbu vertikal menunjukkan struktur karakteristik kualitas yang menjadi perhatian. Bagan kendali merupakan alat untuk menganalisis apakah proses dalam menghasilkan sesuatu sudah terkendali atau tidak secara statistik. Bagan kendali dapat mempermudah untuk mendeteksi adanya penyimpangan dalam proses yang dilakukan [4]. Bagan kendali disertai

dengan tiga garis batas yang terdiri dari garis pusat atau *Center Line* (CL) merupakan garis yang menandakan tidak adanya penyimpangan dari karakteristik sampel. Sepasang garis lainnya ditempatkan di atas garis pusat, yang disebut sebagai batas kendali atas atau *Upper Control Limit* (UCL) dan di bawah garis pusat yang disebut sebagai batas kendali bawah atau *Lower Control Limit* (LCL).

ESTIMASI FUNGSI KEPEKATAN KERNEL TERBAIK

Ada dua faktor yang mempengaruhi estimasi fungsi kepekatan Kernel yaitu fungsi Kernel dan *bandwidth* (h). Fungsi Kernel berperan dalam menentukan bentuk grafik estimasi, sedangkan *bandwidth* berperan dalam menentukan kemulusan dan ketepatan fungsi Kernel. Berdasarkan dua faktor tersebut yang paling berperan dalam menentukan kemulusan keakuratan grafik estimasi fungsi kepekatan Kernel adalah pemilihan *bandwidth* yang optimal [5].

Menurut [5], salah satu metode pemilihan *bandwidth* optimal adalah menggunakan metode *Cross Validation* (CV). Metode ini merupakan metode yang digunakan untuk menduga kesalahan estimasi. Secara numerik *bandwidth* dapat dipilih dan didekati dengan memilih h yang meminimalkan CV seperti pada Persamaan 1.

$$CV(h_p) = \int_{-\infty}^{\infty} f^2(x) dx - \frac{2}{n} \sum_{i=1}^n f_{-i}(X_i) \tag{1}$$

Bandwidth (h) yang optimal dapat dicari dengan meminimalkan CV seperti langkah-langkah sebagai berikut [5]:

- a. Menentukan beberapa nilai awal untuk h yaitu $h_1 < h_2 < \dots < h_k$.
- b. Menghitung $CV(h_p), p=1,2,\dots,k$.
- c. Terdapat tiga kemungkinan yaitu:
 - 1) $CV(h_1) < CV(h_2) < \dots < CV(h_k)$ yang berarti h optimal sebenarnya berada di sebelah kiri h_1 . Langkah a diulangi dengan mengambil nilai-nilai h yang kurang dari h_1 .
 - 2) $CV(h_1) > CV(h_2) > \dots > CV(h_k)$ yang berarti h optimal sebenarnya berada di sebelah kanan h_k . Langkah a diulangi dengan mengambil nilai-nilai h yang lebih dari h_k .
 - 3) $CV(h_{k-1}) > CV(h_k) < CV(h_{k+1})$ yang berarti h optimal sebenarnya berada di sekitar h_k . Langkah a diulangi dengan mengambil nilai-nilai h di sekitar h_k .

FUNGSI KERNEL

Suatu fungsi $K(z)$ untuk $-\infty < z < \infty$ dengan $z = \frac{x_j - X_i}{h}$, dengan x_j merupakan nilai acak ke- j yang akan di estimasi, X_i merupakan titik sampel ke- i dan h merupakan parameter pemulus (*bandwidth*) dengan $h > 0$. Disebut fungsi Kernel jika memenuhi sifat-sifat berikut [6]:

- a. $K(z) \geq 0$ untuk setiap z .
- b. $K(z): \mathfrak{R} \rightarrow \mathfrak{R}$.
- c. $\int_{-\infty}^{\infty} K(z) dz = 1$.
- d. Simetri dimana $K(z) = K(-z)$ untuk setiap z .

Beberapa fungsi Kernel dan grafiknya terdapat pada Tabel 1 antara lain [7]:

Tabel 1 Fungsi Kernel

Nama Kernel	Bentuk Fungsi $K(z)$	Grafik
Epanechnikov	$K(z) = \begin{cases} \frac{3}{4} \frac{1-0,2z^2}{\sqrt{5}}, & \text{untuk } z < \sqrt{5} \\ 0 & , \text{lainnya} \end{cases}$	
Biweight	$K(z) = \begin{cases} \frac{15}{16} (1-z^2)^2, & \text{untuk } z < 1 \\ 0 & , \text{lainnya} \end{cases}$	
Triangular	$K(z) = \begin{cases} 1- z , & \text{untuk } z < 1 \\ 0 & , \text{lainnya} \end{cases}$	
Gaussian	$K(z) = \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{1}{2}\right)z^2}$	

ESTIMASI FUNGSI KEPEKATAN KERNEL

Karakteristik dasar yang menggambarkan sifat atau perilaku dari suatu peubah acak adalah fungsi kepekatan peluang. Dalam statistik, estimasi fungsi kepekatan Kernel merupakan salah satu metode nonparametrik untuk menduga fungsi kepekatan peluang dari suatu peubah acak [1].

Estimasi fungsi kepekatan Kernel dapat dituliskan sebagai berikut:

$$f(x_j) = \frac{1}{hn} \sum_{i=1}^n K\left(\frac{x_j - X_i}{h}\right) \quad (2)$$

dimana:

n = ukuran sampel

h = parameter pemulus (*Bandwidth*)

x_j = nilai acak ke- j yang akan di estimasi, $j=1, 2, 3, \dots, n$

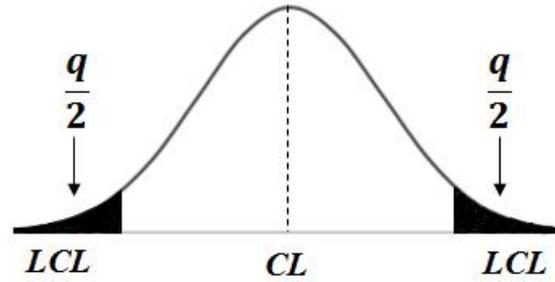
X_i = titik sampel ke- i , $i = 1, 2, 3, \dots, n$

K = fungsi Kernel

Untuk menentukan estimasi fungsi kepekatan Kernel, dapat digunakan berbagai bentuk fungsi Kernel yaitu dengan mengganti fungsi K dengan fungsi-fungsi Kernel lainnya.

BAGAN KENDALI DENGAN ESTIMASI FUNGSI KEPEKATAN KERNEL

Bagan kendali nonparametrik dapat dibuat berdasarkan estimasi fungsi kepekatan Kernel. Dengan mengambil perumpamaan bahwa jika bagan kendali yang digunakan bagan kendali Shewhart, maka batas-batas kendali yang biasa digunakan adalah batas kendali 3 sigma. Hal ini menunjukkan bahwa hanya ada kemungkinan sekitar 0,27% dari titik-titik sampel yang berada diluar batas-batas kendali. Nilai estimasi fungsi kepekatan Kernel untuk setiap x dapat digambarkan dalam bentuk grafik untuk mencari batas kendali seperti gambar berikut [8]:



Gambar 2 Gambaran mengenai hubungan estimasi fungsi kepekatan Kernel dengan batas-batas kendali

Berdasarkan Gambar 2 dapat diketahui bahwa batas kendali bawah (LCL) ditentukan dengan memperoleh luasan grafik bagian kiri sebesar $\frac{q}{2}$, sedangkan batas kendali atas (UCL) ditentukan dengan memperoleh luasan grafik bagian kanan sebesar $\frac{q}{2}$. Oleh karena itu, batas-batas kendali untuk bagan kendali nonparametrik dengan estimasi fungsi kepekatan Kernel dapat ditentukan sebagai berikut:

UCL (*Upper Control Limit*) atau batas kendali atas adalah nilai x , sehingga

$$\int_{-\infty}^{UCL} f(x) dx = \left(1 - \frac{q}{2}\right) 100\% \tag{3}$$

Sebagai garis tengah (*Center Line*) diambil nilai median, yaitu nilai x , sehingga

$$\int_{-\infty}^m f(x) dx = \left(\frac{1}{2}\right) 100\% \tag{4}$$

LCL (*Lower Control Limit*) atau batas kendali bawah adalah nilai x , sehingga

$$\int_{-\infty}^{LCL} f(x) dx = \left(\frac{q}{2}\right) 100\% \tag{5}$$

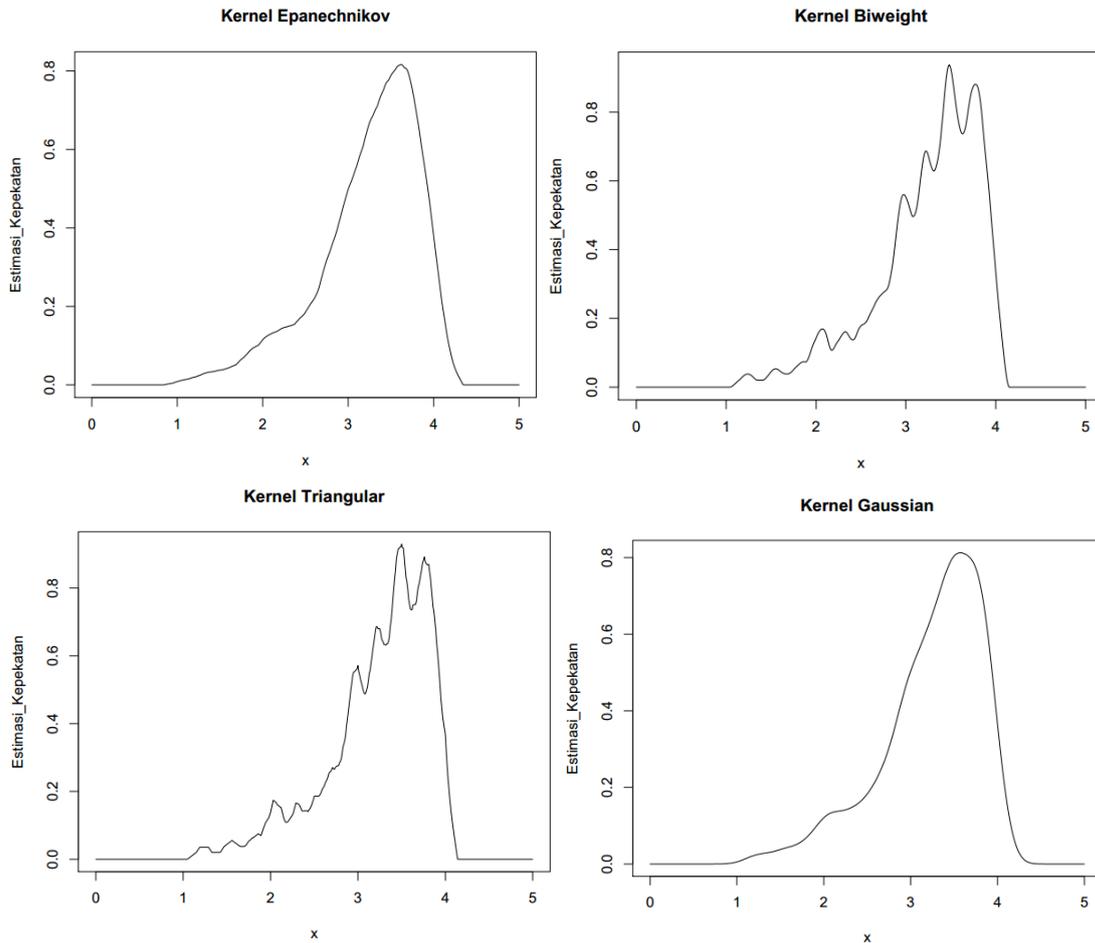
ESTIMASI FUNGSI KEPEKATAN KERNEL

Perhitungan fungsi kepekatan Kernel juga bergantung kepada nilai *Bandwidth*. Hal ini sangat berpengaruh terhadap kemulusan grafik yang akan diperoleh, oleh karena itu diperlukan pencarian *bandwidth* yang tepat (optimal). *Bandwidth* yang optimal dapat dihitung dan dipilih berdasarkan CV yang minimum pada masing-masing fungsi Kernel. Setelah dihitung dengan menggunakan *software R* nilai *bandwidth* yang optimal dapat dilihat pada Tabel 2.

Tabel 2 *Bandwidth* Optimal

No	Kernel	<i>Bandwidth</i> Optimal
1	Epanechnikov	0,1536031
2	Biweight	0,1509103
3	Triangular	0,1403537
4	Gaussian	0,1375388

Setelah diperoleh nilai *Bandwidth* (h) yang optimal pada masing-masing fungsi Kernel berdasarkan CV yang paling minimum, selanjutnya dapat dilakukan perhitungan estimasi fungsi kepekatan pada masing-masing fungsi Kernel dengan menggunakan *software R* untuk 258 sampel. Berdasarkan sampel dan nilai estimasi fungsi kepekatan Kernel yang telah diperoleh, selanjutnya dapat digambarkan grafik estimasi fungsi kepekatan Kernel dengan menggunakan *software R*, sehingga diperoleh grafik estimasi seperti pada Gambar 3 berikut:



Gambar 3 Grafik estimasi fungsi kepekatan Kernel

Berdasarkan Gambar 3 terlihat bahwa grafik estimasi kepekatan Kernel Epanechnikov dan Gaussian memiliki grafik yang paling mulus dibanding Biweight dan Triangular.

BATAS-BATAS KENDALI DENGAN ESTIMASI FUNGSI KEPEKATAN KERNEL

Berdasarkan Persamaan 3, 4, 5 dan $q=0,0027$ dengan bantuan R diperoleh batas kendali UCL, CL dan LCL untuk masing-masing fungsi Kernel seperti pada Tabel 3.

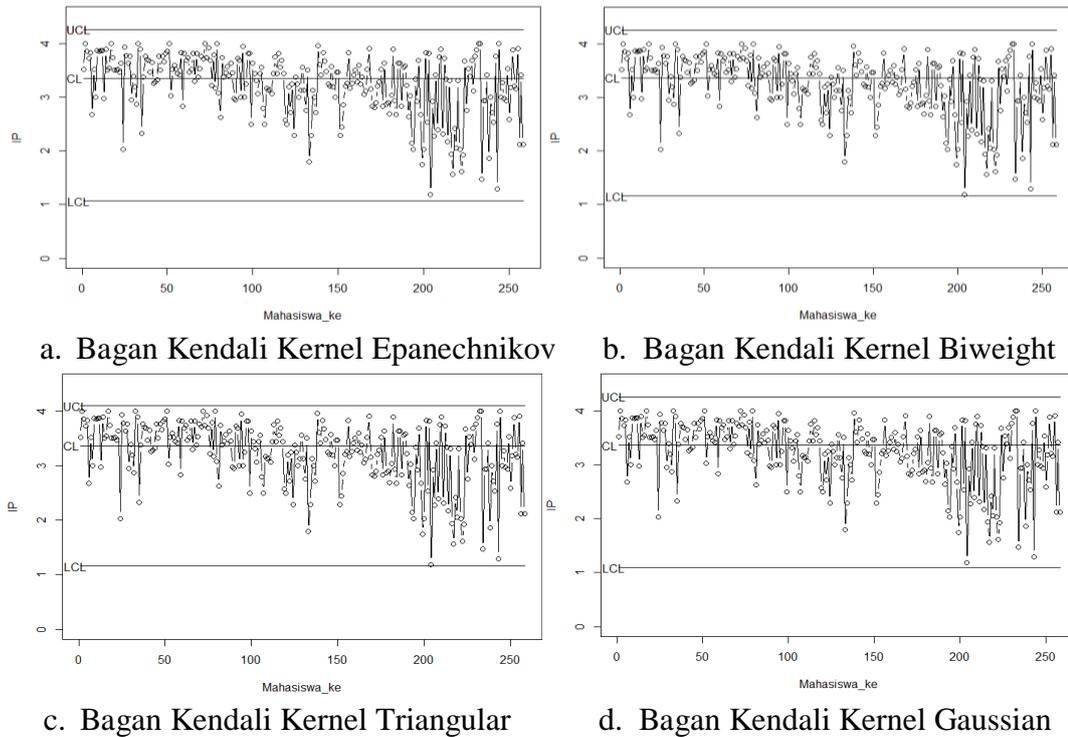
Tabel 3 Batas-batas Kendali

No	Kernel	UCL	CL	LCL
1	Epanechnikov	4,256667	3,353264	1,069231
2	Biweight	4,094637	3,368540	1,159167
3	Triangular	4,093889	3,368657	1,160740
4	Gaussian	4,256667	3,356466	1,096154

Berdasarkan Tabel 3 dapat disimpulkan bahwa estimasi fungsi kepekatan masing-masing Kernel yang digunakan dapat menghasilkan batas kendali yang berbeda-beda.

BAGAN KENDALI NONPARAMETRIK DENGAN ESTIMASI FUNGSI KEPEKATAN KERNEL

Selanjutnya dapat dibangun bagan kendali berdasarkan estimasi fungsi kepekatan Kernel dan batas kendali dengan menggunakan *software* R, sehingga diperoleh empat bagan kendali berdasarkan estimasi fungsi kepekatan Kernel seperti pada Gambar 4 sebagai berikut:



Gambar 4 Bagan kendali dengan estimasi fungsi kepekatan Kernel

Berdasarkan Gambar 4 terlihat bahwa bagan kendali yang dihasilkan dengan estimasi fungsi kepekatan Kernel Epanechnikov, Biweight, Triangular, dan Gaussian menunjukkan semua titik sampel berada di dalam batas kendali. Sehingga dapat disimpulkan bahwa proses perkuliahan di Jurusan Matematika angkatan 2014-2016 FMIPA Untan pada semester genap 2016/2017 berada dalam keadaan terkendali secara statistik.

Penentuan bagan kendali yang paling baik dari keempat bagan kendali yang digunakan, yaitu dapat dibandingkan dari lebar batas kendali yang diperoleh pada masing-masing bagan kendali. Melalui Tabel 3 dapat dicari lebar batas kendali pada bagan kendali dengan cara mencari selisih dari batas kendali atas (UCL) dan batas kendali bawah (LCL), sehingga diperoleh lebar kendali untuk Kernel Epanechnikov sebesar 3,187436, Kernel Biweight sebesar 2,935470, Kernel Triangular sebesar 2,933149, dan Kernel Gaussian sebesar 3,160513. Berdasarkan masing-masing selisih batas kendali yang diperoleh, bagan kendali dengan estimasi fungsi kepekatan Kernel Triangular lebih baik digunakan pada kasus ini.

KESIMPULAN

Berdasarkan hasil pembahasan yang telah dilakukan, dapat diperoleh kesimpulan yaitu dari perbandingan keempat bagan kendali yang digunakan dengan estimasi fungsi kepekatan Kernel Triangular lebih baik. Karena jika dibandingkan dengan bagan kendali Epanechnikov, Biweight, dan Gaussian, lebar batas kendali Kernel Triangular lebih sempit, sehingga lebih sensitif dalam mengidentifikasi titik-titik sampel pada kasus ini.

DAFTAR PUSTAKA

[1] Taungke N, Setiawan A, Parhusip HA. Pengendalian kualitas produk untuk karakteristik pH dengan menggunakan grafik pengendali berdasarkan densitas Kernel, Program Studi Matematika, Fakultas Sains dan Matematika Universitas Kristen Satya Wacana; 2011.

[2] Vermaat MB, RA Ion, RJMM Does, CAJ Klaassen. A comparison of Shewhart individuals control chart based on normal. *Non-Parametric and Extreme-Value Theory*; 2003;19(17):337-353

[3] Oaklan JS. *Statistical Process Control Fifth Edition*. MPG Books Limited. Great Britian; 2003.

[4] Montgomery DC. *Introduction to Statistical Quality Control Ed VI*. Hoboken: John Wiley & Sons; 2009.

- [5] Santoso R. Grafik pengendali nonparametrik empirik, *Media Statistika*; 2008;1(2):83-90
- [6] Jupit E, Maiyastri, Yozza H. Bagan Kendali Nonparametrik Dengan Estimasi Fungsi Kepekatan Kernel (Studi Kasus: Indeks Prestasi Mahasiswa Jurusan Matematika Angkatan 2011-2013 Fmipa Unand pada Semester Ganjil 2015-2016), *Jurnal Matematika UNAND*; 2016;5(2):1-10
- [7] Hardle W. *Smoothing Techniques with Implementation in S*. New York: Springer-Verlag; 2015.
- [8] Montgomery DC. *Pengantar Pengendalian Kualitas Statistik* [Soejoeti Z, trans]. Yogyakarta: Gadjah Mada University Press; 1990.

SANI : Jurusan Matematika FMIPA Untan, Pontianak,
sani.statistika@gmail.com

SETYO WIRA RIZKI : Jurusan Matematika FMIPA Untan, Pontianak,
setyo.wirarizki@math.untan.ac.id

NURFITRI IMRO'AH : Jurusan Matematika FMIPA Untan, Pontianak,
nurfitriimroah18@gmail.com
