



Journal of Economics and Administrative Sciences (JEAS)



Available online at <http://jeasiq.uobaghdad.edu.iq>

استعمال مقدر كثافة اللب في تحديد الحدود للوحات السيطرة متعددة المتغيرات

أ.م.د. أسماء غالب جابر
جامعة بغداد / الإدارة والاقتصاد / الاحصاء

الباحث / فهد حسين عناد
جامعة بغداد / الإدارة والاقتصاد / الاحصاء

Drasmaa.ghalib@coadec.uobaghdad.edu.iq

موبايل/07828940763

fahad89_1@yahoo.com

Received: 14/2/2020

Accepted :20/9/2020

Published :December / 2020

هذا العمل مرخص تحت اتفاقية المشاع الابداعي نسب المُصنّف - غير تجاري - الترخيص العمومي الدولي 4.0 Attribution-NonCommercial 4.0 International (CC BY-NC 4.0)



مستخلص البحث:

تعد السيطرة النوعية من الأدوات الإحصائية الفعالة في مجال السيطرة على الإنتاجية لمراقبة ومطابقة المنتجات المصنعة للمواصفات القياسية والمعايير المعتمدة لبعض السلع والخدمات والغرض منها هو مواكبة التطور الانتاجي والصناعي في سوق العمل والمنافسة وتستعمل لوحات السيطرة النوعية في مراقبة الخصائص النوعية للعمليات الانتاجية فضلا عن الكشف عن الانحرافات غير الطبيعية في العملية الإنتاجية. تم استعمال طريقة مقدر كثافة اللب وهي احدى الطرائق اللامعلمية التي لا تتطلب أي افتراضات حول توزيع البيانات أو تحديد حدود السيطرة في مراقبة العملية الانتاجية في كون البيانات لا تتبع التوزيع الطبيعي أو لها توزيع غير معروف ، اذ يهدف البحث الى مراقبة العملية الانتاجية من خلال عدة متغيرات في أن واحد ليعكس جودة المادة المنتجة . تم استعمال تجارب المحاكاة لمستويات معنوية مختلفة لتوضيح طريقة عمل اسلوب مقدر كثافة اللب باعتماد معيار متوسط طول المدى لبيان اداء وكفاءة الطرائق المستعملة. اظهرت النتائج من خلال المقارنة بين دوال مقدر كثافة اللب ان اسلوب مقدر كثافة اللب لدالة كاوسين اللامعلمية كان أدائها جيد خصوصا عند مستويات معنوية ذات مدى طويل .

المصطلحات الرئيسية للبحث / السيطرة النوعية ، مقدر كثافة اللب ، كاوسين كيرنل ، المربعات الصغرى للعبور الشرعي

بحث مستل من رسالة ماجستير

1- المقدمة (Introduction) :

يعد موضوع السيطرة الاحصائية من اهم المواضيع التي لها دور بارز في تحقيق المواصفات الموضوعية بالنسبة الى المنتج وزيادة القدرة على مراقبة جودة المنتجات الصناعية التي تم انتاجها ورصد العيوب من خلال عملية الانتاج ومن اهم الطرائق المستعملة في السيطرة النوعية هي عمليات السيطرة الاحصائية (SPC) (Statistical Process Control)، والتي تعد وسيلة لمراقبة النوعية باستعمال الاساليب الاحصائية ورصد خصائص الجودة للمواد المصنعة لضمان التزامها بمعايير معينه والحفاظ على العملية الانتاجية بحالة مستقرة ، و من ثم متابعة سير العملية الانتاجية والكشف عن التغيرات غير الطبيعية التي تمكننا من التدخل المبكر لتصحيح العملية الانتاجية ، وتنقسم لوحات السيطرة الى نوعين : النوع الاول يسمى لوحات السيطرة احادية المتغير حيث ان التركيز من خلاله على المنتج باستعمال متغير واحد فقط لتحديد العملية التي تكون تحت السيطرة او خارجها وذلك بافتراض ان متغيرات العملية الانتاجية مستقلة وهذا اجراء غير صحيح في معظم الحالات لان معظم العمليات الانتاجية تتكون من عدة متغيرات ومع التقدم التكنولوجي وتعقد العمليات الصناعية ظهرت الحاجة الى النوع الثاني الذي يتطلب مراقبة العديد من المتغيرات ، ومع استعمال الكمبيوتر وتكنولوجيا المعلومات اصبح من الممكن مراقبة عدة متغيرات في آن واحد لمراقبة العمليات لتحديد العملية اذا كانت تحت السيطرة او خارجها وهذا ما يسمى بالسيطرة الاحصائية متعددة المتغيرات (Multivariate Statistical Process Control) (MSPC) والتي تعتمد على لوحات السيطرة المعلمية متعددة المتغيرات (Parametric Multivariate Control Charts) على افتراض ان التوزيع الاساسي للعملية الانتاجية هو التوزيع الطبيعي (Normal) وفي حقيقة الامر ان هذا الافتراض لا يطبق في جميع الحالات وبذلك سوف تعطي نتائج غير دقيقة اذا كانت البيانات لا تتبع التوزيع الطبيعي في حالة تطبيق الطرائق المعلمية للعملية الانتاجية [Lowry et al:1992]، ولذلك لا يمكن مراقبة العملية الانتاجية بشكل جيد وللتخلص من هذه المشاكل يمكن استعمال لوحات السيطرة اللامعلمية متعددة المتغيرات (Multivariate Non Parametric Control Charts) لانها تعتبر بديل افضل ومن هنا تتضمن مشكلة البحث ان اغلب لوحات السيطرة تستند الى مراقبة صفة واحدة او متغير واحد يعكس جودة المنتج ولكن في الواقع الكثير من العمليات الانتاجية تعتمد على عدة متغيرات تعكس جودة العملية الانتاجية وتتطلب معرفة التوزيع ولذلك تم تطبيق طريقة تقدير كثافة اللب (Kernel Density Estimation). وان الهدف الاساسي للبحث هو عرض لوحة تقدير كثافة اللب

(Kernel Density Estimation) كطريقة لتحديد حدود السيطرة في حالة تعدد المتغيرات للمشاهدات التي لا تتبع التوزيع الطبيعي والمقارنة بينهما من خلال استعمال معيار متوسط طول المدى (ARL) لتوضيح اداء لوحة السيطرة باستعمال المحاكاة. وتكمن اهمية لوحات السيطرة اللامعلمية متعددة المتغيرات انها لا تحتاج الى معرفة التوزيع الاساسي للعملية وتكون اقل تائراً بالقيم الشاذة لذا فان استعمال لوحات السيطرة اللامعلمية متعددة المتغيرات مفيدة في مراقبة العمليات الانتاجية بغض النظر عن معرفة التوزيع الاساسي للعملية وبذلك تقدم خياراً سهلاً لتطبيق الاساليب التقليدية للعمليات الاحصائية [Zou. And Tsung:2011] وفي هذه الدراسة تم التركيز بشكل اساسي على طريقة مقدر الكثافة اللبي.

2- طرائق واساليب

1-2 مقدر الكثافة اللبي (Kernel Density Estimation (KDE)

تعد طريقة مقدر كثافة اللب (Kernel) من الطرائق اللامعلمية التي تم اقتراحها من قبل Rosenblatt) عام (1952)[13] والباحث (Parzen) عام (1962)[11] للحالات احادية البعد وتم تطويرها من قبل Cacoullos) عام (1966)[4] و (Epanechnikov) عام (1969)[5] للحالات متعددة المتغيرات لتقدير دالة الكثافة الاحتمالية للمتغير العشوائي X ويمكن تحديد الاحتمالات المرتبطة مع المتغير العشوائي X من خلال العلاقة الاتية [20]، [17]

$$P(a \leq X \leq b) = \int_a^b f(x) dx \text{ for all } a < b \quad (1)$$

وتعتبر طريقة مقدر كثافة اللب (Kernel Density Estimation) واحدة من الطرائق الأكثر استعمالاً وفائدة لتحليل البيانات بفضل المرونة والسلوك الجيد لها، وان الدافع وراء دراسة Kernel (KD) Density هو عندما تكون المتغيرات لا تتبع التوزيعات الرئيسية مثل التوزيع الطبيعي (Normal) توزيع ذو الحدين (Binomial) وتوزيع ويبل (Weibul)... الخ [14:pp1]. والهدف من استعمال مقدر كثافة اللب (KDE) هو اعادة تعديل البيانات بحيث تمكنها من الحصول على مقدرات ذات صفة تتقارب مع صفات المعلومات الحقيقية .

لنفرض ان X_1, X_2, \dots, X_n هي بيانات عينة عشوائية من المتغير العشوائي X والذي له دالة كثافة احتمال مجهولة $f(X)$ فان مقدر Kernel لهذه الدالة هو [16: pp580]

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n k(u) \quad (2)$$

$$u = \frac{x - X_i}{h}$$

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n k\left(\frac{x - X_i}{h}\right) \quad i = 1, 2, \dots, n \quad (3)$$

حيث ان :

$K(u)$: عرض الحزمة او معلمة التمهيدي او تسمى دالة النافذة وهي دالة حقيقية تقوم بإزالة الانحرافات من المقدار اي (تعديله) ولها عدة خواص منها [17],[18:pp14]:

1. $K(u) \geq 0, \text{ all } u \geq 0$ ، اي ان دالة Kernel تكون غير سالبة (non-negativity).
2. $K(u) = K(-u) \text{ all } u \geq 0$ ، اي ان دالة kernel متماثلة (symmetric).
3. $K(u) \leq 0 \text{ all } u \geq 0$ ، اي ان دالة Kernel هي دالة غير متناقصة (non-decreasing).
4. $\int_{-\infty}^{\infty} K(u) du = 1$
5. $\int_{-\infty}^{\infty} xK(u) du = 0$
6. $\lim_{|z| \rightarrow \infty} |zK(z)| = 0$
7. تحسب العزوم لدالة اللب (Kernel) كالآتي

$$M_j(K) = \int_{-\infty}^{\infty} u^j K(u) du \quad (4)$$

اما العزوم الفردية تساوي صفر [7:pp163]، [6:pp1241]، [3:pp7]

وان قيمة $\hat{f}(x)$ تكون كبيرة اذا كان هنالك عدد كبير من المشاهدات تقترب من (x) اما اذا كان عدد المشاهدات قليل فان قيمة $\hat{f}(x)$ تكون صغيرة، وفيما يأتي نذكر بعض انواع دوال (Kernel) اللامعلمية [17]، [15:pp 3]

Kernel	$K(z)$	Support
Gaussian	$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$	$(-\infty, \infty)$
Uniform	$K(u) = \frac{1}{2}$	$[-1, 1]$
Triangular	$K(u) = (1 - u)$	$[-1, 1]$
Epanchnikov	$K(u) = \frac{3}{4}(1 - u^2)$	$[-1, 1]$

بحيث اختيار دوال (Kernel) يعد مهما وضروريا للحصول على مقدرات تتقارب من الخواص الاحصائية وبما ان اختيار عرض الحزمة (Bandwidth) مهم لأنه يؤثر في التحيز والتباين يجب ان يكون الملائم ولا يقل اهمية عن اختيار دالة اللب الافضل.

1-1-2 اختيار عرض الحزمة (Selection of Band Width)

معلمة التمهيد (Smoothing Parameter) يرمز لها بالرمز (H) في حالة متعدد المتغيرات وبالرمز (h) للأحادية المتغير (البعد) والتي تسمى عرض الحزمة (Bandwidth). وللحصول على مقدر (Kernel) فإن أهم الخطوات هو اختيار عرض الحزمة (Bandwidth) وذلك لأن اختيار عرض الحزمة (h) المناسب يعطي أفضل تقدير لدالة (Kernel) لكونها تؤثر بشكل كبير في التحيز والتباين عندما يزداد عرض الحزمة (h) يزداد مقدار التحيز ويقل التباين والعكس صحيح ولها أهمية كبيرة في حالة متعدد المتغيرات بدلاً من حالة احادية المتغير وتؤثر معلمة التمهيد في درجة التمهيد للمنحنى واقترابه من المنحنى الحقيقي [14:pp75] ومن خصائص عرض الحزمة [9:pp 278]، [18:pp20]،

$$\lim_{n \rightarrow \infty} h = 0 \quad (5)$$

$$\lim_{n \rightarrow \infty} nh = \infty \quad (6)$$

ليجاد عرض الحزمة (BandWidth) توجد طرائق عديدة نستعمل منها طريقة المربعات الصغرى للعبور الشرعي لذلك يعتبر اختيارها مهما في عملية التقريب وللحصول على التقارب الملائم يجب استعمال الطريقة المناسبة لغرض الموازنة بين التباين والتحيز بحيث تكون نسبة الخطأ أقل ما يمكن، هنالك عدة طرائق لحساب عرض الحزمة نستعمل الطريقة الآتية

2-1-2 طريقة المربعات الصغرى للعبور الشرعي (Least Squared Cross Validation) (LSCV) تعد طريقة المربعات الصغرى للعبور الشرعي من الطرائق سهلة التطبيق في ايجاد قيمة معلمة التمهيد (h) واكثرها استعمالا اقترحت من قبل (Rudemo) عام (1982) من خلال تصغير مربع الخطأ التكاملي (ISE) (Integrated Squared Error) وفي عام (1984) من قبل (Bowman) [2] وتعتبر واحدة من افضل الطرائق لإيجاد عرض الحزمة (h) وكما يأتي

[15:pp 12-13],[17:pp.38-40],[10:pp.32-34]، [1:pp 251-252]

$$ISE(h) = ISE(\hat{f}_h) = \int_{-\infty}^{\infty} [\hat{f}(x) - f(x)]^2 dx \quad (7)$$

$$= \int_{-\infty}^{\infty} \hat{f}^2(x) dx - 2 \int_{-\infty}^{\infty} \hat{f}(x)f(x) dx + \int_{-\infty}^{\infty} f^2(x) dx \quad (8)$$

اذ ان

$\int_{-\infty}^{\infty} \hat{f}^2(x) dx$ قيمة تحسب من البيانات والحد المطلوب تقديره في المعادلة اعلاه هو

$$\int_{-\infty}^{\infty} \hat{f}(x)f(x) dx \quad (9)$$

لان الحد الاخير $\int_{-\infty}^{\infty} f^2(x) dx$ مقدار ثابت لا يعتمد على قيمة معلمة التمهيد (h)

ننقل الحد الثابت الى الطرف الاخر لتقليل المقدار وبطرح الحد الثابت نحصل على الاتي

$$ISE - \int_{-\infty}^{\infty} f^2(x) dx = \int_{-\infty}^{\infty} \hat{f}^2(x) dx - 2 \int_{-\infty}^{\infty} \hat{f}(x)f(x) dx \quad (10)$$

حيث ان

$$\int_{-\infty}^{\infty} \hat{f}(x)f(x) dx = E\hat{f}(x) \quad (11)$$

اذ ان

باستعمال (Leave-One-Out Cross validation) لتقدير المقدار $\int_{-\infty}^{\infty} \hat{f}(x)f(x) dx$ يمثل $E\hat{f}(x)$

والذي يساوي

$$E\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n f_{-i}(X_i) \quad (12)$$

ان الغرض الاساسي من طريقة المربعات الصغرى للعبور الشرعي هو ايجاد مقدار $ISE(h)$ من البيانات وتصغيره بواسطة h.

بتعويض معادلة (12) في معادلة (10) نحصل على (LSCV) بالشكل الاتي

$$LSCV(h) = \int_{-\infty}^{\infty} \hat{f}^2(x) dx - \frac{2}{n} \sum_{i=1}^n \hat{f}_{-i}(X_i) \quad (13)$$

في عام (1987) تم تسميت هذه الطريقة من قبل (Terrell) و (Scott) بالعبور الشرعي غير المتحيز لدالة $LSCV(h)$ لذلك

$$E(LSCV(h)) = E(LSCV(h)) - \int_{-\infty}^{\infty} f^2(x) dx \quad (14)$$

$$= MISE(\hat{f}(x)) - \int_{-\infty}^{\infty} f^2(x) dx$$

علماء ان

$$\hat{f}_{-i}(X_i) = \frac{1}{(n-1)h} \sum_{j \neq i}^n K\left(\frac{X_i - X_j}{h}\right) \quad (15)$$

وكذلك يمكن كتابة $\int_{-\infty}^{\infty} \hat{f}(x) dx$ وفق الصيغة الآتية

$$\int_{-\infty}^{\infty} \hat{f}(x) dx = \frac{2}{nh} \sum_{i=1}^n \sum_{j=1}^n K * K\left(\frac{X_i - X_j}{h}\right) \quad (16)$$

حيث ان :

$K * K(u)$: تمثل الالتفاف لدالة اللب (Kernel)

$$u = \left(\frac{X_i - X_j}{h}\right) \quad (17)$$

من خلال تعويض المعادلة (16) ومعادلة (17) في معادلة (13) نحصل على

$$LSCV(h) = \frac{2}{nh} \sum_{i=1}^n \sum_{j=1}^n K * K\left(\frac{X_i - X_j}{h}\right) - \frac{2}{n} \sum_{i=1}^n \hat{f}_{-i}(X_i) \quad (18)$$

$$= \frac{2}{n^2} h^{-1} \left[\frac{n}{2} K * K(0) + \sum_{i=1}^{n-1} \sum_{j=1}^n K * K\left(\frac{X_i - X_j}{h}\right) - \frac{2n}{n-1} \left(\frac{X_i - X_j}{h}\right) \right] \quad (19)$$

علماء ان

$$K * K(0) = \int K(z)K(0 - z) dz = \int K^2(z) dz = \|K\|_2^2 = C_k = R(K)$$

وبذلك فان

$$LSCV(h) + \int_{-\infty}^{\infty} f^2(x) dx \quad (20)$$

هو مقدار غير متحيز الى MISE ولذلك يطلق عليها طريقة العبور الشرعي غير متحيز وان المعلمة التمهيدية (h) والتي تعمل على تقليل الدالة

$$h_{LSCV} = \operatorname{argmin}_h LSCV(h) \quad (21)$$

خوارزمية طريقة المربعات الصغرى للعبور الشرعي (LSCV) [17:pp.41] الخطوة الاولى: افرض قيمة اولية للمعلمة التمهيدية (h) وفق الصيغة الآتية.

$$h = 1.06\sigma n^{-\frac{1}{5}} \quad (22)$$

حيث ان

σ : تمثل الانحراف المعياري

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}} \quad (23)$$

$h = 1.06\sigma n^{-\frac{1}{5}}$: تسمى معلمة التمهيد الثابتة (Fixed Bandwidth) تستخدم معلمة التمهيد الثابتة على طول الخط الحقيقي المستخدم لتقدير دالة الاحتمالية.

الخطوة الثانية: أحسب المقدار (Leave- One - Out) ($\hat{f}_{-i}(X_i)$) كما في المعادلة (15)

$$\hat{f}_{-i}(X_i) = \frac{1}{(n-1)h} \sum_{j \neq i}^n K\left(\frac{X_i - X_j}{h}\right)$$

الخطوة الثالثة: كون دالة العبور الشرعي غير المتحيز حسب المعادلة (18)

$$LSCV(h) = \frac{2}{nh} \sum_{i=1}^n \sum_{j=1}^n K * K\left(\frac{X_i - X_j}{h}\right) - \frac{2}{n} \sum_{i=1}^n \hat{f}_{-i}(X_i)$$

الخطوة الرابعة: كرر الخطوات السابقة وفي كل مره تحذف مشاهدة واحدة

الخطوة الخامسة: اختار المعلمة التمهيدية (h) المقابلة لأصغر LSCV(h)

بحيث ان

$$h_{LSCV} = \operatorname{argmin}_h LSCV(h) \quad (24)$$

الخطوة السادسة: النهاية

3-1-2 الحد الاعلى للسيطرة للمقدر اللبي

يمكن تحديد حد السيطرة (Control Limit) وفق المقدر اللبي لدالة الكثافة الاحتمالية لمشاهدات العينة وبحسب قيمة α المحددة التي تقع قيمتها بين الصفر والواحد اذ ان تم الاعتماد على احصاءات هوتلك متعددة المتغيرات وبذلك يمكن حساب الحد الاعلى للسيطرة وفق الصيغة الاتية: [12:pp 651]

$$CL_{\text{Kernel}} = \hat{F}_h(t)^{-1}(1-\alpha) \quad (25)$$

$$\hat{f}_k(t) = 1/n \sum_{j=1}^n \frac{1}{h} k(u)$$

$$u = \frac{t-T_j^2}{h} \quad (26)$$

$$\hat{f}_h(t) = 1/n \sum_{j=1}^n \frac{1}{h} k\left(\frac{t-T_j^2}{h}\right) \quad (27)$$

حيث ان

$$T_j^2 = (X_j - \bar{X})' s^{-1} (X_j - \bar{X})$$

حيث دالة الكثافة التجميعية للمقدر اللبي وفق الاتي:

$$\hat{F}_h(t) = 1/n \sum_{j=1}^n K\left(\frac{t-T_j^2}{h}\right) \quad (28)$$

اذ ان

$$\hat{F}_h^{-1}(t) = \inf \{ t \mid 1/n \sum_{j=1}^n K(u) \geq q \} \quad 0 < q < 1 \quad (29)$$

حيث ان:

$$q = 1 - \alpha$$

inf: يمثل أصغر قيمة داخل او خارج المجموعة

$$LCL=0$$

(30)

3- الجانب التجريبي

يتضمن أسلوب المحاكاة (Simulation) دراسة أداء لوحة السيطرة مقدر الكثافة اللبي ولنوعين من التوزيعات التوزيع الطبيعي متعدد المتغيرات توزيع t متعدد المتغيرات وبدرجة حرية مختلفة وتم تكرار التجربة (1000) مرة باستعمال عشرة متغيرات وبجسم عينة (120) مشاهدة تمثل خصائص الفيزيائية والكيميائية للمشروع ماء الشطرة القديم التي اجريت عليها المحاكاة ثم تم استعمال معيار متوسط طول المدى (ARL) لبيان كفاءة لوحة السيطرة المستعملة في الكشف عن المتغيرات داخل حدود السيطرة او خارجها بمستويات معنوية مختلفة (0.05,0.02,0.01,0.005,0.002,0.001) على التوالي ولقد تم كتابة برنامج المحاكاة باستعمال برنامج (R 3.5.2) لمعرفة دقة نتائج الطريقة المستعملة في الكشف عن نسب المشاهدات التي تكون داخل لوحات السيطرة و خارجها .

4- مناقشة المحاكاة

سنعرض النتائج في الجدول رقم (1) وجدول رقم (2) حيث يتم بيان معيار متوسط طول المدى للتوزيع الطبيعي وتوزيع (t) متعدد المتغيرات لتوضيح عمل الطرائق المستعملة في السيطرة النوعية .
جدول رقم (1) متوسط طول المدى (ARL) الافتراضي للتوزيع الطبيعي

Distribution	Alpha	$ARL = \frac{1}{\alpha}$	KDE Gaussian	KDE Uniform	KDE Epanchnikov
(Multivariate Normal distribution)	0.001	1000	1000	992.71	989.99
	0.002	500	500	596.71	493.99
	0.005	200	199.92	199.09	198.7
	0.01	100	100	99.91	99.59
	0.02	50	50	50.31	50.29
	0.05	20	20	20.55	20.71

في الجدول رقم (1) اعلاه يبين عند استعمال معيار متوسط طول المدى (ARL) عند القيمة $\text{Alpha}=0.001$ يمكن ملاحظة ان لوحة KDE-Gaussian هي افضل لوحة للسيطرة النوعية والتي كانت قيمة متوسط طول المدى الافتراضية باستعمال هذه الطريقة هي (1000) وهي الاقرب الى القيمة الحقيقية لمعيار متوسط طول المدى ($\text{ARL}=1000$) ثم لوحة KDE-Uniform والتي اعطت قيمة افتراضية (992.71) وتأتي لوحة KDE-Epanchnikov بعدها التي تقترب عن قيمة (ARL) الحقيقية بمتوسط طول المدى الافتراضي (989.99)

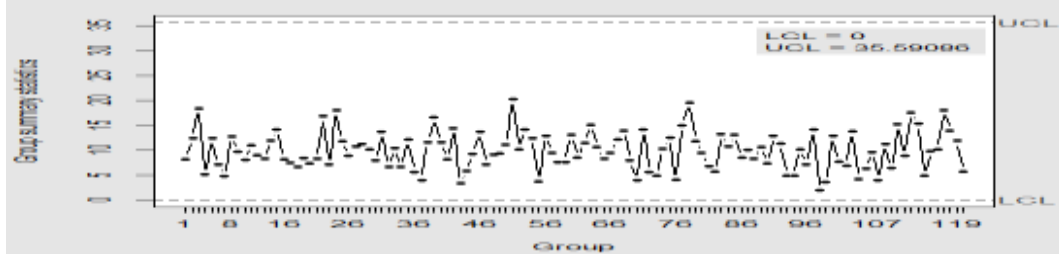
وعند قيمة $\text{Alpha}=0.002$ يمكن ملاحظة ان لوحة KDE-Gaussian هي افضل لوحة للسيطرة النوعية والتي كانت قيمة متوسط طول المدى الافتراضية باستعمال هذه الطريقة هي (500) وهي الاقرب الى القيمة الحقيقية للمعيار متوسط طول المدى ($\text{ARL}=500$) ثم تأتي بعدها لوحة KDE-Epanchnikov والتي اعطت قيمة افتراضية (596.71) ثم لوحة KDE-Uniform التي تقترب عن قيمة (ARL) الحقيقية بمتوسط طول المدى الافتراضي (493.99)

وعند قيمة $\text{Alpha}=0.005$ يمكن ملاحظة ان لوحة KDE-Gaussian هي افضل لوحة للسيطرة النوعية والتي كانت متوسط طول المدى الافتراضية باستعمال هذه الطريقة هي (199.92) وهي الاقرب الى القيمة الحقيقية للمعيار متوسط طول المدى ($\text{ARL}=200$) ثم بعدها لوحة KDE-Uniform. ثاني افضل لوحة للسيطرة والتي اعطت قيمة لمتوسط طول المدى الافتراضية هي (199.09) ومن ثم بعدها لوحة KDE-Epanchnikov التي اعطت قيمة لمتوسط طول المدى الافتراضية هي (198.7)

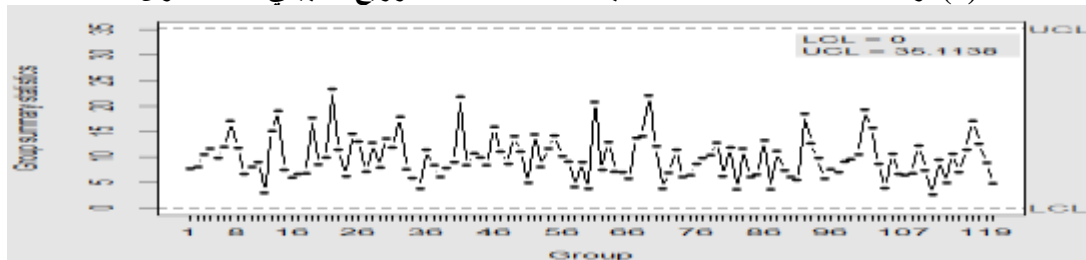
وعند قيمة $\text{Alpha}=0.01$ يمكن ملاحظة ان لوحة KDE-Gaussian هي افضل لوحة للسيطرة النوعية والتي كانت قيمة متوسط طول المدى الافتراضية باستعمال هذه الطريقة هي (100) وهي الاقرب الى القيمة الحقيقية للمعيار متوسط طول المدى ($\text{ARL}=100$) ثم بعدها لوحة السيطرة KDE-Uniform. ثاني افضل لوحة للسيطرة والتي اعطت قيمة لمتوسط طول المدى الافتراضية هي (99.91) ومن ثم بعدها لوحة KDE-Epanchnikov التي اعطت قيمة لمتوسط طول المدى الافتراضية

وعند قيمة $\text{Alpha}=0.02$ يمكن ملاحظة ان لوحة KDE-Gaussian هي افضل لوحة للسيطرة النوعية والتي كانت قيمة متوسط طول المدى الافتراضية باستعمال هذه الطريقة هي (50) وهي الاقرب الى القيمة الحقيقية للمعيار متوسط طول المدى ($\text{ARL}=50$) ثم بعدها لوحة السيطرة KDE-Epanchnikov. ثاني افضل لوحة للسيطرة والتي اعطت قيمة لمتوسط طول المدى الافتراضية هي (50.29) ومن ثم بعدها لوحة KDE-Uniform التي اعطت قيمة لمتوسط طول المدى الافتراضية هي (50.31)

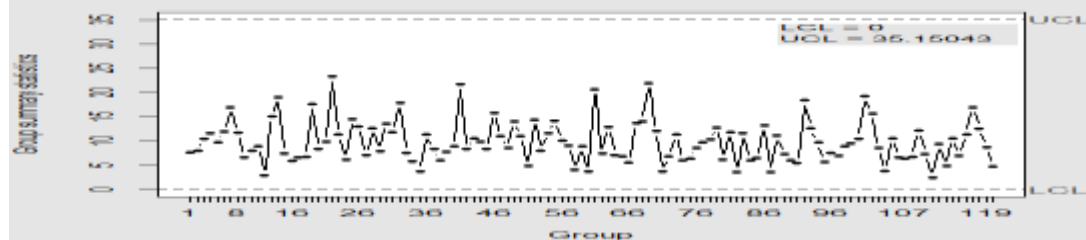
وعند قيمة $\text{Alpha}=0.05$ يمكن ملاحظة ان لوحة KDE-Gaussian هي افضل لوحة للسيطرة النوعية والتي كانت قيمة متوسط طول المدى الافتراضية باستعمال هذه الطريقة هي (20) وهي الاقرب الى القيمة الحقيقية للمعيار متوسط طول المدى ($\text{ARL}=20$) ثم بعدها لوحة السيطرة KDE-Uniform. ثاني افضل لوحة للسيطرة والتي اعطت قيمة لمتوسط طول المدى الافتراضية هي (20.55) ومن ثم بعدها لوحة KDE-Epanchnikov التي اعطت قيمة لمتوسط طول المدى الافتراضية هي (20.71) والاشكال الاتية تبين افضلية لوحة السيطرة KDE-Gaussian على وفق نتائج المحاكاة في جدول رقم (1)



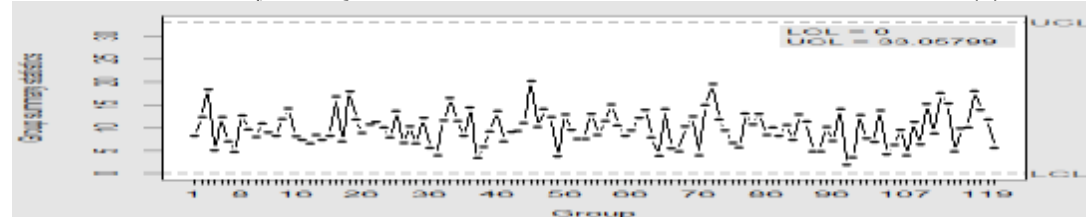
الشكل (1) لوحة KDE-GAUSSIAN باستعمال المحاكاة للتوزيع الطبيعي عند مستوى 0.001



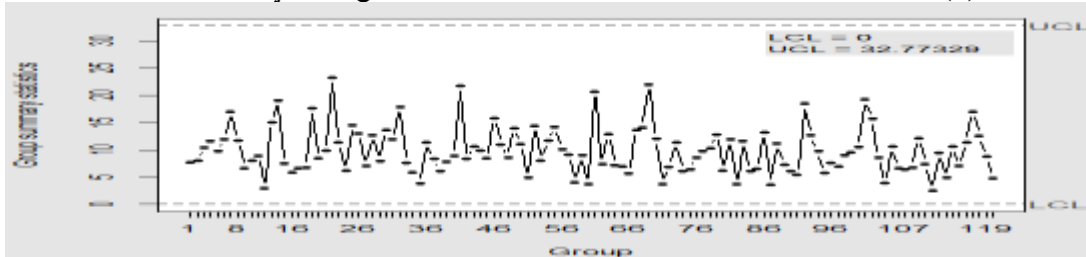
الشكل (2) لوحة السيطرة KDE-Epanchnikov متعددة المتغيرات للتوزيع الطبيعي عند مستوى 0.001



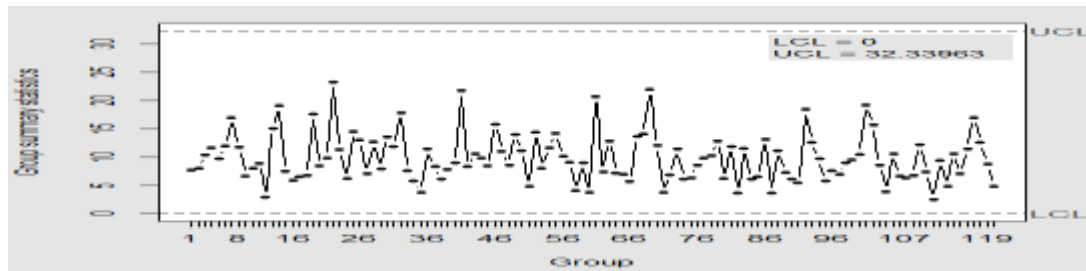
الشكل (3) لوحة السيطرة KDE-Uniform متعددة المتغيرات للتوزيع الطبيعي عند مستوى 0.001



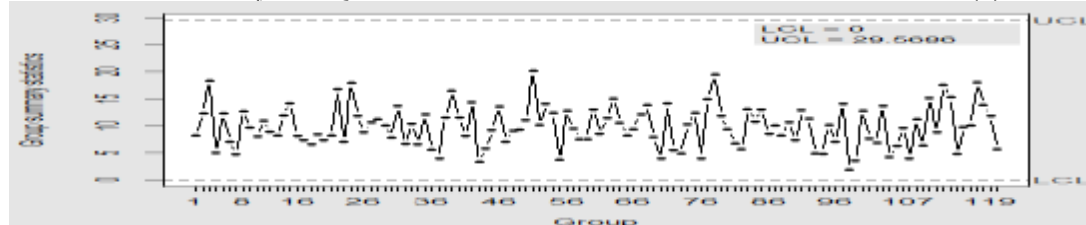
الشكل (4) لوحة KDE-GAUSSIAN باستعمال المحاكاة للتوزيع الطبيعي عند مستوى 0.002



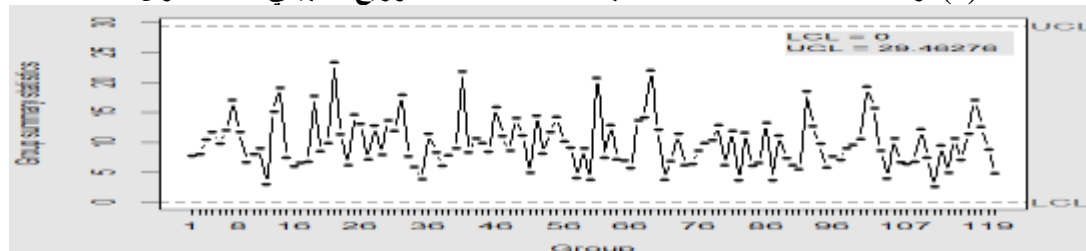
الشكل (5) لوحة السيطرة KDE-Epanchnikov متعددة المتغيرات للتوزيع الطبيعي عند مستوى 0.002



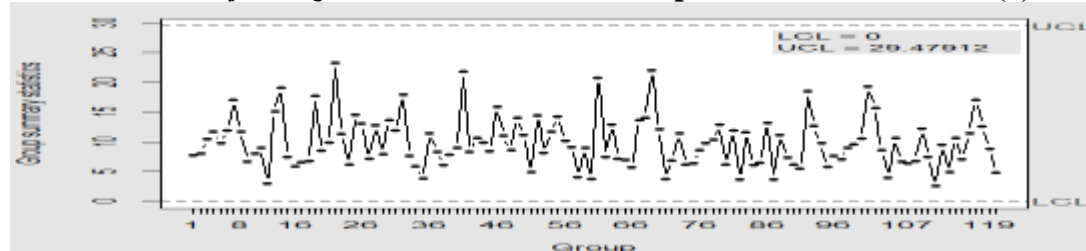
الشكل (6) لوحة السيطرة KDE-Uniform متعددة المتغيرات للتوزيع الطبيعي عند مستوى 0.002



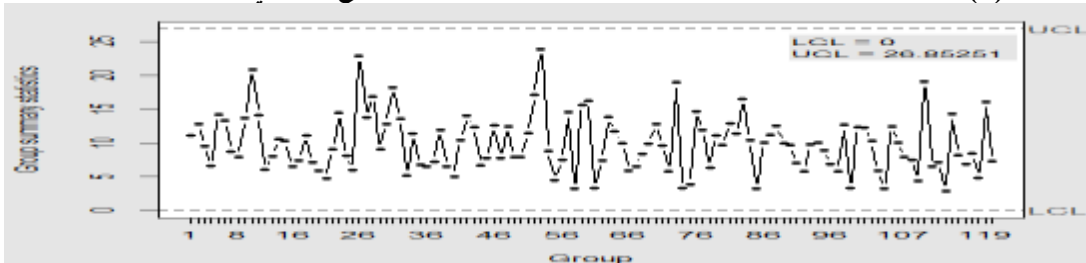
الشكل (7) لوحة KDE-GAUSSIAN باستخدام المحاكاة للتوزيع الطبيعي عند مستوى 0.005



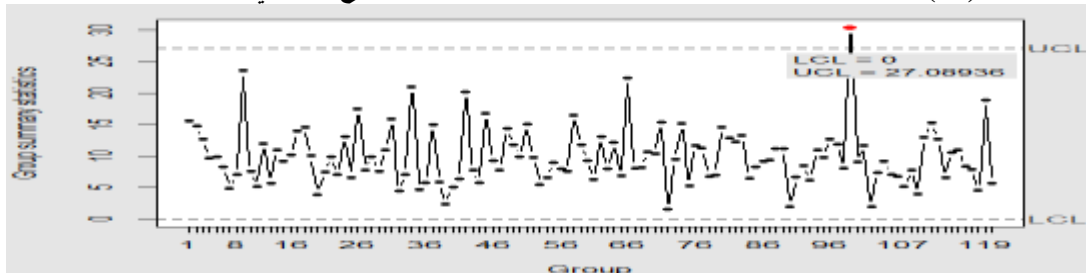
الشكل (8) لوحة السيطرة KDE-Epanchnikov متعددة المتغيرات للتوزيع الطبيعي عند مستوى 0.005



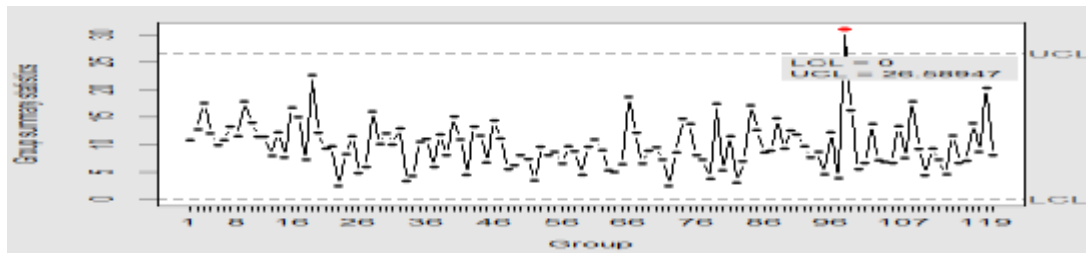
الشكل (9) لوحة السيطرة KDE-Uniform متعددة المتغيرات للتوزيع الطبيعي عند مستوى 0.005



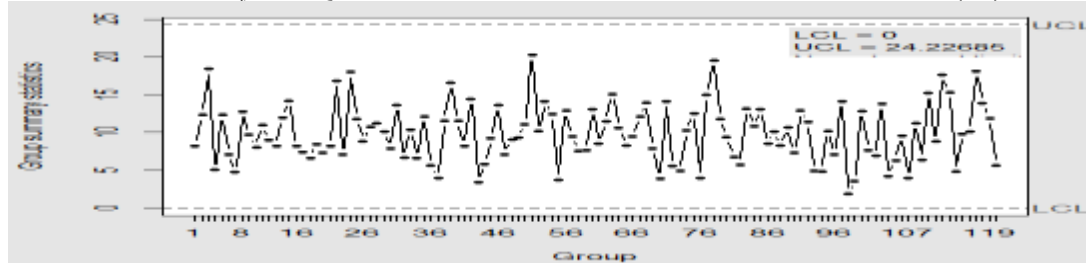
الشكل (10) لوحة KDE-GAUSSIAN باستخدام المحاكاة للتوزيع الطبيعي عند مستوى 0.01



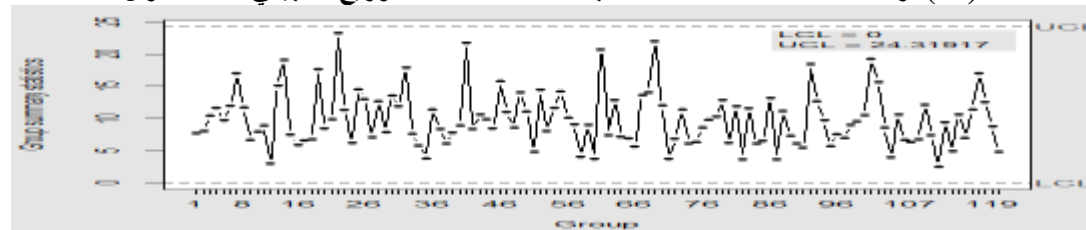
الشكل (11) لوحة السيطرة KDE-Epanchnikov متعددة المتغيرات للتوزيع الطبيعي عند مستوى 0.01



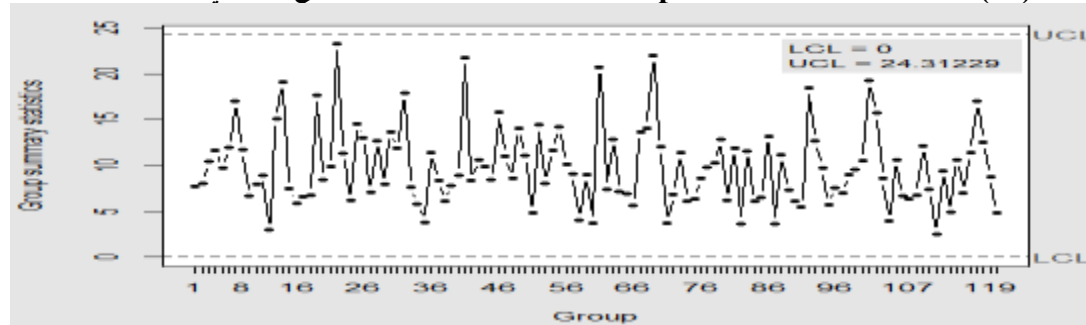
الشكل (12) لوحة السيطرة KDE-Uniform متعددة المتغيرات للتوزيع الطبيعي عند مستوى 0.01



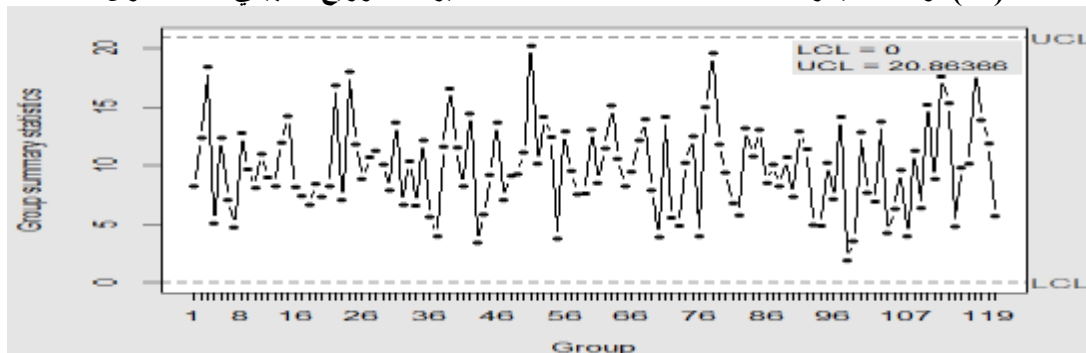
الشكل (13) لوحة KDE-GAUSSIAN باستخدام المحاكاة للتوزيع الطبيعي عند مستوى 0.02



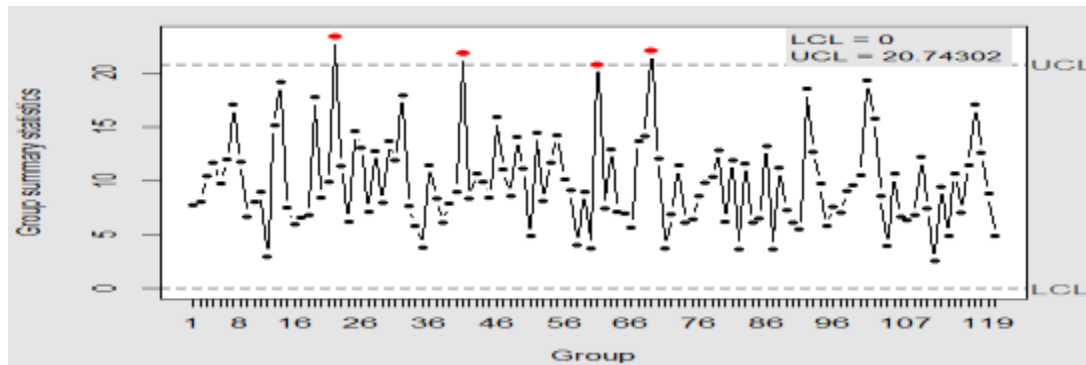
الشكل (14) لوحة السيطرة KDE-Epanchnikov متعددة المتغيرات للتوزيع الطبيعي عند مستوى 0.02



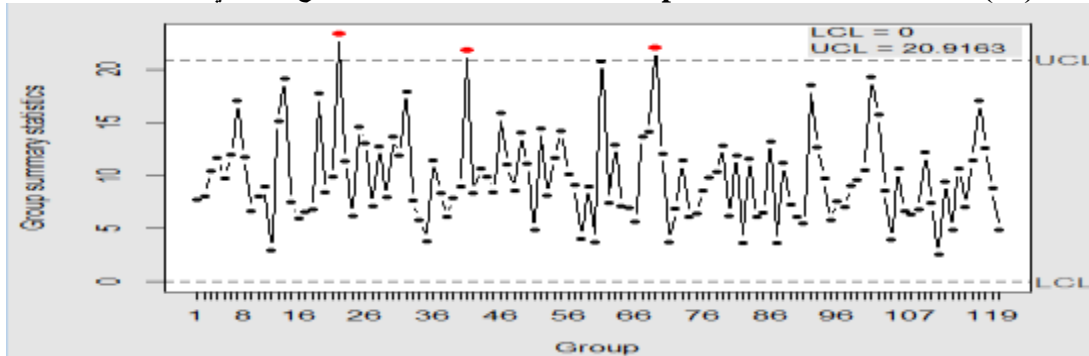
الشكل (15) لوحة السيطرة KDE-Uniform متعددة المتغيرات للتوزيع الطبيعي عند مستوى 0.02



الشكل (16) لوحة KDE-GAUSSIAN باستخدام المحاكاة للتوزيع الطبيعي عند مستوى 0.05



الشكل (17) لوحة السيطرة KDE-Epanchnikov متعددة المتغيرات للتوزيع الطبيعي عند مستوى 0.05



الشكل (18) لوحة السيطرة KDE-Uniform متعددة المتغيرات للتوزيع الطبيعي عند مستوى 0.05

اما جدول رقم (2) يوضح متوسط طول المدى للتوزيع (t) متعدد المتغيرات بدرجة حرية (100) جدول رقم (2) متوسط طول المدى (ARL) الافتراضي للتوزيع (t) بدرجة حرية (100)

Distribution	Alpha	ARL	KDE Gaussian	KDE Uniform	KDE Epanchnikov
t(100) (Multivariate t distribution)	0.001	1000	999.03	986.7	968.7
	0.002	500	499.53	493.7	485.2
	0.005	200	199.87	198.07	196.26
	0.01	100	99.93	99.3	98.4
	0.02	50	50	49.98	50.05
	0.05	20	20.01	20.42	21.04

يبين جدول رقم (2) عند استعمال المعيار متوسط طول المدى (ARL)

عند قيمة $\alpha=0.001$ يمكن ملاحظة لوحة السيطرة KDE-Gaussian هي أفضل لوحة سيطرة والتي اعطت قيمة لمتوسط طول المدى الافتراضية هي (999.03) ثم تأتي بعدها لوحة KDE-Uniform والتي اعطت قيمة افتراضية (986.7) ثم لوحة KDE-Epanchnikov التي تقترب عن قيمة (ARL) الحقيقية بمتوسط طول المدى الافتراضي (969.7)

وعند $\text{Alpha}=0.002$ يمكن ملاحظة ان لوحة KDE-Gaussian هي افضل لوحة للسيطرة النوعية والتي كانت قيمة متوسط طول المدى الافتراضية باستعمال هذه الطريقة هي (499.53) وهي الاقرب الى القيمة الحقيقية للمعيار متوسط طول المدى ($\text{ARL}=500$) ثم تأتي بعدها لوحة KDE-Uniform والتي اعطت قيمة افتراضية (493.7) ثم لوحة KDE-Epanchnikov التي تقترب عن قيمة (ARL) الحقيقية بمتوسط طول المدى الافتراضي (485.2).

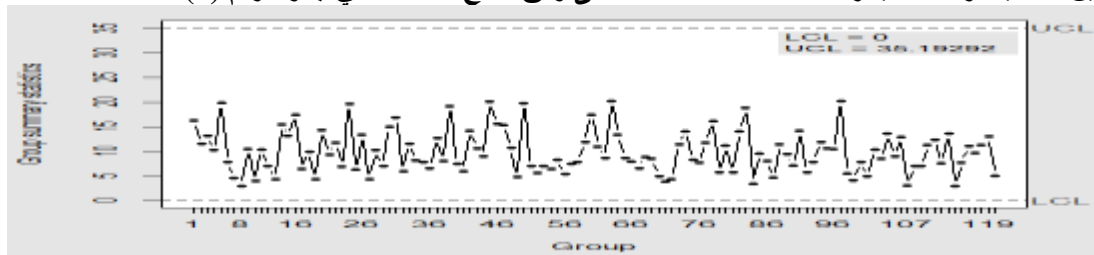
وعند قيمة $\text{Alpha}=0.005$ يمكن ملاحظة ان لوحة KDE-Gaussian هي افضل لوحة للسيطرة النوعية والتي كانت متوسط طول المدى الافتراضية باستعمال هذه الطريقة هي (199.89) وهي الاقرب الى القيمة الحقيقية للمعيار متوسط طول المدى ($\text{ARL}=200$) ثم بعدها لوحة KDE-Uniform. ثاني افضل لوحة للسيطرة والتي اعطت قيمة لمتوسط طول المدى الافتراضية هي (199.07) ثم لوحة KDE-Epanchnikov التي تقترب عن قيمة (ARL) الحقيقية بمتوسط طول المدى الافتراضي (196.26)

وعند قيمة $\text{Alpha}=0.01$ يمكن ملاحظة ان لوحة KDE-Gaussian هي افضل لوحة للسيطرة النوعية والتي كانت قيمة متوسط طول المدى الافتراضية باستعمال هذه الطريقة هي (99.93) وهي الاقرب الى القيمة الحقيقية للمعيار متوسط طول المدى ($\text{ARL}=100$) ثم بعدها لوحة السيطرة KDE-Uniform. ثاني افضل لوحة للسيطرة والتي اعطت قيمة لمتوسط طول المدى الافتراضية هي (99.3) ومن ثم بعدها لوحة KDE-Epanchnikov التي اعطت قيمة لمتوسط طول المدى الافتراضية هي (98.4)

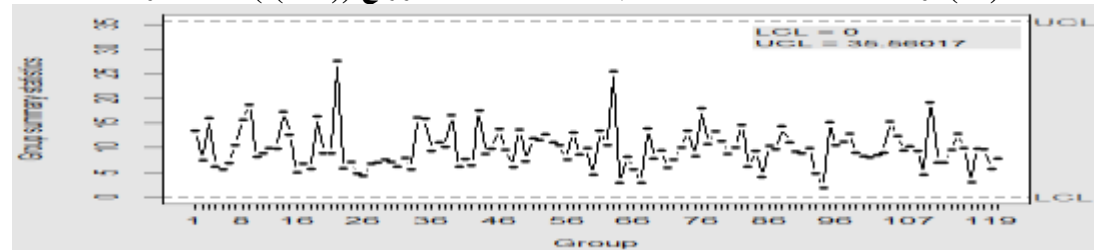
وعند قيمة $\text{Alpha}=0.02$ يمكن ملاحظة ان لوحة KDE-Gaussian هي افضل لوحة للسيطرة النوعية والتي كانت قيمة متوسط طول المدى الافتراضية باستعمال هذه الطريقة هي (50) وهي الاقرب الى القيمة الحقيقية للمعيار متوسط طول المدى ($\text{ARL}=50$) ثم بعدها لوحة السيطرة KDE-Uniform. ثاني افضل لوحة للسيطرة والتي اعطت قيمة لمتوسط طول المدى الافتراضية هي (49.98) ومن ثم بعدها لوحة KDE-Epanchnikov التي اعطت قيمة لمتوسط طول المدى الافتراضية هي (50.05)

وعند قيمة $\text{Alpha}=0.05$ يمكن ملاحظة ان لوحة KDE-Gaussian هي افضل لوحة للسيطرة النوعية والتي كانت قيمة متوسط طول المدى الافتراضية باستعمال هذه الطريقة هي (20) وهي الاقرب الى القيمة الحقيقية للمعيار متوسط طول المدى ($\text{ARL}=20$) ثم بعدها لوحة السيطرة KDE-Uniform. ثاني افضل لوحة للسيطرة والتي اعطت قيمة لمتوسط طول المدى الافتراضية هي (20.42) ومن ثم بعدها لوحة KDE-Epanchnikov التي اعطت قيمة لمتوسط طول المدى الافتراضية وهي تبتعد بقيمة (21.04)

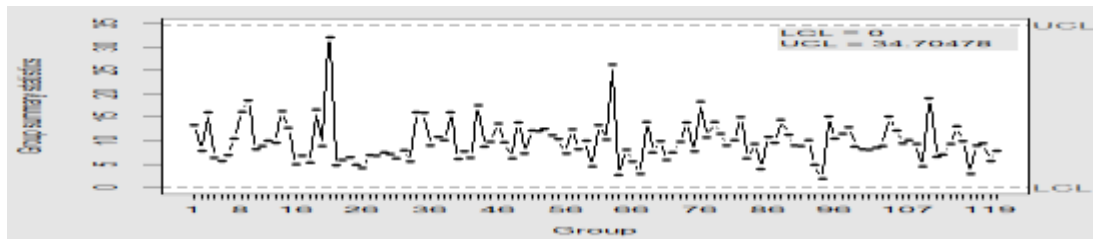
اذ ان تم استعمال درجة حرية واحدة في هذا البحث بحيث كلما كانت درجة الحرية كبيرة تقترب القيم الى معدل طول التشغيل ولذلك السيطرة النوعية تحتاج الى مستويات معنوية ذات المدى الطويل والاشكال الاتية تبين افضلية لوحة السيطرة KDE-Gaussian على وفق نتائج المحاكاة في جدول رقم (2).



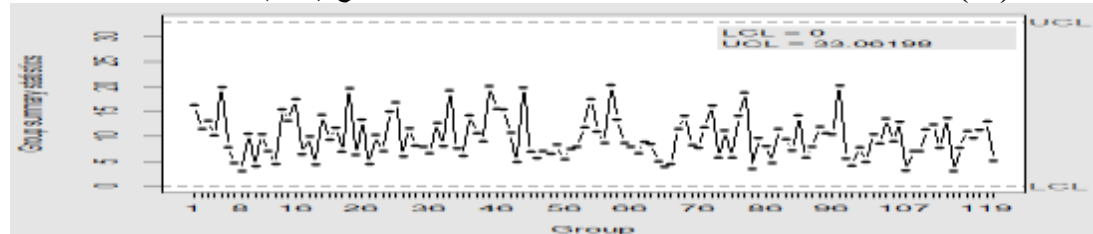
الشكل (19) لوحة KDE-GAUSSIAN باستعمال المحاكاة للتوزيع ($t(100)$) عند مستوى 0.001



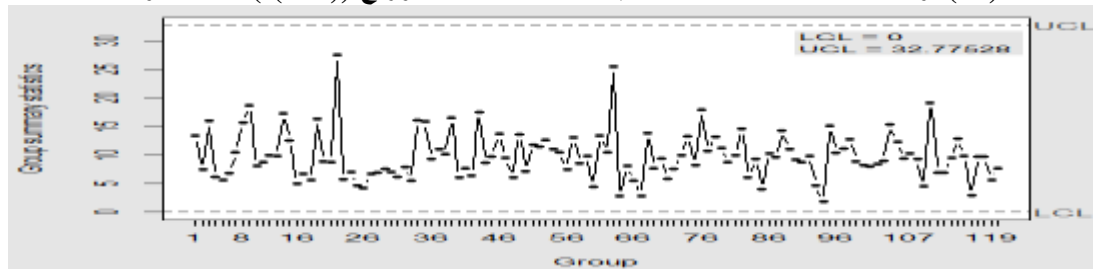
الشكل (20) لوحة السيطرة KDE-Epanchnikov متعددة المتغيرات $t(100)$ عند مستوى 0.001



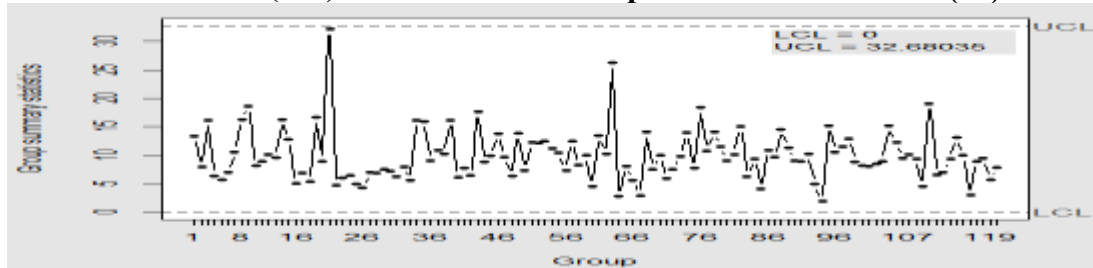
الشكل (21) لوحة السيطرة KDE-Uniform متعددة المتغيرات للتوزيع $t(100)$ عند مستوى 0.001



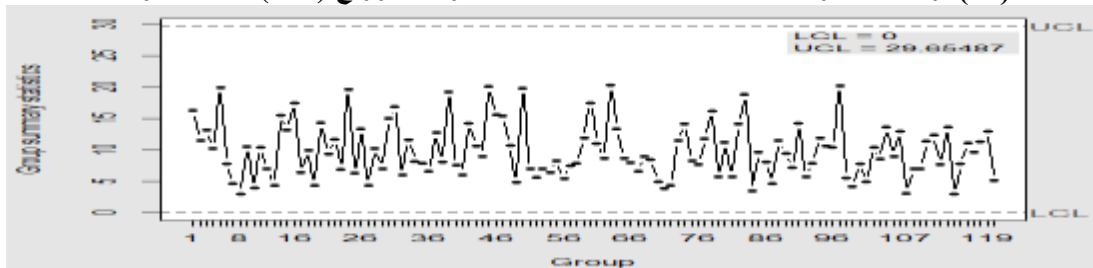
الشكل (22) لوحة KDE-GAUSSIAN باستخدام المحاكاة للتوزيع $t(100)$ عند مستوى 0.002



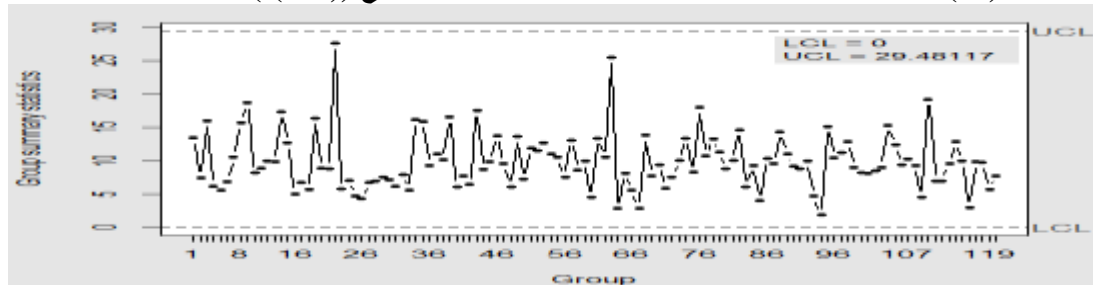
الشكل (23) لوحة السيطرة KDE-Epanchnikov متعددة المتغيرات $t(100)$ عند مستوى 0.002



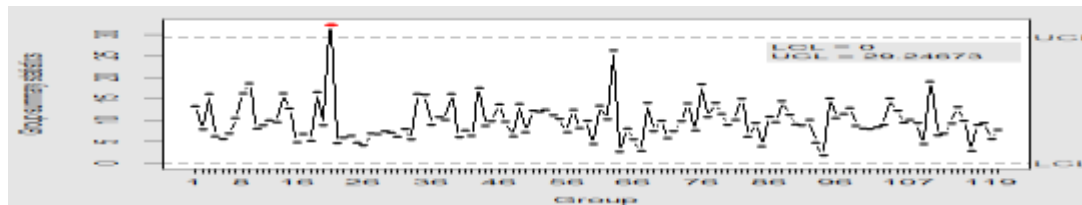
الشكل (24) لوحة السيطرة KDE-Uniform متعددة المتغيرات للتوزيع $t(100)$ عند مستوى 0.002



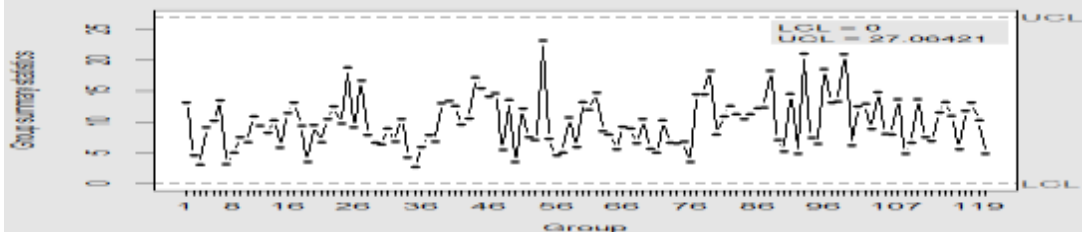
الشكل (25) لوحة KDE-GAUSSIAN باستخدام المحاكاة للتوزيع $t(100)$ عند مستوى 0.005



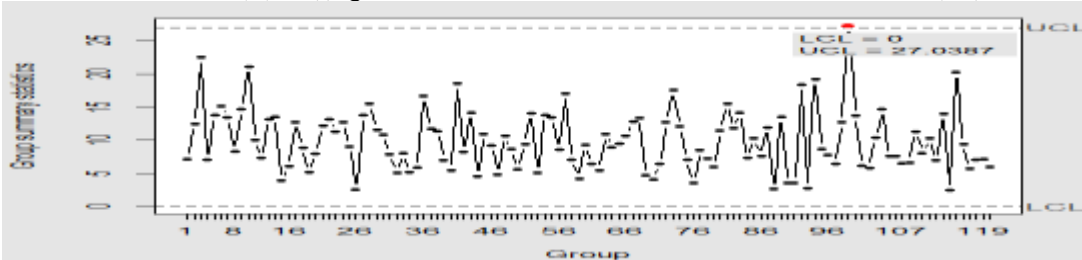
الشكل (26) لوحة السيطرة KDE-Epanchnikov متعددة المتغيرات $t(100)$ عند مستوى 0.005



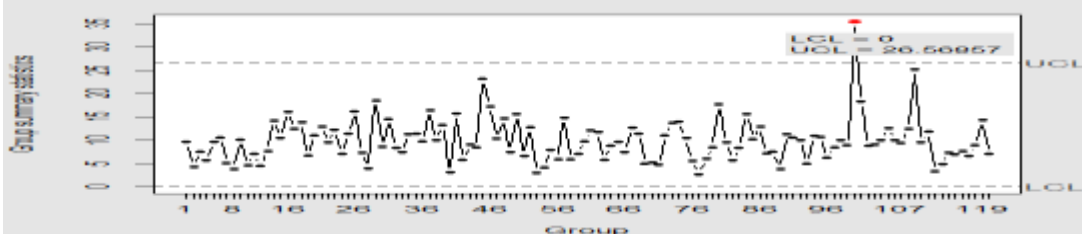
الشكل (27) لوحة السيطرة KDE-Uniform متعددة المتغيرات للتوزيع $t(100)$ عند مستوى 0.005



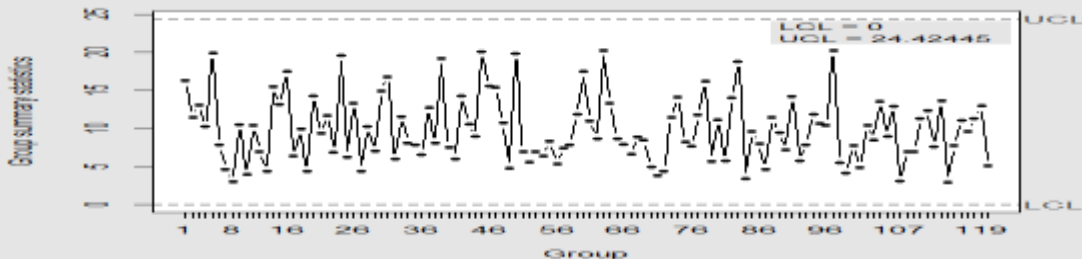
الشكل (28) لوحة KDE-GAUSSIAN باستخدام المحاكاة للتوزيع $t(100)$ عند مستوى 0.01



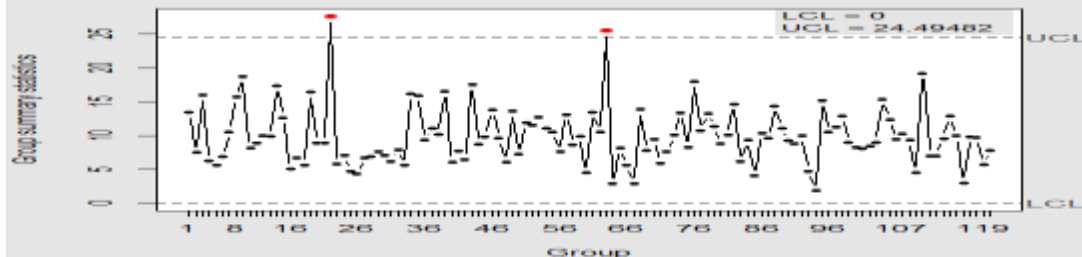
الشكل (29) لوحة السيطرة KDE-Epanchnikov متعددة المتغيرات $t(100)$ عند مستوى 0.01



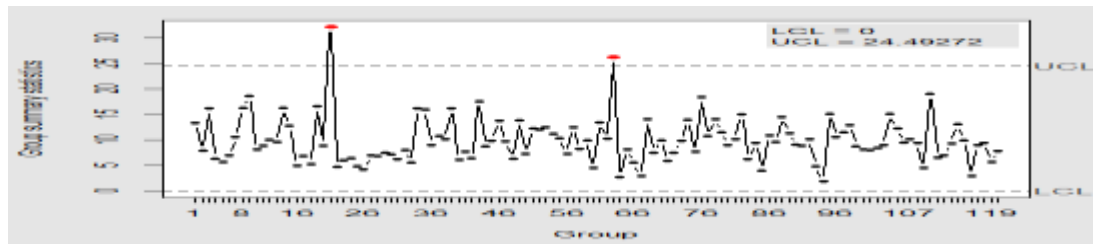
الشكل (30) لوحة السيطرة KDE-Uniform متعددة المتغيرات للتوزيع $t(100)$ عند مستوى 0.01



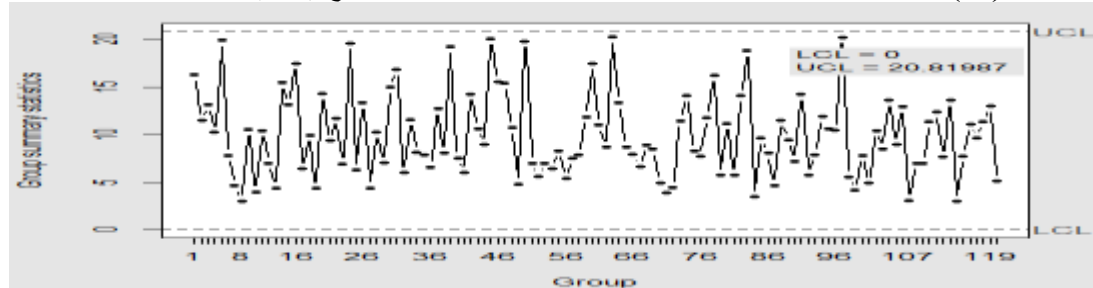
الشكل (31) لوحة KDE-GAUSSIAN باستخدام المحاكاة للتوزيع $t(100)$ عند مستوى 0.02



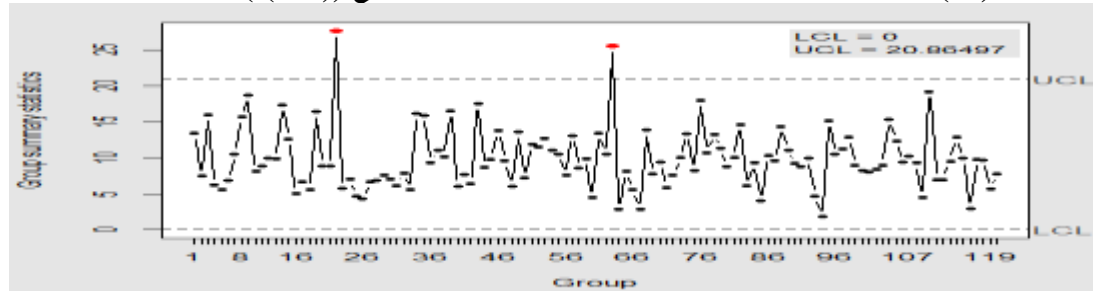
الشكل (32) لوحة السيطرة KDE-Epanchnikov متعددة المتغيرات $t(100)$ عند مستوى 0.02



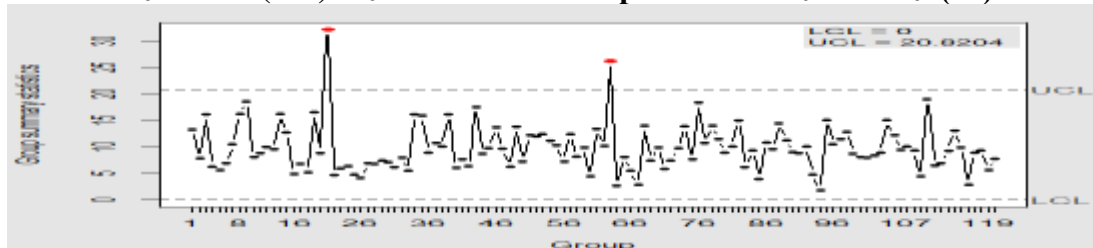
الشكل (33) لوحة السيطرة KDE-Uniform متعددة المتغيرات للتوزيع $t(100)$ عند مستوى 0.02



الشكل (34) لوحة KDE-GAUSSIAN باستخدام المحاكاة للتوزيع $t(100)$ عند مستوى 0.05



الشكل (35) لوحة السيطرة KDE-Epanchnikov متعددة المتغيرات $t(100)$ عند مستوى 0.05



الشكل (36) لوحة السيطرة KDE-Uniform متعددة المتغيرات للتوزيع $t(100)$ عند مستوى 0.05

5- الاستنتاجات والتوصيات

1-5- الاستنتاجات

من خلال استعمال المحاكاة تبين ان اسلوب طريقة مقدر كثافة اللب لدالة كاوسين في حالة التوزيع الطبيعي متعدد المتغيرات والتوزيع t متعدد المتغيرات وفق مقياس متوسط طول المدى انها تعطي نتائج دقيقة وكفاءة وفضلا عن نتائج المحاكاة لدالة كاوسين تبين ان المشاهدات داخل حد السيطرة مطابقة للمواصفات المطلوبة.

2-5- التوصيات

عدم الاعتماد على لوحات احادية المتغير لمراقبة منتج معين بل يجب تطبيق اساليب لوحات متعدد المتغيرات معلمية كاسلوب (هوتلنك) او اللامعلمية كاساليب (البوتستراپ و اشارة الرتب... الخ) لكشف عن الانحرافات والتغيرات التي تحصل في العملية الانتاجية، وكذلك تطبيق اساليب وتقنيات اخرى من لوحات السيطرة متعدد المتغيرات في كافة المجالات فضلا عن تطبيق دوال اخرى لمقدر كثافة اللب

المصادر

1. Abdel-Razzaq and Jetter, J. (2010). "A Simulation Study For The Bandwidth Selection In The Kernel Density Estimation Based On The Exact and the Asymptotic Mise", Pak. J. Statist., Vol. 26, No.(1) ,pp. 239-265.
2. Bowman, Adrian W.(1984). "An Alternative Method of Cross-Validation for the Smoothing of Density Estimates" Biometrika, Vol.71, pp. 353-360.
3. Breheny, P. "Kernel density estimation" ,STA 621: Nonparametric Statistics
4. Cacoullos, T. (1966). "Estimation of a multivariate density", Ann. Inst. Statist. Math. 18, 179-189.
5. Epanechnikov, V. A. (1969). "Non-parametric estimation of a multivariate probability density", Theor. Prob. Appl. 14, 153-158.
6. George, R. T. and David W. S.(1992). "Variable Kernel Density Estimation", The Annals of Statistics, Vol. 20, No. 3, pp. 1236-1265. <http://www.jstor.org/stable/2242011>
7. Jin, Z., & Shao, Y.(1999). "On kernel estimation of a multivariate distribution function", Statistics & Probability Letters 41 ,pp. 163-168
8. Lowry, C.A., Woodall, W.H., Champ, C.W. and Rigdon, S.E., (1992). " A multivariate exponentially weighted moving average control chart", Technometrics, Vol. 34, No.(1), pp.46-53.
9. Molanes-López, E. M. and Cao, R. (2008). "Plug-in bandwidth selector for the kernel relative density estimator" , Annals of the Institute of Statistical Mathematics AISM, Vol. 60, NO. 2, pp. 273-300.
10. Naif, Q. N. and Shakir, A. M (2018). " Experimental Comparison of some of the Classical and Bayesian Nonparametric Estimators for some Reliability Systems", Al-Rafidain University College of Science Magazine, No.(42) ,pp.30-51
11. Parzen, E. (1962). "On estimation of a probability density function and mode", Ann. Stat. Vol.33, PP1065-1076.
12. Phaladianon, P., Kim, S.B., Chen, C.P., Beak, G.J. & Park, S.K., (2011). "Bootstrap-Based T2 Multivariate Control Chart". Journal of Communication in Statistics, Vol 40, No.(5), pp.645-662.
13. Rosenblatt, M. (1956). "Remarks on some nonparametric estimates of a density function", Ann. Math. Stat. Vol. 27, NO. 3, PP.832-837.
14. Silverman, B.W. (1986). "Density estimation for statistics and data analysis", Chapman and Hall, London
15. Turlach, B. A. (1993). "Bandwidth Selection in Kernel Density Estimation: A Review", C.O.R.E. and Institut de Statistique. Université Catholique de Louvain
16. Yang, L. & Marron, J. S. (1999). "Iterated transformation-kernel density estimation", JASA, Vol.94, No.446, PP580-589.
17. Munaf Y.Hmood. (2005). " Comparing Nonparametric Estimators For Probability Density Estimation", PH.D Dissertation University of Baghdad, College of Administration & Economics
18. Munaf Y.Hmood. (2000). " Comparing Nonparametric Kernel Estimators to Estimate Regression Functions", A thesis for degree of MSc in Statistic, University of Baghdad/ College of Administration & Economics.
19. Zou, C. and Tsung, F., (2011). "A multivariate sign EWMA control chart", Technometrics, Vol. 53, No.(1), pp.84-97.
20. Zucchini, W. (2003). "Applied smoothing techniques, part I, kernel density estimation", (Through Internet).

Using Kernel Density Estimator To Determine the Limits of Multivariate Control Charts.

Fahad Hussein Enad

Baghdad University/ College of
Administration & Economics

07828940763

fahad89_1@yahoo.com

D.Asmaa Ghaleeb Jaber

Baghdad University/ College of
Administration & Economics

Drasmaa.ghalib@coadec.uobaghdad.edu.iq

Received: 14/2/2020

Accepted :20/9/2020

Published :December / 2020



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International \(CC BY-NC 4.0\)](https://creativecommons.org/licenses/by-nc/4.0/)

Abstract

Quality control is an effective statistical tool in the field of controlling the productivity to monitor and conform the manufactured products to the standard qualities and the certified criteria for some products and services and its main purpose is to cope with the production and industrial development in the business and competition market. Quality control charts are used to monitor the qualitative properties of the production procedures in addition to detecting the abnormal deviations in the production procedure. The multivariate Kernel Density Estimator control charts method was used which is one of the nonparametric methods that doesn't require any assumptions regarding the distribution of the data or determine the control limits in monitoring the production procedure when the data doesn't follow the normal distribution or has an unknown distribution. The aim of this paper is to monitor the production procedure throughout a number of variables simultaneously to reflect the quality of the produced material. Simulation experiments were used with deferent significance levels to illustrate the way in which the multivariate Kernel Density Estimator charts method work while adopting the average range length criterion to demonstrate the performance and efficiency of the used methods. The results show that the method of nonparametric Kernel Density Estimator charts for the Gaussian Kernel had a good performance especially at significance levels with long range

Keywords \ Quality Control, Kernel Density Estimation, Gaussian Kernel, Least Squares Cross Validation