مجلة تكريت للعلوم الصرفة، 22 (3) 2017
ISSN: 1813 – 1662 (Print)
E-ISSN: 2415 – 1726 (On Line)

تكوين لوحة سيطرة باستخدام نموذج فضاء الحالة لبيانات ذات توزيع t متعدد المتغيرات

هيفاء عبدالجواد سعيد[1] ، دلشاد شاكر اسماعيل بوتاني[2] ، عدنان مصطفى حسين السنجاري[1]

[1]قسم الإحصاء والمعلوماتية ، كلية علوم الحاسوب والرياضيات ، جامعة الموصل ، الموصل ، العراق

[2]قسم الإحصاء ، كلية الإدارة والاقتصاد ، جامعة صلاح الدين ، أربيل ، اربيل ، العراق

الملخص

إن الغرض الرئيسي من هذا البحث هو تكوين لوحة للسيطرة على كل من المعدل والتباين معاً لبيانات تتبع توزيع t متعدد المتغيرات باستخدام نموذج فضاء الحالة (State Space Model) وذلك من خلال استخدام سلسلة عوامل بيز (Bayes' factors)، حيث يتم تكوين اللوحة عندما يكون متجه المعلمات غير معلوم ومصفوفة التباين معلومة. بعد ذلك يتم رسم سلسلة هذه العوامل على اللوحة المحورة لـ EWMA أحادية المتغير بعد أن يتم فرض أن السلسلة تتبع نموذج ARMA(1,1). أخيراً يتم تطبيق النموذج المفترض على البيانات المولدة.

الكلمات المفتاحية: نموذج فضاء الحالة، عوامل بيز، EWMA المحورة، ARMA.

1- المقدمة

يعتبر موضوع السيطرة النوعية (Control quality) من المواضيع بالغة الأهمية في علم الإحصاء الحديث، وذلك لأنه فرع الإحصاء المسؤول عن اتخاذ القرار بوقف العملية الإنتاجية أو استمرارها، ومن أشهر الأساليب الإحصائية في السيطرة النوعية أسلوب لوحات السيطرة (Control chart) [1].

فتح العالم الأمريكي والتر شيوارت (Walter Shewhart) باب تكوين اللوحات الإحصائية في عشرينيات القرن الماضي، حيث درس لوحة السيطرة لمتغير واحد بفرض الاستقلال والتوزيع الطبيعي، ومنذ ذلك الوقت توالت التطورات الكبيرة في هذا الجانب، حيث أن بعض الباحثين طوروا اللوحة لتشمل بيانات فيها نوعاً ما بعض الارتباط، وقام آخرون بتكوين لوحات سيطرة تتضمن عدة متغيرات كما قام آخرون بتكوين لوحات في حالة عدم توفر شرط التوزيع الطبيعي [2].

إن شرط طبيعية البيانات يعد من أكثر الشروط والافتراضات صعوبة وذلك لأن في كثير من الأحيان تحتوي البيانات على قيم شاذة (Outliers)، ولتلافي هذا الخرق في الافتراضات التي بنيت على أساسها اللوحات، سيتم في هذا البحث تكوين لوحات سيطرة لبيانات في حالة عدم توفر شرط الطبيعية وذلك باستخدام نموذج فضاء الحالة (State space) مع أسلوب بيز المتسلسل، حيث سيتم استخدام بيانات تتبع التوزيعات المختلطة (Mixed distributions) والتي تحتوي على جزء من البيانات الطبيعية والجزء الآخر هو عبارة عن بيانات شاذة، ومن هذه التوزيعات توزيع t متعدد المتغيرات.

تم استخدام نموذج فضاء الحالة في تكوين اللوحات من قبل العديد من الباحثين، حيث قام الباحثان سميث وويست في مطلع ثمانينيات القرن الماضي [3] بمراقبة السلاسل الزمنية في الجانب الحياتي عن طريق مرشح كالمن، وقام الباحث ويست [4] باستخدام أسلوب بيز مع النماذج الحركية في تكوين التوزيع التنبؤي لتوزيعات العائلة الأسية ومن ثم مقارنته مع النموذج البديل ومن ثم استخدامه في المراقبة، كما استخدم الباحث جليل [5] النماذج الديناميكية الخطية في دراسة تتعلق بالسيطرة على العملية الانتاجية لغرض اتخاذ القرار، وقام الباحث أبو ناقوس [6] بدراسة تناول فيها تكوين لوحة السيطرة النوعية أطلق عليها لوحة بيز، حيث اعتمد على إحصاء بيز والنماذج الديناميكية الخطية لغرض السيطرة على معدل نوعية المادة المنتجة، كما قامت الباحثة الرسام [7] بتكوين ثلاث لوحات إحداها للسيطرة على الانحراف المعياري وهذه تعتبر مكملة لعمل أبو ناقوس والثانية للسيطرة على الانحراف المعياري والمعدل معاً والثالثة هي لوحة ثنائية البعد للسيطرة على معدلين لصفتين بالاستناد إلى النموذج الديناميكي، وقام الزبيدي [8] بتكوين لوحتين خاصتين بالسيطرة على الصفات النوعية وبالتحديد تكوين لوحة بيز لنسبة المعيب (p-chart)، وتكوين لوحة بيز لعدد العيوب المقابلة للوحة شيوارت لعدد العيوب (c-chart) وذلك من خلال استخدام أسلوب بيز مع النماذج الحركية العامة في تكوين هذه اللوحات الجديدة، كما قام الباحثان سلفادور وكاركالوب [9] بدراسة النموذج الحركي الخطي في حالة كون الأخطاء تتبع توزيع (Matrix Normal) ، وقد اعتمد على عامل بيز (Bayes' factor) في المقارنة بين نموذجين ومن ثم استخدامه في المراقبة (Monitoring)، وفي نفس العام قام الباحثان بان وجاريت [1] بتكوين لوحة متعددة المتغيرات باستخدام نموذج فضاء الحالة وذلك بالاعتماد على لوحة هوتلينغ (Hotelling's) للأخطاء، وقام الباحث كوستاس [2] بتطوير لوحة سيطرة متعددة المتغيرات باستخدام نموذج فضاء الحالة الخطي لبيانات تتبع التوزيع الطبيعي متعدد المتغيرات.

القسم الثاني من هذا البحث يركز على توزيع t متعدد المتغيرات وعلاقته مع التوزيع الطبيعي، بينما وضح القسم الثالث نموذج فضاء الحالة، في حين القسم الرابع اختص بإيجاد مرشحات للمعلمات وكذلك إيجاد توزيعات الأخطاء التنبؤية. بعد ذلك قدم القسم الخامس عامل بيز، كما وضح القسم السادس لوحة EWMA، والقسم السابع فقد خصصت لتطبيق الأقسام السابقة له، فيما تضمن القسم الثامن أهم الاستنتاجات والتوصيات، وكان القسمين الأخيرين للمصادر والملاحق.

2- توزيع t متعدد المتغيرات

في الكثير من الأحيان يعتبر توزيع t متعدد المتغيرات بديلاً جيداً عن التوزيع الطبيعي متعدد المتغيرات وخاصة إذا احتوت البيانات على بعض القيم الشاذة أو لكون المشاهدات تمتلك توزيعات ذات ذيول ثقيلة (Heavy-tailed) بمعنى أن لها نهايات تؤول إلى الصفر بسرعة بالمقارنة مع التوزيع الطبيعي متعدد المتغيرات، على سبيل المثال فيما يتعلق الأمر بالبيانات الاقتصادية والتجارية والصناعية وبالأخص بدراسة الموارد المالية وأسعار الأسهم المالية، فإن النمذجة باستخدام التوزيع الطبيعي ستكون على حساب الدقة في التحليل الإحصائي [10].

يمكن التعبير عن توزيع t متعدد المتغيرات بكونه توزيعاً خليطاً من التوزيع الطبيعي متعدد المتغيرات وتوزيع معكوس كاما (Inverse Gamma) وكالآتي [11]:

$$p(\underline{X}) = \int_0^{\infty} p(\underline{X}|w)p(w)dw \quad \dots (1)$$

حيث أن w متغير عشوائي يتبع توزيع معكوس كاما (IG) بالمعلمتين $(\frac{v}{2}، \frac{v}{2})$ وكما يلي:

$$p(w) = \frac{\left(\frac{v}{2}\right)^{\frac{v}{2}}}{\Gamma\left(\frac{v}{2}\right)} w^{-\frac{v}{2}-1} e^{-\frac{v}{2w}} \quad \dots (2)$$

كما أن المتجه العشوائي $\underline{X}$ المشروط بالمتغير العشوائي w يتبع توزيع طبيعي متعدد المتغيرات والذي يوصف بالشكل التالي:

$$\underline{X}|w \sim N_p(\underline{\theta}, w\Sigma)$$

وبالاعتماد على المعادلة (1)، فإنه بالإمكان إيجاد التوزيع الاحتمالي للمتجه $\underline{X}$ غير المشروط بـ w وكما يلي:

$$p(\underline{X}) = \frac{\left(\frac{v}{2}\right)^{\frac{v}{2}}}{(2\pi)^{\frac{p}{2}}|\Sigma|^{\frac{1}{2}}\Gamma\left(\frac{v}{2}\right)} \int_0^{\infty} w^{-\frac{v}{2}-1} e^{-\frac{v}{2w}} w^{-\frac{p}{2}} e^{-\frac{1}{2w}\left((\underline{X}-\underline{\theta})'\Sigma^{-1}(\underline{X}-\underline{\theta})\right)} dw$$

بعد القيام بعدة عمليات رياضية ينتج:

$$p(\underline{X}) = \frac{\Gamma\left(\frac{v+p}{2}\right)|\Sigma|^{-\frac{1}{2}}}{\Gamma\left(\frac{v}{2}\right)(v\pi)^{\frac{p}{2}}} \left(\frac{(\underline{X}-\underline{\theta})'\Sigma^{-1}(\underline{X}-\underline{\theta})}{v} + 1\right)^{-\frac{v+p}{2}}$$

حيث أن:

$\underline{X}$: متجه عشوائي ذو سعة (p×1) وأن $-\infty < x < \infty$.

$\underline{\theta}$: متجه المتوسط ذو سعة (p×1) وأن $-\infty < \theta < \infty$.

Σ: مصفوفة أكيدة الإيجابية (Positive Definite)، ذات سعة (p × p).

p : تمثل عدد المتغيرات.

v : تمثل درجة الحرية.

ويعبر عن هذا التوزيع وصفياً بالشكل التالي:

$$\underline{X} \sim t_p(\underline{\theta}, \Sigma, v) \quad \dots (3)$$

يمتاز توزيع t متعدد المتغيرات إضافة إلى كونه من التوزيعات المتماثلة (Symmetric Distribution) بأنه يمتلك الخاصية الخطية (Linearity Property)، بمعنى أنه إذا كان $\underline{X}$ موصوفاً كما في العلاقة (3) عندئذ فإن التركيبة الخطية $(C\underline{X} + \underline{d})$ ستكون موصوفة بالشكل التالي [12]:

$$(C\underline{X} + \underline{d}) \sim t_m(C\underline{\theta} + \underline{d}, C\Sigma C', v)$$

حيث أن:

C: مصفوفة ثابتة ذات سعة (m×p) وبرتبة (Rank) مساوية لـ m، إذ أن m عدد صحيح موجب أقل أو يساوي p.

$\underline{d}$: متجه ثابت ذو سعة (m× 1).

3- نموذج فضاء الحالة

يقصد بفضاء الحالة (State space) هو كمية الحالات الممكنة التي يحتويها نظام حركي، وتوصف الحالة (State) بواسطة مجموعة المتغيرات للنظام عند نقطة زمنية معينة، أو هو نظام حركي يشير إلى تقليل مجموعة المتغيرات المعروفة بمتغيرات الحالة (State variables) ويصور تماماً النظام ويستجيب لكل مجموعة من المدخلات (Input)، أما حالة النظام (System state) فتعني أقل مجموعة معلومات عن الماضي والحاضر يمكن من خلالها وصف السلوك المستقبلي لنظام ما وصفاً كاملاً عند معرفة الحالة في الحاضر ومدخلات المستقبل، حيث يتم دمج كل المعلومات الماضية في متغيرات الحالة الحالية، وهذا يجعلها تخضع للخاصية الماركوفية (Markovian property)، أي أن التوزيع الاحتمالي الشرطي لمتغير الحالة عند الزمن t معطى لجميع الحالات الماضية يعتمد فقط على حالته عند الزمن t-1. أما نموذج فضاء الحالة (State Space Model) فهو عبارة عن نموذج ديناميكي يربط بين المتغير المشاهد ومتغير الحالة.

إن أول من قام باستخدام نماذج فضاء الحالة هو العالم الأمريكي كالمن (Kalman) في عام 1961 وقد قام بتكوين هذه النماذج بالاعتماد على المعادلات الفرقية (Difference equations) والمعادلات التفاضلية (Differential equations) واستخدمها وقتئذ في نظم السيطرة (Control system) الكهربائية [13]. أما في المجال الإحصائي فكان أول من استخدم هذه النماذج هو العالم الياباني اكايكي (Akaike) عام 1974، حيث طبق لأول مرة مفهوم نماذج فضاء الحالة لتحليل نموذج الانحدار الذاتي– الأوساط المتحركة (ARMA) [14].

يتكون نموذج فضاء الحالة من معادلتين المعادلة، الأولى تصف متجه المخرجات تسمى معادلة القياس (Measurement equation) أو معادلة المشاهدة (Observed equation)، والثانية تصف متجه المدخلات تسمى معادلة الحالة (State equation) أو معادلة النظام (System equation). تحتوي كل معادلة من هاتين المعادلتين على أخطاء، حيث يسمى الخطأ في معادلة المشاهدة بخطأ المشاهدة، بينما يسمى الخطأ في معادلة الحالة بخطأ الحالة أو خطأ النظام، وتكون هذه الأخطاء غير مترابطة فيما بينها ومستقلة أيضاً عن متجه الحالة.

في هذا البحث سنقوم بتكوين لوحات سيطرة استناداً إلى نموذج فضاء الحالة الخطي ذي الزمن المتقطع عندما تكون الأخطاء العشوائية لها توزيع t متعدد المتغيرات، وفيما يلي النموذج المستخدم:

$$\underline{X}_t = F\underline{\theta}_t + \underline{\varepsilon}_t \quad \dots (4)$$

$$\text{for } t=1,2,3,\dots,n \quad \underline{\theta}_t = G\underline{\theta}_{t-1} + \underline{\omega}_t \quad \dots (5)$$

إذ أن:

$$\underline{\varepsilon}_t \sim t_p(\underline{0}, \Sigma_{\underline{\varepsilon}}, v)$$

$$\underline{\omega}_t \sim t_p\ (\underline{0}, \Sigma_{\underline{\omega}}, v)$$

حيث أن:

$\underline{X}_t$: يمثل متجه متغيرات المشاهدة عند الزمن t، ذو سعة (p×1)، ويمثل p عدد المتغيرات المشاهدة.

$\underline{\theta}_t$: يمثل متجه معلمات النموذج، ذو سعة (p×1).

F: تمثل مصفوفة الانحدار، وتكون عبارة عن مصفوفة ثابتة ذات سعة (p×p).

G: تمثل مصفوفة الحالة، وتكون عبارة عن مصفوفة ثابتة ذات سعة (p×p).

$\underline{\varepsilon}_t$: يمثل متجه خطأ المشاهدة، ذو سعة (p×1).

$\underline{\omega}_t$: يمثل متجه خطأ الحالة، ذو سعة (p×1).

$\Sigma_{\underline{\varepsilon}}$: تمثل مصفوفة التباين لخطأ المشاهدة وتكون أكيدة الإيجابية، ذات سعة (p×p).

$\Sigma_{\underline{\omega}}$: تمثل مصفوفة التباين لخطأ الحالة وتكون أكيدة الإيجابية، ذات سعة (p×p).

v: تمثل درجة الحرية للتوزيع، وتكون قيمة صحيحة موجبة (Scale).

4- ترشيح المعلمات

يعرف الترشيح بأنه التقدير عند الزمن t، ومن أهم المرشحات التي تستخدم مع نموذج فضاء الحالة هو مرشح كالمن (Kalman filter) وهو عبارة عن مرشح كّون خصيصاً من قبل العالم رودولف كالمن في عام 1960 للسيطرة على عدم الاستقرارية التي تعاني منها الكثير من الإشارات الواقعية وتم تطويره على يد ريتشارد بوسي (Richard Bucy) عام 1961، ولذلك سمي في بعض الأحيان بمرشح كالمن – بوسي. يعتبر مرشح كالمن من أهم المرشحات الموجودة في هذا الجانب لكونه يمتاز بالخاصية التعاقبية، بمعنى يمكننا من خلال هذا المرشح الاستفادة من المعلومات والبيانات عند تغيير الزمن، كما أنه لا ينبغي أن تكون جميع البيانات متوفرة، وبذلك فأنه يستخدم في نظم السيطرة والاتصالات وغيرهما [15].

في هذا القسم سيتم ترشيح المعلمات عندما يكون متجه المعلمات غير معلوم ومصفوفة التباين معلومة وكذلك مصفوفة أخطاء النظام $\Sigma_{\underline{\omega}}$ ومعلمة الشكل v معلومتين، وذلك تمهيداً لإيجاد التوزيع التنبؤي الذي يعتبر الأساس في تكوين عوامل بيز.

التوزيع الأولي لمتجه المعلمات $\underline{\theta}$ المشروط بالمتغير العشوائي w عند الزمن صفر يوصف بالشكل الآتي [15]:

$$\underline{\theta}_0|w \sim N_p(\underline{\mu}_{0|0}, w\Sigma_{0|0})$$

حيث أن كل من $\underline{\mu}_{0|0}$ و $\Sigma_{0|0}$ هما على الترتيب متجه المتوسط الابتدائي ومصفوفة التباين الابتدائية وتكونان معلومتين, وأن w يتبع توزيع معكوس كاما (IG). باستخدام الاستقراء الرياضي ونظرية بيز نجد أن التوزيع اللاحق لمتجه المعلمات $\underline{\theta}$ عند الزمن t-1 يعبر عنه وصفياً وفق الصيغة التالية:

$$(\underline{\theta}_{t-1}|w, X^{t-1}) \sim N_p\left(\underline{\mu}_{t-1|t-1}, w\Sigma_{t-1|t-1}\right) \quad \dots (6)$$

إذ أن: $X^{t-1} = \{\underline{X}_1, \underline{X}_2, \dots, \underline{X}_{t-1}\}$

للحصول على التوزيع اللاحق المشروط بـ w لمتجه المعلمات $\underline{\theta}$ عند الزمن t يتم استخدام نظرية بيز وكما يلي:

$$p(\underline{\theta}_t|w, X^t) \propto p(\underline{X}_t|\underline{\theta}_t, w)p(\underline{\theta}_t|w, X^{t-1}) \quad \dots (7)$$

حيث أن $p(\underline{X}_t|\underline{\theta}_t, w)$ تمثل دالة الترجيح (Likelihood function) التي يتم الحصول عليها من المعادلة (4) باستخدام الخاصية الخطية في التوزيع الطبيعي فينتج: $p(\underline{X}_t|\underline{\theta}_t, w) \propto$

$$e^{-\frac{1}{2w}(\underline{X}_t - F\underline{\theta}_t)'\Sigma_{\underline{\varepsilon}}^{-1}(\underline{X}_t - F\underline{\theta}_t)} \quad \dots (8)$$

ويمكن التعبير عن دالة الترجيح المشروطة بـ w عند الزمن t وصفياً بالشكل التالي: $(\underline{X}_t|\underline{\theta}_t, w) \sim N_p(F\underline{\theta}_t, w\Sigma_{\underline{\varepsilon}})$

وأن $p(\underline{\theta}_t|w, X^{t-1})$ تمثل الدالة الاحتمالية للتوزيع السابق المشروط بـ w لمتجه المعلمات $\underline{\theta}$ عند الزمن t والتي يمكن الحصول عليها من المعادلة (5) ومن التوزيع اللاحق عند الزمن t-1 المعرّف في العلاقة (6) فينتج:

$$p(\underline{\theta}_t|w, X^{t-1}) \propto e^{-\frac{1}{2w}(\underline{\theta}_t - \underline{\mu}_{t|t-1})'\Sigma_{t|t-1}^{-1}(\underline{\theta}_t - \underline{\mu}_{t|t-1})} \quad \dots (9)$$

ويمكن التعبير عن الدالة الاحتمالية للتوزيع السابق المشروط بـ w لمتجه المعلمات $\underline{\theta}$ عند الزمن t وصفياً بالشكل التالي:

$$(\underline{\theta}_t|w, X^{t-1}) \sim N_p(\underline{\mu}_{t|t-1}, w\Sigma_{t|t-1})$$

حيث أن: $\underline{\mu}_{t|t-1} = G\underline{\mu}_{t-1|t-1} \quad \dots (10)$

$$\Sigma_{t|t-1} = wG\Sigma_{t-1|t-1}G' + w\Sigma_{\underline{\omega}} \quad \dots (11)$$

الآن ولإيجاد التوزيع الاحتمالي اللاحق المشروط بـ w لمتجه المعلمات $\underline{\theta}$ عند الزمن t نقوم بتعويض كل من المعادلة (8) والمعادلة (9) في المعادلة (7) وكما يلي:

$$p(\underline{\theta}_t|w, X^t) \propto e^{-\frac{1}{2w}(\underline{X}_t - F\underline{\theta}_t)'\Sigma_{\underline{\varepsilon}}^{-1}(\underline{X}_t - F\underline{\theta}_t)} e^{-\frac{1}{2w}(\underline{\theta}_t - \underline{\mu}_{t|t-1})'\Sigma_{t|t-1}^{-1}(\underline{\theta}_t - \underline{\mu}_{t|t-1})}$$

بالقيام ببعض العمليات الرياضية ينتج: $p(\underline{\theta}_t|w, X^t) \propto$

$$e^{-\frac{1}{2w}\left\{(\underline{\theta}_t - F'(FF')^{-1}X_t)'F'\Sigma_{\underline{\varepsilon}}^{-1}F(\underline{\theta}_t - F'(FF')^{-1}X_t) + (\underline{\theta}_t - \underline{\mu}_{t|t-1})'\Sigma_{t|t-1}^{-1}(\underline{\theta}_t - \underline{\mu}_{t|t-1})\right\}}$$

بعد ذلك نقوم باستخدام الصيغة التربيعية التالية [17]:

$$\left(\frac{\underline{Z} -}{\underline{a}}\right)' A(\underline{Z} - \underline{a}) + (\underline{Z} - \underline{b})'B(\underline{Z} - \underline{b}) = (\underline{Z} - \underline{c})'(A + B)(\underline{Z} - \underline{c}) + (\underline{a} - \underline{b})'A(A + B)^{-1}B(\underline{a} - \underline{b}) \quad \dots (12)$$

حيث أن (A و B) مصفوفتان أكيدتا الإيجابية ذواتا سعة $(p \times 1)$ و $(p \times p)$ وأن: $\underline{c} =$ متجهات ذوات سعة (Z و a و b و c)

$$(A + B)^{-1}(A\underline{a} + B\underline{b})$$

$$\underline{Z} = \underline{\theta}_0 \ ; \ \underline{a} = F'(FF')^{-1}X_t \ ; \ A = F'\Sigma_{\underline{\varepsilon}}^{-1}F \ ; \ \underline{b} = \underline{\mu}_{t|t-1} \ ; \ B = \Sigma_{t|t-1}^{-1}$$

وبعد التشبيه والتعويض والقيام بعدة عمليات رياضية تنتج لنا الصيغة التالية:

$$p(\underline{\theta}_t|w, X^t) = \frac{1}{(2\pi)^{\frac{p}{2}}|w\Sigma_{t|t}|^{\frac{1}{2}}} e^{-\frac{1}{2w}\left\{(\underline{\theta}_t - \underline{\mu}_{t|t})'\Sigma_{t|t}^{-1}(\underline{\theta}_t - \underline{\mu}_{t|t})\right\}} \quad \dots (13)$$

وبذلك يمكن التعبير عن دالة كثافة احتمال متجه المعلمات $\underline{\theta}_t$ المشروطة بـ w وصفياً بالشكل التالي:

$$(\underline{\theta}_t|w, X^t) \sim N_p(\underline{\mu}_{t|t}, w\Sigma_{t|t})$$

أما التوزيع اللاحق الحدي غير المشروط لمتجه المعلمات $\underline{\theta}_t$ فيكون وفق الصيغة التالية:

$$p(\underline{\theta}_t|X^t) = \frac{\Gamma\left(\frac{v+p}{2}\right)|\Sigma_{t|t}|^{-\frac{1}{2}}}{\Gamma\left(\frac{v}{2}\right)(v\pi)^{\frac{p}{2}}}\left[\frac{(\underline{\theta}_t-\underline{\mu}_{t|t})'\Sigma_{t|t}^{-1}(\underline{\theta}_t-\underline{\mu}_{t|t})}{v}+1\right]^{-\frac{v+p}{2}} \quad \dots(14)$$

والمعادلة (14) أعلاه هي دالة كثافة احتمال توزيع t متعدد المتغيرات ويمكن التعبير عنه وصفياً بالشكل التالي:

$$(\underline{\theta}_t|X^t) \sim t_p(\underline{\mu}_{t|t}, \Sigma_{t|t}, v)$$

حيث أن:

$$\underline{\mu}_{t|t} = \underline{\mu}_{t|t-1} + K_t\left(\underline{X}_t - F\underline{\mu}_{t|t-1}\right) \quad \dots(15)$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - K_t F\Sigma_{t|t-1} \quad \dots(16)$$

المعادلتان (15) و(16) على التوالي أعلاه تسميان بمعادلتي التحديث فالأولى تقوم بتحديث متجه المتوسط والثانية تقوم بتحديث مصفوفة التباين، و K_t هو عبارة عن معامل كالمن (Kalman factor) ويكون في بحثنا على شكل مصفوفة مربعة عدد صفوفها وأعمدتها تساوي عدد المتغيرات p ويكون معرفاً بالصيغة التالية:

$$K_t = \Sigma_{t|t-1}F'\left(F\Sigma_{t|t-1}F' + \Sigma_\varepsilon\right)^{-1} \quad \dots(17)$$

والمقدار $\left(\underline{X}_t - F\underline{\mu}_{t|t-1}\right)$ الموجود في ذيل المعادلة (15) يسمى الخطأ التنبؤي (Prediction error) أو البواقي (Residual) وهو يمثل الفرق بين المشاهدة الأصلية $\underline{X}_t$ ومعدل التوزيع التنبؤي لتلك المشاهدة حيث أن التوزيع التنبؤي المشروط بـ w للمشاهدة $\underline{X}_t$ بوجود المعلومات X^{t-1} يتم الحصول عليه كالآتي:

$$p(\underline{X}_t|w, X^{t-1}) = \int_{\underline{\theta}_t} p(\underline{X}_t|\underline{\theta}_t, w)\, p(\underline{\theta}_t|w, X^{t-1})d\underline{\theta}_t$$

حيث أن $p(\underline{X}_t|\underline{\theta}_t, w)$ تمثل دالة الترجيح عند النقطة t المعرفة في العلاقة (8) وأن $p(\underline{\theta}_t|w, X^{t-1})$ تمثل التوزيع السابق لمتجه المعلمات عند الزمن t والذي تم الحصول عليه في المعادلة (9)، ومن هذا نستنتج أن:

$$p(\underline{X}_t|w, X^{t-1}) = \frac{1}{(2\pi)^{\frac{p}{2}}|w\Sigma_\varepsilon|^{\frac{1}{2}}}\frac{1}{(2\pi)^{\frac{p}{2}}|w\Sigma_{t|t-1}|^{\frac{1}{2}}} \times \int_{\underline{\theta}_t} e^{-\frac{1}{2w}\left\{(\underline{X}_t-F\underline{\theta}_t)'\Sigma_\varepsilon^{-1}(\underline{X}_t-F\underline{\theta}_t)\right\}} e^{-\frac{1}{2w}\left\{(\underline{\theta}_t-\underline{\mu}_{t|t-1})'\Sigma_{t|t-1}^{-1}(\underline{\theta}_t-\underline{\mu}_{t|t-1})\right\}} d\underline{\theta}_t$$

بعد استخدام المتطابقة (12) والقيام بعملية التكامل وعمليات رياضية أخرى ينتج ما يلي:

$$p(\underline{X}_t|w, X^{t-1}) = \frac{1}{(2\pi)^{\frac{p}{2}}\left|w\left[F\Sigma_{t|t-1}F'+\Sigma_\varepsilon\right]\right|^{\frac{1}{2}}} e^{-\frac{1}{2w}\left\{(\underline{X}_t-F\underline{\mu}_{t|t-1})'\left[F\Sigma_{t|t-1}F'+\Sigma_\varepsilon\right]^{-1}(\underline{X}_t-F\underline{\mu}_{t|t-1})\right\}}$$

وبذلك يمكننا كتابة التوزيع التنبؤي المشروط بـ w وصفياً كالآتي:

$$(\underline{X}_t|w, X^{t-1}) \sim N_p(F\underline{\mu}_{t|t-1}, w\left[(F\Sigma_{t|t-1}F' + \Sigma_\varepsilon)\right])$$

ويمكن كتابة التوزيع التنبؤي الحدي غير المشروط وصفياً بالصيغة الآتية:

$$(\underline{X}_t|X^{t-1}) \sim t_p\left(F\underline{\mu}_{t|t-1}, \left[(F\Sigma_{t|t-1}F' + \Sigma_\varepsilon)\right], v\right) \quad \dots(18)$$

استناداً إلى العلاقة (18) أعلاه وباستخدام الخاصية الخطية في توزيع t متعدد المتغيرات المذكورة في القسم الثاني من هذا البحث يتم الحصول على التوزيع الاحتمالي الحدي غير المشروط لخطأ التنبؤ فيعبر عنه بالصيغة الآتية:

$$\left(\left[\underline{e}_t = \underline{X}_t - F\underline{\mu}_{t|t-1}\right]|X^{t-1}\right) \sim t_p\left(\underline{0}, \left[(F\Sigma_{t|t-1}F' + \Sigma_\varepsilon)\right], v\right)$$

وبذلك فإن الدالة الاحتمالية الحدية لخطأ التنبؤ تكون بالشكل التالي:

$$p(\underline{e}_t|X^{t-1}) = \frac{\Gamma\left(\frac{v+p}{2}\right)|(F\Sigma_{t|t-1}F'+\Sigma_\varepsilon)|^{-\frac{1}{2}}}{\Gamma\left(\frac{v}{2}\right)(v\pi)^{\frac{p}{2}}}\left[\frac{(\underline{e}_t)'(F\Sigma_{t|t-1}F'+\Sigma_\varepsilon)^{-1}(\underline{e}_t)}{v}+1\right]^{-\frac{v+p}{2}} \quad \dots(19)$$

حيث أن:

$\underline{e}_t$: يمثل الخطأ التنبؤي عند الزمن t، $-\infty < e < \infty$.

ولمعرفة مدى ملائمة التوزيع التنبؤي للتنبؤ بالبيانات يتم اختبار الخطأ التنبؤي عن طريق معيار يطلق عليه متوسط مربعات الخطأ التنبؤي القياسي (Mean of squared standard one-step forecast error) والمعطى في الصيغة التالية [9]:

$$MSSE = \frac{1}{n}\sum_{t=1}^{n}\left[(e_{1t}^*)^2 \quad (e_{1t}^*)^2 \quad \cdots \quad (e_{pt}^*)^2\right]' \quad \dots(20)$$

حيث أن e_t^* تمثل الخطأ التنبؤي القياسي وتكون صيغته بالشكل التالي: $e_t^* = \left[\frac{v}{v-2}\left(F\Sigma_{t|t-1}F' + \Sigma_\varepsilon\right)\right]^{-\frac{1}{2}} e_t$

يقال للنموذج بأنه جيد إذا اقترب متجه (MSSE) من المتجه الواحدي (Ones vector).

5- عامل بيز (Bayes' factors)

يعتبر عامل بيز البديل البيزي لاختبار الفرضيات الإحصائية (Statistical Hypotheses)، وتتلخص فكرة هذا العامل في أنه يمثل الأرجحية بين بديلين معينين، بمعنى أنه يمثل النسبة بين شيئين لمعرفة أيهما أصلح أو أفضل ولذلك يسمى في بعض الأحيان بأسلوب التفضيل. يمثل عامل بيز النسبة بين قسمة الاحتمال اللاحق لأحد البديلين (A_1) على الاحتمال السابق له وبين حاصل قسمة الاحتمال اللاحق للبديل الآخر (A_2) على الاحتمال السابق له، وبعد استخدام نظرية بيز تنتج الصيغة العامة لعامل بيز وكما يلي [18]:

$$BF = \frac{p(X^n|A_1)}{p(X^n|A_2)}$$

حيث أن:

$p(X^n|A_1)$: تمثل التوزيع الحدي لدالة الإمكان لنموذج البديل الأول.

$p(X^n|A_2)$: تمثل التوزيع الحدي لدالة الإمكان لنموذج البديل الثاني.

X^n: تمثل جميع البيانات. يستخدم عامل بيز في مجالات عديدة في الاستدلال البيزي (Bayesian Inference) وخاصة في مجالات

المقارنة بين النماذج، ومن هذه الاستعمالات استعماله في تكوين لوحات السيطرة، حيث قام [16] بتعريف عامل بيز بالشكل التالي:

$$BF(t) = \frac{p_1(\underline{e}_t|X^{t-1})}{p_2(\underline{e}_t|X^{t-1})}$$

حيث أن:

$p_1(\underline{e}_t|X^{t-1})$: تمثل دالة الخطأ التنبؤي عند الزمن t للبديل الأول.

$p_2(\underline{e}_t|X^{t-1})$: تمثل دالة الخطأ التنبؤي عند الزمن t للبديل الثاني.

في هذا البحث تم تكوين عوامل بيز على أساس كون البسط يمثل الخطأ التنبؤي للنموذج في الحالة المدروسة والمقام يمثل خطأ الهدف، حيث أن توزيع الهدف (Target distribution) عبارة عن نموذج يتم فرضه لمقارنته مع خطأ التنبؤ وسيتم فرضه بالشكل التالي:

$$\underline{X}_t = \underline{\theta} + \underline{\tau}_t \quad ; \quad \underline{\tau}_t \sim t_p(\underline{0}, \Sigma_{\underline{\tau}}, v)$$

حيث أن:

$\underline{\theta}$: يمثل متجه معدل الهدف (Target mean) ذو سعة (p×1).

عادةً يتم تقديره حسب الصيغة التالية:

$$\underline{\theta} = \frac{1}{n}\sum_{t=1}^{n}[X_{1t} \quad X_{2t} \quad \cdots \quad X_{pt}]' \quad \dots (21)$$

$\underline{\tau}_t$: متجه خطأ الهدف ذو سعة (p×1).

$\Sigma_{\underline{\tau}}$: تمثل مصفوفة تباين الهدف (Target variance matrix) ذو سعة (p×p). يتم الحصول عليها وفق القانون التالي: $\Sigma_{\underline{\tau}} =$

$$\frac{1}{n-1}\sum_{t=1}^{n}(\underline{X}_t - \underline{\theta})(\underline{X}_t - \underline{\theta})' \quad \dots (22)$$

ويمكن كتابة دالة خطأ الهدف (Target error) بالشكل التالي:

$$p(\underline{\tau}_t = \underline{X}_t - \underline{\theta}) = \frac{\Gamma\left(\frac{v+p}{2}\right)|\Sigma_{\underline{\tau}}|^{-\frac{1}{2}}}{\Gamma\left(\frac{v}{2}\right)(v\pi)^{\frac{p}{2}}}\left[\frac{(\underline{\tau}_t)'\Sigma_{\underline{\tau}}^{-1}(\underline{\tau}_t)}{v} + 1\right]^{-\frac{v+p}{2}} \quad \dots (23)$$

وبذلك فإن صيغة عامل بيز النهائية ستكون بالشكل التالي:

$$BF(t) = \frac{p([\underline{e}_t = \underline{X}_t - E(\underline{X}_t|X^{t-1})]|X^{t-1})}{p(\underline{\tau}_t = \underline{X}_t - \underline{\theta})} \quad \dots (24)$$

حيث أن:

$p([\underline{e}_t = \underline{X}_t - E(\underline{X}_t|X^{t-1})]|X^{t-1})$: يمثل الخطأ التنبؤي للنموذج.

$p(\underline{\tau}_t = \underline{X}_t - \underline{\theta})$: يمثل خطأ الهدف.

$E(\underline{X}_t|X^{t-1})$: يمثل المعدل التنبؤي.

بتعويض الدالة الاحتمالية لخطأ التنبؤ المعرفة في المعادلة (19) والدالة الاحتمالية لخطأ الهدف المعرفة في المعادلة (23) في صيغة عامل بيز المعرفة في المعادلة (24) نحصل على صيغة عامل بيز النهائية عندما تكون مصفوفة التباين معلومة ومتجه المعلمات مجهول بالشكل التالي:

$$BF(t) = \frac{|\Sigma|^{\frac{1}{2}}}{|F\Sigma_{t|t-1}F' + \Sigma_{\underline{\varepsilon}}|^{\frac{1}{2}}}\left[\frac{(\underline{X}_t - \underline{\theta})'\Sigma_{\underline{\tau}}^{-1}(\underline{X}_t - \underline{\theta}) + v}{(\underline{X}_t - F\underline{\mu}_{t|t-1})'(F\Sigma_{t|t-1}F' + \Sigma_{\underline{\varepsilon}})^{-1}(\underline{X}_t - F\underline{\mu}_{t|t-1}) + v}\right]^{\frac{v+p}{2}} \quad (25)$$

مع ملاحظة أن قيمة عامل بيز تكون قيمة موجبة لكونها ناتجة من قسمة دالتين احتماليتين, وأنها تكون عبارة عن سلسلة مترابطة وذات توزيع غير طبيعي.

6- تكوين لوحة (EWMA) المحورة

تعتبر لوحة الأوساط المتحركة الموزونة أسياً (Exponentially weighted moving average) من أولى اللوحات التي تم تكوينها للسيطرة على البيانات المرتبطة ذاتياً، وتتلخص فكرة هذه اللوحة برسم المتوسط المتحرك الموزون أسياً ولكافة المشاهدات السابقة والحالية للعملية الإنتاجية وهذا المتوسط يتم الحصول عليه من المعادلة التالية [19]:

$$z_t = \lambda y_t + (1-\lambda)z_{t-1} \quad \text{for } t = 1,2,\dots,n \quad \dots (26)$$

حيث أن:

λ: تمثل معلمة التمهيد (Smoothing parameter) وتكون عبارة عن ثابت اختياري تكون قيمته محصورةً بين الصفر والواحد.

y_t: تمثل قيم مفردات العينة.

z_t: قيمة المتوسط المتحرك الموزون أسياً. و z_0 عادة ما يتم التعويض عنها بقيمة متوسط العينة وهي تمثل أيضاً خط الهدف (Target line) للّوحة في الكثير من الأحيان.

أما حدود اللوحة فتعطى بالصيغة التالية:

$$UCL(t) = E(z_t) + L\sqrt{Var(z_t)} \quad \dots (27)$$

$$TL(t) = E(z_t) \quad \dots (28)$$

$$LCL(t) = E(z_t) - L\sqrt{Var(z_t)} \quad \dots (29)$$

حيث أن:

L: يمثل عرض حد السيطرة وتكون قيمته عادةً تساوي (3).

$UCL(t)$: يمثل الحد الأعلى للّوحة عند الزمن t.

$TL(t)$: يمثل خط الهدف عند الزمن t.

$LCL(t)$: يمثل الحد الأدنى للّوحة عند الزمن t.

وأن $E(z_t)$ و $Var(z_t)$ هما التوقع والتباين على الترتيب.

يعتبر عامل بيز الذي تم الحصول عليه في الجزء السابق من هذا البحث الركيزة الأساسية في تكوين اللوحات، حيث يتم رسم هذه العوامل على اللوحة ذات المتغير الواحد، وقد تم اختيار لوحة الأوساط المتحركة الموزونة أسياً (EWMA) ومن ثم إجراء بعض التعديلات عليها لغرض جعلها ملائمة لعوامل بيز، لكون هذه العوامل تعاني من مشكلتين الأولى هي عدم الطبيعية (Non-normality) والمشكلة الثانية هي وجود الارتباط الذاتي كما مذكور أعلاه، وللتغلب على المشكلة الأولى يتم استخدام معلمة تمهيد صغيرة جداً حيث أن ذلك يساعد على جعل اللوحة أقرب إلى اللامعلمية حسب ما هو مذكور من قبل بعض الباحثين ومنهم [19]، أما المشكلة الثانية فيتم التغلب عليها عن طريق إجراء التعديل على اللوحة بما يناسب البيانات المرتبطة ذاتياً. بالاستفادة من إيجابية قيم عوامل بيز تم إيجاد اللوغاريتم الطبيعي له وذلك لكون اللوغاريتمات أكثر استقراراً من القيم الأصلية لعوامل بيز وقد تم فرض أن السلسلة الناتجة من نوع الانحدار الذاتي الأوساط المتحركة (ARMA(1,1)) وكما يلي:

$$(LBF)_t = c + \phi(LBF)_{t-1} - \vartheta\eta_{t-1} + \eta_t \quad \dots (30)$$

حيث أن:

$(LBF)_t$: تمثل قيمة لوغاريتم عامل بيز.

c: يمثل ثابت السلسلة.

ϕ: يمثل معامل الانحدار الذاتي.

ϑ: يمثل معامل الوسط المتحرك.

η_t: يمثل التشويش أو الخطأ ويكون ذو توزيع طبيعي له وسط صفري وتباين σ^2 ويعبر عنه وصفياً بالتالي: $\eta_t \overset{IID}{\sim} N(0,\sigma^2)$

حيث يتم إيجاد التوقع والتباين للسلسلة $(LBF)_t$ المعرفة في المعادلة (30) ومن ثم إيجاد التوقع والتباين للوحة (EWMA) بالشكل التالي: ... (31) $E(z_t) = \frac{c}{(1-\phi)}$ for $\phi \neq 1$

$$var(z_t) = \frac{\lambda\sigma^2\{((1-\phi(1-\lambda))(1+\vartheta^2-2\phi\vartheta))+2((\phi-\vartheta)(1-\phi\vartheta)(1-\lambda))\}}{\{2-\lambda\}(1-\phi^2)\{1-\phi(1-\lambda)\}} \quad \dots (32)$$

حيث أن قيمة التباين قد تم إيجاده عندما الغاية تقترب من الما–لانهاية وذلك لأن بعض الباحثين أمثال [20] يعد ذلك أفضل من الحالة التي يؤخذ الزمن بنظر الاعتبار.

وبتعويض المعادلتين (31) و(32) في المعادلات (27) و(28) و(29) نحصل على حدود اللوحة الجديدة، وكما يلي: $UCL = \frac{c}{(1-\phi)} + L\sqrt{\frac{\lambda\sigma^2\{((1-\phi(1-\lambda))(1+\vartheta^2-2\phi\vartheta))+2((\phi-\vartheta)(1-\phi\vartheta)(1-\lambda))\}}{\{2-\lambda\}(1-\phi^2)\{1-\phi(1-\lambda)\}}}$ ... (33)

$TL = \frac{c}{(1-\phi)}$ ... (34)

$LCL = \frac{c}{(1-\phi)} - L\sqrt{\frac{\lambda\sigma^2\{((1-\phi(1-\lambda))(1+\vartheta^2-2\phi\vartheta))+2((\phi-\vartheta)(1-\phi\vartheta)(1-\lambda))\}}{\{2-\lambda\}(1-\phi^2)\{1-\phi(1-\lambda)\}}}$ ... (35)

حيث أن قيمة c تقدر بالشكل التالي: $c = (1-\lambda)\overline{(LBF)}$

بينما قيمة σ^2 تقدر بالشكل التالي: $\sigma^2 = \frac{var(LBF)(1-\phi^2)}{(1+\phi^2-2\phi\vartheta)}$

إذ أن $\overline{(LBF)}$ و $var(LBF)$ هما على التوالي يمثلان المعدل والتباين للوغاريتم عامل بيز.

7- الجانب التطبيقي

لغرض تطبيق ما تم الحصول عليه في الأقسام السابقة تم توليد مئة متجه (N=100) بشكل عشوائي باستخدام البرنامج (1) الموجود في الملحق الذي تم كتابته بواسطة برنامج (MATLAB 7.10). لقد تم عملية التوليد لبيانات تتبع توزيع t متعدد المتغيرات لخمسة متغيرات ودرجة حرية (v=6) ومتجه متوسط ومصفوفة تباين معرفتين بالشكل التالي:

$$\underline{\mu} = [10 \quad 15 \quad 20 \quad 25 \quad 30]'$$

$$\Sigma = \begin{bmatrix} 0.6740 & 0.0742 & -0.1873 & -0.0916 & -0.0242 \\ & 0.6300 & -0.0073 & 0.0051 & -0.0997 \\ & & 0.5521 & 0.0683 & -0.0084 \\ & & & 0.6492 & -0.0553 \\ & & & & 0.7713 \end{bmatrix}$$

بفرض أن مصفوفة الانحدار F ومصفوفة الحالة G معلومتين وتم اعتبار أن كل واحدة منهما تمثل مصفوفة الوحدة (Identity matrix) وأن مصفوفة تباين خطأ الحالة معرفة بالشكل التالي:

$$\Sigma_{\underline{\omega}} = \begin{bmatrix} 0.6958 & -0.1605 & -0.2126 & -0.1113 & -0.1834 \\ & 0.3308 & 0.0520 & -0.1262 & 0.0222 \\ & & 0.6132 & -0.0903 & -0.0613 \\ & & & 0.5062 & 0.0858 \\ & & & & 0.3697 \end{bmatrix}$$

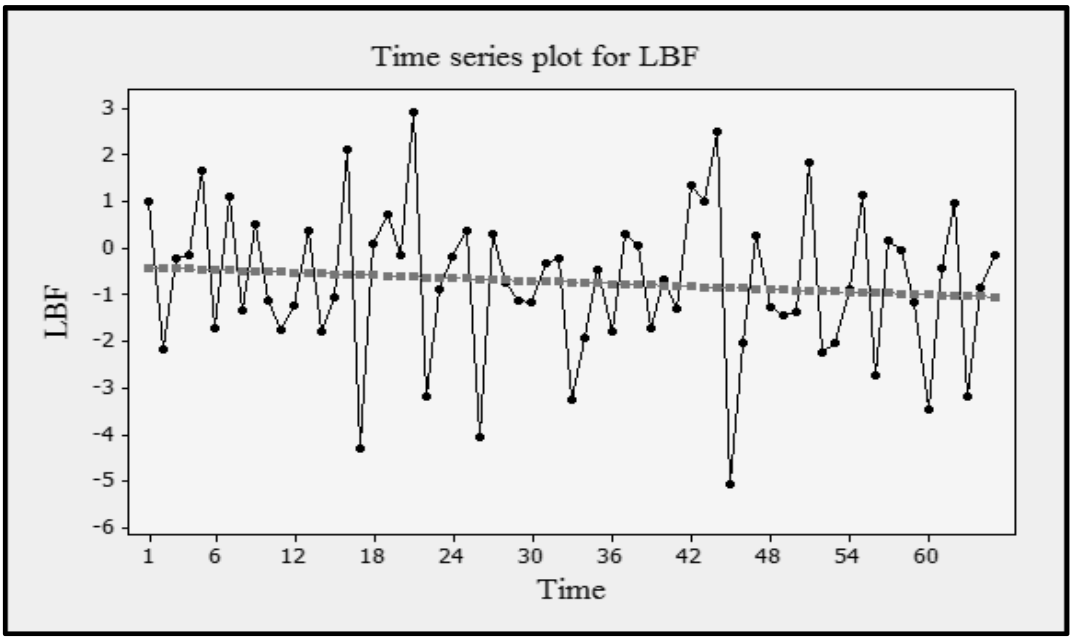
وباعتبار أن متجه معدل الهدف ($\underline{\theta}$) معرف ومساوي لمتجه البيانات المولدة ($\underline{\mu}$) وأن مصفوفة تباين المشاهدة ($\Sigma_{\underline{\varepsilon}}$) ومصفوفة تباين الهدف ($\Sigma_{\underline{\tau}}$) معروفتين وتساويان مصفوفة التباين للبيانات المولدة (Σ).

تم تجزئة العينة إلى جزأين، الجزء الأول (1,2,…,65) والجزء الثاني (66,67,…,100)، حيث تم إيجاد المقدرات عن طريق الجزء الأول وتطبيقها في الجزء الثاني، ولغرض ترشيح المعلمات تم فرض أن متجه متوسط التوزيع اللاحق لمتجه المعلمات ($\underline{\theta}_t$) عند الزمن صفر هو نفسه متجه متوسط البيانات المولدة ($\underline{\mu}$) وأن تباينه عبارة عن مصفوفة الوحدة، وباستخدام البرنامج رقم (2) في الملحق وبتطبيق المعادلات (10) و(11) و(15) و(16) و(17) الواردة في القسم الرابع من هذا البحث تم الحصول على الأخطاء التنبؤية وقد تم اختبارها باستخدام المعيار المعرف في الصيغة (20) فتم التأكد من إمكانية استخدام هذا النموذج في إيجاد عوامل بيز حيث كانت قيمة متجه (MSSE) قريبة من متجه واحدي وكالآتي:

$$MSSE = [0.7544 \quad 1.0000 \quad 1.2083 \quad 1.0487 \quad 1.0108]'$$

قبل إيجاد عوامل بيز يتم إيجاد معدل وتباين وخطأ الهدف وذلك عن طريق المعادلتين (21) و(22) إلا أننا فرضنا أنهما معلومتان كما مر بنا سابقاً، ومن ثم نقوم بإيجاد عامل بيز المعرف وفق المعادلة (25) في القسم الخامس من البحث بواسطة البرنامج رقم (2)، وبذلك نكون قد حصلنا على الركيزة الأساسية لتكوين اللوحات.

يتم التأكد من استقرارية سلسلة عوامل بيز الناتجة من الخطوة السابقة وذلك باستخدام البرنامج الجاهز (Minitab 16.1) وكما موضح في الشكل التالي:



الشكل (1): يوضح استقرارية سلسلة عوامل بيز

أيضاً وباستخدام البرنامج الجاهز (Minitab 16.1) تم الحصول على الجدول التالي:

الجدول (1): يوضح بعض الحالات المختلفة لنموذج (ARMA)

MSE	ϑ	φ	c	ARMA order
2.347	0.6010*	0.2184(n.s)	-0.5877**	ARMA(1,1),c
2.454	-	-0.2832*	-0.9620**	AR (1),c
2.333	0.4236**	-	-0.7519**	MA(1),c
3.226	-0.1609(n.s)	-0.2111(n.s)	-	ARMA(1,1)
3.177	-	-0.0478(n.s)	-	AR (1)
3.178	0.0424(n.s)	-	-	MA(1)

حيث أن (*) تشير إلى أن القيمة معنوية عند (0.05) وأن (**) تشير إلى أن القيمة معنوية عند (0.01) وأن (n.s) تشير إلى أن القيمة غير معنوية.

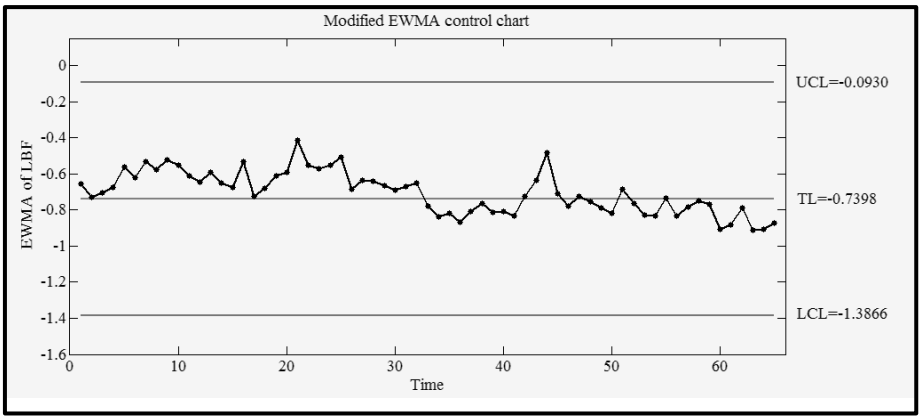
يتبين من الجدول أعلاه أن النموذج MA(1) والذي يحتوي على الثابت c هو النموذج الملائم لتمثيل سلسلة عوامل بيز ولذلك يمكن التعبير رياضياً عنها بالصيغة التالية: $(LBF)_t = -0.7519 - 0.4236\eta_{t-1} + \eta_t$

بالاعتماد على النموذج المذكور وقيم المعلمات التي تم الحصول عليها في الجدول (1) وتعويضها في المعادلات (33) و(43) و(35) الموجودة في القسم السادس من هذا البحث وعلى فرض أن قيمة معلمة التمهيد قيمة صغيرة (λ=0.05) وأن عرض اللوحة (L=3) وبالاستعانة بالبرنامج رقم (3) الموجود في الملحق نحصل على المعلومات التالية:

الجدول (2): يوضح القيم التي تدخل في إيجاد حدود السيطرة

الإحصاءة	القيمة المقدَّرة	الإحصاءة	القيمة المقدَّرة
φ	0	UCL	-0.0930
ϑ	0.4236	TL	-0.7398
c	-0.7519	LCL	-1.3866
σ^2	2.333		

باستخدام البرنامج (MATLAB 7.10) تم رسم لوحة السيطرة للجزء لأول من البيانات كما يلي:



الشكل (2): يوضح لوحة (EWMA) المحورة

نلاحظ أن جميع القيم داخل حدود السيطرة وهذا يدل على إمكانية استخدامها في السيطرة على نوعية هذه البيانات في المستقبل.

لغرض معرفة كون البيانات المتبقية خارج السيطرة أو داخل السيطرة، يتم الاستفادة من النتائج التي تم الحصول عليها من الجزء الأول من البيانات وهذه النتائج بالإضافة إلى السابقة هي كما سسما يلي:

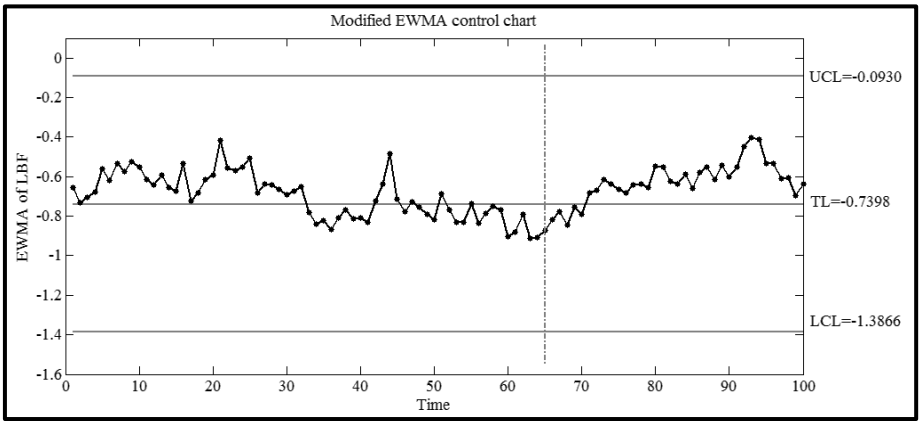
$$\underline{\mu}_{opt} = \underline{\mu}_{65|64} =$$

$$\begin{bmatrix} 10.3454 & 15.3279 & 19.8188 & 25.5234 & 29.1249 \end{bmatrix}'$$

$$\Sigma_{opt} = \Sigma_{65|64}$$

$$\begin{bmatrix} 1.0979 & -0.1712 & -0.3330 & -0.1771 & -0.2333 \\ & 0.6290 & 0.0552 & -0.1645 & -0.0068 \\ & & 0.9596 & -0.0762 & -0.0769 \\ & & & 0.8662 & 0.0882 \\ & & & & 0.7315 \end{bmatrix}$$

حيث أن المتجه أعلاه يمثل المعدل الأمثل لمتجه المعلمات $\underline{\theta}$ كما أن المصفوفة تمثل التباين الأمثل لمتجه المعلمات، وبتطبيق هذه القيم بدلاً من القيم المتغيرة عبر الزمن $\underline{\mu}_{t|t-1}$ و $\Sigma_{t|t-1}$ للجزء الثاني من البيانات فنحصل على عوامل بيز بطريقة مشابهة تماماً إلى تلك الطريقة المستخدمة في الجزء الأول من البيانات، وبرسم سلسلة عوامل بيز وبنفس حدود السيطرة السابقة نحصل على اللوحة للبيانات الحالية وكما موضح في الشكل التالي:



الشكل (3): يوضح لوحة (EWMA) المحورة

نلاحظ أن قيم الجزء الثاني أيضاً داخل حدود السيطرة وهذا يدل على أن العملية تحت السيطرة.

من الجدير بالإشارة إلى أن في حالة كون العملية خارج السيطرة يمكن معرفة المتغير المسبب في ذلك باستخدام نفس الأسلوب السابق مع كل متغير على حدة.

كما تجدر الإشارة إلى أن طريقة تقسيم البيانات إلى جزأين ومن ثم استخدام الجزء الأول من البيانات في التقدير واستخدام الجزء الثاني في التنبؤ من الطرق الجيدة في معرفة العملية الإنتاجية إلا أنه ليس بالضرورة القيام بها.

8- الاستنتاجات والتوصيات

أولاً: الاستنتاجات

نستنتج مما سبق أن:

أ- بالإمكان استخدام لوحة EWMA المحورة في السيطرة على أي سلسلة زمنية غير طبيعة شريطة أن تكون متماثلة.

ب– أن عوامل بيز هي الركيزة الأساسية في تكوين اللوحة المقترحة.

جـ– قيم عامل بيز تحتوي على ارتباط ذاتي سواءً كانت البيانات الأصلية مرتبطة ذاتياً أو غير مرتبطة ذاتياً.

د– يمكن معرفة المتغير المسبب في حالة كون العملية خارج السيطرة وذلك باستخدام نفس الأسلوب السابق مع كل متغير على حدة.

ثانياً: التوصيات

بناءً على نتائج الدراسة نوصي بما يلي:

أ– نوصي بتطبيق اللوحات الناتجة في مجالات الحياة كافة وخاصة تلك المجالات التي تكون توزيعاتها على الأغلب ذات ذيول ثقيلة.

ب– نوصي بتكوين لوحات في حالة كون متجه المعلمات ومصفوفة التباين غير معلومتين.

جـ– نوصي بتعديل اللوحة لتشمل نموذج (ARIMA(p,d,q)) لأي رتبة كانت ولأي عدد من الاختلافات.

الملحق

البرنامج رقم (1)

```
v=input('Input freedom of degree (v):');
N=input('Input size of sample (N):');
theta=input('Input the mean vector (theta):');
Sigma=input('Input the variance matrix:');
data=mvtrnd(Sigma,v,N)';
for t=1:N
   X(:,t)=theta+data(:,t);
end
```

البرنامج رقم (2)

```
%INPUTS
X=input('Input the data:');
F=input('Input F matrix:');
G=input('Input G matrix:');
S_epsilon=input('Input observation variance matrix:');
S_omega=input('Input state variance matrix:');
mu_post(:,1)=input('Input initial mean vector:');
S_post(:,:,1)=input('Input initial variance matrix:');
p=length(mu_post(:,1));
%FILTERING
for t=2:n+1    %n is a part from number of data
   mu_prior(:,t)=G*mu_post(:,t-1);
   S_prior(:,:,t)=G*S_post(:,:,t-1)*G'+S_omega;
   S_error(:,:,t)=F*S_prior(:,:,t)*F'+S_epsilon;
   K(:,:,t)=S_prior(:,:,t)*F'/S_error(:,:,t);
   e(:,t)=X(:,t-1)-F*mu_prior(:,t);
   mu_post(:,t)=mu_prior(:,t)+K(:,:,t)*e(:,t);
   S_post(:,:,t)=(S_prior(:,:,t)-
(K(:,:,t)*F*S_prior(:,:,t)));
   S_e(:,:,t)=(v/(v-
2))*(F*S_prior(:,:,t)*F'+S_epsilon);
   e_star(:,t)=power(S_e(:,:,t),-0.5)*e(:,t);
end
%MSSE
for i=1:p
for j=1:n
   SSE(i,j)= power(e_star(i,j+1),2);
end
end
```

```
MSSE=sum(SSE,2)/n;
%Target distribution
for t=1:n
   T(:,t)=X(:,t);
end
mu_target=mean(T')';
S_target=cov(T');
for t=2:N+1
   tau(:,t)=X(:,t)-mu_target;
end
%Bayes' factor
for t=2:n+1
BF(t)=sqrt(det(S_target/S_error(:,:,t)))*power((v+(tau
(:,t)'/S_target*
tau(:,t)))/(v+(e(:,t)'/S_error(:,:,t)*e(:,t))),(v+p)/2);
end
for t=1:n
LBF(t)=log(BF(t+1));
end
```

البرنامج رقم (3)

```
l= input('Input the smoothing parameter:');
f= input('Input phi value:');
h= input('Input theta value:');
L= input('Input L value:');
s= input('Input standard deviation value:');
c= input('Input c value:');
z(1)=c/(1-f);
for t=1:n
z(t+1)=l*LBF(t)+(1-l)*z(t);
TL(t)=c/(1-f);
Q1(t)=(1-f*(1-l))*(1+power(h,2)-2*f*h);
Q2(t)=(f-h)*(1-f*h)*(1-l);
Q(t)=Q1(t)+Q2(t);
Var(t)=(l*power(s,2)*Q(t))/((2-l)*(1-power(f,2))*(1-
f*(1-l)));
UCL(t)=TL(t)+L*sqrt(Var(t));
LCL(t)=TL(t)-L*sqrt(Var(t));
y(t)=z(t+1);
end
```

المصادر

[1] Pan, Xia and Jarrett Jeffrey, (2004), "**Applying state space to SPC: monitoring multivariate time series**", Journal of Applied Statistics, Vol. 31, pp. 397-418.

[2] Triantafyllopoulos, Kostas, (2006), "**Multivariate control charts based on Bayesian state space models**", Quality and Reliability Engineering International, Vol. 22, pp. 693–707.

[3] Smith, Adrian and West, Mike, (1983), "**Monitoring rental transplants: An application of the multiprocess Kalman filter**", Biometrics, Vol. 39, pp. 867-878.

[4] West, Mike, (1986), "**Bayesian Model Monitoring**", Royal Statistical Society, Vol. 48, pp. 70-78.

[5] Jalil, Talib Sharif, (1988), "**Sequential inspection and intervention policy for a manufacturing process**", Unpublished Ph. D. Thesis, Dept. of Mathematics, Wales University.

[6] أبو ناقوس, أحمد حسن عليان, (1993), "**استخدام إحصاء بيز على النماذج الديناميكية الخطية مع تصميم لوحة بيز**", رسالة ماجستير, قسم الإحصاء, كلية الإدارة والاقتصاد, جامعة الموصل, الموصل, العراق.

[7] الرسام، ريا سالم, (1996)، "**تكوين لوحة بيز ثنائية البعد لغرض مراقبة جودة الإنتاج مع المحاكاة**"، رسالة ماجستير, قسم الإحصاء، كلية الإدارة والاقتصاد، جامعة الموصل, الموصل, العراق.

[8] الزبيدي، طه حسين، **(1997)**، "**تكوين لوحة بيز للسيطرة على الصفات النوعية**"، رسالة ماجستير, قسم الإحصاء، كلية الإدارة والاقتصاد، جامعة الموصل, الموصل, العراق.

[9] Salvador, Manuel and Gargallob, Pilar, (2004), "**Automatic Monitoring and Intervention in Multivariate Dynamic Linear Models**", Computational Statistics & Data Analysis, Vol. 47, pp. 401 – 431.

[10] سعيد, هيفاء عبد الجواد واللهيبي, معتز بالله صلاح غانم, (2011), "**التقدير وبناء الاختبارات البيزية لمعلمتي الموقع والقياس في توزيع t متعدد المتغيرات**", المجلة العراقية في العلوم الإحصائية, عدد خاص بوقائع المؤتمر العلمي الرابع لكلية علوم الحاسوب والرياضيات, ص 297 – 314.

[11] Sorensen, Daniel and Gianola, Daniel, (2002), " **Likelihood, Bayesian, and MCMC methods in quantitative genetics**", Springer-Verlag, New York.

[12] Kotz, Samuel and Nadarajah, Saralees, (2004), "**Multivariate t distributions and their applications**", cambridge university press, U.K.

[13] Kanjilal, Partha Pratim, (1995), "**Adaptive prediction and predictive control**", Peter Peregrinus Ltd., London.

[14] Akaike, Hirotugu, (1974), "**Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes**" Annals of the Institute of Statistical Mathematics, Vol. 26, pp.363-387.

[15] Grewal , Mohinder S. and Andrews, Angus P., (2008), "**Kalman filtering: theory and practice using MATLAB**", 3rd ed., John Wiley & Sons, Inc. .

[16] West, Mike, and Harrison, Jeff, (1997), "**Bayesian forecasting and dynamic models**", 2th ed., Springer-Verlag, New York.

[17] Box, George and Tiao, George, (1973), "**Bayesian inference in statistical analysis**", Addison – Wesley Publishing Company, London, U.K .

[18] Kass, Robert E. and Raftery, Adrian E., (1995), "**Bayes factors**", Journal of the American Statistical Association, Vol. 90, pp. 773-795.

[19] Montgomery, Douglas C., (2009)," **Introduction to statistical quality control** ", 6th ed., John Wiley & Sons, Inc. .

[20] Schmid, Wolfgang, (1997), "**On EWMA charts for time series. in frontiers in statistical quality control**", Vol. 5, pp. 114-137.

Construction of a Control Chart Using SSM for Multivariate t Distribution Data

Hayfa Abdul Jawad Saieed , Delshad Shaker Ismail Botani , Adnan Mustafa Hussein AL-Sinjary

[1] *Statistics and Informatics Department, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq*

[2] *Department of Statistics, College of Management and Economics, University of Salahaddin, Erbil, Erbil, Iraq*

Abstract

The main purpose of this research is to construct a chart for controlling the mean and variance together of a data distributed multivariate **t** distribution using State Space Model (SSM) through applying Bayes' Factors (BF). The constructed control chart will undertake the case when parameters' vector is unknown and variance matrix is known. Then, drawing the series of these factors on the univariate modified EWMA chart after assuming that the series has an ARMA (1,1) model. Finally, applying assumed model on simulated data.

Keywords: State Space Model, Bayes' Factors, Modified EWMA, ARMA.