العنوان:	حول مرشح كالمن الجذر التربيعي لبعض النماذج الخطية الحركية و المقارنة مع مرشح كالمن الاعتيادي
المؤلف الرئيسي:	الخالدى، زيد طارق صالح عباوى
مؤلفين آخرين:	جليل، طالب شريف(مشرف)
التاريخ الميلادي:	2005
موقع:	الموصل
الصفحات:	1 - 79
رقم MD:	559514
نوع المحتوى:	رسائل جامعية
اللغة:	Arabic
الدرجة العلمية:	رسالة ماجستير
الجامعة:	جامعة الموصل
الكلية:	كلية علوم الحاسبات والرياضيات
الدولة:	العراق
قواعد المعلومات:	Dissertations
مواضيع:	الخوارزميات ، الإحصاء الرياضي ، الجذور التربيعية ، مرشح كالمن ، المصفوفات
رابط:	http://search.mandumah.com/Record/559514

حول مرشح كالمن الجذر التربيعي لبعض النماذج الخطية الحركية والمقارنة مع مرشح كالمن الاعتيادي

رسالة تقدّم بها

زيد طارق صالح عبّاوي الخالدي

بكالوريوس إحصاء

إلى

مجلس كلية علوم الحاسبات والرياضيات في جامعة الموصل

وهي جزء من متطلبات نيل شهادة ماجستير علوم في الإحصاء

بإشراف

الدكتور طالب شريف جليل

أستاذ مساعد

كلية علوم الحاسبات والرياضيات – قسم الإحصاء

١٤٢٦ هجري ٢٠٠٥ ميلادي

بسم الله الرحمن الرحيم

اقْرَأْ بِاسْمِ رَبِّكَ الَّذِي خَلَقَ ﴿1﴾ خَلَقَ الْإِنسَانَ مِنْ عَلَقٍ ﴿2﴾
اقْرَأْ وَرَبُّكَ الْأَكْرَمُ ﴿3﴾الَّذِي عَلَّمَ بِالْقَلَمِ ﴿4﴾
عَلَّمَ الْإِنسَانَ مَا لَمْ يَعْلَمْ ﴿5﴾
صدق الله العظيم

سورة العلق – الآيات ١-٥

شكر وتقدير

الحمد لله ربّ العالمين.. والصلاة والسلام على محمد الأمين.. وعلى آله وصحبه أجمعين..

أقدم عظيم امتناني وخالص شكري للأستاذ الفاضل **الدكتور طالب شريف جليل** لما قدمه لي من توجيهات وتوصيات طيلة فترة إشرافه على الرسالة، كما لا أنسى أن أقر بفضله عليّ بعد الله في زرع الرغبة لدي لخوض غمار البحث العلمي.. فجزاه الله عني خير الجزاء..

كما أقدم شكري العميق للأساتذة الأفاضل أعضاء لجنة المناقشة لتفضلهم بمناقشة رسالتي مثمناً في الوقت نفسه التصويبات والآراء المقدمة لإغناء الرسالة بأُسس البحث العلمي وقواعده..

ويسرني أن أتقدم بالشكر الجزيل لأساتذتي الأكارم في قسم الإحصاء الذين كان لهم الفضل الكبير للوصول بي إلى هذه المرحلة، أدامهم الله ذخراً لهذا القسم..

وأقدم امتناني لمنتسبي المكتبة المركزية – قسم الدوريات، ومكتبة الكلية لتعاونهم الدائم، وفقهم الله لخدمة بلدنا العزيز..

ولا أنسى أن أقدم شكري واحترامي إلى الزملاء في الدراسة لما قدموه لي من دعم لوجستي لإتمام هذه الرسالة، فبارك الله فيهم جميعاً..

وجزا الله خيراً كل من دعا لي في ظهر الغيب..

زيـــد

الملخص

وضع كالمن في أوائل ستينات القرن الماضي، مجموعة من المعادلات التعاقبية على هيئة خوارزمية زمنية تعاقبية بغية إيجاد المقدر الخطي الأمثل لحالة (State) النظام الخطي الحركي. عرفت هذه الخوارزمية بمرشح كالمن الاعتيادي. إلا أن هذه الخوارزمية ليست ملائمة لجميع الظروف التي يتطلبها النظام. إن أحد أهم المشاكل التي تواجهنا عند استخدام خوارزمية مرشح كالمن الاعتيادي في التطبيقات العملية - خصوصاً التطبيقات التي تتطلب دقة عالية في النتائج- هي مشكلة فقدان مصفوفة التباين والتباين المشترك لبعض خواصها الضرورية، مثل عدم كونها أكيدة الإيجابية، أو تحولها إلى مصفوفة شاذة، ويحدث ذلك بسبب الخاصية التعاقبية لخوارزمية الترشيح الاعتيادية.

تمّ في هذه الدراسة توضيح بعض الخوارزميات البديلة لخوارزمية مرشح كالمن الاعتيادي، وتعرف هذه الخوارزميات بمرشح كالمن الجذر التربيعي، حيث تكون هذه الخوارزميات أقل تأثراً بالمشاكل العددية. وقد تمّ توضيح ثلاث خوارزميات بشكل نظري وتمّ ذلك على بعض النماذج الخطية الحركية في الأزمنة المتقطعة بالاعتماد على أسلوب بيز، وأيضاً تمّ إجراء تطبيق عملي باستخدام بيانات المحاكاة لمقارنة اثنتين من الخوارزميات الثلاث مع خوارزمية مرشح كالمن الاعتيادي، وتمّ التركيز في المقارنة على مصفوفة التباين والتباين المشترك وربحية المرشح.

المحتويــــــــات

قائمة بالأشكال الواردة في الرسالة

المــقـــدمـــــــة

إن الغاية الأساسية من الترشيح هي إيجاد المقدر الأمثل (Optimal Estimator) للمعلمة. حيث أن مصطلح الترشيح يعني في علم الإحصاء التقدير في الزمن الحاضر، أي إيجاد مقدر أمثل للمعلمة (θ_τ) عندما تكون $(\tau = t)$ حيث أن (t) تمثل الزمن الحاضر [Kalman, 1960]. وكانت الأساسيات التي بنيت عليها نظرية التقدير عند ظهورها للمرة الأولى تعتمد على جعل دالة تباين الخطأ أقل ما يمكن، وكما هو معلوم فإن طريقة المربعات الصغرى وجدت لأول مرة على يد العالم (Gauss) عام ١٧٩٥. ومنذ ذلك الحين تتابعت الدراسات الموسعة في هذا المجال، وخصوصاً فيما يتعلق بالتقدير الخطي الأمثل (Optimal Linear Estimation) وقد ظهرت أولى الدراسات الرائدة في مسألة الترشيح في مطلع الأربعينات من القرن الماضي، حيث قام كلٌ من (Kolmogorov) عام ١٩٤١، و(Wiener) عام ١٩٤٢ بشكل مستقل بتطوير تقنية لإيجاد مقدر خطي ذي أقل متوسط مربع خطأ (Linear Minimum Mean-Square Error Estimator) والذي لقي اهتماماً بالغاً وكان له الأثر الكبير لاحقاً في تطوير فكرة مرشح كالمن.[Sorenson, 1970]، ثم توّج [Wiener, 1949] دراسته تلك بنتائج أخرى تمخضت عن وضع خوارزمية تعاقبية (Recursive Algorithm) لإيجاد المقدر الخطي الأمثل، وكانت هذه الخوارزمية مقيدة بشروط منها أن المشاهدة المدروسة ذات بُعد أحادي (لا توجيهي Scalar) فضلاً عن المعلمة، وأن البيانات شبه غير منتهية $(t_0 \to -\infty)$ فضلاً عن ذلك فإن العملية تتسم بالاستقرار (Stationary Process).

وفي عام ١٩٥٠ قام (Zadeh & Ragazzini) بتطوير النتائج التي توصل إليها (Wiener) وذلك بجعلها أكثر عمومية وذلك من حيث جعل البيانات منتهية $(t_0 > -\infty)$ ولكن العمليات تتسم بعدم الاستقرار (Non-Stationary Processes). تلا ذلك دراسات أخرى في المجال نفسه، نذكر منها على سبيل المثال لا الحصر الدراسة التي قام بها (Dolph & Woodbury) في عام ١٩٥٢ و (Darlington) في عام 1959 والتي لم تتعد النتائج التي توصل إليها (Zadeh & Ragazzini). [Kailath, 1974]

ومع كثرة الدراسات في هذا المجال فقد بقيت عقيمة نوعاً ما عن تفسيرها لظواهر عديدة ظهرت خصوصاً في أواخر الخمسينات بعد دخول عصر الفضاء وظهور مشكلة تحديد مدار القمر الصناعي (Satellite Orbit) من حيث السرعة (Velocity) كمعلمة أولى والتعجيل (Acceleration) كمعلمة ثانية، وذلك لأسباب عدة، مما حدا بالباحثين إلى محاولة إيجاد خوارزمية تتغلب على هذه المشاكل.

وكانت أولى الخوارزميات التي تعالج هذه المشاكل هي الخوارزمية التي وضعها (Swerling) عام (١٩٥٩)، حيث تمكن سويرلنك من وضع أول خوارزمية تعاقبية لحل المشاكل التي تعالج مشكلة تعدد المعلمات، فضلاً عن المشاكل الأخرى، بعد ذلك قام [Kalman, 1960] بوضع خوارزمية تعاقبية أكثر تعقيداً من تلك التي وضعها سويرلنك لكنها تميزت عنها بكونها أول خوارزمية تهتم بدراسة مشاكل تقدير الحالة الحركي Dynamical State Estimation Problems، ويسرت هذه الخوارزمية فهم العديد من المشاكل والظواهر التطبيقية وتحليلها، وقد كانت الخوارزمية التي وضعها كالمن تهتم بدراسة تقدير الحالة الحركي في الأزمنة المتقطعة (Discrete Time) أي أن (t=1,2,…). ثم قام [Kalman & Bucy, 1961] بوضع خوارزمية لتقدير الحالة الحركي في الأزمنة المستمرة (Continuous Time) وهو ما عرف لاحقاً بـ (Kalman-Bucy Filter). [Kailath, 1974]

وقد اعتمد كالمن في إيجاده للمقدر الخطي الأمثل على فكرة الإسقاط المتعامد (Orthogonal Projection)، وقد توصل كالمن إلى أن المقدر الخطي الأمثل للمعلمة (θ_t) هو التوقع الشرطي لـ(θ_t) مشروطاً بكافة المعلومات المتوافرة حتى الزمن (t) أي:

$$E(\theta_t \mid D_t)$$

حيث أن D_t تمثل المعلومات المتوافرة حول المعلمة حتى الزمن الحاضر (t).

وأن هذا المقدر يمكن الحصول عليه إذا توافر أحد الشرطين الآتيين:

١. إذا توزع كلٌ من المعلمة (θ_t) والمشاهدة (y_t) وفق التوزيع الطبيعي.

٢. إذا كانت دالة الخسارة هي دالة الخسارة التربيعية (Quadratic Loss Function)، أي أن: $l(\hat{\theta}_t,\theta_t)=(\hat{\theta}_t-\theta_t)^2$. [Kalman, 1960]

ومن ثمّ لاحظ كل من هارسون وستيفنس [Harrison & Stevens, 1976]، أن الشروط أعلاه مماثلة لإيجاد مقدر بيز ذي أقل متوسط مربع خطأ Minimum Mean Square Error Estimator أو ما يعرف بالمقدر الأمثل والذي يساوي التوقع المشروط لـ (θ_t) بالمعلومات (D_t)، وأثبتا أيضاً أن مقدر بيز ذا أقل متوسط مربع خطأ عند إيجاده بشكل متعاقب (Sequential) فإنه يتساوى مع المقدر الخطي الأمثل الذي وضعه كالمن. إن استخدام نظرية بيز في عملية تكوين مرشح كالمن تعطي سلاسة وسهولة عاليتين في العمليات الحسابية التي يتطلبها تكوين مرشح كالمن مقارنة مع الأسلوب الذي وضعه كالمن للمرة الأولى عام ١٩٦٠ والذي يتسم بالتعقيد بعض الشيء. ومن ابرز من استخدم أسلوب بيز في تكوين مرشح كالمن [Maybeck, 1979] و[Meinhold & Singpurwalla, 1983] و[West & Harrison, 1989] و[Calder, *et al.* 2003] وغيرهم آخرون.

بعد ذلك تتابعت الدراسات النظرية والتطبيقية وبشكل موسع جداً على ما سمي لاحقاً بنظرية مرشح كالمن (Kalman Filter Theory)، تضمنت تطوير خوارزميات محدثة لخوارزمية مرشح كالمن الاعتيادي (Ordinary Kalman Filter Algorithm) وذلك للتغلب على المشاكل التي تواجه المستخدم لهذه الخوارزمية، مثل مشكلة النماذج غير الخطية، أو عندما يكون تباين الأخطاء غير معلوم، أو عندما تكون الأخطاء العشوائية مترابطة، انظر [الحمداني، ١٩٩٦]، وكذلك الخوارزميات التي وضعت للتغلب على المشاكل العددية التي تحصل في مصفوفة التباين* C_t بسبب الإجراء التعاقبي.

إن المشاكل التي تحدث في مصفوفة التباين تؤدي إلى حصول تباعد (Divergence) في خوارزمية مرشح كالمن الاعتيادي، مما يؤدي إلى فشل الإجراء، أو الحصول على نتائج غير دقيقة أو غير واقعية كأن ينخفض التباين بشكل ملحوظ، أو بقاء الخوارزمية غير مستقرة أي عدم الوصول إلى حالة الاستقرار (Steady State). لذلك وضعت صيغ وخوارزميات عدة للتغلب على مشكلة التباعد هذه.

وكانت معظم تلك الخوارزميات تعتمد على فكرة إيجاد مصفوفة الجذر التربيعي لمصفوفة التباين وذلك لأن الدقة المتوافرة تكون أعلى، فضلاً عن أن مصفوفة الجذر التربيعي تكون أكيدة الإيجابية. [Bierman, 1977]، لذلك عرفت هذه الخوارزميات بشكل عام بخوارزمية مرشح كالمن الجذر التربيعي (Square Root Kalman Filter). [Maybeck, 1979]

ويعد بوتر (Potter) أول من وضع خوارزمية مرشح كالمن الجذر التربيعي في عام (١٩٦٣)، عرفت خوارزميته تلك بخوارزمية مرشح الجذر التربيعي لبوتر [Carlson, 1973]، وتشترط هذه الخوارزمية أن تكون المشاهدات مفردة البعد (Scalar) أي أن ($m=1$)، وأن تكون المعلمات خالية من التشويش (Noise) أي تكون ($W_t=0$)، مما جعلها مقيدة بالرغم من كفاءتها.

ثمّ قدم [Bellantoni, & Dodge, 1967] بحثاً تضمن نتائج جديدة في هذا المجال، حيث وضعا خوارزمية جديدة لا تتقيد بشرطي خوارزمية بوتر. مع ذلك فإن هذه الخوارزمية تكون غير كفوءة بالرغم من فعاليتها وذلك في الحالات التي تكون فيها عدد المشاهدات أقل من عدد المعلمات في النموذج المدروس، أي أن ($m<n$). [Kaminski, *et al*. 1971]

وقدم [Andrews, 1968] خوارزمية أخرى لمعالجة مشاكل مصفوفة التباين، حيث تعتمد هذه الخوارزمية في إيجاد مصفوفة الجذر التربيعي على تحليل چولسكي Cholesky Decomposition.

تلا ذلك العديد من الدراسات والبحوث التي تمخضت عن تطوير خوارزميات جديدة لمرشح كالمن الجذر التربيعي تعالج بعض المشاكل العامة، مثل الخوارزمية التي وضعها [Carlson, 1973] والتي تستخدم في النماذج التي تحوي على أخطاء مترابطة (Correlated Errors).

---

* المقصود مصفوفة التباين والتباين المشترك (Covariance Matrix)، وقد كتبت بهذه الصيغة في الرسالة لغرض الاختصار .

أو الخوارزميات التي وضعت لمعالجة مشكلة النماذج غير الخطية مثل خوارزمية مرشح كالمن الجذر التربيعي الموسع (Extended Square Root Kalman Filter)، أو خوارزمية الجذر التربيعي المتوازية (Parallel Square Root Algorithm). [Lu, *et al.* 1992]

لقد استخدم مرشح كالمن الجذر التربيعي على نطاق واسع في الدراسات والتطبيقات الهندسية وخصوصاً هندسة الاتصالات مثل تطبيقات أجهزة الرصد الجوي والفضائي، وفي الصناعات الدوائية، حيث تتطلب هذه التطبيقات دقة عالية في الحسابات.

ولقد وجهنا اهتمامنا في هذه الرسالة بدراسة مرشح كالمن الجذر التربيعي الذي يعد أحد الطرائق لمعالجة المشاكل التي تظهر في مصفوفة التباين في مرشح كالمن الاعتيادي نتيجة للتكرار المتعاقب، حيث تم توضيح بعض خوارزميات مرشح الجذر التربيعي فضلاً عن توضيح الخوارزمية الأصلية لكالمن وذلك في بعض النماذج الخطية الحركية، وذلك بالاعتماد على أسلوب بيز، وتمّ إجراء تطبيق عملي على بيانات المحاكاة Simulation Data التي تمّ توليدها باستخدام نظام ماتلاب Matlab 7.0، وقد وُلّدت هذه البيانات لتكون متوافقة مع النماذج المدروسة مع مراعاة متطلبات الدراسة، وتمت مقارنة النتائج بين مرشح كالمن الجذر التربيعي مع مرشح كالمن الاعتيادي وذلك لبيان الفرق بين الإجراءين.

ولإتمام الهدف من الرسالة بشكل علمي، ولكي يتم الإلمام بمفهوم المواضيع الواردة فيها، فقد جاءت الرسالة بثلاثة فصول، تناول الفصل الأول استدلال بيز ونظرية التقدير والنماذج الخطية الحركية، حيث تم توضيح طبيعة استدلال بيز، فضلاً عن بعض المقدرات البيزية المعتمدة على دالة المخاطرة، كما تمّ توضيح العلاقة بين استدلال بيز والنماذج الخطية الحركية، وتضمن الفصل أيضاً شرحاً مبسطاً لبعض النماذج الخطية الحركية فضلاً عن إيجاد التوزيعين السابق واللاحق لمعلمات تلك النماذج. فيما خصص الفصل الثاني للبحث في نظرية الترشيح وتاريخ تطورها وتوضيح أهم خوارزمية في هذا المجال، وهي خوارزمية مرشح كالمن الاعتيادي، وتم أيضاً توضيح مشكلة التباعد التي تحصل في مصفوفة التباين نتيجة للإجراء التكراري المتعاقب وكيفية التغلب على هذه المشكلة، حيث تم توضيح ثلاث خوارزميات لمرشح كالمن الجذر التربيعي وهن خوارزمية مرشح الجذر التربيعي لبوتر، وخوارزمية مرشح كالمن-شميدت الجذر التربيعي، وخوارزمية مرشح الجذر التربيعي لمصفوفة التباين والتباين المشترك، فضلاً عن بعض الفقرات الأخرى المهمة. واختتمت الرسالة بفصل ثالث تمّ فيه إجراء تطبيق عملي باستخدام بيانات المحاكاة لاثنتين من خوارزميات مرشح الجذر التربيعي، فضلاً عن مقارنة النتائج مع نتائج خوارزمية مرشح كالمن الاعتيادي.

الملخص

وضع كالمن في أوائل ستينات القرن الماضي، مجموعة من المعادلات التعاقبية على هيئة خوارزمية زمنية تعاقبية بغية إيجاد المقدر الخطي الأمثل لحالة (State) النظام الخطي الحركي. عرفت هذه الخوارزمية بمرشح كالمن الاعتيادي. إلا أن هذه الخوارزمية ليست ملائمة لجميع الظروف التي يتطلبها النظام. إن أحد أهم المشاكل التي تواجهنا عند استخدام خوارزمية مرشح كالمن الاعتيادي في التطبيقات العملية - خصوصاً التطبيقات التي تتطلب دقة عالية في النتائج- هي مشكلة فقدان مصفوفة التباين والتباين المشترك لبعض خواصها الضرورية، مثل عدم كونها أكيدة الإيجابية، أو تحولها إلى مصفوفة شاذة، ويحدث ذلك بسبب الخاصية التعاقبية لخوارزمية الترشيح الاعتيادية.

تمّ في هذه الدراسة توضيح بعض الخوارزميات البديلة لخوارزمية مرشح كالمن الاعتيادي، وتعرف هذه الخوارزميات بمرشح كالمن الجذر التربيعي، حيث تكون هذه الخوارزميات أقل تأثراً بالمشاكل العددية. وقد تمّ توضيح ثلاث خوارزميات بشكل نظري وتمّ ذلك على بعض النماذج الخطية الحركية في الأزمنة المتقطعة بالاعتماد على أسلوب بيز، وأيضاً تمّ إجراء تطبيق عملي باستخدام بيانات المحاكاة لمقارنة اثنتين من الخوارزميات الثلاث مع خوارزمية مرشح كالمن الاعتيادي، وتمّ التركيز في المقارنة على مصفوفة التباين والتباين المشترك وربحية المرشح.

الفصل الأول

مقدمة في استدلال بيز
والنماذج الخطية الحركية

Introduction in Bayesian Inference
& The Dynamic Linear Models

الفصل الأول

مقدمة في استدلال بيز والنماذج الخطية الحركية

١-١- مــقـدمـــة في الاستدلال:

للتعرف على استدلال بيز لا بد من أخذ فكرة عن ماهية الاستدلال. حيث يعرف الاستدلال بأنه أحد فروع علم الإحصاء والذي يهتم بجمع المعلومات حول معلمات التوزيعات، ومن أساليبه في ذلك نظرية التقدير واختبار الفرضيات.

وقد شهد القرن العشرون تطوراً متسارعاً في الأساليب (approaches) المختلفة المستخدمة في الاستدلال الإحصائي، حيث يمكن تصنيف الاستدلال الإحصائي إلى مدرستين رئيستين تبعاً لاختلاف وجهة النظر الإحصائية حول ماهية المعلمة، وهاتان المدرستان هما:

١. المدرسة الكلاسيكية أو ما يسمى بمدرسة المعاينة (Classical or Sampling School)، وتفترض هذه المدرسة عند دراستها للمعلمة بأنها -أي المعلمة- ثابت غير معلوم (Unknown Constant)، وتستخلص هذه المدرسة دراستها حول المعلمة من خلال المعلومات التي توفرها العينة، لذا تدعى بمدرسة المعاينة. وممن برز في هذه المدرسة العالم فيشر ونيمان وبيرسون وغيرهم آخرون.

٢. مدرسة بيز(Bayesian School) نسبة إلى الإحصائي البريطاني توماس بيز، حيث تفترض هذه المدرسة بأن المعلمة متغير عشوائي له توزيع احتمالي. وتعتمد هذه المدرسة في دراستها للمعلمة على المعلومات التي توفرها العينة فضلاً عن المعلومات التي تأتي من الخبرة أو الاعتقاد الشخصي (Personal Believes).

١-٢- مبرهنة بيز (Bayes theorem):

توماس بيز (Thomas Bayes 1702-1761) هو راهب بريطاني ساهم في إغناء نظرية الاحتمال ببحثين نشرهما بعد وفاته زميله ريتشارد برايس (Richard Price) في عامي (١٧٦٣ و١٧٦٤)، وقد أضاف إلى البحثين بعض الأمثلة والتعليقات فضلاً عن المقدمة. [Barnett, 1982]

ومن أبرز الذين ساهموا في تطوير أسلوب بيز في المجالين النظري والتطبيقي العلماء (Laplace) و(Lindly) و(Box & Tiao) و(Zellner) و(Jeffrey) وغيرهم آخرون.

وبالإمكان الحصول على صيغة مبرهنة بيز الاحتمالية من المفهوم العام للاحتمال الشرطي، وكالآتي:

علـى فرض أن المتغيـر العشوائي المتقطـع X يأخـذ القيم (x_i's ; $i = 1, 2, \ldots, n$)، ولنفرض أيضاً أن Y متغير عشوائي متقطع أيضـاً ويأخذ القيم (y_j's ; $j = 1, 2, \ldots, m$)، فعندئذ تكـون دالـة الكتلـة الاحتماليـة المشـتركة (Joint Probability Mass Function) لهـذين المتغيـرين هـي $P(X = x_i, Y = y_j)$ ، حيث يمكن كتابة هذه الدالة بأسلوب آخر وذلك بالاعتماد على قانون ضرب الاحتمالات:

$$P(x_i, y_j) = \begin{Bmatrix} P(x_i \mid y_j)\, P(y_j) \\ P(y_j \mid x_i)\, P(x_i) \end{Bmatrix} \qquad \ldots\ldots\ldots(1.1)$$

حيث أن $P(x_i \mid y_j)$ يمثل الاحتمال الشرطي لـ x عندما يعطى y ، ومنه نحصل على:

$$P(x_i \mid y_j) = \frac{P(x_i, y_j)}{P(y_j)}$$

أمـا $P(y_j)$ فهـي تمثـل دالـة الكتلـة الاحتماليـة الحديـة للمتغيـر Y، ويـتم إيجادهـا مـن قـانون الاحتمـال الكلي:

$$P(y_j) = \sum_{i=1}^{n} P(y_j \mid x_i)\, P(x_i) \qquad \ldots\ldots\ldots(1.2)$$

وبـالتعويض بالمعـادلتين (١.١) و(١.٢) نحصـل علـى الصـيغة العامـة لمبرهنـة بيـز [Vaseghi, (2000)] :

$$P(x_i \mid y_j) = \frac{P(y_j \mid x_i)\, P(x_i)}{\sum_{i=1}^{n} P(y_j \mid x_i)\, P(x_i)}$$

وبالطريقة نفسها يمكن الحصول على الصيغة العامة لمبرهنة بيز في حالة المتغيرات المستمرة:

$$P(x \mid y) = \frac{P(y \mid x)\, P(x)}{\int_{\Omega X} P(y \mid x)\, P(x)\, dx}$$

حيث أن $P(\cdot)$ تمثل دالة الكثافة الاحتمالية (Probability Density Function).

١-٣- طبيعة استدلال بيز (Nature of Bayesian Inference) :

افرض أن ($y_1, y_2, \ldots, y_n$) متغيـرات عشـوائية لهـا توزيـع احتمـالي $P(y \mid \theta)$ ، حيـث θ تمثـل معلمة التوزيع، فعندئذٍ:

$$P(\theta \mid y_1, y_2, \ldots, y_n) = \frac{\mathrm{P}(y_1, y_2, \ldots, y_n \mid \theta)\, \mathrm{P}(\theta)}{P(y_1, y_2, \ldots, y_n)}$$

حيث أن:

$$P(y_1, y_2, \ldots, y_n) = c^{-1} = \begin{Bmatrix} \sum_{p=1}^{k} P(y_1, y_2, \ldots, y_n \mid \theta_p)\, P(\theta_p) & \text{متغير متقطع } \theta \\ \\ \int_{\theta \in \Omega} P(y_1, y_2, \ldots, y_n \mid \theta)\, P(\theta) d\theta & \text{متغير مستمر } \theta \end{Bmatrix}$$

إذن:

$$P(\theta \mid y_1, y_2, \ldots, y_n) = c\, \mathrm{P}(y_1, y_2, \ldots, y_n \mid \theta)\, \mathrm{P}(\theta)$$

وبما أن c قيمة ثابتة (ثابت التناسب) - لأنها خالية من θ - فيمكن كتابة صيغة مبرهنة بيز بالصيغة الآتية:

$$P(\theta \mid y_1, y_2, \ldots, y_n) \propto \mathrm{P}(y_1, y_2, \ldots, y_n \mid \theta)\, \mathrm{P}(\theta) \qquad \ldots\ldots\ldots(1.3)$$

وكما نرى فإن الصيغة العامة تتكون من ثلاثة أجزاء، حيث يتكون الطرف الأيمن من عاملين الأول $P(\theta)$ يمثل المعلومات المتوافرة حول المعلمة (خبرة أو اعتقاد شخصي) قبل أخذ أي مشاهدة، ويدعى بالاحتمال الابتدائي أو التوزيع الأولي (Prior Distribution) للمعلمة وهو يلعب دوراً مُهماً في مبرهنة بيز ويصنف إلى صنفين سيجري توضيح كلٌ منهما على حدا لاحقاً.

والجزء الثاني هو $P(y_1, y_2, \ldots, y_n \mid \theta)$ دالة الإمكان (Likelihood Function)، وتعد دالة الإمكان دالة لـ (θ) لأننا نهتم بدراسة (θ) ولذلك تكتب $L(\theta)$، وهي تمثل المعلومات المتوافرة من البيانات حول المعلمة.[Box & Tiao (1973)]

وإذا أخذنا بنظر الاعتبار المعلومات المأخوذة من البيانات فضلاً عن تلك المتوافرة من التوزيع الأولي نحصل على ما يسمى بالتوزيع اللاحق (Posterior Distribution) وهو الجزء الثالث من مبرهنة بيز، ويكتب $P(\theta \mid y_1, y_2, \ldots, y_n)$ ويمثل الطرف الأيسر من المعادلة (1.3). ويجب أن نؤكد هنا أن الثابت c يمثل ثابت التطبيع (Normalization Constant) وهو العامل الذي يجعل تكامل التوزيع اللاحق (أو المجموع) مساوٍ لواحد. وبذلك يمكننا كتابة مبرهنة بيز بالصيغة الآتية:

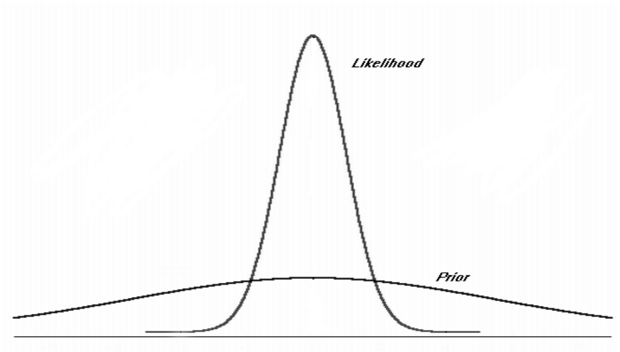
$$P(\theta \mid y_1, y_2, \ldots, y_n) \propto L(\theta)\, P(\theta)$$

وبالرجوع إلى التوزيع الأولي فإنه بالإمكان تصنيفه إلى نوعين وذلك تبعاً لأهمية المعلومات التي يوفرها حول المعلمة θ، وهذان النوعان هما:

١-٣-١- التوزيع الأولي ذو المعلومات القليلة:

Non-Informative Prior Distribution:

عندما تكون دالة الإمكان مهيمنة على التوزيع الأولي فعندئذٍ يعرف ذلك التوزيع بأنه توزيع أولي ذو معلومات قليلة، ويكون هذا التوزيع ذا تأثير قليل على المنطقة التي تكون فيها دالة الإمكان ذات تأثير مدرك (Appreciable) في التوزيع اللاحق بالإضافة إلى ذلك فهو لا يأخذ قيماً كبيرة خارج تلك المنطقة - أنظر الشكل (١-١)- إن التوزيع الذي له الخواص أعلاه يعرف بتوزيع أولي منتظم موضعياً (Locally Uniform Prior Distribution) أو توزيع أولي غير كامل (Improper Prior Distribution). [Box & Tiao, 1973]



الشكل(١-١)
منحنى دالة الإمكان المهيمنة على التوزيع الأولي. [Box & Tiao, 1973]

ويقسم التوزيع الأولي ذو المعلومات القليلة إلى قسمين:

١-٣-١-١- التوزيع الأولي ذو المعلومات القليلة غير القياسي:

Non-Standard Non-Informative Prior Distribution:

هذا النوع من التوزيع الأولي يختلف من شخص إلى آخر، وذلك حسب الخبرة والاعتقاد الشخصي. [الحمداني، ١٩٩٦]

١-٣-١-٢- التوزيع الأولي ذو المعلومات القليلة القياسي:

Standard Non-Informative Prior Distribution:

في هذه الحالة يمكن إيجاد التوزيع الأولي باستخدام قانون جفريز نسبة إلى الإحصائي البريطاني (Harold Jeffreys)، حيث ينص هذا القانون على أن التوزيع الأولي لمعلمة واحدة يتناسب مع الجذر التربيعي لقياس معلومات فيشر (Square Root of Fisher Information Measure)، أي أن:

$$P(\theta) \propto \left(I_y(\theta)\right)^{1/2}$$

حيث أن $I_y(\theta)$ تمثل معلومات فيشر حول المعلمة θ التي يوفرها المتغير العشوائي (y). ويمكن إيجاد معلومات فيشر كالآتي:

$$I_y(\theta) = E\left[\frac{\partial \ln f(y|\theta)}{\partial \theta}\right]^2$$

كما يمكن إيجادها بصيغة بديلة. [Kendall & Stuart, 1976]:

$$I_y(\theta) = -E\left[\frac{\partial^2 \ln f(y|\theta)}{\partial \theta^2}\right]$$

١-٣-٢- التوزيع الأولي ذي المعلومات (Informative Prior Distribution):

هو ذلك التوزيع الذي لا تكون دالة الإمكان مهيمنة عليه، ويكون له تأثير ملحوظ على التوزيع اللاحق. ويبدو من الواضح هنا بأنه إذا كان التوزيع الأولي مهيمناً على دالة الإمكان فهذا يعني أن القسم الأكبر من المعلومات تأتي من هذا التوزيع ولهذا يطلق على التوزيع الأولي في هذه الحالة بذي المعلومات (Informative).

١-٤- العائلات المتآلفة (Conjugate Families):

إذا كان كلٌ من التوزيع الأولي واللاحق ينتميان إلى نفس العائلة من التوزيعات فعندئذٍ يطلق على هذه التوزيعات بالتوزيعات المتآلفة. لمزيد من المعلومات أنظر [Lindgren, 1962]

➢ **مثال:**

على فرض أن $(y_1, y_2, \ldots y_n)$ عينة عشوائية من مجتمع له توزيع طبيعي بمتوسط θ وتباين معلوم σ^2، أي أن:

$$f(y_i|\theta,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \frac{1}{2\sigma^2}(y_i - \theta)^2 \qquad \begin{array}{l} i = 1, 2, \ldots, n \\ -\infty < y_i < \infty \\ -\infty < \theta < \infty \\ \sigma^2 > 0 \end{array}$$

$$\therefore L(\theta) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \theta)^2$$

$$\therefore L(\theta) \propto \exp - \frac{n}{2\sigma^2}(\theta - \bar{y})^2 \qquad , \bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$$

وافرض أيضاً أن θ لها توزيع أولي طبيعي بمتوسط θ_0 وتباين σ_0^2، أي أن:

$$P(\theta)=\frac{1}{\sqrt{2\pi\sigma_0^2}}\exp-\frac{1}{2\sigma_0^2}(\theta-\theta_0)^2$$

وعليه يمكن إيجاد التوزيع اللاحق كالآتي:

$$P(\theta\mid y_1,y_2,\ldots,y_n)\propto P(\theta)\,L(\theta)$$

$$\propto \exp-\frac{1}{2\sigma_0^2}(\theta-\theta_0)^2\times\exp-\frac{n}{2\sigma^2}(\theta-\bar{y})^2$$

وبفتح الأقواس وحذف الحدود التي لا تحتوي على θ نحصل على:

$$P(\theta\mid y_1,y_2,\ldots,y_n)\propto\exp\left[-\frac{\sigma^2+n\sigma_0^2}{2\sigma_0^2\sigma^2}\left(\theta-\frac{\theta_0\sigma^2+n\bar{y}\sigma_0^2}{n\sigma_0^2+\sigma^2}\right)^2\right]$$

والأخير يمثل نواة (Kernel)[1] التوزيع الطبيعي بمتوسط μ :

$$\mu=\frac{n\bar{y}\sigma_0^2+\theta_0\sigma^2}{n\sigma_0^2+\sigma^2}$$

وبقسمة البسط والمقام على $(\sigma^2\sigma_0^2)$ فإن:

$$\mu=\frac{(n\bar{y}/\sigma^2)+(\theta_0/\sigma_0^2)}{(n/\sigma^2)+(1/\sigma_0^2)}$$

وتباين التوزيع اللاحق ν:

$$\nu=\frac{\sigma_0^2\sigma^2}{\sigma^2+n\sigma_0^2}$$

وبذلك فان التوزيع اللاحق لمعدل التوزيع الطبيعي θ يمكن كتابته بالصيغة العامة للتوزيع الطبيعي:

$$P(\theta\mid y_1,y_2,\ldots,y_n)=\frac{1}{\sqrt{2\pi\nu}}\exp-\frac{1}{2\nu}(\theta-\mu)^2$$

إذن كما هو ملاحظ فإن التوزيع الطبيعي الأولي بمتوسط θ_0 وتباين معلوم σ_0^2 أعطى توزيعاً لاحقاً طبيعياً أيضاً، أي أن التوزيعان الأولي واللاحق هما من عائلة متآلفة.

[1] تعرف نواة التوزيع في استدلال بيز (Kernel) بأنها ذلك الجزء من دالة التوزيع الاحتمالي الذي يحتوي على المعلمة المدروسة.

▪ **ملاحظة:**

من الأفضل كتابة متوسط وتباين التوزيع اللاحق بدلالة الدقة (Precision)، حيث دقة العينة هي $\pi_s = n/\sigma^2$ ، ودقة التوزيع الأولي هي $\pi_{\text{O}} = 1/\sigma_{\text{O}}^2$ ، أي أن:

$$\mu = \frac{\pi_s \bar{y} + \pi_{\text{O}} \theta_{\text{O}}}{\pi_s + \pi_{\text{O}}} \qquad\qquad \nu = \frac{1}{\pi_s + \pi_{\text{O}}} = \frac{1}{\pi_p}$$

ويلاحظ أن متوسط التوزيع اللاحق هو المتوسط الموزون لمعدل العينة ومعدل التوزيع الأولي، حيث تكون الأوزان هي الدقة في كل حالة. ويلاحظ أيضاً أن الدقة في التوزيع اللاحق هي الدقة في العينة مضافاً إليها الدقة في التوزيع الأولي $\pi_p = \pi_s + \pi_{\text{O}}$.

والجدول (١-١) يوضح بعض الأمثلة عن العائلات المتآلفة.

الجدول (١-١)

بعض التوزيعات ذات العائلات المتآلفة. [Lindgren, 1962]

Likelihood Fun.	*Prior Dist.*	*Posterior Dist.*
$Bernoulli(\theta)$ $\theta^{\Sigma y_i}(1-\theta)^{n-\Sigma y_i}$	$Beta(\alpha,\beta)$ $\theta^{\alpha-1}(1-\theta)^{\beta-1}$	$Beta(\alpha+\Sigma y_i, \beta+n+\Sigma y_i)$ $\theta^{(\alpha+\Sigma y_i)-1}(1-\theta)^{(\beta+n+\Sigma y_i)-1}$
$Poisson(\theta)$ $e^{-n\theta}\theta^{\Sigma y_i}$	$Gamma(\alpha,\beta)$ $\theta^{\alpha-1}e^{-\beta\theta}$	$Gamma(\alpha+\Sigma y_i, \beta+n)$ $\theta^{(\alpha+\Sigma y_i)-1}e^{-(\beta+n)\theta}$
$Exponential(\theta)$ $\theta^n e^{-\theta\Sigma y_i}$	$Gamma(\alpha,\beta)$ $\theta^{\alpha-1}e^{-\beta\theta}$	$Gamma(\alpha+n, \beta+\Sigma y_i)$ $\theta^{(\alpha+n)-1}e^{-(\beta+\Sigma y_i)\theta}$
$N(\theta,\sigma^2)\,\sigma^2$ *Known* $e^{-\frac{n}{2\sigma^2}(\theta-\bar{y})^2}$	$N(\theta_{\text{O}},\sigma_{\text{O}}^2)$ $e^{-\frac{1}{2\sigma_{\text{O}}^2}(\theta-\theta_{\text{O}})^2}$	$N(\mu,\nu)$ $e^{-\frac{1}{2\nu}(\theta-\mu)^2}$

١-٥- تقدير بيز (Bayes Estimation):

إن الغاية من إيجاد التوزيع اللاحق في مبرهنة بيز هي الحصول على تقدير للمعلمة غني بالمعلومات المعبرة عن واقع حال المعلمة، سواءً كان ذلك التقدير بنقطة أم بفترة. وسوف يكون اهتمامنا في هذه الدراسة حول تقدير بيز بنقطة (Bayesian Point Estimation).

وهناك عدة طرائق لإيجاد مقدر بيز النقطي، حيث تعتمد هذه الطرائق في غالبيتها على إيجادها لمقدر يجعل دالة المخاطرة الشرطية (Conditional Risk Function) أقل ما يمكن، حيث أن دالة

المخـاطرة الشـرطية $\Re(\hat{\theta}\,|\,y_1,y_2,\ldots,y_n)$ ، هـي عبـارة عـن التوقـع الشـرطي لدالـة كلفـة الخطـأ ($C(\hat{\theta},\theta)$ Cost of Error Function)، أو دالة الخسارة ($l(\hat{\theta},\theta)$ Loss Function)، أي أن:

$$\Re(\hat{\theta}\,|\,y) = E[C(\hat{\theta},\theta)\,|\,y_1,y_2,\ldots,y_n]$$

$$= \int_{\theta\in\Omega} C(\hat{\theta},\theta)P(\theta\,|\,y_1,y_2,\ldots,y_n)d\theta$$

أي أن:

$$\hat{\theta}_B = \arg\min \Re(\hat{\theta}\,|\,y_1,y_2,\ldots,y_n)$$

$$= \arg\min\left[\int_{\theta\in\Omega} C(\hat{\theta},\theta)P(\theta\,|\,y_1,y_2,\ldots,y_n)d\theta\right]$$

وتعرف دالة المخاطرة الشرطية بدالة مخاطرة بيز Bayesian Risk Function.

ومن هذه الطرائق:

١-٥-١- تقدير بيز لتعظيم دالة التوزيع اللاحق: Maximum A Posteriori Estimation

هو المقدر الذي يعظم دالة التوزيع اللاحق $P(\theta\,|\,y_1,y_2,\ldots,y_n)$ ، ويرمز له بـ $\hat{\theta}_{MAP}$ ويمكن إيجـاده فقـط عنـدما تكـون دالـة الكلفـة منتظمـة (Uniform Cost Function) – قيمـة ثابتـة لفترات معينـة من قيم θ – [Vaseghi, 2000]، وتعرف:

$$C(\hat{\theta},\theta) = \begin{cases} 0 & if\ \left|\theta-\hat{\theta}\right| < \varepsilon \\ 1 & otherwise \end{cases}$$

حيث أن ε مقدار موجب صغير، إذن تكون دالة المخاطرة كالآتـي:

$$\Re(\hat{\theta}\,|\,y_1,y_2,\ldots,y_n) = \int_{\theta\in\Omega} P(\theta\,|\,y_1,y_2,\ldots,y_n)d\theta$$

وعند التعويض بدالة الخسارة بما يساويها نحصل على الآتي:

$$\Re(\hat{\theta}\,|\,y_1,y_2,\ldots,y_n) = \int_{\left|\theta-\hat{\theta}\right|>\varepsilon} P(\theta\,|\,y_1,y_2,\ldots,y_n)d\theta$$

$$= 1 - \int_{\hat{\theta}-\varepsilon}^{\hat{\theta}+\varepsilon} P(\theta\,|\,y_1,y_2,\ldots,y_n)d\theta$$

وبتطبيـق مبرهنـة القيمـة المتوسـطة (Mean Value Theorem) - الملحـق A- علـى التكامل الأخير نحصل على. [Melsa & Cohn, 1978] :

$$\Re(\hat{\theta} \mid y_1, y_2, \ldots, y_n) = 1 - [(\hat{\theta} + \varepsilon) - (\hat{\theta} - \varepsilon)] \times P(\hat{\theta} \mid y_1, y_2, \ldots, y_n)$$
$$= 1 - 2\varepsilon \times P(\hat{\theta} \mid y_1, y_2, \ldots, y_n)$$

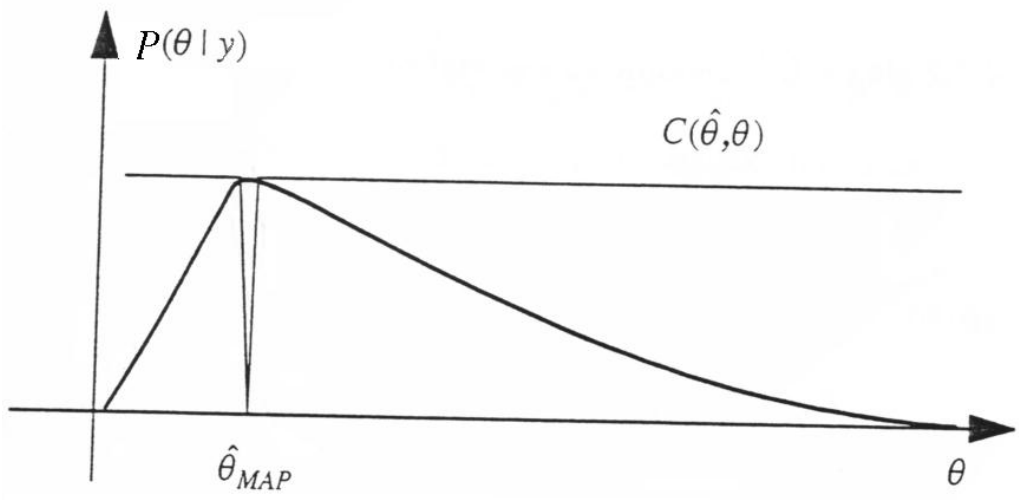
ولأن مقدر بيز بشكل عام يجعل دالة المخاطرة أقل ما يمكن- أنظر الشكل (١-٢) -، إذن ذلك يتطلب أن يكون التوزيع اللاحق أكبر ما يمكن، أي أن:

$$P(\hat{\theta}_B \mid y_1, y_2, \ldots, y_n) \geq P(\hat{\theta} \mid y_1, y_2, \ldots, y_n),\ \forall(\hat{\theta} \in \Omega)$$

حيث أن $\hat{\theta}_B \neq \hat{\theta}$، إذن يكون مقدر بيز لتعظيم التوزيع اللاحق:

$$\hat{\theta}_{MAP} = \arg\max P(\theta \mid y_1, y_2, \ldots, y_n)$$

معنى ذلك أن مقدر بيز لتعظيم التوزيع اللاحق هو منوال التوزيع اللاحق Mode of Posterior Distribution.



الشكل (١-٢)
دالة الكلفة المنتظمة ومنحنى التوزيع اللاحق. [Vaseghi, 2000]

١-٥-٢- تقدير الإمكان الأعظم (Maximum Likelihood Estimation):

وهو التقدير الذي يجعل من دالة الإمكان أعظم ما يمكن ويرمز له بـ $\hat{\theta}_{ML}$، وهذا المقدر لا يمكن إيجاده في استدلال بيز إلا في حالة كون دالة كلفة الخطأ منتظمة وكذلك يجب أن يكون التوزيع الأولي منتظماً أيضاً أو قليل المعلومات، فعندئذ يكون التوزيع اللاحق في تناسب مع دالة الإمكان، فنتمكن من إيجاد مقدر الإمكان الأعظم وذلك بالتعويض عن دالة الخسارة بما يساويها وعن التوزيع اللاحق بما يساويه من مبرهنة بيز ومن ثمّ تطبيق مبرهنة القيمة المتوسطة على التكامل:

$$\Re(\hat{\theta} \mid y_1, y_2, \ldots, y_n) = \int_{\theta\in\Omega} [C(\hat{\theta},\theta)] P(\theta \mid y_1, y_2, \ldots, y_n) d\theta$$
$$= Const. \times \int_{|\theta-\hat{\theta}|>\varepsilon} P(\theta) L(\theta) d\theta$$
$$= Const. \times [1 - 2\varepsilon \times L(\hat{\theta})]$$

من الواضح أن جعل دالة المخاطرة أقل ما يمكن يتطلب جعل دالة الإمكان أكبر ما يمكن:

$$\hat{\theta}_{ML} = \arg\max L(\theta)$$

إن الفرق الأساسي بين مقدر الإمكان الأعظم ومقدر بيز لتعظيم التوزيع اللاحق يكمن في أن مقدر الإمكان الأعظم يفترض أن يكون التوزيع الأولي منتظما- قيمة ثابتة على جميع قيم θ - أو قليل المعلومات. وجدير بالذكر أن مقدر الإمكان الأعظم في هذه الحالة هو نفسه مقدر الإمكان الأعظم في الاستدلال الكلاسيكي. [Vaseghi, 2000]

١-٥-٣- تقدير بيز ذو أقل متوسط مربع خطأ:

Minimum Mean Square Error Bayes Estimation:

يكون مقدر بيز في هذه الطريقة هو المقدر الذي يجعل توقع دالة كلفة مربع الخطأ أو دالة الخسارة التربيعية Quadratic Loss Function أقل ما يمكن، ويرمز له بـ $\hat{\theta}_{MMSE}$. حيث أن دالة كلفة مربع الخطأ هي:

$$C(\hat{\theta},\theta) = (\hat{\theta} - \theta)^2$$

إذن:

$$\Re(\hat{\theta} \mid y_1, y_2, \ldots, y_n) = E[(\hat{\theta} - \theta)^2 \mid y_1, y_2, \ldots, y_n]$$
$$= \int_{\theta\in\Omega} (\hat{\theta} - \theta)^2 P(\theta \mid y_1, y_2, \ldots, y_n) d\theta$$

ولإيجاد مقدر بيز الذي له أقل متوسط مربع خطأ (أو الأمثل Optimal) نشتق دالة المخاطرة الشرطية بالنسبة لـ $\hat{\theta}$ ونساوي المقدار بالصفر وكالآتي:

$$\frac{\partial \Re(\hat{\theta} \mid y_1, y_2, \ldots, y_n)}{\partial \hat{\theta}} = 2\hat{\theta} \int_{\theta\in\Omega} P(\theta \mid y_1, y_2, \ldots, y_n) d\theta - 2 \int_{\theta\in\Omega} \theta P(\theta \mid y_1, y_2, \ldots, y_n) d\theta$$
$$\therefore \hat{\theta}_{MMSE} = \int_{\theta\in\Omega} \theta P(\theta \mid y_1, y_2, \ldots, y_n) d\theta = E(\theta \mid y_1, y_2, \ldots, y_n)$$

وكما نلاحظ فإن تقدير بيز ذو أقل متوسط مربع خطأ هو نفسه معدل التوزيع اللاحق، ويعد هذا المقدر من أكثر مقدرات بيز استخداماً وذلك لسهولة إيجاده.

١-٥-٤- تقدير بيز ذو أقل قيمة متوسط خطأ مطلق:

Minimum Mean Absolute Value of Error Bayes Estimation:

وهو مقدر بيز الذي يجعل توقع دالة كلفة الخطأ المطلق (دالة المخاطرة) أقل ما يمكن، ويرمز له بـ $\hat{\theta}_{MAVE}$، وتعرف دالة كلفة الخطأ المطلق كالآتي:

$$C(\hat{\theta},\theta)=\left|\hat{\theta}-\theta\right|$$

$$\Re(\hat{\theta}\mid y_1,y_2,\ldots,y_n)=\int_{\theta\in\Omega}\left|\hat{\theta}-\theta\right|P(\theta\mid y_1,y_2,\ldots,y_n)d\theta$$

$$=\int_{-\infty}^{\hat{\theta}}(\hat{\theta}-\theta)P(\theta\mid y_1,y_2,\ldots,y_n)d\theta+\int_{\hat{\theta}}^{\infty}(\theta-\hat{\theta})P(\theta\mid y_1,y_2,\ldots,y_n)d\theta$$

وبأخذ المشتقة لدالة مخاطرة بيز بالنسبة لـ $\hat{\theta}$ ومساواة الناتج بالصفر نحصل على الآتي:

$$\frac{\partial\Re(\hat{\theta}\mid y_1,y_2,\ldots,y_n)}{\partial\hat{\theta}}=\int_{-\infty}^{\hat{\theta}}P(\theta\mid y_1,y_2,\ldots,y_n)d\theta-\int_{\hat{\theta}}^{\infty}P(\theta\mid y_1,y_2,\ldots,y_n)d\theta$$

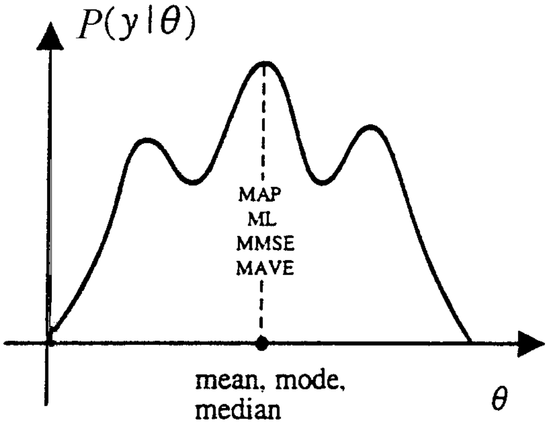
$$\int_{-\infty}^{\hat{\theta}}P(\theta\mid y_1,y_2,\ldots,y_n)d\theta=\int_{\hat{\theta}}^{\infty}P(\theta\mid y_1,y_2,\ldots,y_n)d\theta$$

إذن تكون قيمة $\hat{\theta}_{MAVE}$ هي القيمة التي تحقق المعادلة الأخيرة، وبما أن الوسيط هو الذي يحقق المعادلة أعلاه، فيكون مقدر بيز في هذه الحالة هو وسيط التوزيع اللاحق.

- **ملاحظة:**

عندما يكون التوزيع متماثل (Symmetric) ويكون التوزيع الأولي منتظماً أو قليل المعلومات فإن المقدرات الأربعة تتساوى -أنظر الشكل (١-٣)-، أي أن: $\hat{\theta}_{MAP}=\hat{\theta}_{ML}=\hat{\theta}_{MMSE}=\hat{\theta}_{MAVE}$

وللمزيد من المعلومات أنظر [Vaseghi, 2000 ; Melsa & Cohn, 1978].



الشكل (١-٣)

منحنى التوزيع المتماثل إضافة إلى الوسط والوسيط والمنوال. [Vaseghi, 2000]