تقدير الامكان الأعظم الحصين

د. حمزة اسماعيل شاهين * م. ميسون سالم مجيد **

المقدمة:-

يعد موضوع التقدير (Estimation) فن استدلال واستنتاج المعلومات حول بعض المعلمات غير المعروفة للبيانات الاصلية المتوفرة.وحيث أن التقديرات المستنتجة لابد وان تكون ذات كفاءة جيدة تحت شروط معينة مفروضة على تلك البيانات لذا فقد صممت الطرائق الإحصائية التقليدية لتقدير المعلمات المجهولة في حالة تحقق عدد من الافتراضات الصارمة ولعل أهمها هو توزيع المجتمع قيد الدراسة ومن ثم توزيع العينة التي تمثله سواء كان ذلك المجتمع بمتغير واحد (univariate) أو بعدة متغيرات (Multivariate) .

لكن ومن خلال التجارب والبحوث الكثيرة، توصل الباحثون الى أن هذه الطرائق تكون غير كفوءة في حالة عدم تحقق احد الافتراضات أو الشروط التي تعتمد عليها مثل ابتعاد بيانات العينة عن التوزيع المفترض وذلك اما بسبب توزيع المجتمع المدروس توزيعاً غير التوزيع المفترض أو بسبب وجود القيم الشاذة (Outliers) والتي تعرف بكونها مشاهدات تنحرف بشكل ملحوظ عن بقية المشاهدات وغالباً ما تنشأ من توزيعات ثقيلة الذيل (Heavy Tailed Distribution) من هنا لاحظ الباحثون الحاجة الى ايجاد طرائق ذات كفاءة وموضوعية أكثر لا تتأثر كثيراً بالانحراف

* استاذ مساعد/قسم الاحصاء/كلية الادارة والاقتصاد/الجامعة المستنصرية

** مدرس/قسم الاحصاء/كلية الادارة والاقتصاد/الجامعة المستنصرية

عن الافتراضات المحددة وذلك اما عن طريق تحوير الطرائق التقليدية وتسمى عندئذ بـ(Robustification) أو ايجاد أساليب حصينة بديلة وتسمى عندئذ بـ(Robust substitutes).

ففي عام 1979 وصف Mallows مصطلح الحصانة بأنه يتضمن مجموعة من الصفات المميزة والتي لابد وان تتميز بها الطريقة المستخدمة ومن هذه الصفات المقاومة (Resistance). التمهيد (Smothness) والاتساع (Breadth).

وفي عام 1981 وصف Huber مصطلح الحصانة بأنه يشير لعدم الحساسية للانحرافات الطفيفة عن الافتراضات الواجب تحققها.

وعليه وبصورة عامة تسمى الطريقة الإحصائية (Statistical method) بالطريقة الحصينة (Robust method) اذا كان سلوكها غير حساس نسبياً الى الانحرافات الناتجة عن الافتراضات التي تحقق تلك الطريقة.

ومما سبق ونظراً لكون أي طريقة إحصائية لابد وان تكون ذات كفاءة عالية الى النموذج المفترض لذا فالطريقة الحصينة ينبغي بها أن تكون ذات كفاءة قريبة من كفاءة طرائق التقدير التقليدية في حالة تحقق الافتراضات وأفضل منها في حالة الانحراف عن تلك الافتراضات وان تكون مناسبة لفئة واسعة من التوزيعات.

وعليه فأن هذا البحث يهدف بالدرجة الأساس الى أجراء تحوير لطريقة الامكان الأعظم التقليدية (Maximum Likelihood Method) وذلك من خلال ايجاد مجموعة جديدة من الدوال تعرف بدالة الامكان (Ψ) عن طريق ادخال ما يعرف بالدالة العامة (Ψ) (Generic function) أو ما يطلق عليها بالدالة التوظيفية وايجاد تقديرات حصينة وكفوءة سواء كان ذلك في حالة التوزيع

الطبيعي أو حالة وجود قيم شاذة وعند عدم تحقق الفرضيات المعتمـدة فـي التحليـل الإحصـائي ودراسة الخصائص التقاربية لتلك المقدرات.

2- الجانب النظري:- (2,3,4)

2-1:- طريقة الامكان (Ψ):- (Ψ) Method of Likelihood

بـــافتراض أن $\{f(x,\theta),\theta \in \Theta\}$ نمـــوذج إحصـــائي للمشـــاهدات المعطـــاة $\{y_i, i = 1,2,\dots,n\}$ حيث أن $f(y;\theta)$ هي دالة الكثافـة اعتمـاداً علـى اجـراء ناقـل معين.يرمز له v(dy) وان Θ تمثل فضاء أو مجال المعلمة (θ) .ان لوغـاريتم دالـة الامكـان الأعظم يمكن تعريفه على أساس حاصل جمع لوغاريتمات دوال الكثافة للمشاهدات y_i حيث أن:

$$L(\theta) = \frac{1}{n}\sum^{n} \ell(y_i,\theta)$$

على اعتبار أن $\ell(y,\theta) = \log f(y,\theta)$

وعلى افتراض أن Ψ عبارة عن دالة محدبة،متزايدة وقابلة للاشتقاق ومعرفة على المجـال R فأنه يمكن تعريف دالة الامكان - Ψ وفق الصيغة الاتية:-

$$L_\Psi(\theta) = \frac{1}{n}\sum_{i=1}^{n}\Psi(\ell(y_i,\theta)) - b_\Psi(\theta) \qquad \dots\dots\dots\dots(1)$$

للمشاهدات المعطاة {yi} حيث (i=1,2,...,n)

حيث أن:

$$b_\Psi(\theta) = \int \Psi^*(\ell(y,\theta))\nu(dy) \qquad \dots\dots\dots\dots(2)$$

وان:

$$\Psi^*(z) = \int_0^z \exp(s)\frac{\partial \Psi}{\partial s}(s)\,ds$$

وبالتالي فأن مقدر الامكان - Ψ الأعظم للمعلمة θ يمكن تعريفه كالاتي:

$$\hat{\theta}_{\Psi} = \arg \max_{\theta \in \Theta} \{\overline{L}_{\Psi}(\theta)\} \qquad \text{..........(3)}$$

واستناداً للمعادلة (3) يمكن توليد مجموعة من المقدرات للمعلمة θ بالاعتماد على مجاميع الدوال المحتملة لـ Ψ وسوف نشير في هذا البحث الى Ψ بالدالة العامة (Generic function) وأسلوب التقدير بطريقة الامكان - Ψ .

وان الحد $b_{\Psi}(\theta)$ في المعادلة (3) يلعب دوراً في تصحيح التحيز لمعادلة التقدير المناظرة والتي سيرد توضيحها لاحقاً.

ولابد من الإشارة هنا في حالة $\Psi(z) = z$ فأن طريقة الامكان- Ψ تتحول الى طريقة الامكان الاعتيادية بسبب أن $b_{\Psi}(\theta) = 1$ ولكن بحثنا هذا سوف يتناول دراسة طريقة الامكان- Ψ نحو حالة $\Psi(z) \neq z$.

وكما هو معروف أن طريقة الامكان الأعظم هي نظير لعينة تحقق الحد الأدنى لما يعرف بانحراف(Kullback-Leibler) عن التوزيع الحقيقي للنموذج الإحصائي.

هذا وان انحراف (Kullback-Leibler) يعرف من التكرار أو التردد f(y) الى g(y) وفق الصيغة الاتية:

$$KL(f, g) = \int g(y) \log \frac{g(y)}{f(y)} \nu(dy)$$

أي أن:

$$KL(f(.., \theta), g) = -E_g\{\ell(y, \theta)\}$$

وتحت حالة حذف الحدود الثابتة في θ فأن لوغاريتم دالة الامكان $L(\theta)$ يصبح مكافئ الى $E_g\{\ell(y,\theta)\}$ وإذا تم استبدال التوقع الرياضي بحالة تجريبية من خلال البيانات فأن نتيجة ذلك تصبح طريقة الامكان الأعظم مكافئة لأدنى انحراف (**Kullback-Leibler**).

ومن جانب اخر فأن دالة الامكان - Ψ يمكن أن ترتبط ايضاً بانحراف معين.

وسوف نفترض الاتي ما بين المتغيرين g,f حيث

$$\delta_\Psi(f,g)=\int_f^g \frac{g-\mu}{\mu}\psi(\log\ \mu)d\mu$$

فإذا تم اعتبار أن $\mu/\psi(\log\ \mu)$ هي بمثابة دالة تباين فأن $\delta_\Psi(f,g)$ سوف ينظر على أنها تمثل دالة تباين. علماً أن S_Ψ في دراستنا لا ينظر إليها بمثابة دالة للاستجابة أو المتوسط وإنما دالة للتكرارات (الترددات) ودالياً فأن انحراف - Ψ يمكن أن يعطى وفق الصيغة الاتية:

$$D_\Psi(f,g)=\int \delta_\Psi(f(y),g(y))\,\nu(dy)$$

$$=\left(\int g(y)\Psi(\log\ g(y))\,\nu(dy)-\int\Psi^{*}(\log\ g(y))\,\nu(dy)\right)-\left(\int g(y)\ \Psi(\log f(y))\,\nu(dy)-\int\Psi^{*}(\log f(y))\,\nu(dy)\right) \quad \text{........(4)}$$

وعلى أساس أن $D_\Psi(f,g)$ تكون ذات قيمة موجبة وتساوي صفراً اذا كانت f=g وبشكل خاص،فأنه اذا كانت $\Psi(z)=z$ فأن $D_\Psi(f,g)$ تتحول الى انحراف (-Kullback Leibler) والمعبر عنه بالرمز $KL(f,g)$ كذلك فأن المقدار الذي نحصل عليه من الصيغة (3) سوف يؤدي ايضاً الى تقليل النسخة التجريبية لانحراف- Ψ والمعبر عنه بالرمز $D_\Psi(f(.,\theta),g)$ بسبب انه عندما تحذف الحدود الثابتة في θ نحصل على الاتي:

$$D_\Psi(f(.,\theta),g)=-\{E_g\Psi(\ell(y,\theta))-b_\Psi(\theta)\} \quad \text{................(5)}$$

وتلك المعادلة تمثل دالة الامكان - Ψ بالصيغة (1) عندما يأخذ التوقع التجريبي لمجموعة البيانات $\{yi\}$ مع ملاحظة أن الانحراف - Ψ في الصيغة (5) يكون خطياً في g.

2-2:- الدالة العامة: Generic function

من أحدى الاختبارات الممكنة للدالة العامة هي دالة اللوغاريتم اللوجستية وحسب الصيغة الاتية:

$$\Psi_{\eta}(z) = \log \frac{1 + \exp(z + \eta)}{1 + \exp(\eta)}$$

وان :

$$\Psi_{\eta}^{*}(z) = \exp(z) + \exp(\eta) \log \frac{1}{\exp(z) + \exp(\eta)}$$

ووفقاً لذلك فأن دالة الوزن التي نحصل عليها ستكون:

$$\psi_{\eta}(z) = \frac{\partial}{\partial z}\Psi(z) = \frac{1}{1 + \exp(-z - \eta)}$$

وتستخدم هذه الدالة لمختلف الإغراض فمثلا أن معلمة التوافق η يمكن ان تشير الى قيمة الإشباع في نموذج (perceptron) متعدد الطبقة في الشبكة العصبية وان المعامل (z) يمكن ان يمثل معكوس درجة الحرارة.

ودالة الوزن ψ_{η} عندما η تقترب من (∞)تصبح مساوية الى دالة ثابتة وعند استخدام الدالة $\Psi_{\eta}(z)$ فأن المقدار الذي نحصل عليه وفق الصيغة (3) والانحراف وفق الصيغة (4) تعرف (بمقدر-η)،(وبانحراف-η)على التعاقب.

وهنالك صيغة أخرى للدالة العامة هي:

$$\Psi_{\beta}(z) = \frac{\exp(\beta z) - 1}{\beta}$$

وان

$$\Psi_{\beta}^{*}(z) = \frac{\exp((1-\beta)z)}{1-\beta}$$

حيث $\beta > 0$

والمقدر بالاعتماد على $\Psi_{\beta}(z)$ يعرف بمقدر β – ويتم الحصول عليه عند طريقة تعظيم دالة الهدف:-

$$\frac{1}{n}\sum_{i=1}^{n}\frac{f(y_i,\theta)^{\beta}}{\beta} - \int\frac{f(y,\theta)^{1+\beta}}{1+\beta}\,dv(y)$$

ودالة الانحراف وفق المعادلة (4) والتي تعرف بانحراف β – يمكن توضيحها كالاتي:

$$D_{\beta}=(g,f)=\int\frac{f(y)^{1+\beta}}{1+\beta}dv(y)-\int\frac{g(y)f(y)^{\beta}}{\beta}dv(y)+\int\frac{g(y)^{1+\beta}}{\beta(1+\beta)}dv(y) \qquad ....(6)$$

حيث أن g,f هما دوال الكثافة وفقاً الى الأجراء الناقل V .

وان الأسلوب المطبق في هذا البحث يقوم على استخدام طريقة تقدير M – الحصينةعندما يتم كتابة الاتي:

$$\rho(y,\theta) = \Psi(\ell(y,\theta) - b_{\Psi}(\theta)) \qquad ....(7)$$

فأن مقدر الامكان- Ψ سوف لا يمثل أي شيء ولكنه يرتبط بمقدر- M – بواسطة ρ .

وبالتالي فأن دالة الامكان حسب الصيغة (1) يمكن ان تعتبر انها كوصف جديد للدالة ρ معرفة على دالة الكثافة $f(y,\theta)$ وان توضيح هذه الفكرة سوف تعتمد في ذلك على الفهم الهندسي لمجال البيانات وعلى الوسط الحسابي المبتور وإحصاءات الرتب وغير ذلك.

مثال رقم (1):-

في حالة لدينا نموذج الذي يحتوي على المعلمة θ التي تمثل معلمة موقع (Location parameter) فأن الدالة $b_{\psi}(\theta)$ المعرفة في الصيغة (2) تصبح ثابتة في θ ومن تلك تكون الحالة مكافئة لتقدير- M – ولملاحظة ذلك تم تعريف الدالة العامة كالاتي:

$$\Psi_{H}(z) = \begin{cases} Z + \frac{1}{2}\log(2\pi). & if\ Z \geq \frac{1}{2}(-K^{2} - \log(2\pi)) \\ -K\{2Z - \log(2\pi)\}^{1/2} - \frac{1}{2}K^{2}. & otherwise \end{cases}$$

نتيجة لذلك فأن لوغاريتم دالة الامكان للنموذج الطبيعي بمتوسط θ وتباين 1 سوف يحقق الاتي:-

$$-\Psi_{H}(\ell(y,\theta)) = \begin{cases} \frac{1}{2}(y-\theta)^{2} & if \quad |y-\theta| \leq K \\ K|y-\theta| + \frac{1}{2}K^{2}, & otherwise \end{cases}$$

وهذا الشيء لا يكون له معنى ولكن دالة ρ Huber تكون لمعلمة الموقع.

والخطوة التالية هي افتراض دالة التقدير حيث أن دالة الامكان- Ψ المرتبطة بدالة التقدير والتي تحتاج الى حلها بالنسبة الى θ تكون كالاتي:

$$\frac{1}{n}\sum_{i}\psi(\ell(y,\theta))\,S(y,\theta) - \frac{\partial}{\partial\theta}b_{\psi}(\theta) = 0 \qquad ....(8)$$

حيث أن

$$\psi(z) = (\frac{\partial}{\partial z})\Psi(z)$$

وان

$$S(y,\theta) = (\frac{\partial}{\partial\theta})\ell(y,\theta)$$

مع ملاحظة أن

$$\frac{\partial}{\partial\theta}b_{\psi}(\theta) = \int\psi(\ell(y,\theta))\,S(y,\theta)\,f(y,\theta)$$

وبالتالي يمكن ملاحظة أن الصيغة (8) يمكن أن تكتب كالاتي:

$$\int \psi'(\ell(y,\theta))\, S(y,\theta)\, d(\overline{P}_n - P_\theta)(y) = 0 \qquad \text{..........(9)}$$

حيث أن $\overline{P}_n$ تمثل التوزيع التجريبي و p_θ النموذج الإحصائي المناظر الى $f(..\theta)\,dv$ وهذه هي دالة الامكان الموزونة التي ترتبط مع دالة الوزن ψ' والتي تعتمد على دالة النموذج الإحصائي.

ويمكن بسهولة أن نتوصل الى أن معادلة التقدير (8) تحقق شرط عدم التحيز.

أن دور الدالة العامة Ψ تكون في مشتقتها التي نحصل منها على الدالة ψ' التي تعطي وزناً لكل مشاهدة بحيث أن المشاهدات الأقل وزنا تبتعد عن النموذج المفترض.وهذا ينجز عن طريق عدم تناقص الدالة $\psi'(.)$ والناتج من تحدب الدالة $\Psi(.)$.

ويمكن ملاحظة أن درجة قياس الشواذ لاتتم الإحصاءات التقليدية والتي منها مقياس (Maharanobis) للحالة وانما بمقدار اعتماد لوغاريتم الامكان على النموذج الإحصائي المفترض $f(y,\theta)$.

2-3 خوارزمية الاوزان التكرارية:- Iterativeiy reweighted algorithn

تم افتراض خوارزمية الحصول على مقدر - Ψ والمشار اليه بالرمز $\hat{\theta}_\Psi$ و كالاتي:-

$$\hat{\theta}^{(k+1)} = \arg\max \left\{ \tfrac{1}{n}\sum_{i=1}^{n} \psi'(\ell(y_i,\theta^{(k)}))\, \ell(y_i,\theta) - b_\Psi(\theta) \right\} \qquad \text{...........(10)}$$

حيث K=0,1,.........

ومن الحساب من قيمة مختارة الى θ_0 وان معادلة التقدير للمعلمة $\theta^{(K+1)}$ هي:

$$\frac{1}{n}\sum_{i=1}^{n}\psi\left(\ell(y_i,\theta^{(K)})\right)S(y_i,\theta)=E_\theta\left\{\psi\left(\ell(y,\theta)\right)S(y,\theta)\right\}$$

وهذا يتم من منظور الدالة الموزونة الوهمية وان الوزن $\psi\left(\ell(y,\theta^{(K)})\right)$ يتجدد في كل خطوة وهذا الشيء يمكن أن ينظر اليه على أساس أنه يمثل خوارزمية الاوزان التكرارية وسوف يتم اثبات أن الخوارزمية تحسن من دالة الامكان في كل خطوة مع ملاحظة أن $\Psi(\ell)$ هي محدبة في ℓ .

<u>نظرية (1):-</u>

لمتسلسلة $\{\theta^{(K)};K=0,1,.........\ ..\}$ المعرفة بالصيغة (10) يكون:

$$\overline{L}_\psi(\theta^{(K+1)})\geq\overline{L}_\psi(\theta^{(K)})......... \quad ...(11)$$

البرهان:-

ينتج من افتراض خاصية التحدب للدالة Ψ أن :-

$$\rho(y,\theta)-\rho(y,\theta^{(\cdot)})\geq\psi\left(\ell(y,\theta^{(\cdot)})\right)\left\{\ell(y,\theta)-\ell(y,\theta^{(\cdot)})\right\}-b_\psi(\theta)+b_\psi(\theta^{(\cdot)})$$

حيث أن $\rho(y,\theta)$ كما هو معرف في الصيغة (7) وبالتالي فأن:

$$L_\psi(\theta)-L_\psi(\theta^{(\cdot)})\geq\overline{\tau}(\theta,\theta^{(\cdot)})-\overline{\tau}(\theta^{(\cdot)},\theta^{(\cdot)})$$

حيث أن :

$$\overline{\tau}(\theta,\theta^{(\cdot)})=\sum_{i=1}^{n}\psi\left(\ell(y_i,\theta^{(\cdot)})\right)\ell(y_i,\theta)-b_\psi(\theta)$$

ومن تعريف الصيغة (10) فأن

$$\tau(\theta^{\ \ },\theta^{(\cdot)})\geq\tau(\theta^{(\cdot)},\theta^{(\cdot)})$$

والمتضمن في الصيغة (11) حالة K=1 ويمكن برهنة الحالة العامة K بنفس الطريقة يوضح

$$\theta^{(0)}=\theta^{(K)}$$

مثال (2):-

افترض لدينا حالة التوزيع الطبيعي بمتوسط μ وتباين σ^2. عندئذٍ معادلة التقدير تقودنا الى الاتي:

$$\mu = \frac{\sum \psi_i \, y_i}{\sum \psi_i} \qquad , \qquad \sigma^2 = \frac{\sum \psi_i (y_i - \mu)^2}{\sum \psi_i + c(\sigma^2)} \quad \ldots\ldots\ldots\ldots(12)$$

حيث أن:

$$\psi_i = \psi(\ell(y_i, \mu, \sigma^2))$$

$$C(\sigma^2) = \frac{1}{\sqrt{2\pi}} \int \psi\left(-\frac{z^2}{2} - \frac{1}{2}\log \sigma^2\right)(z^2 - 1)\, e^{-\frac{z^2}{2}}\, dz$$

وهنا خوارزميتنا يمكن تعريفها على أساس التجديد من الطرف الأيسر للصيغة (12) الى الطرف الأيمن.

4-2 التقاربية:- Asympototics

1-4-2 الاتساق:- Consistency

على افتراض أن $\overline{P}_n$ تمثل التوزيع التجريبي من البيانات عندئذٍ مقدر - Ψ يمكن كتابته داليا على أساس دالة التوزيع التجريبي ويعبر عنه بـ $\hat{\theta}_{\psi}(\overline{p}_n)$ وبالتالي فأن معادلة التقدير (9) سوف تقودنا الى ما يعرف باتساق-فشر (Fisher-Consistency)،أي أن:

$$\hat{\theta}_{\psi}(p_{\theta}) = \theta \ldots\ldots\ldots\ldots(13)$$

حيث أن P_{θ} تعبر عن التوزيع الاحتمالي المرتبطة بدالة الكثافة $f(y, \theta)$.

أن اشتقاق معادلة التقدير في الصيغة (9) يمكن توضيحه كالاتي:

$$J(\theta) = -E_{P\theta}\{\psi(\ell(\theta))\, S(y, \theta)\, S(y, \theta)^T\}$$

وعلى اعتبار أن $\overline{P_n} = P_\theta$ حيث تكون تلك محـددة سـالبة **(negative-definite)**، وذلـك لان ψ تكون دالة موجبة ناتجة من افتراض حالة تحدب الدالة Ψ. وباستخدام الصيغة **(13)** ينتج الاتي:

$$\hat{\theta}_\Psi(\overline{p}_n) - \theta \simeq J(\theta)^{-1} \int \psi(\ell(y,\theta))\, S(y,\theta)\, d(\overline{p}_n - p_\theta)(y)$$

وان نتيجة ذلك الحصول على اتساق قوي أو يكافىء اقتراب المقدر $\hat{\theta}_\Psi(\overline{p}_n)$ بشكل مؤكد مـن المعلمة θ.

من جهة ثانية يمكن اثبات الاتساق باستخدام خاصية التعظيم. وذلك بافتراض أن:

$$Q(\theta^*,\theta) = E\, \Psi(\ell(y,\theta)) - b_\Psi(\theta)$$

وبالتالي نحصل على الاتي:

$$Q(\theta^*,\theta) = \int \left[\exp\{\ell(y,\theta^*)\}\Psi(\ell(y,\theta)) - \int e^s \psi(s)\, ds \right] d\nu(y)$$

وبموجب تعريف الدالة Ψ^* فسوف نحصل على بعض النتائج وكالاتي:

$$Q(\theta^*,\theta^*) - Q(\theta^*,\theta) = D_\Psi(f(.,\theta^*) f(.,\theta)) \geq 0$$

وحالة التساوي تحقق اذا وفقط اذا كان :

$$f(y,\theta^*) = f(y,\theta)$$

وعلى الأكثر لجميع حالات y ووفقاً الى:

$$\theta^* = \arg \max_{\theta \in \Theta} Q(\theta^*,\theta)$$

فأن هذا يقود الى اتساق قوي لمقدر - Ψ إذ أن دالة الامكان - Ψ وفق الصـيغة (1) سـوف تقترب على الأكثر وبشكل اكبر الى $Q(\theta^*,\theta)$.

2-4-2 دالة التأثير:- Influnce function

سوف نبحث في هذه الفقرة دالة التأثير للدالة $\hat{\theta}_{\psi}(\overline{p}_n)$ إذ أن دالة التأثير يمكن تعريفها كالاتي:

$$IF(\hat{\theta}_{\psi}, y) = \frac{\partial \hat{\theta}_{\psi}(p_{\in})}{\partial \in} |_{\in=0}$$

حيث أن:

$$p_{\in} = (1-\in)p_{\theta} + \in\delta_{y}$$

إذ أن δ_{y} عدد المشاهدات الملوثة، $\in$ نسبة التلويث.

وبتفاضل دالة التقدير:-

$$\int \{\psi(\ell(z,\theta))\ S(z,\theta)\}(P(z) - P_{\theta}(z)) = 0$$

بالنسبة إلى $\in$ نحصل على الاتي:

$$IF(\hat{\theta}_{\psi}, y) = J_{\psi}(\theta)^{-1}\left\{\psi(\ell(y,\theta))\ S(y,\theta) - \frac{\partial}{\partial\theta}b_{\psi}(\theta)\right\} \qquad \text{..........}\ ..(14)$$

حيث أن:

$$J_{\psi}(\theta) = E_{p_{\theta}}\{\psi(\ell(y,\theta))\ S(y,\theta)\ S(y,\theta)'\}$$

ومن الطبيعي في حالة y قيمتها صفر أو قريبة من الصفر أن تفترض انها تمثل قيمة شاذة أو تكافيء حالة افتراض $z = \ell(y,\theta)$ فاذا كانت قيمة Z سالبة وكبيرة فأن y يشار إليها انها تمثل قيمة شاذة ومن هذه المشاهدة يكون من المهم البحث عن سلوك الدالة $\psi(Z)$ عندما Z تقترب من $(-\infty)$ وان تضمن تحقق الشرطين الآتيين:

$$\lim_{Z\to-\infty} \psi(Z) = 0 \qquad \text{..........}\quad \text{..........}\quad ...(15)$$

وان:

$$\lim_{Z\to-\infty} Z\,\psi(Z) = 0 \qquad \text{..........}\quad ....(16)$$

ويمكن ملاحظة أن الانحراف- β وانحراف- η سوف تحققان الشــرطين (15) و(16) . لأي قيمة $0 < \beta < 1 \ , \ \eta \geq 0$.

2-4-3 متباينة المعلومات:- Information inequality

كما معروف أن مقدرات طريقة الامكان الأعظم تتصف بخاصية الكفــاءة (efficiency) التقاربيــة وهذا يخص مجموعة واسعة من النماذج الإحصائية التي تحقق الشروط التمهيدية. فعلى افتــراض أن المشاهدات تعود للتوزيع المفترض $f(y,\theta)$ عندئذ مقدر الامكان الأعظم للمعلمة θ والذي يرمز له بـــ $\hat{\theta}$ سوف يكون $(\sqrt{n}-Consistent)$ للمعلمة θ وان توزيـــع $\sqrt{n}(\hat{\theta}-\theta)$ يقترب من التوزيع الطبيعي $N(0,I(\theta)^{-1})$ حيث أن $I(\theta)$ تســمى بمصــفوفة معلومــات Fisher إذ أن:

$$I(\theta) = E_{f(.,\theta)}\{S(y,\theta)\,S(y,\theta)'\}$$

وان مقدر دالة الامكان- Ψ الأعظم يكون له نفس نتيجة التقارب أو له نفس توزيع الغاية الــى $\sqrt{n}(\hat{\theta}_{\Psi}-\theta)$ إذ أن:

$$\sqrt{n}(\hat{\theta}_{\Psi}-\theta) \sim N(0,K_{\Psi}(\theta)^{-1})$$

حيث أن :

$$K_{\Psi}(\theta) = J_{\Psi}(\theta)^{-1}\,H_{\Psi}(\theta)\,J_{\Psi}(\theta)^{-1}$$

وحسب الصيغة (14) فأنه:

$$H_{\Psi}(\theta) = E_{f(.,\theta)}\{\psi'(\ell(y,\theta))^{2}\,S(y,\theta)\,S(y,\theta)'\}$$

ووفقاً إلى متباينة (Cauch- Yschwartz) نحصل على:

$$J_{\Psi}(\theta)^{-1} H_{\Psi}(\theta) J_{\Psi}(\theta)^{-1} \geq I(\theta)$$

وبشكل عام فأن جميع مقدرات الامكان- Ψ في حالة $\Psi(z) \neq z$ تكون غير كفـوءة تحـت حالة كون المشاهدات المعطاة تماماً تعود للنموذج الإحصائي.

وعموماً فأن خاصية الكفاءة غالبا ما تنهار عندما يكون النموذج المفترض غير محدد بشكل صحيح أو نتيجة وجوده القيم الشاذة في مجموعة البيانات.

2-4-5 اختبار الفرضيات:- Testing ahypothesis

لاختبار الفرضية:

$$H_{\circ} : \theta = \theta_{\circ}$$

ضد الفرضية:

$$H_{1} : \theta \neq \theta_{\circ}$$

يمكن افتراض احصاءة الاختبار وفق الطريقة الامكان- Ψ وكالاتي:

$$t_{\Psi}(\theta_{\circ}) = (\hat{\theta}_{\Psi} - \theta_{\circ})' K_{\Psi}(\theta_{\circ})^{-1} (\hat{\theta}_{\Psi} - \theta_{\circ})$$

حيث $K_{\Psi}(\theta)$ تم تعريفها نحو الفقرة (2-4-3) وان الاحصاءة $t_{\Psi}(\theta_{\circ})$ يمكـن أن ينظـر اليها على اعتبار انها نوع من إحصاء (wald) ويمكن أعطاء نوع من احصاءة الدرجات كالاتي:

$$S_{\Psi}(\theta_{\circ}) = n\, \bar{S}_{\Psi}(\theta_{\circ})' \, G_{\Psi}(\theta_{\circ})^{-1} \, \bar{S}_{\Psi}(\theta_{\circ})$$

حيث أن:

$$\bar{S}_{\Psi}(\theta) = \frac{1}{n} \sum_{i=1}^{n} \psi(\ell(y_i, \theta))\, S(y_i, \theta) - \frac{\partial}{\partial \theta} b_{\Psi}(\theta)$$

وان:

$$G_{\psi}(\theta) = Var_{f(\cdot,\theta)}\{\psi(\ell(y,\theta))\ S(y,\theta)\}$$

ويتضح أن كل من احصاءة الاختبار (θ) t_{ψ} و $S_{\psi}(\theta_{\circ})$ يتبعان بشكل تقاربي توزيع (مربع كاي) بدرجة حرية تساوي (q) إذ أن (q) تمثل إبعاد المعلمة (θ) علماً أن هذه الطريقة يمكن أن تطبيقها بسهولة لاختبار الفرضية المركبة التي تخص فرضية العدم $H_{\circ}$.

3- الجانب التطبيقي:- Applications

سوف يتناول الجانب التطبيقي بحث مسئلة الحصانة في مجال تحليل متعدد المتغيرات (Multivariate Analysis) وكما معروف أن الجانب النظري للتحليل متعدد المتغيرات مبني بالأساس على حالة افتراض التوزيع الطبيعي متعدد المتغيرات (Multivariate Normal distribution) لذلك فأن دراسة الحصانة في هذا المجال سوف يساعد على بناء استدلال إحصائي لمتعدد المتغيرات يكون حصين للحالات التي يكون فيها المجتمعات لأتتبع التوزيع الطبيعي أو حالة أن البيانات تحتوي على قيم شاذة. وذلك من خلال استخدام طريقة الامكان- Ψ في تقدير موجه المتوسط (mean vector) ومصفوفة التباين والتباين المشترك Σ (Variance- Covariance matrix).

بأتباع الأسلوب الذي تم بحثه في هذه الدراسة.

تقدير موجه المتوسط ومصفوفة التباين المشترك:

بافتراض أن $\{\underline{y}_j \ / \ j = 1,......,\ n\}$ عبارة عن P من المتجهات تضم (n) من المشاهدات وعلى افتراض أن تلك المشاهدات تتبع التوزيع الطبيعي متعدد المتغيرات بمتوسط $\underline{\mu}$ ومصفوفة

التباين –التباين المشترك $\sum$ أي $N_p(\mu, \Sigma)$ وتحت حالة وجود بعض القيم الشاذة في البيانات. معادلة التقدير وفق طريقة الامكان- Ψ يمكن توضيحها كالاتي:

$$\frac{1}{n}\sum^{n}\psi_i\cdot(\underline{y}_i-\underline{\mu})=0 \quad \left(=E\left[\psi\cdot(\underline{y}-\underline{\mu})\right]\right) \qquad ....(17)$$

$$\frac{1}{n}\sum_{1}^{n}\psi_i\cdot\left((\underline{y}_i-\underline{\mu})(\underline{y}_i-\underline{\mu})'-\Sigma\right)=h_\psi(\Sigma)\cdot\Sigma\left(=E\left[\psi\left((\underline{y}-\underline{\mu})(\underline{y}-\underline{\mu})'-\Sigma\right)\right]\right) \qquad ...(18)$$

إذ أن

$$h_\psi(\Sigma)=\frac{1}{p(2\pi)^{p/2}}\int\psi\left(-\frac{1}{2}\log|\Sigma|-\frac{1}{2}\underline{y}'\underline{y}\right)(\underline{y}'\underline{y}-p)\,e^{-\underline{y}'\underline{y}/2}\,dy$$

$$\psi_i=\psi\left(-\frac{1}{2}\log|\Sigma|-\frac{1}{2}(\underline{y}_i-\underline{\mu})'\Sigma^{-1}(\underline{y}_i-\underline{\mu})\right) \qquad ....(19)$$

إذ أن Ψ المعرفة بـــ ψ_i بالنسبة إلى y تكون الى $\cdot\psi$

ويجب ملاحظة أن Σ^{-1} يمكن الاستغناء عنها لطرفي المعادلتين (17) و(18).

فعندما تكون

$$\psi_i(z)=e^{\beta z}$$

فأن

$$h_\psi(\Sigma)=-\frac{\beta|\Sigma|^{-\beta/2}}{(1+\beta)^{1+p/2}}$$

وبالتالي يمكن تبسيط المعادلة (18) كالاتي:

$$\frac{1}{n}\sum^{n}e^{\beta z_i}\left((y_i-\mu)(y_i-\mu)'-\Sigma\right)=-\frac{\beta}{(\beta+1)^{p/2+1}}\Sigma \qquad ....(20)$$

وان

$$Z_i = -\frac{1}{2}(\underline{y}_i - \underline{\mu})' \Sigma^{-1}(\underline{y}_i - \underline{\mu}) \qquad \text{.......... ....(21)}$$

وباستخدام الخوارزمية الموضحة في الفقرة(2– 3) تم الحصول على التقديرات $\hat{\mu}^{(K)}$, $\hat{\Sigma}^{(K)}$ والى (K) من التكرارات الحسابية.

وان تجديد هذه التكرارات يمكن كتابته كالاتي:

$$\begin{cases} \hat{\mu}^{(K+1)} = \dfrac{\sum_{i=1}^{n} e^{\beta \hat{Z}_i^{(K)}} \, y_i}{\sum_{i=1}^{n} e^{\beta \hat{Z}_i^{(K)}}} \\ \hat{\Sigma}^{(K+1)} = \dfrac{\frac{1}{n}\sum_{i=1}^{n} e^{\beta \hat{Z}_i^{(K)}} (\underline{y}_i - \hat{\mu}^{(K+1)})(\underline{y}_i - \hat{\underline{\mu}}^{(K+1)})'}{\frac{1}{n}\sum_{i=1}^{n} e^{\beta \hat{Z}_i^{(K)}} - \beta/(\beta+1)^{p/2+1}} \end{cases}$$

إذ أن:

$$\hat{\underline{Z}}_i^{(K)} = -\frac{1}{2}(\underline{y}_i - \hat{\mu}^{(K)})'(\hat{\Sigma}^{(K)})^{-1}(\underline{y}_i - \hat{\underline{\mu}}^{(K)})$$

أن طريقة الامكان– Ψ بالحقيقة تكون مرتبطة بشكل كبير جداً من الاستدلال الإحصائي المستند على ما يعرف بالتوزيعات الاهليليجية (elliptical distributions) المناسبة للتوزيعات ثقيلة الذيل بدلاً من حالة التوزيع الطبيعي.

فعلى افتراض أن المشاهدات تعود لمجتمع يتبع التوزيع $|\Sigma|^{-1/2} h(-\frac{1}{2}(\underline{y} - \underline{\mu})'\Sigma^{-1}(\underline{y} - \underline{\mu}))$ إذ أن $h(\cdot)$ دالة موجبة غير مرتبطة بالتوزيع $N(\underline{\mu}, \Sigma)$.

عندئذٍ يمكن اشتقاق دوال الامكان كالاتي:

$$\begin{cases} \frac{1}{n}\sum_{i=1}^{n} W_i(\underline{y}_i - \underline{\mu}) = 0 \\ \\ \frac{1}{n}\sum_{i=1}^{n} W_i(\underline{y}_i - \underline{\mu})(\underline{y}_i - \underline{\mu})' = \Sigma \end{cases} \qquad \text{.......... ...}(22)$$

حيث أن:

$$W_i = h'(\underline{Z}_i)/h(\underline{Z}_i)$$

وان Z_i كما معرف في الصيغة (21) ويمكن ملاحظة أن الدالة $h(\cdot)$ في أسلوب الاهليليجية تكون مناظرة للدالة $\psi(\cdot)$ في طريقة الامكان- Ψ .

وبشكل عام في الأسلوب الاهليليجية وطريقة التقدير- M - فأن $\hat{\Sigma}$ لاتمثل تقدير لمصفوفة التباين المشترك للمجتمع وانما تقترب من مصفوفة التباين المشترك مضروبة بعدد ثابت يتبع التوزيع الطبيعي.

وخلافاً لما جاء فأن طريقة الامكان- Ψ تقوم بعملية تقدير لمصفوفة التباين المشترك Σ اعتماداً على الحد $b_\Psi(\Sigma)$ وان الفرق الوحيد ما بين هذه الطرق أن طريقة الامكان- Ψ تقوم بالأصل على نقاط تعود لعينة وليس على مسافة (مهانولويس) وبمقارنة المعادلة (19) مع المعادلة (21) فأنه يمكن ملاحظة الفرق بينهما وبغض النظر عن الحد $\log|\Sigma|$ المتضمن في المعادلتين ويمكننا تقدير تأثير $\log|\Sigma|$ وذلك عن طريق ادخال معلمة التحويل η بحيث أن $\psi(z) = \psi^*(z-\eta)$ تصبح دالة الى لوغاريتم.

ونتيجة لذلك فأنه يمكن أن نشير الى أن طريقة الامكان- Ψ تغطي الاستدلال الإحصائي لأسلوب الاهليليجي. وبالتالي فأن أسلوب التقدير لموجه المتوسط ومصفوفة التباين المشترك يقوم على تحويل كل مشاهدة y_j إلى $Ay_i + b$ وبالتالي فأن التقديرات $(\hat{\mu}, \hat{\Sigma})$ سوف يتم تحويلها

الى $(A\hat{\mu} + b \,.\, A \hat{\Sigma} A')$.ووفقاً لهذا التحويل فأن المعادلة الاتية يجب أن تحقق لأي مصفوفة تحويل A .

$$\hat{\mu} = \sum_{i=1}^{n} \left(\frac{\psi_i(-\log |A| - \hat{\ell}_i)}{\sum_{i=1}^{n} \psi(-\log |A| - \hat{\ell}_i)} \right) y_i$$

إذ أن :

$$\hat{\ell}_i = -\frac{1}{2}\log \left|\hat{\Sigma}\right| - \frac{1}{2}(\underline{y}_i - \hat{\mu})' \hat{\Sigma}^{-1} (\underline{y}_i - \hat{\mu})$$

4ـ الاستنتاجات والتوصيات:-

يتضح من هذه الدراسة أن تطبيق طريقة الامكان التقليدية لتقدير المعلمات المجهولة يتم في حالة تحقق عدد من الافتراضات ولعل أهمها هو توزيع المجتمع قيد الدراسة ومن ثم توزيع العينة التي تمثله وبالحقيقة أن تحقق مثل هذه الحالة عملياً يكون صعب جداً حيث غالباً ما تبتعد بيانات العينة عن التوزيع المفترض بسبب وجود القيم الشاذة. ولذلك فقد تم اجراء تحوير لطريقة الامكان التقليدية وذلك بإدخال دالة التحويل- Ψ إلى لوغاريتم دالة الكثافة $f(y,\theta)$ وقد ساعد على ذلك تحقق بعض الشروط الأساسية في عملية اشتقاق دالة الوزن ψ وتم الحصول على سلسلة من المقدرات الحصينة التي تتميز بكفاءة قريبة من كفاءة طريقة الامكان التقليدية ونوصي بتوسيع تطبيق دالة الامكان Ψ التحويرية لنماذج إحصائية متعددة والحصول على استدلال إحصائي رصين.

المصادر:-

1- Ali.S.M and Silvey (1966)

" A general class of coefficient of dirergence of one distribution from another"

J.Royal statistic.So B 28, 131-142.

2- Campbell, N.A (1980)

" Robust Procedure in multivariate analysis and robust covariance estimation.

Applied statistics 29, 231-237.

3- Huber, P.J (1981)

" Robust statistics" New Yorks;wiley.

4-Mallows, C.L (1979)

" Robust methods; some examples of their use" Amer.statistic 33, 179-184.

5- Shinto, E and Yutaka ,K (2002)

"Robustifing maximum like lihood estimation"

Email; eguchi @ ism.ac. jp

....