

استخدام اسلوب بيز في حساب فترات الثقة

أ. سليم اسماعيل الغرابي/ قسم الاحصاء أ. صباح هادي عبود/ قسم الاحصاء
كلية الادارة والاقتصاد /جامعة بغداد

1. المقدمة:

ان معظم مواضيع الاستدلال الاحصائي تهتم بموضوع تقدير المعالم لتوزيعات احتمالية معلومة الصيغ $f(x|\theta)$ ، حيث ان x هي القيمة المشاهدة للمتغير العشوائي θ, X تمثل المعلمة او متجه المعالم غير المعروفة. ولغرض تقدير المعلمة θ بالطرق المعروفة التي لا تنتمي الى مدرسة بيز. يفترض الاحصائيون ان θ هي كمية ثابتة غير معروفة عندئذ تكون المعلومات المتوفرة عن العينة المسحوبة من المجتمع $X_1, X_2, \dots, X_n$ وبحجم مناسب ($n > 30$) هي كافية لغرض اجراء عملية التقدير. ولكن الذي يحدث في اغلب التطبيقات العملية ان θ في الحقيقة متغيراً عشوائياً وليست كمية ثابتة لذلك ولغرض تقديرها لابد من توفر معلومات اولية عنها بصيغة توزيع احتمالي والذي عادة يقترح من قبل الاحصائي ويسمى بالتوزيع المسبق (Prior Distribution) وعلى هذا الاساس فان اسلوب بيز في التقدير عندما تكون θ متغيراً عشوائياً يكون ملائماً لتقدير θ خلاف الطرق الاخرى في التقدير كطريقة الامكان الاعظم وذلك لان هذا الاسلوب يستخدم المعلومات المتوفرة عن طريق θ او المعلومات المتوفرة عن θ والعينة المسحوبة من المجتمع وذلك بصيغة توزيع احتمالي يسمى التوزيع اللاحق (Prior Distribution) θ وعلى هذا الاساس فان اسلوب بيز في التقدير كونه يستخدم المعلومات المتوفرة عن θ ، هو بدون شك يقود الى نتائج اكثر دقة وواقعية لتخمين θ في مواضيع الاستدلال الاحصائي سواء كانت اختبار فرضيات او تقديرات.

ان اسلوب بيز في التقدير هو مقدر لـ θ يجعل المنفعة المتوقعة اعظم ما يمكن وهو دالة في المتغير x كما انه بالضبط التوقع الشرطي للتوزيع اللاحق في حالة كون دالة المنفعة هي دالة مربع الخطأ علماً بان اسلوب بيز في التقدير هو متحيز ومقبول واكثر كفاءة من أي مقدر اخر نحصل عليه باستخدام طرق التقدير الاخرى التي لا تنتمي الى مدرسة بيز [3].

يهدف هذا البحث الى ايجاد فترة ثقة بيز لـ θ على فرض ان المتغير العشوائي X يتوزع طبيعياً بوسط θ وتباين معلوم σ^2 والتوزيع الاول المقترح لـ θ هو ايضاً التوزيع الطبيعي بوسط وتباين معلوم μ, J^2 على الترتيب.

2. التوزيع اللاحق:

لتكن $X_1, X_2, \dots, X_n$ عينة عشوائية من مجتمع تتوزع مفرداته وفقاً لدالة الكثافة الاحتمالية الشرطية $f(x|\theta)$ والاحصاء Y هي دالة في المتغيرات $X_1, X_2, \dots, X_n$. بفرض ان دالة الكثافة الاحتمالية الشرطية لـ Y هي $g(Y|\theta)$ فان دالة الكثافة الاحتمالية المشتركة لـ θ, Y هي:

$$k(\theta, Y) = h(\theta)g(Y|\theta) \quad \dots\dots\dots(1)$$

حيث $h(\theta)$ هو التوزيع الاول لـ $\theta \in R$, θ علماً ان الحرف R يرمز لمجموعة الاعداد الحقيقية. ان دالة الكثافة الاحتمالية الهامشية لـ Y في هذه الحالة هي الاتية:

$$k_1(Y) = \int_{-\infty}^{\infty} h(\theta) g(Y|\theta) d\theta \quad \dots\dots\dots(2)$$

من (1)، (2) وباستخدام قاعدة بيز العكسية نجد ان دالة الكثافة الاحتمالية الشرطية لـ θ

$$\begin{aligned} k_2(\theta | Y = y) &= \frac{k(\theta | Y)}{k_1(Y)} \\ &= \frac{g(Y|\theta)h(\theta)}{\int_{-\infty}^{\infty} g(Y|\theta)h(\theta)d\theta} \quad \dots\dots\dots(3) \end{aligned}$$

ان (3) يسمى التوزيع اللاحق لـ θ .

افرض ان القيم التجريبية المشاهدة لـ $X_1, X_2, \dots, X_n$ هي $Y_1, Y_2, \dots, Y_n$ التي تعطي القيمة التجريبية y للمتغير العشوائي Y فان الفكرة الاساسية لهذا البحث تكمن في ايجاد دالتين $w_2(y), w_1(y)$ بحيث ان

$$P_1\{w_1(y) < w_2(y)\} = 1 - \alpha \quad \dots\dots\dots(4)$$

ان $1 - \alpha$ يسمى معامل الثقة حيث $0 < \alpha < 1$ يسمى مستوى المعنوية. المعادلة (4) تبين ان الفترة المفتوحة $(w_2(y), w_1(y))$ هي فترة التقدير بحيث يكون الاحتمال الشرطي لكي تكون θ ضمنها مساوياً للعدد $1 - \alpha$ الذي يقع بين الصفر والواحد الصحيح، فاذا كان توزيع X هو التوزيع الطبيعي بوسط θ وتباين معلوم σ^2 عندئذ يكون توزيع $Y = \bar{X}$ ، حيث ان $\bar{X}$ هو وسط العينة العشوائية هو ايضاً التوزيع الطبيعي بوسط θ وتباين σ^2/n ويمكن توضيح ذلك بسهولة باستخدام فكرة الدالة المولدة للعزوم. ولغرض ايجاد التوزيع اللاحق لـ θ نتبع الخطوات التالية علماً بان التوزيع الاول $h(\theta)$ للمتغير العشوائي θ هي التوزيع الطبيعي بوسط وتباين معلومين μ, J^2 على الترتيب

$$k(\theta, Y) = \frac{1}{2\pi J\sigma/\sqrt{n}} \exp \left\{ -\frac{(\theta - \mu)^2}{2J^2} - \frac{(Y - \theta)^2}{2\sigma^2/n} \right\} \quad \dots\dots\dots(5)$$

حيث ان $Y, \theta \in R$

$$\begin{aligned} (Y - \theta)^2 &= [(Y - \mu) + (\mu - \theta)]^2 \\ &= (\theta - \mu)^2 - 2(\theta - \mu)(Y - \mu) + (Y - \mu)^2 \end{aligned} \quad \text{وبما ان} \quad \dots\dots\dots(6)$$

فان المعادلة (5) تصبح بالشكل التالي:

$$\begin{aligned} k(\theta | Y) &= \frac{1}{2\pi J\sigma / \sqrt{n}} \exp \left\{ -\frac{1}{2} \left[\frac{(\theta - \mu)^2}{J^2} + \frac{(\theta - \mu)^2}{\sigma^2 / n} - 2 \frac{(\theta - \mu)(Y - \mu)}{\sigma^2 / n} + \frac{(Y - \mu)^2}{\sigma^2 / n} \right] \right\} \\ &= \frac{1}{2\pi J\sigma / \sqrt{n}} \exp \left[-\frac{1}{2} \left[\left(\frac{1}{J^2} + \frac{1}{\sigma^2 / n} \right) (\theta - \mu)^2 - \frac{2}{\sigma^2 / n} (\theta - \mu)(Y - \mu) \right] \right. \\ &\quad \left. + \frac{1}{\sigma^2 / n} (Y - \mu)^2 \right] \end{aligned} \quad \dots\dots\dots(7)$$

وبمقارنة (7) مع دالة الكثافة الاحتمالية للتوزيع الطبيعي ثنائي المتغيرين Y,X المثالية [1]:

$$F(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-p^2}} \exp \left[-\frac{1}{2(1-p^2)} \left[\left(\frac{x-\mu_1}{\sigma_1} \right)^2 - 2p \left(\frac{x-\mu_1}{\sigma_1} \right) \left(\frac{y-\mu_2}{\sigma_2} \right) + \left(\frac{y-\mu_2}{\sigma_2} \right)^2 \right] \right] \quad \dots\dots\dots(8)$$

نجد ان التوزيع المشترك لـ Y, θ هو التوزيع الطبيعي ثنائي المتغيرين حيث

$$\sigma_2^2 = J^2 + \frac{\sigma^2}{n}, \sigma_1^2 = J^2, \mu = \mu_1 = \mu_2$$

علماً بأن

$$\sigma_y^2 = J^2 + \sigma^2 / n, \sigma_\theta^2 = J^2, \mu_2 = \mu_Y, \mu_1 = \mu_\theta$$

ومعامل الارتباط P هو

$$p = \frac{J}{\sqrt{J^2 + \sigma^2 / n}} \quad \dots\dots\dots(9)$$

وعليه فان التوزيع اللاحق لـ θ باستخدام المعادلة (3) ودالة الكثافة الاحتمالية (8) هو التوزيع

الطبيعي بوسط $\mu_\theta + p \frac{\sigma_\theta}{\sigma_Y} (Y - \mu_Y)$ وتباين $\sigma_\theta^2 (1 - p^2)$ ويمكن توضيح ذلك بسهولة [2].

وبفرض ان دالة المنفعة هي دالة مربع الخطأ $u(\hat{\theta}, \theta) = -(\hat{\theta} - \theta)^2$ ، حيث $\hat{\theta}$ هو مقدر لـ θ ، فان مقدر بيز للمعلمة θ هو التوقع الشرطي للتوزيع اللاحق لـ θ أي ان:

$$\begin{aligned}
 d^B(y) &= E(\theta | y) \\
 &= \mu_\theta + p \frac{\mu_\theta}{\mu_y} (Y - \mu_y) = \mu + \frac{J^2 (Y - \mu)}{J^2 + \sigma^2 / n} \\
 &= \frac{\mu(\sigma^2 / n) + yJ^2}{J^2 + \sigma^2 / n} \dots\dots\dots(10)
 \end{aligned}$$

أما منفعة بيز فهي:

$$\begin{aligned}
 E_{d^B(y)}(u | Y) &= -Var(\theta | Y) \\
 &= \sigma_\theta^2 (1 - p^2) \\
 &= J^2 \left(1 - \frac{J^2}{J^2 + \sigma^2 / n} \right) \\
 &= \frac{J^2 \sigma^2 / n}{J^2 + \sigma^2 / n} \dots\dots\dots(11)
 \end{aligned}$$

4. فئرة ثقتيز

إذا قمنا على التو بإجراء مشاهدة للمتغير العشوائي Y عندئذ يكون المتغير العشوائي θ معبراً عن المتغير العشوائي $\theta | Y$ ، لذلك نتخلّى عن الرمز $\theta | Y$ ونستخدم الحرف θ بدلاً عنه ويكون توزيع المتغير العشوائي المثالي حسب نظرية الغاية المركزية:

(Central Limit Theorem)

$$\frac{\theta - E(\theta)}{\sqrt{\text{var}(\theta)}} = \frac{\theta - \frac{\mu(\sigma^2 / n) + yJ^2}{J^2 + \sigma^2 / n}}{\sqrt{\frac{J^2 \sigma^2 / n}{J^2 + \sigma^2 / n}}} \dots\dots\dots(12)$$

هو المتغير العشوائي الطبيعي القياسي [4]
بفرض $\alpha = 0.05$ وباستخدام جداول التوزيع الطبيعي القياسي نجد ان

$$p_r \left[-2 < \frac{\theta - \frac{\mu(\sigma^2/n) + YJ^2}{J^2 + \sigma^2/n}}{\sqrt{\frac{J^2/n}{J^2 + \sigma^2/n}}} < 2 \right] = 0.95 \quad \dots\dots\dots(13)$$

نأخذ بنظر الاعتبار المتباينة المحصورة بين القوسين [] في المعادلة (13) وبضرب جميع حدودها بالمقدار $\frac{J\sigma/\sqrt{n}}{\sqrt{J^2 + \sigma^2/n}}$ تصبح بالشكل التالي

$$\frac{-2J\sigma/\sqrt{n}}{\sqrt{J^2 + \sigma^2/n}} < \theta - \frac{\mu(\sigma^2/n) + YJ^2}{J^2 + \sigma^2/n} < \frac{2J\sigma/\sqrt{n}}{\sqrt{J^2 + \sigma^2/n}} \quad \dots\dots\dots(14)$$

وبإضافة المقدار $\frac{\mu(\sigma^2/n) + J^2}{J^2 + \sigma^2/n}$ الى جميع حدود المتباينة نجد ان:

$$\frac{\mu(\sigma^2/n) + YJ^2}{J^2 + \sigma^2/n} - \frac{2J\sigma/\sqrt{n}}{\sqrt{J^2 + \sigma^2/n}} < \theta < \frac{\mu(\sigma^2/n) + YJ^2}{J^2 + \sigma^2/n} - \frac{2J\sigma^1/\sqrt{n}}{\sqrt{J^2 + \sigma^2/n}} \quad \dots\dots\dots(15)$$

وهي فترة بيز للمعلمة θ .

إذا كان حجم العينة كبيراً بمافيه الكفاية فان $Y = \bar{X}$ تقرب من μ كفاية ويكون σ^2/n صغيراً جداً ويقرب من الصفر عندما تقترب n من ما لانهاية وتكون عندئذ غاية فترة بيز هي الفترة التالية

$$(Y - 2\sigma/\sqrt{n}, Y + 2\sigma/\sqrt{n}) \quad \dots\dots\dots(16)$$

5. النصيات والنائج:

لقد تم التوصل من هذا البحث الى ايجاد فترة تقدير بيز للمعلمة θ بمعامل ثقة 0.95 على فرض ان توزيع المشاهد X هو التوزيع الطبيعي بوسط θ وتباين معلوم وباقتراح توزيع مسبق لـ θ ، التوزيع الطبيعي بوسط وتباين معلومين. لما لهذا التوزيع من تطبيقات عملية كثيرة. نوصي من هذا البحث باستخدام فترات التقدير البيزية في التطبيقات العملية لأن معالم التوزيعات في هذه التطبيقات هي في الغالب معالم عشوائية وليست كميات ثابتة غير معروفة. كما جرت العادة بأساليب التقدير بالطرائق التقليدية.

المصادر:

1. Anderson, T.W., (1972): An Introduction to Multivariate statistical Analysis, Wiley Eastern private Limited, New Delhi.
2. الجاسم، صباح ود. الصراف، زكي (1992) : نظرية القرارات الاحصائية دار الحكمة للطباعة والنشر- جامعة بغداد.
3. الجاسم، صباح (1996) : تقدير معلمة توزيع لابلاس، مجلة العلوم الاقتصادية والادارية- كلية الادارة والاقتصاد- جامعة بغداد.
4. الناصر، د. عبد المجيد ود. ظافر حسين رشيد (1988): الاستدلال الاحصائي، مطبعة التعليم العالي- بغداد.