مجلة ابن الهيثم للعلوم الصرفة والتطبيقية المجلد23 (1) 2010

مقارنة تجريبية بين المقدرات التقليدية ومقدر بيز لمعلمة التوزيع الاسي

نادية هاشم النور

قسم الرياضيات ، كلية العلوم ، الجامعة المستنصرية

الخلاصه

اهتم البحث بمقارنة اداء المقدرات التقليدية لمعلمة التوزيع الاسي المتمثلة بـ (مقدر الامكان الاعظم ، المقدر المنتظم غير المتحيز ذو الاصغر تباين) ومقدر بيز في حالة خلو البيانات من المشاهدات الشاذة وفي حالة احتواء البيانات على نسبة من تلك المشاهدات .

من خلال توظيف اسلوب المحاكاة (طريقة *Monte Carlo*) واعتماد متوسط مربعات الخطأ (*MSE*) مقياس احصائي للمقارنة بين اداء المقدرات الثلاثة عند احجام عينات تراوحت مابين الصغيرة ، المتوسطة والكبيرة (n=5,10,25,50,100) ولحالات مختلفة (بتكرار 1000 تجربة لكل حالة) تم التوصل الى ان مقدر بيز هو الافضل لجميع حالات التوزيع الاسي المدروسة عند قيم محددة لـ α,β .

المقدمة

ان الهدف الاساس من دراسة اي مجتمع يكمن في استدلال او تقدير بعض خصائصه او معالمه التي غالباً ماتكون مجهولة ونرغب في تقديرها .

اتخذت عملية التقدير اتجاهين مختلفين يعرف الاتجاه الاول بالمدرسة التقليدية (الكلاسيكية) ويضم طرائق متعددة في التقدير، منها: (طريقة العزوم *Moment Method* ، و طريقة الامكان الاعظم *Maximum Likelihood Method* ، و طريقة المقدر المنتظم غير المتحيز ذي الاصغر تباين *Uniformly Minimum Variance Unbiased Estimator* ،.....)، اذ يفترض هذا الاتجاه أن المعلمة التي يراد تقديرها لمجتمع معين عبارة عن كمية ثابتة غير معلومة تقدر في ضوء معطيات العينة بوصفها المصدر الوحيد للمعلومات . اما الاتجاه الثاني الذي يعرف بالمدرسة البيزية فيرى ان الاعتماد على المعلومات المتوافرة من العينة غير كاف لعمل التحليلات الاحصائية؛ لذا يجد ضرورة الاعتماد على افتراض أن المعلمة (المعلمات) التي يراد تقديرها عبارة عن متغير (متغيرات) عشوائية لابد من الحصول على معلومات اولية عنها من خلال البيانات والتجارب السابقة وتصاغ تلك المعلومات بشكل توزيع احتمالي اولي (سابق) *Prior Probability Distribution* ومن ثم يتم التقدير بالاعتماد على التوزيع الاحتمالي اللاحق *Posterior Probability Distribution*.

نظراً لاهمية التوزيع الاسي وتطبيقاته المتعددة في شتى مجالات الحياة ، فقد استحوذ موضوع تقدير معلمة هذا التوزيع على اهتمام العديد من الباحثين فضلاً عن الاهتمام بالتوصل الى مقدرات حصينة تكون ذا اداء قريب من اداء الطرائق التقليدية عند تحقق الافتراضات المحددة "منها خلو البيانات من المشاهدات الشاذة [1] التي تعرف بكونها اما مشاهدات مخالفة (*Discordant*) او مشاهدات ملوثة (*Contaminant*) ويقصد بالمخالفة تلك المشاهدة التي تظهر

بشكل غير منسجم مع بقية البيانات اما الملوثة فهي المشاهدة التي لا تكون ضمن المجتمع المحدد [2] " وافضل منها في حالة الانحراف عن تلك الافتراضات. ومن البحوث الاكثر حداثة التي تناولت موضوع تقدير معلمة التوزيع الاسي نذكر :

في عام 1996 اوجد كل من *Siu* و*Tso* [3] معلمة التوزيع من خلال اعتماد مقدرات الامكان الاعظم، ومقدرات *Shrinkage* .

في عام 2001 قدم كل من *Elfessi* و *Reineke* [4] المقدرات التقليدية لمعلمة التوزيع الاسي من وجهة نظر بيزية وتم تقدير تلك المقدرات عددياً من خلال اعتماد بيانات لـ 20 مشاهدة كما اوضحا امكانية استخلاص صيغ المقدرات التقليدية من صيغة مقدر بيز.

في عام 2005 ناقش كل من *Ahmed* ، *Volodin* و *Hussein* [5] مسألة تقدير معلمة التوزيع الاسي باعتماد طريقة الامكان الموزونة *Weighted Likelihood* وبتوظيف اسلوب المحاكاة بطريقة *Monte Carlo* عند احتواء البيانات على نسبة من المشاهدات الشاذة مع اقتراح مقدر حصين للمعلمة.

في عام 2009 اهتم *Asgharzadeh* [6] باشتقاق مقدرات بيز وبيز التجريبي *Bayes and Empirical Bayes Estimators* لمعلمة التوزيع الاسي .

وقد اهتم البحث الحالي بمقارنة بعض المقدرات التقليدية لمعلمة التوزيع الاسي والمتمثلة بـ (مقدر الامكان الاعظم ، المقدر المنتظم غير المتحيز ذو الاصغر تباين) ومقدر بيز في حالة خلو البيانات من المشاهدات الشاذة وفي حالة احتواء البيانات على نسبة من تلك المشاهدات وتم توظيف اسلوب المحاكاة (طريقة *Monte Carlo*) واعتماد متوسط مربعات الخطأ *Mean square Error (MSE)* مقياساً احصائياً لتحديد الافضل اداءً من بين المقدرات الثلاثة المدروسة .

التوزيع الاسي

يعد التوزيع الاسي احد التوزيعات المستمرة الواسعة الاستخدام لاسيما في مجال تجارب الحياة وتتمثل دالة كثافته الاحتمالية بالصيغة الاتية [4،7،8] :

$$f(x;\theta)=\theta\, e^{-\theta x} \qquad ; \quad 0<x<\infty, \theta>0 \qquad \dots (1)$$

اذ ان : θ معلمة التوزيع .

كما ذكر آنفاً هناك طرائق متعددة لتقدير معلمة التوزيع سيتناول البحث الحالي كل من :

- الطرائق التقليدية المتمثلة بطريقة الامكان الاعظم وطريقة المقدر المنتظم غير المتحيز ذي الاصغر تباين .
- طريقة بيز .

مقدر الامكان الاعظم

ليكن $x_1,x_2,\dots,x_n$ عينة عشوائية بحجم n مسحوبة من مجتمع معين مع دالة كثافة احتمالية $f(x;\theta)\ ;\ \theta\in\Omega$ ، (Ω : فضاء المعلمة *Space Parameter*) [8].

إن مبدأ طريقة الامكان يتمثل بايجاد مقدر للمعلمة θ بما يعظم دالة الامكان التي هي عبارة عن دالة الكثافة الاحتمالية المشتركة للعينة العشوائية والتي تعطى بالصيغة الاتية [8]:

$$L(x_1,x_2,\dots,x_n;\theta)=f(x_1;\theta).f(x_2;\theta).\dots.f(x_n;\theta)=\prod_{i=1}^{n} f(x_i;\theta) \qquad \dots(2)$$

وبما ان قيمة المعلمة التي تجعل دالة الامكان اكبر مايمكن هي نفسها التي تجعل لوغاريتم الدالة اكبر مايمكن استناداً الى حقيقة أن الدالة اللوغارتمية ذو صفة تزايدية رتيبة *Monotonic Increasing* لذا يمكن التوصل الى مقدر الامكان الاعظم للمعلمة θ من حل المعادلة الموضحة بالصيغة (3) ادناه والمتضمنة اخذ لوغاريتم دالة الامكان والتفاضل نسبة الى θ والمساواة بالصفر .

$$\frac{\partial \ell(x_1, x_2, \ldots, x_n; \theta)}{\partial \theta} = 0 \qquad \ldots(3)$$

اذ ان :

$$\ell(x_1, x_2, \ldots, x_n; \theta) = \ln L(x_1, x_2, \ldots, x_n; \theta)$$

وبافتراض دالة التوزيع الاسي فأن دالة الامكان تكون :

$$L(x_1, x_2, \ldots, x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta) = \theta^n \, e^{-\theta \sum_{i=1}^{n} x_i} \qquad \ldots(4)$$

بأخذ لوغاريتم الطرفين ومن ثم التفاضل نحصل على *MLE* للمعلمة θ و على وفق الخطوات الاتية :

$$\ell(x_1, x_2, \ldots, x_n; \theta) = \ln L(x_1, x_2, \ldots, x_n; \theta) = n \ln \theta - \theta \sum_{i=1}^{n} x_i$$

$$\frac{\partial \ell(x_1, x_2, \ldots, x_n; \theta)}{\partial \theta} = \frac{n}{\theta} - \sum_{i=1}^{n} x_i$$

$$\frac{\partial \ell(x_1, x_2, \ldots, x_n; \theta)}{\partial \theta} = 0 \quad \Rightarrow \frac{n}{\hat{\theta}} - \sum_{i=1}^{n} x_i = 0$$

$$\Rightarrow \hat{\theta}_{\mathrm{ML}} = \frac{n}{\sum_{i=1}^{n} x_i} \qquad \ldots\ldots 5$$

المقدر المنتظم غير المتحيز ذو الاصغر تباين

تعد دالة التوزيع الاسي احد الدوال التي تنتمي الى العائلة الاسية *Exponential Family* والمتمثلة بالصيغة : [8]

$$f(x; \theta) = e^{[p(\theta)k(x) + s(x) + q(\theta)]} \qquad \ldots(6)$$

اذ يمكن اعادة كتابة دالة التوزيع الاسي بالشكل الاتي :

$$f(x; \theta) = \theta \, e^{-\theta x} = e^{\ln \theta - \theta x}$$

$$p(\theta) = -\theta, k(x) = x, s(x) = 0, q(\theta) = \ln \theta$$

و من ثم فأن: $t = \sum_{i=1}^{n} k(x_i) = \sum_{i=1}^{n} x_i$ يمثل احصاءة كافية كاملة *Complete Sufficient Statistic* للمعلمة θ .

بأخذ التوقع الرياضي للاحصاءة الكافية الكاملة ($t = \sum_{i=1}^{n} x_i$) وبالرجوع الى نظريـة ليمان– شيفيه *Lehmann - Scheffe* فأن *UMVUE* للمعلمة θ يمكن ايجاده على وفق الخطوات الاتية :

$$\mathrm{E}(t) = \mathrm{E}(\sum_{i=1}^{n} x_i) = \frac{n}{\theta} \Rightarrow \mathrm{E}(\frac{t}{n}) = \frac{1}{\theta}$$

$$\mathrm{E}(\frac{n}{t}) = \int_t \frac{n}{t} f(t;\theta)dt = \int_0^{\infty} \frac{n}{t} \frac{t^{n-1}\theta^n e^{-\theta t}}{\Gamma(n)} dt$$

$$= \frac{n}{n-1}\int_0^{\infty} \frac{t^{(n-1)-1}\theta^n e^{-\theta t}}{\Gamma(n-1)} dt$$

$$= \frac{n}{n-1}\theta \quad \Rightarrow \quad \mathrm{E}(\frac{n-1}{t}) = \theta$$

$$\Rightarrow \quad \hat{\theta}_{\mathrm{UMVU}} = \frac{n-1}{t} = \frac{n-1}{\sum_{i=1}^{n} x_i} \qquad \ldots(7)$$

مقدر بيز

يعتمد اسلوب بيز كما ذكر آنفاً على استخدام معلومات مسبقة عن المعلمة غير المعروفة بوصفها متغيراً عشوائياً وبافتراض ان هذه المعلومات المسبقة يمكن صياغتها على شكل توزيع احتمالي للمعلمة يعرف بدالة الكثافة الاحتمالية الاولية (السابقة) ويجري التعرف على هذه المعلومات من بيانات وتجارب سابقة أو من النظرية التي تحكم تلك الظاهرة، وكذلك يعتمد اسلوب بيز على معلومات العينة الحالية المتمثلة بدالة الأمكان الخاصة بالمشاهدات، وعليه

بدمج دالة الكثافة الاحتمالية الاولية للمعلمة $\pi(\theta)$ مع دالة الأمكان $f(\mathrm{x}/\theta)$ ، يتم الحصول على دالة الكثافة الاحتمالية اللاحقة للمعلمة $f(\theta/x)$ التي تمثل توزيع احتمالي مشروط للمعلمة θ بشرط الحصول على العينة [9].
ويمكن تلخيص ما تقدم رياضياً على وفق الاتي:

$$f(\theta/\mathrm{x}) \;\; f(\mathrm{x}) = f(\mathrm{x}/\theta) \;\; \pi(\theta)$$

اذ ان :

$$f(\mathrm{x}) = \int_\theta f(\mathrm{x},\theta)d\theta = \int_\theta f(\mathrm{x}/\theta) \;\; \pi(\theta)d\theta$$

وعليه فان:

$$f(\theta/\mathrm{x}) = \frac{f(\mathrm{x}/\theta)\,\pi(\theta)}{f(\mathrm{x})} = \frac{f(\mathrm{x}/\theta)\,\pi(\theta)}{\int_\theta f(\mathrm{x}/\theta) \;\; \pi(\theta)d\theta}$$

$$f(\theta/\mathrm{x}) \;\; \propto \;\; f(\mathrm{x}/\theta) \;\; \pi(\theta)$$

إذ إن العلاقة $\propto$ تشير إلى إن الكمية تناسبية .

عموماً ، عند وجود عينة عشوائية $x_1, x_2, \ldots, x_n$ فأن دالة الكثافة الاحتمالية اللاحقة تعطى بالصيغة : **مجلة**

$$f(\theta/x_1,x_2,...,x_n)=\frac{\prod_{i=1}^{n} f(x_i;\theta)\pi(\theta)}{\int_\theta \prod_{i=1}^{n} f(x_i;\theta)\pi(\theta)d\theta}=\frac{L(x_1,x_2,...,x_n;\theta)\pi(\theta)}{\int_\theta L(x_1,x_2,...,x_n;\theta)\pi(\theta)d\theta}$$

بعد الحصول على دالة الكثافة الاحتمالية اللاحقة وللوصول إلى الهدف الرئيس المتمثل بتقدير المعلمة ، يتم تحديد ما يعرف بدالة الخسارة *Loss function* التي يرمز لها عادةً بـ $\ell(\hat{\theta}-\theta)$ وتعرف كما يأتي [10] :

$$\ell(\hat{\theta}-\theta)\geq 0 \qquad \forall\ \hat{\theta},\forall\theta$$
$$\ell(\hat{\theta}-\theta)=0 \qquad \forall\ \hat{\theta}=\theta$$

ان التقدير النقطي بأسلوب بيز يعتمد على إيجاد قيمة $\hat{\theta}$ التي تقلل توقع الخسارة ، بمعنى :

$$\underset{\hat{\theta}}{Min}\, E\big(\ell(\hat{\theta}-\theta)\big)=\underset{\hat{\theta}}{Min}\int \ell(\hat{\theta}-\theta)\ f(\theta/x)\ d\theta$$

هناك أنواع عديدة من دوال الخسارة، والنوع الأكثر شيوعاً واستخداماً هو دالة الخسارة التربيعية *Quadratic loss function* [10]:

$$\ell(\hat{\theta},\theta)=(\hat{\theta}-\theta)^2$$
$$E[\ell(\hat{\theta},\theta)/x]=E[(\hat{\theta}-\theta)/x]^2$$
$$=\int(\hat{\theta}-\theta)^2.f(\theta/x)d\theta$$
$$=\hat{\theta}^2\int f(\theta/x)d\theta-2\hat{\theta}\int\theta\ f(\theta/x)d\theta+\int\theta^2 f(\theta/x)d\theta$$

للحصول على أقل توقع خسارة يتم أخذ المشتقة الأولى للمعادلة اعلاه بالنسبة إلى $\hat{\theta}$ ومساواتها إلى الصفر نحصل على :

$$2\hat{\theta}\int f(\theta/x)d\theta-2\int\theta f(\theta/x)d\theta=0$$

$$\hat{\theta}.1-E(\theta/x)=0\Rightarrow\hat{\theta}=E(\theta/x)$$

وعليه فالصيغة السابقة توضح التقدير النقطي للمعلمة θ الذي يكون امثل عندما يكون مساوياً لتوقع دالة الكثافة الاحتمالية اللاحقة للمعلمة θ ، وبعبارة أخرى إن تقدير المعلمة بموجب أسلوب بيز، وبالاعتماد على دالة خسارة تربيعية يكون مساوياً إلى الوسط الحسابي للتوزيع اللاحق للمعلمة .

لتكن دالة الكثافة الاحتمالية الاولية بالصيغة [4] :

$$\pi(\theta)=\theta^{\alpha-1}e^{-\beta\theta}\qquad;\theta>0,\ -\infty<\alpha<\infty,\beta\geq 0$$

فأن دالة التوزيع اللاحق ومقدر بيز للمعلمة θ يمكن ايجادهما وفق الخطوات الاتية :

$$f(\theta / x_1, x_2, \dots, x_n) = \frac{\theta^n e^{-\theta \sum_{i=1}^{n} x_i} \theta^{\alpha-1} e^{-\beta\theta}}{\int_0^\infty \theta^n e^{-\theta \sum_{i=1}^{n} x_i} \theta^{\alpha-1} e^{-\beta\theta}\, d\theta}$$

$$= \frac{\theta^{\alpha+n-1} e^{-\theta(\beta + \sum_{i=1}^{n} x_i)}}{\int_0^\infty \theta^{\alpha+n-1} e^{-\theta(\beta + \sum_{i=1}^{n} x_i)}\, d\theta}$$

$$= \frac{\theta^{\alpha+n-1} e^{-\theta(\beta + \sum_{i=1}^{n} x_i)}}{\Gamma(\alpha+n) \frac{1}{(\beta + \sum_{i=1}^{n} x_i)^{\alpha+n}}}$$

$$= \frac{(\beta + \sum_{i=1}^{n} x_i)^{\alpha+n}}{\Gamma(\alpha+n)} \theta^{\alpha+n-1} e^{-\theta(\beta + \sum_{i=1}^{n} x_i)}$$

$$\Rightarrow \hat{\theta}_{Bayes} = \mathrm{E}(\theta / x_1, x_2, \dots, x_n) = \int_\theta \theta\, f(\theta / x_1, x_2, \dots, x_n)\, d\theta$$

$$= \int_0^\infty \theta \frac{(\beta + \sum_{i=1}^{n} x_i)^{\alpha+n}}{\Gamma(\alpha+n)} \theta^{\alpha+n-1} e^{-\theta(\beta + \sum_{i=1}^{n} x_i)}\, d\theta$$

$$= \frac{(\beta + \sum_{i=1}^{n} x_i)^{\alpha+n}}{\Gamma(\alpha+n)} \int_0^\infty \theta^{(\alpha+n+1)-1} e^{-\theta(\beta + \sum_{i=1}^{n} x_i)}\, d\theta$$

$$= \frac{(\beta + \sum_{i=1}^{n} x_i)^{\alpha+n}}{\Gamma(\alpha+n)} \Gamma(\alpha+n+1) \frac{1}{(\beta + \sum_{i=1}^{n} x_i)^{\alpha+n+1}}$$

$$\Rightarrow \hat{\theta}_{Bayes} = \frac{\alpha+n}{\beta + \sum_{i=1}^{n} x_i} \qquad \dots(8)]$$

ومن صيغة مقدر بيز اعلاه يمكن استخلاص صيغ المقدرات التقليدية وذلك باعتماد قيم معينة لكل من α, β فعند $\alpha = \beta = 0$ ينتج MLE وعند $\alpha = -1, \beta = 0$ ينتج $UMVUE$.

تم اعتماد اسلوب المحاكاة "طريقة *Monte Carlo*" للمقارنة بين المقدرات الثلاثة الاتية :

1. مقدر الامكان الاعظم ، $\hat{\theta}_{ML}$ (الصيغة (5)) .
2. المقدر المنتظم غير المتحيز ذو الاصغر تباين ، $\hat{\theta}_{UMVU}$ (الصيغة (7)) .
3. مقدر بيز ، $\hat{\theta}_{Bayes}$ (الصيغة (8)) ، بأختيار قيم مختلفة لـ α,β وكما يأتي :

$\alpha = \beta$: $\alpha=\beta=1$ ، $\alpha=\beta=4$

$\alpha < \beta$: $\alpha=0,\beta=1$ ، $\alpha=0,\beta=4$ ، $\alpha=1,\beta=4$

$\alpha > \beta$: $\alpha=1,\beta=0$ ، $\alpha=4,\beta=0$ ، $\alpha=4,\beta=1$

تم تحليل النتائج بالاعتماد على قيمة متوسط مربعات الخطأ (*MSE*) *Mean Square Error* الذي يساوي مربع مقدار التحيز (B) مضافاً اليه مقدار التباين $\text{var}(\hat{\theta})$ ، بمعنى :

$$MSE = B^2 + \text{var}(\hat{\theta})$$

اذ ان :

$$B = E(\hat{\theta}) - \theta \quad ; \quad E(\hat{\theta}) = \frac{\sum_{i=1}^{I} \hat{\theta}_i}{\mathrm{I}}$$

$$\text{var}(\hat{\theta}) = \sum_{i=1}^{\mathrm{I}} (\hat{\theta}_i - E(\hat{\theta}))^2 / (\mathrm{I}-1)$$

I: عدد مرات تكرار تجربة المحاكاة.

اما الحالات المدروسة فتمثلت بما يأتي :

1. التوزيع الاسي على وفق قيم مختارة للمعلمة θ وكما يأتي :

$X \sim Exp(\theta = 0.5)$

$X \sim Exp(\theta = 2)$

2. التوزيع الاسي الملوث بنسبة 20% وكما يأتي :

$X \sim 80\%\ Exp(\theta = 0.5) + 20\%\ Exp(\theta = 10)$

$X \sim 80\%\ Exp(\theta = 2) + 20\%\ Exp(\theta = 0.5)$

3. التوزيع الاسي الملوث بنسبة 40% وكما يأتي :

$X \sim 60\%\ Exp(\theta = 0.5) + 40\%\ Exp(\theta = 10)$

$X \sim 60\%\ Exp(\theta = 2) + 40\%\ Exp(\theta = 0.5)$

وتم تطبيق الحالات الست المذكورة لاحجام العينات المتراوحة مابين الصغيرة (n =5,10) ، و المتوسطة (n = 25) والكبيرة (n = 50,100) وبتكرار مقداره (I=1000) .

تحليل النتائج

يمكن تحليل النتائج التي توصلنا اليها خلال الجانب التجريبي والواردة في الجداول (1)...(6) بالنقاط الآتية :

1. عند مقارنة اداء المقدرات التقليدية المتمثلة بـ MLE و $UMVUE$ بعضهما مع بعض نلاحظ تقارب اداء كل منهما عند حجوم العينات المتوسطة والكبيرة ($n \geq 25$) ونلاحظ تفوق اداء $UMVUE$ على MLE عند التعامل مع حالة التوزيع الاسي غير الملوث عند $\theta = 0.5$ (جدول(1))، وعند $\theta = 2$ (جدول(2)) كذلك الحال عند التعامل مع التوزيع الملوث بنسبة 20% (جدول(3))، و 40% (جدول(4)) عندما تكون قيمة معلمة التوزيع الملوث اكبر من قيمة المعلمة للتوزيع الاساسي في حين يتضح العكس عند التعامل مع حالة التوزيع الملوث عندما تكون قيمة المعلمة للتوزيع الملوث اصغر من قيمة المعلمة للتوزيع الاساسي بمعنى تراجع اداء $UMVUE$ ليتفوق عليه اداء MLE (الجدولين(5) و(6)) .
2. عند مقارنة اداء المقدرات البيزية بعضها مع بعض يتضح ان اداء تلك المقدرات يتغير تبعاً لطبيعة الحالة المدروسة ، حجم العينة والقيمة المحددة لـ α, β فضلاً عن قيمة معلمة التوزيع وكما موضح ادناه :

- عند التعامل مع الحالة غير الملوثة للتوزيع الاسي وعند $\theta = 0.5$ (جدول(1)) ولحجوم العينات الصغيرة (n=5,10) نلاحظ تفوق اداء مقدر بيز عند $\alpha = 1, \beta = 4$ ، اذ تنخفض قيمة MSE مقارنة ببقية التقديرات، في حين يتفوق ذلك المقدر عند $\alpha = 0, \beta = 4$ لحجوم العينات المتوسطة والكبيرة $n \geq 25$.
- عند التعامل مع الحالة غير الملوثة للتوزيع الاسي وعند $\theta = 2$ (جدول(2)) ولحجوم العينات الصغيرة (n=5,10) نلاحظ تفوق اداء مقدر بيز عند $\alpha = \beta = 1$ ، اذ تنخفض قيمة MSE مقارنة ببقية التقديرات، في حين يتفوق ذلك المقدر عند $\alpha = \beta = 4$ لحجوم العينات (n=25,50) اما عند n=*100* فيلاحظ تفوق الاداء عند $\alpha = 1, \beta = 4$.
- عند التعامل مع الحالة الملوثة للتوزيع الاسي بنسبة 20% وعندما تكون معلمة التوزيع الملوث اكبر من قيمة معلمة التوزيع الاسي الاساسي (جدول(3)) نلاحظ تفوق اداء مقدر بيز عند $\alpha = 0, \beta = 4$ لجميع حجوم العينات بأستثناء n=*5* وعند مضاعفة نسبة التلويث لتكون 40% (جدول(4)) نجد ان هذا المقدر قد ازداد في تفوقه ليشمل جميع حجوم العينات من دون استثناء .
- عند التعامل مع الحالة الملوثة للتوزيع بنسبة 20% وعندما تكون قيمة معلمة التوزيع الملوث اصغر من قيمة معلمة التوزيع الاساسي (جدول(5)) نلاحظ تفوق اداء مقدر بيز عند $\alpha = 4, \beta = 0$ لحجوم العينات المتوسطة والكبيرة وعند مضاعفة نسبة التلويث لتكون 40% (جدول(6)) نجد ان هذا المقدر قد ازداد في تفوقه ليشمل جميع حجوم العينات بأستثناء n=*5* .
- عموماً ، ان افضل اداء لمقدر بيز عند التعامل مع التوزيع غير الملوث وبقيمة معلمة مساوية الى 0.5 يكون عندما $\alpha < \beta$ كذلك الحال عند التعامل مع التوزيع الملوث وبقيمة معلمة اكبر من قيمة معلمة التوزيع الاساسي وينعكس الحال عند التعامل مع التوزيع الملوث بقيمة معلمة اصغر من قيمة معلمة التوزيع الاساسي ليكون افضل اداء عندما $\alpha > \beta$.

3. عند مقارنة اداء المقدرات المدروسة بعضها مع بعض نجد ان مقدر بيز قد احتل دائماً المرتبة الاولى والثانية في افضلية الاداء عند التعامل مع الحالة غير الملوثة والحالة الملوثة على حد سواء تبعاً للقيمة المحددة لـ α, β . اما المقدرات التقليدية فقد احتلت المرتبة الثالثة في بعض الاحيان .

الاستنتاجات والتوصيات

يمكن تلخيص أهم الإستنتاجات التي اوضحتها نتائج المحاكاة بما يأتي:

1. تنخفض قيم MSE عموماً بزيادة حجم العينة .

2. غالباً ماتكون قيم MSE لمقدرات بيز اصغر من قيمها المناظرة للمقدرات الكلاسيكية .
3. يتقارب اداء المقدرات التقليدية المتمثلة بـ MLE و $UMVUE$ بعضهما مع بعض عند حجوم العينات المتوسطة والكبيرة .
4. يتفوق اداء $UMVUE$ على MLE عند التعامل مع البيانات الخالية من المشاهدات الشاذة، في حين يتباين

n	***ML***	***UMVU***	***Bayes***

ادائهما عند التعامل مع البيانات بوجود تلك المشاهدات تبعاً لقيمة معلمة التوزيع الملوث .

5. يتغير اداء المقدرات البيزية تبعاً لقيمة معلمة التوزيع الاسي ، حالة التوزيع المدروسة (احتواء/خلو البيانات من المشاهدات الشاذة) ، حجم العينة والقيمة المحددة لـ α,β .
6. ان مقدر بيز هو الافضل لجميع حالات التوزيع الاسي المدروسة تبعاً للقيم المحددة لـ α,β .

في ضوء ماورد .. نوصي بدراسة سلوك المقدرات عند حجوم العينات الصغيرة بشكل منفرد كما نوصي باعتماد مقدر بيز لتقدير معلمة التوزيع الاسي عند قيم محددة لـ α,β .

المصادر

1. Barnett, V. and Lewis, T. (1984). "Outliers in Statistical Data". 2nd Edition, New York, John Wiley.
2. Beckman, R. J. and Cook, R. D. (1983). "Outlier…….s". Technometrics, 25: 119-149.
3. Siu, K. T. and Tso, G. (1996). "Shrinkage Estimation of Reliability for Exponentially Distribution Lifetimes". Comm. Statist. Simula., 25: 415-430.
4. Elfessi, A. and Reineke, D. M. (2001). "A Bayesian Look at Classical Estimation: The Exponential Distribution". Journal of Statistics Education, 9: 1.
5. Ahmed, E. S.; Volodin, A. I. and Hussein A. A. (2005). "Robust Weighted Likelihood Estimation of Exponential Parameter". IEEE Transactions on Reliability, 54: 389-395.
6. Asgharzadeh, A. (2009). "On Bayesian estimation from exponential distribution based on records". Journal of the Korean Statistical Society, 38: 125-130.
7. Cohen, A. C. and Helm F. R. (1973). "Estimation in Exponential Distribution". Technometrics, 15: 415-418.
8. Hogg, R. V. and Craig, A. T. (1978). "Introduction to Mathematical Statistics". 4th Edition, New York.
9. Klugman, S. A. (1992). "Bayesian Statistics in Actuarial Science". Kluwer Academic Publishers, USA.
10. Mood, A. M.; Graybil, F. A. and Boes, D. C. (1985). "Introduction to the Theory of Statistics". 3rd Edition, Megraw-Hill.

n	*ML*	*UMVU*	*Bayes*							
			$\alpha=\beta=1$	$\alpha=\beta=4$	$\alpha=0,\beta=1$	$\alpha=0,\beta=4$	$\alpha=1,\beta=0$	$\alpha=4,\beta=0$	$\alpha=1,\beta=4$	$\alpha=4,\beta=1$
5	0.04546	0.03615	0.04301	0.04828	0.02883	0.02883	0.079875	0.333745	0.01733	0.20238
10	0.01843	0.01712	0.01890	0.02391	0.01555	0.01590	0.025171	0.078002	0.01246	0.05815
25	0.00952	0.00808	0.01009	0.01254	0.00829	0.00643	0.01189	0.024506	0.00679	0.02077
50	0.00235	0.00198	0.00256	0.00344	0.00209	0.00169	0.00294	0.006006	0.00183	0.00524
100	0.00278	0.00231	0.00300	0.00372	0.00251	0.00182	0.003306	0.005241	0.00221	0.00482

n	*ML*	*UMVU*	*Bayes*							
			$\alpha=\beta=1$	$\alpha=\beta=4$	$\alpha=0,\beta=1$	$\alpha=0,\beta=4$	$\alpha=1,\beta=0$	$\alpha=4,\beta=0$	$\alpha=1,\beta=4$	$\alpha=4,\beta=1$
5	0.09092	0.04036	0.09241	0.10483	0.04267	0.01509	0.18267	0.70502	0.01439	0.42406
10	0.08129	0.05030	0.08470	0.09540	0.05412	0.01655	0.12292	0.31161	0.02704	0.23118
25	0.01172	0.00849	0.01306	0.01761	0.00949	0.00528	0.01601	0.03526	0.00703	0.02986
50	0.02697	0.02286	0.02839	0.03265	0.02421	0.01719	0.03143	0.04692	0.02058	0.04296
100	0.01023	0.00937	0.01055	0.01154	0.00967	0.00815	0.01116	0.01433	0.00890	0.01358

n	*ML*	*UMVU*	*Bayes*							
			$\alpha=\beta=1$	$\alpha=\beta=4$	$\alpha=0,\beta=1$	$\alpha=0,\beta=4$	$\alpha=1,\beta=0$	$\alpha=4,\beta=0$	$\alpha=1,\beta=4$	$\alpha=4,\beta=1$
5	0.74176	0.34543	0.15382	0.30664	0.24916	1.41131	1.66207	7.56693	1.05461	1.17749
10	0.45062	0.37855	0.24354	0.25518	0.27793	0.81732	0.61758	1.68781	0.63870	0.50889
25	0.15686	0.14108	0.11829	0.10552	0.11836	0.25404	0.18677	0.36126	0.20525	0.18972
50	0.13091	0.12558	0.11512	0.09683	0.11368	0.13948	0.13961	0.18596	0.12512	0.13807
100	0.04462	0.03973	0.03796	0.02564	0.03416	0.02611	0.05041	0.07317	0.02485	0.05454

5	0.24327	0.09676	0.23239	0.22221	0.10791	0.00753	0.46372	1.56863	0.02715	0.91162
10	0.22640	0.14758	0.22222	0.21678	0.15028	0.04165	0.32326	0.72208	0.07159	0.52741
25	0.07673	0.06181	0.07948	0.08734	0.06464	0.03785	0.09346	0.15459	0.04806	0.13426
50	0.09236	0.08292	0.09418	0.09941	0.08479	0.06521	0.10231	0.13525	0.07308	0.12533
100	0.07500	0.07098	0.07583	0.07827	0.07181	0.06293	0.07914	0.09227	0.06660	0.08857

n	***ML***	***UMVU***	***Bayes***							
			$\alpha=\beta=1$	$\alpha=\beta=4$	$\alpha=0,\beta=1$	$\alpha=0,\beta=4$	$\alpha=1,\beta=0$	$\alpha=4,\beta=0$	$\alpha=1,\beta=4$	$\alpha=4,\beta=1$
5	1.83187	1.89192	1.28825	1.11952	1.44158	2.08106	2.25861	5.25954	1.79287	1.54498
10	1.22314	1.39948	1.20980	1.18049	1.37456	1.78244	1.06932	0.74299	1.61669	0.82218
25	1.24481	1.32274	1.2371	1.21660	1.31218	1.49939	1.16957	0.96000	1.42566	1.02680
50	1.34009	1.37922	1.33416	1.31757	1.37256	1.46596	1.30153	1.18930	1.42811	1.22230
100	1.05624	1.07590	1.05585	1.05473	1.07532	1.13153	1.03678	0.97958	1.11206	0.99860

n	***ML***	***UMVU***	***Bayes***							
			$\alpha=\beta=1$	$\alpha=\beta=4$	$\alpha=0,\beta=1$	$\alpha=0,\beta=4$	$\alpha=1,\beta=0$	$\alpha=4,\beta=0$	$\alpha=1,\beta=4$	$\alpha=4,\beta=1$
5	0.56486	0.93073	0.61652	0.74126	0.94525	1.85484	0.35957	0.70719	1.52624	0.19479
10	1.76622	1.85502	1.24465	0.93430	1.27743	1.47242	1.95178	2.95456	1.31256	1.40757
25	0.57245	0.63725	0.58490	0.62199	0.64911	0.89022	0.51366	0.37346	0.81723	0.42448
50	0.65801	0.69454	0.66440	0.68251	0.70042	0.82668	0.62272	0.52428	0.78911	0.56339
100	0.60474	0.62165	0.60794	0.61749	0.62479	0.68582	0.58817	0.54054	0.66827	0.55942

Experimental Comparison between Classical and Bayes Estimators for the Parameter of Exponential Distribution

N. H. Al-Noor
Department of Mathematics, College of Sciences, University of Al-Mustanseriya

Abstract

This paper is interested in comparing the performance of the traditional methods to estimate parameter of exponential distribution (Maximum Likelihood Estimator, Uniformly Minimum Variance Unbiased Estimator) and the Bayes Estimator in the case of data to meet the requirement of exponential distribution and in the case away from the distribution due to the presence of outliers (contaminated values).

Through the employment of simulation (Monte Carlo method) and the adoption of the mean square error (MSE) as criterion of statistical comparison between the performance of the three estimators for different sample sizes ranged between small, medium and large (n=5,10,25,50,100) and different cases (with 1000 replications), it was reached that the Bayes Estimator " with specific values of α and β " is the best for all cases studied .

جدول (1) قيم MSE : $\theta = 0.5$ $X \sim Exp\ ()$

جدول (2) قيم MSE : $\theta = 2$ $X \sim Exp()$

جدول(3) قيم MSE ($\theta = 0.5$) $X \sim 80\%\ Exp\ ):\theta = 10\ ) + 20\%\ Exp\ ($

جدول (4) قيم MSE : $\theta = 0.5$ ($X \sim 60\%\ Exp\ ()\ \theta = 10) + 40\%\ Exp\ ($

جدول (5) قيم MSE : $\theta = 2$ ($X \sim 80\%\ Exp\ ()\ \theta = 0.5\ ) + 20\%\ Exp\ ($

جدول(6) قيم MSE : ($\theta = 2$ ($X \sim 60\%\ Exp:)\ \theta = 0.5) + 40\%\ Exp\ ($