# التقدير الحصين لمعلمة التوزيع الاسي على المدى الجزئي


هناء حسن عبدالله

كلية العلوم / قسم الرياضيات

الجامعة المستنصرية


# Robust Estimates for Parameter of Exponential Distribution on Subrange


Hana'a H. Abdullah

Department of Mathematics

College of Science

AL-Mustansereyah University

# التقدير الحصين لمعلمة التوزيع الاسي على المدى الجزئي

**هناء حسن عبدالله**

**كلية العلوم / قسم الرياضيات**

**الجامعة المستنصرية**

**<u>المستخلص:</u>**

ان التقديرات التقليدية الشائعة لمعلمات التوزيعات تتأثر كثيرا بوجود القيم الشاذة (الشواذ) وللتخلص من تأثيرها يتم البحث عن تقديرات حصينة للمعلمات ، حيث تعطي هذه التقديرات قيم قريبة من التقديرات التقليدية عند عدم وجود الشواذ وقيم جيدة عند وجودها.

في هذا البحث تم اقتراح التقدير الحصين ( $MAD_{\propto}^{*}/D$ $D > 0,$) لمعلمة التوزيع الاسي، ثم عمل دراسة باستخدام اسلوب المحاكاه لتوليد بيانات باحجام مختلفة والمقارنة بين التقديرات الحصينة عند تغير قيم ($\propto$) حيث ان ( $\propto= 0.1, 0.25, 0.4$) مع الانحراف المعياري ($SD$) بالاعتماد على متوسط مربعات الخطأ ($MSE$) معياراً للمقارنة واتضح من خلال الدراسة ان التقدير ($MAD_{0.4}^{*}/0.327$) هو افضل تقدير حصين لمعلمة التوزيع الاسي عند احجام العينات المختلفة ولقيم مختلفة للمعلمة.

## <u>Abstract:</u>

The popular classical estimates for parameters of distributions are highly influenced by outliers, to avoid the effect of these outliers; robust estimates for parameters should be located gives approximate value of the classical estimation when the data are free from outliers.

In this paper, suggestion was made to use ($MAD_{\propto}^{*}/D,\ D > 0$) as a robust estimates for the parameter of the exponential distribution. A simulation technique was used to generate data in different size based on random number generation to compare between the suggested robust estimates for different value of ($\propto$) with the standard deviation depending on the mean square error (MSE). Results indicate that ($MAD_{0.4}^{*}/0.327$) is the best robust estimator for the parameter of exponential distribution of different size for the samples generated and for different value of the parameter.

## ١) المقدمة:

خلال جمع البيانات الخاصة بالتطبيقات العملية في المجتمع الذي يتم دراسته تظهر بعض القيم (المشاهدات) تنحرف بشكل ملحوظ عن بقية مشاهدات العينة وتسمى هذه القيم بالقيم الشاذة او الشواذ (outliers) او الملوثة (contaminant) ...الخ، يعزى سبب ظهور هذه الشواذ الى احتواء المجتمع لمثل هذه القيم او اخطاء يتعرض لها الباحث وهذه القيم الشاذة تؤثر كثيرا في التقديرات التقليدية حتى لو كانت قيمة واحدة مثل معدل العينة Sample mean وتباين العينة Sample variance وغيرها، لهذا فان التقديرات الحصينة تعطي تقدير جيد لهذه المعلمات لانها لاتتاثر بهذه القيم الشاذة في حالة وجودها [1] .

في هذا البحث تم اقتراح مقدر حصين لمعلمة التوزيع الاسي $EXP(\lambda)$ بالشكل $(MAD^{*}_{\propto}(x)/D)$، حيث ان D قيمة عددية موجبة تم ايجادها بحل المعادلة $(e^{D}-e^{-D}=\frac{\alpha}{1-\alpha})$ ، $(\propto = 0.1, 0.25, 0.4)$ وهذا ماسنتناوله في الجانب النظري، اما الجانب العملي حيث تم عمل دراسة باستخدام اسلوب المحاكاه (Simulation) لتوليد البيانات باستخدام طريقة التحويل العكسي ولاحجام عينات مختلفة وقيم $\propto, \lambda$ المختلفة وتم تكرار التجربة ١٠٠٠ مرة. واخيرا تم تحليل النتائج التي حصلنا عليها من هذه الدراسة وايضا تضمن البحث اهم الاستنتاجات التي توصلنا اليها.

## ٢) الجانب النظري:

### ٢.١) تعريف المقدرات $MAD_{\propto}$ , $M_{\propto}$

لتكن $X=\{x_1,..,x_n\}$ مجموعة من المشاهدات ، $x_{(1)} \le \cdots \le x_{(n)}$ احصاءات مرتبية order statistics . ان متوسط المدى $\propto -$ $(\propto -midrange)$ $M_{\propto}(x)$ هو مقدر لمعلمات الموقع [2] ويعرف بالشكل :

$$M_{\propto}(x)=\frac{x_{([n\propto]+1)}+x_{(n-[n\propto])}}{2} \qquad (2.1.1)$$

حيث ان $[n\propto]$ تمثل اكبر عدد صحيح اقل او يساوي $n\propto$ ، $(0<\propto<0.5)$ على الفترة

$$\mathrm{I}_{\propto}=[x_{([n\propto]+1)};x_{(n-[n\propto])}]$$

اما $MAD_{\propto}$ فهو مقدار لمعلمات القياس [1] ويعرف بانه الانحراف المطلق لمتوسط المدى $\propto -$ $(\propto$ $-midrange\ absolute\ deviation)$ حيث ان :

$$MAD_{\propto}(x)=M_{\propto}\{|x_i-M_{\propto}(x)|\} \qquad (2.1.2)$$

لنفرض ان x متغير عشوائي مستمر له دالة كثافة احتمال تجميعية cdf- (Cumulative distribution function) وله قيم موجبة فان:

$$F(x_\propto) = \propto \qquad , \qquad 0 < \propto < 1 \qquad (2.1.3)$$

$x_\propto$ يمثل قيمة التقسيم ال $\propto$ الى $x$ ($x_\propto be \; the \; \propto -percentile \; of \; x$) وبذلك يكون:

$$F(M_\propto) = \propto \qquad (2.1.4)$$

وتستخدم العلاقة اعلاه لايجاد $M_\propto$.

## ٢.٢) المقدر الحصين المقترح $\mathbf{MAD}_\propto^*$/D لمعلمة التوزيع الاسي :

ليكن x له توزيع اسي بالمعلمة $\lambda$ فان pdf:

$$f(x) = \frac{1}{\lambda} e^{-\frac{x}{\lambda}} \; , \quad x > 0, \quad \lambda > 0$$

وان cdf للمتغير X ستكون:

$$F(x) = \begin{cases} 1 - \exp(-x/\lambda) & , \quad x > 0 \\ 0 & , \quad otherwise \end{cases}$$

باستخدام الصيغة الاخيرة و (2.1.4) نحصل على:

$$1 - \exp(-M_\propto/\lambda) = \propto$$

بذلك سيكون:

$$M_\propto = -\lambda \, ln(1 - \propto) \qquad (2.2.1)$$

بما انه $\lambda$ تمثل معلمة قياس في التوزيع $Exp(\lambda)$ حيث ان $SD = \lambda$ سيتم استخدام مقدر القياس $MAD_\propto$ [3] في المقدر الحصين المقترح للمعلمة $\lambda$ وسيكون بالشكل:

$$\hat{\lambda}_\propto = MAD_\propto^*(x)/D \quad , \quad 0 < \propto < 0.5 \qquad (2.2.2)$$

حيث ان $D$ قيمة عددية اكبر من الصفر يتم ايجادها كما يلي:

بما انه $MAD_{\propto} = M_{\propto}\{|x - M_{\propto}|\}$ وباستخدام الصيغة (2.1.4) نحصل على :

$$F(M_{\propto}\{|x - M_{\propto}|\} = \propto$$

اذن:

$$P(|X - M_{\propto}| \leq M_{\propto}\{|x - M_{\propto}|\}) = \propto$$

اي ان:

$$P(|X - M_{\propto}| \leq MAD_{\propto}) = \propto$$

او:

$$P(-MAD_{\propto} \leq X - M_{\propto} \leq MAD_{\propto}) = \propto$$

بذلك نحصل على:

$$P(X \leq M_{\propto} + MAD_{\propto}) - P(X \leq \ M_{\propto} - MAD_{\propto}) = \propto$$

$$F(M_{\propto} + MAD_{\propto}) - F(M_{\propto} - MAD_{\propto}) = \propto$$

باستخدام الصيغ (2.2.1) ,(2.2.2) نجد ان:

$$F(-\lambda \ln(1-\propto) + \lambda D) - F(-\lambda \ln(1-\propto) - \lambda D) = \propto$$

$$1 - exp(\ln(1-\propto) - D) - 1 + exp(\ln(1-\propto) + D) = \propto$$

$$exp(\ln(1-\propto))\,(e^{D} - e^{-D}) = \propto$$

لذلك يكون:

$$e^{D} - e^{-D} = \frac{\alpha}{1-\alpha}\,, \qquad 0 <\propto< 0.5$$

وبحل المعادلة اعلاه ينتج:

$$D = ln\sqrt{1 + (\frac{\propto}{2(1-\propto)})^2} \quad + \frac{\alpha}{2(1-\alpha)} \qquad , \ D > 0 \qquad\qquad (2.2.3)$$

## ٣) **الجانب العملي:**

في هذا البحث تم دراسة اسلوب المحاكاه (Monte Carlo) من خلال فرض حالات مختلفة مشابهة للواقع العملي، حيث تم افتراض:

١. احجام مختلفة للعينة التي يتم دراستها ($n = 10, 50, 75$) وذلك لدراسة مدى تاثيرحجم العينة على كفاءة التقديرات الحصينة المقترحة ولقيم $\propto$ المختلفة .
٢. تم تطبيق اسلوب المحاكاه لقيم $\propto$ المختلفة ($\propto = 0.1, 0.25, 0.4$) وذلك لدراسة سلوك التقديرات الحصينة المقترحة عندما يتغير $\propto$ .
٣. توليد بيانات للاحجام المختلفة ولقيم $\lambda$ المختلفة ($\lambda = 1,2,3$) وذلك بتطبيق طريقة التحويل العكسي [4] والتي يمكن تطبيقها على النحو الاتي:

أ. ليكن R متغير عشوائي ذو توزيع منتظم $R \sim U(0,1)$ .
ب. استخدام المتغير R في التعبير عن داله التوزيع التراكمي للمتغير X كالاتي:

$$0 \le u = F(x) \le 1$$

$$u = 1 - \exp(-x/\lambda)$$

جـ. تحويل الصيغة الاخيرة لتكون بدلالة المتغير X بالشكل:

$$\exp(-x/\lambda) = 1 - u$$

وحيث ان $0 \le u \le 1$ فان $0 \le 1 - u \le 1$ اي ان كل من $1-u, u$ عبارة عن متغير عشوائي ذو توزيع منتظم (0,1) لذا فان:

$$\exp(-x/\lambda) = r$$

وبذلك سيكون

$$x = -\lambda \ln r \qquad (3.1)$$

حيث ان $r = 1 - u \quad , \quad -\lambda \ln r \sim \exp(\lambda)$

٤. ايجاد التقديرات الحصينة لمعلمة القياس $\lambda$ ( $MAD^*_{\propto}$ /$D$ ) ولقيم $\propto$ المختلفة ($\propto=$ 0.1, 0.25, 0.4) من خلال تطبيق الخطوات الاتية:

*a.* توليد القيم $r$ حيث ان $r \sim U(0,1)$

*b.* ترتيب القيم التي تم توليدها باستخدام الصيغة (3.1) تصاعديا.

*c.* استخدام الصيغة (2.1.1) لايجاد $M_{\propto}(x)$

*d.* ايجاد $y_i$ حيث ان : $y_i = |x_i - M_{\propto}(x)|$

*e.* *ترتيب* $y_i$ تصاعديا.

*f.* تطبيق الصيغة (٢.١.٢) على المتغير $y_i$ لايجاد $MAD^*_{\propto}$

*g.* ايجاد قيمة D *حسب الصيغة* (2.2.3)

٥. المقارنة بين التقديرات الحصينة المختلفة عند قيم $\propto$ المختلفة باستخدام متوسط مربعات الخطاء (MSE) حيث ان:

$$MSE(\hat{\lambda}) = (E(\hat{\lambda}) - \lambda)^2 + Var(\hat{\lambda})$$

$(E(\hat{\lambda}) - \lambda)$ يمثل التحيز biased

$Var(\hat{\lambda})$ يمثل تباين التقدير $\hat{\lambda}$

٦. تم تكرار التجربة لكل حالة من الحالات المدروسة (١٠٠٠) مرة .

٧. نفذ الجانب من البحث باستخدام الحاسوب من خلال برنامج (**Mathcad**) ، حيث تم استخدام الايعازات الاساسية التالية:

أ. الايعاز runif(n,a,b) لتوليد بيانات للاحجام المختلفة ولقيم $\lambda$ المختلفة ($\lambda = 1,2,3$) حيث يولد هذا الايعاز مصفوفة بيانات عشوائية احادية البعد بحجم n للمتغير العشوائي R (نقطة ١ اعلاه) وذات توزيع منتظم ومن خلال استخدام صيغة التحويل في اعلاه لتوليد قيم العنصر X ذو التوزيع الاسي .

ب. تم توليد حلقات داخلية inner loops لتوليد البيانات والمصفوفات حسب التكرار المحدد (١٠٠٠ تكرار) باستخدام الايعاز for loop وذكمجها ضمن مصفوفة ثنائية البعد لسهولة التعامل حيث بالامكان استدعاء اي عنصر من عناصر المصفوفة من خلال الامكانيات المتاحة من قبل بيئة البرنامج.

ت. تم التعامل مع المصفوفات ثنائية البعد المولدة لايجاد المعاملات الاحصائية المختلفة المذكورة اعلاه (النقط ٤، ٥) باستخدام الايعازات mean(x) لايجاد الوسط الحسابي، var(x) لايجاد التباين.

ث. تم ادراج هيكلية البرنامج والايعازات المستخدمة بالتفصيل في الملحق (أ) الخاص بالبحث.

## ٤) الاستنتاجات:

من خلال ملاحظة الجداول (١) ، (٢)، (٣) يمكن تلخيص النتائج الاتية:

١. اظهر التقدير الحصين المقترح ($MAD^*_{0.4}$ /0.327) افضل النتائج للتوزيع الاسي ولقيم $\lambda$ المختلفة ($\lambda = 1, 2, 3$) وعند قيم $n$ المختلفة (صغيرة، متوسطة ، كبيرة) من حيث انخفاض قيمة متوسط مربعات الخطأ (MSE) مقارنة بالتقديرات الاخرى المدروسة.

٢. ارتفاع كفاءة التقدير المقترح ($MAD^*_{0.4}$ /0.327) من خلال انخفاض (MSE) مقارنة بالتقدير(SD) التقليدي ولاحجام العينات المختلفة.

٣. يقل تاثير القيم الشاذة على التقدير ($MAD^*_{0.4}$ /0.327) من خلال انخفاض قيمة (MSE) له ولاحجام العينات المختلفة ولقيم $\lambda$ المختلفة.

جدول (١) :النتائج للتقديرات الحصينة المقترحة $MAD^*_{\propto}/D$ عندما $\lambda = 1$

| n | METHOD | ∝=0.1 | 0.25 | 0.4 | *SD* |
|---|---|---|---|---|---|
| | λ= 1 | | | | |
| 10 | *E($\hat{\lambda}$)* | 11.02 | 3.256 | 1.345 | 1.617 |
| | *Var($\hat{\lambda}$)* | 24.625 | 2.008 | 0.388 | 0.362 |
| | *biased* | 10.020 | 2.256 | 0.345 | 0.617 |
| | *MSE* | 125.036 | 7.099 | 0.506 | 0.742 |
| | | | | | |
| 50 | *E($\hat{\lambda}$)* | 12.018 | 3.221 | 1.453 | 1.857 |
| | *Var($\hat{\lambda}$)* | 5.549 | 0.362 | 0.072 | 0.078 |
| | *biased* | 11.018 | 2.221 | 0.453 | 0.278 |
| | *MSE* | 126.952 | 5.295 | 0.278 | 0.812 |
| | | | | | |
| 75 | *E($\hat{\lambda}$)* | 12.270 | 3.211 | 1.466 | 2.407 |
| | *Var($\hat{\lambda}$)* | 4.117 | 0.235 | 0.053 | 0.083 |
| | *biased* | 11.270 | 2.211 | 0.466 | 1.403 |
| | *MSE* | 131.130 | 5.123 | 0.270 | 2.063 |

جدول (2): النتائج للتقديرات الحصينة المقترحة $MAD^*_{\propto}/D$ عندما $\lambda = 2$

| n | λ= 2 | | | | |
|---|---|---|---|---|---|
| | METHOD | ∝=0.1 | 0.25 | 0.4 | *SD* |
| 10 | $E(\hat{\lambda})$ | 22.097 | 6.499 | 2.716 | 3.081 |
| | $Var(\hat{\lambda})$ | 91.029 | 7.780 | 1.729 | 1.899 |
| | *biased* | 20.097 | 4.499 | 0.716 | 1.081 |
| | *MSE* | 494.929 | 28.022 | 2.241 | 3.068 |
| | | | | | |
| 50 | $E(\hat{\lambda})$ | 23.925 | 6.481 | 2.937 | 3.732 |
| | $Var(\hat{\lambda})$ | 21.372 | 1.417 | 0.313 | 0.318 |
| | *biased* | 21.925 | 4.481 | 0.937 | 1.732 |
| | *MSE* | 502.087 | 21.492 | 1.192 | 3.316 |
| | | | | | |
| 75 | $E(\hat{\lambda})$ | 24.488 | 6.475 | 2.929 | 4.828 |
| | $Var(\hat{\lambda})$ | 17.408 | 0.957 | 0.216 | 0.346 |
| | *biased* | 22.488 | 4.475 | 0.929 | 2.828 |
| | *MSE* | 523.133 | 20.979 | 1.078 | 8.345 |

جدول (٣): النتائج للتقديرات الحصينة المقترحة $MAD^*_{\propto}/D$ عندما $\lambda = 3$

| n | λ= 3 | | | | |
|---|---|---|---|---|---|
| | METHOD | ∝=0.1 | 0.25 | 0.4 | *SD* |
| 10 | $E(\hat{\lambda})$ | 34.070 | 9.912 | 4.069 | 4.804 |
| | $Var(\hat{\lambda})$ | 235.397 | 18.115 | 3.344 | 3.535 |
| | *biased* | 31.070 | 6.912 | 1.069 | 1.804 |
| | *MSE* | 1200.744 | 65.889 | 4.486 | 6.789 |
| | | | | | |
| 50 | $E(\hat{\lambda})$ | 36.112 | 9.650 | 4.369 | 5.586 |
| | $Var(\hat{\lambda})$ | 54.402 | 3.261 | 0.721 | 0.757 |
| | *biased* | 33.112 | 6.650 | 1.369 | 2.586 |
| | *MSE* | 1150.820 | 47.481 | 2.596 | 7.444 |
| | | | | | |
| 75 | $E(\hat{\lambda})$ | 36.732 | 9.712 | 4.393 | 7.242 |
| | $Var(\hat{\lambda})$ | 39.168 | 2.154 | 0.485 | 0.779 |
| | *biased* | 33.732 | 6.712 | 1.393 | 4.242 |
| | *MSE* | 1177.05 | 47.202 | 2.426 | 18.777 |

## 5) **Reference**


1. Ricardo A. Maronna. R. Douglas Martin and Victor J. Yahai (2006) “Robust statistics – theory and methods”, John Wiley and Sons. LTD.
2. Croux, C. and Haesbroeck, G.(2002) “Maxbias Curves of robust location estimators based on subranges”, Nonparametric statistics, Vol. 14(3), PP.295-306.
3. David J. Olive, (2006) “ Robust estimators for transformed location scale families “, Southern Illinois University. Mailcode 4408 Carbondale. IL 62901-4408, USA.

٤. رشيد عبدالله هدى (٢٠٠٧) "استخدام تصميم التجارب في تحليل حساسية نماذج المحاكاه " اطروحة دكتوراه مقدمة الى كلية الادارة والاقتصاد / الجامعة المستنصرية.