

مقارنة بعض طرق تقدير معالم توزيع ويبل

*السعيد المهدى الطاهر و عبدالسلام كامل سليمان و نوري عمر علي

جميلة أبوالنيران و إيمان أبوغرارة

قسم الإحصاء - كلية العلوم - جامعة سبها، ليبيا

*للمراسلة: als.altaher@sebhau.edu.ly

الملخص تهدف هذه الدراسة إلى مقارنة الأداء العملي لثلاث طرق لتقدير معالم توزيع ويبل الاحتمالي للقيم العظمى، وهذه الطرق هي: طريقة الإمكان الأعظم و طريقة العزوم وطريقة المربعات الصغرى الاعتيادية، حيث تم استخدام أسلوب المحاكاة الإحصائية للمقارنة بين هذه الطرق في ظل اختلاف قيم معالم التوزيع الحقيقية وحجم العينة ، بغية الوصول إلى التقدير الأمثل ومحاولة تعميم النتائج. دلت النتائج على أن طريقة الإمكان الأعظم تبدو أكثر ملائمة لتقدير المعالم المجهولة في العينات الكبيرة. بينما طريقة العزوم والمربعات الصغرى الاعتيادية يمكن استخدامها في حالة العينات الصغيرة.

الكلمات المفتاحية: معالم توزيع وايبل، المحاكاة، الامكان الاعظم، العزوم ، المربعات الصغرى الاعتيادية.

A comparison of some methods for estimating the Weibull distribution parameters

*Alsaidi Altaher , Abdslam Suliman , Nuri Ali

Jamela Abo-Neran and Iman Abo-Grarah

Department of Statistics , faculty of Science, University of Sebha, Libya

*Corresponding author: als.altaher@sebhau.edu.ly

Abstract This study aims to compare the practical performance of three methods to estimate the parameters of Weibull distribution. These methods are: the method of maximum likelihood, the method of moments and the method of Ordinary least squares. Simulation experiments have been conducted for evaluation purposes under different real distribution parameters and different sample sizes. Results showed that the method of maximum likelihood seems to be more appropriate to estimate the unknown parameters for large sample size whereas the method of moments and the method of Ordinary least squares can be used for small sample sizes.

Keywords: Weibull distribution , simulation , maximum likelihood , moments , Ordinary Least Squares.

المقدمة

استخدامات أخرى. لذلك يجب أخذ الحيطة والحذر في اختيار أفضل الطرق لتقدير معالم النموذج وهنا تكمن أهمية هذه الدراسة.

توزيع ويبل ذو المعلمتين

إذا كان X متغير عشوائي يتبع توزيع ويبل بالمعلمتين α, β فإن دالة كثافة الاحتمال تأخذ الشكل التالي:

$$f(x; \alpha, \beta) = \frac{\beta}{\alpha^\beta} (x)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^\beta}; 0 < x < \infty \dots (1)$$

حيث α تسمى معلمة المقياس بينما β معلمة الشكل.

أما الدالة التوزيعية فتكون بالشكل

$$F(x) = \begin{cases} 1 - e^{-\left(\frac{x}{\alpha}\right)^\beta} & , \quad x > 0 \dots \dots (2) \\ 0 & , \quad x \leq 0 \end{cases}$$

يعتبر توزيع ويبل أحد التوزيعات الاحتمالية المستمرة واسعة الاستخدام حيث أن لهذا التوزيع مكانة وأهمية في حقل المعولية (Reliability) واختبار الحياة (Life testing) والرقابة على الجودة (Quantity control) ويعتبر أحد نماذج الفشل الشائعة الاستخدام فضلا عن استخداماته في التنبؤ في مجالات الأرصاد والأحوال الجوية. أنظر [1] و [2]. ويعتبر هذا التوزيع حالة خاصة من التوزيع العام للقيم المتطرفة عندما تكون معلمة الشكل أقل من الصفر.

تعد مسألة تقدير المعالم المجهولة لتوزيع ويبل من المسائل المهمة التي حظيت وتحظى باهتمام الباحثين والمهتمين بالإحصاء الرياضي نظراً إلى تطور طرائق التقدير وإيجاد مقدرات لهذه المعلمات أنظر [3] و [4] و [5] و [6]. ومما تجدر الإشارة إليه أنه إذا تم تقدير المعالم بشكل خاطئ أو غير دقيق فإن ذلك سوف يؤثر سلباً عند استخدام توزيع ويبل في عملية التنبؤ أو عند استخدامه للرقابة على الجودة أو أي

$$L = \ln(L(x_i; \alpha, \beta)) = n \ln \beta - n \ln \alpha + \sum_{i=1}^n \ln(x_i) - \left(\frac{\sum x_i}{\alpha}\right)^\beta$$

$$\frac{\partial \ln L}{\partial \beta} = \frac{n}{\beta} - n \ln \alpha + \sum_{i=1}^n \ln(x_i) - \frac{\sum_{i=1}^n x_i^\beta \ln(x_i) - \ln \alpha \sum x_i^\beta}{\alpha^\beta}$$

$$\frac{\partial \ln L}{\partial \beta} = 0 \text{ ويجعل}$$

لا يمكن حل المعادلة بالطرائق الاعتيادية وذلك بسبب ارتفاع درجة اللاخطية فيها لذلك تستخدم الطرق العددية لحل المعادلة غير الخطية.

$$\frac{\partial \ln L}{\partial \beta} = \frac{n\beta}{\alpha} + \frac{\beta}{\alpha^{\beta+1}} \sum_{i=1}^n x_i^\beta = 0$$

$$\alpha = \frac{1}{n} \left(\sum_{i=1}^n x_i^\beta\right)^{\frac{1}{\beta}} \dots\dots(10)$$

حيث أن معظم البرامج الإحصائية مزودة بدوال لإيجاد هذه التقديرات

Moments Method(MOM) طريقة العزوم

تعتبر طريقة العزوم من أقدم الطرق المستخدمة في تقدير معالم مجتمع معلوم التوزيع، حيث تستخدم العينة لتقدير معلمة المجتمع المجهولة، ولا نملك معلومات عن المعلمة إلا أنها ثابتة.

تعتمد طريقة العزوم على فرضية مساواة عزوم المجتمع μ_r مع عزوم العينة m_k ثم نتحصل على معادلات بالنسبة لمعالم المجتمع المجهولة ونحصل على التقديرات المطلوبة.

نوجد عزوم المجتمع حول الصفر

$$\mu'_r = E(x^r) = \alpha \left(1 + \frac{r}{\beta}\right) \dots\dots(11)$$

$$m_k = \frac{\sum_{i=1}^n x_i^k}{n} \quad i = 1, 2, \dots, n \quad \text{نوجد عزوم العينة}$$

نوجد التقدير وذلك بمساواة عزوم العينة بعزوم المجتمع

$$\alpha \left(1 + \frac{1}{\beta}\right) = \bar{x} \dots\dots(12)$$

$$\alpha^2 \left(1 + \frac{2}{\beta}\right) = \frac{1}{n} \sum_{i=1}^n x_i^2 \dots\dots(13)$$

بقسمة المعادلة (13) على مربع معادلة (12) نحصل على

$$\frac{\alpha^2 \left(1 + \frac{2}{\beta}\right)}{\left(\alpha \left(1 + \frac{1}{\beta}\right)\right)^2} = \frac{\frac{1}{n} \sum_{i=1}^n x_i^2}{\bar{x}^2}$$

العزم الرائي

$$M_x^r = E(x^r) = \int_0^\infty x^r \frac{\beta}{\alpha^\beta} x^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^\beta} dx \dots\dots(3)$$

التباين

$$V(x) = \alpha^2 \left[(2\beta^{-1} + 1) - \left(\beta^{-1} + 1\right)^2 \right] \dots\dots(4)$$

الوسيط

$$m = \alpha (Ln(2))^{\frac{1}{\beta}} \dots\dots(5)$$

المنوال

$$Mo = \alpha \left(\frac{\beta-1}{\beta}\right)^{\frac{1}{\alpha}} \quad \alpha > 1 \dots\dots(6)$$

الدالة المولدة للعزوم

$$M_x(t) = Ee^{tx} = \int_0^\infty e^{tx} f(x) d(x)$$

$$= \sum_{n=0}^\infty \frac{x^n \alpha^n}{n!} (1 + n/\beta), \beta \geq 1 \dots\dots(7)$$

الدالة المميزة

$$Q_x(t) = Ee^{itx} = \int_0^\infty e^{itx} f(x) d(x)$$

$$= \sum_{n=0}^\infty \frac{(it)^n \alpha^n}{n!} (1 + n/\beta), \beta \geq 1 \dots\dots(8)$$

طرق تقدير معالم توزيع ويبيل

من طرق التقدير لمعالم توزيع ويبيل طريقة الإمكان الأعظم و العزوم والمربعات الصغرى.

• طريقة الإمكان الأعظم Method Of Maximum Likelihood(MLE)

تستخدم دالة الإمكان الأعظم في تقدير المعلمات وهي دالة احتمال مشترك لـ n من المتغيرات العشوائية المستقلة x_1, x_2, \dots, x_n جميعها تتبع توزيع ويبيل ذي المعلمتين ويرمز لها بالرمز $L(x_i; \alpha, \beta)$. ولإيجاد المعالم المجهولة نقوم بمفاضلة دالة الإمكان الأعظم جزئياً بالنسبة لمعالم التوزيع (α, β) المراد تقديرها ومساواة المشتقة بالصفر ثم حل منظومة المعادلات الناتجة من عملية التفاضل وكما يلي:

$$L(x_i; \alpha, \beta) = \prod_{i=1}^n f(x_i; \alpha, \beta)$$

$$f(x_i; \alpha, \beta) = \frac{\beta}{\alpha^\beta} x_i^{\beta-1} e^{-\left(\frac{x_i}{\alpha}\right)^\beta}; i = 1, \dots, n$$

$$L(x_i; \alpha, \beta) = \left(\frac{\beta}{\alpha^\beta}\right)^n \prod_{i=1}^n x_i^{\beta-1} e^{-\left(\frac{x_i}{\alpha}\right)^\beta} \dots\dots(9)$$

ولما كانت دالة الإمكان الأعظم دالة غير تناقصية فإن القيمة التي تعظمها هي ذاتها التي تعظم لوغاريتمها

المقدرات $\hat{\alpha}, \hat{\beta}$ من المعلمات α, β

$$\hat{\beta} = \hat{\beta}_1$$

على أساس دراسة [7] يمكن أن تعطي أفضل تقدير .

ولما كانت القيمة المقدره لدالة التوزيع التراكمي باستخدام رتبة

$$\text{المتوسط تأخذ الشكل } \hat{F}(x) = \frac{1}{n+1} \text{ فإن}$$

$$\hat{\alpha} = \exp \left(- \frac{\sum_{i=1}^n \ln [-\ln (1 - \hat{F}(x_i - 1))] - \hat{\beta} \sum_{i=1}^n \ln x_i}{\hat{\beta} n} \right)$$

الجانب التطبيقي

• المحاكاة الإحصائية

في هذه الدراسة تم استخدام أسلوب المحاكاة باستخدام البرنامج الإحصائي R حيث تم توليد 1000 عينة عشوائية (عدد مرات تكرار التجربة يساوي 1000) مأخوذة من توزيع ويبل، ثم استخدمت هذه العينات في تقدير معلمي التوزيع باستخدام الطرق الثلاثة (الإمكان الأعظم-العزوم- المربعات الصغرى

الاعتيادية) وتم حساب $\hat{\beta}_1, \dots, \hat{\beta}_{1000}$ وهي

تقديرات لمعلمة الشكل β و $\hat{\alpha}_1, \dots, \hat{\alpha}_{1000}$ وهي تقديرات لمعلمة المقياس α .

يمكن حساب قيمة الجذر التربيعي لمتوسط مربعات الخطأ

من الصيغة

$$RMSE = \sqrt{\frac{1}{1000} \sum_{i=1}^{1000} \left[(\hat{\beta}_i - \beta)^2 + (\hat{\alpha}_i - \alpha)^2 \right]}$$

حيث

α : القيمة الافتراضية لمعلمة المقياس.

β : القيمة الافتراضية لمعلمة الشكل.

واعتمدت تجارب المحاكاة المنجزة على ركيزتين هما: الشكل الرياضي للنموذج وحجم العينة.

حيث تم اقتراح ثلاث نماذج بكل نموذج ثلاث خيارات مختلفة لقيم β, α كما تم توليد عينات بأحجام مختلفة

$$n=5, n=10, n=20, n=40, n=80, n=160, n=320$$

()

والنماذج موضحة بالجدول رقم 1.

كما بين [7] انه لا توجد صيغة رياضية محددة لإيجاد قيمة المعلمة وإنما يتم استخدام الطرق العددية بتقدير مبدئي.

$$\hat{\beta} = \left(\frac{\bar{x}}{\sqrt{s_x^2}} \right)^{1.086}$$

و نتحصل على $\hat{\alpha}$ من المعادلة (12)

$$\hat{\alpha} = \frac{\bar{x}}{1 + \frac{1}{\beta}}$$

• طريقة المربعات الصغرى الاعتيادية Ordinary Least Squares Method(OLS)

تعتمد طريقة المربعات الصغرى على تصغير مجموع مربعات الأخطاء $Q(\beta_0, \beta_1)$ لأقل ما يمكن. و يتم ذلك بالتفاضل الجزئي لمجموع مربعات الأخطاء بالنسبة لمعالم التوزيع (α, β) المراد تقديرها ومساواة المشتقة بالصفر. ولاستخدام هذا المفهوم سوف يتم التعامل مع دالة التوزيع وتحويلها إلى دالة خطية بأخذ لوغاريتمها مرتين وكما يلي:

$$\ln [-\ln (1 - F(x))] = \beta \ln x - \beta \ln \alpha$$

وبفرض

$$y = \ln [-\ln (1 - F(x))]$$

$$X = \ln x, \beta_1 = \beta \text{ and } \beta_0 = -\beta \ln \alpha$$

نتحصل على المعادلة الخطية التالية

$$y = \beta_1 X + \beta_0$$

وعليه يمكن كتابة مجموع مربعات الأخطاء بالشكل

$$Q(\beta_0, \beta_1) = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 \ln x_i)^2$$

توجد β_0 بالاشتقاق بالنسبة لـ β_0 ويساوي ناتجها بالصفر.

توجد β_1 بالاشتقاق بالنسبة لـ β_1 ويساوي ناتجها بالصفر،

حيث تكون الصيغ النهائية كما يلي:

$$\hat{\beta}_1 = \frac{n \sum_{i=1}^n \ln x_i \ln [-\ln (1 - \hat{F}(x))] - \sum_{i=1}^n \ln x_i \sum_{i=1}^n \ln [-\ln (1 - \hat{F}(x))]}{n \sum_{i=1}^n \ln^2 x_i - \left(\sum_{i=1}^n \ln x_i \right)^2}$$

$$\hat{\beta}_0 = \frac{1}{n} \sum_{i=1}^n \ln [-\ln (1 - \hat{F}(x))] - \hat{\beta}_1 \frac{1}{n} \sum_{i=1}^n \ln x_i$$

جدول 1: القيم المختارة لكل من α, β في النماذج الثلاث

النموذج	α	β	α	β
الأول عندما تكون $\alpha = \beta$	1.5	1.5	2	2
الثاني عندما تكون $\alpha > \beta$	2	1.5	3	1.5
الثالث عندما تكون $\alpha < \beta$	1.5	2	3	1.5

تم تلخيص نتائج تجربة المحاكاة في ثلاث جداول (جدول رقم 2,3,4)

جدول 2: قيم الجذر التربيعي لمتوسط مربعات الأخطاء للنموذج الأول عندما $\alpha = \beta$

$(\alpha = 1.5, \beta = 1.5)$							
Sample size	n=5	n=10	n=20	n=40	n=80	n=160	n=320
OLS	1.25372	0.5786057	0.3890187	0.2664391	0.1881570	0.1347265	0.10180786
MLE	1.35377	0.6684998	0.4189295	0.2709250	0.1851704	0.1258091	0.08815118
MOM	1.066293	0.6041243	0.4083334	0.2785726	0.1948610	0.1331146	0.09511848
$(\alpha = 2, \beta = 2)$							
Sample size	n=5	n=10	n=20	n=40	n=80	n=160	n=320
OLS	1.494216	0.7086802	0.5013929	0.3412484	0.2349292	0.1727216	0.1353269
MLE	2.041944	0.8240362	0.5437327	0.3270547	0.2134222	0.1511007	0.1043571
MOM	1.727030	0.7214114	0.5066986	0.3207617	0.2137686	0.1527734	0.1061556
$(\alpha = 3, \beta = 3)$							
Sample size	n=5	n=10	n=20	n=40	n=80	n=160	n=320
OLS	5.979296	1.035304	0.7014717	0.4849029	0.3728166	0.2836420	0.2273132
MLE	3.163430	1.258900	0.7041905	0.4366264	0.2947084	0.2044125	0.1457250
MOM	2.798804	1.103014	0.6588304	0.4178693	0.2885410	0.2024300	0.1440870

(OLS): Ordinary Least Square, (MLE): Method of Maximum Likelihood, (MOM): Moment method

جدول 3: قيم الجذر التربيعي لمتوسط مربعات الأخطاء للنموذج الثاني عندما $\alpha > \beta$

$(\alpha = 2, \beta = 1.5)$							
Sample size	n=5	n=10	n=20	n=40	n=80	n=160	n=320
OLS	2.242830	0.6629456	0.4277344	0.3028590	0.2169418	0.1566391	0.1168433
MLE	1.426903	0.7240611	0.4394645	0.3036341	0.2092124	0.1479246	0.1042642
MOM	1.218536	0.6682072	0.432744	0.3059127	0.2160542	0.1558408	0.1117576
$(\alpha = 3, \beta = 1.5)$							
Sample size	n=5	n=10	n=20	n=40	n=80	n=160	n=320
OLS	2.095469	0.8480828	0.5710179	0.3916320	0.2824100	0.2039095	0.1519014
MLE	1.580022	0.8569191	0.5775862	0.3874310	0.2698896	0.1923998	0.1334148
MOM	1.383011	0.8107386	0.5717190	0.3938284	0.2754932	0.1999351	0.1386853
$(\alpha = 3, \beta = 2)$							
Sample size	n=5	n=10	n=20	n=40	n=80	n=160	n=320
OLS	2.993450	0.8283708	0.5586029	0.3889933	0.2785819	0.2078228	0.1565096
MLE	1.947929	0.9291427	0.5680295	0.3697883	0.2582327	0.1775238	0.1249831
MOM	1.630419	0.8254948	0.5406679	0.3641797	0.2579495	0.1786948	0.1278773

(OLS): Ordinary Least Square, (MLE): Method of Maximum Likelihood, (MOM): Moment method

جدول 4: قيم الجذر التربيعي لمتوسط مربعات الأخطاء للنموذج الثالث عندما $\alpha < \beta$

$(\alpha = 1.5, \beta = 2)$							
Sample size	n=5	n=10	n=20	n=40	n=80	n=160	n=320
OLS	1.242002	0.6985435	0.4627116	0.3172034	0.2297467	0.1640064	0.12448404
MLE	1.806261	0.8347313	0.5040713	0.2974919	0.2081276	0.1414971	0.09516309
MOM	1.484603	0.7298035	0.4676182	0.2873300	0.2057035	0.1444059	0.09777457
$(\alpha = 1.5, \beta = 3)$							
Sample size	n=5	n=10	n=20	n=40	n=80	n=160	n=320
OLS	1.713758	0.9182383	0.6554770	0.4756952	0.3502980	0.2697818	0.2155671
MLE	2.734297	1.1625351	0.6622421	0.4407805	0.2726895	0.1900678	0.1332885
MOM	2.233775	0.9951178	0.6122496	0.4154747	0.2652205	0.1870457	0.1314570
$(\alpha = 2, \beta = 3)$							
Sample size	n=5	n=10	n=20	n=40	n=80	n=160	n=320
OLS	2.717854	0.9997282	0.6406848	0.4766599	0.3790559	0.2804748	0.2193144
MLE	2.814721	1.1891836	0.6807315	0.4262937	0.2986720	0.2004228	0.1391894
MOM	2.424556	1.0439613	0.6156232	0.4057976	0.2900616	0.1958315	0.1379172

(OLS): Ordinary Least Square, (MLE): Method of Maximum Likelihood, (MOM): Moment method

بعد تفحص هذه الجداول تم ملاحظة الآتي:

[3]- ضوية سلمان حسين و سميرة خليل إبراهيم و بيداء إسماعيل عبد الوهاب (2008)، "استخدام المقدار المخلص في تقدير معلمه الشكل لتوزيع وايبل"، مجلة جامعه النهريين-قسم الإحصاء-كلية الإدارة والاقتصاد-جامعة بغداد المجلد 11(3) كانون الأول، ص [73-78].

[4]- عبد الجبار خضر بخيث و سعد احمد عبد الرحمن النعمي و خيريه سلمان حسن (2011)، "مقارنة ثلاثة مقدرات مختلفة لمعلمه القياس لتوزيع وايبل ذي المعلمتين وقياس كفاءة المقدرات باستخدام المحاكاة"، المجلة العراقية للعلوم الإحصائية (20) ص [104-124].

[5]- I. Pobočíková, Z. Sedliačková, (2012) The least square and the weighted least square methods for estimating the Weibull distribution parameters - a comparative study, *Communications- Scientific Letters of the University of Žilina*, Vol. 14, no. 4 , 88-93.

[6]- M. Zaindin, A. M. Sarhan, Parameters Estimation of the Modified Weibull Distribution, (2009). *Applied Mathematical Sciences*, Vol. 3, no., 541-550

[7]- I. Pobočíková, Z. Sedliačková (2014). Comparison of Four Methods for Estimating the Weibull Distribution Parameters. *Applied Mathematical Sciences*, Vol. 8, no. 83, 4137 - 4149

1. كلما زادت حجم العينة كلما قلت قيمة جذر متوسط مربعات الأخطاء في جميع الطرق مع أحجام العينة المختلفة بما يتوافق مع نظرية الغاية المركزية.

2. كلما زادت كل من قيمة α و β معا أدى إلى تزايد في قيمة جذر متوسط مربعات الأخطاء وهذا ملاحظ في الطرق الثلاث مع اختلاف أحجام العينة.

3. عند تثبيت قيمة β وزيادة قيمة α أو بالعكس أدى أيضا إلى زيادة قيمة جذر متوسط مربعات الأخطاء.

4. لم يكن هناك سلوكا ثابتا من حيث كون أي الطرق أفضل فتارة تكون طريقة الإمكان الأعظم وأخرى تكون طريقة العزوم وأحيانا طريقة المربعات الصغرى هي الأفضل.

5. طريقة العزوم تبدو أنها أكثر ملائمة عندما يكون حجم العينة صغير (n=5 , n=10) ثم تأتي بعدها طريقة المربعات الصغرى وأخيرا طريقة الإمكان الأعظم

6. إذا كان حجم العينة متوسط (n=20 , n=40) تعتبر طريقة العزوم أفضل ثم طريقة الإمكان الأعظم ثم طريقة المربعات الصغرى.

7. طريقة الإمكان الأعظم تبدو أنها أفضل من طريقة العزوم والمربعات الصغرى عندما يكون حجم العينة كبير (n=80, n=160, n=320).

الخلاصة

تم تنفيذ تجارب محاكاة لغرض مقارنة أداء ثلاثة طرق لتقدير معالم توزيع وايبل ذي المعلمتين وهي طريقة الإمكان الأعظم، وطريقة المربعات الصغرى الاعتيادية، طريقة العزوم؛ حيث دلت النتائج على انه كلما زاد حجم العينة كلما اقتربت القيمة التقديرية لمعلمة الشكل والقياس من القيمة الحقيقية للمعلم كما أن طريقة العزوم والمربعات الصغرى تبدو أفضل للعينات ذوات الحجم الصغير أما طريقة الإمكان الأعظم فهي أفضل للعينات ذوات الحجم الكبير.

المراجع

[1]- مطانيوس مغول و عدنان غانم (2011)، "فعالية استخدام توزيع وايبل الاحتمالي في التنبؤ"، مجلة جامعة-دمشق للعلوم الاقتصادية والقانونية، المجلد 27-العدد الرابع.

[2]- صادق مولي جعفر و بيداء إسماعيل عبد الوهاب و انتصار عبيد حسون (2009)، "أفضل تقدير لمعوليه توزيع وايبل ذي معلمتين"، مجلة بغداد للعلوم مجلد 6(4) 2009.