المقارنة بين طريقتين لاختبار مطابقة بيانات لتوزيع ويبل ذي المعلمتين

د. مكي أكرم محمد صالح
علي إبراهيم منصور
مهدي صالح نايف
كلية التربية – جامعة بغداد

الملخص: استخدم أسلوب المحاكاة لتقدير معلمتي توزيع ويبل (معلمة الشكل ومعلمة القياس) ،وذلك من خلال تطبيق طريقة المربعات الصغرى (Least Square method) وباخذ عدة قيم حقيقية لمعلمة الشكل $(\beta = 5,10,15)$ ومعلمة القياس $(\alpha = 10)$ وتم توليد بيانات حقيقية بحجوم عينة مختلفة (n=10,20,30,40) ، وتوليد بيانات تقديرية بالاعتماد على قيم مقدرات المعلمتين وقورنت حسن مطابقة مجموعتي البيانات الحقيقة والبيانات التقديرية لتوزيع ويبل ذي المعلمتين باستخدام طريقتين مختلفتين وهما:-

- طريقة الرسم البياني (Graphical Method)
- طريقة الزمن الكلي في الاختبار (Total Time on Test)

وقد ظهرت النتائج بان طريقة الرسم البياني افضل من طريقة الزمن الكلي في الاختبار لبيان مدى اتباع بيانات لتوزيع ويبل ذي المعلمتين .

1- المقدمة:يعد توزيع ويبل من التوزيعات المستخدمة بشكل واسع كنموذج لانواع الفشل نتيجة للاستهلاك او الاعياء والكلل[6]، فاذا كان T متغيراً عشوائياً يتبع توزيع ويبل ذي المعلمتين فان صيغة دالتاه الاحتمالية والتوزيعية هي على الترتيب :

$$f(t;\alpha,\beta)=\frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1}\exp\left(-\left(\frac{t}{\alpha}\right)^{\beta}\right)$$

$$F(t;\alpha,\beta)=1-\exp\left(-\left(\frac{t}{\alpha}\right)^{\beta}\right)$$

حيث تمثل :-

f الدالة الاحتمالية للتوزيع

F الدالة التوزيعية للتوزيع

t قيمة المتغير العشوائي T ويمثل زمن حدوث الفشل

α معلمة القياس (Scale parameter)

β معلمة الشكل (Shape parameter)

ليكن Y متغيراً عشوائيا يتبع توزيع القيمة المتطرفة ، فأن صيغة دالتاه الاحتمالية والتوزيعية هي على الترتيب[7]:-

$$g(y;\theta,\zeta)=\frac{1}{\theta}\exp\left(\frac{y-\zeta}{\theta}\right)\exp\left(-\exp\left(\frac{y-\zeta}{\theta}\right)\right)$$

$$G(y;\theta,\zeta)=1-\exp\left(-\exp\left(\frac{y-\zeta}{\theta}\right)\right) \qquad \ldots\ (1)$$

حيث يمثل:-

y قيمة المتغير العشوائي Y وان $-\infty < y < \infty$

θ معلمة القياس (Scale parameter)

ζ معلمة الموقع (Location parameter)

وان العلاقة بين التوزيعين اعلاه هي :-

$$Y=LogT$$

$$Y=Lnt\ ,\ \theta=\frac{1}{\beta}\ ,\ \zeta=Ln\alpha$$

2- طريقة المربعات الصغرى (LS)

تم اعتماد اسلوبي (Pazren) (1979) [4]،و(Kindermenn, Lariccia) (1983) [3] لتحويل النموذج غير الخطي لنموذج خطي في تطبيق طريقة (LS) وذلك باخذ معكوس الدالة التوزيعية.

من الصيغة (1) فان الدالة المئينية لتوزيع القيمة المتطرفة هي:-

$$y(G)=\zeta+\theta(Ln(-Ln(1-G))) \qquad \ldots\ (2)$$

$$Z=Ln(-Ln(1-G))$$

بالتعبير عن قيمة G بقيمة P حيث ان:-

$$P_i = \frac{i}{n+1} \qquad i = 1,2,\ldots,n$$

فان الصيغة (2) تصبح كالأتي :-

$$y(P_i) = \zeta + \theta Z_i \qquad \ldots (3)$$

وبتفاضل الصيغة (3) لكل من ζ و θ على التوالي نحصل على:-

$$\frac{\partial y}{\partial \zeta} = 1 \qquad \ldots (4)$$

$$\frac{\partial y}{\partial \theta} = 1 \qquad \ldots (5)$$

وباستخدام الرتبة الاولى لسلسة تيلر للصيغ (3) ،(4) و(5) نحصل على الاتي

$$E(y_i) = \zeta_\circ + \theta_\circ Z_i + (\hat{\zeta} - \zeta_\circ) + (\hat{\theta} - \theta_\circ) Z_i$$

$$E(y_i) = \hat{\zeta} + \hat{\theta} Z_i \qquad \ldots (6)$$

ليكن $Q(P_i) = Z_i$

فان الصيغة (6) يمكن كتابتها بالشكل الاتي :-

$$E(y_i) = \hat{\zeta} + \hat{\theta} Q(P_i)$$

لتكن K مصفوفة من الرتبة $n \times 2$ ومعرفة بالشكل الاتي:-

$$K = \begin{bmatrix} 1 & Q(P_i) \\ 1 & Q(P_i) \\ . & . \\ . & . \\ . & . \\ 1 & Q(P_n) \end{bmatrix}$$

فان استخراج القيم التقديرية لـ θ ، ζ يمكن ان يستحصل من الصيغة الاتية:-

$$\lambda^T = (K^T K)^{-1} K^T Y_i$$

حيث ان:-

$$\lambda = \left(\hat{\zeta} \quad \hat{\theta}\right)_{1\times 1}$$

وتمثل كل من:

$\hat{\zeta}$ القيمة التقديرية لمعلمة الموقع

$\hat{\theta}$ القيمة التقديرية لمعلمة القياس

Y_i قيم المشاهدات حيث ان

$$Y_i = \left[y_1 y_2 \ldots y_n\right]^T$$

ومن $\hat{\zeta}$ و $\hat{\theta}$ يمكننا الحصول على القيم التقديرية لمعلمتي توزيع ويبل α و β وعلى النحو الاتي:

$$\hat{\alpha} = \exp\left(\hat{\zeta}\right) \qquad \ldots(7)$$

$$\hat{\beta} = \frac{1}{\hat{\theta}} \qquad ...(8)$$

3-) طرائق اختبار حسن مطابقة البيانات Goodness of fit Methods (

1-3 طريقة الرسم البياني:

طبقت هذه الطريقة لاختبار حسن مطابقة البيانات لتوزيع ويبل ذي المعلمتين ، لتكن

$t_1, t_2, \dots, t_n$ عينة مرتبة تصاعديا تمثل البيانات الحقيقية التي تتبع توزيع ويبل ذا المعلمتين

اولاً:-

تحديد قيم t_i مقابل قيم P_i لكل i حيث ان : $i=1,2,\dots,n$

$$P_i = \frac{i - 0.5}{n}$$

ومن خلال انتشار النقاط الناتجة من الازواج المرتبة (P_i، t_i) ينتج رسماً بيانياً يتبع توزيع ويبل ذو المعلمتين .

ثانياً:-

1- تقدير معلمتي توزيع ويبل α و β

2- توليد بيانات مقدرة $\hat{t}_i$ من القيم التقديرية لـ α و β .

3- رسم قيم $\hat{t}_i$ مقابل قيم P_i لكل i.

ومن خلال انتشار النقاط الناتجة من الازواج المرتبة (P_i، $\hat{t}_i$) ينتج رسمياً بيانياً تابعاً لتوزيع ويبل ذي المعلمتين.

ويتم الاختبار من خلال ملاحظة مدى مطابقة الرسمين البيانيين الناتجين (كما موضح في الاشكال 1,2,3,4) ، فان تطابق الرسمين البيانيين كليا (شبه كلي) تكون البيانات تتبع توزيع ويبل وبخلافه فان البيانات لاتتبع هذا التوزيع.

2-3 طريقة الزمن الكلي في الاختبار (TTT)

ان التمثيل البياني باستخدام طريقة (TTT) خطوة فعالة في تعيين مدى ملائمة البيانات المعطاة لتوزيع ويبل ذي المعلمتين . ان طريقة (TTT) تطبق فقط على التوزيع الاسي، فاذا كان X متغيراً عشوائياً يتبع التوزيع الاسي فيمكن تعريف T_i بالشكل الاتي :

$$T_i = \sum_{j=1}^{i} (n - j + 1)(x_j - x_{j-1})$$

ليكن

$$C_i = \frac{T_i}{T_n}$$

فان تحديد الازواج المرتبة $\left(\frac{i}{n}, C_i\right)$ يجب ان يقع على (بالقرب من) الخط المستقيم الواصل بين النقطتين (0,0) و (1,1) في شكل مربع الوحدة.

طبقت طريقة (TTT) لاختبار حسن مطابقة البيانات لتوزيع ويبل ذي المعلمتين وكما ياتي:-

ليكن t قيمة للمتغير العشوائي T الذي يتبع توزيع ويبل ذا المعلمتين فان الصيغة

$$Q(t_i) = t_i^{\beta}$$

ستحول المتغير T الذي يتبع توزيع ويبل الى المتغير $Q(T_i)$ الذي يتبع التوزيع الاسي ، فان التمثيل البياني يكون بحساب قيم $\frac{Q_i}{Q_n}$ لكل i ، وبعد ذلك تحدد هذه القيم مقابل قيم $\frac{i}{n}$ لكل i ، حيث ان:

$$Q_i = \sum_{j=1}^{i}(n-j+1)\{Q(t_j)-Q(t_{j-1})\} \qquad i=1,2,\dots,n$$

ينتج رسماً بيانياً يقع على (بالقرب من) الخط المستقيم في مربع الوحدة.

وبأخذ قيمة $\hat{\beta}$ ثم حساب قيم $\frac{\hat{Q}_i}{\hat{Q}_n}$ لكل i ، وتحديد هذه القيم مقابل قيم $\frac{i}{n}$ ينتج رسماً بيانياً اخر داخل مربع الوحدة حيث ان :

$$\hat{Q}_i = \sum_{j=1}^{i}(n-j+1)\{\hat{Q}(t_j)-\hat{Q}(t_{j-1})\} \qquad i=1,2,\dots,n$$

فاذا تطابق الرسمين في (كما موضحة في الاشكال5,6,7,8)، تكون البيانات تتبع توزيع ويبل ذا المعلمتين .

4- المحاكاة Simulation

تم استخدام اسلوب المحاكاة لتوليد بيانات باعتماد اسلوب F. Sayler [5] (1977) ، حيث تم توليد بيانات حقيقية تتبع توزيع ويبل ذا المعلمتين من قيم حقيقية متعددة للمعلمتين α وβ وحجوم عينات مختلفة كما مبنية في الجدول (1) ، ومن خلال هذه البيانات تم تقدير معلمتين توزيع ويبل بطريقة المربعات الصغرى (LS) وباعتماد القيم التقديرية لمعلمتي توزيع ويبل تم توليد بيانات تقديرية جديدة تتبع هذا التوزيع، وتمت مقارنة تطابق هاتين المجموعتين من البيانات بطريقتي الرسم البياني والزمن الكلي في الاختبار.

α	β	Sample Size			
10	5 10 15	10	20	30	40

جدول (1) يبين القيم الحقيقية لمعلمتي توزيع وبيل وحجوم العينات

5- النتائج Results

5-1 نتائج تقدير المعلمتين

بعد تطبيق طريقة (LS) على البيانات المولدة لتقدير معلمتي توزيع ويبل وباستخدام كل من الصيغتين (6) و (7) ظهرت نتائج تقدير معلمتي الشكل القياس كما في الجداول (2)،(3) و(4)

	n			
	10	20	30	40
$\hat{\alpha}$	10.239	10.768	10.437	10.157
$\hat{\beta}$	3.952	5.105	5.335	5.501

جدول (2)

يبين القيم التقديرية لمعلمتي توزيع ويبل عندما $(\alpha = 10, \beta = 5)$

sampl / Est.	n			
	10	20	30	40

$\hat{\alpha}$	10.119	10.377	10.216	10.078
$\hat{\beta}$	7.904	10.210	10.670	10.002

جدول (3)

يبين القيم التقديرية لمعلمتي توزيع ويبل عندما $(\alpha = 10, \beta = 10)$

	sample n			
Est.	10	20	30	40
$\hat{\alpha}$	10.079	10.250	10.144	10.052
$\hat{\beta}$	11.856	15.315	16.004	16.503

جدول (4)

يبين القيم التقديرية لمعلمتي توزيع ويبل عندما $(\alpha = 10, \beta = 15)$

2-5 نتائج الرسوم البيانية :

بعد تقدير معلمتي توزيع ويبل تم توليد بيانات تقديرية والممثلة في الاشكال (1,2,3,4,5,6,7,8) بالرمز ◘ ،اما البيانات الحقيقية فمتمثلة بالرمز * . وكانت النتائج كما في الاشكال الاتية:

6- الاستنتاجات:

استحصلت الرسوم البيانية المبينة في الاشكال (1)،(2) ،(3) و(4) من خلال طريقة الرسم البياني ، كما استحصلت الرسوم البيانية المبينة في الاشكال (5)،(6) ،(7) و(8) بواسطة طريقة (TTT) ، وقد استنتج الاتي:-

بمقارن الشكل (1) مع الشكل (5) يتبين ان طريقة (TTT) هي افضل من طريقة الرسم البياني ،اما عند مقارنة الاشكال (2),(3),(4) مع الاشكال (6),(7),(8) على الترتيب يتبين بان تطابق البيانات التقديرية المولدة مع البيانات الحقيقية المولدة في طريقة الرسم البياني هي افضل من طريقة (TTT) في معظم الحالات المأخوذة ، مما يدل على ان استخدام طريقة الرسم البياتي لبيان مدى ملائمة بيانات لتوزيع ويبل ذي المعلمتين هي افضل من طريقة الزمن الكلي في الاختبار .

المصادر:

1. **Barlow, R.E. and Campo, R,(1975) "Total Time on Test Processes and Application to Failure Data Analysis ".In .Reliability and Fault Tree Analysis, SIAM. Philadelphia.**

2. D Agostion, R.B .and Stephens, M.A., (1986)."Goodness of Fit Techniques". Marcel Dakkre, Inc., New Yourk.
3. LaRiccia, V.N. and Kindermenn, R.P., (1983)."An Asymptotically Efficient Closed Form Estimator for the Three – Parameter Log-Normal Distribution ". Communications in Statistics Theory and Methods.12, 243-261.
4. Parzen, E.C. (1979). "Nonparametric Statistical Data Modeling" Jour., Amer. Statistic. Assoc., 74.
5. Sayler, F., (1977)."A Monte –Cairo Comparison of the Method of Moments the Maximum-Likelihood Estimates of W.P. For Cos Data ".In. The Theory and Applications of Reliability.Vol.2.
6. Weibull, W. (1993)."A statistical Theory of the Strength of Materials". Ingenious Vetens kaps Academia-Handilingar, No.151 Generalstabens Litografiska Anstalts, Furlag,Stocholm ,Sweden.
7. White, J.S., (1995)."Linear Estimation for Log –Weibull Distribution ".Annals of Reliability and Maintainability.4, 915-923.

Comparison between the two methods to test for the distribution of data matching Whipple with two parameters

Dr. Akram M..S. Meki
Ali. I. Mansour
Mahdi .S. Nayef
college of Education
University of Baghdad

Abstract:Use of Job Simulations to assess my teacher distribution and Whipple (teacher form and a parameter measurement), and through the application of the least squares method (Least Square method) and took several real values for the parameter form and a parameter measurement has been generating real data sizes in a different sample (n = 10,20, 30.40), and generate estimated data based on the values of parameters and compared the capabilities of good matching data sets truth and estimated data for the distribution of all parameters and reported using two different methods, namely: -

- How to graph (Graphical Method)
- How the total time in the test (Total Time on Test)

The results appeared that the way the graph is better than the way the total time in the test to demonstrate the extent to which the data for the distribution of Whipple with two parameters