تعريف التوزيع الاسي العام وتقدير معلماته باعتماد ثلاث طرائق وبواسطة المحاكاة

م . م اياد رفعت رؤوف
كلية الادارة والاقتصاد / جامعة بغداد
مركز الحاسبة الالكترونية

المستخلص :

يتناول البحث تعريف التوزيع الاسي العام ذي المعلمتين (α ، λ)، والذي ادخل كتوزيع بديل لتوزيع كاما وتوزيع ويبل واللوغارتمي الطبيعي في الكثير من توزيعات وقت الحياة للوحدات المنتجة، وهو توزيع ذو التواء الى جهة اليمين، وله منوال وحيد. وقد تضمن البحث المقارنة بين ثلاث طرائق لتقدير معلمتي القياس λ والشكل α ، هما الامكان الاعظم، العزوم، والمربعات الصغرى بواسطة المحاكاة، حيث نفذت المحاكاة عند حجوم عينات مختلفة وباعتماد قيم اولية مختلفة وعرضت النتائج في جداول خاصة.

Abstract

The two parameters Generalized Exponential distribution is a particular member of the three-parameter Exponential Weibull distribution. The $GE(\alpha, \lambda)$ has a right skewed uni model density function and monotone Hazard function similar to the density functions and Hazard function of Gamma and Weibull distribution. It is observed that the GE (α, λ) can be used quite effectively to analyze life time data in place of Gamma, Weibull and Log-Normal distribution. Here we present three methods for estimating α, λ these are maximum likelihood, and moment and least squares estimators. The comparison between them is done by simulation all the results are explained in tables.

Keywords: Generalized Exponential Distribution, MLE, MME, LSE Density Function, Hazard Function MSE.

المقدمة

يعتبرالتوزيع الاسي العام ذي المعلمتين، حالة خاصة من توزيع (ويبل–الاسي) ذي الثلاث معلمات. وان التوزيع الاسي العام والذي سوف نرمز له GE(λ , α) هو دالة احتمالية لها منوال وحيد و ذات التواء الى جهة اليمين، اما دالة المخاطرة له فهي منتظمة،

ويستخدم لتحليل البيانات المتعلقة بأوقات الحياة للوحدات بدلاً من توزيع كاما، وبدلاً من توزيع ويبل والتوزيع الطبيعي اللوغاريتمي، وقد قدرت معلماته بطرائق مختلفة. وسوف نحاول التعرف على بعضها تحقيقاً لهدف البحث.

هدف البحث

يهدف البحث الى التعرف على GE(α , λ) وكيفية تقدير معلماته القياس والشكل بطريقة الامكان الاعظم والعزوم والمربعات الصغرى الاعتيادية، وسوف نعتمد على المحاكاة في المقارنة بين المقدرات لعدة حجوم عينات مختلفة، ولعدة قيم معلمات، واعتماد مقياس متوسط مربعات الخطأ (MSE) للمقارنة بين المقدرات، وسوف تنفذ تجارب بالمحاكاة لحجوم عينات (N=10, 20, 50, 100)، وتكرار كل تجربة R=500.

الجانب النظري

ليكن x متغير عشوائي يتبع التوزيع GE(λ, α) (حيث ان α تمثل معلمة الشكل، λ معلمة القياس) ذو دالة الكثافة الاحتمالية التالية:

$$f(x,\alpha,\lambda) = \alpha\lambda(1-e^{-\lambda x})^{\alpha-1} e^{-\lambda x} \qquad x>0 \qquad .....(1)$$

$$0 \qquad o/w$$

وان الدالة الاحتمالية التجميعية c.d.f هي:

$$F(x,\alpha,\lambda) = (1-e^{-\lambda x})^{\alpha} \qquad x>0 \quad \alpha,\lambda>0 \quad ..... (2)$$

وبالامكان اشتقاق الدالة المولدة للعزوم لهذا التوزيع من تطبيق التعريف:

$$M_x(t) = E(e^{tx})$$

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

$$= \int_0^{\infty} e^{tx} \alpha\lambda(1-e^{-\lambda x})^{\alpha-1} e^{-\lambda x} dx$$

$$= \alpha\lambda \int_0^{\infty} e^{-x(\lambda-t)} (1-e^{-\lambda x})^{\alpha-1} dx \qquad ..... (3)$$

وبادخال التحويلات التالية على المعادلة (3) نجد ان:

لتكن

$$Z = 1 - e^{-\lambda x}$$

ومنها

$$e^{-\lambda x} = 1 - Z \longrightarrow -\lambda e^{-\lambda x}\, dx = -\, dz$$

وعندما x=0 فان z=0،

وعندما $\infty \longrightarrow x$ فان $1 \longrightarrow z$، وعليه فان:

$$\because Z = 1 - e^{-\lambda x}$$

$$-\lambda x = \ln(1-z)$$

$$-x = \frac{1}{\lambda} \ln(1-z)$$

$$x = \ln(1-z)^{-\frac{1}{\lambda}}$$

$$x\, t = \ln(1-z)^{-\frac{t}{\lambda}}$$

وعندئذ تصبح المعادلة (3):

$$M_x(t) = \alpha \int_0^{\infty} e^{tx} (1 - e^{-\lambda x})^{\alpha-1} \lambda e^{-\lambda x}\, dx$$

$$= \alpha \int_0^{1} (1-z)^{-\frac{t}{\lambda}} z^{\alpha-1}\, dz \qquad \ldots\ldots (4)$$

وباعتماد قاعدة بيتا للتكامل تصبح الدالة المولدة للعزوم هي:

$$M_x(t) = \alpha\ beta(\alpha, 1 - \frac{t}{\lambda})$$

$$= \alpha \frac{\Gamma(\alpha)\,\Gamma(1-\frac{t}{\lambda})}{\Gamma(\alpha+1-\frac{t}{\lambda})} = \frac{\Gamma(\alpha+1)\,\Gamma(1-\frac{t}{\lambda})}{\Gamma(\alpha+1-\frac{t}{\lambda})} \qquad \ldots\ldots (5)$$

حيث ان $t < \lambda$.

ومن المعادلة (5) نجد ان:

$$\mu = E(x) = \frac{1}{\lambda}[\psi(\alpha+1) - \psi(1)] \quad \ldots\ldots (6)$$

$$\sigma^2 = V(x) = \frac{1}{\lambda^2}\left[\Psi'(1) - \Psi'(\alpha+1)\right] \quad \ldots\ldots (7)$$

وتعرف الدالة $\Psi(\alpha) = \frac{d}{d\alpha}\Gamma(\alpha)$ بـ digamma.

اما معامل الالتواء (skewness) للتوزيع الاسي العام فهو يعتمد على العزم الثالث والثاني، ويعرف بالمقدار β_1 حيث ان:

$$\sqrt{\beta_1} = \frac{M_3}{M_2^{3/2}}$$

وكذلك معامل التفلطح (Kurtosis) يعتمد على العزم الثاني والرابع، ويعرف بـ β_2 :

$$\beta_2 = \frac{M_4}{M_2^2}$$

$$M_2 = \frac{1}{\lambda^2}\left[\Psi'(1) - \Psi'(\alpha+1) + (\Psi(\alpha+1) - \Psi(1))^2\right] \quad \ldots\ldots (8)$$

ويمكن استخراج العزوم الاخرى M_3 , M_4 .

اما دالة المخاطرة (Hazard) للتوزيع الاسي العام فهي:

$$h(x,\alpha,\lambda) = \frac{f(x,\alpha,\lambda)}{1-F(x,\alpha,\lambda)} = \frac{\alpha\lambda e^{-\lambda}(1-e^{-\lambda x})^{\alpha-1}}{1-(1-e^{-\lambda x})^{\alpha}} \quad \ldots\ldots (9).$$

وحيث ان λ هي معلمة القياس، فان معلمة الشكل لدالة المخاطرة لاتعتمد على λ ، بل فقط على α ، ولأي قيمة ثابتة الى λ فان المعادلة (9) تكون متزايدة عندما ($\lambda > 1$)، ومتناقصة عندما ($\alpha < 1$)، وعندما ($\alpha = 1$) فهي ثابتة، وهي مقعرة لوغارتمياً عندما ($\alpha > 1$)، ومحدبة

لوغارتمياً عندما ($\alpha \leq 1$)، وهي كما هو الحال في دالة المخاطرة لتوزيع ويبل او كاما، فهي صعبة في الحصول على مصفوفة معلومات فيشر المتعلقة بالتباين، لذلك سوف نعتمد على دالة المخاطرة النسبية والمعرفة بالمعادلة (10) ادناه:

$$r(x,\alpha,\lambda) = \frac{f(x,\alpha,\lambda)}{F(x,\alpha,\lambda)} = \frac{\alpha \lambda e^{-\lambda x}}{1-e^{-\lambda x}} \qquad \ldots\ldots (10)$$

ولجميع قيم α فهي دالة متناقصة بالنسبة لـ x. وتعتمد المعادلة (10) في حساب مصفوفة فيشر للمعلومات Fisher Information Matrix حول حول المعلمات غير المعلومةunknown parameters.

بعض طرائق تقدير معلمات التوزيع الاسي العام :

1- طريقة الامكان الاعظم

تعتمد هذه الطريقة في ايجاد مقدري الامكان الاعظم للمعلمتين (α , λ) على تعظيم لوغاريتم دالة الامكان الاعظم المعرف بالمعادلة (11):

$$LnL(\alpha,\lambda) = n\,Ln(\alpha) + nLn(\lambda) + (\alpha-1)\sum_{i=1}^{n} Ln(1-e^{-\lambda x_i}) - \lambda \sum_{i=1}^{n} x_i \qquad \ldots\ldots (11)$$

وعند اشتقاق المعادلة (11) بالنسبة لـ λ , α نحصل على:

$$\frac{\partial LnL}{\partial \lambda} = \frac{n}{\lambda} + (\alpha-1)\sum_{i=1}^{n} \frac{x_i e^{-\lambda x_i}}{(1-e^{-\lambda x_i})} - \sum_{i=1}^{n} x_i$$

$$\frac{\partial LnL}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n} Ln(1-e^{-\lambda x_i})$$

$$\hat{\alpha}(\lambda) = -\frac{n}{\sum_{i=1}^{n} Ln(1-e^{-\lambda x_i})} \qquad \ldots\ldots (12)$$

وعند اعادة تعويض المعادلة (12) في المعادلة (5) نحصل على معادلة تسمى:

$$g(\lambda) = L(\hat{\alpha}(\lambda), L)$$

$$= K- n\,Ln\sum_{i=1}^{n}(-Ln(1-e^{-\lambda xi})+ n\,Ln(\lambda)-\sum_{i=1}^{n}Ln(1-e^{-\lambda xi})-\lambda\sum_{i=1}^{n}xi \quad \ldots\ldots (13)$$

وبتطبيق طريقة النقطة الصامدة على المعادلة (13) نحصل :

$$\lambda = h(\lambda)$$

الى ان تتقارب النتائج في تكرارين متتاليين.

علماً بان:

$$h(\lambda)=\left[\frac{\sum_{i=1}^{n}\left(\frac{x_i e^{-\lambda x_i}}{(1-e^{-\lambda x_i})}\right)}{\sum_{i=1}^{n}Ln(1-e^{-\lambda x_i})}+\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{(1-e^{-\lambda x_i})}\right)\right]^{-1} \quad \ldots\ldots (14)$$

وبعد ايجاد مقدار $\hat{\lambda}_{MLE}$ ، نعوض لأيجاد $\hat{\alpha}_{MLE}$.

2- مقدرات العزوم MLI

تستخرج مقدرات العزوم لأي توزيع من مساواة عزوم العينة مع عزوم المجتمع، ثم حل المعادلات الناتجة، وتعرف عزوم العينة من:

$$m_r=\sum_{i=1}^{n}\frac{x_i^r}{n}$$

اما عزوم المجتمع فتعرف من $\mu_r = E(x^r)$ ، وبالامكان الحصول على العزوم المختلفة μ_r من اشتقاق الدالة المولدة للعزوم للتوزيع $\alpha\,GE(,\lambda)$ ، والتي تم استخراجها في المعادلة (5)، حيث ان:

$$\mu = E(x) = \frac{1}{\lambda}[\,\psi(\alpha+1)-\psi(1)] \quad \ldots\ldots (15)$$

ومن المعلوم ان معامل الاختلاف (C.V) مستقل عن معلمة القياس (λ) لذلك عند مساواة (C.V) للعينة مع (C.V) للمجتمع نحصل على المعادلة غير الخطية (16):

$$\frac{S}{\overline{X}} = \frac{\sqrt{\psi'(1) - \psi'(\alpha+1)}}{\psi(\alpha+1) - \psi(1)} \qquad \ldots\ldots (16)$$

علماً بأن:

$$\overline{X} = \sum_{i=1}^{n} xi / n$$

$$S^2 = \sum_{i=1}^{n} (X_i - \overline{X})^2 / (n-1)$$

$$\psi(\alpha) = \frac{d\,\Gamma(\alpha)}{d(\alpha)}$$

وتسمى digamma.

$$\psi'(\alpha) = \frac{d}{d\alpha}\psi(\alpha)$$

وعند حل المعادلة (16) الغير خطية بأية طريقة من طرق التحليل العددي مثل نيوتن رافسن او النقطة الصامدة نحصل على مقدار العزوم للمعلمة α وهو $\hat{\alpha}_{MME}$، ثم يعوض هذا المقدار للحصول على مقدار $\hat{\lambda}_{MME}$ وهو:

$$\hat{\lambda}_{MME}(\alpha \text{ ثابت}) = \frac{\psi(\alpha+1) - \psi(1)}{\overline{X}} \qquad \ldots\ldots (17)$$

3- طريقة المربعات الصغرى Least Squares Estimators

تعتمد طريقة المربعات الصغرى على توزيع القيم المرتبة لمفردات العينة وهو ($G(y_{(j)})$، فاذا كانت لدينا عينة عشوائية حجمها n ($y_1, y_2, \ldots, y_n$) من توزع له c.d.f معلومة، فان:

$$E[G(y_{(j)})] = \frac{j}{n+1}$$

$$V[G(y_{(j)})] = \frac{j(n-j+1)}{(n+1)^2(n+2)}$$

$$\mathrm{Cov}[G(y_{(j)}),G(y_{(k)})] = \frac{j(n-k+1)}{(n+1)^2(n+2)} \qquad \text{for } j < k$$

فان مقدرات المربعات الصغرى نحصل عليها من تصغير المقدار k وهو:

$$k = \sum_{j=1}^{n}\left(G(y_{(j)}) - \frac{j}{n+1}\right)^2$$

وبالنسبة للتوزيع $GE(\alpha, \lambda)$ فأن :

$$k = \sum_{j=1}^{n}\left[\left(1-e^{-\lambda x_{(j)}}\right)^{\alpha} - \frac{j}{n+1}\right]^2 \qquad \ldots\ldots (18)$$

وعند اشتقاق المعادلة (18) بالنسبة لـ λ , α نحصل على:

$$\frac{\partial k}{\partial \alpha} = 2\sum_{j=1}^{n}\left[\left(1-e^{-\lambda x_{(j)}}\right)^{\alpha} - \frac{j}{n+1}\right]\left(1-e^{-\lambda x_{(j)}}\right)^{\alpha}(1)Ln\left(1-e^{-\lambda x_{(j)}}\right)$$

$$\frac{\partial k}{\partial \lambda} = 2\sum_{j=1}^{n}\left[\left(1-e^{-\lambda x_{(j)}}\right)^{\alpha} - \frac{j}{n+1}\right]\alpha\left(1-e^{-\lambda x_{(j)}}\right)^{\alpha-1} e^{-\lambda x_{(j)}} x_{(j)}$$

$$\longrightarrow \quad 2\sum_{j=1}^{n} x_{(j)} e^{-\lambda x_{(j)}}\left[\alpha\left(1-e^{-\lambda x_{(j)}}\right)^{\alpha-1}\right]\left[\left(1-e^{-\lambda x_{(j)}}\right)^{\alpha} - \frac{j}{n+1}\right]$$

وهي معادلات غير خطية وتحتاج الى طرق تكرارية خاصة لحلها.

واذا كانت معلمة القياس (λ) معلومة فان مقدر العزوم للمعلمة (α) يمكن الحصول عليه من حل المعادلة (17) مباشرة، حيث ان:

$$\lambda \bar{X} = \psi(\alpha+1) - \psi(1)$$

وان مقدار المربعات الصغرى نحصل عليه من تصغير المقدار k المعرف بالمعادلة (18) بالنسبة لـ α فقط (لأن λ معلومة).

فيما يلي نتائج مقدرات الطرائق الثلاث، الامكان الاعظم، العزوم، والمربعات الصغرى، وقد كتبت فيه كل مقدر اضافة الى $MSE(\hat{\alpha})$ داخل قوسين عندما ($\lambda = 1$)، وقد تم توليد قيم X من المعادلة:

$$X=((- Ln(1-U^{(\frac{1}{\alpha})})/\lambda)$$

علماً بأن $U \in [0,1]$، وان ($\lambda=1$)، ($\alpha = 0.3, 0.6, 1.0, 2.0$) كقيم اولية، واخذت حجوم عينات هي ($n = 10,20,50,100$)، وكررت كل تجربة (R=500). والجدول رقم (1) يلخص النتائج لطرائق التقدير الثلاث.

جدول رقم (1) القيم التقديرية للمعلمة α مع قيم $MSE(\hat{\alpha})$ عندما λ معلومة

n	Method	α=0.3	α=0.6	α=1.0	α=2.0
10	MLE	1.108(0168)	1.115(0.175)	1.115(0.172)	1.114(0.171)
	MME	1.096(0.826)	1.079(0.425)	1.093(0.375)	1.093(0.129)
	LSE	1.099(0.230)	0.928(0.129)	0.921(0.127)	0.928(0.237)
20	MLE	1.056(0.066)	1.049(0.062)	1.052(0.063)	1.052(0.064)
	MME	1.027(0.329)	1.037(0.179)	1.046(0.142)	1.037(0.120)
	LSE	1.033(0.080)	1.036(0.082)	1.038(0.095)	1.035(0.086)
50	MLE	1.020(0.022)	1.020(0.022)	1.020(0.023)	1.020(0.023)
	MME	1.012(0.125)	1.016(0.063)	1.014(0.051)	1.014(0.043)
	LSE	1.014(0.029)	1.012(0.028)	1.044(0.026)	1.016(0.029)
100	MLE	1.008(0.010)	1.009(0.011)	1.009(0.010)	1.009(0.010)
	MME	1.008(0.060)	1.009(0.031)	1.007(0.025)	1.009(0.021)
	LSE	1.008(0.014)	1.007(0.014)	1.005(0.014)	1.008(0.014)

يتضح من الجدول ان معدل الخطأ النسبي لمقدر α وبالطرائق الثلاث يتناقص كلما ازداد حجم العينة، وان جميع مقدرات α ($\hat{\alpha}_{MLE}$, $\hat{\alpha}_{MME}$, $\hat{\alpha}_{LSE}$) عندما ($\lambda=1$) كانت متسعة (Consistent)، ونتائج المقدرات متقاربة.

والان نأتي الى مقارنة طرائق تقدير λ عندما تكون معلمة الشكل α معلومة حيث افترض:

α = 0.5, 1.5, 2.5

وقدرت λبطريقة (MLE , MME , LSE) ولحجوم عينات (n=10,20,50,100)، وكررت كل تجربة (R=500)، والنتائج موضحة في الجدول رقم (2).

جدول رقم (2) القيم التقديرية للمعلمة λمع قيم $MSE(\hat{\lambda})$ عندما α معلومة

n	Method	α=0.5	α=1.5	α=2.5
10	MLE	1.244(0.523)	1.073(0.101)	1.043(0.057)
	MME	1.234(0.512)	1.076(0,102)	1.046(0.058)
	LSE	1.234(1.174)	1.053(0.122)	1.029(0.065)
20	MLE	1.111(0.163)	1.042(0.041)	1.023(0.026)
	MME	1.107(0.161)	1.034(0.044)	1.027(0.026)
	LSE	1.091(0.251)	0.960(0.040)	1.016(0.029)
50	MLE	1.043(0.046)	1.013(0.016)	1.008(0.009)
	MME	1.041(0.046)	1.013(0.016)	1.009(0.009)
	LSE	1.037(0.074)	1.007(0.018)	1.005(0.010)
100	MLE	1.022(0.021)	1.006(0.007)	1.004(0.005)
	MME	1.022(0.021)	1.007(0.007)	1.005(0.004)
	LSE	1.020(0.033)	1.004(0.009)	1.003(0.005)

نلاحظ ان الفروقات بين المقدرات صغيرة جداً، وخاصة عندما يكبر حجم العينة، وكذلك لوحظ تناقص قيم $MSE(\hat{\lambda})$، وعندما $\alpha < 1$ يكون $MSE(\hat{\lambda})$ كبير مقارنة بقيم $MSE(\hat{\lambda})$ عندما $\alpha > 1$.

الاستنتاجات والتوصيات

الاستنتاجات

1- يمكن استخدام التوزيع الاسي العام $GE(\alpha, \lambda)$ بدلاً من توزيع ويبل او توزيع كاما لتحليل بيانات اوقات الحياة، حيث انه يحقق افضل توفيق لمنحنى البيانات مقارنة بتوزيع ويبل ذي المعلمتين او كاما ذي المعلمتين.

2- تحتاج طريقة الامكان الاعظم الى حسابات تكرارية مطولة ناتجة عن حل المعادلات غيرالخطية، لذلك تم اولاً تقدير α عندما λ معلومة، وبالامكان تقدير $\hat{\alpha}$ ، $\hat{\lambda}$ عندما تكون كلاهما مجهولة.

3- بالامكان الاستفادة من مقدرات المعلمات للتوزيع $GE(\alpha, \lambda)$ في وضع تصاميم لخطط عينات القبول سيما وان C.D.F لهذا التوزيع تعتمد على المعلمتين (α , λ)، وان احتمال قبول الانتاج L(p) يعتمد على C.D.F، فبعد الحصول على افضل المقدرات يمكن وضع تصاميم مختلفة لخطط عينات القبول اعتماداً على تحقيق المتراجحة:

$$\sum_{i=0}^{c} \subset_{i}^{n} \mathrm{p}^{\mathrm{i}}\, \mathrm{q}^{\mathrm{n-i}} \geq 1-\alpha$$

حيث $P = F(X, \hat{\alpha}, \hat{\lambda})$.

4- بالامكان ايجاد التوزيع الاحتمالي للمقدرين، حيث وجد ان:

$$\left[\sqrt{\mathrm{n}}\,(\hat{\alpha}_{MLE} - \alpha), \sqrt{\mathrm{n}}\,(\hat{\lambda}_{MLE} - \lambda)\right] \xrightarrow{d} N_2(0, I^{-1}(\alpha, \lambda))$$

وان مصفوفة Fisher للمعلومات يمكن ايجادها من صيغة $r(x, \alpha, \lambda)$ المعرفة في المعادلة (10) .

التوصيات

1- نوصي باستخدام طرائق اخرى مثل المربعات الصغرى الموزونة، ونقاط التجزئة لتقدير المعلمات (α λ ,) للتوزيع الاسي العام.

2- نوصي باعتماد مقدرات بيز باعتبار ان كل من λ α , ايضاً متغيرات عشوائية لها توزيع اولي يحدد من الخبرة والبيانات السابقة.

3- نوصي بتعميم نتائج المقدرات والاستفادة منها من بناء تصاميم لخطط عينات القبول عندما يكون الوقت المستغرق في الفحص متغير عشوائي يتبع التوزيع الاسي العام.

References

1- Gupta R. D. & Kundu, D. (1997). "Exponentiated exponential Family: An alternative to gamma and Weibull distribution"; Technical report Dept of Math., University of New Brusnwick, Saint – John, NB, Canada .

2- Gupta, R. D. & Kundu, D. (1999). "Generalized exponential distribution"; Australian and Newzealand Journal of statistics, vol. 173-188.

3- Gupta R. D. & Kundu, D. (2004). "Discriminating between gamma and generalized exponential distributions"; Journal of statistical computation and simulation vol. 74, No.2 107-121.

4- Kundu, D. Gupta, R. D. & Manglick, A. (2005), "Discriminating between the log normal and generalized exponential distribution"; Journal of statistical planning and Inference, vol. 127, 213-227.