روش های مختلف برآورد پارامترهای توزیع بر نوع ۱۲ دو پارامتری

هانیه پناهی*
گروه ریاضی دانشگاه آزاد اسلامی واحد لاهیجان

چکیده

در برآورد پارامترهای یک توزیع، همیشه انتخاب بهترین برآورد پارامتر و به تبع آن بهترین روش برآورد مورد نظر بوده است. در این مقاله ابتدا پارامترهای توزیع بر نوع ۱۲ دو پارامتری با استفاده از روش های کلاسیک معروفی همانند درستنمایی ماکزیمم و گشتاوری برآورد شده اند، سپس پارامتر p از این توزیع با در نظر گرفتن تابع پیشین گاما و تحت توابع زیان مربع خطا و لگاریتمی به روش بیزی برآورد شده است. درآخر برای هر قسمت با استفاده از روش شبیه سازی نمونه هایی تصادفی از توزیع بر نوع ۱۲ دو پارامتری تولید شده است، و روش های مختلف برآورد با استناد به اریبی و ریشه میانگین مربع خطا مقایسه شده اند.

کلمات کلیدی: توزیع بر نوع ۱۲، برآوردگر درستنمایی ماکزیمم، برآوردگر گشتاوری، اطلاع فیشر، برآوردگر بیزی، تابع زیان.

۱ مقدمه

بر[1] در سال۱۹۴۲ با در نظر گرفتن معادله دیفرانسیل زیر ۱۲ نوع توزیع مختلف معرفی کرد[۳].

$$F'(x)=F(x)(1-F(x))\,g(x)$$

از بین این ۱۲ نوع توزیع، بر نوع ۱۰ و بر نوع ۱۲ بیشتر مورد توجه قرار گرفته است. محققین زیادی از جمله، وینگو [۱۱و۱۰]، زیمر و همکاران [۱۲]، رودریگز[۸]، وانگ [۹]، راکب و کاندو [۷] جاهین [۶] و احمد و همکاران [۱] به مطالعه این دو نوع توزیع پرداختند.

در این مقاله توزیع بر نوع ۱۲ دو پارامتری با تابع توزیع تجمعی زیر مطالعه گردیده است:

$$F(x)=1-\left(1-x^{b}\right)^{-p} \qquad x>0 \qquad b,p>0 \tag{۱}$$

بنابراین تابع چگالی احتمال آن عبارتست از:

$$f(x,p,b)=pbx^{b-1}(1+x^{b})^{-(p+1)} \qquad p,b>0 \qquad x>0 \tag{۲}$$

* عهده دار مکاتبات

[1] Burr

که p و b پارامترهای توزیع هستند. این توزیع کاربرد فراوانی در کنترل کیفیت، مدل بندی توزیع درآمد و آزمون فرض دارد. هدف اصلی این مقاله بررسی روش های مختلف برآورد پارامترهای توزیع بر نوع ۱۲ دو پارامتری است. تفاوت اساسی بین فلسفه برآورد بیزی و کلاسیک، این است که در برآورد بیزی، پارامتر توزیع همانند یک متغیر تصادفی است، در حالی که در برآورد کلاسیک، پارامتر، همانند یک نقطه ثابت در نظر گرفته می شود. هرگاه اطلاعات اضافی درباره پارامتر قابل دسترسی باشد، رهیافت بیزی بهتر از روش کلاسیک برای نمونه با حجم کوچک می باشد. در این بررسی برآوردهای بیزی و کلاسیک پارامتر(پارامترهای) توزیع بر نوع ۱۲ با تابع احتمال (۲) مطالعه شده است. ابتدا برآورد پارامترها به روش درست نمایی ماکزیمم و گشتاوری مورد مطالعه و مقایسه قرار گرفته است. سپس با معلوم فرض کردن پارامتر b، پارامتر p در توزیع بر نوع ۱۲ دو پارامتری با تابع پیشین گاما و تحت برخی توابع زیان به روش بیزی برآورد شده و نتایج با استفاده از روش شبیه سازی مورد بررسی قرار گرفته اند.

۲ برآورد پارامترهای توزیع بر نوع ۱۲ دو پارامتری به روش کلاسیک

۲-۱ برآورد پارامترها به روش درستنمایی ماکزیمم

روش درستنمایی ماکزیمم عبارتست از دستورالعملی برای به دست آوردن برآوردگری به نام برآوردگر درستنمایی ماکزیمم (MLE) و مبتنی بر یک تابع مهم به نام تابع درستنمایی است، تابع درستنمایی، تابعی از پارامترهاست و هدف روش درستنمایی ماکزیمم عبارتست از حداکثر کردن تابع درستنمایی برحسب پارامترها. بنابراین، با این فرض که X_1، ...، X_n نمونه ای تصادفی از توزیع بر نوع ۱۲ دو پارامتری با پارامترهای (p,b) باشد، تابع درستنمایی آن عبارتست از:

$$L(x,p,b) = p^n\ b^n \prod_{i=1}^{n}\left(x_i^{b-1}\right)\prod_{i=1}^{n}\left(1+x_i^b\right)^{-(p+1)} \tag{۳}$$

با در نظر گرفتن لگاریتم این تابع درستنمایی داریم:

$$ln\,L(x,p,b) = n\,ln\,p + n\,ln\,b + (b-1)\sum_{i=1}^{n} ln\,x_i - (p+1)\sum_{i=1}^{n} ln\left(1+x_i^b\right) \tag{۴}$$

با استفاده از نتایج آشور و همکاران [۲] برآوردگرهای درستنمایی ماکزیمم با مشتق گیری از لگاریتم تابع درستنمایی نسبت به پارامترهای مجهول p و b و مساوی قرار دادن معادلات حاصل با صفر و حل آن ها، به صورت زیر حاصل می شوند:

$$\hat{p} = \frac{n}{\sum_{i=1}^{n} ln\left(1+x_i^{\hat{b}}\right)} \tag{۵}$$

$$\frac{n}{\sum_{i=1}^{n} ln\left(1+x_i^{\hat{b}}\right)}=\frac{n+\hat{b}\sum_{i=1}^{n} ln\, x_i}{\hat{b}\sum_{i=1}^{n}\frac{x_i^{\hat{b}}\, ln\, x_i}{1+x_i^{\hat{b}}}}-1 \qquad (۶)$$

$\hat{b}_{MLE}$ با استفاده از روش تکرار از رابطه (۶) به دست می آید، سپس با قرار دادن آن در (۵)، $\hat{P}_{MLE}$ محاسبه خواهد شد.

اکنون با در نظر گرفتن نتایج مجانبی نرمال برای $\hat{P}_{MLE}$ و $\hat{b}_{MLE}$ داریم: [۴]

$$\left(\sqrt{n}\left(\hat{p}_{MLE}-p\right),\sqrt{n}\left(\hat{b}_{MLE}-b\right)\right)\xrightarrow{d} N_2\left(0,I^{-1}(p,b)\right)$$

که $I(p,b)$ ماتریس اطلاع فیشر است.

$$I(p,b)=-\frac{1}{n}\begin{bmatrix} E\left(\frac{\partial^2 ln\,L}{\partial p^2}\right) & E\left(\frac{\partial^2 ln\,L}{\partial p\partial b}\right)\\ E\left(\frac{\partial^2 ln\,L}{\partial b\partial p}\right) & E\left(\frac{\partial^2 ln\,L}{\partial b^2}\right)\end{bmatrix}$$

عناصر ماتریس اطلاع فیشر به صورت زیر می باشند:

$$E\left(\frac{\partial^2 ln\,L}{\partial p^2}\right)=-\frac{n}{p^2} \qquad (۷)$$

$$\begin{aligned} E\left(\frac{\partial^2 ln\,L}{\partial b^2}\right)&=-\frac{n}{b^2}-n(p+1)E\left(\frac{X^b(ln\,X)^2}{\left(1+X^b\right)^2}\right)\\ &=-\frac{n}{b^2}-n(p+1)\left(\frac{p}{b^2(p+1)(p+2)}\left(\Gamma''(2)-2\Gamma'(2)\psi(p+1)+\frac{\Gamma''(p+1)}{\Gamma(p+1)}\right)\right)\end{aligned} \qquad (۸)$$

$$\begin{aligned} E\left(\frac{\partial^2 ln\,L}{\partial p\partial b}\right)&=-nE\left(\frac{X^b\, ln\,X}{1+X^b}\right)\\ &=-\frac{n}{b\,(p+1)}\left(\Gamma'(2)-\psi(p)\right)\end{aligned} \qquad (۹)$$

همچنین $\psi(p)=\frac{\Gamma'(p)}{\Gamma(p)}$

۲-۲ برآورد پارامترها به روش گشتاوری

روش گشتاورها که، یکی از قدیمی ترین روش های برآوردیابی می باشد، متشکل از برابر گرفتن چند گشتاور اول جامعه با گشتاورهای متناظر یک نمونه و به دست آوردن تعدادی معادله، برای برآورد پارامترهای مجهول جامعه است. به عبارتی این روش، عبارت است از حل معادلات $M_r = E(X^r)$ برای یافتن پارامتر ها، که $E(X^r)$ گشتاور r ام جامعه حول مبدا است، این گشتاورها به صورت توابعی از پارامترهای مجهول جامعه هستند. همچنین $M_r = \sum_{i=1}^{n} X_i^r$ نمایانگرگشتاورr ام نمونه بر پایه نمونه تصادفی داده شده می باشد. بنابراین با فرض اینکه X_1، ...، X_n نمونه ای تصادفی از توزیع بر نوع ۱۲ دو پارامتری با پارامترهای (p,b) باشد، آن گاه برآورد گشتاوری (برآورد MM) پارامترهای مجهول p و b از تشکیل وحل دو معادله زیر حاصل خواهد شد:

$$\frac{\Gamma\left(\frac{1}{b}+1\right)\Gamma\left(p-\frac{1}{b}\right)}{\Gamma(p)} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad (۱۰)$$

$$\frac{\Gamma\left(\frac{2}{b}+1\right)\Gamma\left(p-\frac{2}{b}\right)}{\Gamma(p)} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \qquad (۱۱)$$

اکنون برای مقایسه برآوردهای حاصل از روش گشتاورها و روش درستنمایی ماکزیمم با استفاده از شبیه سازی [5] نمونه هایی به حجم ۱۰ ، ۲۵ ، ۵۰ و۸۵ از توزیع بر نوع ۱۲ دو پارامتری با پارامترهای۲=p و ۳=b به دست می آوریم و در هر نمونه تولید شده، p و b را به روش گشتاورها و روش درستنمایی ماکزیمم برآورد می کنیم. میانگین مقادیر برآورد شده p و b واریبی آن ها در جدول (۱) بیان شده است. تمام نتایج بیان شده بر اساس ۱۰۰۰ بار تکرار می باشند. مشاهده می شود که اریبی در روش درستنمایی ماکزیمم کمتر از روش گشتاورها است. همچنین با افزایش حجم نمونه اریبی کاهش می یابد.

۳ برآورد پارامتر توزیع بر نوع ۱۲ دو پارامتری به روش بیزی

فرض کنید X_1، ...، X_n نمونه ای تصادفی از توزیع بر نوع ۱۲ دو پارامتری با پارامترهای (p,b) باشد، با فرض معلوم بودن b برای محاسبه برآوردگر بیزی p ، آن را همانند متغیری تصادفی در نظر می گیریم و تابع پیشین آن را به صورت زیر تعریف می کنیم:

$$\pi(p) = \frac{m^{\nu+1}}{\Gamma(\nu+1)} p^{\nu} e^{-mp} \qquad m>0, \nu>0 \qquad (۱۲)$$

با استفاده از تابع درستنمایی (۳) و تابع پیشین(۱۲)، تابع پسین p بر اساس نمونه تصادفی $X=(x_1,...,x_n)$ برابر است با:

$$\pi(p|x)=\frac{(\sum_{i=1}^{n} ln(1+x_i^b)+m)^{n+\nu+1}}{\Gamma(n+\nu+1)}\ p^{n+\nu}\ e^{-(\sum_{i=1}^{n} ln(1+x_i^b)+m)\,p} \qquad (۱۳)$$

بنابراین با در نظر گرفتن رابطه (۱۳) برآوردگر بیزی p از توزیع بر نوع ۱۲ دو پارامتری

الف) تحت تابع زیان مربع خطا $L(p,d)=(d-p)^2$ به صورت

$$\begin{aligned} d_{SE} &= E(p|x)=\int_0^\infty p\,\frac{(\sum_{i=1}^{n} ln(1+x_i^b)+m)^{n+\nu+1}}{\Gamma(n+\nu+1)}\ p^{\nu+1}\ e^{-(\sum_{i=1}^{n} ln(1+x_i^b)+m)\,p}\,dp \\ &= \frac{n+\nu+1}{\sum_{i=1}^{n} ln\left(1+x_i^b\right)+m} \end{aligned} \qquad (۱۴)$$

ب) تحت تابع زیان لگاریتمی $L(p,d)=ln\left|\left(\frac{d}{p}\right)\right|^2$ به صورت

$$d_{LN}=exp\left(\int ln\,p\ \frac{\left(\sum_{i=1}^{n} ln\left(1+x_i^b\right)+m\right)^{n+\nu+1}}{\Gamma\left(n+\nu+1\right)}\ p^{n+\nu} e^{-\left(\sum_{i=1}^{n} ln\left(1+x_i^b\right)+m\right)p}\ dp\right) \qquad (۱۵)$$

می باشد.

از نرم افزار فرترن برای مقایسه برآوردهای بیزی p با تابع پیشین گاما و توابع زیان مربع خطا و لگاریتمی استفاده شده است. مقادیر p از تابع پیشین گاما با پارامترهای مشخص $m=6,\nu=4$ تولید شده اند. سپس با استفاده از مقادیر p نمونه هایی به حجم $n=15,30$ از توزیع بر نوع ۱۲ دو پارامتری با پارامتر معلوم $b=5$ تولید شده است. نتایج حاصل از میانگین برآوردها و ریشه میانگین مربع خطا در جدول (۲) بیان شده است. ملاحظه می شود که با انتخاب $n=15$ ریشه میانگین مربع خطا (RMSE) برآوردها، تحت تابع زیان مربع خطا، نسبت به تابع زیان لگاریتمی کمتر می باشد اما اگر حجم نمونه را به $n=30$ افزایش دهیم، تابع زیان لگاریتمی، RMSE کمتری دارد. همچنین برآوردهای حاصل تحت تابع زیان لگاریتمی از مقدار p کمتر می باشند. و با افزایش حجم نمونه ریشه میانگین مربع خطا کاهش می یابد.

٤ نتیجه گیری

در این مقاله، به بررسی روش های مختلف برآورد پارامترهای توزیع بر نوع ۱۲ دو پارامتری پرداخته شده است. نتایج نشان می دهند که در برآورد پارامترهای این توزیع به روش کلاسیک، اریبی برآوردگرهای درستنمایی ماکزیمم در مقایسه با برآوردگرهای گشتاوری کمتر است. همچنین برآوردگرهای درستنمایی

ماکزیمم دارای توزیع نرمال مجانبی می باشند و دنباله ای از برآوردگرهای درستنمایی ماکزیمم، بهترین نرمال مجانبی (BAN) است. در ادامه، برآوردگرهای بیزی توزیع بر نوع ۱۲ دو پارامتری برای تابع پیشین مزدوج گاما و تحت دو تابع زیان مربع خطا و لگاریتمی، مطالعه شده است. مزیت در نظر گرفتن توزیع های پیشین مزدوج در این است که توزیع پسین، متعلق به خانواده توزیع پیشین است و چگالی پسین و چگالی پیشین همگی شکل تابعی مشابهی دارند و از این رو، انعطاف پذیری را تضمین می کنند. با مقایسه ریشه میانگین مربع خطا (RMSE) برآوردها، مشاهده گردیده است که تابع زیان مربع خطا همیشه دارای کمترین، RMSE نمی باشد.

حجم نمونه	معیار	$\hat{b}_{MME}$	$\hat{b}_{MLE}$	$\hat{p}_{MME}$	$\hat{p}_{MLE}$
۱۰	مقدار برآورد شده (اریبی)	۳/۴۷۷۱ (۰/۴۷۷۱)	۳/۳۰۸۴ (۰/۳۰۸۴)	۲/۳۱۷۲ (۰/۳۱۷۲)	۲/۲۷۸۱ (۰/۲۷۸۱)
۲۵	مقدار برآورد شده (اریبی)	۳/۱۶۸۵ (۰/۱۶۸۵)	۳/۰۷۴۷ (۰/۰۷۴۷)	۲/۱۲۲۳ (۰/۱۲۲۳)	۲/۰۹۹۶ (۰/۰۹۹۶)
۵۰	مقدار برآورد شده (اریبی)	۳/۱۴۰۴ (۰/۱۴۰۴)	۳/۰۵۶۱ (۰/۰۵۶۱)	۲/۰۵۲۸ (۰/۰۵۲۸)	۲/۰۳۶۵ (۰/۰۳۶۵)
۸۵	مقدار برآورد شده (اریبی)	۳/۰۷۶۵ (۰/۰۷۶۵)	۳/۰۰۸۴۹ (۰/۰۰۸۴۹)	۲/۰۱۲۴ (۰/۰۱۲۴)	۲/۰۰۸۶ (۰/۰۰۸۶)

جدول۱

حجم نمونه	مقدار p	زیان مربع خطا	ریشه میانگین مربع خطا	زیان لگاریتمی	ریشه میانگین مربع خطا
	۰/۵۷۶۸۳	۰/۵۷۷۹۳	۰/۰۰۴۶۸	۰/۵۷۳۳۱	۰/۰۰۴۶۹
	۰/۶۷۱۱۲	۰/۶۷۲۰۸	۰/۰۰۳۴۶	۰/۶۶۹۱۹	۰/۰۰۳۴۷
n=۱۵	۰/۷۷۷۵۳	۰/۷۷۸۲۶	۰/۰۰۱۹۴	۰/۷۷۶۹۰	۰/۰۰۱۹۴
	۰/۸۸۴۲۱	۰/۸۸۴۶۳	۰/۰۰۰۶۳	۰/۸۸۴۲۶	۰/۰۰۰۶۳
	۰/۹۶۷۷۲	۰/۹۶۷۸۴	۰/۰۰۰۰۶	۰/۹۶۷۸۱	۰/۰۰۰۰۶
	۰/۵۷۷۸۵	۰/۵۷۸۹۳	۰/۰۰۲۴۹	۰/۵۷۶۲۸	۰/۰۰۲۴۸
	۰/۶۷۲۰۷	۰/۶۷۲۹۹	۰/۰۰۱۸۵	۰/۶۷۱۳۴	۰/۰۰۱۸۵
n=۳۰	۰/۷۷۸۳۰	۰/۷۷۸۹۹	۰/۰۰۱۰۴	۰/۷۷۸۲۱	۰/۰۰۱۰۳
	۰/۸۸۴۶۸	۰/۸۸۵۰۷	۰/۰۰۰۳۴	۰/۸۸۴۸۵	۰/۰۰۰۳۳
	۰/۹۶۷۸۶	۰/۹۶۷۹۷	۰/۰۰۰۰۳	۰/۹۶۷۸۶	۰/۰۰۰۰۳

جدول۲

منابع

[1] Ahmad, K. E., Fakhry, M. E., Jaheen, Z. F. (1997), Empirical Bayes estimation of P(Y < X) and characterization of Burr-type X model. Journal of Statistical Planning and Inference, vol. 64, 297-308.

[2] Ashour, S. K. , Abd-Elfattah, A. M., Mandouh, R. M. (2005), Bayesian and Non-Bayesian Estimation of Pr{Y<X} in Two Parameters Lomax Distribution.

[3] Burr, I. W. (1942), Cumulative frequency functions. Ann. Math. Statist. , 13: 215-22.

[4] Gupta, R. D. Kundu, D. (1999), Generalized Exponential Distributions Statistical Inferences, Technical Report, The University of New Brunswick, Saint.

[5] Gupta, R. D. , Kundu, D. (2001), Exponentiated Exponential FamilyAn Alternative to Gamma and Weibull Distributions, Technical Report, The University of New Brunswick, Saint John.

[6] Jaheen. Z. F. (1995), Bayesian approach to prediction with outliers from the Burr type X. model. Microelectronic. Reliability, 35, 45-47. 118.

[7] Raqab, M. Z., and Kundu, D. (2003), Burr Type X distribution; revisited. submitted for publication.

[8] Rodriguez, R. N. (1977), A guide to Burr Type XII distributions, Biometrika, vol. 64, 129-134.

[9] Wang, F. K., Keats, J. B., Zimmer, W. J. (1996), Maximum likelihood distribution with censored and uncensored data. Microelectron. Reliability, 36: 359–362.

[10] Wingo, D. R. (1983), Maximum likelihood methods for fitting the Burr Type XII distribution parameters to life test data. Biometrical J. , 25: 77–84.

[11] Wingo, D. R. (1993), Maximum likelihood methods for fitting the Burr type XII distribution to multiply (progressively) censored life test data. *Metrika*, 40: 203- 210.

[12] Zimmer, W. J., Keats, J. B., Wang, F. K. (1998), The Burr XII distribution in reliability analysis. Journal of Quality Technology, 30: 386.