Statistical Sciences Extension
Vol. 2, No. 1, Winter 2017, (P 9-14)

گسترهٔ علوم آماری
سال دوم، شماره اول، پاییز و زمستان 1395، (ص 9-14)

برآورد پارامتر توزیع نمایی سانسور شده از راست

آمنه سادات میرنیام[*1]**، زهرا شناوری**[2]**، عبدالرسول برهانی حقیقی**[3]

1. دانشجوی دکتری، گروه آمار، دانشگاه شیراز
2. دانشجوی دکتری، گروه آمار، دانشگاه شیراز
3. استادیار، گروه آمار، دانشگاه شیراز

تاریخ دریافت: 1394/10/08 **تاریخ پذیرش**: 1395/09/01

The Estimation of the Right Censored Exponential Distribution Parameter

A. Mirniam[*1]**, Z. Shenavari**[2]**, A. Borhani Haghighi**[3]

1. Ph.D Student, Department of Statistics, Shiraz University
2. Ph.D Student, Department of Statistics, Shiraz University
3. Assistant Professor, Department of Statistics, Shiraz University

Received: 2015/15/28 **Accepted:** 2016/11/21

Abstract

In this article, the generalized progressive type II censoring design (right censoring) is introduced. Then the likelihood function for such censored variables is derived and it is precisely determined for the exponential distribution case. The derived maximum likelihood estimator has no closed form, so the estimate is achieved by the numerical "False Position" method. Finally, a suitable confidence interval for the parameter of the exponential distribution is constructed in the form of a theorem.

Keywords

Maximum Likelihood Estimator, Exponential Distribution, Progressive Type II Censoring, Survival Function, False Position Method.

چکیده

در این مقاله ابتدا طرح سانسور نوع دوم پیشرو تعمیم‌یافته (سانسور از راست) معرفی می‌شود. سپس تابع درست‌نمایی را برای این‌گونه متغیرها به دست آورده شده و در حالت توزیع نمایی، تابع درست‌نمایی به صورت دقیق محاسبه گردیده است. از آنجا که برآوردگر درست‌نمایی ماکسیمم حاصل از این تابع صورت تحلیلی ندارد، لذا با استفاده از روش عددی «موقعیت خطا»، برآورد پارامتر نمایی را به دست می‌آوریم. در پایان یک بازه اطمینان مناسب برای پارامتر توزیع نمایی در این طرح معرفی می‌شود.

واژگان کلیدی

تابع درست‌نمایی، توزیع نمایی، سانسور نوع دوم پیشرو تعمیم‌یافته، تابع بقا، روش موقعیت خطا.

*Corresponding Author: bahareh104@gmail.com

*** نویسنده مسئول**: آمنه سادات میرنیام

مقدمه

متغیرهای تصادفی سانسور شده (نمونه‌های سانسور شده) را دیویس و فلدستین[1] (1979)، سن[2] (1979)، هالپرین و همکاران[3] (1989)، ویورس و بالاکریشنان[4] (1994) باندیوپادهیای و چاتوپادهیای[5] (1995)، ین وتسه[6] (1996)، آگاروالاو بالاکریشنان[7] (1998)، بالاکریشنان و همکاران[8] (2001) و گیلباد[9] (2001) مطالعه کرده‌اند.

در اینجا تابع درست‌نمایی را برای متغیرهای تصادفی سانسور شده نوع دوم پیشرو تعمیم‌یافته به دست آورده است و در حالت خاصی که متغیرها دارای توزیع نمایی هستند، این تابع را ارائه می‌کند. سپس برآورد ماکسیمم درست‌نمایی پارامتر نمایی با استفاده از این تابع به دست می‌آید. از آنجا که این برآوردکننده دارای فرمول بسته‌ای نیست، با استفاده از روش موقعیت خطا به برآورد آن می‌پردازد. در پایان یک بازه اطمینان برای این برآورد ارائه خواهد کرد.

طرح سانسور نوع دوم پیشرو و تابع درست‌نمایی

در ابتدا یک طرح سانسور شده نوع دوم پیشرو را معرفی می‌کنیم. تحت این طرح سانسور، n مشابه در یک آزمون طول عمر قرار داده می‌شوند. پس از مشاهدهٔ اولین شکست، به تعداد R_1، از واحدهای مانده به‌طور تصادفی از بین مشاهدات دیگر حذف می‌شوند. پس از شکست بعدی، R_2 واحد از باقیمانده‌ها به‌طور تصادفی از آزمون بیرون کشیده می‌شوند و در نهایت پس از شکست mام، R_m واحد مانده نیز کنار می‌روند. بنابراین ما در این نوع نمونه‌گیری مشاهده می‌کنیم که m شکست صورت گرفته و $R_1+R_2+\dots+R_m$ واحد پشت‌سرهم سانسور می‌شوند که نتیجه می‌دهد:

$$n = m + \left(R_1 + R_2 + \dots + R_m\right)$$

1. Davis & Feldstain
2. Sen
3. Halperin, et al.
4. Viveros & Balakrishnan
5. Bandyopadhyay & Chatopadhyay
6. Yuen & Tse
7. Aggarwala & Balakrishnan
8. Balakrishnan, et al.
9. Guilbaud

به این نوع سانسور، سانسور از راست هم گفته می‌شود. حال اگر فرض شود که به تعداد r شکست اولیه مشاهده نشده‌اند و عمل سانسور از زمان شکست واحد $r+1$ام آغاز شود، سانسور نوع دوم پیشرو تعمیم‌یافته را خواهیم داشت. لازم به ذکر است که مقادیر r، m و R_iها، مقادیر از پیش تعیین‌شده هستند. در مطالعات آزمون عمر و قابلیت اعتماد، توزیع نمایی، $E(\theta)$، نقش مهمی را بازی می‌کند و استنباط‌هایی که بر اساس این مدل به دست می‌آیند، به‌طور گسترده‌ای مورد استفاده قرار می‌گیرند (فرناندز[10]، 2004).

تابع چگالی احتمال و تابع قابلیت اعتماد (بقا) برای یک متغیر تصادفی X به ازای $X > 0$ و $\theta > 0$، به‌صورت زیر است.

$$f(x|\theta) = \theta^{-1}\exp(-x/\theta)$$
$$S(x|\theta) = \exp(-x/\theta)$$

که در آن $X \sim E(\theta)$ و θ طول عمر متوسط است. همچنین $\lambda = \theta^{-1}$ را نرخ شکست می‌نامیم. حال فرض کنید n واحد که به‌طور تصادفی از یک جمعیت $E(\theta)$ انتخاب شده‌اند و θ در آن نامعلوم است، تحت یک طرح سانسور نوع دوم پیشرو تعمیم‌یافته در یک آزمون قرار داده می‌شوند و فرض کنید $X = (x_{r+1}, \dots, x_m)$ نمونه‌ی مشاهده‌شده باشد. در این صورت برای به دست آوردن تابع درست‌نمایی مراحل زیر را دنبال می‌کنیم:

. احتمال مشاهده نشدن r زمان شکست اول برابر است با:

$$\binom{n}{r}\{P(X \le x_{r+1}|\theta)\}^r = \binom{n}{r}\{F(x_{r+1}|\theta)\}^r$$

2. احتمال مشاهده x_{r+1} و حذف R_{r+1} تا از مؤلفه‌ها عبارت است از:

$$\binom{n_r}{1} f(x_{r+1}|\theta)\{P(X \ge x_{r+1})\}^{R_{r+1}}$$
$$= n_r\, f(x_{r+1}|\theta)\{1 - F(x_{r+1}|\theta)\}^{R_{r+1}}$$

3. احتمال مشاهده x_{r+2} و حذف R_{r+2} تا از مؤلفه‌ها عبارت است از:

10. Fernandez

$$\binom{n_{r+1}}{1} f(x_{r+2}|\theta)\{P(X \ge x_{r+2})\}^{R_{r+2}} = n_{r+1} f(x_{r+2}|\theta)\{1-F(x_{r+2}|\theta)\}^{R_{r+2}}$$

به همین ترتیب در مرحله‌ی $m-r+1$، احتمال مشاهده x_m و حذف باقی‌مانده مؤلفه‌ها (تا R_m) عبارت است از:

$$\binom{n_{m-1}}{1} f(x_m|\theta)\{P(X \ge x_m)\}^{R_m} = n_{m-1} f(x_m|\theta)\{1-F(x_m|\theta)\}^{R_m}$$

بنابراین تابع درست‌نمایی برای θ به‌صورت زیر خواهد بود:

$$L(\theta|X) = \binom{n}{r}\prod_{j=r}^{m-1} n_j (1-e^{-x_{r+1}/\theta})^r \times \theta^{-(m-r)} e^{-\sum_{i=r+1}^{m}\frac{x_i(1+R_i)}{\theta}} \qquad (1)$$

در نتیجه:

$$L(\theta|\mathbf{X}) \propto \{1-e^{-x_{r+1}/\theta}\}^r \theta^{-(m-r)} e^{-\frac{w}{\theta}}$$

که در آن w مقدار مشاهده‌شده آمارهٔ زیر است:

$$W \equiv W(\mathbf{x}) = \sum_{i=r+1}^{m} X_i(1+R_i)$$

نکته: اگر $r=0$ باشد واضح است که W آماره‌ی بسندهٔ مینیمال است. در غیر این صورت (X_{r+1}, W) آمارهٔ بسندهٔ مینیمال خواهد بود. از طرفی داریم

$$W = (n-r)X_{r+1} + V$$

که در آن

$$V = V(\mathbf{x}) = \sum_{i=r+2}^{m} n_{i-1}(X_i - X_{i-1}).$$

روشن است که (X_{r+1}, V) نیز آمارهٔ بسنده مینیمال است.

برآورد درست‌نمایی ماکسیمم

برآورد ماکسیمم درستنمایی θ، (MLE)، با $\hat{\theta} \equiv \hat{\theta}_n$ نشان داده می‌شود و با استفاده از لگاریتم رابطهٔ (1)، یعنی از رابطه زیر به دست می‌آید:

$$\frac{1}{\theta^2}\left\{w-(m-r)\theta - \frac{r x_{r+1}}{\exp(\frac{x_{r+1}}{\theta})-1}\right\} = 0 \qquad (2)$$

همان‌طوری که از رابطه (2) مشاهده می‌شود، به دست آوردن یک عبارت با صورت بسته برای $\hat{\theta}$ امکان‌پذیر نیست، (مگر جایی که $r=0$ باشد که در این حالت: $\hat{\theta} = w/m$ می‌شود.)؛ بنابراین باید با استفاده از روش‌های موجود این MLE را تعدیل کرد و محاسبه عددی کرد. یک روش یافتن این MLEی تعدیل‌شده به‌وسیله خطی‌کردن معادلهٔ درست‌نمایی به شکل مناسبی است. با استفاده از قضیهٔ 1 که در زیر آمده است، نشان داده می‌شود که برآوردکننده درست‌نمایی ماکسیمم یکتا است.

قضیه 1: فرض کنید:

1. X دارای چگالی نمایی دو پارامتری به صورت

$$X \sim E\left(\frac{x-\mu}{u}\right) \quad \mu \in R, u>0, x>0$$

باشد، که در آن μ معلوم است.

2. در آزمون طول عمر، r شکست اولیه مشاهده نمی‌شوند.

3. $T^{R}_{r+1:m,n} \le \dots \le T^{R}_{m:m,n}$ آماره‌های ترتیبی $r+1$ام تا mام از X در آزمون طول عمر موردنظر با n واحد اولیه و طرح سانسور $(R_{r+1},\dots,R_m)$ هستند.

4. از s بار آزمون، مقادیر طول عمر مربوط مشاهده می‌شود که در آن داده‌ها با متغیرهای تصادفی زیر توصیف می‌شوند:

$$\left(T^{R_i}_{i:r_i+1:m_i,n_i}, \dots, T^{R_i}_{i:m_i:m_i,n_i}\right)$$

مقادیر مشاهده‌شده t_{ij} و طرح سانسور $R_i = (R_{ij})$ می‌باشند که در آن:

$$1 \le i \le s \quad \text{و} \quad r_i + 1 \le j \le m_i$$

آنگاه MLEی ν جواب یکتای معادله زیر است:

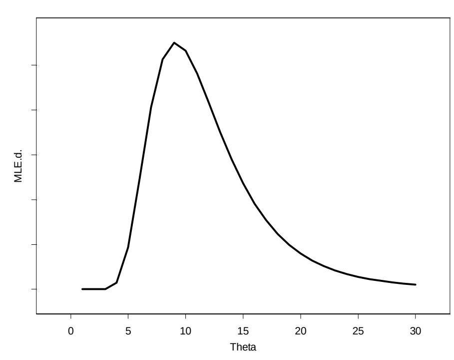
$$\sum_{i=1}^{s} \frac{r_i \left(T^{\tilde{R}_i}_{i;r_i+1:m_i,n_i} - \mu\right)}{\exp\{\left(T^{\tilde{R}_i}_{i;r_i+1:m_i,n_i} - \mu\right)/\nu\} - 1} + \nu \sum_{i=1}^{s} (m_i - r_i) = \sum_{i=1}^{s} \sum_{j=r_i+1}^{m_i} (R_{ij}+1)(T^{\tilde{R}_i}_{i;j:m_i,n_i} - \mu) \tag{3}$$

اثبات در منبع [1].

با ترسیم شکل رابطۀ (2) نیز می‌توان یکتا بودن ماکسیمم تابع درست‌نمایی را تبیین کرد. (شکل 1)

محاسبه عددی

یکی از روش‌های عددی برای به دست آوردن مقدار برآورد درست‌نمایی، روش نیوتن است. با وجود این روش‌های الحاقی مانند موقعیت خطا[1] ترجیح داده می‌شوند، زیرا می‌تواند کرانه‌ای صریحی ایجاد کند. در زیر این روش را توضیح داده و سپس قضیه‌ای جهت به دست آوردن فاصله اطمینان ارائه می‌شود.



شکل 1. نمودار تابع درست‌نمایی رابطه (2)

روش موقعیت خطا

قضیه 2: فرض کنید که f یک تابع پیوسته روی فاصله $[a,b]$ باشد و همچنین فرض کنید که عددی مانند r در این بازه وجود دارد به طوری که $f(r) = 0$.

اگر $f(a)$ و $f(b)$ علامت‌های متفاوت داشته باشند، و $c_n = \frac{a_n f(b_n) - b_n f(a_n)}{f(b_n) - f(a_n)}$ نشان‌دهندۀ دنبالۀ نقاطی باشد که به‌وسیلۀ فرآیند موقعیت خطا تولید می‌شوند، آنگاه دنبالۀ $\{c_n\}$ به سمت ریشۀ $x = r$ میل می‌کند. یعنی $\lim_{n \to \infty} c_n = r$.

اگر $r \neq 0$ قانون موقعیت خطا فرآیند ساده و راحتی برای تعیین $\hat{\theta}$ است. بدین‌صورت که این قضیه کرانه‌ای صریحی برای مقدار $\hat{\theta}$ ایجاد می‌کند. به‌علاوه این روش همیشه خوب کار می‌کند. (مثلاً روش نیوتون، اگر مقدار اولیه نزدیک به جواب نباشد، ممکن است در هم‌گرایی‌اش زمان‌بر باشد و یا در یک سیکل قرار گیرد.) و همچنین آزمایش‌ها نشان داده‌اند که میزان هم‌گرایی این روش بسیار سریع است. در نهایت هزینه محاسبات نیز کاهش می‌یابد. (در این روش نیازی به محاسبه مشتقات نیست.) بالاتر از همه اینکه خطاها با دقت بالایی می‌توانند کراندار شوند. برای اثبات می‌توان به وب‌گاه‌هایی که در منابع [13] و [14] آمده است مراجعه کرد.

فاصله اطمینان برای پارامتر θ

قضیه 3: برآورد درست‌نمایی ماکسیمم θ، یعنی $\hat{\theta}$، در فاصلۀ $\hat{\theta}_L \le \hat{\theta} \le \hat{\theta}_U$ قرار دارد، جایی که

$$\hat{\theta}_L = \left(u + \sqrt{u^2 + v^3}\right)^{\frac{1}{3}} + \left(u - \sqrt{u^2 + v^3}\right)^{\frac{1}{3}} - \frac{\alpha}{3}$$

$$\hat{\theta}_U = \min\left(\frac{\omega}{\mu - \rho}, \frac{\rho + \sqrt{\rho^2 + 24\mu\omega x_{r+1}}}{12\mu}\right)$$

که در آن:

1. False Position

$$\rho = 6w - (m-3r)x_{r+1}$$
$$u = ab/6 - c/2 - (a/3)^3$$
$$v = b/3 - (a/3)^2$$
$$a = \frac{3(m-r)x_{r+1} - 6w}{6m}$$
$$b = \frac{(m-r)x_{r+1}^2 - 3w\ x_{r+1}}{6m}$$
$$c = -\frac{w\ x_{r+1}^2}{6m}$$

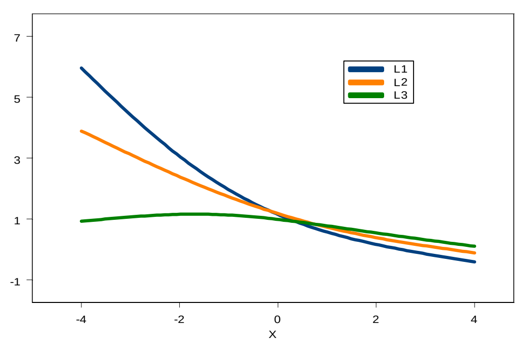
اثبات: فرض کنید که برای $x > 0$،

$$L_1 = (6-2x)/(6+x),$$
$$L_2 = (x)/(exp(x)-1),$$
$$L_3 = 1/(1+(x/2)+(x^2/6)).$$

ابتدا درستی نامساوی روبه‌رو را اثبات می‌کنیم:

$$\max\{0, L1\} < L2 < L3$$

برای این کار ابتدا $L_1 = \frac{6-2x}{6+x} = 0$ را قرار می‌دهیم که مقدار $x = 3$ به دست می‌آید.



شکل 2. نمودارهای L1، L2 و L3 روی یک محور مختصات

اگر $x < 3$ باشد، آنگاه $\max\{0,L1\} = L1$ و اگر $x > 3$ باشد، $\max\{0,L1\} = 0$. نمودارهای L1، L2 و L3 روی یک محور مختصات در شکل (2) رسم شده‌اند. حال برای اثبات قضیه از این نامساوی کمک می‌گیریم و سمت راست رابطهٔ (2) را می‌سازیم. به‌سادگی دیده می‌شود که:

$$\rho\theta.\max\left\{0, \frac{6-2\frac{x_{r+1}}{\theta}}{6+\frac{x_{r+1}}{\theta}}\right\} < \frac{\rho\theta.\frac{x_{r+1}}{\theta}}{\exp\left(\frac{x_{r+1}}{\theta}\right)-1} < \frac{\rho\theta}{1+\frac{x_{r+1}}{2}+\frac{x_{r+1}^2}{6}}$$

یا

$$w-(m-r)-\frac{\rho\theta}{1+\frac{x_{r+1}}{2}+\frac{x_{r+1}^2}{6}}$$
$$< w-(m-r)\theta-\frac{\rho x_{r+1}}{\exp\left(\frac{x_{r+1}}{\theta}\right)-1}$$
$$< w-(m-r)\theta+\min\left\{0, \frac{6-2\frac{x_{r+1}}{\theta}}{6+\frac{x_{r+1}}{\theta}}\right\}$$

حال با ضرب نامساوی بالا در θ^{-2} و در نظر گرفتن تعاریف ρ، c، b، a، v و u، قضیه اثبات می‌شود.

نکته 1: اگر $r = 0$ باشد، نامساوی‌ها به تساوی تبدیل می‌شود و خواهیم داشت: $\hat{\theta} = \frac{w}{m}$.

نتیجه‌گیری

طرح سانسور نوع دوم پیشرو طرح بسیار پرکاربردی در بسیاری از رشته‌ها و مشاغل می‌باشد ولی از آنجا که در اکثر موارد فرم بسته‌ای برای تابع درست‌نمایی وجود ندارد، روش‌های عددی می‌توانند کارساز باشند.

منابع

[1] Aggarwala, R., Balakrishnan, N. (1998). Some properties of progressive censored order statistics from arbitrary and uniform distribiutions with applications to inference and simulation. Journal of statistical planning and inference, **70**, 35-49.

[2] Balakrishnan, N., Cramer, E., kamps, U., sckenk, N. (2001). Progressive type II

censored order statistics from exponential distributions. Statistics, **35**, 537-556.

[3] Balakrishnan, N., Sandhu, R.A. (1995). A simple simulational algorithm for generating progressive type II censored samples .The American Statistician, Vol. **49**, No.2. pp.229-230.

[4] Balakrishnan, N., Sandhu, R.A. (1996). Best linear unbiased and maximum likelihood estimation for exponential distribution under general progressive type II censored samples. Sankhya B, **58**, 1-9.

[5] Bandyopadhyay, U., Chattopadhyay, G., (1995). Progressive censoring under inverse sampling for nonparametric two-sample problems. Sequential Anal., **14**, 1-28.

[6] Davis, H.T., Feldstein, M.L. (1979). The generalized pareto law as a model for progressively censored survival data. Biometrica, **66**, 299-306 .

[7] Fernandez, A.J. (2004). On estimating exponential parameters with general type II progressive censoring. Journal of statistical planning and inference, **121**, 135- 147.

[8] Guilbaud, O. (2001). Exact non-parametric confidence intervals for quantiles with progressive type II censoring. Scand. J. Statis, **28**, 699-713.

[9] Halperin, M., Hamdy, M.I., Thall, P.F. (1989). Distribution - free confiedence interval for a parameter of Wilcoxon-Mann-Whitney type for ordered categories and progressive censoring. Biometrics, **45**, 509-521.

[10] Sen, P.K. (1979). Weak convergence of som quantile processes arising in progressively censored tests. Ann. Statist, **7**, 414-431.

[11] Viveros, R., Balakrishnan, N. (1994). Interval estimation of life characteristics from progressively censored samples. Technometrics, **36**, 84-91.

[12] Yuen, H.K., Tse, S.K. (1996). Parameters-estimation for Weibull distributed lifetimes under progressive censoring with random removals. J. Statistics. Comput. Simulation, **55**, 57-71.

[13] Website: math.fullerton.edu/mathews /n2003 /RegulaFalsiMod.html.

[14] Website: mat.iitm.ac.in /~sryedida /caimna/ transcendental/ bracketing%20 methods/regula-falsi/regula-falsi.html.