

## چند برآوردگر ناریب با کمترین واریانس سهل الوصول\*

ویلیام گونتر

ترجمه عادل محمدپور و سید محمود طاهری<sup>†</sup>

## چکیده

وجود دارد. در وهله اول، شاید قادر باشیم شکل  $u(Y)$  را به طور صحیح حدس بزنیم که معمولاً چنانچه  $h(\theta) = \theta$ ، این کار شدنی است. اگر  $E(Y) = K\theta$  آنگاه می‌گیریم  $u(Y) = Y/k$ ، گاهی  $E(Y) = k_1/\theta$  آنگاه و چنین به ذهن می‌آید که با  $1/Y$  کار کنیم. اگر  $E(1/Y) = k_2\theta$  آنگاه  $u(Y) = \frac{1}{k_2 Y}$ . در بسیاری موارد یک حدس خوب برای  $u(Y)$ ، برآورد درست‌نمایی ماکسیمم است. در وهله دوم یک برآوردگر ناریب  $h(\theta)$  را حدس بزنید یا پیدا کنید و سپس امید شرطی برآوردگر ناریب به شرط  $Y$  را محاسبه کنید. این رهیافت اغلب وقتی مفید است که در پی یافتن یک برآوردگر ناریب برای  $P(X \leq c)$  باشیم. برای مطالعه مثالهای مقدماتی می‌توانید به [۵] ص ۲۳۵، و [۷] ص ۳۲۷ رجوع کنید. مثالهای پیچیده‌تر را می‌توان در مقاله‌های مجلات علمی یافت. سومین شیوه برای یافتن MVUE، حل معادله انتگرال زیر است:

$$\int u(y)g(y; \theta)dy = h(\theta) \quad (1)$$

که در حالت گسسته، انتگرالگیری به مجموعیایی تبدیل می‌شود. دو روش اول به حد کافی در کتابهای درسی مورد بحث قرار گرفته‌اند. شیوه سوم کمتر مورد توجه قرار می‌گیرد و هدف این مقاله جلب توجه به مواردی است که برای آنها (۱) به سادگی قابل حل است. معادله (۱) را با روشهای گوناگونی می‌توان حل کرد. تیت (۱۹۵۹) روشی را با استفاده از تبدیلهای انتگرال پیشنهاد کرده است. چون این روش

چندین روش به منظور یافتن برآوردگرهای ناریب با کمترین واریانس (MVUE) برای تابعی از پارامترهای توزیع در دسترس است. در این مقاله به دو روش که به ندرت استفاده می‌شود اما در صورت قابل کاربرد بودن، ساده‌اند می‌پردازیم. اولین روش که قبلاً توسط دیویس<sup>۱</sup> [۲] و تیت<sup>۲</sup> [۹] مورد بحث قرار گرفته‌اند برآوردگرهایی را از طریق مشتقگیری به دست می‌دهد، در صورتی که تکیه‌گاه یک متغیر تصادفی پیوسته به پارامتری مجهول بستگی داشته باشد. دومین روش که قابلیت کاربرد گسترده‌تری دارد و این امکان را می‌دهد که با آن برآوردگرهایی برای بعضی توابع پیچیده، با استفاده از نتایج شناخته شده از نظریه توزیعها، به دست آوریم. مثالهایی ارائه شده‌اند، که بسیاری از آنها مناسب تمرینهای کلاسی‌اند.

## ۱ مقدمه

فرض کنید  $X_1, X_2, \dots, X_n$  نمونه‌ای تصادفی از توزیعی با تابع چگالی  $f(x; \theta)$ ،  $a \leq x \leq b$ ،  $\theta_1 \leq \theta \leq \theta_2$ ، باشد. مسأله یافتن یک MVUE برای تابعی از  $\theta$ ، مثلاً  $h(\theta)$ ، در بسیاری از کتابهای درسی بحث شده است. روش متداول این است که نخست یک آماره بسنده کامل مانند  $Y$  با تابع چگالی  $g(y; \theta)$  را تشخیص می‌دهیم و سپس سعی می‌کنیم تابعی از  $Y$  مانند  $u(Y)$  بیابیم به قسمی که  $E(u(Y)) = h(\theta)$  (برای نمونه ن. ک. به [۷]، ص ۳۲۶). برای این کار شیوه‌های کار گوناگونی

† دانشجویان دوره دکتری آمار دانشگاه شیراز

\*) Minimum Variance Unbiased Estimator

1) DAVIS 2) TATE

مثال ۱.

$$f(x; \theta) = \frac{1}{\theta}, \quad 0 < x < \theta$$

در این صورت

$$Q_1(\theta) = \frac{1}{\theta}, \quad M_1(x) = 1, \quad Y = Y_n$$

اگر  $h(\theta) = \theta^r$ ، آنگاه بنابر (۴-الف)  $u(y) = y^r(1 + \frac{r}{n})$ ، بک MVUE برای  $h(\theta)$  است.

$$\text{اگر } h(\theta) = P(X \leq c) = \frac{c}{\theta}$$

$$u(y) = \left(\frac{c}{y}\right)\left(1 - \frac{1}{n}\right)$$

مثال ۲.

$$f(x; \theta) = e^{-(x-\theta)}; \quad x > \theta$$

در این صورت

$$Q_2(\theta) = e^\theta, \quad M_2(x) = e^{-x}, \quad Y = Y_1$$

اگر  $h(\theta) = \theta^r$  آنگاه

$$u(y) = y^r - ry^{r-1}/n$$

اگر  $h(\theta) = P(X \leq c) = 1 - e^{-(c-\theta)}$  آنگاه

$$u(y) = 1 - \left(1 - \frac{1}{n}\right)e^{-(c-y)}$$

مثال ۳ [۹] (صص ۳۶۱-۳۶۲)

$$f(x; \theta) = e^{-x}/(e^{-\theta} - e^{-b}); \quad \theta < x < b$$

در این صورت،

$$Q_2(\theta) = \frac{1}{(e^{-\theta} - e^{-b})}, \quad M_2(x) = e^{-x}, \quad Y = Y_1$$

اگر  $h(\theta) = \theta^r$  آنگاه با کمی محاسبه داریم

$$u(y) = y^r - ry^{r-1}[1 - e^{-(b-y)}]/n$$

اگر  $h(\theta) = P(X \leq c) = \frac{e^{-\theta} - e^{-c}}{e^{-\theta} - e^{-b}}$  آنگاه

$$u(y) = \frac{1}{n} + \left(1 - \frac{1}{n}\right) \frac{[1 - e^{-(c-y)}]}{[1 - e^{-(b-y)}]}$$

دو تابع چگالی دیگر که نتایج مشابهی با مثالهای بالا دارند عبارتند از [۹]:

$$f(x; \theta) = \frac{b\theta}{(b-\theta)x^2}, \quad 0 < \theta < x < b$$

غالباً پیچیده است و به جدول محدودی نیاز دارد، ما حتی الامکان از آن استفاده نمی‌کنیم مگر اینکه نتوان از روشهای دیگر استفاده کرد. در بخش ۲ بعضی حالتها که در آنها معادله (۱) را می‌توان با مشتقگیری حل کرد، بحث شده‌اند. در بخش ۳ توضیح می‌دهیم که برای بعضی از  $h(\theta)$ ها،  $u(Y)$  را می‌توان با چند عمل و آگاهی مقدماتی از نظریه توزیعها پیدا نمود.

## ۲ حل معادله‌های انتگرالی به وسیله مشتقگیری

اگر  $X$  متغیر تصادفی بیوسته‌ای با تابع چگالی

$$f_1(x; \theta) = Q_1(\theta)M_1(x) \quad a < x < \theta \quad (2)$$

یا

$$f_2(x; \theta) = Q_2(\theta)M_2(x) \quad \theta < x < b \quad (3)$$

باشد، آنگاه  $u(Y)$  را می‌توان با مشتقگیری به دست آورد. برای حالت اول  $Y_n = Y$  ( $n$  امین آماره ترتیبی)، و

$$g(y; \theta) = nM_1(y)\{[Q_1(\theta)]^n/[Q_1(y)]^{n-1}\}, \quad a < y < \theta$$

با مشتقگیری از

$$\int_a^\theta f(x; \theta) dx = 1$$

و

$$\int_a^\theta u(y)g(y; \theta) dy = h(\theta)$$

(و نقل عبارتهای شامل  $\theta$  از سمت چپ هر معادله به سمت راست)، خواهیم داشت

$$M_1(\theta) = -\{Q_1'(\theta)/[Q_1(\theta)]^2\}$$

$$u(\theta) = \{Q_1(\theta)h'(\theta) - nh(\theta)Q_1'(\theta)\}/nM_1(\theta)[Q_1(\theta)]^2$$

که از آنها نتیجه می‌شود که

$$u(y) = h(y) + h'(y)/nQ_1(y)M_1(y) \quad (\text{الف} - 4)$$

که در حالت خاص  $h(\theta) = \theta$ ، تمرین ۷-۲۷ از [۵] خواهد شد. برای حالت دوم  $Y = Y_1$  (اولین آماره ترتیبی) و با روشی مشابه حالت اول نتیجه زیر به دست می‌آید:

$$u(y) = h(y) - [h'(y)/nQ_2(y)M_2(y)] \quad (\text{ب} - 4)$$

شایان ذکر است که هر دو حالت فوق توسط دیویس [۲] و تیت [۹] بحث و بررسی شده‌اند.



$$u(y) = \begin{cases} \sum_{x=k}^c \frac{\binom{x-1}{k-1} \binom{y-x-1}{n-k-1}}{\binom{y-1}{n-k-1}} & ; y > c + k(n-1) \\ 1 & ; y \leq c + k(n-1) \end{cases}$$

مثال ۷ (توزیع یکنواخت گسسته)

$$f(x; \theta) = \frac{1}{\theta} \quad x = 1, 2, \dots, \theta$$

$Y = Y_n$  دارای چگالی زیر است

$$g(y; \theta) = \left(\frac{y}{\theta}\right)^n - \left(\frac{y-1}{\theta}\right)^n \quad y = 1, 2, \dots, \theta$$

برای  $h(\theta) = \theta^r$  ( $r > -n$  یک عدد صحیح)، از معادله (۱) داریم

$$\sum_{y=1}^{\theta} u(y) \left[ \left(\frac{y}{\theta}\right)^n - \left(\frac{y-1}{\theta}\right)^n \right] = \theta^r$$

که می‌توان آن را به صورت زیر بازنویسی کرد

$$\sum_{y=1}^{\theta} u(y) \left[ \frac{y^n - (y-1)^n}{y^{n+r} - (y-1)^{n+r}} \right] \left[ \left(\frac{y}{\theta}\right)^{n+r} - \left(\frac{y-1}{\theta}\right)^{n+r} \right] = 1$$

معادله فوق شکل معادله (۶) را دارد که در آن  $g^*(y; \theta^*)$  همان  $g(y; \theta)$  است با این تفاوت که به جای  $n$ ،  $n+r$  قرار داده شده است. بنابراین

$$u(y) = \frac{y^{n+r} - (y-1)^{n+r}}{y^n - (y-1)^n}$$

برای حالت خاص جالب  $h(\theta) = \theta$  داریم

$$u(y) = \frac{y^{n+1} - (y-1)^{n+1}}{y^n - (y-1)^n}$$

که در [۵، ص ۲۳۰] به آن اشاره شده است. حالت خاص دیگر عبارت است از  $h(\theta) = P[X \leq c] = \frac{c(c+1)}{2\theta}$  که برای آن داریم

$$u(y) = \frac{c(c+1)}{2} \times \frac{y^{n-1} - (y-1)^{n-1}}{y^n - (y-1)^n}$$

مثال ۸ (توزیع گاما با  $\alpha$  معلوم)

$$f(x; \theta) = \frac{1}{\theta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\theta}, \quad x > 0, \quad \theta > 0$$

می‌دانیم که  $Y = \sum_{i=1}^n X_i$  توزیع گاما با پارامترهای  $n\alpha$  و  $\theta$  دارد. برای  $h(\theta) = \theta^r e^{-k/\theta}$  ( $r > -n\alpha$ ) از معادله (۱) داریم،

$$\int_0^{\infty} u(y) \frac{y^{n\alpha-1} e^{-y/\theta}}{\theta^{n\alpha} \Gamma(n\alpha)} dy = \theta^r e^{-k/\theta}$$

حالت خاص جالب دیگری عبارت است از

$$h(\theta) = P(X \leq c) = \sum_{x=0}^c \binom{k}{x} \theta^x (1-\theta)^{k-x}$$

که برای آن، مجموع فوق هندسی زیر را به دست می‌آوریم:

$$u(y) = \begin{cases} \sum_{x=0}^c \binom{k}{x} \frac{\binom{n-k-k}{y-x}}{\binom{n-k}{y}} & ; y \geq c \\ 1 & ; y < c \end{cases}$$

این نتیجه در [۳] با استفاده از توزیع شرطی به دست آمده است.

مثال ۶ (توزیع دوجمله‌ای منفی)

$$f(x; \theta) = \binom{x-1}{k-1} \theta^k (1-\theta)^{x-k}$$

$$x = k, k+1, \dots; 0 < \theta < 1$$

می‌دانیم که  $Y = \sum_{i=1}^n X_i$  دارای توزیع دوجمله‌ای منفی با پارامترهای  $n$  و  $\theta$  است. برای  $h(\theta) = \theta^r (1-\theta)^s$  و  $s$  و  $r$  اعداد صحیح و  $r < nk$ ، از معادله (۱) داریم

$$\sum_{y=nk}^{\infty} u(y) \binom{y-1}{nk-1} \theta^{nk} (1-\theta)^{y-nk} = \theta^r (1-\theta)^s$$

که، با در نظر گرفتن  $u(y) = 0$  برای  $y = nk, \dots, nk+s-1$  می‌توان آن را به صورت زیر بازنویسی کرد:

$$\sum_{y=nk+s}^{\infty} \left[ u(y) \frac{\binom{y-1}{nk-1}}{\binom{y-r-s-1}{nk-r-1}} \right] \binom{y-r-s-1}{nk-r-1} \times \theta^{nk-r} (1-\theta)^{y-r-s-(nk-r)} = 1$$

با قرار دادن  $w = y - r - s$ ، رابطه بالا را می‌توان به صورت رابطه (۶) نوشت که در آن  $g^*(w; \theta^*)$  توزیع دوجمله‌ای منفی با پارامترهای  $nk - r$  و  $\theta$  است. بنابراین

$$u(y) = \begin{cases} \frac{\binom{y-r-s-1}{nk-r-1}}{\binom{y-1}{nk-1}} & ; y \geq nk+s \\ 0 & ; y < nk+s \end{cases}$$

یک حالت خاص جالب توجه عبارت است از  $h(\theta) = \theta$ ، که برای آن  $u(y) = \frac{nk-1}{y-1}$  همچنین اگر

$$h(\theta) = P(X \leq c) = \sum_{x=k}^c \binom{x-1}{k-1} \theta^k (1-\theta)^{x-k}$$

آنگاه  $u(y)$  به صورت زیر به دست می‌آید

(۵) توزیع پارتو فرض کنید

$$f(x; \theta) = \frac{\theta}{(1+x)^{\theta+1}} \quad x > 0, \quad \theta > 0$$

برای  $h(\theta) = \theta^{-r} e^{-k/\theta}$  ( $r < n$ ) بیابید.

راهنمایی:  $Y = \sum_{i=1}^n \ln(1+X_i) \sim G(n, \frac{1}{\theta})$  و مشابه مثال ۸ عمل کنید.

(۶) توزیع بتا، حالت خاص فرض کنید

$$f(x; \theta) = \theta x^{\theta-1} \quad 0 < x < 1, \quad 0 < \theta$$

برای  $h(\theta) = \theta^{-r} e^{-k/\theta}$  ( $r < n$ ) بیابید.

راهنمایی:  $Y = -\sum_{i=1}^n \ln X_i \sim G(n, 1/\theta)$  و مشابه مثال ۸ عمل کنید.

(۷) توزیع لاپلاس فرض کنید

$$f(x; \theta) = \frac{1}{2\theta} e^{-\frac{|x|}{\theta}} \quad -\infty < x < \infty, \quad \theta > 0$$

برای  $h(\theta) = \theta^r e^{-k/\theta}$  ( $r > -n$ ) بیابید.

راهنمایی:  $Y = \sum_{i=1}^n |X_i| \sim G(n, \theta)$  و مشابه مثال ۸ عمل کنید.

(۸) واریانس توزیع نرمال فرض کنید  $X \sim N(\mu, \sigma^2)$  و  $\mu, \sigma^2$

هر دو مجهول باشند. با فرض  $\theta = 2\sigma^2$ ،  $h(\theta) = \theta^r e^{-k/\theta}$  برای

برای  $h(\theta) = \theta^r e^{-k/\theta}$  بیابید. از جواب به دست آمده استفاده کنید و

راهنمایی:  $V(s^2) = \frac{2\sigma^4}{n-1}$  برای  $Y = (n-1)s^2 \sim G(\frac{n-1}{2}, 2\sigma^2)$  و مشابه مثال ۸ عمل کنید.

که، با قرار دادن  $u(y) = 0$  برای  $y \leq k$ ، می‌توان آن را به صورت زیر بازنویسی کرد:

$$\int_k^\infty \frac{u(y) \Gamma(n\alpha + r) y^{n\alpha-1} (y-k)^{n\alpha+r-1} e^{-(y-k)\theta}}{\Gamma(n\alpha)(y-k)^{n\alpha+r-1} \theta^{n\alpha+r} \Gamma(n\alpha+r)} dy = 1$$

با جانشینی  $w = y - k$  معادله فوق به صورت معادله (۶) درمی‌آید که در آن  $g^*(w, \theta^*)$  توزیع گاما با پارامترهای  $n\alpha + r - 1$  و  $\theta$  است. بنابراین

$$u(y) = \frac{\Gamma(n\alpha)}{\Gamma(n\alpha+r)} \cdot \frac{(y-k)^{n\alpha+r-1}}{y^{n\alpha-1}}, \quad y > k$$

برای دو حالت خاص به تمرینهای ۳ و ۴ مراجعه کنید.

## چند تمرین

بعضی مطالب در متن اصلی مقاله به کوتاهی اشاره شده‌اند. این مطالب را در قالب چند تمرین که برای کار کلاسی، نیز مفید باشد می‌آوریم.

(۱) روابط (۴-الف) و (۴-ب) را ثابت کنید.

(۲) برای هر یک از توابع چگالی انتهای بخش ۲، و برای  $h(\theta) = \theta^r$   $h(\theta) = P(X \leq c)$ ،  $h(\theta) = P(X \leq c)$  را بیابید.

(۳) در مثال ۸،  $h(\theta) = V(X) = \alpha\theta^r$  را برای  $h(\theta) = V(X)$  بیابید.

(۴) در مثال ۸،  $h(\theta) = E(e^{-tx}) = \frac{1}{(1+\theta t)^\alpha}$  را برای  $h(\theta) = E(e^{-tx})$  بیابید [۸، ص ۴۹۸].

راهنمایی:  $\sum_{t=0}^\infty \frac{\Gamma(\alpha+i)}{i! \Gamma(\alpha)} (-t)^i \theta^i = \frac{1}{(1+\theta t)^\alpha}$

## مراجع

- [1] Barton, D. E. (1961), "Unbiased Estimation of a Set of Probabilities," *Biometrika*, 48, 227-229.
- [2] Davis, R. C. (1951), "On Minimum Variance in Nonregular Estimation," *Annals of Mathematical Statistics*, 22, 43-57.
- [3] Folks, John Leroy, Pierce, Donald A., and Stewart, Charles (1965), "Estimating the Fraction of Acceptable Product," *Technometrics*, 7, No. 1. 43-50.
- [4] Glasser, Gerald J. (1962), "Minimum Variance Unbiased Estimators for Poisson Probabilities," *Technometrics*, 4, No. 3, 409-418.
- [5] Hogg, Robert V., and Craig, Allen T. (1970). *Introduction to Mathematical Statistics*, 3rd. ed., London: The MacMillan Co.

- [6] Laurent, Andre G. (1963), "Conditional Distribution of Order Statistics and the Distribution of the Reduced  $i$ th order Statistic of the Exponential Model," *Annals of Mathematical Statistics*, 34, 652-657.
- [7] Mood, Alexander M., Graybill, Franklin A., and Boes, Duane C. (1974), *Introduction to the Theory of Statistics*, 3rd ed., New York: McGraw-Hill.
- [8] Schaefer, Richard L. (1976), "On the Computation of Certain Minimum Variance Unbiased Estimators," *Technometrics*, 18, No. 4, 497-499.
- [9] Tate, R. F. (1959), "Unbiased Estimation: Functions of Location and Scale Parameters," *Annals of Mathematical Statistics*, 30, 341-366.
- [10] Tenenbein, Richard I. (1971), "The Racing Car Problem," *The American Statistician*, 25, No. 1, 38-40.

---

\* این مقاله، ترجمه (همراه با توضیح و تلخیص) از مقاله‌ای با مشخصات

زیر است:

Guenther, W. C., (1978), "Some Easily Found Minimum Variance Unbiased Estimators", *The American Statistician*, 32, 30-34.

---