Mathematical Researches

Print ISSN: 2588-2546
Online ISSN: 2588-2554

Homepage: https://jemr.khu.ac.ir/

The best unbiased estimators in the absence of sufficiency and completeness

M. Shams[1], **Gh. Hesamian**[2]

1. Corresponding Author, Department of Statistics, Faculty of Mathematical Sciences, University of Kashan, Kashan, Iran, E-mail: mehdishams@kashanu.ac.ir
2. Department of Statistics, Payame Noor University, Tehran 193953697, Iran, E-mail: gh.hesamian@pnu.ac.ir

Article Info

Article type:
Research Article

Article history:
Received: 15 August 2021
Received in revised form: 22 September 2021
Accepted: 21 February 2022
Published online: 6 February 2024

Keywords:
Complete sufficient statistics, minimal sufficient statistics, uniformly minimum variance unbiased estimator, Fisher information, Binary operation.

ABSTRACT

Introduction
One of the goals of mathematical statistics is to estimate and find a good estimate for an unknown parameter of population. In estimation problem, finding a good estimator is relative, and a criterion must be chosen under which an estimator is good. One of these methods is to find the best estimator under a certain class. If we are looking for the best estimator (the minimum risk estimator under given loss) in the class of unbiased estimators, the estimator is called the minimum risk unbiased estimator, and in the special case under the squared error loss, the risk is converted to variance and the uniformly minimum variance unbiased estimator (UMVUE) is found as an optimal estimator. In some cases, the complete sufficient statistic (CSS) does not exist, there may also be non-fixed parameter functions, which are UMVU estimates.

Material and Methods
In this study, a simple generalization of the Lehmann-Scheffe theorem (Lehmann and Scheffe, 1950) is proposed in cases where UMVUEs exist but a CSS does not exist. Also, another method is introduced based on the group action. In this method, UMVUE for the unknown parameter is found using a commutative and associative binary operation. Finally, the motivation for using the words "*completeness*" and "*unbiasedness*" is expressed.

Conclusion
If a CSS exists, every estimable parametric function are UMVUE. Conversely, if each estimable parametric function is an UMVUE, then there will be a CSS under certain conditions. In some cases, although there is not CSS, there may also be non-fixed parametric functions, which are UMVUE. In this paper, generalizations and simple examples are given that UMVUE exists, but not CSS are available. In case there is no CSS and a suitable statistic that is independent of the ancillary statistic is not easily found, conditioning the sufficient statistic on the ancillary statistic will recover the lost information. A UMVUE for an unknown parameter can also be found using a commutative and associative binary operation. It should be noted that completeness and unbiasedness are not characteristic of a statistic or its parametric form, but are characteristic of the family of distributions of a statistic, and deleting even one point of the parameter space can lose the characteristic of completeness.

How to cite: Shams, Mehdi, & Hesamian, Gholamreza, (2023). The best unbiased estimators in the absence of sufficiency and completeness. *Mathematical Researches*, 9 (3), 35 - 55.

 Publisher: Kharazmi University

شاپا چاپی: ۲۵۴۶-۲۵۸۸
شاپا الکترونیکی: ۲۵۵۴-۲۵۸۸

پژوهش‌های ریاضی

Homepage: https://mmr.khu.ac.ir/

بهترین براوردگرهای ناریب در صورت عدم وجود بسندگی و کامل بودن

مهدی شمس[1]✉، غلامرضا حسامیان[2]

۱. نویسنده مسئول، گروه آمار، دانشکده علوم ریاضی، دانشگاه کاشان، کاشان، ایران. ✉رایانامه: mehdishams@kashanu.ac.ir
۲. گروه آمار، دانشگاه پیام نور، تهران، ایران. رایانامه: gh.hesamian@pnu.ac.ir

اطلاعات مقاله

نوع مقاله: مقاله پژوهشی

تاریخ دریافت: ۱۴۰۰/۵/۲۴
تاریخ بازنگری: ۱۴۰۰/۶/۳۱
تاریخ پذیرش: ۱۴۰۰/۱۲/۲
تاریخ انتشار: ۱۴۰۲/۱۱/۱۷

چکیده

در این تحقیق، تعمیم ساده‌ای از قضیه لی‌من-شفه در مواردی که UMVUEها وجود داشته باشند اما آماره بسنده کامل موجود نباشد، مطرح شده است. همچنین روش دیگری بر اساس عمل گروه معرفی می‌شود. در این روش به کمک یک عمل دوتایی جابه‌جایی و شرکت‌پذیر، UMVUE برای پارامتر مجهول پیدا می‌شود. در پایان انگیزه استفاده از واژه‌های «*کامل بودن*» و «*ناریبی*» بیان می‌شود به این صورت که کامل بودن و ناریبی ویژگی آماره یا صورت پارامتری آن نیست، بلکه ویژگی خانواده توزیع‌های یک آماره است و حذف حتی یک نقطه از فضای پارامتر ممکن است کامل بودن را تغییر دهد.

واژه‌های کلیدی:
آماره بسنده کامل،
آماره بسنده مینیمال،
براورد ناریب یکنواخت با کمترین واریانس،
اطلاع فیشر،
عمل دوتایی.

استناد: شمس، مهدی؛ و حسامیان، غلامرضا (۱۴۰۲). بهترین براوردگرهای ناریب در صورت عدم وجود بسندگی و کامل بودن. *پژوهش‌های ریاضی*، ۹ (۳)، ۳۵ – ۵۵.

ناشر: دانشگاه خوارزمی

مقدمه

یکی از اهداف شاخه آمار ریاضی مسئله براورد و یافتن براورد خوب برای پارامتر مجهول جامعه است. در مسئله براورد، خوب بودن یک امر نسبی است و به ناچار باید معیاری انتخاب کرد که تحت آن یک براوردگر خوب باشد. یکی از این روش‌ها یافتن بهترین براوردگر تحت یک کلاس خاص است. اگر در بین کلاس براوردگرهای ناااریب دنبال بهترین براوردگر (براوردگر با کمترین مخاطره تحت زیان داده شده) باشیم، براوردگر حاصل یک براوردگر نااریب با کمترین مخاطره UMRUE[۱] نامیده می‌شود و در حالت خاص که تابع زیان مربع خطا هست، مخاطره به واریانس تبدیل شده و بهترین براوردگر نارایب با کمترین واریانس (UMVUE[۲]) به عنوان یک براوردگر بهینه پیدا می‌شود. گاسپرونی و همکاران (۲۰۲۰) مطالب آموزشی مفیدی در مورد UMVUE ارائه کردند. کوبروسلی (۲۰۲۲)، UMVUE برای تابع توزیع گاما با پارامتر مقیاس صحیح و معلوم پیدا کرد. چیچاگوف (۲۰۲۲)، یک مقایسه مجانبی بین MLE[۳] و UMVUE، برای خانواده توزیع‌های تک‌پارامتری بریده ارائه کرد. هیروسه و مانو (۲۰۲۳)، UMVUE مجانبی و کاربرد آن در نمونه‌گیری را تحلیل کردند. در استادلارد و کیمانی (۲۰۱۸) کاربردهای UMVUE در آزمایش‌های بالینی چندمرحله‌ای مطرح شد. وربیک و همکاران (۲۰۲۱) در مورد نااریبی، کارایی و همچنین یافتن UMVUE پارامترهای احتمالی نظیر $P(X>Y)+\frac{1}{2}P(X=Y)$ و کاربرد آن در نظریه قابلیت اطمینان، دقت تشخیص و آزمایش‌های بالینی تحقیقات قابل توجهی انجام دادند. کاویتا و شارما (۲۰۲۲) بر پایه براورد به روش‌های MLE و همچنین UMVUE، براورد بهینه برای پارامترهای قابلیت اطمینان نرم‌افزارها پیدا کردند. موریال و همکاران (۲۰۲۱)، UMVUE برای ابعاد فراکتال را محاسبه کردند. سرینیواس و کیل (۲۰۱۶)، MLE و UMVUE برای پارامترهای مدل صف‌بندی یک‌باجه‌ای با ورودی مارکف و سرویس ثابت پیدا کردند.

قضیه لی‌من-شفه (لی‌من و شفه، ۱۹۵۰) که یکی از قضایای معروف در استنباط آماری است، یکی از روش‌های یافتن این برآورگرهای بهینه است و بیان می‌کند، اگر آماره بسنده کامل وجود داشته باشد، همه توابع پارامتری براوردپذیر، UMVUE هستند (لی‌من و رومانو، ۲۰۰۵). وارون قضیه نیز بیان می‌کند، اگر هر تابع پارامتری براوردپذیر یک UMVUE باشد، آن‌گاه آماره بسنده کامل (یا به‌طور دقیق‌تر میدان سیگمایی آماره بسنده و کامل) به شرطی که خانواده اندازه‌های احتمال، توسط برخی اندازه‌های $\sigma-$متناهی مغلوب شده باشد، وجود خواهد داشت (بهادر، ۱۹۵۷). با این حال در بعضی مواقع اگر آماره بسنده کامل موجود نباشد نیز ممکن است توابع پارامتری غیرثابتی وجود داشته باشند، که براورد UMVU باشند. در متون درسی کارشناسی و تحصیلات تکمیلی، به ندرت به این موضوع اشاره شده و یا مثالی راجع به آن آورده شده است. در این مقاله، تعمیم‌های ساده‌ای از قضیه لی‌من-شفه (لی‌من و شفه، ۱۹۵۰) مطرح شده است. در مواردی که UMVUE وجود داشته باشد، اما آماره بسنده کامل موجود نباشد، مثال‌های کاربردی از قضیه لی‌من-شفه (لی‌من و شفه، ۱۹۵۰) قابل بیان است. نتیجه بیان شده در فراسر (۱۹۵۶) حالت خاصی از این موضوع است. برای مواردی که UMVUE وجود دارد، اما آماره بسنده کامل موجود نیست، روشی دیگر نیز ارائه شده است. در بخش اول به حالتی که آماره بسنده کامل وجود داشته

[1] Uniformly Minimum Risk Unbiased Estimator
[2] Uniformly Minimum Variance Unbiased Estimator
[3] Maximum Likelihood Estimator

باشد پرداخته در انتها به این مورد اشاره می‌شود که شرطی کردن باعث بازیابی اطلاعات از دست رفته خواهد شد. در انتها چند مثال ذکر می‌شود که UMVUE موجود است، اما آماره بسنده کامل موجود نیست. در بخش دوم روش دیگری بر اساس عمل گروه معرفی می‌شود. در این روش به کمک یک عمل دوتایی جابه‌جایی و شرکت‌پذیر، UMVUE برای پارامتر مجهول پیدا می‌شود. در بخش سوم انگیزه استفاده از واژه‌های «*کامل بودن*» و «*ناریبی*» بیان می‌شود. همچنین این حقیقت بررسی می‌شود که کامل بودن یک ویژگی خانواده توزیع‌ها است تا متغیر تصادفی یا صورت پارامتری و حذف حتی یک نقطه از فضای پارامتر ممکن است کامل بودن را تغییر دهد. همچنین ناریبی، همانند کامل بودن، ویژگی آماره یا صورت پارامتری آن نیست، بلکه ویژگی خانواده توزیع‌های یک آماره است.

۱. براورد ناریب یکنواخت، با کمترین واریانس وقتی آماره بسنده کامل وجود ندارد

همانطور که اشاره شد در برخی مواقع که آماره بسنده کامل موجود نباشد نیز ممکن است توابع پارامتری غیرثابتی وجود داشته باشند، که براورد UMVU باشند. در روهاتگی (۱۹۷۶، ص۳۵۶) یک مثال برای شرح این حقیقت آورده شده است که در ذیل به آن اشاره می‌کنیم:

مثال ۱.۱. فرض کنید $X_1,\dots,X_n$ یک نمونه تصادفی از توزیع نمایی با میانگین θ و $Y_1,\dots\dots,Y_n$ نمونه تصادفی از توزیع نمایی با میانگین $\frac{1}{\theta}$ است و X_iها و Y_iها مستقل‌اند. براوردی از θ بر اساس نمونه $(X_1,\dots,X_n;Y_1,\dots,Y_n)$ را در نظر بگیرید. قرار دهید $S_1(\mathbf{x})=\sum_{i=1}^{n}x_i$ و $S_2(\mathbf{y})=\sum_{i=1}^{n}y_i$. بنابراین:

$$\begin{aligned} f_\theta(x_1,\dots,x_n,y_1,\dots,y_n) &= \prod_{i=1}^{n} f_\theta(x_i)\prod_{i=1}^{n} f_\theta(y_i) \\ &= (\tfrac{1}{\theta})^n e^{-\frac{1}{\theta}\sum_{i=1}^{n}x_i}\theta^n e^{-\theta\sum_{i=1}^{n}y_i} \\ &= e^{-(\frac{1}{\theta}S_1(\mathbf{x})+\theta S_2(\mathbf{y}))}. \end{aligned}$$

در این صورت $(S_1(\mathbf{X}),S_2(\mathbf{Y}))$ برای θ بسنده است. از این که نسبت درستنمایی برای $\theta_0=1$، یعنی

$$\frac{f_\theta(\mathbf{x},\mathbf{y})}{f_{\theta_0}(\mathbf{x},\mathbf{y})}=\frac{e^{-(\frac{1}{\theta}S_1(\mathbf{x})+\theta S_2(\mathbf{y}))}}{e^{S_1(\mathbf{x})+S_2(\mathbf{y})}}=e^{-(\frac{1}{\theta}+1)S_1(\mathbf{x})-(\theta+1)S_2(\mathbf{y})}$$

فقط از طریق آماره $(S_1(\mathbf{X}),S_2(\mathbf{Y}))$ به θ بستگی دارد، لذا (S_1,S_2) آماره بسنده مینیمال برای θ است. یاداوری این نکته ضروری است که انتخاب $\theta_0=1$ با توجه به این حقیقت است که در توزیع نمایی، θ پارامتر مکان است و ۱ عضو خنثی در یک مجموعه که نسبت به عمل ضرب گروه است. آماره‌های

$$S(\mathbf{X},\mathbf{Y})=\left(\frac{S_1(\mathbf{X})}{S_2(\mathbf{Y})}\right)^{\frac{1}{2}}$$
$$A(\mathbf{X},\mathbf{Y})=(S_1(\mathbf{X})S_2(\mathbf{Y}))^{\frac{1}{2}}$$

را در نظر بگیرید. وارون تبدیل دستگاه معادلات بالا به صورت زیر خواهد بود:

$$\begin{cases} S_1(\mathbf{X})=S(\mathbf{X},\mathbf{Y})A(\mathbf{X},\mathbf{Y}) \\ S_2(\mathbf{Y})=\dfrac{A(\mathbf{X},\mathbf{Y})}{S(\mathbf{X},\mathbf{Y})}. \end{cases}$$

ژاکوبی تبدیل اخیر برابر است با:

$$J=\begin{vmatrix} \dfrac{\partial S_1(\mathbf{X})}{\partial A(\mathbf{X},\mathbf{Y})} & \dfrac{\partial S_1(\mathbf{X})}{\partial S(\mathbf{X},\mathbf{Y})} \\ \dfrac{\partial S_2(\mathbf{Y})}{\partial A(\mathbf{X},\mathbf{Y})} & \dfrac{\partial S_2(\mathbf{Y})}{\partial S(\mathbf{X},\mathbf{Y})} \end{vmatrix}=\begin{vmatrix} S(\mathbf{X},\mathbf{Y}) & A(\mathbf{X},\mathbf{Y}) \\ \dfrac{1}{S(\mathbf{X},\mathbf{Y})} & \dfrac{-A(\mathbf{X},\mathbf{Y})}{S^2(\mathbf{X},\mathbf{Y})} \end{vmatrix}$$
$$=\frac{-A(\mathbf{X},\mathbf{Y})}{S(\mathbf{X},\mathbf{Y})}-\frac{A(\mathbf{X},\mathbf{Y})}{S(\mathbf{X},\mathbf{Y})}=\frac{-2A(\mathbf{X},\mathbf{Y})}{S(\mathbf{X},\mathbf{Y})}.$$

از این رو با توجه به استقلال $S_1(\mathbf{X})$ و $S_2(\mathbf{Y})$، تابع چگالی احتمال توأم A و S به صورت زیر است:

$$f_{A(\mathbf{X},\mathbf{Y}),\mathbf{S}(\mathbf{X},\mathbf{Y})}(A(\mathbf{x},\mathbf{y}),S(\mathbf{x},\mathbf{y}))=f_{S_1(\mathbf{X})}(S(\mathbf{x},\mathbf{y})A(\mathbf{x},\mathbf{y}))f_{S_2(\mathbf{Y})}\left(\frac{A(\mathbf{x},\mathbf{y})}{S(\mathbf{x},\mathbf{y})}\right)\left|\frac{-2A(\mathbf{x},\mathbf{y})}{S(\mathbf{x},\mathbf{y})}\right|$$
$$=\frac{(S(\mathbf{x},\mathbf{y})A(\mathbf{x},\mathbf{y}))^{n-1}\quad e^{-S(\mathbf{x},\mathbf{y})A(\mathbf{x},\mathbf{y})\theta}}{\Gamma(n)\theta^n}\times\frac{\theta^n\left(\frac{A(\mathbf{x},\mathbf{y})}{S(\mathbf{x},\mathbf{y})}\right)^{n-1}\quad e^{\frac{-A(\mathbf{x},\mathbf{y})}{S(\mathbf{x},\mathbf{y})}\theta}}{\Gamma(n)}\times\frac{2A(\mathbf{x},\mathbf{y})}{S(\mathbf{x},\mathbf{y})}$$
$$=\frac{2}{[\Gamma(n)]^2}\exp\left\{-A(\mathbf{x},\mathbf{y})\left(\frac{S(\mathbf{x},\mathbf{y})}{\theta}+\frac{\theta}{S(\mathbf{x},\mathbf{y})}\right)\right\}\frac{[A(\mathbf{x},\mathbf{y})]^{2n-1}}{S(\mathbf{x},\mathbf{y})}.$$

واضح است که A و S از هم مستقل نیستند. توزیع A دارای تابع چگالی احتمال $C(\mathbf{x},\mathbf{y})[A(\mathbf{x},\mathbf{y})]^{2n-1}$ است که در آن $C(\mathbf{x},\mathbf{y})$ ثابت انتگرال‌گیری است که به $\mathbf{x}$، $\mathbf{y}$ و n وابسته است، اما به θ بستگی ندارد. در حقیقت $C(\mathbf{x},\mathbf{y})=4K_0[2A(\mathbf{x},\mathbf{y})]/[\Gamma(n)]^2$ به طوری که K_0 صورت استانداردی از تابع بسل است. در نتیجه A یک آماره کمکی برای θ است. تابع چگالی احتمال شرطی S به شرط $A=a$ به صورت زیر داده شده است:

$$f_{S(\mathbf{X},\mathbf{Y})|A(\mathbf{X},\mathbf{Y})=a}(S(\mathbf{x},\mathbf{y})\,|\,A(\mathbf{x},\mathbf{y})=a)=\frac{f_{A(\mathbf{X},\mathbf{Y}),S(\mathbf{X},\mathbf{Y})}(a,S(\mathbf{x},\mathbf{y}))}{f_{A(\mathbf{X},\mathbf{Y})}(a)}$$
$$=\frac{\dfrac{2\exp\left\{-a\left[\frac{S(\mathbf{x},\mathbf{y})}{\theta}+\frac{\theta}{S(\mathbf{x},\mathbf{y})}\right]\right\}a^{2n-1}}{\Gamma(n)^2 S(\mathbf{x},\mathbf{y})}}{\dfrac{4K_0[2a]a^{2n-1}}{\Gamma(n)^2}}$$
$$=\frac{\exp\left\{-a\left(\frac{S(\mathbf{x},\mathbf{y})}{\theta}+\frac{\theta}{S(\mathbf{x},\mathbf{y})}\right)\right\}}{2K_0[2a]S(\mathbf{x},\mathbf{y})}.$$

مثال ۲.۱. فرض کنید (X, Y) دارای تابع چگالی احتمال

$$f_\theta(x,y)=\exp\left\{-\left(\tfrac{x}{\theta}+\theta y\right)\right\},\qquad x>0, y>0$$

است. برای تک‌مشاهده (x,y) داریم:

$$\frac{\partial}{\partial\theta}\log f_\theta(x,y)=\frac{\partial}{\partial\theta}\left[-\left(\tfrac{x}{\theta}+\theta y\right)\right]=\frac{x}{\theta^2}-y.$$

از این که:

$$E_\theta(Y^2)=Var(Y)+(E(Y))^2=\frac{1}{\theta^2}+\frac{1}{\theta^2}=\frac{2}{\theta^2},$$
$$E_\theta(X^2)=Var(X)+(E(X))^2=\theta^2+\theta^2=2\theta^2,$$
$$E(XY)=E(X)E(Y)=\theta.\frac{1}{\theta}=1$$

می‌توان نتیجه گرفت که اطلاعات نهفته در (X, Y) راجع به پارامتر مجهول θ برابر است با:

$$I(\theta)=E_\theta(Y^2)+\frac{E(X^2)}{\theta^4}-\frac{2E(XY)}{\theta^2}=\frac{2}{\theta^2}+\frac{2}{\theta^2}-\frac{2}{\theta^2}=\frac{2}{\theta^2}$$

از این‌رو اطلاع فیشر در یک نمونه n تایی از توزیع داده شده برابر است با $\frac{2n}{\theta^2}$.

در مثال ۱،۱، $X_1,\ldots,X_n \overset{i.i.d.}{\sim} \Gamma(1,\frac{1}{\theta})$ و $Y_1,\ldots,Y_n \overset{i.i.d.}{\sim} \Gamma(1,\theta)$ و X_iها و Y_iها از هم مستقل‌اند. از این رو (X_1,Y_1) دارای توزیع $f_\theta(x,y)$ است. میزان اطلاع فیشر برای θ را در خانواده‌ای با تابع چگالی احتمال متناظر با $S(\mathbf{X},\mathbf{Y})=\sqrt{\sum_{i=1}^n X_i \Big/ \sum_{i=1}^n Y_i}$ محاسبه می‌کنیم. با استفاده از توابع چگالی احتمال $\sum_{i=1}^n X_i \sim \Gamma(n,\frac{1}{\theta})$ و

$\sum_{i=1}^{n} Y_i \sim \Gamma(n,\theta)$ و با استفاده از تکنیک تبدیل احتمال به راحتی می‌توان دید، $S(\mathbf{X},\mathbf{Y})$ دارای تابع چگالی احتمال $g_\theta(s)=\frac{2\Gamma(2n)}{[\Gamma(n)]^2}s^{-1}\left(\frac{s}{\theta}+\frac{\theta}{s}\right)^{-1}$ است. بنابراین با توجه به رابطه $\frac{\partial \log g_\theta(s)}{\partial\theta}=-2n\left(-\frac{s}{\theta^2}+\frac{1}{s}\right)\left(\frac{s}{\theta}+\frac{\theta}{s}\right)^{-1}$ داریم:

$$E_\theta\left\{\frac{\partial}{\partial\theta}\log g_\theta(S)\right\}^2=\frac{4n^2}{\theta^2}E_\theta\left\{1-4\left(\frac{S}{\theta}+\frac{\theta}{S}\right)^{-2}\right\}$$
$$=\frac{4n^2}{\theta^2}\left\{1-4\frac{n}{2(2n+1)}\right\}$$
$$=\frac{2n}{\theta^2}\left(\frac{2n}{2n+1}\right)<\frac{2n}{\theta^2}$$

که عبارت بالا بیان‌کننده اطلاعات θ در S است که از اطلاعات موجود در نمونه کمتر است. می‌توان نشان داد اطلاع فیشر در تابع چگالی احتمال شرطی S به شرط $A=a$ که $A(\mathbf{X},\mathbf{Y})=\sqrt{S_1(\mathbf{X})S_2(\mathbf{Y})}$ برابر با $\frac{2a}{\theta^2}\frac{K_1(2a)}{K_0(2a)}$ است، که K_0 و K_1 به ترتیب تابع بسل مرتبه 0 و 1 هستند.

قضیه باسو (باسو، ۱۹۵۵ و ۱۹۵۸) بیان می‌کند آماره بسنده کران‌دار از آماره کمی مستقل است. روهاتگی (۱۹۷۶، ص۳۵۸-۳۵۶) در مورد یکی از نتایج عکس این قضیه مطالبی را به صورت زیر بازنویسی می‌کند:

عکس قضیه باسو (باسو، ۱۹۵۵) درست نیست. باسو (۱۹۵۸) به اشتباه نشان داد که عکس قضیه‌اش برقرار است. برای این که آماره S از هر آماره کمکی مستقل باشد، نیازی به کامل بودن آماره نیست (برای مثال لی‌من، ۱۹۸۱ را ببینید). مثال ۱٫۱ که توسط فیشر بیان شده است، نشان می‌دهد، اگر آماره بسنده کامل برای θ وجود نداشته باشد، اما نتوانیم آماره مناسبی پیدا کنیم که از آماره کمکی A مستقل باشد، در برخی موارد توسط شرطی کردن می‌توان از روی آماره کمکی اطلاعات مفقود شده را بازیابی کرد.

مثالی در رائو (۱۹۷۳، ص۳۷۹) بیان شده است که در این بخش به بررسی آن می‌پردازیم.

مثال ۳.۱. متغیر تصادفی X با تابع جرم احتمال زیر را در نظر بگیرید:

$$\begin{cases} P(X=-1)=\alpha \\ P(X=n)=(1-\alpha)^2\alpha^n, \quad n=0,1,2,\ldots,0<\alpha<1. \end{cases}$$

الف) متغیر تصادفی X آماره بسنده مینیمال برای α است، اگر چه کامل نیست.

تابع احتمال توأم $\mathbf{X}=(X_1,\ldots,X_n)$ را می‌توان به صورت زیر نوشت:

$$\begin{aligned} f_{\alpha}(x) &= \exp\{[(\ln\alpha)c(x+1)]+[2\ln(1-\alpha)+x\ \ln\alpha]u(x)\} \\ &= \exp\{[(\ln\alpha-2\ln(1-\alpha)]c(x+1)+2\ln(1-\alpha)c(x+1)+2\ln(1-\alpha)u(x)+(x\ \ln\alpha)u(x)\} \\ &= \exp\{c_1(\alpha)c(x+1)+xc_2(\alpha)u(x)+2\ln(1-\alpha)\} \end{aligned}$$

که در آن

$$\begin{aligned} &c_1(\alpha)=\ln\alpha-2\ln(1-\alpha), \\ &c_2(\alpha)=\ln\alpha, \\ &c(t)=\begin{cases}1 & t=0\\ 0 & t\neq 0,\end{cases} \\ &u(t)=\begin{cases}1 & t\geq 0\\ 0 & t<0.\end{cases} \end{aligned}$$

برای نمونه تصادفی $X_1,\ldots,X_n$، تابع چگالی احتمال توأم به صورت زیر است

$$\begin{aligned} f_{\alpha}(\mathbf{x}) &= \exp\{\sum_{i=1}^{n}c_1(\alpha)c(x_i+1)+\sum_{i=1}^{n}c_2(\alpha)x_i u(x_i)+2n\ \ln(1-\alpha)\} \\ &= g(T_1(\mathbf{x}),T_2(\mathbf{x}),\alpha)h(\mathbf{x}). \end{aligned}$$

طبق قضیه دسته‌بندی نیمن-فیشر (لی‌من و رومانو، ۲۰۰۵) ، $(T_1(\mathbf{x}),T_2(\mathbf{x}))$ آماره بسنده برای α است که در آن $T_1(\mathbf{x})=\sum_{i=1}^{n}c(x_i+1)$ و $T_2(\mathbf{x})=\sum_{i=1}^{n}x_i u(x_i)$. مشاهده می‌شود که نسبت

$$\frac{f_{\alpha}(\mathbf{x})}{f_{\alpha}(\mathbf{y})}=\exp\{c_1(\alpha)[T_1(\mathbf{x})-T_1(\mathbf{y})]+c_2(\alpha)[T_2(\mathbf{x})-T_2(\mathbf{y})]\}$$

زمانی به α بستگی ندارد که $(T_1(\mathbf{x}),T_2(\mathbf{x}))=(T_1(\mathbf{y}),T_2(\mathbf{y}))$ و از این رو طبق روش لی‌من-شفه (لی‌من و شفه، ۱۹۵۰) ، $(T_1(\mathbf{X}),T_2(\mathbf{X}))$ یک آماره بسنده مینیمال برای α است. اما این آماره کامل نیست. برای اثبات این حقیقت برای هر $\alpha\in(0,1)$، داریم:

$$\begin{aligned} E_{\alpha}(X_i) &= (-1)\alpha+\sum_{i=0}^{\infty}x\,\alpha^{x}(1-\alpha)^2 \\ &= -\alpha+(1-\alpha)\sum_{i=0}^{\infty}x\,\alpha^{x}(1-\alpha) \qquad (۱) \\ &= -\alpha+(1-\alpha)\frac{\alpha}{(1-\alpha)}=0. \end{aligned}$$

از طرف دیگر اتحاد $x_i = x_i u(x_i) - c(x_i + 1)$ به راحتی ثابت می‌شود. برای بررسی این اتحاد مشاهده می‌شود در حالتی که $x_i = -1$، هر دو طرف معادله برابر هستند ($RHS = (-1)u(-1) - c(0) = -1 = LHS$) و برای مقادیر $x_i = 0,1,2,...$ نیز دو طرف معادله با هم مساوی هستند ($RHS = x_i u(x_i) - c(x_i + 1) = x_i = LHS$). با گرفتن مجموع از دو طرف این اتحاد داریم:

$$\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} X_i u(X_i) - \sum_{i=1}^{n} c(X_i + 1) = T_2 - T_1.$$

بنابراین نتیجه می‌شود که برای هر $\alpha \in (0,1)$، $E\left(\sum_{i=1}^{n} X_i\right) = E_\alpha(T_2 - T_1)$ و در پی آن به کمک رابطه (1):

$$E_\alpha(T_2 - T_1) = \sum_{i=1}^{n} E_\alpha(X_i) = 0.$$

اما برای هر $\alpha \in (0,1)$، $P_\alpha(T_2 - T_1 = 0) = P_\alpha\left(\sum_{i=1}^{n} X_i = 0\right) < 1$. پس آماره بسنده برای α کامل نیست.

ب) تابعی از X که برای صفر ناریب است، لزوماً به صورت cX است. نشان می‌دهیم آماره

$$T(X) = \begin{cases} 1 & X = 0 \\ 0 & o.w. \end{cases}$$

براوردگر ناریب با کمترین واریانس برای $(1-\alpha)^2$ است.

همان‌طور که در قسمت (الف) نشان دادیم، خانواده توزیع‌های داده شده کامل نیست. با استفاده از قضیه لی‌من-شفه (لی‌من و شفه، ۱۹۵۰) نشان می‌دهیم که آماره $T(X)$ براوردگر $UMVU$ پارامتر $(1-\alpha)^2$ است. برای این منظور ابتدا نشان می‌دهیم کلاس براوردگرهای ناریب صفر برابر با $\mathcal{U}_0 = \{g(x) : g(x) = cx, c \neq 0, x = -1,0,1,2,...\}$ است. برای هر $0 < \alpha < 1$ داریم:

$$0 = E(g(X)) = \alpha g(-1) + \sum_{x=0}^{\infty} g(x)(1-\alpha)^2 \alpha^x$$

و یا $-g(-1)(1-\alpha)^{-2}\alpha = \sum_{x=0}^{\infty} g(x)\alpha^x$. از طرفی

$$\begin{aligned}(1-\alpha)^{-2} &= \sum_{r=0}^{\infty}\binom{-2}{r}(-\alpha)^r \\ &= \sum_{r=0}^{\infty}(-1)^r(r+1)(-\alpha)^r \qquad (۲) \\ &= \sum_{r=0}^{\infty}(r+1)\alpha^r .\end{aligned}$$

پس

$$\sum_{x=0}^{\infty} g(x)\alpha^x = -g(-1)\sum_{r=0}^{\infty}(r+1)\alpha^{r+1} = -g(-1)\sum_{u=1}^{\infty}u\,\alpha^u$$

که در تساوی آخر از تغییر متغیر $r+1=u$ استفاده شده است. با مساوی قرار دادن ضرایب سری‌های توانی، داریم:

$$\begin{aligned}&g(x)=cx \qquad x=1,2,\ldots\\ &g(0)=0\\ &g(-1)=-c\end{aligned}$$

و از این رو $g(x)=cx$ ، $x=-1,0,1,2,\ldots$ که کلاس براوردگرهای نااریب صفر را مشخص می‌کند. بنابراین مسئله تعیین براوردگر نااریب بهینه که دارای کمترین واریانس در $\alpha=\alpha_0$ باشد، منجر به تعیین مقدار b می‌شود که کمیت

$$E_{\alpha_0}[(T(X)+bg(X))^2]=\sum_{x=-1}^{\infty}(T(x)+bcx)^2 f_{\alpha_0}(x).$$

را مینیمم کند. با انتخاب $a=bc$، سمت راست عبارت بالا به صورت زیر ساده می‌شود:

$$\sum_{x=-1}^{\infty}(T(x)+ax)^2 f_{\alpha_0}(x)=a^2\alpha_0+a^2\sum_{x=-1}^{\infty}x^2(1-\alpha_0)^2\alpha_0^x = k_1(a).$$

بنابراین:

$$\begin{aligned}k_1'(a)&=2a\alpha_0+2a(1-\alpha_0)^2\sum_{x=0}^{\infty}x^2\alpha_0^{\,x},\\ k_1''(a)&=2\alpha_0+2(1-\alpha_0)^2\sum_{x=0}^{\infty}x^2\alpha_0^{\,x}>0.\end{aligned}$$

از آن‌جا که $k_1''(a)>0$، تابع صعودی $k_1(a)$ مینیمم خود را در $a_0^2=0$ اختیار می‌کند که واریانس $T(X)+bg(X)$ را در α_0 مینیمم می‌کند و به مقدار α_0 وابسته نیست، بنابراین $T(X)$، براوردگر نااریب با کمترین واریانس برای $(1-\alpha)^2$ خواهد بود.

ج) هیچ $UMVUE$ برای α وجود ندارد، اگر چه براوردگر ناااریب برای α موجود است.

فرض می‌کنیم $\delta(X)$ براوردگر نااریب برای α باشد. پس

$$\alpha = E(\delta(X)) = \delta(-1)\alpha + \sum_{x=0}^{\infty} \alpha^x \delta(x)(1-\alpha)^2$$

و در پی آن طبق اتحاد (۲) و تغییر متغیر $r+1=u$ داریم:

$$\begin{aligned}
\sum_{x=0}^{\infty} \delta(x)\alpha^x &= (1-\delta(-1))\alpha(1-\alpha)^{-2} \\
&= \alpha(1-\delta(-1))\sum_{r=0}^{\infty}(r+1)\alpha^r \\
&= (1-\delta(-1))\sum_{r=0}^{\infty}(r+1)\alpha^{r+1} \\
&= (1-\delta(-1))\sum_{u=1}^{\infty}(u)\alpha^u .
\end{aligned}$$

با مساوی قرار دادن ضرایب توانی، داریم:

$$\begin{aligned}
&\delta(0) = 0 \\
&\delta(-1) = 1-a \\
&\delta(x) = ax, \quad x = 1,2,\ldots
\end{aligned}$$

در نتیجه براوردگر نااریب برای α وجود دارد، اما اگر $UMVUE$ مثل $\delta(X)$ برای α وجود داشته باشد، طبق قضیه لی‌من-شفه (لی‌من و شفه، ۱۹۵۰)، باید برای هر عضو aX (متعلق به مجموعه براوردگرهای نااریب α)، $\delta(X)$ و X ناهمبسته باشند. بنابراین

$$\begin{aligned}
0 = Cov\left(\delta(X), aX\right) &= aE\left(X\,\delta(X)\right) \\
&= \sum_{x=-1}^{\infty} x\,\delta(x) f_\alpha(x) \\
&= a[-\alpha\delta(-1) + (1-\alpha)\sum_{x=0}^{\infty} x\,\alpha^x(1-\alpha)] \\
&= a[-\alpha\delta(-1) + (1-\alpha)\frac{\alpha}{(1-\alpha)}] \\
&= a[\alpha(1-\delta(-1))] = a^2\alpha
\end{aligned}$$

که تناقض است، زیرا $a^2\alpha \neq 0$. لذا طبق قضیه لی‌من-شفه (لی‌من و شفه، ۱۹۵۰)، $\delta(X)$ براوردگر نااریب با کمترین واریانس برای α نیست.

در ادامه این بخش، حالتی که آماره بسنده کامل وجود نداشته باشد، ولی UMVUE موجود است بررسی می‌شود. همه نتایج این قسمت مشابه قضیه لی‌من-شفه (لی‌من و شفه، ۱۹۵۰)، برای هر تابع زیان محدب دلخواه معتبر هستند، اما برای سادگی فقط برای تابع زیان مربع خطا بیان می‌شوند. همه براوردگرهایی که در نظر می‌گیریم، دارای واریانس متناهی هستند، بنابراین عبارت «*با واریانس متناهی*» در ادامه بخش حذف شده است. یک آماره را UMVUE گوییم، هرگاه برای امید ریاضی خودش، UMVUE باشد (لی‌من و شفه، ۱۹۵۰).

فرض کنیم خانواده اندازه‌های احتمال دارای فضای نمونه $\mathcal{X}$ به صورت $\{P_{\theta,\phi}, \theta \in \Theta, \phi \in \Phi\}$ هستند. یک عضو تصادفی $\mathcal{X}$ را با X نشان می‌دهیم.

لم ۱.۱. اگر برای هر ϕ ثابت، آماره $S(X)$ یک UMVUE باشد، آن‌گاه $S(X)$، UMVUE خواهد بود.

اثبات. می‌دانیم اگر آماره $S_1(X,\phi)$ تابعی از ϕ باشد و برای هر مقدار ثابت ϕ، دارای امید ریاضی باشد که برابر با امید ریاضی S است، به ازای هر θ,ϕ داریم:

$$Var_{\theta,\phi}[S] \leq Var_{\theta,\phi}[S_1(X,\phi)].$$

بنابراین واریانس هر مقدار مشخص S، کمتر از واریانس براوردگر نااریب آن است. پس S یک UMVUE است. ■

قضیه ۱.۱. اگر $T(X)$ به ازای هرمقدار ثابت ϕ، آماره بسنده کامل برای θ باشد، آن‌گاه:

الف) اگر h به عنوان تابعی از T، براوردگر نااریب برای $\Psi(\theta,\phi)$ باشد، آن‌گاه h یکتا است و $h(T)$ یک UMVUE برای $\Psi(\theta,\phi)$ است.

ب) علاوه بر آن اگر توزیع T وابسته به ϕ نباشد، آن‌گاه هر تابع پارامتری براوردپذیر $\Psi(\theta)$ می‌تواند براوردگر UMVU باشد و UMVUE از امید شرطی مستقل از θ، یعنی $E_{\phi_0}(S|T)$ به‌دست می‌آید، به طوری که S هر براوردگر نااریب از $\Psi(\theta)$ و ϕ_0 دلخواه است.

اثبات. قسمت الف) اگر T به ازای هر مقدار ثابت ϕ، برای θ کامل باشد، آن‌گاه برای (θ,ϕ) نیز چنین است. اثبات بلافاصله از قضیه لی‌من-شفه (لی‌من و شفه، ۱۹۵۰) که ادعا می‌کند هر تابعی از آماره بسنده کامل یک UMVUE است و لم ۱,۱ به دست می‌آید.

قسمت ب) با توجه به این که توزیع $E_{\phi_0}(S|T)$ مستقل از ϕ است، از این رو:

$$E_{\theta,\phi}[E_{\phi_0}[S \mid T]] = E_{\theta,\phi}[E_{\theta,\phi_0}[S \mid T]] = E_{\theta,\phi_0}[S] = \psi(\theta)$$

و حکم اثبات می‌شود. ■

قضیه ۱،۱، از قضیه لی‌من-شفه (لی‌من و شفه، ۱۹۵۰) قوی‌تر است، زیرا یک آماره می‌تواند به ازای هر ϕ ثابت برای θ بسنده باشد، بدون این که برای (θ,ϕ) بسنده باشد.

مهم‌ترین بخش قضیه ۱،۱، یعنی قسمت (الف) می‌تواند به صورت زیر گسترش یابد که اثبات آن مشابه اثبات قضیه ۱،۱ است:

اگر خانواده اندازه احتمال $\mathcal{P}$ روی $\mathcal{X}$ را بتوان به صورت اجتماعی از مجموعه‌های $\{\mathrm{P}_\phi, \phi \in \Phi\}$ نوشت که P_ϕها لزوماً مجزا نیستند و آماره $T\ (X)$ به ازای هر ϕ ثابت، برای P_ϕ بسنده و کامل باشد، آن‌گاه هر تابعی از T (نسبت به $\mathcal{P}$) یک UMVUE است.

با توجه به قضیه ۱،۱ یاداوری می‌شود که هر تابعی از UMVUE همیشه UMVUE نیست (بهادر، ۱۹۵۷). آزمایش تصادفی X که دارای دو نتیجه است را در نظر بگیرید. برآمد اول T، یک روند تصادفی مستقل از θ، و برآمد دوم X، یک روند تصادفی مستقل از ϕ و مشاهده T است که شامل هیچ‌گونه اطلاعات اضافی درمورد θ نیست. بدیهی است که آزمایش دوم مستقل از آزمایش اول است. دو مثال غیربدیهی با مفروضات اضافه قسمت (ب) در بارندورف- نیلسن (۱۹۷۸) آورده شده که در ادامه به ذکر آن‌ها می‌پردازیم.

مثال ۴.۱. (فرایند پواسون با روند لُگ- خطی). برای یک فرایند پواسون با تابع شدت $\lambda(t) = e^{\alpha+\beta t}$ که مشاهدات در فاصله زمانی $[0,T]$ روی می‌دهند، n را تعداد رویدادها در این فاصله زمانی و $t_1 < t_2 < \ldots < t_n$ را زمان‌هایی که این مشاهدات اتفاق می‌افتد در نظر بگیرید. در این حالت n نیز دارای توزیع پواسون با میانگین $e^x \int_0^T e^{\beta t} dt$ است و برای n معلوم، بردار t_* دارای تابع چگالی احتمال

$$n!\left(\frac{\beta}{e^{\beta t}-1}\right)^n e^{\beta t_*}$$

است. در ضمن آماره ترتیبی نمونه n تایی دارای توزیع نمایی بریده در بازه $[0,T]$ است. در نتیجه (α,β) در $\mathbb{R}^2$ تغییر می‌کنند و تعداد رویدادهای n گسسته است، از این رو استنباط روی β، باید با شرطی کردن روی n انجام شود (کاکس و لوئیس، ۱۹۶۶)

مثال ۵.۱. (رگرسیون پواسون). معمولاً در رگرسیون لگاریتمی خطی پواسون همه مقادیر گسسته هستند. مشاهدات با پارامتر رگرسیون t، و توزیع پواسون با میانگین $e^{\alpha+\beta t}$، یکی از مثال‌های ساده در این مبحث است. نمونه مستقل $X_1,\ldots,X_n$ که $t_1 < t_2 < \ldots < t_n$ را در نظر بگیرید. در این صورت تابع احتمال $(X_1,\ldots,X_n)$ به صورت:

$$\exp\left\{e^{-n\alpha-\beta t_{.}}\right\}\frac{1}{x_1!\ldots x_n!}\exp\left\{\alpha x_{.}+\beta\sum_{j=1}^{n}t_j x_j\right\}$$

خواهد بود که $x_{.}$ دارای توزیع پواسون با پارامتر $e^{\alpha}\sum_{j=1}^{n}e^{\beta t_j}$ است. با این حال با شرطی کردن روی $x_{.}$ هر یک از متغیرهای $(X_1,\ldots,X_n)$ دارای توزیع چندجمله‌ای خواهند بود، به طوری‌که هر کدام از احتمال‌ها فقط وابسته به β هستند. بنابراین اگر α و β به‌طور مستقل، و همچنین α در $\mathbb{R}$ تغییر کنند، آن‌گاه $x_{.}$ بریده است. به‌طور کلی اگر m دنباله‌ای از مشاهدات به صورت $t_1<t_2<\ldots<t_n$ باشد و پارامتر α از دنباله‌ای به دنباله دیگر تغییر کند، x_{ij}، مشاهده iام از دنباله jام دارای توزیع پواسون با میانگین $\exp\left\{\alpha_i+\beta t_j\right\}$، و توزیع توأم این مشاهدات را که می‌توان توزیع حاشیه‌ای $x_{..}$ در نظر گرفت، پواسون با میانگین $\sum_{i=1}^{m}e^{\alpha_i}\sum_{j=1}^{n}e^{\beta t_j}$ و همچنین توزیع شرطی $x_{.*}$ به شرط $x_{..}$، توزیع چندجمله‌ای با پارامتر $\left(\sum_{i=1}^{m}e^{\alpha_i}\right)^{-1}\left(e^{\alpha_1},\ldots,e^{\alpha_m}\right)$ است. و در نهایت، توزیع شرطی x_{**} به شرط $x_{.*}$ که توزیع اخیر فقط به β بستگی دارد، بیان‌گر این نتیجه است که مشروط بر این که تغییرات α_* و β از هم مستقل، و دامنه تغییرات α_*، $\mathbb{R}^m$ باشد، هر دوی $x_{..}$و $x_{.*}$ بریده هستند.

در انتهای این بخش با استفاده از قسمت (الف) در قضیه ۱،۱، مثال‌های ساده‌ای ارائه می‌شود که UMVUE وجود دارد، اما آماره بسنده کامل موجود نیست.

مثال ۶.۱. فرض کنید $X_1,X_2,\ldots,X_n$ نمونه تصادفی از متغیر تصادفی $X=\mu+\sigma Y$ باشد، به طوری که Y دارای توزیع گاما، با پارامتر شکل p و پارامتر مقیاس معلوم است. در این صورت هر تابعی از $\bar{X}$ یک UMVUE است، به خصوص $\bar{X}$ یک UMVUE برای $E(X)$ نیز هست. برای مشاهده این موضوع، ابتدا توجه کنید که برای هر μ ثابت، $\bar{X}-\mu$ آماره بسنده کامل برای σ است. لذا برای هر μ ثابت، $\bar{X}$ نیز یک آماره بسنده برای σ است و لذا طبق قضیه ۱،۱، $\bar{X}$ یک UMVUE برای σ خواهد بود. همان‌طور که مشاهده می‌کنید برای $p\neq 1$ هیچ آماره بسنده کاملی برای (μ,σ) وجود ندارد، اما اگر در مواردی آماره بسنده به‌طور کران‌دار کامل وجود داشت، آن آماره، آماره بسنده مینیمال نیز هست.

اکنون $\mathcal{P}$ را به عنوان خانواده اندازه‌های احتمال P در نظر بگیرید. این نتیجه که میدان سیگمایی $\mathcal{P}$-بسنده و به‌طور کران‌دار کامل، لزوماً یک میدان سیگمایی $\mathcal{P}$-بسنده مینیمال است، یک نتیجه معروف است که به سختی می‌توان آن را در متون آماری متعارف یافت. به عنوان مثال، این موضوع در اشمترر (۱۹۷۴، ص۲۲۰) و لی‌من و شفه (لی‌من و شفه، ۱۹۵۰) با فرض اضافه‌تر که میدان سیگمایی آماره بسنده مینیمال موجود باشد بیان شده و همچنین زاکس (۱۹۷۳)، ص۷۳) این فرض را به‌طور ضمنی در نظر گرفته است. برای نشان دادن حقیقت فوق کافی است گزاره زیر را ثابت کنیم که به راحتی بر اساس قضیه لی‌من-شفه (لی‌من و شفه، ۱۹۵۰) اثبات می‌شود.

گزاره ۱.۱. اگر $\mathcal{A}$ میدان سیگمایی $\mathcal{P}$-بسنده و به‌طور کران‌دار کامل و $\mathcal{A}_1$ میدان سیگمایی $\mathcal{P}$-بسنده مینیمال باشد، آن‌گاه تقریباً همه جا $\mathcal{P}$، $\mathcal{A} = \mathcal{A}_1$.

اثبات (روش اول). به وضوح تقریباً همه جا $\mathcal{P}$، $\mathcal{A}_1 \subseteq \mathcal{A}$. کافی است ثابت کنیم تقریباً همه جا $\mathcal{P}$، $\mathcal{A} \subseteq \mathcal{A}_1$. برای این منظور، فرض کنید $A \in \mathcal{A}$. در این صورت $E\left(I_A | \mathcal{A}_1\right)$ وجود دارد و $\mathcal{A}_1$-اندازه‌پذیر است و در پس آن $\mathcal{A}$-اندازه‌پذیر نیز هست. بنابراین طبق خاصیت امید مکرر، $E\left[I_A - E\left(I_A | \mathcal{A}_1\right)\right] = 0$ و چون تقریباً همه جا $\mathcal{P}$ $\left|I_A - E\left(I_A | \mathcal{A}_1\right)\right| \le 2$، در این صورت طبق به‌طور کران‌دار کامل‌دار بودن میدان سیگمایی $\mathcal{A}$، تقریباً همه جا $\mathcal{P}$ $I_A = E\left(I_A | \mathcal{A}_1\right)$. برابری اخیر و همچنین این که $E\left(I_A | \mathcal{A}_1\right)$، $\mathcal{A}_1$-اندازه‌پذیر است، منجر به $\mathcal{A}_1$-اندازه‌پذیر بودن I_A نیز خواهد شد و در پی آن $A \in \mathcal{A}_1$ و از این رو تقریباً همه جا $\mathcal{P}$، $\mathcal{A} \subseteq \mathcal{A}_1$ که منجر به اثبات می‌شود.

(روش دوم). برای $A \in \mathcal{A}$ دلخواه، طبق قضیه لی‌من-شفه (لی‌من و شفه، ۱۹۵۰)، تابع نشانگر I_A یک UMVUE است. اما طبق قضیه رائو-بلکول (لی‌من و کسلا، ۱۹۹۸)، $E\left(I_A | \mathcal{A}_1\right)$ نیز یک براوردگر ناریب برای $E_P[I_A]$ با واریانس کمتر یا مساوی با I_A است. از آن‌جا که $E\left(I_A | \mathcal{A}_1\right)$ نیز یک UMVUE است و با توجه به یکتا بودن UMVUE، تقریباً همه‌جا $\mathcal{P}$، $I_A = E\left(I_A | \mathcal{A}_1\right)$ و در نتیجه تقریباً همه جا $\mathcal{P}$، $\mathcal{A} \subseteq \mathcal{A}_1$ و در پی آن تقریباً همه جا $\mathcal{P}$ $\mathcal{A} = \mathcal{A}_1$ و از این رو $\mathcal{A}$ یک میدان سیگمایی بسنده مینیمال است. ■

به راحتی می‌توان دید که آماره بسنده مینیمال، یک آماره بدیهی مثل آماره ترتیبی $(X_{1:1}, X_{1:2}, \ldots, X_{1:n})$ است، اما کامل نیست. با این حال اگر $p = 1$، آن‌گاه $(\bar{X}, X_{1:n})$ آماره بسنده کامل خواهد بود.

مثال ۷.۱. در مثال ۶.۱، توزیع Y به صورت $P(Y \le y) = y^p$، $0 \le y \le 1$ به طوری که $p > 0$ را در نظر بگیرید، با استدلال مشابه مثال ۶.۱، به جای $\bar{X}$، $X_{n:n}$ را به کار می‌بریم. به این ترتیب هر تابعی از $X_{n:n}$ یک UMVUE برای $\mu + \sigma[(pn)/(pn+1)]$ است. برای $p \ne 1$ آماره بسنده مینیمال برای (μ, σ) بدیهی است ولی کامل نیست. اگر $p = 1$ باشد، $(X_{1:n}, X_{n:n})$ آماره بسنده کامل برای (μ, σ) است.

مثال‌های دیگر نیز به سادگی با توجه به اصل زیر ساخته می‌شود:

اگر $P_{\theta,\phi}$ دارای تابع چگالی به صورت $\mathrm{f(x;(q,f)) = \exp\{A(q,f)\,g(T(x),f) + B(x,f) + C(q,f)\}}$ باشد، به طوری که T معلوم و برای هر ϕ ثابت، $g(.,\phi)$ یک تابع وارون‌پذیر باشد، آن‌گاه $T(X)$ یک آماره بسنده کامل برای θ به ازای هر ϕ ثابت است. بدیهی است که برای هر ϕ مشخص، مجموعه $\{A(\theta,\phi), \theta \in \Theta\}$ تهی نیست.

۲. روش‌های دیگر

در این بخش به کمک یک عمل دوتایی جابه‌جایی و شرکت‌پذیر، UMVUE برای پارامتر مجهول پیدا می‌شود.

عمل $*$ را به عنوان عمل دوتایی، جابه‌جایی و شرکت‌پذیر در گروه $\mathbb{R}\setminus\{0\}$ در نظر بگیرید. برای مثال اعمال $+$ ، $\times$ و همچنین اعمال دوتایی $x*y=(x+y)/(1+xy)$ که $|x|,|y|<1$ و $x*y=\max(x,y)$ جابه‌جایی و شرکت‌پذیر هستند.

تبدیل $T(x_1,x_2,\dots,x_n)=x_1*x_2*\dots*x_n$ را در نظر بگیرید. اگر $X_1,\dots,X_p$ یک نمونه تصادفی از جامعه ای با تابع توزیع تجمعی $F(x;\theta)$، $\theta\in\Theta$ باشد، تابع توزیع تجمعی برای $T(X_1,X_2,\dots,X_p)$ را با $F^{p*}(x;\theta)$ نمایش می‌دهیم.

قضیه ۱.۲. فرض کنید برای هر نمونه n تایی از جامعه‌ای با تابع توزیع تجمعی $F(x;\theta)$، آماره $S(X_1,X_2,\dots,X_n)$ وجود داشته باشد به طوری که (T,S) آماره بسنده کامل برای θ است. آن‌گاه برای یک نمونه دلخواه از جامعه با تابع توزیع تجمعی $F^{p*}(x;\theta)$، هر تابعی از $T(X_1,X_2,\dots,X_n)$ یک UMVUE است.

اثبات. می‌توان نوشت $X_j=X_{j1}*X_{j2}*\dots X_{jp}$، $j=1,2,\dots,p$ به طوری که برای هر j، $X_{j1},X_{j2},\dots,X_{jp}$ نمونه‌ای از $F(x;\theta)$ است. پس طبق قضیه لی‌من-شفه (لی‌من و شفه، ۱۹۵۰) برای نمونه فرضی $X_{11},X_{12},\dots,X_{np}$، براوردگر $T(X_1,\dots,X_n)=X_{11}*X_{12}*\dots*X_{np}$ یک UMVUE است. همچنین T بر اساس نمونه واقعی و با اطلاعات کمتر $X_1,X_2,\dots,X_n$ نیز یک UMVUE است.■

برای مشاهده کاربرد قضیه ۱،۲، دوباره مثال‌های ۶،۱ و ۷،۱ را در نظر بگیرید. برای $p=1$ آماره $(T,X_{1:n})$ برای (μ,σ) بسنده کامل است (در مثال ۶،۱، $T=\sum_{j=1}^{n}X_j$ و در مثال ۷،۱، $T=X_{n:n}$). قرار دهید $Y_j=\frac{X_j-\mu}{\sigma}$ و در پی آن برای مقادیر صحیح p به‌دست می‌آوریم:

$$\sum_{j=1}^{p}X_j=p\mu+\sigma\sum_{j=1}^{p}Y_j,$$

$$X_{p:p}=\mu+\sigma Y_{p:p}.$$

اگر Y_j دارای توزیع نمایی باشد، آن‌گاه $\sum_{j=1}^{p}Y_j$ دارای توزیع گاما با پارامتر شکل p خواهد بود و اگر Y_j دارای توزیع $U(0,1)$ باشد، آن‌گاه تابع توزیع تجمعی $Y_{p:p}$ به صورت y^p خواهد بود که $0\le y\le 1$.

از این رو نتایج UMVUE در مثال ۶،۱ و ۷،۱ ، همان نتایج قضیه ۱،۲ خواهند بود. ایده اصلی قضیه ۱،۲ را می‌توان با مثال ساده زیر نیز تحلیل کرد.

مثال ۱.۲. قرار دهید $X_1, X_2, ..., X_n$ مشاهدات مستقل از متغیر تصادفی $X = Z + Y$ که Z و Y مستقل‌اند. متغیر تصادفی Z دارای توزیع $N(\mu, 1)$ با پارامتر مجهول μ و Y یک مقدار صحیح با توزیع معلوم است. تابع g را یک تابع متناوب با دوره تناوب $\frac{1}{n}$ در نظر بگیرید، یعنی برای هر n داریم $g(x) = g(x + \frac{1}{n})$. در این صورت $X_i = Z_i + Y$ و از این رو $\overline{X} = \overline{Z} + \frac{1}{n}Y$ و در پی آن از این که g دارای دوره تناوب $\frac{1}{n}$ است، داریم:

$$g(\overline{X}) = g(\overline{Z} + \tfrac{1}{n}Y) = g(\overline{Z}).$$

با استقراء می‌توان نشان داد برای هر $m \in \mathbb{N}$، $g(x) = g(x + \frac{m}{n})$. زیرا اگر حکم برای m برقرار باشد، طبق فرض استقراء و همچنین متناوب بودن تابع g داریم:

$$\begin{aligned} g(x + \tfrac{m+1}{n}) &= g(x + \tfrac{m}{n} + \tfrac{1}{n}) \\ &= g(x + \tfrac{m}{n}) \\ &= g(x). \end{aligned}$$

در این صورت $g(\overline{X})$ یک UMVUE است که نتیجه ساده‌ای از $g(\overline{X}) = g(\overline{Z})$ است. در اینجا Z_jها با X_jها متناظرند و $\overline{Z}$ آماره بسنده کامل برای μ بر اساس نمونه ساختگی $(X_1, Z_1, X_2, Z_2, ..., X_n, Z_n)$ است.

۳. کامل بودن و براوردگر ناریب

در این بخش انگیزه استفاده از واژه‌های «*کامل بودن*» و «*ناریبی*» بیان می‌شود.

استیگلر (۱۹۷۲) مفهوم کامل بودن و ارتباط آن با براوردگر ناریب را به خوبی توضیح داده است. همان‌طور که لی‌من و شفه (لی‌من و شفه، ۱۹۵۰) در سال ۱۹۵۰ این موضوع را عنوان کردند، مفهوم کامل بودن خانواده اندازه‌های احتمال کاربرد بسیاری در مسائل براورد دارد و تبدیل به یک موضوع اصلی در مباحث آمار ریاضی شده است. دلیل کارساز بودن استفاده از خاصیت کامل بودن در مباحث آمار ریاضی این است که تعریف آن به راحتی قابل بیان است و در بسیاری از موارد، مفهوم آن به‌طور مستقیم به واسطه قضایای رائو-بلکول و لی‌من-شفه (لی‌من و کسلا، ۱۹۹۸)، در یافتن بهترین براوردگر ناریب کاربرد دارد. اگر چه وقتی دانشجو با مفهوم کامل بودن روبه‌رو می‌شود، بهتر می‌تواند با کاربردهای آن آشنا شود، اما به همان اندازه احتمال دارد از نام آن متعجب شود و تعجب کند که چه ارتباطی میان کاربرد آماری «*کامل بودن*» و تعریف آن وجود دارد.

بررسی تعدادی از متون که «*کامل بودن*» در آنها تعریف شده است، نشان می‌دهد که هیچ یک از نویسندگان درباره وجه تسمیه «*کامل بودن*» اظهار نظری نکرده و ممکن است از آن اجتناب کنند.

هدف این بخش، روشن‌سازی انگیزه استفاده از کلمه «*کامل بودن*» و علاوه بر آن، وجه برجسته‌ای از تعریف «*ناریبی*» با طرح یک مثال ساده است که اغلب در دوره‌های مقدماتی نادیده گرفته شده است. با یک تعریف ساده بحث را آغاز می‌کنیم.

تعریف ۱.۳. خانواده $\mathscr{P} = \{P_\theta : \theta \in \Theta\}$ را به عنوان خانواده توزیع‌های احتمال متغیر تصادفی (یا آماره) X در نظر بگیرید، که با مجموعه اندیس‌گذار Θ نشانه‌گذاری شده است. گوییم $\mathscr{P}$ کامل است، هرگاه برای هر تابع ϕ و $\theta \in \Theta$ که در آن

$$E_\theta\left[\phi(X)\right] = 0 \qquad (۳)$$

نتیجه بگیریم که برای هر $\theta \in \Theta$ تقریباً همه جا $\phi(x) = 0$، یعنی جاهایی که این تساوی رخ نمی‌دهد دارای احتمال صفر است (لی‌من و شفه، ۱۹۵۰). (E_θ نشان می‌دهد که مقدار مورد انتظار با توجه به توزیع متناظر با θ است).

به‌طور شهودی تعریف ۱,۳ نشان می‌دهد $\mathscr{P}$ کامل است، اگر هیچ براوردگر نااریبی برای صفر، به جز براوردگر بدیهی $\phi(x) \equiv 0$ وجود نداشته باشد. انگیزه شهودی انتخاب کلمه کامل بودن می‌تواند به صورت زیر بیان شود:

تابع ϕ باید شرایط (۳) را داشته باشد که نشان‌دهنده یک محدودیت به روی ϕ است. خانواده بزرگ‌تر $\mathscr{P}$ (یا Θ) محدودیت بیشتری روی ϕ ایجاد می‌کند. وقتی که $\mathscr{P}$ آنقدر بزرگ شود که تحت شرایط (۳) همه ϕها را به جز مورد بدیهی $\phi(x) \equiv 0$ رد کند، گوییم $\mathscr{P}$ کامل است.

به عنوان مثال، خانواده توزیع $\mathscr{P} = \left\{P_N : N \geq 1\right\}$ را در نظر بگیرید به طوری‌که:

$$P_N(X = k) = \begin{cases} \frac{1}{N} & k = 1, 2, \dots, N \\ 0 & o.w. \end{cases}$$

می‌توان نشان داد برای این که برای هر $N \geq 1$، $E_N\left[\phi(X)\right]$ برابر با صفر باشد، باید برای $k = 1, 2, \dots, N$، $\phi(k) = 0$. برای دیدن جزئیات فرض کنید:

$$E_N(\phi(X)) = \sum_{k=1}^{N} \phi(k) P(X = k) = \sum_{k=1}^{N} \phi(k) \tfrac{1}{N} = 0.$$

بنابراین برای هر $N \geq 1$، $\sum_{k=1}^{N} \phi(k) = 0$. با قرار دادن $N = 1$ داریم $\phi(1) = 0$ و به همین صورت با استقراء قوی نشان می‌دهیم که برای $k = 1, 2, \dots, N$، $\phi(k) = 0$ و به عبارت دیگر $\mathscr{P}$ کامل است. با این حال $\mathscr{P}$ به سختی کامل است، یعنی اگر یکی از بی‌نهایت توزیع احتمال را حذف کنیم، آن‌گاه $\mathscr{P}$ کامل نخواهد بود. برای اثبات این موضوع، تابع ϕ_0 با ضابطه

$$\phi_0(k) = \begin{cases} 0 & k = 1, 2, \dots, n-1, n+2, n+3, \dots \\ a & k = n \\ -a & k = n+1 \end{cases} \qquad (۴)$$

را در نظر بگیرید به‌طوری که a هر مقدار ثابت ناصفر است. به راحتی می‌توان مشاهده کرد که

$$E_N(\phi_0(X)) = \sum_{k=1}^{\infty}\phi_0(k)P(X=k)$$
$$= \begin{cases} aP(X=n) - aP(X=n+1) + 0 = 0 & N \neq n \\ aP(X=n) - 0 + 0 = \dfrac{a}{N} & N = n. \end{cases}$$

بنابراین ϕ_0 یک براوردگر ناریب صفر، برای خانواده $\mathcal{P}-\{P_n\}$ است (تنها براوردگر غیربدیهی صفر که با توجه به (۴) برای هر a متفاوت است).

با یک مثال ساده نشان دادیم کامل بودن یک ویژگی خانواده توزیع‌ها است تا متغیرتصادفی یا صورت پارامتری، به طوری که تعریف آماری کامل بودن با کاربرد روزمره آن مرتبط است و حذف حتی یک نقطه از فضای پارامتر ممکن است کامل بودن را تغییر دهد. به عنوان اطلاعات ضمنی در مورد این مثال، قضیه لی‌من-شفه (لی‌من و شفه، ۱۹۵۰) بیان می‌کند از آن‌جا که $\mathcal{P}$ کامل است، براوردگر ناریب $\phi_1(X) = 2X - 1$ براوردگر ناریب با کمترین واریانس N برای خانواده $\mathcal{P}$ است. اما از آن‌جا که $\mathcal{P}-\{P_n\}$ کامل نیست، این قضیه در مورد این خانواده کاربردی ندارد و در حقیقت ϕ_1 براوردگر ناریب با کمترین واریانس N برای خانواده $\mathcal{P}-\{P_n\}$ نیست. براوردگر ناریب با کمترین واریانس برای خانواده $\mathcal{P}-\{P_n\}$ به صورت:

$$\phi_2(k) = \begin{cases} 2k-1 & k \neq n, k \neq n+1 \\ 2n & k = n, n+1 \end{cases}$$

است. این حقیقت را می‌توان مستقیماً اثبات کرد و یا از تعمیم قضیه لی‌من-شفه (لی‌من و کسلا، ۱۹۹۸) استفاده کرد که بیان می‌کند شرط لازم و کافی برای این که ϕ_2 دارای کمترین واریانس باشد این است که با همه براوردگرهای ناریب صفر ناهمبسته باشد.

یک محاسبه ساده نشان می‌دهد که

$$Var_N\left[\phi_2(X)\right] = \begin{cases} Var_N\left[\phi_1(X)\right] & N < n \\ Var_N\left[\phi_1(X)\right] - \frac{2}{N} & N > n \end{cases}$$

در حالی که ϕ_2 در بین همه براوردگرهای ناریب N برای خانواده $\mathcal{P}-\{P_n\}$ دارای کمترین واریانس است، ولی برای خانواده $\mathcal{P}$ ناریب نیست، زیرا

$$E_n\left[\phi_2(X)\right] = \frac{n+1}{n}.$$

این مسئله، نشان‌دهنده این واقعیت است که ناریبی، همانند کامل بودن ویژگی آماره یا صورت پارامتری آن نیست، بلکه ویژگی خانواده توزیع‌های یک آماره است.

استیگلر (۱۹۷۲)، خانواده $P_0 = \{P_N : N \geq 1, N \neq N_0\}$ را در نظر گرفت به‌طوری که

$$P_N(X = x) = \tfrac{1}{N} \qquad x = 1, 2, , , N$$

و نشان داد که:

$$T(X) = \begin{cases} 2X - 1 & X \neq N_0, X \neq N_0 + 1 \\ 2N_0 & X = N_0, N_0 + 1 \end{cases}$$

یک UMVUE برای N است، در حالی که X برای P_0 یا N آماره بسنده کامل نیست.

۴. بحث و نتیجه‌گیری

در حالتی که آماره بسنده کامل وجود داشته باشد، همه توابع پارامتری براوردپذیر، UMVUE هستند. بر عکس، اگر هر تابع پارامتری براوردپذیر یک UMVUE باشد، آن‌گاه آماره بسنده کامل تحت شرایط خاص وجود خواهد داشت. در برخی حالات با این که آماره بسنده کامل وجود ندارد، نیز ممکن است توابع پارامتری غیرثابتی وجود داشته باشند، که UMVUE باشند. در این مقاله، تعمیم‌ها و مثال‌های ساده‌ای از قضیه لی‌من-شفه (لی‌من و شفه، ۱۹۵۰) مطرح شد که UMVUE وجود دارد، اما آماره بسنده کامل موجود نیست. در حالتی که آماره بسنده کامل وجود نداشته باشد و آماره مناسبی که از آماره کمکی مستقل باشد به راحتی یافت نشود، شرطی کردن آماره بسنده روی آماره کمکی باعث بازیابی اطلاعات از دست رفته خواهد شد. به کمک یک عمل دوتایی گروه که جابه‌جایی و شرکت‌پذیر هست، نیز می‌توان UMVUE برای پارامتر مجهول پیدا کرد. قابل ذکر است، کامل بودن و ناریبی ویژگی آماره یا صورت پارامتری آن نیست، بلکه ویژگی خانواده توزیع‌های یک آماره است و حذف حتی یک نقطه از فضای پارامتر می‌تواند خصوصیت کامل بودن را از دست بدهد.

References

1. R. R. Bahadur, On unbiased estimates of uniformly minimum variance, Sankhya **18** (1957), 211-224.
2. O. Barndorff-Nielsen, Information and Exponential Families in Statistical Theory, Chichester, 1978.
3. D. Basu, On statistics independent of a complete sufficient statistic. Sankhya, **15** (1955), 377-380.
4. D. Basu, On statistics independent of a sufficient statistic. Sankhya, **20** (1958), 223-226.
5. V. V. Chichagov, Asymptotic Comparison of MLE and UMVUE in the Case of a Truncated One-Parameter Family of Distributions, Journal of Mathematical Sciences, **267** (2022), 170-179.
6. D. R. Cox and P. A. W. Lewis, The statistical analysis of series of events, Methuen, London. 1966.

7. D. A. S. Fraser, Sufficient statistics with nuisance parameters, Ann. Math. Stat. **27** (1956), 838-842.
8. F. Gasperoni, F. Ieva, A. M. Paganoni, F. Gasperoni, F. Ieva, A. M. Paganoni, Uniform Minimum Variance Unbiased Estimators (UMVUE), Eserciziario di Statistica Inferenziale, Springer Milan, (2022), 45-61.
9. M. Y. Hirose and S. Mano, Asymptotic UMVUE: Asymptotic Moments Matching the UMVUE under the Ewens Sampling Formula, Calcutta Statistical Association Bulletin (2023) (Accepted).
10. K. Thukral and S. K. Sharma, Estimation of Software Reliability Using Lindley Distribution Based on MLE and UMVUE, International workshop of Mathematical Modelling, Applied Analysis and Computation, Springer Nature Switzerland, (2022), 299-316.
11. J. Kubrusly, Uniformly Minimum-Variance Unbiased Estimator (UMVUE) for the Gamma Cumulative Distribution Function with Known and Integer Scale Parameter, Open Journal of Statistics, **12** (2022), 168-174.
12. E. L. Lehmann, An Interpretation of Completeness and Basu’s Theorem, J. Am. Stat. Assoc. **76** (1981), 335-340.
13. E. L. Lehmann and G. Casella, Theory of Point Estimation, 2nd ed., Springer-Verlag, New York, 1998.
14. E. L. Lehmann and J. P. Romano, Testing Statistical Hypotheses, Springer, New York, 3rd ed., 2005.
15. E. L. Lehmann and H. Scheffe, Completeness, similar regions and unbiased estimation, Part I. Sankhya **10** (1950), 305 -340.
16. Z. L. Maureal, E. C. Castillano and R. N. Padua, Uniform Minimum Variance Unbiased Estimator of Fractal Dimension, Recoletos Multidisciplinary Research Journal, **9**(1) (2021), 63-68.
17. C. R., Rao, Linear Statistical Inference and Its Applications, New York, 2nd ed., 1973.
18. V. K. Rohatgi, An Introduction to Probablity Theory and Mathematical Statistics, New York, 1976.
19. L. Schmetterer, Introduction to Mathematical Statistics, Berlin - Hiedelberg – New York, 1974.
20. V. Srinivas and B. K. Kale, ML and UMVU estimation in the M/D/1 queuing system, Commun. Statist. Theor. Meth., **45**(19) (2016), 5826-5834.
21. N. Stallard and P. K. Kimani, Uniformly minimum variance conditionally unbiased estimation in multi-arm multi-stage clinical trials, Biometrika, **105**(2) (2018), 495-501.
22. S. M. Stigler, Completeness and unbiased estimation, Amer. Stat. **26** (1972), 28-29.
23. J. Verbeeck, V. Deltuvaite-Thomas, B. Berckmoes, T. Burzykowski, M. Aerts, O. Thas, M. Buyse and G. Molenberghs, Unbiasedness and efficiency of non-parametric and UMVUE estimators of the probabilistic index and related statistics, Statistical Methods in Medical Research, **30**(3) (2021), 747-768.
24. S. Zacks, The Theory of Statistical Inference, New York, 1973.