Kharazmi University

Mathematical Researches

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

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

The Role of Statistical Independence in Finding Optimal Estimators

Mehdi Shams[1]; Leila Nasiri[2] ✉

1. Department of Statistics, Faculty of Mathematical Sciences, University of Kashan, Kashan, Iran.

E-mail: mehdishams@kashanu.ac.ir

2. Department of Mathematics and Computer Science, Faculty of Science, Lorestan University, Khorramabad, Iran.

✉E-mail: nasiri.le@fs.lu.ac.ir

Article Info

Article type:

Research Article

Article history:

Received:
22 February 2020
Revised form:
18 JulyMay 2020
Accepted:
15 September 2020
Published online:
14 May 2022

Keywords:

Estimation;
Risk function;
Independence;
Invariant statistic;
Equivariant estimator.

ABSTRACT

Introduction

Naturally in choosing a point estimator we are interested in choosing an estimator that minimizes the risk function for all parameter space values. In practice, this is not possible due to the large number of estimators. One way to solve this problem (find optimal estimators) is to limit the range of estimators and find the best estimator in the finite range. This limitation leads us to two types of optimal estimators, namely the best equivariant estimator and the minimum risk unbiased estimator, respectively, in terms of limiting ourselves to the class of equivariant or unbiased estimators. In this paper, the role of independence in simplifying the calculation of these estimators is examined. We also deal with the stochastic independence of an invariant function and its comparison with the Basu's theorem.

To find the optimal estimators, the class of estimators can be limited. This limitation can be applied to the class of equivariant or unbiased estimators, which leads to two types of optimal estimators, namely the best equivariant estimator and the minimum risk unbiased estimator, respectively. For this purpose, in addition to the Rao-Blackwell-Lehmann-Scheffé theorem (*Casella and Berger*, 2001), two other methods have been proposed by *Sathe and Varde (1969)* and *Eaton and Morris (1970)* which can be useful to achieve this goal. In these two methods, by limiting the class of equivariant or unbiased estimators, the estimator with the minimum risk is considered as the optimal estimator. Independence can play a key role in making it easier to calculate the risk function of the best equivariant estimator and the minimum risk unbiased estimator. In fact, the role of independence is to eliminate the conditional probability in calculating the risk function of the best equivariant estimator and the minimum risk unbiased estimator, which in most cases results from Basu's theorem and the independence of ancillary statistic from complete sufficient statistics. Similar to Basu's theorem, it can be shown that in certain circumstances an invariant statistic and equivariant function are independent of each other, which can play a role in eliminating the conditional probability by the independence of the maximum invariant statistic from the equivariant sufficient statistic. It is noteworthy that in this case, the completeness assumption of

Basu's theorem has been replaced by equivariance assumption and the ancillarity assumption of Basu's theorem has been replaced by invariance.

Material and methods

We first introduce the definitions that are needed. In the second part, by limiting the class of equivariant estimators, we create a type of optimal estimator called the best equivariant estimator and show that in groups that act as transitive on the parameter space, an invariant function is independent of the equivariant sufficient statistic. In the third part, by limiting the class of unbiased estimators, we make a type of optimal estimator called minimum risk unbiased estimator, which in a special case where the square error loss function is the same as the minimum variance unbiased estimator, which in the fourth part, in addition to Rao-Blackwell-Lehmann- Scheffé theorem (*Casella and Berger*, 2001), introduce two other methods proposed *by Sathe and Varde (1969)* and *Eaton and Morris (1970)* which, with the help of independence, provide a simpler method for finding a minimum variance unbiased estimator.

Conclusion

The following conclusions were drawn from this research.

- In order to find the optimal estimators by limiting the class of estimators to the class of equivariant or unbiased estimators, the independence of complete sufficient statistics from ancillary statistics and applying Basu's theorem can be a way to simplify calculations.
- In some statistical problems with the transitive transformation group, the equivariant function can be used instead of the complete sufficient statistic. In this case, instead of using the Basu's theorem, the independence of an invariant function and the equivariant sufficient statistic can be inferred. Hence, the assumption of completeness for establishing the Basu's theorem is replaced by the equivariance and having a transitive transformation group.
- Having a transitive group, any invariant function is also ancillary, but the incompleteness of a sufficient statistic can also result in its independence from an invariant statistic. If the group of transitive transformations and complete sufficient statistic is also equivariant, the case of Basu's theorem concludes this view. The opposite is not always true and can be corrected in such a way that if a complete statistic is also equivariant with a transitive group, it is not necessary that each ancillary statistic be independent of it. Rather, it is possible to find an ancillary statistic that is also invariant, which is independent of the given equivariant complete sufficient statistic. By finding the condition that an ancillary statistic is also invariant, these results can be extended, in which case Basu's theorem is the result. Of course, this open problem needs further research and it is hoped that researchers will be diligent in generalizing it.

How to cite: Shams, M., Nasiri, L; (2022). The Role of Statistical Independence in Finding Optimal Estimator. *Mathematical Researches*, 8 (1), 127-147

 Publisher: Kharazmi University

شاپا چاپی: ۲۵۴۶-۲۵۸۸
شاپا الکترونیکی: ۲۵۵٤-۲۵۸۸

پژوهش‌های ریاضی

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

نقش استقلال آماری در یافتن برآوردگرهای بهینه

مهدی شمس[۱]، لیلا نصیری[۲] ✉

۱. گروه آمار، دانشگاه کاشان، کاشان، ایران. پست الکترونیکی: mehdishams@kashanu.ac.ir
۲. نویسندهٔ مسئول، گروه ریاضی، دانشگاه لرستان، خرم‌آباد، ایران. پست الکترونیکی: nasiri.le@fs.lu.ac.ir

اطلاعات مقاله

نوع مقاله: مقاله پژوهشی

تاریخ دریافت: ۱۳۹۸/۱۲/۰۳
تاریخ بازنگری: ۱۳۹۹/۰۶/۰۵
تاریخ پذیرش: ۱۳۹۹/۰۶/۲۵
تاریخ انتشار: ۱۴۰۱/۰۲/۲۴

واژه‌های کلیدی:
برآورد،
تابع مخاطره،
استقلال،
آمارهٔ ناوردا،
برآوردگر هم‌وردا.،

چکیده

به طور طبیعی در انتخاب یک برآوردگر نقطه‌ای علاقه‌مند به انتخاب برآوردگری هستیم که برای تمام مقادیر فضای پارامتر، تابع مخاطره را مینیمم کند. در عمل با توجه به بزرگی ردهٔ برآوردگرها چنین امکانی وجود ندارد. یکی از راه‌ها برای حل این مشکل (پیدا کردن برآوردگرهای بهینه)، محدود کردن ردهٔ برآوردگرها و پیدا کردن بهترین برآوردگر در ردهٔ محدود شده است. این محدودیت بر حسب این که خود را به ردهٔ برآوردگرهای هم‌وردا یا نااریب محدود کنیم به ترتیب منجر به دو نوع برآوردگر بهینه یعنی بهترین برآوردگر هم‌وردا و برآوردگر نااریب با کمترین مخاطره می‌شود که در این مقاله نقش استقلال در ساده‌تر محاسبه کردن این برآوردگرها مورد بررسی قرار می‌گیرد. همچنین به استقلال تصادفی یک تابع ناوردا و هم‌وردا و مقایسه آن با قضیهٔ باسو می‌پردازیم.

استناد: شمس، مهدی؛ نصیری، لیلا؛ (۱۴۰۱). نقش استقلال آماری در یافتن برآوردگرهای بهینه. *پژوهش‌های ریاضی*، ۸ (۱)، ۱۴۷-۱۲۷.

ناشر: دانشگاه خوارزمی

۱. مقدمه

برای پیدا کردن برآوردگرهای بهینه می‌توان ردهٔ برآوردگرها را محدود کرد. این محدودیت می‌تواند روی ردهٔ برآوردگرهای هم‌وردا یا نااریب انجام شود که به ترتیب منجر به دو نوع برآوردگر بهینه یعنی بهترین برآوردگر هم‌وردا و برآوردگر نااریب با کمترین مخاطره می‌شود. برای این منظور به جز قضیهٔ رائو-بلاکول-لی‌من-شفه (کسلا و برگر، ۲۰۰۱) دو روش دیگر توسط ست و وارد (۱۹۶۹) و ایتون و موریس (۱۹۷۰) مطرح شده است که می‌تواند برای رسیدن به این هدف مفید باشد. در این دو روش با محدود کردن روی ردهٔ برآوردگرهای هم‌وردا یا نااریب، برآوردگری که دارای کمترین مخاطره است به عنوان برآوردگر بهینه در نظر گرفته می‌شود. استقلال می‌تواند در ساده‌تر محاسبه کردن تابع مخاطرهٔ بهترین برآوردگر هم‌وردا و برآوردگر نااریب با کمترین مخاطره نقش اساسی داشته باشد. در حقیقت نقش استقلال حذف احتمال شرطی در محاسبه تابع مخاطرهٔ بهترین برآوردگر هم‌وردا و برآوردگر نااریب با کمترین مخاطره است، که این احتمال شرطی در اکثر موارد به کمک قضیهٔ باسو و استقلال آمارهٔ کمکی از آمارهٔ بسندهٔ کامل نتیجه می‌شود. همچنین مشابه با قضیهٔ باسو می‌توان نشان داد در شرایط خاص یک تابع ناوردا و هم‌وردا از هم مستقل هستند که این نتیجه می‌تواند در حذف احتمال شرطی به کمک استقلال آمارهٔ ناوردای ماکسیمال از برآوردگر هم‌وردای بسنده ایفای نقش کند. قابل ذکر است که در این حالت فرض کامل بودن برای برقراری قضیهٔ باسو با هم‌وردایی و فرض کمکی بودن با ناوردایی جابه‌جا شده است. برای قابل فهم بودن بهتر مقاله در ابتدا تعاریفی که مورد نیاز است را مطرح می‌کنیم. در بخش دوم با محدودکردن روی ردهٔ برآوردگرهای هم‌وردا یک نوع برآوردگر بهینه به نام بهترین برآوردگر هم‌وردا می‌سازیم و نشان می‌دهیم در گروه‌هایی که روی فضای پارامتر به صورت انتقالی عمل می‌کنند، یک تابع ناوردا از یک برآوردگر هم‌وردای بسنده مستقل است. در بخش سوم با محدودکردن روی ردهٔ برآوردگرهای نااریب یک نوع برآوردگر بهینه به نام برآوردگر نااریب با کمترین مخاطره می‌سازیم که در حالت خاص که تابع زیان مربع خطا باشد همان برآوردگر نااریب با کمترین واریانس است که در بخش چهارم علاوه بر قضیهٔ رائو-بلاکول-لی‌من-شفه (کسلا و برگر،۲۰۰۱) دو روش دیگر که توسط ست و وارد (۱۹۶۹) و ایتون و موریس (۱۹۷۰) مطرح شده‌اند را معرفی می‌کنیم که به کمک استقلال روش ساده‌تری برای یافتن برآوردگر نااریب با کمترین واریانس ارائه می‌دهند.

ابتدا چند مفهوم مقدماتی را تعریف می‌کنیم.

تعریف ۱-۱: مجموعه دلخواه S را در نظر می‌گیریم، هر تابع از $S\times S$ به S را یک عمل دوتایی روی مجموعهٔ S می‌نامیم (رومن، ۲۰۱۲).

تعریف ۱-۲: مجموعهٔ ناتهی $\mathcal{G}$ را همراه با عمل دوتایی $*$ یک گروه گویند، هرگاه:

الف) $\mathcal{G}$ نسبت به عمل $*$ بسته باشد، یعنی برای هر دو عضو $a,b\in\mathcal{G}$، $a*b\in\mathcal{G}$.

ب) $\mathcal{G}$ نسبت به عمل $*$ شرکت‌پذیر باشد، یعنی برای هر $a,b,c\in\mathcal{G}$،

$$a*(b*c)=(a*b)*c.$$

ج) عنصر $e\in\mathcal{G}$ (عضو همانی) وجود داشته باشد به طوری که به ازای هر $a\in\mathcal{G}$،

$$a * e = e * a = a.$$

د) به ازای هر $a \in \mathcal{G}$ یک $a^{-1} \in \mathcal{G}$ (عضو وارون) موجود باشد به قسمی که

$a * a^{-1} = a^{-1} * a = e$ (رومن، ۲۰۱۲).

گروه‌هایی که بیشتر در علم آمار مورد توجه هستند گروه تبدیل‌ها هستند.

تعریف ۱-۳: اگر $\mathcal{X}$ یک مجموعهٔ ناتهی و $\mathcal{G}$ مجموعهٔ تمام توابع یک به یک از $\mathcal{X}$ به $\mathcal{X}$ باشد و عمل گروه را برای هر $g_1, g_2 \in \mathcal{G}$ به صورت $(g_1 \circ g_2)(x) = g_1(g_2(x))$ تعریف کنیم، در این صورت گروه $\mathcal{G}$ را گروه تبدیل‌ها روی $\mathcal{X}$ می‌گویند، (لی‌من و کسلا، ۱۹۹۸).

یک گروه یک رابطه هم‌ارزی میان اعضای فضای نمونه می‌سازد که رده‌های هم‌ارزی حاصل را مدار می‌نامیم. مدار متناظر با نقطه x را با $O_x = \{gx : g \in \mathcal{G}\}$ نمایش می‌دهیم. در آمار با گروه‌های توپولوژیکی سر و کار داریم. گروه $\mathcal{G}$ را یک گروه توپولوژیکی گویند، هرگاه تابع $l : \mathcal{G} \times \mathcal{G} \to \mathcal{G}$ با ضابطه $l(g_1, g_2) = g_1^{-1} g_2$ پیوسته باشد (فولند، ۱۹۹۵).

اگر $\sigma(\mathcal{X})$ و $\sigma(\mathcal{Y})$ به ترتیب دو σ-جبر روی فضای $\mathcal{X}$ و $\mathcal{Y}$ باشند، تابع $f : \mathcal{X} \to \mathcal{Y}$ را اندازه‌پذیر گویند، هرگاه برای هر $A \in \sigma(\mathcal{Y})$، $f^{-1}(A) \in \sigma(\mathcal{X})$ (فولند، ۱۹۹۵). در این مقاله در مورد دو تابع اندازه‌پذیر ناوردا و هم‌وردا بحث خواهد شد که به صورت زیر تعریف می‌شوند.

تعریف ۱-۴: اگر $\mathcal{G}$ گروه تبدیل‌ها روی $\mathcal{X}$ باشد، آمارهٔ T را تحت گروه $\mathcal{G}$، ناوردا گویند، هرگاه برای هر $x \in \mathcal{X}$ و $g \in \mathcal{G}$، $T(x) = T(gx)$ و آمارهٔ S را هم‌وردا گویند، هرگاه برای هر $x \in \mathcal{X}$ و $g \in \mathcal{G}$، $S(gx) = gS(x)$ (زکز، ۱۹۷۱).

تعریف ۱-۵: آمارهٔ ناوردای T را ناوردای بیشین گویند، اگر برای هر $x, y \in \mathcal{X}$ که $T(x) = T(y)$، $g \in \mathcal{G}$ وجود داشته باشد به قسمی که $y = g(x)$ (زکز، ۱۹۷۱).

تابع ناوردا روی مدارها ثابت است و تابع ناوردای بیشین علاوه بر آن مقادیر مختلف را به مدارهای مجزا اختصاص می‌دهد و لذا بهترین حالتی است که فضای نمونه را به مدارهایی افراز می‌شود که تحت $\mathcal{G}$ معادل هستند.

تعریف ۱-۶: گروه $\mathcal{G}$ روی فضای $\mathcal{X}$ انتقالی است، هرگاه برای هر $x, y \in \mathcal{X}$ یک $g \in \mathcal{G}$ وجود داشته باشد به قسمی که $y = g(x)$ (لی‌من و کسلا، ۱۹۹۸).

آماره‌ای که برای استنباط به کار برده می‌شود نباید باعث تلخیص نمونهٔ اولیه به قسمی شود که دیگر برای بررسی مسئله کفایت نکند. برای این منظور از آمارهٔ بسنده استفاده می‌کنیم که به صورت زیر تعریف می‌شود.

تعریف ۱-۷: آمارهٔ T را برای θ بسنده گویند، هرگاه به ازای هر θ توزیع شرطی $\mathbf{X}|T = t$، برای همه مقادیر t به پارامتر θ بستگی نداشته باشد (لی‌من و کسلا، ۱۹۹۸).

آمارهٔ U را کمکی (فرعی یا غیرآگاهی بخش) برای یک پارامتر گویند، هرگاه توزیع U به آن پارامتر مجهول بستگی نداشته باشد (لی‌من و کسلا، ۱۹۹۸). آماره‌های کمکی به خودی خود هیچ اطلاعی راجع به پارامتر ارائه نمی‌دهند، ولی در

ترکیب با بقیه داده‌ها می‌توانند اطلاعات کاملی راجع به پارامتر داشته باشند و همچنین با شرطی کردن روی آماره‌های کمکی این تضمین وجود دارد که هیچ اطلاعی راجع به پارامتر با این تقلیل از دست نمی‌رود.

تعریف ۱-۸: آمارهٔ T را کامل گویند، هرگاه برای هر تابع g، اگر به ازای هر $\theta \in \Theta$، $E_\theta\left(g(T)\right)=0$، آن‌گاه به ازای هر $\theta \in \Theta$، $P_\theta\left[g(T)=0\right]=1$ (لی‌من و کسلا، ۱۹۹۸).

خاصیت کامل بودن یک آماره به تنهایی کاربردی ندارد ولی وقتی این خاصیت مربوط به آمارهٔ بسنده باشد، موجب سادگی برخی از مسائل آمار و احتمال می‌شود (پارسیان، ۱۳۹۵).

متغیرهای تصادفی X و Y را مستقل گویند، هرگاه برای هر دو مجموعهٔ A و B از اعداد حقیقی داشته باشیم:

$$P\left(X \in A, Y \in B\right)=P\left(X \in A\right)P\left(Y \in B\right)$$

و آن را با نماد $X \perp Y$ نمایش می‌دهیم. در بسیاری از مسائل آماری نیاز به وجود اثبات استقلال دو آماره داریم. با استفاده از قضیهٔ باسو بدون این که توزیع توأم دو آماره محاسبه شوند، وجود این استقلال ثابت می‌شود.

قضیه ۱-۱ *(قضیهٔ باسو):* اگر آمارهٔ T برای خانوادهٔ توزیع‌های $\mathcal{P}$ بسنده و کامل باشد و آمارهٔ U یک آمارهٔ کمکی باشد، آن‌گاه T و U به ازای هر $\theta \in \Theta$ مستقل هستند (باسو، ۱۹۵۵).

می‌دانیم که سه تایی $(\Theta, \mathcal{A}, L)$ که در آن Θ فضای پارامتر، $\mathcal{A}$ فضای عمل‌ها و L نمایان‌گر تابع زیان است اساس کار نظریه تصمیم را تشکیل می‌دهند. تابع زیان به صورت زیر تعریف می‌شود.

تعریف ۱-۹: تابع $L:\Theta\times\mathcal{A}\rightarrow\mathbb{R}$ را یک تابع زیان گویند، هرگاه برای هر $a\in\mathcal{A}$ و $\theta\in\Theta$، $L\left(\theta,a\right)$ برابر مقدار جریمه ناشی از انتخاب عمل a باشد وقتی که θ پارامتر حقیقی است.

تابع‌های زیان معروف عبارتند از:

الف) تابع زیان مربع خطا (SEL[۱]):

$$L\left(\theta,a\right)=\left(a-\theta\right)^2.$$

ب) تابع زیان قدر مطلق خطا (AEL[۲]):

$$L\left(\theta,a\right)=\left|a-\theta\right|.$$

ج) تابع زیان نمایی خطی (LINEX[۳]):

$$L\left(\theta,a\right)=e^{c(a-\theta)}-c\left(a-\theta\right)-1\,;c\neq 0.$$

حال اگر وقتی θ پارامتر حقیقی باشد، براوردگر $\delta:\mathcal{X}\rightarrow\mathcal{A}$ را در نظر بگیریم و $X=x$ مشاهده شود، مقدار زیان ناشی از انتخاب برآورد $\delta(x)$ برابر $L\left(\theta,\delta(x)\right)$ خواهد بود. به قاعده δ یک تابع تصمیم گویند و ردهٔ تمام تصمیم‌ها را با $\mathcal{D}$ نمایش می‌دهند. بنا براین در سه تایی $(\Theta,\mathcal{D},R)$، $\mathcal{D}$ فضای تصمیم و R تابع مخاطره است که به صورت زیر تعریف می‌شود.

۱ Squared Error Loss

۲ Absolute error loss

۳ Linear-Exponential

تعریف ۱-۱۰: تابع $R:\Theta\times\mathcal{D}\to\mathbb{R}$ که به صورت زیر تعریف می‌شود را یک تابع مخاطره گویند:

$$R(\theta,\delta)=E_\theta\left(L\left(\theta,\delta(X)\right)\right).$$

این تابع میزان دقت یا عدم دقت برآوردگر را نشان می‌دهد. به این معنی که در این جا بهترین برآوردگر δ آن است که برای هر $\theta\in\Theta$، $R\left(\theta,\delta\right)$ مینیمم شود.

تعریف ۱-۱۱: اگر X یک متغیر تصادفی روی فضای اندازه‌پذیر $\mathcal{X}$ با توزیعی از خانوادهٔ $\mathcal{P}=\left\{P_\theta:\theta\in\Theta\right\}$ باشد، $\delta(X)$ برای $g(\theta)$ نااریب (کلاسیک) است، اگر برای هر $\theta\in\Theta$، $E_\theta\left(\delta(X)\right)=g(\theta)$ (لی‌من و کسلا، ۱۹۹۸).

تعریف عمومی زیر از نااریبی توسط لی‌من (۱۹۵۱) بیان شده است.

تعریف ۱-۱۲: اگر X یک متغیر تصادفی با توزیعی از خانوادهٔ $\mathcal{P}$ باشد، آمارهٔ $\delta(X)$ برای $g(\theta)$، L-نااریب است، اگر برای هر $\theta,\theta'\in\Theta$ که $\theta\neq\theta'$:

$$E_\theta\left(L\left(\theta,\delta(X)\right)\right)\le E_\theta\left(L\left(\theta',\delta(X)\right)\right).$$

در حالت خاص که تابع زیان SEL باشد نااریبی کلاسیک نتیجه می‌شود و نیز اگر تابع زیان AEL باشد، $g(\theta)$ میانه توزیع متغیر تصادفی $\delta(X)$ است. در حالتی که تابع زیان LINEX باشد، $\delta(X)$ یک برآوردگر LINEX-نااریب $g(\theta)$ است، هرگاه $\delta(X)$ برای هر $\theta\in\Theta$ در رابطهٔ $E_\theta\left(e^{a\delta(X)}\right)=e^{ag(\theta)}$ صدق کند. حال اگر یک برآوردگر نااریب مثل $h(X)$ برای $e^{ag(\theta)}$ پیدا کنیم، یعنی $E_\theta\left(h\left(X\right)\right)=e^{ag(\theta)}$ به شرط آن که با احتمال یک آمارهٔ $h(X)$ یک آمارهٔ مثبت باشد، برآوردگر $\delta(X)=\frac{1}{a}\ln\left(h(X)\right)$ یک برآوردگر LINEX-نااریب برای $g(\theta)$ خواهد بود و لذا برای $g(\theta)=\theta$ داریم:

$$\begin{aligned}R\left(\theta,\delta\right)&=e^{-a\theta}E_\theta\left(e^{a\delta(X)}\right)-aE_\theta\left(\delta(X)-\theta\right)-1\\&=e^{-a\theta}e^{a\theta}-aE_\theta\left(\delta(X)-\theta\right)-1\\&=-aE_\theta\left(\delta(X)-\theta\right).\end{aligned}$$

پس از تعریف مفاهیم نظریه تصمیم به دنبال برآوردگرهای بهینه هستیم که یکی از روش‌ها برای رسیدن به این منظور محدود کردن روی ردهٔ برآوردگرهای هم‌وردا یا نااریب (خطی نااریب) است که به ترتیب منجر به دو نوع برآوردگر بهینه یعنی بهترین برآوردگر هم‌وردا یعنی MREE[۱] و براوردگرهای نااریب با کمترین مخاطره یا واریانس به طور یکنواخت یعنی UMRUE[۲] یا UMVUE[۳] (بهترین براوردگرهای خطی نااریب یعنی BLUE[۴]) می‌شود.

[۱] Minimum risk equivariant estimator

[۲] Uniformly minimum risk unbiased estimator

[۳] Uniformly minimum variance unbiased estimator

[۴] Best linear unbiased estimator

تعریف ۱-۱۳: آمارهٔ $\delta(\mathbf{X})$ را براوردگر ناارایب با کمترین مخاطره به طور یکنواخت (موضعی) برای $g(\theta)$ گویند، هرگاه برای هر $\theta \in \Theta$، $E_\theta\left(\delta(\mathbf{X})\right) = g(\theta)$ و برای هر براوردگر ناارایب دیگر $T(\mathbf{X})$ از $g(\theta)$ و برای هر (یک) $\theta \in \Theta$ داشته باشیم $R(\theta,\delta) \le R(\theta,T)$ (پارسیان، ۱۳۹۵).

در حالت خاص که تابع زیان SEL باشد، تعریف ۱-۱۱ در حالتی که برای هر (یک) $\theta \in \Theta$ داشته باشیم

$$Var_\theta\left(\delta(\mathbf{X})\right) \le Var_\theta\left(T(\mathbf{X})\right)$$

منجر به تعریف UMVUE (براوردگرهای ناارایب با کمترین واریانس به طور موضعی یعنی LMVUE[۱]) می‌شود. قضیه زیر شرط وجود UMRUE را بیان می‌کند.

قضیه ۱-۲ ***(قضیهٔ رائو-بلاکول-لی‌من-شفه):*** اگر X یک متغیر تصادفی روی فضای اندازه‌پذیر $\mathcal{X}$ با توزیعی از خانوادهٔ $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ و T یک آمارهٔ بسندهٔ کامل برای $\theta \in \Theta$ باشد و همچنین تابع زیان $L(\theta,a)$ برای هر $\theta \in \Theta$ ثابت در a محدب باشد، آن‌گاه برای هر پارامتر براوردپذیر $g(\theta)$، یک UMRUE یکتا بر پایهٔ T وجود دارد (پارسیان، ۱۳۹۵).

۲. یافتن بهترین برآوردگرهای هم‌وردا

در این بخش بـا محـدودکردن روی ردهٔ برآوردگرهـای هم‌وردا یـک نوع برآوردگر بهینـه بـه نـام بهترین برآوردگر هم‌وردا می‌سازیم و برای راحتی خود را به خانواده با پارامترهای مکان از توزیع‌ها، محدود می‌کنیم. اگر

$$\mathcal{G} = \{g_c : g_c(\mathbf{x}) = (x_1 + c, \ldots, x_n + c); c \in \mathbb{R}\}$$

گروه تبدیل‌های خطی روی $\mathcal{X}$ و بردار تصادفی $\mathbf{X} = (X_1, \ldots, X_n)$ دارای تابع چگالی توأم

$$f_\theta(\mathbf{x}) = f(x_1 - \theta, \ldots, x_n - \theta)$$

باشد، در این صورت آمارهٔ ناوردای بیشین به صورت $\mathbf{Y} = (Y_1, \ldots, Y_{n-1})^T$ خواهد بود به طوری که برای هر $i = 1, \ldots, n-1$، $Y_i = X_i - X_n$. زیرا برای هر $\mathbf{x} = (x_1, \ldots, x_n) \in \mathcal{X}$ و هر $g_c \in \mathcal{G}$ داریم:

$$\mathbf{Y}\left(g_c(\mathbf{x})\right) = \mathbf{Y}\left((x_1 - x_n, \ldots, x_{n-1} - x_n)\right) = \mathbf{Y}(\mathbf{x}).$$

پس $\mathbf{Y}$ تحت گروه $\mathcal{G}$، نـاوردا اسـت. از طرف دیگر اگر برای هر $\mathbf{x} = (x_1, \ldots, x_n), \mathbf{x}' = (x'_1, \ldots, x'_n) \in \chi^n$، $\mathbf{Y}(\mathbf{x}) = \mathbf{Y}(\mathbf{x}')$، آن‌گاه برای هر $i = 1, \ldots, n-1$، $x_i - x_n = x'_i - x'_n$. قرار می‌دهیم $c = x'_n - x_n$. بنابراین برای هر $i = 1, \ldots, n-1$، $x'_i = x_i + c$ و یا $\mathbf{x}' = (x_1 + c, \ldots, x_{n-1} + c) = g_c(\mathbf{x})$.

بنابراین $\mathbf{Y}$ تحت گروه $\mathcal{G}$، ناوردای بیشین است.

اکنون با توجه به فرضیات بالا و استفاده ار حالت خاص تعریف ۱-۴ برای گروه‌های جمعی، تعریف زیر را مطرح می‌کنیم.

تعریف۲-۱: برآوردگر δ از θ را هم‌وردای مکانی گویند، هرگاه برای هر عدد حقیقی c و همچنین هر $x_1, \ldots, x_n \in \mathcal{X}$، $\delta(x_1 + c, \ldots, x_n + c) = \delta(x_1, \ldots, x_n) + c$ (لی‌من و کسلا، ۱۹۹۸).

---

۱ Locally minimum variance unbiased estimator

ماهیت برآوردگرهای هم‌وردای مکانی را می‌توان با در نظر گرفتن مفاهیم برآورد نقطه‌ای به این صورت توجیه کرد که اگر فرض کنیم $\delta(\mathbf{X})$ یک برآوردگر θ باشد و $g_c \in \mathcal{G}$ یک تبدیل مکانی باشد، یعنی $g_c(\mathbf{x}) = \mathbf{x} + c\mathbf{1_n}$ (که در آن $\mathbf{1_n}=(1,...,1)$)، در این صورت $\delta(g_c(\mathbf{X})) = \delta(\mathbf{X}+c\mathbf{1_n})$ یک برآوردگر برای $\theta+c$ خواهد بود، بنابراین:

$$\hat{\theta} = \delta(\mathbf{X}),$$
$$\hat{\theta} + c = \delta(\mathbf{X}+c\mathbf{1_n})$$

که با کم کردن طرفین این دو تساوی از هم داریم:

$$\delta(\mathbf{X}+c\mathbf{1_n}) = \delta(\mathbf{X}) + c\,.$$

برآوردگری مثل δ که در تساوی بالا صدق کند را هم‌وردای مکانی نامیم. اگر تابع زیان SEL باشد، آن‌گاه بهترین برآوردگر هم‌وردای مکانی θ به شکل برآوردگر پیتمن در می‌آید:

$$\hat{\theta}_p = \frac{\int \theta f\left(X_1-\theta,...,X_n-\theta\right)d\theta}{\int f\left(X_1-\theta,...,X_n-\theta\right)d\theta}.$$

نتیجهٔ عمومی زیر برای یافتن بهترین برآوردگر هم‌وردای مکانی مفید است.

قضیهٔ ۲-۱: فرض می‌کنیم بردار تصادفی $\mathbf{X}=(X_1,...,X_n)$ از خانوادهٔ توزیع‌های با پارامتر مکان $\mathcal{P}$ و $\mathcal{G}$ گروه تبدیل‌ها روی $\mathcal{X}$ و $L(\theta,a)=\rho(a-\theta)$ تابع زیان باشد و همچنین $\mathbf{Y}=(Y_1,...,Y_{n-1})^T$ که $Y_i = X_i - X_n$، $i=1,...,n-1$. اگر فرض کنیم که برآوردگر هم‌وردای مکانی $\delta_\circ$ برای θ با مخاطرهٔ متناهی $R(\theta,\delta_\circ)<\infty$ وجود داشته باشد و نیز برای هر $\mathbf{y}$، $\eta(\mathbf{y})=\eta^*(\mathbf{y})$ وجود داشته باشد که عبارت

$$E_\circ\left\{\rho\left(\delta_\circ(\mathbf{X})-\eta(\mathbf{y})\right)\middle|\mathbf{Y}=\mathbf{y}\right\}$$

را مینیمم کند، آن‌گاه بهترین برآوردگر هم‌وردای مکانی δ^* برای θ با مینیمم مخاطره وجود دارد که به صورت $\delta^*(\mathbf{X}) = \delta_\circ(\mathbf{X}) - \eta^*(\mathbf{Y})$ است (لی‌من و کسلا، ۱۹۹۸).

نتیجهٔ ۲-۱: تحت فرضیات قضیهٔ ۲-۱:

الف) اگر $\rho(a-\theta)=\left(a-\theta\right)^2$، آن‌گاه $\eta^*(\mathbf{y}) = E_\circ\left[\delta_\circ(\mathbf{X})\middle|\mathbf{Y}=\mathbf{y}\right]$.

ب) اگر $\rho(a-\theta)=\left|a-\theta\right|$، آن‌گاه $\eta^*(\mathbf{y})$ میانه $\delta_\circ(\mathbf{X})$ تحت توزیع شرطی $\mathbf{X}|\mathbf{Y}=\mathbf{y}$ هست (لی‌من و کسلا، ۱۹۹۸).

در صورت وجود آمارهٔ بسندهٔ کامل T برای θ، اگر بتوانیم تابعی از T مثل $\delta_\circ$ پیدا کنیم، به شرطی که $\mathbf{Y}$ یک آمارهٔ کمکی باشد، پیدا کردن برآوردگر هم‌وردا با مینیمم مخاطره ساده می‌شود. زیرا در این صورت می‌توانیم با به کارگیری قضیهٔ باسو (باسو، ۱۹۵۵) از توزیع غیر شرطی از T استفاده کنیم و بنابراین $\eta(\mathbf{y})$ ثابت می‌شود. مطلب اخیر را تحت سه مثال زیر بیان می‌کنیم.

مثال ۲-۱: فرض می‌کنیم $X_1,...,X_n$ یک نمونهٔ تصادفی از توزیع $N(\theta,\sigma^2)$ باشد که در آن $\theta\in\mathbb{R}$ مجهول و σ^2 معلوم است. دراین صورت برآوردگر هم‌وردای مکانی $\delta_\circ(X_1,...,X_n)=\overline{X}$ بسندهٔ کامل است و آمارهٔ ناوردای بیشین

$\mathbf{Y}$ یک آمارهٔ کمکی است. بنابراین با به کارگیری قضیهٔ باسو $\overline{X} \perp \mathbf{Y}$. اکنون برای هر تابع زیان به صورت $L(\theta,a)=\rho\left(|a-\theta|\right)$ که ρ نسبت به هر یک از مؤلفه‌هایش یک تابع اکیداً غیر نزولی است، $\eta^*(\mathbf{y})=0$ عبارت

$$\psi(\mathbf{y}) = E_\circ\left(\rho\left|\overline{X} - \eta(\mathbf{y})\right|\right)$$

را مینیمم می‌کند و لذا برآوردگر هم‌وردای مکانی با مینیمم مخاطره از θ برابر $\delta^*(\mathbf{X})=\overline{X}$ می‌شود. □

مثال ۲-۲: فرض می‌کنیم $X_1,\ldots,X_n$ یک نمونهٔ تصادفی از توزیع نمایی دم‌بریده $E(\theta,\sigma^2)$ با تابع چگالی احتمال مشترک زیر باشند:

$$f_\theta(x) = \tfrac{1}{\sigma^2} e^{-(x-\theta)/\sigma^2} I_{\{x:\, x\geq\theta\}}(x) \qquad \theta\in\mathbb{R}\,;(\text{معلوم } \sigma^2).$$

دراین جا $X_{(1)}$ برآوردگر هم‌وردای مکانی و همچنین یک آمارهٔ بسندهٔ کامل برای θ است. همچنین آماره $\mathbf{Y}$ که در قضیهٔ ۲-۱ تعریف شد، یک آمارهٔ کمکی است. پس طبق قضیهٔ باسو $X_{(1)} \perp \mathbf{Y}$ و لذا $\eta(\mathbf{y})=\eta^*(\mathbf{y})$ باید طوری تعیین شود که عبارت

$$\psi(\mathbf{y}) = E_\circ\left(\rho\left|X_{(1)} - \eta(\mathbf{y})\right|\right)$$

را مینیمم کند. بنابراین تحت SEL بهترین برآوردگر هم‌وردای مکانی θ به صورت

$$X_{(1)} - E_\circ(X_{(1)}) = X_{(1)} - \tfrac{\sigma^2}{n}$$

و تحت AEL به صورت

$$X_{(1)} - med_\circ(X_{(1)}) = X_{(1)} - \tfrac{\sigma^2 \ln 2}{n}$$

نوشته می‌شود.□

مثال ۲-۳: فرض می‌کنیم $X_1,\ldots,X_n$ یک نمونهٔ تصادفی از توزیع یکنواخت $U\left(\frac{\theta-1}{2\sigma^2},\frac{\theta+1}{2\sigma^2}\right)$ باشد که در آن σ^2 معلوم است. اختیار می‌کنیم $\delta_\circ(\mathbf{X})=\frac{X_{(1)}+X_{(n)}}{2}$ و سپس $\eta(\mathbf{y})$ را طوری پیدا می‌کنیم که

$$E_\circ\left[\rho\left(\delta_\circ(\mathbf{X})-\eta(\mathbf{y})\right)\middle|\mathbf{Y}=\mathbf{y}\right]$$

را مینیمم کند. توزیع شرطی $\delta_\circ(\mathbf{X})\,|\,\mathbf{Y}=\mathbf{y}$ فقط از طریق تفاضل $X_{(i)}-X_{(1)}\,;(i=2,\ldots,n)$ به $\mathbf{y}$ بستگی پیدا می‌کند. از طرف دیگر با استفاده از قضیهٔ باسو آمارهٔ بسندهٔ کامل $T=\left(X_{(1)},X_{(n)}\right)$ و آمارهٔ کمکی

$$Z_i = \frac{X_{(i)}-X_{(1)}}{X_{(n)}-X_{(1)}},\, i=2,\ldots,n-1$$

از هم مستقل هستند. بنابراین توزیع شرطی $\delta_\circ(\mathbf{X})\,|\,X_{(i)}-X_{(1)}$ معادل توزیع شرطی $\delta_\circ(\mathbf{X})\,|\,X_{(n)}-X_{(1)},Z_i$ خواهد بود که با توجه به استقلال Z_iها و T، معادل توزیع شرطی $\delta_\circ(\mathbf{X})\,|\,V=X_{(n)}-X_{(1)}$ خواهد بود که این توزیع حول صفر متقارن است و لذا $\delta_\circ(\mathbf{X})=\frac{X_{(1)}+X_{(n)}}{2}$ یک برآوردگر هم‌وردای مکانی با مینیمم مخاطره برای θ هست.

□

در انتهای این بخش به استقلال تصادفی یک تابع ناوردا و هموردا می‌پردازیم. به طور شهودی اگر آمارهٔ هموردای $T(X)$ بسنده باشد، شامل تمام اطلاعات X درباره پارامتر است و لذا می‌توان راجع به استقلال آن از آمارهٔ کمکی $U(X)$ که توزیعش به پارامتر بستگی ندارد صحبت کرد. از طرف دیگر می‌دانیم اگر $U(X)$ یک تابع ناوردا از X باشد، توزیع آن یک تابع ناوردا از پارامتر خواهد بود و اگر گروه $\mathcal{G}$ روی فضای پارامتر Θ به طور انتقالی عمل کند، $U(X)$ کمکی است. با توجه به این که $U(X)$ ناورداست اگر و تنها اگر تابعی از ناوردای بیشین باشد، پیشنهاد می‌شود که آمارهٔ ناوردای بیشین از $T(X)$ مستقل است.

قضیهٔ ۲-۲: فرض کنید گروه‌های $\mathcal{G}$ و $\bar{\mathcal{G}}$ روی $\mathcal{X}$ و $\mathcal{Y}$ عمل می‌کنند و تابع $(\bar{g},x)\rightarrow \bar{g}x$ از $\bar{\mathcal{G}}\times\mathcal{Y}$ به $\mathcal{Y}$ اندازه‌پذیر باشد و $\bar{\mathcal{G}}$ به طور انتقالی روی $\mathcal{Y}$ عمل کند و همچنین تابع $\tau:\mathcal{X}\rightarrow\mathcal{Y}$ اندازه‌پذیر و هموردا و تابع $h:\mathcal{X}\rightarrow\mathcal{Z}$ اندازه‌پذیر و نسبت به گروه $\mathcal{G}$ ناوردا باشد. برای متغیر تصادفی $X\in\mathcal{X}$ با توزیع P_o قرار دهید $Y=\tau(X)$ و $Z=h(X)$ و فرض کنید $\tau(X)$ یک آمارهٔ بسنده برای خانوادهٔ $\{gP_o:g\in\mathcal{G}\}$ از توزیع‌های روی $\mathcal{X}$ باشد. بنابراین وقتی برای $g\in\mathcal{G}$، X دارای توزیع gP_o است، آن‌گاه $Y\perp Z$ (ایتون، ۱۹۸۳).

از این که $\mathcal{G}$ به طور انتقالی روی $\{gP_o:g\in\mathcal{G}\}$ عمل می‌کند و $Z=h(X)$، نسبت به گروه $\mathcal{G}$ ناوردا است، توزیع آن تحت هر gP_o (برای هر $g\in\mathcal{G}$) یکسان است و لذا $Z=h(X)$ آمارهٔ کمکی است. بنابراین طبق قضیهٔ باسو از استقلال آن از آمارهٔ بسندهٔ کامل نتیجه می‌شود. در قضیهٔ ۲-۲ فرض کامل بودن برای برقراری قضیهٔ باسو با هم‌وردایی جابه‌جا شده است.

مثال ۲-۴: اگر $X_1,\dots,X_n$ یک نمونهٔ تصادفی از توزیع $N(\mu,\sigma^2)$ ($\mu\in\mathbb{R}$ و $\sigma>0$ نامعلوم) باشد. با در نظر گرفتن عمل گروه $(a,b)x=ax+b$ آمارهٔ بسندهٔ $\tau(X)=(S,\bar{X})$ هموردا است که در آن $S^2=\sum_{i=1}^n(X_i-\bar{X})^2$، زیرا $\tau\left((a,b)X\right)=(a,b)(S,\bar{X})$. آمارهٔ

$$h(X)=(\tau(X))^{-1}X=\frac{X-\bar{X}}{S}$$

ناوردای بیشین است، زیرا

$$h\left((a,b)X\right)=((a,b)\tau(X))^{-1}(a,b)X=(\tau(X))^{-1}X=h(X)$$

و همچنین اگر $\left(\tau(X)\right)^{-1}X=\left(\tau(Y)\right)^{-1}Y$، آن‌گاه $Y=[\tau(Y)(\tau(X))^{-1}]X$ و لذا X و Y در یک مدار یکسان قرار دارند. با توجه به انتقالی بودن گروه هر تابع ناوردا مثل $h(X)$ کمکی نیز هست که در این مثال خاص با توجه به این که آمارهٔ بسندهٔ هموردای $\tau(X)$، کامل نیز هست استقلال آن از آمارهٔ ناوردای $h(X)$ که کمکی نیز هست واضح است. ولی با توجه به برقراری شرایط قضیهٔ ۲-۲ در حالت کلی‌تر با نداشتن فرض کامل بودن $\tau(X)$ و کمکی بودن $h(X)$ نیز توانستیم استقلال آمارهٔ بسندهٔ هموردا را از یک آمارهٔ ناوردا نتیجه بگیریم.□

نتیجه ۲-۲: از این که $\mathcal{G}$ به طور انتقالی روی $\{gP_o:g\in\mathcal{G}\}$ عمل می‌کند و $Z=h(X)$، نسبت به گروه $\mathcal{G}$ ناوردا است توزیع آن تحت هر gP_o (برای هر $g\in\mathcal{G}$) یکسان است و لذا $Z=h(X)$ آمارهٔ کمکی است. بنابراین طبق قضیهٔ باسو از استقلال آن از آمارهٔ بسندهٔ کامل نتیجه می‌شود. در قضیهٔ ۲-۲ فرض کامل بودن برای برقراری قضیهٔ باسو با ناوردایی جابه‌جا شده است.

۳. به دست آوردن برآوردگرهای ناریب با کمترین مخاطره یکنواخت(UMRUE)

در بخش ۲ برای بــه دســـت آوردن برآوردگرهــای بهینه خود را به ردهٔ برآوردگرهای هم‌وردا محدود کردیم. در این بخش خود را بــه ردهٔ برآوردگرهــای نــاریــب محــدود می‌کنیم. خواص بهینــه متــداول در میــان برآوردگرهــای نــاریــب عبارتند از براوردگرهــای نــاریــب بــا کمترین واریــانس به طور یکنواخت (UMVUE)، براوردگرهای ناریب با کمترین واریانس به طور موضـــعی (LMVUE)، براوردگرهــای نــاریــب بــا کمترین مخاطره به طور یکنواخت (UMRUE) و بهترین براوردگرهای خطی ناریب (BLUE). حال اگر تابع زیان مربع خطا باشـد، مسئله به پیدا کردن UMVUE منتهی می‌شود که در بخش ۴ به آن می‌پردازیم. درحالتی که تابع زیان کلی باشــد، مســئله به پیدا کردن UMRUE منجر می‌شــود، که در این بخش خود را بــه یافتن UMRUE تحت تابع زیان نمایی خطی (LINEX) محدود می‌کنیم. با به کارگیری قضـــیهٔ رائو-بلاکول-لی‌من-شــفه (کســلا و برگر، ۲۰۰۱) می‌توان برای هر پارامتر LINEX-برآوردپذیر $g(\theta)$ برآورد LINEX-ناریب یکتا با کمترین مخاطره به دســت آورد، به این صــورت که اگر T یک آمارهٔ بســندهٔ کامل برای θ و $h(T)$ یک آمارهٔ مثبت باشــد که برای هر $\theta\in\Theta$، $E_\theta\left(h(T)\right)=e^{ag(\theta)}$، آن‌گاه $\delta^*(T)=\frac{1}{a}\ln h(T)$، UMRUE پارامتر $g(\theta)$ است (پارسیان و کرمانی، ۲۰۰۱). حال به ذکر مثال‌هایی در این زمینه می‌پردازیم.

مثال ۳-۱: فرض می‌کنیم $X_1,\dots,X_n$ یک نمونهٔ تصادفی از توزیع نمایی منفی $E(\mu,\sigma)$ با تابع چگالی زیر باشد:

$$f_{\mu,\sigma}(x)=\tfrac{1}{\sigma}\exp(-\tfrac{1}{\sigma}(x-\mu))I_{\{x:\ x>\mu\}}(x)$$

که μ و σ پارامترهای نامعلوم هستند. می‌خواهیم UMRUE پارامتر μ را پیدا کنیم. می‌دانیم $\left(X_{(1)},S\right)$ یک آمارهٔ بسندهٔ کامل برای (μ,σ) است که در آن $S=\sum_{i=1}^n(X_i-X_{(1)})$. از طرف دیگر از این که $X_{(1)}$ دارای توزیع $E(\mu,n\sigma)$ است، نتیجه می‌گیریم

$$E_{\mu,\sigma}\left[(1-a\sigma/n)\exp(aX_{(1)})\right]=(1-a\sigma/n)M_{X_{(1)}}(a)=e^{a\mu}.$$

همچنین با توجه به این که $2S/\sigma$ دارای توزیع $\chi^2_{2(n-1)}$ است، S کمکی است و $E_\sigma(S)=\sigma(n-1)$. لذا با توجه به استقلال $X_{(1)}$ و S (قضیهٔ باسو) داریم:

$$E_{\mu,\sigma}e^{aX_{(1)}+\ln\left(1-\frac{aS}{n(n-1)}\right)}=E_{\mu,\sigma}e^{aX_{(1)}}E_{\mu,\sigma}\left(1-\frac{aS}{n(n-1)}\right)=\frac{e^{a\mu}}{1-a\sigma/n}(1-a\sigma/n)=e^{a\mu}.$$

بنابراین UMRUE پارامتر μ (به شرط $a<0$) برابر مقدار زیر است:

$$\delta^*(X)=X_{(1)}+\tfrac{1}{a}\ln\left(1-\frac{aS}{n(n-1)}\right)$$

و تابع مخاطرهٔ این برآوردگر به صورت زیر به دست می‌آید:

$$R(\delta^*,\mu)=-aE\left[X_{(1)}-\mu+\tfrac{1}{a}\ln\left(1-\frac{aS}{n(n-1)}\right)\right].\square$$

مثال ۳-۲: فرض می‌کنیم $X_1,\dots,X_n$ یک نمونهٔ تصادفی از توزیع نمایی با میانگین σ باشد. می‌خواهیم UMRUE پارامتر σ را پیدا کنیم. می‌دانیم $T=n\bar{X}$ یک آمارهٔ بسندهٔ کامل برای σ است. با در نظر گرفتن تابع بسل به صورت

$$J_{\circ}(x)=\sum_{k=0}^{\infty}\left(\tfrac{x}{2}\right)^{2k}/\left(k!\right)^{2}$$

برای هر $\sigma>0$ داریم:

$$E_{\sigma}\left(J_{\circ}(2\sqrt{aX_1})\right)=\sum_{k=0}^{\infty}a^{k}E_{\sigma}(X_1^{\,k})/\left(k!\right)^{2}=\sum_{k=0}^{\infty}a^{k}\sigma^{k}k!/\left(k!\right)^{2}=e^{a\sigma}.$$

با توجه به مثبت بودن مقادیر آمارهٔ $J_{\circ}(2\sqrt{aX_1})$، می‌توان گفت برآوردگر LINEX-نااریب (به شرط $a>0$) برابر با

$$\delta(X_1)=\tfrac{1}{a}\ln\left(J_{\circ}(2\sqrt{aX_1})\right)$$

است، بنابراین UMRUE پارامتر σ (که $a>0$) به صورت

$$\delta^{*}(T)=\tfrac{1}{a}\ln\left[E\left(J_{\circ}(2\sqrt{aX_1})\middle|T\right)\right]$$

به دست می‌آید. از طرف دیگر چون آمارهٔ $U=X_1/T$ دارای توزیع $Beta(1,n-1)$ است، توزیع U به پارامتر مجهول بستگی ندارد (U آمارهٔ کمکی است) و لذا طبق قضیهٔ باسو U و T مستقل هستند. بنابراین:

$$\begin{aligned}
g(T)&=E\left(J_{\circ}(2\sqrt{aX_1})\middle|T\right)\\
&=E\left(J_{\circ}\left(2\sqrt{\frac{aX_1}{T}}\sqrt{T}\right)\middle|T\right)\\
&=E\left(J_{\circ}(2\sqrt{aUT})\right)\\
&=\int_0^1\sum_{k=0}^{\infty}\frac{(\sqrt{auT})^{2k}}{\left(k!\right)^{2}}(n-1)(1-u)^{n-2}du\\
&=\sum_{k=0}^{\infty}\frac{(aT)^{k}}{\left(k!\right)^{2}}(n-1)\int_0^1u^{k}(1-u)^{n-2}du\\
&=\sum_{k=0}^{\infty}\frac{(aT)^{k}}{\left(k!\right)^{2}}(n-1)\frac{\Gamma(n-1)\Gamma(k+1)}{\Gamma(n+k)}\\
&=\sum_{k=0}^{\infty}\frac{\Gamma(n)}{\Gamma(n+k)}\frac{(aT)^{k}}{k!}=H_n(aT)
\end{aligned}$$

که در آن تابع $H_{\alpha}(.)$ به صورت زیر تعریف شده است:

$$H_{\alpha}(\lambda)=\sum_{k=0}^{\infty}\frac{\Gamma(\alpha)}{\Gamma(\alpha+k)}\frac{\lambda^{k}}{k!}.$$

بنابراین UMRUE پارامتر σ (به شرط $a>0$) به صورت زیر به دست می‌آید:

$$\delta^{*}(T)=\tfrac{1}{a}\ln g(T)=\tfrac{1}{a}\ln H_n(aT).$$

حال اگر $a<0$ باشد، با در نظر گرفتن تابع بسل

$$I_{\circ}(x)=\sum_{k=0}^{\infty}(-1)^{k}\left(\tfrac{x}{2}\right)^{2k}/\left(k!\right)^{2}$$

به سادگی ثابت می‌شود که

$$\tfrac{1}{a}\ln\left(I_{\circ}(2\sqrt{-aX_1})\right)$$

یک برآوردگر LINEX-نااریب σ (به شرط $a<0$) است و به طور مشابه داریم:

$$h(T)=E\left(I_{\circ}(2\sqrt{aUT})\right)=\sum_{k=0}^{\infty}\frac{\Gamma(n)}{\Gamma(n+k)}\frac{(-aT)^k}{k!}=H_n(-aT)$$

و بنابراین

$$\delta^{**}(T)=\tfrac{1}{a}\ln H_n(-aT)$$

یک UMRUE پارامتر σ (به شرط $a<0$) است.□

مثال ۳-۳: فرض می‌کنیم $X_1,\dots,X_n$ یک نمونهٔ تصادفی n تایی ($n\ge 5$) از توزیع $N(\mu,\sigma^2)$ باشد، که در آن هر دو پارامتر μ و σ^2 نامعلوم هستند. می‌خواهیم UMRUE پارامتر σ^2 را پیدا کنیم. می‌دانیم که $(\bar{X},S^2)$ یک آمارهٔ بسندهٔ کامل برای (μ,σ^2) است و $S^2=\sum_{i=1}^{\infty}(X_i-\bar{X})^2$ دارای توزیع $\Gamma(\frac{n-1}{2},2\sigma^2)$ یا همان $\sigma^2\chi^2_{n-1}$ است. مشابه مثال‌های قبلی به راحتی می‌توان نشان داد که

$$\tfrac{1}{a}\ln H_{\frac{n-1}{2}}(\tfrac{a}{2}\sigma^2)$$

یک UMRUE پارامتر σ^2 (به شرط $a<0$) است. برای حالت $a>0$ از این که برای هر $i=1,\dots,n$:

$$M_{X_i}(t)=E(e^{tX_i})=\exp(\mu t+\sigma^2t^2/2)$$

داریم:

$$M_{X_1-X_2}(t)=M_{X_1}(t)M_{-X_2}(t)=M_{X_1}(t)M_{X_2}(-t)=e^{\sigma^2t^2}.$$

حال با انتخاب $h(X)=a^{-\frac{1}{2}}(X_1-X_2)$ به سادگی تحقیق می‌شود که برای هر μ و σ^2:

$$E_{\mu,\sigma^2}\left[e^{ah(X)}\right]=M_{X_1-X_2}(a^{-\frac{1}{2}}a)=M_{X_1-X_2}(\sqrt{a})=e^{a\sigma^2}$$

یعنی $h(X)$ یک برآوردگر LINEX-نااریب σ^2 (به شـرط $a>0$) اسـت. در نتیجه با به کارگیری قضـیهٔ رائو-بلاکول-لی‌من-شـفه و با اختیار کردن $g(S^2)=E\left[e^{ah(X)}\middle|(\bar{X},S^2)\right]$، $\frac{1}{a}\ln g(S^2)$ یک UMRUE برای پارامتر σ^2 (به شرط $a>0$) است (پارسیان، ۱۳۸۰).□

۴. یافتن برآوردگرهای نااریب با کمترین واریانس به طور یکنواخت(UMVUE)

اگر برای بـه دســـت آوردن برآوردگرهای بهینه رده را به برآوردگرهای نااریب محدود کنیم و تابع زیان مربع خطا باشـــد، مســـئله به پیدا کردن برآوردگر نااریب با کمترین واریانس به طور یکنواخت (UMVUE) منتهی می‌شــود. همانند بخش ۳ اســتقلال آمارهٔ بســندهٔ کامل از آمارهٔ کمکی و به کار گیری قضـــیهٔ باســو در پیدا کردن UMVUEها می‌تواند مفید باشــد. قضـیهٔ رائو-بلاکول-لی‌من-شـفه (کسـلا و برگر، ۲۰۰۱) می‌تواند ابزار مناسـبی برای پیدا کردن UMVUEها باشـد. علاوه بر آن از دو روش دیگر که توسـط سـت و وارد (۱۹۶۹) و ایتون و موریس (۱۹۷۰) مطرح شـده است نیز می‌توان به هدف مورد نظر رسـید. در پایان این بخش به مقایســه این دو روش نیز می‌پردازیم. برای دیدن کاربرد قضـیهٔ رائو-بلاکول-لی‌من-شفه در پیدا کردن UMVUEها به مثال زیر می‌پردازیم.

مثال ۴-۱: فرض می‌کنیم $X_1,\dots,X_n$ یک نمونهٔ تصـادفی n تایی از توزیع $LN(\mu,\sigma^2)$ باشـد که هر دو پارامتر μ و σ^2 مجهول هستند. بنا براین میانگین و واریانس متغیر تصادفی X از توزیع بالا به صورت زیر قابل محاسبه است:

$$\theta = E_{\mu,\sigma^2}(X) = e^{\mu+\sigma^2/2}, \qquad (۱)$$

$$\eta^2 = Var_{\mu,\sigma^2}(X) = e^{2\mu+\sigma^2}\left(e^{\sigma^2}-1\right). \qquad (۲)$$

برای محاســبه برآورد θ و η^2، شیوهٔ معمول استفاده از تبدیل $Y_i = \ln X_i$، $i=1,...,n$ با توزیع $N(\mu,\sigma^2)$ است که با این تبدیل مســئله به برآورد پارامترهای توزیع نرمال منجر می‌شــود. پس فرض می‌کنیم Y_iها دارای توزیع $N(\mu,\sigma^2)$ هســتند به طوری که روابط (۱) و (۲) برای آن‌ها برقرار باشــد. به وضوح $(\overline{Y}, S_Y^2)$ آمارهٔ بسندهٔ توأم کامل برای (μ,σ^2) هست که در آن $S_Y^2 = \sum_{i=1}^n (Y_i - \overline{Y})^2$. براوردگر حداکثر درستنمایی (MLE[۱]) برای θ و η^2 به صورت زیر است:

$$\hat{\theta} = \exp\left(\overline{Y} + S_Y^2/2n\right),$$

$$\hat{\eta}^2 = \exp\left(2\overline{Y} + S_Y^2/n\right)\left\{\exp\left(S_Y^2/n\right)-1\right\}$$

که هر دو به ترتیب برای θ و η^2 اریب هستند، زیرا:

$$E(\hat{\theta}) = \theta \exp\left\{-\tfrac{n-1}{2n}\sigma^2\right\}\left(1-\tfrac{1}{n}\sigma^2\right)^{-(n-1)/2},$$

$$E(\hat{\eta}^2) = \eta^2 \frac{\exp\left\{(\frac{2}{n}-1)\sigma^2\right\}}{\exp\left\{\sigma^2-1\right\}}\left((1-\tfrac{4}{n}\sigma^2)^{-\frac{n-1}{2}} - (1-\tfrac{2}{n}\sigma^2)^{-\frac{n-1}{2}}\right).$$

فینی (۱۹۴۴) با تعریف سری زیر:

$$f(t) = 1 + t + \frac{n-1}{n+1}\frac{t^2}{2!} + \frac{(n-1)^2}{(n+1)(n+3)}\frac{t^3}{3!} + \dots$$

MLE برای θ و η^2 را به صورت زیر ساده کرد:

$$\hat{\theta} = \exp(\overline{Y}) f\left(\tfrac{1}{2n}S_Y^2\right),$$

$$\hat{\eta}^2 = \exp\left(2\overline{Y}\right)\left\{f\left(\tfrac{2}{n}S_Y^2\right) - f\left(\tfrac{n-2}{n(n-1)}S_Y^2\right)\right\}$$

که برای θ و η^2 نااریب هســتند و لذا با توجه به قضــیهٔ رائو-بلاکول-لی‌من-شــفه ، $\hat{\theta}$ و $\hat{\eta}^2$ به ترتیب UMVUE برای θ و η^2 هســتند. اکنون می‌خواهیم با اســتفاده از قضــیهٔ باســو و قضــیهٔ رائو-بلاکول-لی‌من-شــفه و تبدیلات متعامد یک صــورت انتگرالی برای UMVUE پارامترهای θ و η^2 به دست آوریم.

از این که $(\overline{Y}, S_Y^2)$ آمارهٔ بســندهٔ کامل توأم برای (μ,σ^2) و $\exp(Y_n)$ یک برآوردگر نااریب برای θ اســت، می‌توان با اســتفاده از قضــیهٔ رائو-بلاکول نتیجه گرفت که $E\{\exp(Y_n)|\overline{Y}, S_Y^2\}$ یک UMVUE برای θ اســت. حال اگر اختیار کنیم

$$\frac{Y_n - \overline{Y}}{S_Y} = -\sqrt{\tfrac{n-1}{n}}U$$

با استفاده از تبدیل متعامد هلمرت به سادگی نتیجه می‌شود که تابع چگالی U به صورت زیر است:

$$f_U(u) = \frac{1}{B\left(\frac{1}{2},\frac{n-2}{2}\right)}(1-u^2)^{\frac{n-4}{2}} \quad ; \quad -1<u<1$$

[۱] Maximum likelihood estimator

که بستگی به پارامتر مجهول ندارد و لذا U یک آمارهٔ کمکی است. بنابراین با توجه به قضیهٔ باسو و در نتیجه U $(Y_n-\bar{Y})/S_Y$ از $\left(\bar{Y},S_Y^2\right)$ مستقل است و لذا:

$$\begin{aligned}\hat{\theta}&=E\left(\exp(Y_n)\middle|\bar{Y},S_Y^2\right)\\&=E\left\{\exp\left(\bar{Y}+\frac{Y_n-\bar{Y}}{S_Y}S_Y\right)\middle|\bar{Y},S_Y^2\right\}\\&=E\left\{\exp(\bar{Y})\middle|\bar{Y},S_Y^2\right\}E\left\{\exp\left(-\sqrt{\tfrac{n-1}{n}}S_YU\right)\middle|\bar{Y},S_Y^2\right\}\\&=\exp(\bar{Y})E\left\{\exp\left(-\sqrt{\tfrac{n-1}{n}}S_YU\right)\right\}.\end{aligned}\tag{۳}$$

شن (۱۹۹۸) با انجام محاسباتی مقدار بالا را به صورت زیر ساده کرد:

$$\hat{\theta}=\frac{\exp(\bar{Y})}{B\left(\frac{1}{2},\frac{n-2}{2}\right)}\int_{-1}^{1}\exp\left(-\sqrt{\tfrac{n-1}{n}}S_Yu\right)(1-u^2)^{\frac{n-4}{2}}du\tag{۴}$$

که یک UMVUE به صورت انتگرالی برای θ است. البته شن (۱۹۹۸) در همان مقاله به سادگی معادل بودن UMVUE به دو صورت بالا یعنی صورت انتگرالی (۴) و صورت سری (۳) را اثبات می‌کند.

اکنون با توجه به این که $\exp(2Y_n)-\exp\left(Y_n+Y_{n-1}\right)$ یک برآوردگر ناریب برای η^2 اسـت، با اسـتفاده از قضـیهٔ رائو-بلاکول-لی‌من-شفه می‌توان نتیجه گرفت که UMVUE برای η^2 به صورت زیر است:

$$\hat{\eta}^2=E\left\{e^{2Y_n}-e^{Y_n+Y_{n-1}}\middle|\bar{Y},S_Y^2\right\}.\tag{۵}$$

اکنون مشابه با قسمت قبل، UMVUE پارامتر η^2 را پیدا می‌کنیم:

$$\begin{aligned}E\left(\exp(2Y_n)\middle|\bar{Y},S_Y^2\right)&=E\left\{\exp\left(2\bar{Y}-2\sqrt{\tfrac{n-1}{n}}S_YU\right)\middle|\bar{Y},S_Y^2\right\}\\&=\exp(2\bar{Y})E\left\{\exp\left(-2\sqrt{\tfrac{n-1}{n}}S_YU\right)\right\}\\&=\frac{\exp(2\bar{Y})}{B\left(\frac{1}{2},\frac{n-2}{2}\right)}\int_{-1}^{1}\exp\left(-2\sqrt{\tfrac{n-1}{n}}S_Yu\right)(1-u^2)^{\frac{n-4}{2}}du.\end{aligned}$$

از طرف دیگر با توجه به قضیهٔ باسو، آمارهٔ کمکی $V=(Y_n+Y_{n-1}-2\bar{Y})/S_Y$ از $\bar{Y}$ و S_Y مستقل است، لذا:

$$\begin{aligned}E\left\{\exp\left(Y_n+Y_{n-1}\right)\middle|\bar{Y},S_Y^2\right\}&=\exp(2\bar{Y})E\left\{\exp(S_YV)\middle|\bar{Y},S_Y^2\right\}\\&=\exp(2\bar{Y})E\left\{\exp(S_YV)\right\}\\&=\frac{\exp(2\bar{Y})}{B\left(\frac{1}{2},\frac{n-2}{2}\right)}\int_{-1}^{1}\exp\left(\sqrt{\tfrac{2(n-2)}{n}}S_Yv\right)(1-v^2)^{\frac{n-4}{2}}dv.\end{aligned}$$

با کم کردن طرفین دو رابطهٔ اخیر داریم:

(۶)

$$\hat{\eta}^2=\frac{\mathrm{e}^{2\bar{Y}}}{B\left(\frac{1}{2},\frac{n-2}{2}\right)}\int_{-1}^{1}\left[e^{-2\sqrt{\frac{n-1}{n}}S_Y v}-\mathrm{e}^{\sqrt{\frac{2(n-2)}{n}}S_Y v}\right](1-v^2)^{\frac{n-4}{2}}dv$$

و بنابراین $\hat{\eta}^2$ یک UMVUE به صورت انتگرالی (۶) برای η^2 هست، که به طور مشابه برابر صورت سری به دست آمده دررابطهٔ (۵) هست.□

اکنون به بیان قضیهٔ ست و وارد (۱۹۶۹) می‌پردازیم.

قضـــیـهٔ۴-۱: متغیر تصـــادفی Z با مقادیر حقیقی و با تابع توزیع $F_\theta(z)$ را در نظر می‌گیریم و فرض می‌کنیم T آمارهٔ بسندهٔ کامل برای $\theta\in\Theta$ باشد. تابع $V(z,t)$ که در شرایط زیر صدق می‌کند را در نظر می‌گیریم:

الف) $V(z,t)$ نسبت به مؤلفهٔ اول (برای هر مقدار ثابت t) یک تابع اکیداً صعودی است.

ب) $U\equiv V(Z,T)$ یک آمارهٔ کمکی با تابع توزیع $H(u)$ است.

در این صورت UMVUE برای $F_\theta(z)$ به صورت $H(V(z,T))$ خواهد بود (ست و وارد، ۱۹۶۹).

اثبات: با استـــفاده از قضـــیهٔ باســو T و U مســـتقل هســـتند. همچنین با توجه به این که $I_{\{z:\,Z\le z\}}(z)$ یک برآوردگر نااریب برای $F_\theta(z)$ اســت، با به کارگیری قضیهٔ رائو-بلاکول-لی‌من-شفه، UMVUE برای $F_\theta(z)$ به صورت زیر به دست می‌آید:

$$\begin{aligned}E\left[I_{\{z:Z\le z\}}(z)\middle|T=t\right]&=P\left(Z\le z\middle|T=t\right)\\&=P\left(V(z,T)\le V(z,T)\middle|T=t\right)\\&=P\left(U\le V(z,t)\right)\qquad(T\perp U)\\&=H\left(V(z,t)\right).\square\end{aligned}$$

مثال ۴-۲: فرض می‌کنیم $X_1,\dots,X_n$ یک نمونهٔ تصادفی با اندازه $n\ge 3$ از توزیع $N(\mu,\sigma^2)$ باشد که $\mu\in\mathbb{R}$ و $\sigma>0$. می‌خواهیم برای $P_{\mu,\sigma^2}(X_1\le x)$ یک UMVUE پیدا کنیم. کولموگوروف (۱۹۵۰)، لیبرمن و رزنیکوف (۱۹۵۵) و همچنین بارتون (۱۹۶۱) و باسو (۱۹۶۴)، برای حل این مسئله روش‌هایی ارائه دادند.

اگر قرار دهیم $S^2=\sum_{i=1}^n(X_i-\bar{X})^2$، آن‌گاه $T=(\bar{X},S)$ آمارهٔ بســـندهٔ کامل برای (μ,σ^2) هســـت و همچنین $U=(X_1-\bar{X})/S$ یک آمارهٔ کمکی اســـت، لذا با اســـتفاده از قضـــیهٔ ۴-۱، UMVUE برای $F_{\mu,\sigma^2}(x)$ به صـــورت $H\left((x-\bar{X})/S\right)$ داده می‌شـــود کـه H تـابع توزیع کنـاری U اســـت و تابع چگالی احتمال کناری U در کتاب رائو (۱۹۷۳، صفحهٔ۳۲۳) به دست آمده است. رائو برای این منظور تعریف می‌کند:

$$V=\frac{\sqrt{n}}{n-1}U.$$

بنابراین V دارای تابع چگالی احتمال زیر است:

$$f_V(v)=k(1-v^2)^{\frac{n-4}{2}}\quad;\ |v|\le 1$$

و در نتیجه UMVUE برای $F_{\mu,\sigma^2}(x)$ برابر است با

$$\int_{-1}^{\frac{\sqrt{n}}{n-1}\cdot\frac{x-\bar{X}}{S}}f_V(v)dv\quad.\square$$

مثـال ۴-۳: فرض می‌کنیم $X_1,...,X_n$ یک نمونهٔ تصـادفی با اندازه $n \geq 2$ از توزیع وایبل $W(p,\theta)$ با تابع چگالی زیر باشد:

$$f_\theta(x)=\tfrac{p}{\theta}x^{p-1}\exp\left(-x^p/\theta\right) \quad ; x>0, \theta>0$$

به طوری که $p>0$ معلوم است. در این حالت $T=\sum_{i=1}^n X_i^p$ آمارهٔ بسندهٔ کامل برای θ خواهد بود. همچنین از این که $X_1^p,...,X_n^p$ یک نمونهٔ تصادفی از توزیع نمایی $E(\theta)$ هستند، بنابراین $U=X_1^p/T$ دارای توزیع $Beta(1,n-1)$ است، لذا U یک آمارهٔ کمکی با تابع توزیع $F(u)=1-(1-u)^{n-1}$ است. بنابراین با توجه به قضیهٔ ۴-۱، UMVUE برای

$$P_\theta(X_1\leq x)=P_\theta(X_1^p\leq x^p)$$

برابر است با:

$$k(T)=\begin{cases}1-(1-x^p/T)^{n-1} & T>x^p\\ 0 & T\leq x^p.\end{cases}$$

در حالت خاص که $p=1$ باشد، توزیع وایبل برابر توزیع نمایی $E(\theta)$ خواهد شد و UMVUE برای

$$P_\theta(X_1\leq x)=1-\exp(-x/\theta)$$

به شرط این که $T>x$ باشد برابر

$$1-\left(1-\frac{x}{T}\right)^{n-1}$$

خواهد بود.□

مثال ۴-۴: فرض می‌کنیم $X_1,...,X_n$ یک نمونهٔ تصادفی از توزیع $\Gamma(\alpha,\beta)$، با تابع چگالی زیر باشد:

$$f_\beta(x)=\tfrac{1}{\beta^\alpha\Gamma(\alpha)}e^{-x/\beta}x^{\alpha-1} \quad x\geq 0 \ ; \ \alpha,\beta>0$$

کـه در آن α معلوم اســـت، در این صـــورت $T=\sum_{i=1}^n X_i$ آمـارهٔ بســـنـدهٔ کـامـل برای β اســـت و چون توزیع $U=X_1/T$ یعنی $B(\alpha,(n-1)\alpha)$ به پارامتر مجهول β بســتگی ندارد، لذا U یک آمارهٔ کمکی اسـت. بنابراین با استفاده از قضیهٔ ۴-۱ با فرض این که $0\leq x/t\leq 1$، UMVUE برای $P_\theta\left(X_1\leq x\right)$ برابر است با:

$$\tfrac{1}{B(\alpha,(n-1)\alpha)}\int_0^{x/T} y^{\alpha-1}(1-y)^{(n-1)\alpha-1}dy \ .\square$$

همچنین ســـت و وارد (۱۹۶۹) برای توزیع نمـایی دو پـارامتری در حـالت‌های مختلف که یکی از پارامترها و یا هر دو آن‌ها مجهول باشند به روش ذکر شده UMVUE پیدا کرده‌اند.

ایتون و موریس (۱۹۷۰) به منظور تسـهیل در به دست آوردن UMVUE مشابه قضیهٔ ۴-۱ قضیه‌ای را مطرح کردند. روش کار به این صـورت اسـت که اگر یک برآوردگر ناریب از یک پارامتر را بتوان به صورت تابعی از آمارهٔ بسندهٔ کامل و یک آمارهٔ کمکی بیان کرد، آن‌گاه UMVUE پارامتر را می‌توان به صـورت امید ریاضـی همان تابع نوشـت که در آن امید ریاضـی بر حســب توزیع کناری آمارهٔ کمکی به دسـت می‌آید. بنابراین چون با توجه به قضـیهٔ باسـو آمارهٔ بســندهٔ کامل از آمارهٔ کمکی

مســتقل است، این آماره در محاسبهٔ UMVUE هیچ نقشی ندارد. اکنون به بیان قضیهٔ ایتون و موریس (۱۹۷۰) می‌پردازیم که در اثبات آن به طور مستقیم از قضیهٔ باسو استفاده می‌شود.

قضـــیهٔ ۴-۲: فرض می‌کنیم $h(\mathbf{X})=W(T,U)$ یک برآوردگر ناریب برای پارامتر $\gamma(\theta)$ باشــد و وقتی که T آمارهٔ بسندهٔ کامل برای θ و U یک آمارهٔ کمکی باشد. در این صورت UMVUE برای $\gamma(\theta)$ برابر است با:

$$\gamma^*(T)=E_U\left[W(T,U)\right]$$

که در آن E_U امید ریاضی بر حسب توزیع کناری U هست (ایتون و موریس، ۱۹۷۰).

اثبات: با اســتفاده از قضــیهٔ باسو T و U مستقل هستند. پس به کمک قضیهٔ رائو-بلاکول-لی‌من-شفه، UMVUE برای $\gamma(\theta)$ برابر است با:

$$\begin{aligned} E\left(h(\mathbf{X})\middle|T=t\right) &= E\left(W(T,U)\middle|T=t\right)\\ &= E\left(W(t,U)\middle|T=t\right) \qquad (T\perp U)\\ &= E_U\left(W(t,U)\right)\\ &= \gamma^*(t). \square \end{aligned}$$

ایتون و موریس (۱۹۷۰) در قضیهٔ ۴-۲ فرضی که ست و وارد (۱۹۷۰) در قضیهٔ ۴-۱ در نظر گرفته بودند یعنی وجود یک تابع $V(z,t)$ که برای هر مقدار ثابت t روی مؤلفهٔ اول نانزولی باشد را لازم ندانسته‌اند، بلکه آن‌ها یک برآوردگر ناریب که تابعی از آمارهٔ بسندهٔ کامل و آمارهٔ کمکی باشد را در نظر گرفته‌اند که این کار نسبت به روش ست و وارد (۱۹۶۹) از نظر کاربرد بیشتر مورد استفاده قرار می‌گیرد. به طور نمونه مورد اخیر در زمینهٔ گروه‌های ناوردا به کار می‌رود و منجر به کاربردهای جالبی شامل تعمیم چند متغیره از نتیجهٔ کولموگوروف می‌گردد که در مثال ۴-۲ همین بخش به آن اشاره شده است.

۶. نتیجه گیری

برای پیدا کردن برآوردگرهای بهینه از روش محدود کردن ردهٔ برآوردگرها به ردهٔ برآوردگرهای هم‌وردا یا ناریب (که به ترتیب منجر به بهترین برآوردگر هم‌وردا و برآوردگر ناریب با کمترین مخاطره می‌شود)، استقلال آمارهٔ بسندهٔ کامل از آمارهٔ کمکی و به کارگیری قضیهٔ باسو می‌تواند در ساده‌تر کردن محاسبات راه گشا باشد. ولی در برخی مسائل آماری با گروه تبدیلات انتقالی، می‌توان به جای آمارهٔ بسندهٔ کامل از تابع هم‌وردا استفاده کرد و در این حالت به جای استفاده از قضیهٔ باسو، استقلال یک تابع ناوردا و یک تابع بسندهٔ هم‌وردا را می‌توان نتیجه گرفت. در این حالت فرض کامل بودن برای برقراری قضیهٔ باسو با هم‌وردایی و انتقالی بودن گروه تبدیلات جابه‌جا شده است. لازم به یادآوری است که انتقالی بودن گروه موجب می‌شود هر تابع ناوردا کمکی نیز باشد، ولی با این حال کامل نبودن یک آمارهٔ بسنده نیز می‌تواند استقلال آن از یک آمارهٔ ناوردا را نتیجه دهد. واضح است که در حالتی که گروه تبدیلات انتقالی و آمارهٔ بسندهٔ کامل، هم‌وردا نیز باشد، قضیهٔ باسو این دیدگاه را نتیجه می‌دهد. عکس این مطلب همواره صحیح نیست و به این صورت قابل اصلاح است که اگر در یک آمارهٔ بسندهٔ کامل نسبت به یک گروه انتقالی هم‌وردا نیز باشد لزومی ندارد هر آمارهٔ کمکی از آن مستقل باشد بلکه می‌توان آمارهٔ کمکی که ناوردا نیز باشد پیدا کرد که از آمارهٔ بسندهٔ کامل هم‌وردای داده شده مستقل باشد. با پیدا کردن شرایطی که یک آمارهٔ

کمکی، ناوردا نیز باشد می‌توان این نتایج را گسترش داد که در این حالت قضیهٔ باسو نتیجه می‌شود. البته این مسئله نیاز به بررسی و تحقیقات بیشتری دارد و امید است محققان در پیشبرد و تعمیم آن کوشا باشند.

References

۱. پارسیان، ا.، ۱۳۸۰، *برآورد نقطه‌ای نااریب*، ندا، انجمن آمار ایران، سال اول، شماره اول، ص (۱۴-۹).

۲. پارسیان، ا.، ۱۳۹۵، استنباط آماری، جلد اول، انتشارات علمی پارسیان.

3. Barton, D. E., ۱۹۶۱, Unbiased Estimation of a Set of Probabilities, *Biometrika*, 49, pp.227-229.

4. Basu, D., 1955, On Statistics Independent of a Complete Sufficient Statistic, *Sankhya*, 15, pp.377-380.

5. Basu, A. P., 1964, Estimates of Reliability for some Distributions Useful in Life Testing, *Technometrics*, 6, pp.215-219.

6. Casella, G. and Berger, R., 2001, *Statistical Inference*, ۲nd Edition, Wadsworth, Pacific Groves, California.

7. Eaton, M. L. , 1983, Multivariate Statistics, A Vector Space Approach. Wiley, New York.

8. Eaton, M. L. and Morris, C.N., 1970, The Application of Invariance to Unbiased Estimation, *Ann. Math. Statist.*, 41, pp.1708-1716.

9. Finney, D. J., 1941, On the Distribution of a Variate whose Logarithm is Normally Distributed, *J. R. Stat. Soc. Suppl.*, 7, pp. 155-161.

10. Folland, G. B., 1995, A Course in Abstract Harmonic Analysis. CRC Press. Boca Raton.

11. Kolmogorov, A. N. 1950, Unbiased Estimators, *Izvestiaa Akad. Nauk SSSR Series Math.*, 14, pp.303-326.

12. Lehmann, E. L. 1951, A General Concept of Unbiasedness, *Ann. Math. Statist*., 22, pp.587–592.

13. Lehmann, E. L. and Casella, G., 1998, *Theory of Point Estimation*, 2nd Edition, Springer-Verlag, New York.

14. Lieberman, G. J. and Resnikoff, G. J., 1955, Sampling Plans for Inspection by Variables, *J. Amer. Statist. Assoc.*, 50, pp.457- 516.

15. Parsian, A and Kirmani, S.N.U.A., 2001, Estimation Under LINEX Loss Function, *Handbook of Applied Econometrics and Statistical Inference*, eds. Uplah et al, Ch. 4, pp. 53-76.

16. Rao, C. R., 1973, *Linear Statistical Inference and its Applications*, 2nd Edition, Wiley, New York.

17. Roman, S., 2012, Fundamentals of Group Theory: An Advanced Approach. Birkhauser, Basel.

18. Sathe, Y. S. and Varde, S. R., 1969, On Minimum Variance Unbiased Estimation of Reliability, *Ann. Math. Statist.*, 40, pp.710-714.

19. Shen, W. H., 1998, Estimation of Parameters of a Lognormal Distribution, *Taiwanese J. of Math*., 2, pp.243-250.

20. Zacks, S., 1971, *The Theory of Statistical Inference*, Johns Wiley and Sons, New York.