ریاضی و جامعه
شاپا (چاپی): ۶۴۹۳-۲۳۴۵، شاپا(الکترونیکی): ۶۵۰۷-۲۳۴۵
جلد ۴، شماره ۲ (۱۳۹۸)، صص. ۲۳-۱۱

http://www.ui.ac.ir
http://math-sci.iranjournals.ir

آماره‌های کمکی چگونه می‌توانند آگاهی بخش باشند؟

مهدی شمس

چکیده. آماره‌های کمکی از این که توزیعشان به پارامتر مجهول بستگی ندارد ظاهراً هیچ اطلاعی راجع به پارامتر مجهول ارائه نمی‌دهند. اگر یک آماره بسنده، کامل نیز باشد، اطلاعات غیر لازم از نمونه درباره پارامتر مجهول حذف می‌شوند و در این حالت کامل بودن آماره بسنده منجر به بیشترین فشردگی در داده‌ها می‌شود و در حقیقت این آماره بسنده کامل از هر آماره کمکی دلخواه مستقل است. در این مقاله نشان می‌دهیم با این که ظاهراً آماره کمکی هیچ اطلاعی راجع به پارامتر نمی‌دهد، ولی وقتی توزیع توأم آن با یك براوردگر حداکثر درستنمایی در نظر گرفته شود، یك آماره بسنده مینیمال حاصل می‌شود و با این روش اطلاعات از دست رفته در آماره کمکی دوباره بازیابی می‌شوند و در نهایت این آماره کمکی در تلفیق با یك براوردگر حداکثر درستنمایی، آگاهی بخش بوده است. همچنین نشان می‌دهیم با شرطی کردن روی آماره‌های کمکی این تضمین وجود دارد که هیچ اطلاعی راجع به پارامتر با این تقلیل از دست نمی‌رود و در این حالت نیز شرطی کردن روی آماره‌های کمکی، آگاهی بخش خواهد بود.

۱. مقدمه

در تلخیص داده‌های نمونه تصادفی[۱] $\mathbf{X}$ (شامل برداری از متغیرهای تصادفی[۲] مستقل و هم‌توزیع[۳])، از تابعی از نمونه که به پارامتر[۴] مجهول بستگی ندارد استفاده می‌شود که اصطلاحاً آن را آماره[۵] گویند. در حقیقت آماره اطلاعات چند متغیر تصادفی را در یک متغیر تصادفی فشرده می‌کند. در تلخیص داده‌ها ممکن است کمی از اطلاعات مسئله مورد مطالعه از بین رود، لذا آماره‌ای که برای استنباط به کار برده می‌شود نباید باعث تلخیص نمونه اولیه به قسمی شود که دیگر برای بررسی مسئله کفایت نکند و این خاصیت منجر به تعریف بسندگی[۶] می‌شود. آماره T را برای θ بسنده[۷] می‌نامیم هرگاه به ازای هر $\theta \in \Theta$ (که Θ فضای پارامتر[۸] است) توزیع شرطی $\mathbf{X}\,|\,T = t$ برای هر t به پارامتر θ بستگی نداشته باشد[۱۰]. علت مطالعه بسندگی در استنباط آماری این است که آماره بسنده داده‌ها را به گونه‌ای خلاصه و فشرده می‌کند که اطلاعات راجع به پارامترهای جامعه از دست نرود و این خلاصه کردن در یک آماره شامل تمام اطلاعات نمونه پیرامون پارامتر باشد. اصطلاح آماره و همچنین آماره بسنده برای اولین بار در سال ۱۹۲۲ توسط فیشر معرفی شد[۷]. پس از تعریف بسندگی این سؤال مطرح می‌شود که چقدر می‌توانیم داده‌ها را خلاصه کنیم که هنوز بسنده باشند. واضح است خلاصه کردن داده‌ها تا جایی جایز است که اطلاعات نهفته در مورد پارامتر مجهول را حفظ کنند. آماره بسنده

عبارات و کلمات کلیدی. آماره کمکی، آماره بسنده مینیمال، قضیه بهادر، قضیه باسو، برآوردگر حداکثر درستنمایی.
دبیر تخصصی رابط: مجید اسدی
تاریخ دریافت: ۱۳۹۶/۰۵/۱۶   تاریخ پذیرش: ۱۳۹۸/۰۹/۱۷
DOI: http://dx.doi.org/10.22108/msci.2019.105827.1235

[1]Random Sample [2]Random Variables [3]Identically and Independently Distributed [4]Parameter [5]Statistic [6]Sufficiency [7]Sufficient [8]Parameter Space

مینیمال[9] حداکثر تلخیص را روی نمونه انجام می‌دهد. در حقیقت آماره بسنده T را بسنده مینیمال گوئیم، هرگاه تابعی از هر آماره بسنده دیگر باشد. آماره T را کامل[10] گوئیم، هرگاه برای هر تابع g، اگر به ازای هر $\theta \in \Theta$، $E_\theta\left(g(T)\right) = 0$، آنگاه به ازای هر $\theta \in \Theta$، $P_\theta\left[g(T) = 0\right] = 1$ [۱۰]. به عبارت دیگر خانواده توزیع‌های تولید شده توسط T کامل است، اگر تنها برآوردگر نااریب[11] برای صفر در این خانواده خود صفر باشد. اصطلاح کامل بودن برای اولین بار در سال ۱۹۵۰ توسط لهمن و شفه[12] معرفی شد[۱۰]. خاصیت کامل بودن یك آماره به تنهایی کاربردی ندارد، ولی وقتی این خاصیت مربوط به آماره بسنده باشد، خواص یکتایی مشخصی را تضمین می‌کند و موجب سادگی برخی از مسائل آمار و احتمال می‌شود. اصطلاح آماره کمکی[13] (فرعی، ناآگاهنده یا آگاهی نابخش[14]) برای اولین بار در سال ۱۹۲۵ توسط فیشر [۷] در زمینه تخمین حداکثر درستنمایی ارائه و معرفی شد. کلمه ancillary از واژه لاتین ancilla به معنی همراه و ملازم گرفته شده است. فیشر تابع درستنمایی را به عنوان تابعی در نظر گرفت که تمام اطلاعاتی را که داده‌ها می‌بایست در مورد پارامتر مجهول ارائه کنند در بر می‌گرفت، ولی وی یک تعریف رسمی از آماره کمکی بیان نکرد. پس از فیشر افرادی چون باسو [۳] آماره کمکی را به این صورت معرفی کردند که توزیع آن نباید به پارامترهای مدل بستگی داشته باشد. برخی افراد مانند لهمن و اسچولز [۲] علاوه بر داشتن شرط قبلی، آماره کمکی را به عنوان تابعی از آماره بسنده مینیمال در نظر می‌گیرند و با این کار رده آماره‌های کمکی را محدود می‌کنند. از این که توزیع آماره کمکی به پارامتر مجهول بستگی ندارد، اولین نکته‌ای که به ذهن می‌رسد این است که این آماره شامل هیچ اطلاعی پیرامون پارامتر نیست. در بخش ۲ نشان می‌دهیم که چگونه کامل بودن[15] آماره بسنده منجر به خارج کردن اطلاعات غیر لازم درباره پارامتر می‌شود. همچنین در بخش ۳ با تلفیق آماره کمکی با یک براوردگر حداکثر درستنمایی[16] (MLE) اطلاعات از دست رفته توسط آماره کمکی را دوباره بازیابی[17] می‌کنیم.

۲. نقش کامل بودن آماره بسنده در حذف اطلاعات غیر مفید در مورد پارامتر

آماره کمکی تابعی از آماره بسنده کامل نیست. همچنین با توجه به این که آماره بسنده کامل در فشرده کردن داده‌ها موثر است، شگفت آور نیست که آماره بسنده کامل همیشه مینیمال است (قضیه بهادر[18] [۱۰]). یك روش برای اثبات عدم وجود آماره بسنده کامل، استفاده از عکس نقیض قضیه بهادر است، زیرا اگر آماره T بسنده مینیمال نباشد، آنگاه T کامل نیست. پس اگر آماره مورد نظر مینیمال نباشد، جستجو برای یافتن آماره بسنده کامل بی‌نتیجه است، چون طبق عکس نقیض قضیه بهادر، چنین آماره‌ای وجود ندارد [۱۰].

مفهوم بسندگی ارتباط نزدیکی با مفهوم کمکی بودن دارد. اگر نمونه در هنگام شرطی شدن توسط آماره T، یک آماره کمکی شود، آماره T بسنده خواهد بود. بوس و همکاران [۵] تفاوت اساسی بین یك آماره بسنده (بسنده مینیمال) و یك آماره بسنده کامل را در این دانسته‌اند که آماره بسنده ممکن است شامل اطلاعات اضافی (اطلاعات فرعی، کمکی یا آگاهی نابخش) باشد که برای تخمین پارامتر مجهول مفید نیست که در اصطلاح گوئیم آماره بسنده هنوز به طور کامل فشرده نشده است و این حقیقت نشان می‌دهد، ممکن است یك تابع از T که کمکی است وجود داشته باشد. اما همان طور که لهمن [۹] و همچنین بوس و همکاران [۵] اشاره کرده‌اند، وجود آماره بسنده مینیمال به خودی خود، عدم وجود تابعی از T که کمکی باشد را تضمین نمی‌کند. در صورتی که با به کارگیری قضیه باسو[19][۱۰]، که بیان می‌کند آماره بسنده کامل از هر آماره کمکی مستقل است، وجود آماره بسنده کامل T تضمین می‌کند که تابعی از T که کمکی باشد وجود ندارد و این بدین معنی است که T شامل اطلاعات کمکی و غیر لازم درباره پارامتر مجهول نیست. به عبارت

[9]Minimal Sufficient Statistic [10]Complete [11]Unbiased Estimator [12]Scheffé [13]Ancillary [14]Non-Informative [15]Completeness [16]Maximum Likelihood Estimator [17]Recovery [18]Bahadur's Theorem [19]Basu's Theorem

DOI: http://dx.doi.org/10.22108/msci.2019.105827.1235

دیگر کامل بودن یك آماره بسنده، منجر به بیشترین فشردگی در داده‌ها می‌شود. نتیجه زیر خلاصه بحث بالا را جمع‌بندی می‌کند.

نتیجه ۱.۲. *هیچ آماره کمکی غیر ثابت که تابعی از آماره بسنده کامل T باشد وجود ندارد.*

اثبات. *اگر $U = f(T)$ کمکی باشد، آنگاه $E_\theta(f(T)) = E(U) = c$ که c یك مقدار ثابت قابل محاسبه است. با توجه به کامل بودن آماره T، با احتمال یك $f(T) = c$، یعنی U تابع ثابتی از T است.*

بنابراین با توجه به نتیجه ۱.۲، اگر یك آماره بسنده، کامل نیز باشد، در این صورت حاوی هیچ اطلاعات کمکی نیست. پس کامل بودن یك آماره بسنده، مزیتی برای این آماره در جداسازی قسمت مفید داده‌ها (که حاوی اطلاعات است) از قسمتی که حاوی هیچ اطلاعات مفیدی نیست (اطلاعـاتی که آماره کمکی می‌دهد)، فراهم می‌کند. بنابراین کامل بودن یك آماره بسنده به طور ذاتی منجر به بیشترین فشردگی در داده‌ها می‌شود.

تعریف ۲.۲. *آماره V را کمکی مرتبه اول[۲۰] گوییم، اگر $E_\theta(V)$ به $\theta \in \Theta$ بستگی نداشته باشد*[۱۱].

بدیهی است که اگر V یك آماره کمکی باشد، آنگاه $E_\theta(V)$ مقداری ثابت است و به θ بستگی ندارد و بنابراین V آماره کمکی مرتبه اول نیز هست، بنابراین نتیجه ۱.۲ برای آماره کمکی مرتبه اول هم صدق می‌کند ولی عکس نتیجه ۱.۲ فقط برای آماره کمکی مرتبه اول برقرار است. به عبارت دیگر اگر بتوان تابعی از آماره بسنده T ساخت که آماره کمکی و لذا آماره کمکی مرتبه اول باشد، لزومی ندارد T به طورکامل فشرده باشد (نتیجه ۱.۲)، زیرا اطلاعاتی در T وجود دارد که درباره θ نیست. برعکس یك آماره بسنده T در فشردگی داده‌ها تواناست، اگر هیچ تابع غیر ثابتی از T یافت نشود که آماره کمکی مرتبه اول باشد. بنابراین نتیجه زیر را خواهیم داشت:

نتیجه ۳.۲. *اگر هیچ تابع غیر ثابتی از آماره بسنده T که آماره کمکی مرتبه اول باشد وجود نداشته باشد، آنگاه T کامل است.*

در ادامه نتیجه ۱.۲ که نشان دادیم در صورت وجود آماره بسنده کامل T، هیچ آماره کمکی (غیر از ثابت‌ها) از T به دست نمی‌آید، می‌توانیم با ارائه تعریف آماره کمکی مشروط نتیجه مشابهی به دست آوریم:

تعریف ۴.۲. *آماره U را کمکی مشروط [۲۱] به شرط V گوییم، هرگاه توزیع $U \mid V = v$ برای هر v به پارامتر مجهول بستگی نداشته باشد.*

در حالت خاص اگر $\mathbf{X}$ کمکی مشروط به شرط $V = v$ باشد (برای هر v)، آنگاه V بسنده است و اگر U کمکی مشروط به شرط متغیر تصادفی ثابت V باشد، آنگاه U کمکی است.

نتیجه ۵.۲. *اگر T کامل باشد، آماره کمکی مشروط غیر بدیهی به شرط T وجود ندارد.*

اثبات. *فرض می‌کنیم آماره کمکی مشروط U به شرط T وجود داشته باشد، حال اگر A پیشامدی وابسته به U باشد، آنگاه داریم:*

$$\begin{aligned} E\left(P\left(A\,|T\right) - I_A(T)\right) &= E_\theta\left(E(I_A|T)\right) - P_\theta(A) \\ &= E_\theta(I_A) - P_\theta(A) \\ &= 0 \end{aligned}$$

[20]First-Order Ancillary [21]Conditionally Ancillary

DOI: http://dx.doi.org/10.22108/msci.2019.105827.1235

که با توجه به کامل بودن T، نتیجه می‌گیریم، با احتمال یك $P(A \mid T) = I_A(T)$ و این یعنی U با احتمال یك ثابت است.

مثال ۶.۲. اگر $X_1, \dots, X_n$، n متغیر تصادفی مستقل و هم‌توزیع با تابع چگالی احتمال توأم متعلق به خانواده مکان[22]

$$\mathcal{P} = \left\{ \prod_{i=1}^{n} f_\theta(x_i) \; : f_\theta(x) = f(x-\theta); x, \theta \in \mathbb{R} \right\}$$

باشند، دو حالت مجزای زیر را بررسی می‌کنیم:

الف) f متعلق به یکی از خانواده‌های کوشی[23]، لجستیك[24] یا نمایی دوگانه[25] باشد.

ب) f متعلق به یکی از خانواده‌های نرمال[26]، نمایی دم بریده[27] یا یکنواخت دو پارامتری[28] باشد.

حالت الف)

۱ـ فرض می‌کنیم چگالی f متعلق به خانواده کوشی $C(0,1)$ یعنی $f(x) = 1/\pi(1+x^2)$، $x \in \mathbb{R}$ باشد. آماره‌های مرتب[29] $\mathbf{T} = (X_{(1)}, \dots, X_{(n)})$ برای خانواده $\mathcal{P}$ بسنده مینیمال است، زیرا نسبت

$$\begin{aligned}\frac{f_\theta(\mathbf{x})}{f_\theta(\mathbf{y})} &= \prod_{i=1}^{n} \frac{f(x_i-\theta)}{f(y_i-\theta)} \\ &= \prod_{i=1}^{n} \frac{f(x_{(i)}-\theta)}{f(y_{(i)}-\theta)} \\ &= \prod_{i=1}^{n} \frac{1+(y_{(i)}-\theta)^2}{1+(x_{(i)}-\theta)^2}\end{aligned}$$

به θ بستگی ندارد، اگر و تنها اگر $(y_{(1)}, \dots, y_{(n)}) = (x_{(1)}, \dots, x_{(n)})$. همچنین $\mathbf{T}$ کامل نیست، زیرا تابع چگالی توأم

$$f_{\mathbf{T}}(\mathbf{y}) = n! \prod_{i=1}^{n} \frac{1}{\pi\left(1+(x_{(i)}-\theta)^2\right)} \; ; \; x_{(1)} \leq \dots \leq x_{(n)}$$

که در آن $\mathbf{y} = (x_{(1)}, \dots, x_{(n)})$ تابعی زوج بر حسب $x_{(i)} - \theta$ است، پس با انتخاب تابع فرد غیر صفر h با ضابطه $h(\mathbf{y}) = x_{(n)} - x_{(1)}$، اگر برای هر $\theta \in \mathbb{R}$، $E_\theta(h(\mathbf{T})) = 0$ در این صورت داریم:

$$\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} h(\mathbf{y}) f_{\mathbf{T}}(\mathbf{y}) \, dx_{(1)} \cdots dx_{(n)} = 0$$

در صورتی که

$$P_\theta(h(\mathbf{T}) = 0) = P_\theta\left(X_{(n)} - X_{(1)} = 0\right) \neq 1.$$

---

[22]Location Family [23]Cauchy [24]Logistic [25]Double Exponential [26]Normal [27]Truncated Exponential [28]Two-Parameter Uniform
[29]Order Statistics

DOI: http://dx.doi.org/10.22108/msci.2019.105827.1235

۲ــ فرض می‌کنیم چگالی f از خانواده لجستیك $Lo(\circ, ۱)$ یعنی $f(x) = e^x(۱+e^x)^{-۲}$، $x \in \mathbb{R}$ باشد. مشابه قسمت (۱) مشاهده می‌شود که $\mathbf{T}$ برای θ بسنده مینیمال است. همچنین $\mathbf{T}$ کامل نیست، زیرا برای هر $\theta \in \mathbb{R}$،

$$\begin{aligned} E_\theta(X_{(n)} - X_{(۱)}) &= E_\theta(X_{(n)} - \theta) - E_\theta(X_{(۱)} - \theta) \\ &= E_\circ(X_{(n)}) - E_\circ(X_{(۱)}) \\ &= c \end{aligned}$$

در صورتی که

$$P_\theta(X_{(n)} - X_{(۱)} - c \neq \circ) = ۱.$$

همان طور که در این حالت می‌بینیم، آماره کمکی $X_{(n)} - X_{(۱)}$ تابعی از آماره بسنده مینیمال $\mathbf{T}$ هست و این یعنی $X_{(n)} - X_{(۱)}$ و $\mathbf{T}$ از هم مستقل نیستند. بنابراین ادعای لهمن [۹] و بوس و همکاران [۵] مبنی بر این که وجود آماره بسنده مینیمال $\mathbf{T}$ به خودی خود، عدم وجود تابعی از $\mathbf{T}$ که کمکی باشد را تضمین نمی‌کند، تأیید می‌شود.

۳ــ فرض می‌کنیم چگالی f از خانواده نمایی دوگانه (لاپلاس[۳۰]) $L(\circ, ۱, ۱)$ یعنی $f(x) = \frac{۱}{۲}e^{-|x|}$، $x \in \mathbb{R}$ باشد. به طور مشابه $\mathbf{T}$ آماره بسنده مینیمال برای $\mathcal{P}$ است و $\mathbf{U} = (X_{(n)} - X_{(۱)}, \ldots, X_{(n)} - X_{(n-۱)})$ کمکی و تابعی از آماره بسنده $\mathbf{T}$ است، پس طبق عکس نقیض قضیه باسو، $\mathbf{T}$ بسنده کامل نیست.

حالت ب) فرض می‌کنیم چگالی f متعلق به یکی از خانواده‌های نرمال $N(\theta, ۱)$، نمایی دم بریده $E(\lambda, ۱)$ یا یکنواخت دو پارامتری $U(\theta_۱, \theta_۲)$ با توابع چگالی به ترتیب

$$\begin{aligned} f(x) &= \tfrac{۱}{\sqrt{۲\pi}} \exp\left\{-\tfrac{۱}{۲}(x-\theta)^۲\right\}; && x \in \mathbb{R} \\ f(x) &= \exp\{-(x-\lambda)\}; && x > \lambda \\ f(x) &= \tfrac{۱}{\theta_۲ - \theta_۱}; && \theta_۱ < x < \theta_۲ \end{aligned}$$

باشد. در این حالت به ترتیب $\bar{X}$، $X_{(۱)}$ یا $(X_{(۱)}, X_{(n)})$ آماره‌های بسنده کامل هستند. در حالی که در حالت (الف)، $\mathbf{T} = (X_{(۱)}, \ldots, X_{(n)})$ آماره بسنده مینیمال بود. اطلاعات نهفته در آماره کمکی دلیل اصلی این اختلاف بین خانواده چگالی‌های ذکر شده در قسمت الف و ب است. برای روشن شدن این تفاوت با توجه به این که در توزیع‌های با پارامتر مکان، تفاضل‌های $Y_i = X_{(n)} - X_{(i)}$ برای $i = ۱, \ldots, n-۱$ یك آماره کمکی $n-۱$ بعدی تشکیل می‌دهد، در حالت نرمال و نمایی دم بریده به دلیل این که آماره بسنده مینیمال، کامل است پس طبق قضیه باسو، آماره‌های کمکی Y_i از آماره بسنده مینیمال مستقل‌اند. بنابراین در این دو حالت Y_iها و یا ترکیب آن‌ها با آماره بسنده مینیمال، هیچ اطلاعاتی درباره θ نمی‌دهند. ولی هنگامی که n آماره ترتیبی، مینیمال هستند، آماره‌های کمکی Y_i تابعی از آماره بسنده مینیمال هستند و در این حالت بسندگی در فشردن آماره‌های کمکی موفق نیست و همچنین Y_iها به خودی خود هیچ اطلاعی را ارائه نمی‌دهند، اما در بخش ۳ مشاهده می‌کنیم که در ترکیب با باقیمانده داده‌ها می‌توانند حاوی اطلاعات شوند.

در حالتی که f متعلق به خانواده یکنواخت دو پارامتری $U(\theta_۱, \theta_۲)$ باشد، Y_iها مستقل از آماره بسنده مینیمال نیستند، زیرا تابعی از آماره بسنده مینیمال $(X_{(۱)}, X_{(n)})$ هستند، اما نسبت‌های $Z_i = Y_i / Y_۱$ $(i = ۱, \ldots, n-۱)$ این خاصیت را دارند. حال اگر نمونه مرتب شده را به صورت $(Y_\circ, Y_۱, Z_۲, Z_۳, \ldots, Z_{n-۱})$ نمایش دهیم که در آن $Y_\circ = \frac{X_{(۱)} + X_{(n)}}{۲}$ و $Y_۱ = X_{(n)} - X_{(۱)}$، مشاهده می‌کنیم که $X_{(۱)}$ و $X_{(n)}$ را می‌توان فقط با استفاده از $Y_\circ$ و $Y_۱$ به دست آورد و $Z_۲$ تا $Z_{n-۱}$ در به دست آوردن $X_{(۱)}$ و $X_{(n)}$ نقشی ندارند. به عبارت دیگر همان طور که قبلاً اشاره کردیم، آماره بسنده کامل $\mathbf{T} = (X_{(۱)}, X_{(n)})$ قسمت مفید داده‌ها یعنی $Y_\circ$ و $Y_۱$ را از قسمتی که به خودی خود و

[30] Laplace

DOI: http://dx.doi.org/10.22108/msci.2019.105827.1235

یا ترکیب با بقیه داده‌ها هیچ اطلاعی را ارائه نمی‌دهند (یعنی $Z_۲$ تا $Z_{n-۱}$)، جدا می‌سازد و لذا $\mathbf{T}$ حاوی هیچ اطلاع کمکی نیست. در این جا آماره $Y_۱$ کمکی نیست (اصطلاحاً گوییم آگاهنده یا آگاهی بخش[31] است)، زیرا اگر باشد، آن‌گاه برای هر $\theta \in \mathbb{R}$، $E_\theta(Y_۱) = c$ درصورتی که برای هر $\theta \in \mathbb{R}$، $P_\theta(Y_۱ = c) \neq ۱$ که با کامل بودن $\mathbf{T}$ تناقض دارد.

در حالتی که f متعلق به خانواده یکنواخت یك پارامتری مثل $U(\theta, \theta+۱)$ با تابع چگالی $f(x) = ۱$، $\theta < x < \theta+۱$ باشد، $\mathbf{T}$ آماره بسنده مینیمال است و کامل نیست، زیرا برای هر $\theta \in \mathbb{R}$، $E_\theta(Y_۱) = (n-۱)(n+۱)^{-۱}$ ولی

$$P_\theta\left(Y_۱ = (n-۱)(n+۱)^{-۱}\right) \neq ۱.$$

لذا در حالتی که f متعلق به خانواده یکنواخت دو پارامتری باشد، $\mathbf{T}$ در پی نیاز ساختن اطلاع کمکی نسبت به حالت یك پارامتری موفق‌تر است که این خاصیت فقط به خاطر کامل بودن $\mathbf{T}$ در حالت دو پارامتری است [۱۱].

همگی توزیع‌ها در مثال ۶.۲ متعلق به خانواده توزیع‌های با پارامتر مکان هستند که آن‌ها را در دو گروه زیر مورد بررسی قرار دادیم:

گروه اول، شامل توزیع یکنواخت دو پارامتری است و همچنین توزیع‌هایی که آماره بسنده مینیمال یك بعدی دارند مانند نرمال و نمایی دم بریده که در این خانواده‌ها آماره بسنده مینیمال، کامل نیز هست، زیرا بعد آماره بسنده با بعد فضای پارامتر یکسان است و لذا از آماره کمکی $Y_i = X_{(n)} - X_{(i)}$ مستقل است که در این حالت، کامل بودن در کنار گذاشتن آماره کمکی موفق بوده است.

گروه دوم، شامل توزیع یکنواخت یك پارامتری و همچنین توزیع‌هایی است که برای آن‌ها آماره بسنده مینیمال آماره‌های ترتیبی حاصل از n مشاهده است مانند کوشی، لجستیك و نمایی دوگانه، که در این خانواده‌ها آماره بسنده مینیمال، کامل نیست و از آماره کمکی $Y_i = X_{(n)} - X_{(i)}$ مستقل نیست، زیرا Y_i تابعی از آماره بسنده مینیمال است. بنابراین در این حالت بسندگی موفقیتی در کنار گذاشتن آماره کمکی نداشته است.

مثال ۷.۲. *اگر خانواده تمام توزیع‌های پیوسته*

$$\mathcal{P} = \left\{\prod_{i=۱}^{n} f(x_i) : f \text{ پیوسته}\right\}$$

را در نظر بگیریم، آن‌گاه $\mathbf{T} = (X_{(۱)}, .., X_{(n)})$ کامل است[۱۰]. لذا طبق قضیه باسو، هیچ تابع غیر ثابتی از $\mathbf{T}$ که کمکی باشد وجود ندارد. بنابراین همان طور که لهمن [۹] اشاره کرده، برای خانواده تمام توزیع‌های پیوسته بر خلاف خانواده‌های کوشی، لجستیك یا نمایی دوگانه، بسندگی در فشردن همه آماره‌های کمکی موفق بوده است [۸].

با توجه به دو مثال بالا می‌توان نتیجه گرفت که تنها در مدل‌هایی که آماره بسنده مینیمال، کامل نیز باشد، این آماره به طور کامل در کنار گذاشتن آماره‌های کمکی موفق عمل می‌کند و قضیه باسو نیز بر همین موضوع تأکید دارد. زیرا همان طور که قضیه باسو و همچنین مثال‌های اخیر نشان می‌دهد، استقلال[32] آماره کمکی از آماره بسنده مینیمال به دلیل ویژگی‌های این دو آماره نیست، بلکه معلول خاصیت آماری دیگری است که آن خاصیت کامل بودن آماره بسنده مینیمال است.

۳. تلفیق براوردگر حداکثر درستنمایی با آماره کمکی برای بازیابی اطلاعات مفقود شده در داده‌ها

آماره‌های کمکی به تنهایی هیچ اطلاعی راجع به پارامتر ندارند، این بدین مفهوم است که با تقلیل داده‌ها به برآوردگر حداکثر درستنمایی، اطلاعاتی از نمونه مفقود می‌شوند. فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال یا

[31] Informative [32] Independence

تابع جرم احتمال $f_\theta(x)$ باشد و M یك MLE برای θ باشد که دارای تابع چگالی احتمال یا تابع جرم احتمال $g_\theta(m)$ است. با برقراری شرایط نظم، اطلاع فیشر[۳۳] راجع به θ مربوط به متغیرهای تصادفی X و M به ترتیب به صورت $I(\theta) = E_\theta(\frac{-\partial^2}{\partial\theta^2}\log f_\theta(X))$ و $J(\theta) = E_\theta(\frac{-\partial^2}{\partial\theta^2}\log g_\theta(M))$ تعریف می‌شوند. فرض می‌کنیم شرایط نظم برای مسئله مورد نظر وجود داشته باشد. به راحتی نشان داده می‌شود $I(\theta) \geq J(\theta)$ و شرط لازم و کافی برای برقراری تساوی این است که M بسنده باشد. فیشر [۱۰]، $\lambda(\theta) = I(\theta) - J(\theta)$ را به عنوان یك معیار منطقی برای اندازه‌گیری اطلاع موجود در آماره کمکی U پیشنهاد می‌کند که آن را اطلاع کمکی[۳۴] می‌نامیم. در واقع $\lambda(\theta)$ میزان اطلاع از دست رفته در استفاده از آماره M به عنوان برآوردگر θ است.

در بسیاری از کتاب‌های آماری ادعا می‌شود که یک MLE برای یک پارامتر نامعلوم θ باید تابعی از هر آماره بسنده باشد. این ادعا ناصحیح است. بلکه اگر T یک آماره بسنده برای θ باشد و MLE یکتایی مثل $\hat{\theta} = M$ برای θ وجود داشته باشد، آنگاه $\hat{\theta}$ باید تابعی از T باشد و اگر MLE منحصر به فردی وجود نداشته باشد، آنگاه MLE پارامتر θ را می‌توان طوری انتخاب کرد که تابعی از T باشد. پس صحیح‌تر است بگوییم که اگر MLE یکتا باشد، آنگاه تابعی از آماره بسنده مینیمال است، اما لزومی ندارد MLE به خودی خود بسنده باشد. در این حالت اغلب یك MLE مثل M در تلفیق با یك آماره کمکی مثل U، بسنده است [۸]. یعنی همان‌طور که قبلاً اشاره کرده بودیم، آماره‌های کمکی به خودی خود هیچ اطلاعی را ارائه نمی‌دهند و آگاهی بخش نیستند، ولی در ترکیب با بقیه داده‌ها می‌توانند اطلاعات کاملی راجع به پارامتر داشته باشند. به عبارت دیگر اگر MLE، بسنده باشد، هیچ اطلاعی از دست نمی‌رود، ولی در غیر این صورت باید به فکر بازیابی اطلاعات از دست رفته باشیم. فیشر میزان اطلاع از دست رفته در استفاده از آماره M به عنوان برآوردگر θ را با ملاك $\lambda(\theta) = I_X(\theta) - I_M(\theta)$ نشان می‌دهد. در حالتی که MLE، بسنده باشد، نتیجه می‌گیریم که $\lambda(\theta) = 0$، به این معنی که اگر MLE را به عنوان برآورد کننده θ به کار گیریم، هیچ اطلاعی از دست نمی‌رود [۳]. برای این منظور فیشر معتقد است که یك آماره کمکی در تلفیق با یك MLE، یك آماره بسنده مینیمال تشکیل می‌دهد که با این کار اطلاعات از دست رفته دوباره بازیابی می‌شوند [۱]. اساس پی جویی اطلاعات از دست رفته توسط فیشر پی ریزی شد. از دیدگاه فیشر در چنین شرایطی، اگر آماره کمکی U وجود داشته باشد که آماره توأم $(M,\ U)$ برای θ بسنده باشد، در این صورت برای بازیابی اطلاعات و کامل شدن استنباط به جای توزیع M از توزیع شرطی $M|U$ استفاده می‌کنیم [۳].

برای روشن شدن مطلب مثال‌های زیر می‌توانند مفید باشند.

مثال ۱.۳. *(مسئله نیل[۳۵]) نمونه تصادفی $(X_1, Y_1), \ldots, (X_n, Y_n)$ با تابع چگالی*

$$f_\theta(x, y) = e^{-(\theta x + y/\theta)} I_{\mathbb{R}^+ \times \mathbb{R}^+}(x, y)$$

و فضای پارامتر $\Theta = (0, \infty)$ را در نظر می‌گیریم. می‌توان نشان داد $M = \sqrt{\sum_i Y_i / \sum_i X_i}$ یك MLE برای θ است که بسنده نیست. پس اگر برای یك مسئله استنباطی از M استفاده کنیم، برخی اطلاعات مفقود می‌شوند. همچنین به سادگی می‌توان نشان داد که آماره $U = \sqrt{\sum_i X_i \sum_i Y_i}$ کمکی است. اگر ترکیب M با آماره کمکی U را در نظر بگیریم، با این کار $(M,\ U)$ یك آماره بسنده توأم برای θ با تابع چگالی زیر است:

$$f(m, u) = \begin{cases} \frac{2}{(\Gamma(n))^2} u^{2n-1} \frac{1}{m} e^{-u\left(\frac{\theta}{m} + \frac{m}{\theta}\right)} & m, u > 0 \\ 0 & \text{سایر جاها} \end{cases}$$

[33] Fisher Information [34] Ancillary Information [35] Nail Problem

DOI: http://dx.doi.org/10.22108/msci.2019.105827.1235

اما از بسط دوجمله‌ای[36]

$$(\frac{\theta}{m} + \frac{m}{\theta})^n = \sum_{k=0}^{n} \binom{n}{k} (\frac{m}{\theta})^{-n+2k},$$

و به کمک بسط مک‌لورن[37] تابع نمایی، داریم:

$$e^{-u(\frac{\theta}{m}+\frac{m}{\theta})} = \sum_{n=0}^{\infty} \frac{(-1)^n u^n}{n!} \sum_{k=0}^{n} \binom{n}{k} (\frac{m}{\theta})^{-n+2k}.$$

بنابراین

$$\begin{aligned} f_U(u) &= \frac{2}{(\Gamma(n))^2} u^{2n-1} \int_0^\infty \frac{1}{m} \sum_{n=0}^{\infty} \frac{(-1)^n u^n}{n!} \sum_{k=0}^{n} \binom{n}{k} (\frac{m}{\theta})^{-n+2k} dm \\ &= \frac{4}{(\Gamma(n))^2} u^{2n-1} \sum_{n=0}^{\infty} \frac{(-1)^n u^n}{(n!)^2} \\ &= \frac{4}{(\Gamma(n))^2} u^{2n-1} K_0(2u) \end{aligned}$$

که در آن

$$K_0(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(n!)^2 2^{2n}}$$

تابع بسل[38] است. بنابراین توزیع آماره U به پارامتر بستگی ندارد و لذا کمکی است. در این حالت $I(\theta) = 2n/\theta^2$. به طور مشابه تابع چگالی احتمال M برابر است با:

$$f_M(m) = \frac{2\Gamma(2n)}{(\Gamma(n))^2} m (\frac{\theta}{m} + \frac{m}{\theta})^{-2n}; m, \theta > 0.$$

از طرفی

$$\begin{aligned} (\frac{\theta}{M} + \frac{M}{\theta})^{-2} &= \int_0^\infty \frac{2\Gamma(2n)}{m(\Gamma(n))^2} (\frac{\theta}{m} + \frac{m}{\theta})^{-2(n+1)} dm \\ &= \frac{2(\Gamma(n+1))^2 \Gamma(2n)}{2(\Gamma(n))^2 \Gamma(2n+2)} \\ &= \frac{n}{2(2n+1)}. \end{aligned}$$

لذا میزان اطلاع موجود در M برابر است با:

$$\begin{aligned} J(\theta) &= \left(\frac{\partial \log f_\theta(M)}{\partial \theta}\right)^2 \\ &= E\left(\frac{4n^2}{\theta^2} \frac{(\frac{M}{\theta} - \frac{\theta}{M})^2}{(\frac{\theta}{M} + \frac{M}{\theta})^2}\right) \\ &= \frac{4n^2}{\theta^2} E\left(\frac{(\frac{M}{\theta} + \frac{\theta}{M})^2 - 4}{(\frac{\theta}{M} + \frac{M}{\theta})^2}\right) \\ &= \frac{4n^2}{\theta^2} \left(1 - 4E(\frac{\theta}{M} + \frac{M}{\theta})^{-2}\right) \\ &= \frac{2n}{\theta^2} \frac{2n}{2n+1}. \end{aligned}$$

[36]Binomial Expansion [37]Maclaurin Expansion [38]Bessel Function

بنابراین زیان اطلاع برابر با

$$\lambda(\theta) = I(\theta) - J(\theta) = \frac{۲n}{(۲n+۱)\theta^۲}$$

خواهد بود. با توجه به دیدگاه فیشر نباید اطلاع در M را با $J(\theta)$ اندازه گرفت، بلکه به جای آن از

$$\begin{aligned} J(\theta\,|U) &= -\left(\tfrac{\partial^۲}{\partial\theta^۲}\log f(M|U,\theta)|\theta,U\right) \\ &= -\left(\tfrac{\partial^۲}{\partial\theta^۲}\log\left[\tfrac{۱}{۲K_{\circ}(۲U)}e^{-U(\frac{M}{\theta}+\frac{\theta}{M})}\tfrac{۱}{M}\right]\right) \\ &= -E(\tfrac{\partial^۲}{\partial\theta^۲}[-\log M - \log ۲K.(۲U) - U(\tfrac{M}{\theta}+\tfrac{\theta}{M})]) \\ &= \tfrac{۲n}{\theta^۲}\tfrac{K_۱(۲U)}{K_{\circ}(۲U)} \end{aligned}$$

استفاده می‌کنیم که در آن $K_{\circ}$ و $K_۱$ توابع بسل هستند و

$$K_۱(x) = \sum\nolimits_{n=\circ}^{\infty}\tfrac{(-۱)^n x^{۲n+۱}}{n!(n+۱)!۲^{۲n+۱}}.$$

بنابراین داریم

$$E\left(J(\theta\,|U)\right) = \frac{۲n}{\theta^۲} = I(\theta)$$

و استنباط براساس $f(M|U,\theta)$ منطقی‌تر به نظر می‌رسد، زیرا اطلاعی که براساس توزیع شرطی M به شرط آماره کمکی U محاسبه می‌شود، تمام اطلاعات موجود در نمونه را به ما می‌دهد و هیچ اطلاعی از دست نمی‌رود و این مطلب نشان می‌دهد که مقداری اطلاع در U موجود است. جالب است که اگر $f(T|U,\theta)$ را برای برآورد θ به کار بگیریم، باز هم برآوردگر درستنمایی θ برابر با M خواهد بود، با این تفاوت که وقتی آماره کمکی را نیز در استنباط دخالت می‌دهیم هیچ اطلاعی از دست نمی‌رود. بنابراین اطلاعات از دست رفته از طریق آماره کمکی بازیابی شده و استنباط کامل می‌شود [۳].

مثال ۲.۳. فرض می‌کنیم متغیر تصادفی X دارای تابع احتمال زیر و فضای پارامتر $\Theta = [\circ, ۱]$ باشد:

$$P_\theta(X=i) = \begin{cases} \frac{i-\theta}{۱۲} & i = ۱,۲,۳ \\ \frac{i+\theta}{۱۲} & i = ۴,۵,۶. \end{cases}$$

به راحتی می‌توان نشان داد که MLE پارامتر θ به صورت زیر است:

$$M \equiv M(X) = \begin{cases} \circ & X = ۱,۲,۳ \\ ۱ & X = ۴,۵,۶ \end{cases}$$

که دو افراز $\{۴,۵,۶\}$ و $\{۱,۲,۳\}$ را روی فضای نمونه بوجود می آورد. آماره M بسنده مینیمال نیست، زیرا مثلاً

$$P_\theta(X=۱\,|\,M=\circ\,) = \begin{cases} \frac{P_\theta(X=۱)}{P_\theta(M=\circ)} = \frac{۱-\theta}{۶-۳\theta}, & M=\circ \\ \circ & M \neq \circ \end{cases}$$

به θ بستگی دارد. از طرف دیگر آماره U که به صورت زیر تعریف می‌شود کمکی است:

$$U(X) = \begin{cases} \circ & X = ۱,۴ \\ ۱ & X = ۲,۵ \\ ۲ & X = ۳,۶ \end{cases}$$

DOI: http://dx.doi.org/10.22108/msci.2019.105827.1235

که افراز $\{۱,۴\}$، $\{۲,۵\}$ و $\{۳,۶\}$ را روی فضای نمونه بوجود می‌آورد.

مشاهده می‌شود که هر U–افراز، هر M–افراز را در یک مجموعه تک عضوی قطع می‌کند و با پیدا کردن تابع احتمال توأم (M,U) (جدول۱)، دیده می‌شود که (M,U) و X معادل‌اند، زیرا هر دو یک افراز روی فضای نمونه ایجاد می‌کنند. لذا چون X بسنده مینیمال است، (M,U) نیز چنین است.

U / M	۰	۱	۲	
۰	$\frac{۱-\theta}{۱۲}$	$\frac{۲-\theta}{۱۲}$	$\frac{۳-\theta}{۱۲}$	$\frac{۲-\theta}{۴}$
۱	$\frac{۱+\theta}{۱۲}$	$\frac{۲+\theta}{۱۲}$	$\frac{۳+\theta}{۱۲}$	$\frac{۲+\theta}{۴}$
	$\frac{۱}{۶}$	$\frac{۱}{۳}$	$\frac{۱}{۲}$	۱

جدول ۱: تابع احتمال توأم (M,U)

اکنون با توجه به دیدگاه فیشر که استنباط باید بر پایه توزیع شرطی $M\,|U$ باشد، برای استنباط تابع احتمال $M\,|U$ را در نظر می‌گیریم (جدول۲).

U / M	۰	۱	۲
۰	$\frac{۱-\theta}{۲}$	$\frac{۲-\theta}{۴}$	$\frac{۳-\theta}{۶}$
۱	$\frac{۱+\theta}{۲}$	$\frac{۲+\theta}{۴}$	$\frac{۳+\theta}{۶}$

جدول ۲: تابع احتمال شرطی $M \mid U$

در این مثال هم می‌بینیم که چون MLE، بسنده نیست، نمی‌تواند حاوی تمام اطلاعات جامعه باشد، ولی وقتی با آماره کمکی U تلفیق می‌شود، (M,U) شامل همان اطلاعاتی است که X به ما می‌دهد [۳].

بنابراین زمانی که یک MLE مثل T خود بسنده نباشد، استفاده از آن موجب از دست رفتن اطلاعاتی به اندازه $\lambda(\theta)$ می‌شود که این اطلاعات می‌تواند توسط شرطی کردن روی آماره کمکی U بازیابی شوند. به عبارت دیگر اگر آماره کمکی U وجود داشته باشد به طوری که زوج (T,U) برای θ بسنده باشد، کافی است توزیع شرطی T به شرط آماره کمکی U را در نظر بگیریم تا هیچ اطلاعی از دست نرود. برای این منظور اگر اطلاع فیشر راجع به θ حاوی در $h_\theta(T|U)$ به صورت زیر تعریف شود:

$$J(\theta\,|U) = E_\theta\left[-\{\tfrac{\partial^۲}{\partial\theta^۲}\log h_\theta(T\,|U)\}\,|U\right],$$

در این صورت $I(\theta) = E_\theta[J(\theta\,|U)]$. در این تعریف فرض می‌کنیم (T,U) بسنده است و در این صورت U یک آماره کمکی است که مکمل T است. با توجه به دیدگاه فیشر اندازه مناسب برای اطلاع در T برابر $J(\theta\,|U)$ هست و $J(\theta)$ معیار مناسبی برای این منظور نیست.

DOI: http://dx.doi.org/10.22108/msci.2019.105827.1235

مثال ۳.۳. *فرض می‌کنیم متغیر تصادفی X دارای توزیع $U[\theta, \theta + ۱)$ با تابع چگالی $f(x) = ۱$ ، $\theta < x < \theta + ۱$ باشد که در آن $\theta \geq ۰$. به راحتی دیده می‌شود که تابع درستنمایی برای مشاهده X به صورت زیر است:*

$$L(\theta) = \begin{cases} ۱ & X - ۱ < \theta \leq X \\ ۰ & \text{سایر جاها} \end{cases}$$

لذا $M = [X]$ یك MLE برای θ است که بسنده نیست. قسمت اعشاری X یعنی تابع $\phi(X) = X - [X]$ دارای توزیع $U[۰, ۱)$ است و لذا $\phi(X)$ یك آماره کمکی است. از طرف دیگر با توجه به تساوی $X = [X] + \phi(X)$ جفت $([X], \phi(X))$ با X معادل است و لذا چون X بسنده مینیمال است، $([X], \phi(X))$ نیز چنین است. در این مثال هم مشاهده می‌کنیم که آماره کمکی در تلفیق با MLE می‌تواند برای اطلاع دادن راجع به نمونه مفید باشد [۳].

مثال ۴.۳. *N آزمایش برنولی[۳۹] را تکرار می‌کنیم که در آن N برابر ۵ یا ۱۰۰ است و مقدار N با پرتاب یك سکه سالم به دست می‌آید. در این مثال نمونه $\mathbf{X} = (Y, N)$ است که Y تعداد موفقیت‌ها در N آزمایش است. برآورد ML پارامتر θ عبارت است از $\hat{\theta} = M = Y/N$ و همچنین آماره N کمکی است. اکنون می‌خواهیم راجع به داده $\mathbf{x} = (۳, ۵)$ تحلیل کنیم. به وضوح این داده از لحاظ کیفیت از داده $(۶۰, ۱۰۰)$ خیلی متفاوت است، با این که برآورد ML هر دو برابر $\frac{۳}{۵}$ است. چون MLE در این حالت بسنده نیست، باز در این جا استدلال فیشر مبنی بر این که استنباط باید بر پایه مدل شرطی $P(.|N, \theta)$ باشد تا MLE بسنده باشد، معنی پیدا می‌کند. در این مثال هم می‌بینیم که MLE در ترکیب با آماره کمکی N بسنده است، یعنی به عبارت دیگر (Y, N) یک آماره بسنده است.*

در ادامه اگر مقدار N برحسب این که اولین آزمایش موفقیت یا شکست باشد، برابر ۵ یا ۱۰۰ تعیین شود، در این حالت N دیگر آماره کمکی نیست. تابع درستنمایی تولید شده به وسیله دنباله مشاهدات از N شکست و موفقیت برابر $\theta^y(۱ - \theta)^{N-y}$ است که y تعداد موفقیت‌ها در کل آزمایش‌ها است. لذا (Y, N) آماره بسنده مینیمال و $\hat{\theta} = M = Y/N$ برآوردگر ML برای θ است. ارزیابی برآورد ML برحسب مدل اصلی (غیر شرطی) $P_{\hat{\theta}|\theta}(.)$ مستلزم از دست دادن مقدار قابل توجهی از اطلاعات است که بازیابی آن‌ها مقدور نیست، زیرا برای تکمیل MLE هیچ آماره کمکی وجود ندارد [۴].

مثال ۵.۳. *اگر نمونه تصادفی $X_۱, \ldots, X_n$ داری توزیع $N(\theta, ۱)$ با تابع چگالی $f(x) = \frac{۱}{\sqrt{۲\pi}} \exp\left\{-\frac{۱}{۲}(x - \theta)^۲\right\}$ ، $x \in \mathbb{R}$ باشد، در این حالت $T = \overline{X}$ یک آماره بسنده کامل و MLE است و $\mathbf{U} = (X_۱ - X_۲, .., X_۱ - X_n)$ آماره کمکی است. برخلاف مثال‌های قبل MLE بسنده کامل است. حال می‌خواهیم ببینیم در این حالت چه استنباطی می‌توان داشت؟*

زوج $(T, \mathbf{U})$ نیز در این جا اطلاعات راجع به کل نمونه را به ما می‌دهد، ولی چون T بسنده کامل است، خود به تنهایی تمام اطلاعات را در بر دارد و نیازی نیست که با یك آماره کمکی ترکیب شود. پس در صورتی که MLE یك آماره بسنده مینیمال نبود، می‌توانستیم برای اطلاع یافتن در مورد کل نمونه از ترکیب آماره کمکی $\mathbf{U}$ با یك آماره دیگر (اغلب یك MLE) استفاده کنیم، ولی در این جا که MLE یك آماره بسنده مینیمال نیز هست، هیچ اطلاعی از دست نمی‌رود که بخواهیم آن را دوباره بازیابی کنیم [۲].

از دیدگاه بیزی[۴۰]، این که یك آماره کمکی به تنهایی امکان ندارد راجع به پارامتر به ما اطلاعاتی بدهد یك امر بدیهی است. زیرا اگر U یك آماره کمکی و q چگالی پیشین[۴۱] برای متغیر تصادفی Θ باشد، q چگالی پسین[۴۲] $\Theta \,|\, U = u$ نیز

[39] Bernoulli Experiment [40] Bayesian Perspective [41] Prior Density [42] Posterior Density

DOI: http://dx.doi.org/10.22108/msci.2019.105827.1235

خواهد بود، زیرا با توجه به کمکی بودن U، $P_{U|\,\Theta=\theta}(u)$ به θ بستگی ندارد و بنابراین برای هر θ و u داریم:

$$\begin{aligned} q_{\Theta|\,U=u}(\theta) &= \frac{q(\theta)P_{U|\,\Theta=\theta}(u)}{\sum\limits_{\theta} q(\theta)P_{U|\,\Theta=\theta}(u)} \\ &= \frac{q(\theta)P_{U|\,\Theta=\theta}(u)}{P_{U|\,\Theta=\theta}(u)\sum\limits_{\theta} q(\theta)} \\ &= \frac{q(\theta)}{\sum\limits_{\theta} q(\theta)} \\ &= q(\theta). \end{aligned}$$

پس توزیع پسین Θ برای مشاهده U نسبت به توزیع پیشین Θ تغییر نکرده است [۶]. به عبارت دیگر با داشتن U نمی‌توان عقیده راجع به Θ را تغییر داد. لذا از دیدگاه بیزی، آماره کمکی U آماره‌ای است که به تنهایی کمکی است.

۴. نتیجه گیری

در این مقاله نشان دادیم هیچ آماره کمکی (یا آماره کمکی مرتبه اول) غیر ثابت که تابعی از آماره بسنده کامل باشد وجود ندارد و بر عکس اگر بتوان تابعی از آماره بسنده ساخت که آماره کمکی و لذا آماره کمکی مرتبه اول باشد، لزومی ندارد این آماره بسنده به طورکامل فشرده شده باشد. ولی اگر هیچ تابع غیر ثابتی از آماره بسنده که آماره کمکی مرتبه اول باشد وجود نداشته باشد، آنگاه این آماره بسنده، کامل نیز هست. همچنین ثابت می‌شود آماره کمکی مشروط غیر بدیهی به شرط یک آماره بسنده کامل وجود ندارد. این حقیقت که آماره کمکی غیر ثابت، تابعی از آماره بسنده کامل نیست و همچنین این نکته که از دیدگاه بیزی یك آماره کمکی به تنهایی امکان ندارد راجع به پارامتر به ما اطلاعاتی بدهد، ظاهراً نشان می‌دهد یک آماره کمکی نباید حاوی اطلاعات مفیدی راجع به پارامتر باشد. ولی آماره کمکی بر خلاف ظاهرش که حاوی اطلاعات پیرامون پارامتر نیست، به همراه برآوردگر حداکثر درستنمایی می‌تواند اطلاعات از دست رفته از نمونه را دوباره بازیابی کند. همچنین یک آماره بسنده مینیمال حاوی اطلاعات اضافی است که برای تخمین پارامتر مفید نیست و به عبارت دیگر آماره بسنده مینیمال هنوز به طور کامل فشرده نشده است. در صورت کامل بودن آماره بسنده، در آماره هیچ اطلاعات اضافی وجود نخواهد داشت و در حقیقت هیچ تابعی از این آماره، یک آماره کمکی نیست که البته این نشان دهنده استقلال آماره بسنده کامل از آماره کمکی است. بنابراین در حالتی که برآوردگر حداکثر درستنمایی یک آماره بسنده کامل است، خود به تنهایی تمام اطلاعات راجع به پارامتر را در بر دارد و نیازی به این که برای بازیابی اطلاعات از دست رفته با آماره کمکی تلفیق شود نیست. در حالتی که برآوردگر حداکثر درستنمایی، کامل نیست با استفاده از یک آماره کمکی می‌توان اطلاعات از دست رفته در آن را دوباره بازیابی کرد. این بازیابی اطلاعات مفقود شده به دو روش انجام می‌شود، یکی این که توزیع توأم آماره کمکی با یك براوردگر حداکثر درستنمایی در نظر گرفته شود که در این حالت یك آماره بسنده مینیمال حاصل می‌شود و با این روش اطلاعات از دست رفته در آماره کمکی دوباره بازیابی شده و منجر به یک آماره آگاهی بخش می‌شود. در روش دیگر با شرطی کردن براوردگر حداکثر درستنمایی روی آماره‌های کمکی، هیچ اطلاعی راجع به پارامتر از دست نمی‌رود و در این حالت نیز شرطی کردن، موجب بازیابی اطلاعات مفقود شده راجع به پارامتر شده و در حقیقت آگاهی بخش بوده است.

DOI: http://dx.doi.org/10.22108/msci.2019.105827.1235

مراجع

[1] G. A. Barnard and D. A. Sprott, A Note On Basu's Examples of Anomalous Ancillary Statistics (with discussion), *Fundation of Statistical Inference*, (1971) 176–163.

[2] D. Basu, The Family of Ancillary Statistics, *Sankhya*, **21** (1959) 247–256.

[3] D. Basu, Recovery of Ancillary Information, *Sankhya, A.*, **26** (1964) 3–16.

[4] D. Basu, On Ancillary Statistics, Pivotal Quantities and Confidence Statements, *Topics in Applied Statistics*, (Y. P. Chaubey and T. D. Dwivedi, eds.) Concordia University, Montreal, (1981) 1–29.

[5] D. D. Boos and J. M. Hughes-Oliver, Applications of Basu's Theorem, *Amer. Statist.*, **52** (1998) 218–221.

[6] A. P. Dawid, Conditional Independence for Statistical Operations, *Ann. Statist.*, **8** (1980) 598–617.

[7] R. A. Fisher, Theory of Statistical Estimation, *Proc. Camb. Phil. Soc.*, **22** (1925) 700–725.

[8] M. Ghosh, Basu's Theorem with Applications, *Sankhya, A.*, **64** (2002) 509–531.

[9] E. L. Lehmann, An Interpretation of Completeness and Basu's Theorem, *J. Amer. Statist. Assoc.*, **76** (1981) 335–340.

[10] E. L. Lehmann and J. P. Romano, Testing Statistical Hypotheses, 3rd edition, Springer, New York, 2005.

[11] E. L. Lehmann and F. W. Scholz, Ancillarity. In Current Issues in Statistical Inference: Essays in Honor of D. Basu. Eds. M. Ghosh and P. K. Pathak. *Ims Lecture Notes and Monograph Series*, **17** (1992) 32-51.

مهدی شمس
اصفهان، دانشگاه کاشان، گروه آمار
mehdishams@kashanu.ac.ir

مهدی شمس متولد آذر ماه ۱۳۵۷ در شهر اصفهان است. وی در سال ۱۳۷۵ وارد مقطع کارشناسی رشته ریاضی کاربردی در دانشگاه اصفهان شد. در سال ۱۳۸۰ وارد مقطع کارشناسی ارشد رشته آمار محض در دانشگاه صنعتی اصفهان شد. وی در سال ۱۳۸۷ وارد مقطع دکتری رشته آمار، گرایش استنباط آماری در دانشگاه فردوسی مشهد شد. او از سال ۱۳۹۲ تاکنون عضو هیات علمی گروه آمار دانشگاه کاشان است.

DOI: http://dx.doi.org/10.22108/msci.2019.105827.1235