Kharazmi University

Mathematical Research

Year 2026, Volume 11, Issue 4, pp. 8–30

Print ISSN: 2588-2546

Online ISSN: 2588-2554

DOI: xxxx

Independence of sufficient equivariant statistics and invariant statistics under topological groups

Mehdi Shams[1]

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

Received: 22 November 2024 Accepted:2 November 2025 Published online:25 February 2026

Abstract: In statistics, the independence between statistics is of particular importance. Without calculating the joint distribution of two statistics, given the necessary conditions such as the completeness and sufficiency of one statistic and the ancillarity of the other, this independence can be proven. This paper examines generalizations of this result, as well as cases involving bounded completeness and Bayesian sufficiency. Finally, under the action of topological groups, this fact is generalized, and the independence of an invariant function from a sufficient equivariant statistic is established.

Keywords: Group action, Countably completeness, Pairwise sufficiency, Bayesian viewpoint, Invariance, Equivariance.

©2026 Kharazmi University, Tehran, Iran. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution-NonComertial 4.0 International (CC BY-NC 4.0 license) (http://creativecommons.org/licenses/by-nc/4.0/).

[1]Corresponding author

E-mail addresses: (Mehdi Shams) mehdishams@kashanu.ac.ir

پژوهش‌های ریاضی

سال ۱۴۰۴، دوره ۱۱، شماره ۴، ص ۸—۳۰

شاپا چاپی: ۲۵۸۸-۲۵۴۶

شاپا الکترونیکی: ۲۲۸۸-۲۵۵۴

DOI: xxxx

استقلال آماره هم‌وردای بسنده و آماره ناوردا تحت گروه‌های توپولوژیکی

مهدی شمس[۱]

گروه آمار، دانشکدهٔ علوم ریاضی، دانشگاه کاشان، کاشان، ایران.

تاریخ دریافت: ۱۴۰۳/۹/۲ تاریخ پذیرش: ۱۴۰۴/۸/۱۱ تاریخ انتشار: ۱۴۰۴/۱۲/۷

چکیده: در علم آمار وجود استقلال بین آماره‌ها از اهمیت ویژه‌ای برخوردار است. بدون محاسبه توزیع توأم دو آماره، با داشتن شرایط لازم از جمله بسنده کامل بودن یک آماره و کمکی بودن آماره دیگر، وجود این استقلال ثابت می‌شود. در این مقاله تعمیم‌های این موضوع و همچنین حالت‌های به‌طور شمارا کامل بودن و بیزی مورد بررسی قرار خواهند گرفت. در پایان، تحت عمل گروه‌های توپولوژیکی این حقیقت تعمیم داده شده و استقلال یک تابع ناوردا با یک آماره هم‌وردای بسنده ثابت می‌شود.

واژه‌های کلیدی: عمل گروه، به‌طور شمارا کامل بودن، بسندگی جفتی، دیدگاه بیزی، ناوردایی، هم‌وردایی.

۱ پیش‌گفتار

قضیه باسو جزو معدود نتیجه‌های ساده در علم آمار است که با گذشت زمان هنوز از اهمیت ویژه‌ای برخوردار است و مثل لم نیمن–پیرسن، نامساوی کرامر–رائو و قضیه رائو–بلاکول هسته اصلی استنباط آمار کلاسیک را تشکیل می‌دهد [۱۲]. این قضیه یکی از قضیه‌های مشهور آمار است که در سال ۱۹۵۵ توسط باسو [۲] اثبات شد و با توجه به کاربردهای متنوع آن سال‌ها پس از اثباتش مشهور و کارا باقی ماند. در واقع این قضیه در بیشتر کتاب‌های مهم استنباط آماری از جمله لی‌من و رومانو [۲۰]، لی‌من و کسلا [۱۹]، ماخوپادیای [۲۱] و کسلا و برگر [۶] بیان شده

[۱] نویسنده مسئول مقاله

(Mehdi Shams) mehdishams@kashanu.ac.ir

است. قضیه باسو به‌صورت یک نتیجه کاملاً کاربردی ظاهر می‌شود و این جنبه آن باعث شده که کاربردهای فراوانی در زمینه‌های مختلف داشته باشد.

آماره، تابعی از نمونه تصادفی است که اطلاعات مربوط به چند متغیر تصادفی را در یک متغیر تصادفی فشرده می‌سازد. در تلخیص داده‌ها ممکن است کمی از اطلاعات مربوط به مسئله مورد مطالعه از بین برود، لذا آماره‌ای که برای استنباط به‌کار برده می‌شود، نباید باعث تلخیص نمونه اولیه به‌قسمی شود که دیگر برای بررسی مسئله کفایت نکند. علت مطالعه بسندگی در استنباط آماری این است که خلاصه کردن و فشرده کردن داده‌ها توسط یک آماره مورد علاقه است به‌طوری که بدون از دست دادن اطلاعات، شامل تمام اطلاعات نمونه پیرامون پارامتر باشد. پس از تعریف بسندگی، این سؤال مطرح می‌شود که چقدر می‌توان داده‌ها را خلاصه کرد که هنوز بسنده باشند. واضح است خلاصه کردن داده‌ها تا جایی جایز است که اطلاعات نهفته در داده‌ها در مورد پارامتر مجهول حفظ شود. آماره‌ای که این خصوصیت را دارد، آماره بسنده مینیمال نامیده می‌شود و این آماره بسنده، حداکثر تلخیص را روی نمونه انجام می‌دهد. تفاوت اساسی بین یک آماره بسنده مینیمال و یک آماره بسنده کامل در این است که آماره بسنده ممکن است شامل اطلاعات اضافی (اطلاعات فرعی، کمکی یا ناآگاهی‌بخش) باشد که برای تخمین پارامتر مجهول مفید نیست، (در اصطلاح گویند آماره بسنده هنوز به‌طور کامل فشرده نشده است) و این یعنی ممکن است یک تابع از آماره بسنده مینیمال وجود داشته باشد که یک آماره کمکی است؛ در صورتی که اگر یک آمار بسنده، کامل نیز باشد، وجود یک تابع از آن که کمکی باشد امکان‌پذیر نیست. آماره کمکی، آماره‌ای است که توزیع آن به پارامتر مجهول بستگی ندارد. قابل ذکر است که وجود آماره بسنده مینیمال به خودی خود، عدم وجود تابعی از آن که کمکی باشد را تضمین نمی‌کند. در صورتی که با به‌کارگیری قضیه باسو، وجود آماره بسنده کامل تضمین می‌کند که تابعی از این آماره که کمکی باشد وجود ندارد، و این بدین معنی است که آماره بسنده کامل شامل اطلاعات کمکی و غیر لازم درباره پارامتر مجهول نیست. به‌عبارت عامیانه‌تر کامل بودن یک آماره بسنده، منجر به بیشترین فشردگی در داده‌ها می‌شود و به‌عبارت دیگر آماره بسنده کامل در بیان و فشردگی ممکن داده‌ها موفق است. اولین ارتباط بین مفاهیم فوق در سال ۱۹۵۵ توسط باسو اثبات شد که باعث کشف ارتباط بین بسندگی، آماره‌های کمکی و استقلال گردید [۱۳]. به‌نظر می‌رسد ایده‌ای که باسو را به فکر مطرح کردن این قضیه انداخت به‌شرح زیر باشد:

«چنانچه آماره بسنده‌ای، کامل باشد علاوه بر داشتن همه اطلاعات لازم درباره پارامتر، هیچ اطلاعات بیشتری درباره آن ندارد. پس چنین آماره‌ای نمی‌تواند ارتباطی با یک آماره کمکی که شامل هیچ اطلاعی درباره پارامتر نیست (چون توزیع آن به پارامتر بستگی ندارد) داشته باشد. از این رو طبیعی به‌نظر می‌رسد که هر آماره کمکی مستقل از آماره بسنده کامل باشد.»

نسخه اصلی قضیه باسو، در [۲] اثبات شد. این قضیه یکی از قضیه‌های اساسی استنباط آماری است که کاربردهای فراوانی در شاخه‌های متعدد آمار دارد و استقلال آماره بسنده کامل از آماره کمکی را تضمین می‌کند. لذا از مراجع متعدد، اثبات‌های مختلفی که برای آن ارائه شده به زبان ساده بیان خواهند شد. در [۵] به این نکته اشاره می‌شود که اگر از دیدگاه بیزی به مسئله نگاه شود، آنگاه مفهوم بسندگی و کمکی بودن مستقیماً به نظریه استقلال شرطی وابسته می‌شود [۹]. در [۴]، نسخه بیزی قضیه باسو مطرح شد، یعنی برای آماره U و T که در چارچوب بیزی، به‌ترتیب کمکی و بسنده هستند، یعنی متغیرهای تصادفی U و Θ مستقل باشند و متغیرهای تصادفی X و Θ مستقل شرطی به شرط T باشند و علاوه بر آن T به‌طور کران‌دار کامل باشد، ثابت شد T و U مستقل شرطی به شرط Θ هستند. تعریف بیزی بسندگی، به نظریه بسندگی بیزی که توسط کولموگوروف [۱۶] ارائه شد، مربوط می‌شود. جرگنسن [۱۵] یک تعمیم از قضیه باسو را مطرح کرد که اگر آماره توأم (T,V) بسنده و توزیع شرطی $T|V$ کامل و (V,U) یک آماره کمکی باشد، آنگاه T و U مستقل شرطی روی V هستند [۱۸]. در [۱۷]، نشان داده شد که یک آماره بسنده جفتی و به‌طور شمارا کامل برای یک خانواده از توزیع‌های احتمال از یک آماره کمکی مستقل است. در [۱۴]، مفهوم کامل بودن به کامل بودن شرطی نسبت به یک میدان سیگمایی تعمیم داده می‌شود و در این حالت نیز می‌توان تعمیم دیگری از قضیه باسو را مطرح کرد.

خطوط اصلی بخش‌های این مقاله به‌شرح ذیل است:
در بخش اول، قضیه باسو و اثبات‌های مختلف آن مرور می‌شود. در بخش دوم، تعمیم‌های این قضیه برای آماره کمکی مشروط یا آماره بسنده جفتی و به‌طور شمارا کامل معرفی خواهد شد. در بخش سوم، حالت بیزی این قضیه معرفی می‌شود. بالاخره در بخش چهارم، ایده استقلال یک تابع ناوردا و هم‌وردا که تعمیمی بر قضیه باسو تحت گروه‌های توپولوژیکی هست بیان و اثبات می‌شود. در بخش پنجم، نتیجه‌گیری و پیشنهادات برای تحقیقات آینده مطرح می‌شود.

۲ استقلال آمار بسنده کامل از آماره کمکی

فرض کنید توزیع متغیر تصادفی X به پارامتر θ بستگی دارد و بر این اساس خانواده توزیع‌های X که یک خانواده پارامتری است با $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ نمایش داده می‌شود. در این جا برای راحتی استقلال دو متغیر تصادفی X و Y با نماد $X \perp\!\!\!\perp Y$ نمایش داده می‌شود [۱۳].

ابتدا برای آشنایی خوانندگان با مفاهیم پایه‌ای مورد نیاز، چند تعریف مقدماتی در این رابطه مرور می‌شوند.

هر تابع از متغیر تصادفی نظیر $T \equiv T(X)$ که به پارامتر مجهولی بستگی نداشته باشد یک آماره نامیده می‌شود [۶]. معمولاً $T(\mathbf{X})$ به‌صورت تابعی از نمونه تصادفی $\mathbf{X} = (X_1, \ldots, X_n)$ تعریف می‌شود که اطلاعات مربوط به چند متغیر تصادفی را در یک متغیر تصادفی فشرده می‌سازد [۶]. اصطلاح آماره برای اولین بار در سال ۱۹۲۲ توسط فیشر معرفی شد [۷، ۸].

آماره $T \equiv T(\mathbf{X})$ را برای θ (یا به‌طور معادل، برای خانواده $\mathcal{P}$) بسنده گویند، هرگاه به‌ازای هر θ، توزیع شرطی $\mathbf{X}|T = t$، برای همه مقادیر t به پارامتر $\theta \in \Theta$ بستگی نداشته باشد. آماره بسنده T را بسنده مینیمال گویند، هرگاه تابعی از هر آماره بسنده دیگر باشد [۱۹].

آماره U را کمکی (فرعی یا ناآگاهی‌بخش) برای یک پارامتر گویند، هرگاه توزیع U به آن پارامتر مجهول بستگی نداشته باشد [۵]. بنابراین آماره کمکی شامل هیچ اطلاعی پیرامون پارامتر نیست [۵]. اصطلاح آماره کمکی برای اولین بار در سال ۱۹۲۵ توسط فیشر [۱۱] معرفی شد [۳]. آماره U را کمکی مرتبه اول گویند، اگر $\mathbb{E}_\theta[U]$ به $\theta \in \Theta$ بستگی نداشته باشد. بدیهی است که اگر U یک آماره کمکی باشد، آنگاه $\mathbb{E}_\theta[U]$ مقداری ثابت است و به $\theta \in \Theta$ بستگی ندارد و بنابراین U آماره کمکی مرتبه اول نیز هست.

آماره T را $\mathcal{G}-$کامل ($\mathcal{G}$ یک خانواده از توابع با مقادیر حقیقی است) گویند، هرگاه برای هر تابع $g \in \mathcal{G}$، اگر به‌ازای هر $\theta \in \Theta$، $\mathbb{E}_\theta[g(T)] = 0$، آنگاه به‌ازای هر $\theta \in \Theta$، $P_\theta(g(T) = 0) = 1$ [۱۸]. در حالت خاص اگر $\mathcal{G}$ خانواده توابع انتگرال‌پذیر یا توابع کران‌دار باشد، به‌ترتیب آماره T را کامل و به‌طور کران‌دار کامل می‌نامند. به‌عبارت دیگر خانواده توزیع‌های تولید شده توسط T کامل است، اگر تنها برآوردگر ناریب برای صفر در این خانواده خود صفر باشد. واضح است که هر آماره کامل، به‌طور کران‌دار کامل است، ولی عکس این مطلب صحیح نیست [۱۹]. اصطلاح کامل بودن برای اولین بار در سال ۱۹۵۰ توسط لی‌من و شفه [۱۹] معرفی شد [۷، ۸]. ویژگی کامل بودن یک آماره به تنهایی کاربردی ندارد، ولی وقتی این ویژگی مربوط به آماره بسنده باشد، خواص یکتایی مشخصی را تضمین می‌کند و موجب سادگی برخی از مسائل آمار و احتمال می‌شود. لذا از این به بعد بحث، به آماره بسنده کامل (و یا به‌طور کران‌دار کامل) محدود می‌شود.

اکنون قضیه باسو و اثبات‌های متفاوت آن در مراجع [۱، ۱۰، ۱۳، ۱۶] بیان می‌شود.

قضیه ۱.۲. (قضیه باسو [۲]) *اگر آماره T برای خانواده توزیع‌های $\mathcal{P}$ بسنده و به‌طور کران‌دار کامل باشد و آماره U یک آماره کمکی باشد، آنگاه T و U به‌ازای هر θ مستقل (شرطی) هستند.*

اثبات. **برهان اول:** با توجه به کمکی بودن آماره U، به‌ازای یک مقدار ثابت u، $p = F_U(u) = P_\theta(U \leq u)$ به θ

بستگی ندارد. اکنون اگر تابع کران‌دار g به‌صورت $g(t) = P(U \leq u \,|\, T = t)$ تعریف شود، به‌سادگی می‌توان نوشت

$$\begin{aligned} \mathbb{E}_\theta\left(g(T) - p\right) &= \mathbb{E}_\theta \mathbb{E}\left(I_{[-\infty,u]}(U) \,\middle|\, T\right) - p \\ &= \mathbb{E}_\theta\left(I_{[-\infty,u]}(U)\right) - p \\ &= P_\theta\left(U \leq u\right) - p = 0 \end{aligned} \tag{۱.۲}$$

که در تساوی دوم از ویژگی امید مکرر استفاده شده است.

از اینکه g یک تابع کران‌دار با مقادیر حقیقی است و همچنین با توجه به ویژگی به‌طور کران‌دار کامل بودن T، از (۱.۲) می‌توان نتیجه گرفت که برای هر $\theta \in \Theta$، $P_\theta\left(g(T) = p\right) = 1$ و از این رو با احتمال ۱، $P(U \leq u) = P\left(U \leq u \,|\, T = t\right)$ که این تساوی با توجه به کمکی بودن U و بسندگی T به‌صورت $P_\theta(U \leq u) = P_\theta\left(U \leq u \,|\, T = t\right)$ بازنویسی می‌شود و این معادل استقلال (شرطی) T و U است، زیرا برای هر $\theta \in \Theta$:

$$\begin{aligned} P_\theta(U \leq u)P_\theta(T \leq t) &= pP_\theta(T \leq t) \\ &= p\,\mathbb{E}_\theta\left(I_{[-\infty,t]}(T)\right) \\ &= \mathbb{E}_\theta\left(p\, I_{[-\infty,t]}(T)\right) \\ &= \mathbb{E}_\theta\left(I_{[-\infty,t]}(T)\, g(T)\right) \\ &= \mathbb{E}_\theta\left(I_{[-\infty,t]}(T)\, \mathbb{E}\left(I_{[-\infty,u]}(U) \mid T\right)\right) \\ &= \mathbb{E}_\theta\left(\mathbb{E}\left(I_{[-\infty,t]}(T)\, I_{[-\infty,u]}(U) \mid T\right)\right) \\ &= \mathbb{E}_\theta\left(\mathbb{E}\left(I_{[-\infty,t]\times[-\infty,u]}(T,U) \mid T\right)\right) \\ &= \mathbb{E}_\theta\left(I_{[-\infty,t]\times[-\infty,u]}(T,U)\right) \\ &= P_\theta(U \leq u, T \leq t) \end{aligned} \tag{۲.۲}$$

که در تساوی سوم و ششم از ویژگی خطی بودن امید ریاضی و در تساوی هشتم از ویژگی امید مکرر استفاده شده است. در پایان تساوی (۲.۲) استقلال (شرطی) T و U را اثبات می‌کند.

برهان دوم: تابع دلخواه کران‌دار و با مقادیر حقیقی h را در نظر بگیرید. اکنون برای هر $\theta \in \Theta$ می‌توان نوشت:

$$\begin{aligned} \mathbb{E}_\theta\left[\mathbb{E}(g(U) \mid T) - \mathbb{E}g(U)\right] &= \mathbb{E}_\theta\mathbb{E}(g(U) \mid T) - \mathbb{E}_\theta\mathbb{E}g(U) \\ &= \mathbb{E}g(U) - \mathbb{E}g(U) = 0. \end{aligned} \tag{۳.۲}$$

حال با توجه به اینکه T بسنده و U کمکی است، می‌توان نتیجه گرفت که $\mathbb{E}_\theta[h(U)|T]$ و $\mathbb{E}_\theta[h(U)]$ بستگی به θ ندارند و لذا $h(T) = \mathbb{E}_\theta[h(U)|T] - E_\theta[h(U)]$ به پارامتر مجهول بستگی ندارد، یعنی یک آماره است. پس از (۳.۲) می‌توان نتیجه گرفت، برای هر $\theta \in \Theta$، $\mathbb{E}_\theta[h(T)] = 0$ که با توجه به کامل بودن T و اینکه h تابعی با مقادیر حقیقی است، با احتمال یک $\mathbb{E}_\theta[g(U)|T] = \mathbb{E}_\theta[g(U)]$ که با انتخاب $g(U) = I_{[-\infty,u]}(U)$ به‌صورت $P_\theta(U \in A|T) = P_\theta(U \in A)$ بازنویسی می‌شود و معادل استقلال T و U است.

برهان سوم: برای هر مجموعه بورل A تعریف کنید، $p_A = P(U \in A)$ و $\eta_A(t) = P(U \in A \mid T = t)$ که هر دو احتمال به پارامتر مجهول، یعنی θ بستگی ندارد (زیرا U کمکی و T بسنده است). لذا به‌ازای هر $\theta \in \Theta$،

$\mathbb{E}_\theta(\eta_A(T)) = \mathbb{E}_\theta(P(U \in A \mid T)) = \mathbb{E}_\theta[\mathbb{E}(I_A(U) \mid T)] = \mathbb{E}_\theta(I_A(U)) = P_\theta(U \in A) = p_A$.
حال با استفاده از کامل بودن T، می‌توان نتیجه گرفت که با احتمال یک $\eta_A(T) = p_A$، یعنی به‌ازای هر $\theta \in \Theta$، $P_\theta(U \in A \mid T = t) = P_\theta(U \in A)$ که این معادل استقلال T و U است.

برهان چهارم: برای استقلال T و U کافی است ثابت شود که به‌ازای هر u و t، $f_U(u) = f_{U|T=t}(u)$. تابع $f_U(u)$ به θ بستگی ندارد (چون U کمکی است). همچنین تابع $f_{U|T=t}(u)$ به θ بستگی ندارد (زیرا T بسنده است). لذا اگر برای هر u ثابت، تابع $h_u(t) = f_U(u) - f_{U|T=t}(u)$ تعریف شود، در این صورت تابع $h_u(t)$ فقط به t بستگی دارد و به θ بستگی ندارد (یعنی $h_u(T)$ یک آماره است). حال برای هر $\theta \in \Theta$ می‌توان نوشت:

$$\begin{aligned}\mathbb{E}_\theta(h_u(T)) &= \mathbb{E}_\theta(f_U(u)) - \mathbb{E}_\theta(f_{U|T=t}(u)) \\ &= f_U(u) - \int_{-\infty}^{\infty} f_{U|T=t}(u)\, f_\theta(t)\, dt \\ &= f_U(u) - \int_{-\infty}^{\infty} f_{U,T}(u,t)\, dt = 0\end{aligned} \tag{۴.۲}$$

که با توجه به ویژگی کامل بودن آماره T، از (۴.۲) می‌توان نتیجه گرفت با احتمال یک، $h_u(T) = 0$، یعنی به‌ازای هر t، $f_U(u) = f_{U|T=t}(u)$ که این هم معادل استقلال T و U است. □

چون هر آماره کامل، یک آماره به‌طور کران‌دار کامل نیز هست، قضیه ۱.۲، برای آماره‌های کامل نیز صحیح است. نسخه اصلی قضیه باسو [۲] صرفاً بر اساس چارچوب آمار کلاسیک نوشته شده و در صورت قضیه، کلمه شرطی آورده نشده بود. مورد اخیر در سال ۱۹۸۲ توسط باسو [۳] اضافه شد که احتمالاً برای تمایز چارچوب آمار کلاسیک از چارچوب آمار بیزی بوده و با این کار θ به‌عنوان متغیر تصادفی در نظر گرفته شده و لذا استقلال T و U مشروط (به شرط θ) خواهد شد. البته همان طور که در برهان اول در قضیه ۱.۲ مشاهده شد، دو تساوی $P(U \le u) = P(U \le u \,|\, T = t)$ و نیز $P_\theta(U \le u) = P_\theta(U \le u \,|\, T = t)$ (برای هر $\theta \in \Theta$) هر دو نتیجه گرفته می‌شوند و این یعنی در هر دو چارچوب کلاسیک و بیزی، قضیه ۱.۲ صحیح است. قضیه ۱.۲ علاوه بر خانواده‌های پارامتری، برای خانواده‌های ناپارامتری نیز برقرار است [۱].

اگر T آماره بسنده به‌طور کران‌دار کامل و U هر آماره کمکی دلخواه باشد، آنگاه با به‌کارگیری قضیه ۱.۲، می‌توان نتیجه گرفت که برای هر دو تابع دلخواه g و h، $g(T)$ و $h(U)$ ناهمبسته‌اند، یعنی $\mathbb{COV}(f(T), g(U)) = 0$ [۲۲].

با استفاده از قضیه ۱.۲ و این مطلب که میزان اطلاع فیشر در یک آماره کمکی برابر صفر است، می‌توان نتیجه گرفت که اگر T آماره بسنده به‌طور کران‌دار کامل و U هر آماره کمکی دلخواه باشد، آنگاه میزان اطلاع فیشر در دو آماره (T, U) و T یکسان است، یعنی $I_{(T,U)}(\theta) = I_T(\theta)$ که در آن $I_T(\theta) = \mathbb{E}_\theta\left(\frac{\partial}{\partial\theta} \ln g_\theta(T)\right)^2$ میزان اطلاع فیشر نسبت به پارامتر θ در آماره T و g_θ تابع چگالی یا تابع احتمال متغیر تصادفی T است. برای اثبات این حقیقت از این نکته استفاده می‌شود که طبق قضیه ۱.۲، $T \perp\!\!\!\perp U$، و از این‌رو $I_{(T,U)}(\theta) = I_T(\theta) + I_U(\theta)$ در پی آن $I_{(T,U)}(\theta) = I_T(\theta)$. البته چنین نتیجه‌ای به‌طور شهودی نیز قابل قبول است، زیرا با توجه به اینکه یک آماره کمکی اطلاعاتی درباره پارامتر مجهول نمی‌دهد (و برای همین منظور یکی از اسامی انتخابی برای آن آماره ناآگاهی‌بخش است)، بنابراین با ترکیب این آماره با یک آماره بسنده کامل (که خود به‌تنهایی حاوی تمام اطلاعات درباره پارامتر است)، هیچ اطلاع جدیدی درباره پارامتر به‌دست نمی‌آید.

در پایان این بخش به این نکته اشاره می‌شود که قضیه ۱.۲ برای آماره کمکی مرتبه اول نیز صحیح است، یعنی آماره بسنده و به‌طور کران‌دار کامل T و آماره کمکی مرتبه اول U به‌ازای هر $\theta \in \Theta$ مستقل (شرطی) هستند. ذکر این نکته لازم است که اگر T آماره بسنده کامل باشد، آنگاه هیچ آماره کمکی غیر از آماره ثابت به‌عنوان تابعی از T

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

۳ تعمیم در حالت‌های استقلال شرطی و همچنین استقلال آماره بسنده جفتی و به‌طور شمارا کامل از آماره کمکی

در این بخش چند تعمیم از قضیه ۱.۲ معرفی می‌شود که در دو مورد استقلال شرطی نتیجه‌گیری می‌شود و در یک مورد استقلال آماره بسنده جفتی و به‌طور شمارا کامل از آماره کمکی مطرح می‌شود.

جرگنسن [۱۵] طی یک سخنرانی در سال ۱۹۹۸ یک تعمیم کلی از قضیه ۱.۲ را مطرح کرد که در ذیل به آن اشاره شده است.

قضیه ۱.۳. ([۱۸]) *اگر آماره توأم (T,V) بسنده و توزیع شرطی $T\,|\,V$ کامل و (V,U) یک آماره کمکی باشد، آنگاه T و U مستقل شرطی روی V هستند، یعنی $T \perp\!\!\!\perp U\,|\,V$.*

اثبات. فرض کنید g تابعی کران‌دار باشد و تعریف کنید $h(T,V)=\mathbb{E}[g(U)\mid T,V]-\mathbb{E}[g(U)\mid V]$. اگر V آماره بسنده برای پارامتر θ باشد، آنگاه $\mathbb{E}[g(U)\mid V]$ تنها تابعی از V هست که به θ بستگی ندارد. از طرف دیگر بسندگی (T,V) منجر به این می‌شود که امید ریاضی شرطی $\mathbb{E}[g(U)\mid T,V]$ نیز به θ بستگی نداشته باشد، پس $h(T,V)$ یک آماره است. حال با استفاده از ویژگی امید مکرر داریم:

$$\mathbb{E}[h(T,V)\mid V]=\mathbb{E}[\mathbb{E}[g(U)\mid T,V]\mid V]-\mathbb{E}[g(U)\mid V]=0. \qquad (۱.۳)$$

از رابطه (۳.۳) و کامل بودن $T\mid V$ نتیجه می‌گیریم که با احتمال یک، $h(T,V)=0$ یا به‌طور معادل داریم $\mathbb{E}[g(U)\mid T,V]=\mathbb{E}[g(U)\mid V]$.

حال اگر $g(U)=I_{[-\infty,u]}(U)$ انتخاب شود، $P[U\le u\mid T=t,V=v]=P[U\le u\mid V=v]$ که این نشان‌دهنده استقلال شرطی U و T نسبت به V است. □

مثال ۲.۳. *فرض کنید $X_1,\dots,X_n \overset{i.i.d.}{\sim} E(\theta)$ که در آن $\theta>0$. قرار دهید $T=\sum_{i=1}^n X_i$، $V=\frac{X_1}{T}$ و $U=\frac{X_2}{X_1}$. تابع چگالی توأم $(X_1,\dots,X_n)$ برابر است با $f_X(\mathbf{x};\theta)=\theta^n\exp(-\theta\sum_{i=1}^n x_i)I_{(0,\infty)^n}(\mathbf{x})$ که در آن $\mathbf{x}=(x_1,\dots,x_n)$. تعریف کنید $W_i=\frac{X_i}{T}$، $i=1,\dots,n-1$.*

در این حالت تبدیل $(x_1,\dots,x_n)\mapsto(t,w_1,\dots,w_{n-1})$ را به‌کار برده که وارون این تبدیل به‌صورت $x_i=tw_i$، $i=1,\dots,n-1$ و $x_n=t(1-\sum_{i=1}^{n-1}w_i)$ و ژاکوبی تبدیل t^{n-1} خواهد بود. دامنه مقادیر برای $(t,\mathbf{w})$، $\mathbf{w}=(w_1,\dots,w_{n-1})\in\Delta_{n-1}=\left\{(w_1,\dots w_{n-1}):w_i\ge 0,\ \sum_{i=1}^{n-1}w_i<1\right\}$ هست. لذا با تعریف بردار تصادفی $\mathbf{W}=(W_1,\dots,W_{n-1})$، تابع چگالی احتمال توأم $(T,\mathbf{W})$ برابر است با $f_{T,\mathbf{W}}(t,\mathbf{w};\theta)=\theta^n e^{-\theta t}t^{n-1}I_{\{t>0,\mathbf{w}\in\Delta_{n-1}\}}$ که تابع $h(\mathbf{w})=I_{\{t>0,\mathbf{w}\in\Delta_{n-1}\}}$ به پارامتر θ بستگی ندارد. اکنون از اینکه $V=W_1=\frac{x_1}{T}$، تابع چگالی احتمال توأم (T,V) با انتگرال‌گیری روی $w_2,\dots,w_{n-1}$ برابر خواهد بود با $f_{T,V}(t,v;\theta)=\theta^n e^{-\theta t}t^{n-1}c(v)$ که در آن $c(v)=\int_{w_2,\dots,w_{n-1}} I_{\{(v,w_2,\dots,w_{n-1})\in\Delta_{n-1}\}}\,dw_2\cdots dw_{n-1}$ برای مقادیر $v\in(0,1)$، متناهی و مستقل از θ است. بنابراین، با استفاده از قضیه تجزیه نیمن [۱۹]، (T,V) یک آماره بسنده توأم برای θ خواهد بود.

از اینکه $f_{\mathbf{W}}(\mathbf{w})=\int_0^\infty f_{T,\mathbf{W}}(t,\mathbf{w};\theta)\,dt=I_{\{\mathbf{w}\in\Delta_{n-1}\}}\int_0^\infty\theta^n e^{-\theta t}t^{n-1}\,dt=\Gamma(n)I_{\{\mathbf{w}\in\Delta_{n-1}\}}$ به پارامتر θ بستگی ندارد، و در حقیقت بردار تصادفی $(\frac{X_1}{T},\dots,\frac{X_n}{T})$ دارای توزیع دریخله $Dirichlet(1,\dots,1)$

هست، لذا هر تابع اندازه‌پذیر از این نسبت‌ها نظیر U و V نیز آماره کمکی هستند. بنابراین (V,U) یک آماره کمکی است.

از اینکه $f_{T|V}(t|v;\theta)=\frac{f_{T,V}(t,v;\theta)}{f_V(v)}\propto\theta^n t^{n-1}e^{-\theta t}$، $t>0$، داریم $T|V=v\sim Gamma(n,\theta)$ که پارامتر شکل یعنی n معلوم است و از اینکه خانواده گاما متعلق به خانواده نمایی تک‌پارامتری هست، آماره $T|V$ کامل است. (در حقیقت با توجه به یکتایی تبدیل‌های لاپلاس، اثبات کامل بودن خانواده گاما واضح است.)

با توجه به برقراری شرایط قضیه ۱.۲، حکم قضیه ۱.۳، یعنی $T\perp\!\!\!\perp U\mid V$ صادق است که آن را می‌توان به‌صورت مستقیم نیز نتیجه گرفت. برای هر تابع اندازه‌پذیر کران‌دار b، تابع $m(v)=\mathbb{E}[b(U)\mid V=v]$ خوش‌تعریف است. اکنون تعریف می‌کنیم $h_v(t)=\mathbb{E}[b(U)\mid T=t,V=v]-m(v)$ و لذا طبق ویژگی امید مکرر، داریم $\mathbb{E}_\theta[h_v(T)\mid V=v]=0$. اما می‌دانیم $T\mid V$ کامل است که نتیجه می‌دهد تقریباً همه جا $h_v(T)=0$ و به‌عبارت دیگر $\mathbb{E}[b(U)\mid T,V]=\mathbb{E}[b(U)\mid V]$. تساوی اخیر نشان می‌دهد برای تمام توابع اندازه‌پذیر کران‌دار a و b، $\mathbb{E}[a(T)b(U)\mid V]=\mathbb{E}[a(T)\mid V]\mathbb{E}[b(U)\mid V]$ که معادل است با $T\perp\!\!\!\perp U\mid V$.

تعمیم مشابهی از قضیه ۱.۲ را می‌توان با استفاده از تعریف ۳.۳ بیان کرد.

تعریف ۳.۳. ([۲۳]) *آماره U را کمکی مشروط به شرط V گوییم، هرگاه توزیع $U|V=v$ 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 Uکمکی است.

نتیجه ۴.۳. ([۲۳]) *اگر T یک آماره بسنده کامل باشد، آماره کمکی مشروط نابدیهی به شرط T وجود ندارد.*

اثبات. فرض کنید آماره کمکی مشروط U به شرط T وجود داشته باشد، حال اگر A پیشامدی وابسته به U باشد، آنگاه $\mathbb{E}\left(P\left(A\mid T\right)-I_A(T)\right)=\mathbb{E}_\theta\left(\mathbb{E}\left(I_A\mid T\right)-P_\theta(A)\right)=\mathbb{E}_\theta(I_A)-P_\theta(A)=0$. با توجه به کامل بودن T، این نتیجه حاصل می‌شود که با احتمال یك $P\left(A\mid T\right)=I_A(T)$ و این یعنی U با احتمال یك ثابت است.

□

اکنون قضیه زیر برای استقلال شرطی آمار کمکی مشروط و آماره دلخواه دیگر که توأم آن آماره دلخواه با آماره‌ای که روی آن مشروط می‌شود، یک آماره بسنده کامل هست بیان می‌شود.

قضیه ۵.۳. *اگر آماره توأم (T,V) بسنده و T کامل باشد و U یك آماره کمکی مشروط به شرط V باشد، آنگاه T و U مستقل شرطی به شرط V هستند، یعنی $T\perp\!\!\!\perp U\mid V$.*

اثبات. چون آماره (T,V) بسنده است، توزیع شرطی $U\mid T=t,V=v$ برای هر t و v به پارامتر θ بستگی ندارد و به‌طور مشابه چون U کمکی مشروط به شرط $V=v$ است، بنابراین برای هر v، توزیع شرطی $U\mid V=v$ به پارامتر θ بستگی ندارد. در نتیجه $h(U,V)=P(U\leq u\mid T,V)-P(U\leq u\mid V)$ یک آماره است. بنابراین:

$$\begin{aligned}\mathbb{E}_\theta\{h(U,V)\}&=\mathbb{E}_\theta\{P(U\leq u\mid T,V)-P(U\leq u\mid V)\}\\&=\mathbb{E}_\theta\mathbb{E}(I_{[-\infty,u]}(U)\mid T,V)-\mathbb{E}_\theta\mathbb{E}(I_{[-\infty,u]}(U)\mid V)\\&=P_\theta(U\leq u)-P_\theta(U\leq u)=0\end{aligned}\qquad(۲.۳)$$

با توجه به (۲.۳) و کامل بودن آماره T، با احتمال یک $h(U,V)=0$. لذا برای هر u و t و v:

$$P(U\leq u\mid T=t,V=v)=P(U\leq u\mid V=v)\qquad(۳.۳)$$

که (۳.۳) معادل استقلال مشروط U و T به شرط V است، یعنی $T \perp\!\!\!\perp U \mid V$. □

مثال ۶.۳. *فرض کنید $X_1, X_2, X_3 \overset{i.i.d.}{\sim} N(\mu, 1)$ که در آن $\mu \in \mathbb{R}$. قرار دهید $T = \bar{X}$، $V = X_1 - X_2$ و $U = X_2 - X_3$. با استفاده از قضیه تجزیه نیمن [۱۹]، T یک آماره بسنده کامل برای μ خواهد بود. مشاهدات اصلی نمونه را می‌توان بر حسب (T, V, U) نوشت، یعنی $X_1 = \bar{X} + \frac{2}{3}V + \frac{1}{3}U$، $X_2 = \overline{X} - \frac{1}{3}V + \frac{1}{3}U$، $X_3 = \overline{X} - \frac{1}{3}V - \frac{2}{3}U$. این حقیقت نشان می‌دهد، (T, V, U) یک تبدیل خطی (وارون‌پذیر) از داده‌های (X_1, X_2, X_3) است. با توجه به اینکه T یک آماره بسنده برای μ بوده و (T, V) بخشی از تبدیل خطی وارون‌پذیر (T, V, U) است، (T, V) نیز یک آماره بسنده برای μ خواهد بود.*

اکنون به اثبات کمکی مشروط بودن $U \mid V$ می‌پردازیم. از اینکه $Y_i = X_i - \mu \sim N(0, 1)$، $i = 1, 2, 3$، متغیرهای تصادفی V و U را می‌توان بر حسب Y_iها به‌صورت $V = (Y_1 + \mu) - (Y_2 + \mu) = Y_1 - Y_2$ و $U = (Y_2 + \mu) - (Y_3 + \mu) = Y_2 - Y_3$ بازنویسی کرد. بنابراین

$$\begin{pmatrix} V \\ U \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} \sim N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \right). \tag{۴.۳}$$

با توجه به توزیع داده شده در ۴.۳، داریم $U \mid V = v \sim N\left(-\frac{v}{2}, \frac{3}{2}\right)$ که نشان می‌دهد U یک آماره کمکی مشروط به شرط V است و به‌طور دقیق‌تر و از دیدگاه بیزی، $\mu \perp\!\!\!\perp U \mid V$ (رجوع کنید به بخش ۴).

با توجه به قضیه ۵.۳، داریم که $T \perp\!\!\!\perp U \mid V$ که می‌خواهیم این ادعا را به‌طور مستقیم نیز اثبات کنیم؛ به‌راحتی ثابت می‌شود:

$$\begin{pmatrix} T \\ V \\ U \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim N_3 \left(\begin{pmatrix} \mu \\ 0 \\ 0 \end{pmatrix}, \Sigma = (\sigma_{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \right). \tag{۵.۳}$$

اما از ۴.۳، می‌دانیم $\mathbb{COV}(T, U \mid V) = \sigma_{13} - \sigma_{12}\sigma_{22}^{-1}\sigma_{23} = 0 - 0 \times \frac{1}{2} \times (-1) = 0$ که نتیجه می‌دهد $T \perp\!\!\!\perp U \mid V$.

حالت خاص دو تعمیم، یعنی قضیه‌های ۱.۳ و ۵.۳، در حالتی که آماره V یک متغیر تصادفی ثابت باشد، قضیه ۱.۲ را نتیجه می‌دهد؛ زیرا یك آماره کمکی مشروط به یك متغیر تصادفی ثابت، خود یك آماره کمکی است و همچنین استقلال مشروط به یک متغیر تصادفی ثابت، استقلال آن دو آماره را نتیجه می‌دهد. به‌عبارت دقیق‌تر، فرض کنید V یک متغیر تصادفی ثابت باشد. قضیه ۱.۳ بیان می‌کند که اگر آماره‌ی توأم (T, V) بسنده باشد و توزیع شرطی $T \mid V$ کامل و (V, U) یك آماره کمکی باشد، آنگاه $T \perp\!\!\!\perp U \mid V$. همچنین قضیه ۵.۳، بیان می‌کند که اگر آماره‌ی توأم (T, V) بسنده کامل باشد و U یك آماره کمکی مشروط به شرط V باشد، آنگاه $T \perp\!\!\!\perp U \mid V$. قضیه‌های ۱.۳ و ۵.۳، را با فرض جدید می‌توان به این صورت بازنویسی کرد که اگر T بسنده کامل باشد و U آماره کمکی باشد، آنگاه $T \perp\!\!\!\perp U$؛ که همان قضیه ۱.۲ است. توجه کنید که اگر V یک متغیر تصادفی ثابت باشد، بسندگی (T, V) منجر به بسندگی T می‌شود؛ کامل بودن $T \mid V$ منجر به کامل بودن T می‌شود؛ کمکی بودن (V, U) منجر به کمکی بودن U می‌شود؛ بسنده کامل بودن (T, V) منجر به بسنده کامل بودن T می‌شود؛ کمکی مشروط بودن U به شرط V منجر به کمکی بودن U می‌شود و در نهایت حکم قضیه‌های ۱.۳ و ۵.۳، یعنی $T \perp\!\!\!\perp U \mid V$، منجر به $T \perp\!\!\!\perp U$ می‌شود.

با در نظر گرفتن تعریف‌های ۷.۳ و ۹.۳ تعمیم دیگری از قضیه ۱.۲ به‌دست می‌آید.

تعریف ۷.۳. ([۱۷]) *آماره T را برای خانواده $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ بسنده جفتی گویند، هرگاه T برای همه زیرمجموعه‌های دو عضوی $\mathcal{P}_\circ \subset \mathcal{P}$ بسنده باشد. به‌طور معادل، هرگاه به‌ازای هر زیرمجموعه شمارای $\Theta_\circ \subset \Theta$، آماره T برای خانواده $\{P_\theta : \theta \in \Theta_\circ\}$ بسنده باشد.*

بسندگی، بسندگی جفتی را نتیجه می‌دهد، ولی عکس آن صحیح نیست [۱۷]. برای مشاهده این حقیقت بررسی مثال زیر می‌تواند مفید باشد.

مثال ۸.۳. *فرض کنید $\mathcal{P} = \{P_\theta : \theta \in \Theta\} = \{\delta_{x_\circ} : x_\circ \in [\circ,1] \times \{\circ\}\} \cup \{\lambda_{[\circ,1]\times\{1\}}\}$ که در آن $\mathcal{X} = [\circ,1] \times \{\circ,1\}$، تکیه‌گاه متغیر تصادفی، $\{\delta_{x_\circ} : x_\circ \in [\circ,1] \times \{\circ\}\}$ خانواده اندازه‌های دیراک و $\lambda_{[\circ,1]\times\{1\}}$ اندازه لبگ یک‌بعدی هستند. در حقیقت، $\delta_{x_\circ}(A) = I_A(x_\circ)$ و برای هر $A \subseteq [\circ,1] \times \{1\}$، $\lambda_{[\circ,1]\times\{1\}}(A)$ طول A را نشان می‌دهد. آماره S که به‌صورت $S(x,y) = x$ تعریف شده باشد، بسنده جفتی است، اما بسنده نیست. برای بررسی جزئیات باید سه حالت جداگانه برای همه زیرمجموعه‌های دو عضوی $\mathcal{P}_\circ \subset \mathcal{P}$ بررسی شوند:*

حالت اول: *هر دو پارامتر، اندازه‌های دیراک δ_{x_1} و δ_{x_2} باشند. در این حالت برای $i = 1, 2$ داریم:*

$$f_{\delta_{x_i}}(x,y) = \begin{cases} 1 & (x,y) = (x_i, \circ) \\ \circ & o.w. \end{cases} = g_i(x)h(x,y) \tag{۶.۳}$$

که در آن $h_1(x,y) = I_{\{\circ\}}(y)$.

حالت دوم: *یک پارامتر، اندازه دیراک $\delta_{x_\circ}$ و دیگری اندازه لبگ λ باشد. در این حالت داریم:*

$$\begin{aligned} f_{\delta_{x_\circ}}(x,y) &= \begin{cases} 1 & (x,y) = (x_\circ, \circ) \\ \circ & o.w. \end{cases} = g_1(x)h_1(x,y), \\ f_\lambda(x,y) &= \begin{cases} 1 & y = 1 \\ \circ & y = \circ \end{cases} = g_2(x)h_2(x,y). \end{aligned} \tag{۷.۳}$$

که در آن $h_1(x,y) = I_{\{\circ\}}(y)$ و $h_2(x,y) = I_{\{1\}}(y)$.

حالت سوم: *هر دو پارامتر اندازه‌های لبگ λ باشند. در این حالت داریم:*

$$f_\lambda(x,y) = \begin{cases} 1 & y = 1 \\ \circ & y = \circ \end{cases} = g_2(x)h_2(x,y) \tag{۸.۳}$$

که در آن $h_2(x,y) = I_{\{1\}}(y)$

از سه حالت بررسی شده و تساوی‌های (۶.۳)، (۷.۳) و (۸.۳) و قضیه تجزیه نیمن [۱۹]، بسندگی جفتی آماره S، یعنی $S(x,y) = x$ را ثابت می‌کند.

برای اثبات عدم بسندگی آماره S به برهان خلف عمل می‌کنیم. فرض می‌کنیم آماره S برای خانواده $\mathcal{P}$ بسنده باشد. در این حالت طبق قضیه تجزیه نیمن [۱۹]، توابع نامنفی g_θ و h وجود دارند که برای هر $\theta \in \Theta$ داریم $f_\theta(x,y) = g_\theta(S(x,y)) \cdot h(x,y) = g_\theta(x) \cdot h(x,y)$. دو حالت ویژه را مورد بررسی قرار می‌دهیم. اندازه

دیراک در $(x_۰, ۰)$، یعنی $\theta = \delta_{(x_۰,۰)}$ را در نظر می‌گیریم. بنابراین:

$$f_{\delta_{(x_۰,۰)}}(x,y) = \begin{cases} ۱ & (x,y) = (x_۰, ۰) \\ ۰ & o.w. \end{cases} = g_\theta(x) \cdot h(x,y) \qquad (۹.۳)$$

بنابراین (۹.۳) نتیجه می‌دهد:

$$g_{\delta_{(x_۰,۰)}}(x_۰) \cdot h(x_۰, ۰) = ۱ \qquad (۱۰.۳)$$

و برای هر $(x,y) \neq (x_۰, ۰)$، $g_{\delta_{(x_۰,۰)}}(x) \cdot h(x,y) = ۰$.

در حالت دیگر، اندازه لبگ روی $[۰, ۱] \times \{۱\}$ یعنی $\theta = \lambda$ را در نظر می‌گیریم. بنابراین

$$f_\lambda(x,y) = \begin{cases} ۱ & y = ۱ \\ ۰ & y = ۰ \end{cases} = g_\theta(x) \cdot h(x,y) \qquad (۱۱.۳)$$

لذا با توجه به (۱۱.۳) برای هر x، $g_\lambda(x) \cdot h(x, ۰) = ۰$ و $g_\lambda(x) \cdot h(x, ۱) = ۱$ که در تساوی دوم این نکته تضمین می‌شود که برای هر x، $g_\lambda(x) \neq ۰$ و $h(x, ۱) \neq ۰$. اما در حالت اندازه‌گیری دیراک، برای هر $x \neq x_۰$، $g_{\delta_{(x_۰,۰)}}(x) \cdot h(x, ۰) = ۰$ که با توجه به ناصفر بودن تابع g، برای هر x داریم $h(x, ۰) = ۰$ و به‌ویژه $h(x_۰, ۰) = ۰$. با قرار دادن این مقدار در (۱۰.۳) به تناقض $۱ = ۰$ می‌رسیم که نشان می‌دهد آماره S برای خانواده $\mathcal{P}$ بسنده نیست.

به زبان ساده‌تر، در دو حالت مختلف از فضای پارامتری، یعنی برای تابع دیراک (لبگ)، تابع چگالی احتمال تنها برای مقدار $y = ۰$ ($y = ۱$) مثبت است و لذا هیچ تابعی مانند $h(x,y)$ که مستقل از θ باشد وجود ندارد که برای تمام توزیع‌های متعلق به خانواده $\mathcal{P}$ بتوانیم به کمک قضیه تجزیه نیمن [۱۹]، تابع چگالی $f_\theta(x,y)$ را به حاصل‌ضرب تابعی از آماره S و θ، مثل $g_\theta(S(x,y)) = g_\theta(x)$ در تابعی از نمونه نظیر $h(x,y)$ نوشت؛ بنابراین آماره S در خانواده زیرمجموعه‌های دو عضوی $\mathcal{P}_۰ \subseteq \mathcal{P}$ بسنده و لذا بسندگی جفتی است، ولی در خانواده $\mathcal{P}$ بسنده نیست. □

تعریف ۹.۳. ([۱۷]) آماره T را به‌طور شمارا کامل گویند، هرگاه یک زیرمجموعه شمارای $\Theta_۰ \subset \Theta$ وجود داشته باشد که برای هر زیرمجموعه شمارای $\Theta_۰ \subset \Theta_۱ \subset \Theta$، خانواده $\mathcal{P} = \{P_\theta(T^{-۱}(.)) : \theta \in \Theta_۱\}$ کامل باشد.

به‌طور شمارا کامل بودن، کامل بودن را نتیجه می‌دهد، ولی عکس آن صحیح نیست [۱۷]. در زیر یک مثال نقض برای برقرار نبودن عکس ارائه می‌شود.

مثال ۱۰.۳. فرض کنید $\mathcal{P} = \{P_\theta = \frac{۱}{۲}\lambda_{[۰,۱]} + \frac{۱}{۲}\delta_\theta : \theta \in \Theta = [۰, ۱]\}$ که در آن $\mathcal{X} = [۰, ۱] \times \{۰, ۱\}$ تکیه‌گاه متغیر تصادفی، δ_θ اندازه دیراک و $\lambda_{[۰,۱]}$ اندازه لبگ یک‌بعدی هستند. فرض کنید $f : [۰, ۱] \to \mathbb{R}$ یک تابع بورل‌ـ اندازه‌پذیر باشد به‌طوری که برای هر $\theta \in [۰, ۱]$، $\int f \, dP_\theta = \frac{۱}{۲}\int_۰^۱ f(x)\, d\lambda(x) + \frac{۱}{۲}f(\theta) = ۰$. با قرار دادن $c = \int_۰^۱ f(x) d\lambda(x)$، داریم $\frac{۱}{۲}c + \frac{۱}{۲}f(\theta) = ۰$ که نتیجه می‌دهد برای هر $\theta \in [۰, ۱]$، $f(\theta) = -c$. بنابراین $c = \int_۰^۱ f(x) d\lambda(x) = \int_۰^۱ (-c) d\lambda(x) = -c \cdot \lambda([۰, ۱]) = -c$ که نتیجه می‌دهد برای هر $\theta \in [۰, ۱]$، $f(\theta) = ۰$. پس تقریباً همه‌جا روی $[۰, ۱]$، $f = ۰$ که کامل بودن خانواده $\mathcal{P}$ را نشان می‌دهد.

اکنون نشان می‌دهیم خانواده $\mathcal{P}$ به‌طور شمارا کامل نیست. با برهان خلف، فرض کنید خانواده $\mathcal{P}$ به‌طور شمارا کامل باشد. خانواده شمارای $\Theta_۰ = \{\theta_۱, \theta_۲, \ldots\}$ و تابع بورل‌ـ اندازه‌پذیر $g : [۰, ۱] \to \mathbb{R}$ را در نظر بگیرید که

به‌صورت زیر تعریف می‌شود:

$$g(x)=\begin{cases}۱ & x\in\Theta_{۰}\\ -۱ & x\in[۰,۱]\setminus\Theta_{۰}\end{cases} \tag{۱۲.۳}$$

بنابراین برای مقادیر $\theta_i\in\Theta_{۰}$ داریم:

$$\begin{aligned}\int g\,dP_{\theta_i}&=\frac{۱}{۲}\int_{۰}^{۱}g(x)\,d\lambda(x)+\frac{۱}{۲}g(\theta_i)\\&=\frac{۱}{۲}[۱\cdot\lambda(\Theta_{۰})+(-۱)\cdot\lambda([۰,۱]\setminus\Theta_{۰})]+\frac{۱}{۲}g(\theta_i)\\&=\frac{۱}{۲}[۱\times۰+(-۱)\times۱]+\frac{۱}{۲}\times۱=۰\end{aligned} \tag{۱۳.۳}$$

که در تساوی آخر از شمارا بودن $\Theta_{۰}$ و در پی آن $\lambda(\Theta_{۰})=۰$ و $\lambda([۰,۱]\setminus\Theta_{۰})=۱$ و همچنین تعریف تابع g در (۱۲.۳) استفاده شد. تساوی (۱۳.۳) نشان می‌دهد برای هر $\theta_i\in\Theta_{۰}$، $\int g\,dP_{\theta_i}=۰$ در صورتی که تقریباً همه‌جا $g\neq۰$؛ به‌عبارت دقیق‌تر، $P_{\theta_i}(\{g\neq۰\})=P_{\theta_i}([۰,۱])=۱$ و لذا زیرخانواده $\mathcal{P}_{۰}=\{P_{\theta_{۰}}:\theta_{۰}\in\Theta_{۰}\}$ کامل نیست که تناقض با فرض است؛ پس خانواده $\{P_\theta:\theta\in[۰,۱]\}$ به‌طور شمارا کامل نیست.

قضیه ۱۱.۳. ([۱۷]) *اگر T یک آماره بسنده جفتی و به‌طور شمارا کامل برای خانواده $\mathcal{P}=\{P_\theta:\theta\in\Theta\}$ و U یک آماره کمکی باشد، آن‌گاه $T\perp\!\!\!\perp U$.*

اثبات. فرض کنید A یک مجموعه بورل در فضای مقادیر U باشد. چون U یک آماره کمکی است، $P_\theta(U\in A)$ به θ بستگی ندارد، لذا مشابه برهان سوم از قضیه ۱.۲، داریم $\mathbb{E}_\theta[P_\theta(U\in A|T)]=P_\theta(U\in A)$ و در پی آن $\mathbb{E}_\theta[P_\theta(U\in A|T)-P_\theta(U\in A)]=۰$. از طرف دیگر $g(T)=P_\theta(U\in A|T)-P_\theta(U\in A)$ یک آماره است، زیرا T یک آماره بسنده جفتی و U یک آماره کمکی هست. چون آماره T به‌طور شمارا کامل نیز هست، می‌توان نتیجه گرفت برای هر زیرمجموعه شمارای $\Theta_{۰}\subset\Theta$ و به‌ازای هر $\theta\in\Theta$، $g(T)=۰$ و لذا به‌ازای هر $\theta\in\Theta$، $P_\theta(U\in A|T)=P_\theta(U\in A)$ که این معادل استقلال T و U است. □

تعمیم‌های دیگری نیز از قضیه ۱.۲ وجود دارد که به ذکر آنها نمی‌پردازیم. به‌عنوان نمونه در [۱۴] مفهوم کامل بودن به کامل بودن شرطی نسبت به یک میدان سیگمایی تعمیم داده می‌شود و تعمیم قضیه ۱.۲ در این حالت مطرح می‌شود.

۴ بررسی استقلال از دیدگاه بیزی

باسو [۳] در سال ۱۹۸۲ خاطر نشان می‌کند که که اگر از دیدگاه بیزی به مسئله نگاه شود و Θ به‌عنوان یك متغیر تصادفی با توزیع پیشین ξ در نظر گرفته شود و مدل $\mathcal{P}=\{P_\theta:\theta\in\Theta\}$ به‌صورت مجموعه توزیع‌های شرطی $X|\Theta=\theta$ باشد، آن‌گاه مفهوم بسندگی و کمکی بودن مستقیماً به نظریه استقلال شرطی وابسته می‌شود [۹].

اکنون برای یک مدل $\mathcal{P}$ و برای هر توزیع پیشین ξ، توزیع احتمال توأم (X,Θ) در نظر گرفته می‌شود. در این چارچوب آماره U را کمکی گویند، اگر توزیع شرطی $U|\Theta=\theta$ به θ بستگی نداشته باشد که همان تعریف آماره کمکی مشروط، یعنی تعریف ۳.۳ است.

همچنین آماره T بسنده است، اگر توزیع شرطی X به شرط T و Θ فقط به T وابسته باشد. علاوه بر آن U کمکی است، اگر برای هر توزیع توأم Q_ξ از (X,Θ)، دو متغیر تصادفی U و Θ مستقل باشند و همچنین T بسنده است، اگر برای هر Q_ξ، X وΘ مستقل شرطی به شرط T باشند [۳، ۱۰]. تعریف بیزی بسندگی به‌طور مستقیم به نظریه بسندگی بیزی که در سال۱۹۴۲ توسط کولموگوروف [۱۶] ارائه شد، مربوط می‌شود [۶].

ارتباط بسندگی به‌طور معمول و بسندگی بیزی به این صورت است که اگر T به‌طور معمول بسنده باشد، آنگاه به‌ازای هر توزیع پیشین ξ برای متغیر تصادفی Θ، توزیع پسین $\Theta|X$ فقط از طریق $T(X)$ به X بستگی دارد و لذا با توجه به تعریف کولموگوروف [۱۶]، T بسنده بیزی است. برعکس، اگر تابع چگالی پیشین Θ به‌ازای هر θ مثبت باشد و توزیع شرطی $\Theta|X$ فقط به $T(X)$ بستگی داشته باشد، آنگاه T برای θ بسنده به‌طور معمول است [۱، ۴].

اکنون می‌توان قضیه ۱.۲ را با فرمول‌بندی بیزی به‌صورت زیر نوشت. (اثبات را برای حالتی که همگی متغیرها گسسته هستند مطرح می‌کنیم و برای حالت پیوسته قضیه به‌طور مشابه اثبات می‌شود).

قضیه ۱.۴. ([۴]) *فرض کنید برای هر Q_ξ، متغیرهای تصادفی U و Θ مستقل باشند (یعنی U در چارچوب بیزی، کمکی است) و متغیرهای تصادفی X و Θ مستقل شرطی به شرط T باشند (یعنی T در چارچوب بیزی، بسنده است)، آنگاه T و U مستقل شرطی به شرط Θ هستند، مشروط بر آن که T به‌طور کران‌دار کامل باشد.*

اثبات. برای هر پیشامد A، احتمال‌های زیر در نظر گرفته می‌شوند:

$$\eta_A(t,\theta)=P(U\in A\mid T=t,\Theta=\theta),\quad \lambda_A(\theta)=P(U\in A\mid \Theta=\theta). \tag{۱.۴}$$

با توجه به اینکه T یک آماره کمکی بیزی است (یعنی $U \perp\!\!\!\perp \Theta$)، احتمال شرطی داده شده در (۱.۴) به احتمال غیرشرطی $\lambda_A(\theta)=P(U\in A)$ بدل می‌شود که به θ بستگی ندارد. همچنین از اینکه T یک آماره بسنده بیزی است (یعنی $X \perp\!\!\!\perp \Theta \mid T$) و U تابعی از X است، می‌توان نتیجه گرفت که $U \perp\!\!\!\perp \Theta \mid T$ و در نتیجه:

$$\begin{aligned}\eta_A(t,\theta)&=\frac{P(U\in A,T=t,\Theta=\theta)}{P(T=t,\Theta=\theta)}\\&=\frac{P(U\in A,\Theta=\theta\mid T=t)P(T=t)}{P(\Theta=\theta\mid T=t)P(T=t)}\\&=\frac{P(U\in A\mid T=t)P(\Theta=\theta\mid T=t)}{P(\Theta=\theta\mid T=t)}\\&=P(U\in A\mid T=t)\end{aligned} \tag{۲.۴}$$

که در تساوی (۲.۴) از $U \perp\!\!\!\perp \Theta \mid T$ استفاده شده است. بنابراین $\eta_A(t,\theta)$ نیز به θ بستگی ندارد و بنابراین $\psi_A(T)=\eta_A(T,\theta)-\lambda_A(\theta)$ یك آماره است و لذا به‌ازای هر $\theta\in\Theta$ داریم:

$$\begin{aligned}\mathbb{E}_\theta(\psi_A(T))&=E_\theta(\eta_A(T,\theta))-\mathbb{E}_\theta(\lambda_A(\theta))\\&=\mathbb{E}_\theta(P(U\in A\mid T))-P(U\in A)\\&=\mathbb{E}_\theta\mathbb{E}(I_A(U)\mid T)-\mathbb{E}_\theta(I_A(U))=0\end{aligned} \tag{۳.۴}$$

که در تساوی آخر از ویژگی امید مکرر استفاده شده است.

حال با استفاده از کامل بودن آماره T، از تساوی (۳.۴) می‌توان نتیجه گرفت که با احتمال یک $\psi_A(T)=0$ یا $\eta_A(T,\theta)=\lambda_A(\theta)$ و یا $P(U\in A\mid T=t,\Theta=\theta)=P(U\in A\mid\Theta=\theta)$ که این معادل استقلال شرطی T و U به شرط Θ است (یعنی $T \perp\!\!\!\perp U \mid \Theta$). □

مثال ۲.۴. *فرض کنید* $\Theta \sim N(۰, \tau^۲)$ *و* $X_۱, X_۲ \mid \Theta \overset{indep}{\sim} N(\Theta, \sigma^۲)$ *. از دیدگاه بیزی،* $T = \frac{X_۱+X_۲}{۲}$ *آماره بسنده و* $U = X_۱ - X_۲$ *یک آماره کمکی است. به‌عبارت دیگر از دیدگاه بیزی داریم* $T \mid \Theta \sim N\left(\Theta, \frac{\sigma^۲}{۲}\right)$ *و* $U \mid \Theta \sim N(۰, ۲\sigma^۲)$ *.*

از اینکه T *به‌طور معمول بسنده است، توزیع پسین* $\Theta \mid X \sim N\left(\frac{\tau^۲}{\sigma^۲+\tau^۲}T, \left(\frac{۱}{\tau^۲} + \frac{۲}{\sigma^۲}\right)^{-۱}\right)$ *فقط از طریق* T *به نمونه بستگی دارد و لذا* T *بسنده بیزی است. آماره* T *به‌طور کران‌دار کامل نیز هست، لذا طبق قضیه ۱.۴ ،* T *و* U *مستقل شرطی به شرط* Θ *هستند که این حقیقت را به‌طور مستقیم نیز می‌توان نتیجه گرفت. داریم*

(۴.۴)

$$\begin{aligned}\mathbb{COV}(T, U \mid \Theta) &= \mathbb{COV}\left(\frac{X_۱ + X_۲}{۲}, X_۱ - X_۲ \mid \Theta\right) \\ &= \frac{۱}{۲}\left[\mathbb{COV}(X_۱, X_۱) - \mathbb{COV}(X_۱, X_۲) + Cov(X_۲, X_۱) - \mathbb{COV}(X_۲, X_۲)\right] \\ &= \frac{۱}{۲}\left[\sigma^۲ - ۰ + ۰ - \sigma^۲\right] = ۰.\end{aligned}$$

تساوی (۴.۴) نشان می‌دهد که $T \mid \Theta$ *و* $U \mid \Theta$ *ناهبسته شرطی هستند. اما نرمال بودن* $T \mid \Theta$ *و* $U \mid \Theta$ *تضمین می‌کند که* $U \perp\!\!\!\perp \Theta \mid T$ *.*

۵ استقلال آماره هم‌وردای بسنده و آماره ناوردا

در این بخش، قضیه ۱.۲ تحت یک گروه توپولوژیکی که روی متغیر تصادفی عمل می‌کند مورد بررسی قرار می‌گیرد و استقلال یک تابع ناوردا با یک آماره بسنده هم‌وردا اثبات می‌شود. در حقیقت، در قضیه ۱.۲ ، فرض بسندگی حفظ می‌شود، ولی فرض کامل بودن با هم‌وردایی و فرض کمکی بودن با ناوردایی جابه‌جا خواهد شد.

تابع $f : \mathcal{X} \to \mathcal{Y}$ را ناوردا $(G-$ناوردا) گوییم، هرگاه برای هر $g \in G$ و $x \in \mathcal{X}$، $f(gx) = f(x)$ و آن را ناوردای ماکسیمال گوییم، هرگاه ناوردا باشد و $f(x_۱) = f(x_۲)$ نتیجه دهد برای یک $g \in G$، $x_۲ = gx_۱$. تابع ناوردا روی مدارها ثابت است و تابع ناوردای ماکسیمال علاوه بر آن، مقادیر مختلف را به مدارهای مجزا اختصاص می‌دهد و لذا بهترین حالتی است که $\mathcal{X}$ به مدارهایی افراز می‌شود که تحت G معادل‌اند. تابع $f : \mathcal{X} \to \mathcal{Y}$ را هم‌وردا $(G-$هم‌وردا) گوییم، هرگاه برای هر $g \in G$ و $x \in \mathcal{X}$، $f(gx) = gf(x)$. گروه G به‌طور انتقالی روی فضای پارامتر Θ عمل می‌کند، هرگاه برای هر $\theta_۱, \theta_۲ \in \Theta$ یک عضو $g \in G$ وجود داشته باشد به‌طوری که $\theta_۲ = g\theta_۱$ [۱۹]. قابل ذکر است که مجموعه $\mathcal{X}$ می‌تواند زیرمجموعه $\mathbb{R}^n$ باشد و در این حالت باز عضو g روی $\mathbf{x}$ به‌صورت $g\mathbf{x}$ عمل خواهد کرد که در آن باید شرط $e\mathbf{x} = \mathbf{x}$ و $(g_۱g_۲)\mathbf{x} = g_۱(g_۲\mathbf{x})$ برای هر $\mathbf{x} \in \mathcal{X}$ و $g_۱, g_۲ \in G$ برقرار باشد که e عضو همانی گروه G است.

به‌طور شهودی اگر آماره هم‌وردای $\tau(\mathbf{X})$ بسنده نیز باشد، شامل تمام اطلاعات $\mathbf{X}$ درباره پارامتر است و لذا می‌توان راجع به استقلال آن از آماره کمکی $h(\mathbf{X})$ که توزیعش به پارامتر بستگی ندارد صحبت کرد. از طرف دیگر می‌دانیم اگر $h(\mathbf{X})$ یک تابع ناوردا باشد، توزیع آن یک تابع ناوردا از پارامتر خواهد بود و اگر گروه G روی فضای پارامتر Θ به‌طور انتقالی عمل کند، $h(\mathbf{X})$ کمکی است. با توجه به این که $h(\mathbf{X})$ ناورداست اگر و تنها اگر تابعی از ناوردای ماکسیمال باشد، پیشنهاد می‌شود که آماره ناوردای ماکسیمال از $\tau(\mathbf{X})$ مستقل است.

قضیه ۱.۵. *فضاهای اندازه‌پذیر* $(\mathcal{X}, \sigma(\mathcal{X}))$ *و* $(\mathcal{Y}, \sigma(\mathcal{Y}))$ *که به‌ترتیب گروه توپولوژیکی* G *ـ فشرده و موضعاً فشرده* G *روی آن‌ها عمل می‌کنند را در نظر بگیرید. فرض کنید* G *به‌طور انتقالی روی* $\mathcal{Y}$ *عمل کند. تابع* $\tau : \mathcal{X} \to \mathcal{Y}$ *اندازه‌پذیر و* $G-$*هم‌وردا، یک فضای اندازه‌پذیر و تابع* $h : \mathcal{X} \to \mathcal{Z}$ *اندازه‌پذیر و* $G-$*ناوردا باشد. برای متغیر تصادفی* $X \in \mathcal{X}$ *با* $\mathcal{L}(X) = P_o$ *(یعنی* X *دارای توزیع* P_o *است)، قرار دهید* $Z = h(X)$ *و فرض*

کنید $Y = \tau(X)$ یک آماره بسنده برای خانواده $\{gP_\circ : g \in G\}$ از توزیع‌ها روی $(\mathcal{X}, \sigma(\mathcal{X}))$ باشد. بنابراین اگر برای هر $g \in G$، $\mathcal{L}(X) = gP_\circ$، آن‌گاه $Z \perp\!\!\!\perp Y$.

اثبات. فرض کنید $X \in \mathcal{X}$ دارای توزیع $P_\circ$ و $Q_\circ$ توزیع القایی آن باشد. برای یک تابع اندازه‌پذیر کران‌دار مثل f روی $\mathcal{Z}$، امید شرطی $H_g(y) = \mathbb{E}_{gP_\circ}(f(h(X)) \mid \tau(X) = y)$ به‌خاطر کران‌دار بودن تابع f، خوش‌تعریف است. از اینکه $\tau(X)$ یک آماره بسنده برای خانواده $\{gP_\circ : g \in G\}$ است، یک تابع اندازه‌پذیر H روی $\mathcal{Y}$ وجود دارد به‌طوری‌که برای هر $g \in G$ و $y \notin N_g$ ، $H_g(y) = H(y)$ که در آن N_g یک مجموعه با اندازه صفر نسبت به تابع احتمال $gQ_\circ$ است، یعنی $gQ_\circ(N_g) = Q_\circ(g^{-1}N_g) = 0$. همچنین تقریباً همه‌جا یک تابع $H_e(y)$ وجود دارد که برای هر تابع اندازه‌پذیر و کران‌دار k در معادله زیر صادق می‌کند:

$$\begin{aligned}\int_{\mathcal{Y}} k(y)H_e(y)Q_\circ(dy) &= \mathbb{E}_{Q_\circ}(k(Y)H_e(Y)) \\ &= \mathbb{E}_{Q_\circ}\mathbb{E}_{P_\circ}(k(Y)f(h(X)) \mid \tau(X) = y) \\ &= \mathbb{E}_{P_\circ}(k(\tau(X))f(h(X))) \\ &= \int_{\mathcal{X}} k(\tau(x))f(h(x))P_\circ(dx).\end{aligned} \tag{۱.۵}$$

اندازه احتمال $gP_\circ$، برای هر تابع اندازه‌پذیر و کران‌دار l در معادله زیر صادق می‌کند:

$$\int_{\mathcal{X}} l(x)gP_\circ(dx) = \int_{\mathcal{X}} l(gx)P_\circ(dx). \tag{۲.۵}$$

با توجه به این که τ، $G-$هم‌وردا‌ست، توزیع $\tau(X)$، $gQ_\circ$ خواهد بود، یعنی $\mathcal{L}(\tau(X)) = gQ_\circ$ که در آن برای $g \in G$، $\mathcal{L}(X) = gP_\circ$. با توجه به این حقیقت و ناوردایی $f \circ h$، برای هر تابع کران‌دار k داریم:

$$\begin{aligned}\int_{\mathcal{Y}} k(y)H_e(y)Q_\circ(dy) &= \int_{\mathcal{X}} k(\tau(x))f(h(x))P_\circ(dx) \\ &= \int_{\mathcal{X}} k(g^{-1}\tau(gx))f(h(gx))P_\circ(dx) \\ &= \int_{\mathcal{X}} k(g^{-1}\tau(x))f(h(x))gP_\circ(dx) \\ &= \int_{\mathcal{Y}} k(g^{-1}y)H_g(y)(gQ_\circ)(dy) \\ &= \int_{\mathcal{Y}} k(y)H_g(gy)Q_\circ(dy)\end{aligned} \tag{۳.۵}$$

که در تساوی اول از (۱.۵)؛ در تساوی دوم از ویژگی $G-$هم‌وردایی τ و $G-$ناوردایی h؛ در تساوی سوم از (۲.۵)؛ در تساوی چهارم از (۱.۵) و در تساوی پنجم از (۲.۵) استفاده شده است. بنابراین معادله (۳.۵) نتیجه می‌دهد، تقریباً همه‌جا $H_e(y) = H_g(gy)$ $Q_\circ$. اکنون، از اینکه تقریباً همه‌جا $Q_\circ$، $H(y) = H_e(y)$ و برای هر $g \in G$ که در آن $\alpha(g)y \notin N_g$ ، داریم $H_g(gy) = H(gy)$، می‌توان نتیجه گرفت که تقریباً همه‌جا $Q_\circ$، $H_g(gy) = H(gy)$ و این نتیجه می‌دهد تقریباً همه‌جا $Q_\circ$، $H(y) = H(gy)$. لذا یک تابع اندازه‌پذیر $G-$ناوردا مثل $\tilde{H}$ وجود دارد به‌طوری که تقریباً همه‌جا $Q_\circ$، $H = \tilde{H}$. اما از اینکه G روی $\mathcal{Y}$ به‌صورت انتقالی

عمل می‌کند، $\tilde{H}$ باید ثابت باشد و در پی آن تقریباً همه‌جا $Q_\circ$، H ثابت است. پس، تقریباً همه‌جا $Q_\circ$،

$$H_e(y) = E_{P_\circ}(f(h(X))\,|\tau(X) = y) \tag{۴.۵}$$

ثابت است. اگر تابع k روی $(\mathcal{Y}, \sigma(\mathcal{Y}))$ کران‌دار باشد و برای هر $y \in \mathcal{Y}$،

$$H(y) = H(y_\circ) \tag{۵.۵}$$

داریم

$$\begin{aligned}
\mathbb{E}_{P_\circ}\left[k\left(\tau(X)\right) f\left(h(X)\right)\right] &= \mathbb{E}_{P_\circ}\mathbb{E}_{Q_\circ}\left[k\left(\tau(X)\right) f\left(h(X)\right) \mid \tau(X)\right] \\
&= \int_{\mathcal{Y}} \mathbb{E}_{P_\circ}\left[k\left(\tau(X)\right) f\left(h(X)\right) \mid \tau(X) = y\right] Q_\circ(dy) \\
&= \int_{\mathcal{Y}} k(y)\mathbb{E}_{P_\circ}\left[f\left(h(X)\right) \mid \tau(X) = y\right] Q_\circ(dy) \\
&= \int_{\mathcal{Y}} k(y)H(y)Q_\circ(dy) \\
&= H(y_\circ)\int_{\mathcal{Y}} k(y)Q_\circ(dy) \\
&= \int_{\mathcal{Y}} H(y)Q_\circ(dy)\cdot \mathbb{E}_{P_\circ}\left[k\left(\tau(X)\right)\right] \\
&= \mathbb{E}_{P_\circ}\left[f\left(h(X)\right)\right]\cdot \mathbb{E}_{P_\circ}\left[k\left(\tau(X)\right)\right]
\end{aligned} \tag{۶.۵}$$

که در تساوی اول از ویژگی امید مکرر؛ در تساوی دوم از تعریف امید ریاضی؛ در تساوی سوم از ویژگی خطی‌بودن امید ریاضی؛ در تساوی چهارم از رابطه (۴.۵)؛ در تساوی پنجم از رابطه (۵.۵)؛ در تساوی ششم از رابطه (۱.۵) و تعریف امید ریاضی و در تساوی هفتم نیز از رابطه (۱.۵) و تعریف امید ریاضی استفاده شده است.

تساوی (۶.۵) نشان می‌دهد که برای متغیر تصادفی $X \in \mathcal{X}$ با $\mathcal{L}(X) = P_\circ$، آماره‌های $Z = h(X)$ و $Y = \tau(X)$ مستقل هستند. وقتی $\mathcal{L}(X) = \tilde{P}_\circ = g_1 P_\circ$ داریم $\{gP_\circ : g \in G\} = \{g\tilde{P}_\circ : g \in G\}$ و در پی آن $Y = \tau(X)$ برای خانواده $\{g\tilde{P}_\circ : g \in G\}$ بسنده است. استدلالی که برای $P_\circ$ انجام شد را برای $\tilde{P}_\circ$ نیز می‌توان انجام داد. بنابراین برای متغیر تصادفی $X \in \mathcal{X}$ با $\mathcal{L}(X) = g_1 P_\circ$، آماره‌های $Z = h(X)$ و $Y = \tau(X)$ مستقل هستند. □

مثال ۲.۵. *فرض کنید $X_1, \ldots, X_n$ یک نمونه تصادفی از توزیع $N(\mu, \sigma^2)$ باشد که در آن $\mu \in \mathbb{R}$ و $\sigma > 0$ نامعلومند. عمل گروه توپولوژیکی (گروه آفین یا گروه مکان‌ـ مقیاس) را به‌صورت $G = \{(a, b) : a > 0, b \in \mathbb{R}\}$ تعریف کنید که در آن برای هر $(a, b) \in G$ و $\mathbf{x} = (x_1, \ldots x_n)$، $(a, b)\mathbf{x} = a\mathbf{x}+b = (ax_1+b, \ldots ax_n+b)$ و عمل گروه القایی $(a, b)\cdot(\mu, \sigma^2) = (a\mu + b, a^2\sigma^2)$ است. آماره بسنده $\tau(\mathbf{X}) = (S, \bar{X})$ هم‌وردا هست*

(که در آن $S^۲ = \sum_{i=۱}^{n}(X_i - \bar{X})^۲$)، زیرا داریم:

$$\begin{aligned}\tau((a,b)\mathbf{X}) &= \tau((aX_۱ + b, \ldots, aX_n + b)) \\ &= \left(\sqrt{\sum_{i=۱}^{n}(aX_i + b - a\overline{X} - b)^۲},\ a\overline{X} + b\right) \\ &= (aS,\ a\overline{X} + b) \\ &= (a,b)(S,\ \overline{X}) \\ &= (a,b)\tau(\mathbf{X})\end{aligned} \tag{۷.۵}$$

که عمل گروه روی فضای برد آماره $\tau(\mathbf{X})$ به‌صورت $(a,b)(\delta_۱,\delta_۲) = (a\delta_۱,\ a\delta_۲ + b)$ تعریف شده است. در حقیقت رابطه (۷.۵) نشان می‌دهد، برای هر عضو $g = (a,b) \in G$، $\tau(g\mathbf{X}) = g\tau(\mathbf{X})$ که نشان می‌دهد، آماره $\tau(\mathbf{X})$، هم‌وردا نیز هست.

آماره $h(\mathbf{X}) = (\tau(\mathbf{X}))^{-۱}\mathbf{X} = \frac{\mathbf{X}-\overline{X}}{S}$ ناوردای ماکسیمال است، زیرا:

$$\begin{aligned}h(g\mathbf{X}) &= h((a,b)\mathbf{X}) \\ &= (\tau((a,b)X))^{-۱}(a,b)\mathbf{X} \\ &= ((a,b)\tau(\mathbf{X}))^{-۱}(a,b)\mathbf{X} \\ &= \tau(\mathbf{X})^{-۱}(a,b)^{-۱}(a,b)\mathbf{X} \\ &= \tau(\mathbf{X})^{-۱}\mathbf{X} \\ &= h(\mathbf{X})\end{aligned} \tag{۸.۵}$$

که (۸.۵) نشان می‌دهد، $h(\mathbf{X})$ ناوردا هست؛ همچنین اگر برای دو نمونه مختلف مثل $\mathbf{X}$ و $\mathbf{Y}$، $h(\mathbf{X}) = h(\mathbf{Y})$، آنگاه $(\tau(\mathbf{X}))^{-۱}\mathbf{X} = (\tau(\mathbf{Y}))^{-۱}\mathbf{Y}$ که با ضرب دو طرف تساوی از سمت چپ در $\tau(\mathbf{Y})$ داریم $\tau(\mathbf{Y})(\tau(\mathbf{X}))^{-۱}\mathbf{X} = \mathbf{Y}$، یعنی وجود دارد $g = \tau(\mathbf{Y})(\tau(\mathbf{X}))^{-۱}$ که $\mathbf{Y} = g\mathbf{X}$ و لذا $\mathbf{X}$ و $\mathbf{Y}$ در یک مدار یکسان قرار دارند. بنابراین $h(\mathbf{X})$ ناوردای ماکسیمال است و برچسب‌هایی که روی افرازهای ناوردا می‌سازد یکتا است. لازم به ذکر است که در این حالت، آماره $h(\mathbf{X}) = \frac{\mathbf{X}-\bar{X}}{S}$ کمکی نیز هست، زیرا:

$$\begin{aligned}h(\mathbf{X}) &= \frac{\mathbf{X} - \bar{X}}{S} \\ &= \left(\frac{X_۱ - \bar{X}}{S}, \ldots, \frac{X_n - \bar{X}}{S}\right) \\ &= \left(\frac{\frac{X_۱-\mu}{\sigma} - \frac{۱}{n}\sum_{i=۱}^{n}\frac{X_i-\mu}{\sigma}}{\sqrt{\sum_{i=۱}^{n}\left(\frac{X_i-\mu}{\sigma} - \frac{\bar{X}-\mu}{\sigma}\right)^۲}}, \ldots, \frac{\frac{X_n-\mu}{\sigma} - \frac{۱}{n}\sum_{i=۱}^{n}\frac{X_i-\mu}{\sigma}}{\sqrt{\sum_{i=۱}^{n}\left(\frac{X_i-\mu}{\sigma} - \frac{\bar{X}-\mu}{\sigma}\right)^۲}}\right) \\ &= \left(\frac{Z_۱ - \bar{Z}}{\sqrt{\sum_{i=۱}^{n}(Z_i - \bar{Z})^۲}}, \ldots, \frac{Z_n - \bar{Z}}{\sqrt{\sum_{i=۱}^{n}(Z_i - \bar{Z})^۲}}\right)\end{aligned} \tag{۹.۵}$$

که در آن $Z_i = \frac{X_i-\mu}{\sigma} \sim N(۰,۱)$، $i = ۱, \ldots, n$ متغیرهای تصادفی مستقل و هم‌توزیع هستند. با توجه به اینکه

Z_iها به پارامترهای مجهول یعنی $\mu \in \mathbb{R}$ و $\sigma > 0$ بستگی ندارند، طبق ساختار (۹.۵) آماره $h(\mathbf{X})$ کمکی نیز هست.

با توجه به انتقالی بودن گروه، هر تابع ناوردا مثل $h(\mathbf{X})$ کمکی نیز هست که در این مثال خاص با توجه به اینکه آماره بسنده هم‌وردای $\tau(\mathbf{X})$، کامل نیز هست، استقلال آن از آماره ناوردای $h(\mathbf{X})$ که کمکی نیز هست، واضح است. ولی با توجه به برقراری شرایط قضیه ۱.۵، در حالت کلی‌تر با نداشتن فرض کامل بودن $\tau(\mathbf{X})$ و کمکی بودن $h(\mathbf{X})$ نیز توانستیم استقلال آماره بسنده هم‌وردا و یک آماره ناوردا را نتیجه بگیریم.

لازم به ذکر است که آماره‌های ناوردا تحت شرایط خاص، کمکی نیز هستند [۲۰]، ولی عکس این موضوع صحیح نیست که در زیر مثال نقضی برای آن ارائه می‌شود.

مثال ۳.۵. فرض کنید $X_1, X_2 \overset{i.i.d.}{\sim} N(\mu, 1)$. آماره $U = U(X_1, X_2)$ را به‌صورت زیر تعریف می‌کنیم:

$$U = \begin{cases} X_1 - X_2 & X_1 + X_2 \geq 1 \\ X_2 - X_1 & X_1 + X_2 < 1 \end{cases} \tag{۱۰.۵}$$

آماره U تحت گروه جمعی یا مکانی ناوردا است، زیرا برای عمل گروه انتقال $g_\alpha(X_1, X_2) = (X_1 + \alpha, X_2 + \alpha)$، $\alpha \in \mathbb{R}$ داریم $U(g_\alpha(X_1, X_2)) = U(X_1 + \alpha, X_2 + \alpha) \neq U(X_1, X_2)$، ولی U یک آماره کمکی است، زیرا توزیع آن به پارامتر مجهول μ بستگی ندارد. برای دیدن این حقیقت، با استفاده از اینکه $Y_1 = X_1 - X_2 \sim N(0, 2)$ و $Y_2 = X_1 + X_2 \sim N(2\mu, 1)$، آماره U در (۱۰.۵) به‌صورت زیر بازنویسی می‌شود:

$$U = \begin{cases} Y_1 & Y_2 \geq 1 \\ -Y_1 & Y_2 < 1 \end{cases} \tag{۱۱.۵}$$

با توجه به نرمال بودن متغیرهای تصادفی Y_1 و Y_2 و ناهمبستگی آن‌ها، یعنی $\mathbb{COV}(Y_1, Y_2) = 0$، استقلال Y_1 و Y_2 نتیجه می‌شود که با استفاده کامل بودن قضیه ۱.۲ از بسنده کامل بودن Y_2 و کمکی بودن Y_1 نیز می‌توان به همین نتیجه رسید. اکنون تابع توزیع متغیر تصادفی U بازنویسی شده در (۱۱.۵) را محاسبه می‌کنیم:

$$\begin{aligned} F_U(t) &= P(U \leq t) \\ &= P(Y_2 \geq 1 \mid Y_1 \leq t)P(Y_1 \leq t) + P(Y_2 < 1)P(-Y_1 \leq t \mid Y_2 < 1) \\ &= P(Y_2 \geq 1)P(Y_1 \leq t) + P(Y_2 < 1)P(-Y_1 \leq t) \\ &= P(Y_2 \geq 1)P(Y_1 \leq t) + P(Y_2 < 1)P(Y_1 \leq t) \\ &= F_{Y_1}(t) \end{aligned} \tag{۱۲.۵}$$

که در تساوی دوم از قانون احتمال کل؛ در تساوی سوم از استقلال Y_1 و Y_2 و در تساوی چهارم از هم‌توزیع بودن Y_1 و $-Y_1$ استفاده شده است. لذا (۱۲.۵) نشان می‌دهد $U \sim N(0, 2)$ و لذا U یک آماره کمکی است.

ملاحظه ۴.۵. از اینکه G به‌طور انتقالی روی $\{gP_o : g \in G\}$ عمل می‌کند و $Z = h(\mathbf{X})$، Gـ ناوردا‌ست، توزیع آن تحت هر gP_o (برای هر $g \in G$) یکسان است و لذا $Z = h(\mathbf{X})$ آماره کمکی است. بنابراین طبق قضیه ۱.۲، استقلال آن از آماره بسنده کامل نتیجه می‌شود. در قضیه ۱.۵، فرض کامل بودن برای برقراری قضیه ۱.۲، با هم‌وردایی جابه‌جا شده است.

مثال ۵.۵. *فرض کنید $X_1,\dots,X_n$ یک نمونه تصادفی از توزیع یکنواخت $U(0,\theta)$، $\theta>0$ باشد. آماره بسنده $\tau(\mathbf{X})=X_{(n)}$ هموردا نسبت به گروه ضربی (گروه مقیاس) $g_c(x_1,\dots,x_n)=(cx_1,\dots,cx_n)$، $c>0$، است، زیرا $\tau(g_c\mathbf{X})=\tau(cX_1,\dots,cX_n)=cX_{(n)}=c\tau(\mathbf{X})$ آماره $h(\mathbf{X})=\left(\frac{X_1}{X_{(n)}},\dots,\frac{X_{n-1}}{X_{(n)}}\right)$ تحت گروه ضربی، ناوردا است، زیرا $h(g_c\mathbf{X})=h(cX_1,\dots,cX_n)=\left(\frac{cX_1}{cX_{(n)}},\dots,\frac{cX_{n-1}}{cX_{(n)}}\right)=h(\mathbf{X})$. لذا طبق قضیه ۱.۵، $h(\mathbf{X})$ از $\tau(\mathbf{X})$ مستقل است. لازم به ذکر است که در این حالت کامل بودن و کمکی بودن در قضیه ۱.۲ به‌ترتیب نقششان را به هم‌وردایی و ناوردایی می‌دهند.*

لم ۶.۵. *اگر $f:\mathcal{X}\to\mathcal{Y}$ یک‌به‌یک و $f:\mathcal{Y}\to\mathcal{Z}$ ناوردای ماکسیمال باشد، آنگاه $k=f\circ h:\mathcal{Z}\to\mathcal{Y}$ نیز ناوردای ماکسیمال است.*

اثبات. برای هر $g\in G$ و $x\in\mathcal{X}$ داریم $k(gx)=f(h(gx))=f(h(x))=k(x)$ و بنابراین k ناورداست. همچنین اگر $k(x_1)=k(x_2)$، آنگاه $f(h(x_1))=f(h(x_2))$ و با توجه به یک‌به‌یک بودن f، ناوردای ماکسیمال بودن h وجود دارد $g\in G$ به‌طوری که $x_1=gx_2$. بنابراین h ناوردای ماکسیمال است. □

مثال ۷.۵. *فرض کنید $X_1,\dots,X_n$ نمونه‌های تصادفی مستقل و هم‌توزیع از توزیع $\Gamma(a,b)$ با تابع چگالی احتمال $f(x)=\frac{x^{a-1}e^{-x/b}}{\Gamma(a)b^a}$، $x>0$ باشد که $a,b>0$. فضای نمونه $\mathcal{X}=(\mathbb{R}^+)^n$ و گروه توپولوژیکی $G=\mathbb{R}^+$ را با عمل $g_c(x_1,\dots,x_n)=(cx_1,\dots,cx_n)$، $c>0$ در نظر بگیرید.*

آماره $\tau(\mathbf{X})=\sum_{i=1}^n X_i$ یک آماره بسنده کامل برای پارامتر b بوده که تحت عمل گروه G هم‌وردا نیز هست، زیرا $\tau(g_c\mathbf{X})=\sum_{i=1}^n cX_i=c\tau(\mathbf{X})$. آماره ناوردای ماکسیمال به‌صورت $h(\mathbf{X})=\left(\frac{X_1}{\sum_{i=1}^n X_i},\dots,\frac{X_n}{\sum_{i=1}^n X_i}\right)$ است، زیرا $h(g_c\mathbf{X})=\left(\frac{cX_1}{c\sum_{i=1}^n X_i},\dots,\frac{cX_n}{c\sum_{i=1}^n X_i}\right)=h(\mathbf{X})$ نشان‌دهنده ناوردایی آن هست و اگر برای دو نمونه مختلف مثل $\mathbf{X}$ و $\mathbf{Y}$، $h(\mathbf{X})=h(\mathbf{Y})$، آنگاه وجود دارد $c=\frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n X_i}$ که $Y=g_c\mathbf{X}$ که نشان می‌دهد $\mathbf{X}$ و $\mathbf{Y}$ در یک مدار یکسان قرار دارند و $h(\mathbf{X})$ ناوردای ماکسیمال است و برچسب‌هایی که روی افرازه‌های ناوردا می‌سازند یکتا است. با استفاده از قضیه ۱.۵، آماره ناوردای ماکسیمال $h(\mathbf{X})$ از آماره بسنده هم‌وردای $\tau(\mathbf{X})$ مستقل است.

با در نظر گرفتن تابع یک‌به‌یک $g(a_1,\dots,a_n)=\left(\frac{a_1\sum_{i=1}^n a_i}{a_n},\dots,\frac{a_{n-1}\sum_{i=1}^n a_i}{a_n},\frac{a_n\sum_{i=1}^n a_i}{|a_n|}\right)$ و استفاده از لم ۶.۵، آماره $g\circ h(\mathbf{X})=\left(\frac{X_1}{X_n},\dots,\frac{X_{n-1}}{X_n},\frac{X_n}{|X_n|}\right)$ و در پی آن $k(\mathbf{X})=\left(\frac{X_1}{X_n},\dots,\frac{X_{n-1}}{X_n}\right)$ نیز ناوردای ماکسیمال بوده و طبق قضیه ۱.۵ آماره ناوردای ماکسیمال $k(\mathbf{X})$ از آماره بسنده هم‌وردای $\tau(\mathbf{X})$ مستقل است.

همان‌طور که در مثال ۷.۵ مشاهده می‌شود، با توجه به کامل بودن برآوردگر هم‌وردا، یعنی برآوردگر $\tau(\mathbf{X})$، قضیه‌های ۱.۲ و ۱.۵ هر دو به نتیجه یکسانی منجر می‌شوند که اولی از کامل بودن آماره بسنده $\tau(\mathbf{X})$ و کمکی بودن آماره $h(\mathbf{X})$ و دومی از هم‌وردایی آماره بسنده $\tau(\mathbf{X})$و ناوردایی آماره $h(\mathbf{X})$ استفاده می‌کند. مثال زیر نشان می‌دهد که قضیه ۱.۵ حتی در مواقعی که قضیه ۱.۲ قابل اعمال نیست نیز می‌تواند مورد استفاده قرار گیرد. یعنی حالت‌های خاصی هستند که آماره بسنده هم‌وردا، ویژگی کامل بودن را ندارد یا اینکه آماره ناوردا، یک آماره کمکی نباشد و در این حالت‌ها، قضیه ۱.۵ می‌تواند ابزار قوی‌تری نسبت به که قضیه ۱.۲ برای استقلال آماره بسنده هم‌وردا و آماره ناوردا ارائه دهد.

مثال ۸.۵. *فرض کنید X یک متغیر تصادفی گسسته با نیکه‌گاه $\mathcal{X}=D=\{\pm1,\pm2,\pm3,\pm4,\pm5\}$ و تابع*

احتمال زیر باشد:

$$P_\theta(X=x)=\begin{cases}\theta^{۲}(۱-\theta)/۲ & x=\pm ۵\\ \theta(۱-\theta)^{۲}/۲ & x=\pm ۴\\ \theta^{۳}/۲ & x=\pm ۳\\ (۱-\theta)^{۳}/۲ & x=\pm ۲\\ \theta(۱-\theta)/۲ & x=\pm ۱\end{cases} \tag{۱۳.۵}$$

به‌راحتی ثابت می‌شود، تمام توابع کمکی برای تابع احتمال داده شده در (۱۳.۵) به‌صورت زیر هستند

$$h_۰(a)=\begin{cases}u_۱ & a\in D'\\ u_۲ & a\in D\setminus D'\end{cases} \tag{۱۴.۵}$$

که $D'=\{d_۱,d_۲,d_۳,d_۴,d_۵\}\subseteq D$ به‌گونه‌ای انتخاب شده که $\pm a$ به‌طور هم‌زمان در D' نباشند. برای اینکه آماره دلخواه $h_۰$ کمکی باشد، باید برای هر u $P_\theta(h_۰=u)$ به پارامتر θ بستگی نداشته باشد. از طرف دیگر

$$\begin{aligned}P_\theta(h_\theta=u)&=\sum\nolimits_{i\in D'}P_\theta(X=i)\\ &=\sum\nolimits_{i\in D'}A_iP_\theta(X=i)\\ &=\tfrac{۱}{۲}\theta^{۲}(۱-\theta)B_۵+\tfrac{۱}{۲}\theta(۱-\theta)^{۲}B_۴+\tfrac{\theta^{۳}}{۲}B_۳+\tfrac{(۱-\theta)^{۳}}{۲}B_۲+\theta(۱-\theta)B_۱\\ &=\left[\frac{B_۴-B_۵}{۲}+\frac{B_۲-B_۳}{۲}\right]\theta^{۳}+\left[\frac{B_۵}{۲}-B_۴+\frac{۳B_۲}{۲}-\frac{B_۱}{۲}\right]\theta^{۲}\\ &\quad+\left[\frac{B_۴}{۲}-\frac{۳B_۲}{۲}+\frac{B_۱}{۲}\right]\theta+\frac{B_۲}{۲}\end{aligned} \tag{۱۵.۵}$$

که ضرایب A_i مقادیر ۰ و ۱ را به ترتیب برای حالت $i\in D'$ و $i\in D\setminus D'$ اختیار می‌کنند و $B_i=A_i+A_{i+۱}$. بنابراین، تابع جرم احتمال (۱۵.۵) به پارامتر θ بستگی ندارد اگر و تنها اگر:

$$\begin{cases}B_۴-B_۵=B_۲-B_۳\\ B_۵-۲B_۴+B_۱=۳(B_۱-B_۲)\\ B_۱-B_۴=۳(B_۱-B_۲)\end{cases} \tag{۱۶.۵}$$

به‌راحتی دیده می‌شود که تنها جواب معادله (۱۶.۵)، $B_۱=B_۲=B_۳=B_۴=B_۵=\delta$ است. با توجه به اینکه A_iها فقط مقادیر ۰ و ۱ را اختیار می‌کنند، $\delta\in\{۰,۱,۲\}$. اکنون مسئله را در هر یک از این سه حالت بررسی می‌کنیم:

حالت اول: اگر $\delta=۰$، در این صورت برای هر i داریم $A_i=۰$ و لذا $D'=\varnothing$ که غیر ممکن است.

حالت دوم: اگر $\delta=۲$، در این صورت برای هر i داریم $A_i=۱$ و در این حالت آماره کمکی، آماره بدیهی است، یعنی برای هر $a\in D$ داریم $h_۰(a)=u$.

حالت سوم: اگر $\delta=۱$، در این صورت برای هر i، یکی از حالت‌های $A_i=۰\wedge A_{i+۱}=۱$ یا

$A_i = ۱ \wedge A_{i+۱} = ۰$ رخ می‌دهد.

به‌عبارت دیگر D' باید زیرمجموعه‌های پنج عضوی از D را انتخاب کند که $+a$ و $-a$ را به‌طور همزمان نداشته باشد. لذا تنها آماره‌های کمکی به شکل $h_۰$ ارائه شده در (۱۴.۵) هستند که آماره بدیهی (حالت دوم) را نیز شامل می‌شوند. $h_۰$ به‌گونه‌ای تعریف شود که $h_۰(-۱) = h_۰(-۲) = h_۰(-۳) = h_۰(-۴) = h_۰(-۵) = ۰$ و $h_۰(۱) = h_۰(۲) = h_\theta(۳) = h_۰(۴) = h_۰(۵) = ۱$ نتیجه می‌شود که $P(h_۰ = ۱) = P(h_۰ = ۰) = \frac{۱}{۲}$ و در پی آن $h_۰$ یک آماره کمکی است.

اکنون اگر $G = \{g_۱, g_۲\}$ روی D عمل کند که در آن $g_۱(x) = x$ و $g_۲(x) = -x$، یک گروه روی فضای پارامتر $\Theta = (۰, ۱)$ به‌صورت $\theta \mapsto g(\theta) = \theta$ برای هر $\theta \in \Theta$ و $g \in G$ القا می‌کند. به‌وضوح مدل تحت این گروه ناوردا‌ست. آماره‌های کمکی بدیهی همگی $G-$ناوردا هستند. بنابراین اگر برای هر $a \in D$، $h(a) = v$، $h(X)$ یک آماره کمکی است که $G-$ناوردا نیز هست.

به‌راحتی ثابت می‌شود $\tau(X) = |X|$ برای θ بسنده مینیمال است، اما کامل نیست. بنابراین در این حالت نمی‌توان از قضیه ۱.۲ استفاده کرد. اما $\tau(X)$، $G-$هم‌وردا‌ست که $\bar{G} = \{g_۱\} \leq G$ (نماد $\leq$، نشان‌دهنده زیرگروه است) روی $D_۱ = \{۱, ..., ۵\}$، به‌طور بدیهی عمل می‌کند (در حقیقت $\tau(X)$، Gـ ناوردا‌ست). از اینکه برای هر $t = ۱, \ldots, ۵$، $P(h(X) = v|\tau(X) = t) = P(h(X) = v) = \frac{۱}{۲}$ نتیجه گرفته می‌شود که $h(X)$ و $\tau(X)$ مستقل‌اند که این نتیجه از قضیه ۱.۵ نیز به‌دست می‌آید. به‌طور مشابه $U(X) = X$، $G-$هم‌وردا‌ست و لذا با استفاده از قضیه ۱.۵، U از $h(X)$ مستقل است.

۶ بحث و نتیجه‌گیری

در این مقاله به بررسی استقلال دو آماره، بدون محاسبه توزیع توأم آنها پرداخته شد. اولین نسخه توسط باسو [۲] معرفی شد که در آن استقلال آمار بسنده کامل با آمار کمکی را تضمین می‌کرد. حالت دیگری توسط جرگنسن [۱۵] به این صورت مطرح شد که اگر آماره (T, V) بسنده و توزیع شرطی $T \mid V$ کامل و (V, U) یك آماره کمکی باشند، آنگاه $T \perp\!\!\!\perp U \mid V$. حالت دیگر بیان می‌کند، اگر آماره (T, V) بسنده کامل و U یك آماره کمکی مشروط به شرط V باشد، آنگاه $T \perp\!\!\!\perp U \mid V$. در [۱۷] ثابت شد، اگر T یك آماره بسنده جفتی و به‌طور شمارا کامل برای خانواده $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ و V یك آماره کمکی باشد، آنگاه $T \perp\!\!\!\perp V$. در نسخه بیزی قضیه باسو [۴]، فرض می‌شود برای هر Q_ξ، $U \perp\!\!\!\perp \Theta$ (یعنی در چارچوب بیزی، U یک آماره کمکی است) و $X \perp\!\!\!\perp \Theta \mid T$ (یعنی، در چارچوب بیزی، T بسنده است) و همچنین T به‌طور کران‌دار کامل است و نتیجه گرفته می‌شود $T \perp\!\!\!\perp U \mid \Theta$.

در پایان قضیه ۱.۵ برای حالتی که گروه توپولوژیکی روی فضای متغیر تصادفی عمل می‌کند، مطرح و ثابت شد زمانی که گروه روی فضا به‌صورت انتقالی عمل می‌کند، آماره بسنده هم‌وردا از آماره ناوردا مستقل است. در مثال‌هایی که گروه انتقالی است، هر تابع ناوردا (که کمکی نیز هست) از آماره بسنده هم‌وردای مستقل است، اما در بیشتر مثال‌ها مشاهده می‌شود که آماره بسنده هم‌وردا، کامل نیز هست و لذا قضیه ۱.۲ استقلال آن از آماره ناوردا که کمکی نیز هست را تضمین می‌کند. ولی با توجه به برقراری شرایط قضیه ۱.۵ در حالت کلی‌تر با نداشتن فرض کامل بودن آماره بسنده و همچنین کمکی بودن آماره دیگر نیز توانستیم استقلال آماره بسنده هم‌وردا را از یک آماره ناوردا نتیجه بگیریم. در انتها نشان دادیم که قضیه ۱.۵ حتی در مواقعی که قضیه ۱.۲ قابل اعمال نیست نیز می‌تواند مورد استفاده قرار گیرد. یعنی حالت‌هایی هستند که آماره بسنده هم‌وردا، ویژگی کامل بودن را ندارد یا اینکه آماره ناوردا، یک آماره کمکی نیست، در اینجا قضیه ۱.۵ می‌تواند ابزار قوی‌تری نسبت به قضیه ۱.۲ برای استقلال آماره بسنده هم‌وردا و آماره ناوردا ارائه دهد. قضیه ۱.۲، آنقدر قابل انعطاف است که محققان علاقه‌مند می‌توانند تعمیم‌های دیگری از آن به‌دست آورند.

مراجع

[1] Arnold, S.F., 1988. Sufficient Statistics. *Encyclopedia of Statistical Sciences*, 9, pp.72-80.

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

[3] Basu, D., 1981. On Ancillary Statistics, Pivotal Quantities and Confidence Statements. *Topics in Applied Statistics*, pp.1-29.

[4] Basu, D. and Pereira, C.A.B., 1983. Conditional Independence In Statistics. *Sankhya, A*, 45, pp.324-337.

[5] Boos, D.D. and Hughes-Oliver, J.M., 1998. Applications of Basu's Theorem. *The American Statistician*, 52(3), pp.218-221.

[6] Casella, G. and Berger, R., 2001. Statistical Inference. 2nd edition. Pacific Groves: Wadsworth.

[7] David, H.A., 1995. First Occurrence of Common Terms in Mathematical Statistics History. *The American Statistician*, 49(1), pp.121-133.

[8] David, H.A., 1998. First Occurrence of Common Terms in Probability and Statistics, A Second List, with Corrections. *Journal of the American Statistical Association*, 93(441), pp.36-40.

[9] Dawid, A.P., 1979. Conditional Independence in Statistical Theory. *Journal of the Royal Statistical Society: Series B (Methodological)*, 41(1), pp.1-15.

[10] Ferguson, T.S., 1967. Mathematical Statistics: A Decision Theoretic Approach. New York: Academic Press.

[11] Fisher, R., 1925. Theory of Statistical Estimation. *Proceedings of the Cambridge Philosophical Society*, 22, pp.700-725.

[12] Ghosh, J.K., 2002. Debabrata Basu: A Brief Life-Sketch. *Sankhya, A*, 64(1), pp.1-2.

[13] Ghosh, M., 2002. Basu's Theorem with Applications. *Sankhya, A*, 64(3), pp.509-531.

[14] Goossen, K., 1986. On Sufficiency, Conditional Completeness and the Theorems of Basu. *Statistics and Decisions*, 4(1-2), pp.85-96.

[15] Jorgensen, B., 1998. Basu's Theorem and Generalizations. *Seminar in Mathematical Statistics and Probability*.

[16] Kolmogorov, A.N., 1942. Sur l'Estimation Statistique des Parametres de la loi de Gauss. *Bulletin of the Academy of Sciences of the USSR: Mathematical Series*, 6, pp.3-32.

[17] Kusmierek, A., 2001. On Minimal Pairwise Sufficient Statistics. *Journal of Applied Analysis*, 7(2), pp.285-292.

[18] Lehmann, E.L., 1981. An Interpretation of Completeness and Basu's Theorem. *Journal of the American Statistical Association*, 76(374), pp.335-340.

[19] Lehmann, E.L. and Casella, G., 1998. Theory of Point Estimation. 2nd edition. New York: Springer-Verlag.

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

[21] Mukhopadhyay, N., 2000. Probability and Statistical Inference. New York: Marcel Dekker.

[22] Small, C.G. and McLeish, D.L., 1988. Generalizations of Ancillarity, Completeness and Sufficiency in an Inference Function Space. *The Annals of Statistics*, 16(2), pp.534-551.

[23] Weschler, S., 2002. Carlinhos Pereira. *The ISBA Bulletin*, 9(2), pp.6-9.