استقلال توابع هم‌وردا و ناوردا در خانواده نرمال تعمیم‌یافته

مهدی شمس[1]

تاریخ دریافت: ۱۳۹۵/۱۱/۲۶
تاریخ پذیرش: ۱۳۹۷/۶/۲۶

چکیده:

در این مقاله با فراخوانی توزیع نرمال تعمیم‌یافته که یک خانوادهٔ نمایی عمومی با پارامتر مکان و مقیاس است، یک شرط لازم و کافی برای استقلال یک آمارهٔ ناوردای مکانی از آمارهٔ بسندهٔ هم‌وردای مکانی که یک برآوردگر ماکسیمم درست‌نمایی است، ارائه می‌شود. در پایان نشان داده می‌شود که عکس این مطلب به جز در حالت‌های مجانبی صحیح است.

واژه‌های کلیدی: قضیهٔ باسو، خانواده نمایی، پارامترهای مکان و مقیاس.

۱ مقدمه

قضیهٔ باسو [۱] یکی از قضیه‌های مشهور آمار است که باعث کشف ارتباط بین بسندگی، آماره‌های کمکی و استقلال شد. با استفاده از این قضیه وجود استقلال بدون محاسبه توزیع توأم دو آماره حاصل می‌شود. [۷] حالات خاصی از قضیهٔ باسو (و عکس آن) در خانواده‌هایی با چگالی خاص را بیان کرده‌اند که در آنها استقلال آمارهٔ بسندهٔ کامل و آمارهٔ کمکی نتیجه می‌شود. در این مقاله در ابتدا یک خانواده نمایی عمومی به نام توزیع نرمال تعمیم‌یافته را معرفی می‌کنیم که بیشتر توزیع‌های معروف آماری نظیر نرمال، نمایی دم‌بریده، گاما، لگ‌نرمال، وایبل و یکنواخت را در بر می‌گیرد و سپس در این خانواده بر اساس یک نمونهٔ تصادفی، یک شرط لازم و کافی برای استقلال آمارهٔ بسندهٔ هم‌وردای مکانی که یک برآوردگر ماکسیمم درست‌نمایی است از هر آمارهٔ دلخواه ناوردای مکانی مطرح می‌شود (که به نوعی یک حالت از قضیهٔ باسو و عکس آن است) و در حالت خاص نتایج ثابت شده در [۱۰، ۱۱] را شامل می‌شود. به‌ازای یک پارامتر خاص در توزیع نرمال تعمیم‌یافته مشاهده می‌شود که برآوردگر ماکسیمم درست‌نمایی برای پارامتر مکان برابر میانگین نمونه است که این نتیجه با نتایج [۸، ۱۳] مطابقت دارد. [۱۶] نشان داد که اگر گشتاور اول (دوم) نمونه، یک برآوردگر ماکسیمم درست‌نمایی برای پارامتر مقیاس باشد، آن‌گاه تحت شرایط نظم توزیع نمایی (نرمال) با میانگین یک (صفر) است که [۱۳] حقیقت اخیر را در حالت چندمتغیره تعمیم دادند. [۹] تحت شرایط نظم نشان دادند که اگر برای اندازه نمونه $n = ۴$ برآوردگر ماکسیمم درست‌نمایی برای یک پارامتر مکان برابر میانه نمونه شود، آن‌گاه توزیع حاصل، توزیع لاپلاس است. در مورد شرایط وجود برآوردگر ماکسیمم درست‌نمایی در خانواده‌های گروه یک پارامتری [۴]، توزیع‌های کروی [۳] و توزیع‌های گسسته [۱۴، ۱۵] تحقیقات ارزنده‌ای انجام شده است که برای اطمینان بیشتر می‌توان به منابع اشاره شده رجوع کرد. هم‌چنین در این مقاله نشان داده می‌شود که اگر در یک نمونهٔ تصادفی به اندازهٔ $n = ۲$، تفاضل دو نمونه از صورت کلّی برآوردگر ماکسیمم درست‌نمایی متناظر با پارامتر مکان توزیع نرمال تعمیم‌یافته مستقل باشد، آن‌گاه دو متغیر دارای توزیع نرمال تعمیم‌یافته هستند و در پایان به‌کمک قضیهٔ باسو مثال نقضی برای عدم برقراری این گزاره در حالت‌های مجانبی به دست آورده می‌شود. در این مثال نقض از دو متغیر تصادفی مستقل و هم‌توزیع هندسی استفاده می‌شود، [۶] نشان داد عکس این حالت برقرار است و استقلال تفاضل دو نمونهٔ تصادفی از صورت کلّی برآوردگر ماکسیمم درست‌نمایی متناظر با پارامتر مکان توزیع

[1] عضو هیئت علمی گروه آمار، دانشگاه کاشان، ایران

نرمال تعمیم‌یافته منجر به این حقیقت می‌شود که متغیرها دارای توزیع هندسی هستند. در ابتدا به حالات خاصی از قضیهٔ باسو (و عکس آن) مطرح شد [۷] ، اشاره می‌شود. قضیه‌ها بدون اثبات ارائه می‌شود. خواننده برای مشاهده اثبات قضایا به [۷] مراجعه کند.

قضیهٔ ۱.۱. [۷] فرض کنید $X_1, ..., X_n$ یک نمونهٔ تصادفی از جمعیتی با تابع چگالی

$$f_{\boldsymbol{\theta}}\left(x\right) = Q\left(\boldsymbol{\theta}\right) M\left(x\right) e^{\sum_{j=1}^{q} P_j(\boldsymbol{\theta}) K_j(x)} I_{(a,\,b)}(x) \quad ; q < n$$

باشد، به‌طوری که $\boldsymbol{\theta} = (\theta_1, ..., \theta_q)$ بردار پارامترهای نامعلوم است و شرایط نظم به‌صورت زیر برقرار است:

۱. M تابعی نامنفی و پیوسته است.

۲. K'_j ها پیوسته و مستقل خطی هستند.

۳. P_j ها پیوسته‌اند و برد $(P_1, ..., P_q)$ که زیرمجموعهٔ $\mathbb{R}^q$ است، شامل فواصل q-بعدی غیرتهی است.

۴. a و b به θ_j ها بستگی ندارند.

در این صورت اگر $Z_j = \sum_{i=1}^{n} K_j(X_i)$ که در آن $j = 1, ..., q$، شرط لازم و کافی برای استقلال آمارهٔ دلخواه U از آمارهٔ بسندهٔ کامل توأم $\mathbf{Z} = (Z_1, ..., Z_q)$ این است، U یک آمارهٔ کمکی باشد.

قضیهٔ ۲.۱. [۷] فرض کنید $X_1, ..., X_n$ یک نمونهٔ تصادفی از جمعیتی با تابع چگالی

$$f_{\boldsymbol{\theta}}\left(x\right) = Q_1\left(\boldsymbol{\theta}\right) M_1\left(x\right) \exp\left(\sum_{j=1}^{q} P_j(\boldsymbol{\theta}) K_j\left(x\right)\right); \quad q < n$$

باشد، به‌طوری که $\boldsymbol{\theta} = (\theta_1, ..., \theta_q)$ و علاوه بر شرط ۳ قضیهٔ ۱.۱، شرایط نظم زیر نیز برقرار باشد:

۱. K_j ها مستقل‌اند و مشتقات جزئی آنها نسبت به هر یک از اجزای نمونه پیوسته هستند.

۲. دامنهٔ $f_{\boldsymbol{\theta}}$ به پارامتر بستگی نداشته باشد.

در این صورت $\mathbf{Z} = (Z_1, ..., Z_q)$ یک آمارهٔ بسندهٔ کامل توأم برای $\boldsymbol{\theta}$ است و شرط لازم و کافی برای استقلال آمارهٔ دلخواه U و $\mathbf{Z}$ این است که U آمارهٔ کمکی باشد.

قضیهٔ ۳.۱. [۷] فرض کنید $X_1, ..., X_n$ یک نمونهٔ تصادفی از جمعیتی با تابع چگالی

$$f_\theta\left(x\right) = Q\left(\theta\right) M\left(x\right) I_{(a(\theta),\, b(\theta))}\left(x\right)$$

باشد به‌طوری که پارامتر نامعلوم θ متعلق به فواصل غیرتهی است و علاوه بر برقراری شرط ۱ قضیهٔ ۱.۱، a' و b' پیوسته هستند و $sup\, a\left(\theta\right) = \inf\ b\left(\theta\right)$. همچنین یکی از شرایط زیر برقرار باشد:

۱. a ثابت و b اکیداً یکنوا است (یا برعکس).

۲. a اکیداً نزولی و b اکیداً صعودی است (یا برعکس).

در صورت برقراری شرط ۱ یا ۲، به ترتیب آماره‌های

$$Z = X_{(n)}$$
$$Z^* = X_{(1)}$$

یا

$$Z = \max\left\{a^{-1}(X_{(1)}), b^{-1}(X_{(n)})\right\}$$
$$Z^* = \min\left\{a^{-1}(X_{(1)}), b^{-1}(X_{(n)})\right\}$$

آمارهٔ بسندهٔ کامل برای θ است و شرط لازم و کافی برای استقلال از آماره U آن است که U آمارهٔ کمکی باشد.

قضیهٔ ۴.۱. [۷] فرض کنید $X_1, ..., X_n$ یک نمونهٔ تصادفی از جمعیتی با تابع چگالی

$$f_{\boldsymbol{\theta}}\left(x\right) = Q\left(\boldsymbol{\theta}\right) M\left(x\right) I_{(a(\boldsymbol{\theta}),\, b(\boldsymbol{\theta}))}(x)$$

باشد به‌طوری که در بردار پارامتر نامعلوم $\boldsymbol{\theta} = (\theta_1, \theta_2)$ اعضای θ_1 و θ_2 متعلق به بازه‌های غیرتهی هستند و علاوه بر شرط ۱ قضیهٔ ۱.۱، شرط زیر برقرار هستند:

a و b نسبت به هر یک از مؤلفه‌ها یکنوا و پیوسته باشند و برد (a, b) غیرتهی باشد.

در این صورت $\mathbf{Z} = \left(X_{(1)}, X_{(n)}\right)$ آمارهٔ بسندهٔ کامل توأم برای $\boldsymbol{\theta}$ است و شرط لازم و کافی برای استقلال آمارهٔ دلخواه U و $\mathbf{Z}$ این است که U یک آمارهٔ کمکی باشد.

قضیهٔ ۵.۱. [۷] فرض کنید $X_1, ..., X_n$ یک نمونهٔ تصادفی از جمعیتی با تابع چگالی

$$f_{\boldsymbol{\theta}}(x) = Q(\boldsymbol{\theta})\, M(x) \exp\left(\sum_{j=3}^{q} P_j(\theta_3, ...\theta_q) K_j(x)\right)$$

$$I_{(a(\theta_1,\theta_2),\, b(\theta_1,\theta_2))}(x), \quad q < n$$

باشد به‌طوری که هر یک از مختصات بردار پارامتر $\boldsymbol{\theta} = (\theta_1, ..., \theta_q)$ متعلق به بازه‌های غیرتهی است، شرایط ۱ تا ۴ قضیهٔ ۱.۱ و شرط قضیهٔ ۴.۱ برقرار هستند و $Z_1 = X_{(1)}$، $Z_2 = X_{(n)}$ و $Z_j = \sum_{i=1}^{n} K_j(X_i)$ که در آن $j = 3, ..., q$. در این صورت $\mathbf{Z} = (Z_1, ..., Z_q)$ یک آمارهٔ بسندهٔ کامل توأم برای $\boldsymbol{\theta}$ است و شرط لازم و کافی برای استقلال آمارهٔ دلخواه U و $\mathbf{Z}$ این است که U یک آمارهٔ کمکی باشد.

۲ استقلال در خانواده‌های نمایی با پارامترهای مکان و مقیاس

فرض کنید متغیر تصادفی X دارای توزیع گاما با تابع چگالی

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

است. در این صورت تابع چگالی متغیر تصادفی $Y = \frac{1}{\gamma}\ln(X)$ برای $\gamma \neq 0$ به‌صورت

$$f_Y(y) = \frac{|\gamma|\, \alpha^{\alpha}}{\Gamma(\alpha)} \exp\left\{-\alpha e^{\gamma(y-\theta)} + \alpha\gamma(y-\theta)\right\}$$

خواهد بود، که در آن $\theta = \ln \sqrt[\gamma]{\alpha\beta}$ پارامتر مکان برای خانواده توابع چگالی $f_Y(y)$ است [۵]. میانگین و واریانس Y را می‌توان به‌صورت

$$E(Y) = \theta + \tfrac{1}{\gamma}\left[\Psi(\alpha) - \ln(\alpha)\right]$$

$$Var(Y) = \sigma^2 = \tfrac{1}{\gamma^2}\Psi'(\alpha)$$

محاسبه کرد که در آن $\Psi(\alpha)$ و $\Psi'(\alpha)$ اولین و دومین مشتق $\ln(\Gamma(\alpha))$ هستند. توزیع Y با نماد $N(\theta, \sigma^2, \gamma)$ نشان داده می‌شود که در آن $\theta \in \mathbb{R}$، $\sigma^2 > 0$ و $\gamma \neq 0$، در حالت خاص این توزیع وقتی $\gamma \to 0$ یا $\gamma \to -\infty$ ($\gamma \to +\infty$) به ترتیب به $N(\theta, \sigma^2)$ یا نمایی دم‌بریده از چپ (راست) با توابع چگالی

$$f_Y(y) = \tfrac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\tfrac{1}{2\sigma^2}(y-\theta)^2\right\}; \; y \in \mathbb{R} \qquad \gamma \to 0,$$

$$f_Y(y) = \tfrac{1}{\sigma} \exp\left\{-\tfrac{1}{\sigma}(y-\theta)\right\}; \qquad y > \theta \qquad \gamma \to -\infty,$$

$$f_Y(y) = \tfrac{1}{\sigma} \exp\left\{+\tfrac{1}{\sigma}(y-\theta)\right\}; \qquad y < \theta \qquad \gamma \to +\infty$$

میل می‌کند و برای سادگی به‌ترتیب از نمادهـــای $N(\theta, \sigma^2, 0)$، $N(\theta, \sigma^2, -\infty)$ و $N(\theta, \sigma^2, +\infty)$ استفاده می‌شود. در نتیجه می‌توان نماد $N(\theta, \sigma^2, \gamma)$ را روی $\theta \in \mathbb{R}$ و $\sigma^2 > 0$ و $\gamma \in \mathbb{R}^* = \mathbb{R} \cup \{-\infty, +\infty\}$ نیز تعریف کرد. این خانواده از توزیع‌ها یک خانوادهٔ نمایی یک‌پارامتری در θ (و یا σ^2) و یک خانوادهٔ نمایی دوپارامتری در (θ, σ^2) هستند. حال اگر $X_1, ..., X_n$ یک نمونهٔ تصادفی از توزیع $N(\theta, \sigma^2, \gamma)$ باشد، آن‌گاه آمارهٔ بسنده (برآوردگر ماکسیمم درست‌نمایی) برای θ که در آن پارامترهای σ^2 و γ ثابت‌اند برابر است با:

$$M_\gamma(X_1, ..., X_n) = \begin{cases} \frac{1}{\gamma}\ln\left(\frac{1}{n}\sum_{i=1}^{n} e^{\gamma x_i}\right), & \gamma \in \mathbb{R} - \{0\} \\ \bar{X}, & \gamma = 0 \\ X_{(n)}, & \gamma = +\infty \\ X_{(1)}, & \gamma = -\infty \end{cases}$$

که آن را اصطلاحاً میانگین مرتبه γ گوییم [۴]. این تابع برای مقادیر ثابت نمونه یک تابع اکیداً صعودی و پیوسته از γ است، مگر این که همهٔ X_iها مساوی باشند که در این حالت $M_\gamma(X_1, ..., X_n)$ اثبات است. برای یک نمونهٔ تصادفی به اندازهٔ $n \geq 2$ ثابت وجود و یا عدم وجود یک آمارهٔ بسندهٔ کامل برای پارامتر مکان θ در خانوادهٔ $N(\theta, \sigma^2, \gamma)$ یک مسئله حل نشده است [۵]. حالت خاص توزیع $N(\theta, \sigma^2, 0)$ با نتیجه [۸] (و در حالت چندمتغیره [۱۳]) مبنی بر این‌که اگر برآوردگر ماکسیمم درست‌نمایی برای یک پارامتر مکان برابر میانگین نمونه شود، آن‌گاه توزیع نرمال (نرمال چندمتغیره) است منطبق هستند.

تذکر ۱.۲. اگر متغیر تصادفی Y دارای توزیع $N(\theta, \sigma^2, \gamma)$ باشد، آن‌گاه توزیع $X = e^Y$ با نماد $L(\theta, \sigma^2, \gamma)$ نمایش داده می‌شود که در این توزیع θ پارامتر مقیاس است. در حالت خاص $L(\theta, \sigma^2, 1)$ توزیع گاما، $L(\theta, \sigma^2, 0)$ توزیع لگ‌نرمال $L(\theta, \frac{1}{\sigma^2}\Psi'(\frac{1}{\gamma}), \gamma)$، توزیع وایبل $(\gamma > 0)$ و $L(\theta, 1, +\infty)$

توزیع یکنواخت $U(\circ, \theta)$ هستند.

تحقیقات متنوع راجع به پارامترهای مکان و مقیاس در خانواده‌های نمایی باعث شده که ایده‌ای برای عمومیت قضیهٔ باسو پیدا شود که در این‌جا به آن اشاره می‌شود. می‌دانیم در خانواده‌های نمایی وقتی بعد فضای پارامتر و بعد آمارهٔ بسنده یکسان باشد، آمارهٔ بسنده، کامل است [۱۲]. از طرف دیگر در خانواده‌های با پارامترهای مکان، مقیاس و ترکیبی از این دو، آماره‌های کمکی به راحتی به دست می‌آیند. به‌عنوان نمونه اگر $X_1, ..., X_n$ یک نمونهٔ تصادفی از توزیعی با پارامتر مکان θ باشد، آن‌گاه توزیع توأم تفاضل‌های $X_1 - X_2, ..., X_{n-1} - X_n$ به پارامتر θ بستگی ندارد. بنا بر این طبق قضیهٔ باسو آمارهٔ بسندهٔ کامل T برای θ از هر تابعی از تفاضل‌های X_iها مستقل است [۲]. به‌طور مشابه توزیع توأم نسبت‌های X_i بستگی به پارامتر مقیاس ندارد و توزیع توأم نسبت‌های تفاضل‌ها به پارامتر مکان-مقیاس بستگی ندارد. بنا بر این عمده‌ترین کاربردهای قضیهٔ باسو در حالت یک‌متغیره، در خانواده‌های نمایی(با پارامترهای مکان، مقیاس و مکان-مقیاس) است، زیرا در این خانواده‌ها بعد فضای پارامتر با بعد آمارهٔ بسنده یکسان است و در نتیجه آمارهٔ بسنده، کامل است و لذا می‌توان از قضیهٔ باسو استفاده کرد. به‌عنوان مثال اگر $X_1, ..., X_n$ یک نمونهٔ تصادفی از توزیع لگ‌نرمال $L(1, \sigma^2, \circ)$ باشد، آن‌گاه $\sum_{i=1}^{n} \ln {X_i}^2$ از هر تابع $g(X_1, ..., X_n)$ که در آن به‌ازای هر λ داشته باشیم $g(X_1, ..., X_n) = g({X_1}^\lambda, ..., {X_n}^\lambda)$ مستقل است، زیرا σ پارامتر مقیاس در توزیع$(N(\circ, \sigma^2, \circ$ است. شایان ذکر است که قضیهٔ باسو علاوه بر موارد بالا در خانواده‌هایی که پارامتر مکان و یا مقیاس ندارند نیز استفاده می‌شود.

در این مقاله ابتدا ناوردایی مکانی و مقیاسی تعریف می‌شود. فرض کنید $X_1, ..., X_n$ یک نمونهٔ تصادفی از یک مدل ناوردای مکانی (مقیاسی) است هرگاه متغیرهای مستقل و هم‌توزیع $W_1, ..., W_n$ وجود داشته باشند به‌طوری که به‌ازای یک θ داشته باشیم:

$$(X_1, ..., X_n) \stackrel{d}{=} (W_1 + \theta, ..., W_n + \theta)$$
$$((X_1, ..., X_n) \stackrel{d}{=} (\theta W_1, ..., \theta W_n)).$$

در این‌جا توزیع W_iها ثابت و به پارامتر θ بستگی ندارد. همچنین آمارهٔ U را ناوردای مکانی (مقیاسی) گویند هرگاه به‌ازای λ و هر $x_1, ..., x_n$ داشته باشیم:

$$U(\lambda + x_1, ..., \lambda + x_n) = U(x_1, ..., x_n)$$
$$(U(\lambda + x_1, ..., \lambda + x_n) = U(x_1, ..., x_n))$$

و آمارهٔ T را هم‌وردای مکانی (مقیاسی) گویند هرگاه به‌ازای λ و هر $x_1, ..., x_n$ داشته باشیم:

$$U(\lambda + x_1, ..., \lambda + x_n) = U(x_1, ..., x_n) + \lambda$$
$$(U(\lambda x_1, ..., \lambda x_n) = \lambda U(x_1, ..., x_n))[۱۲].$$

گزارهٔ ۲.۲. [۱۲]. آمارهٔ ناوردای مکانی (مقیاسی) برای مدل ناوردای مکانی (مقیاسی)، یک آمارهٔ کمکی است.

اثبات. گزاره در حالت مکان ثابت می‌شود (در حالت مقیاس مشابه است). متغیرهای مستقل و هم‌توزیع $W_1, ..., W_n$ وجود دارند به‌طوری که به‌ازای یک θ، داریم

$$(X_1, ..., X_n) \stackrel{d}{=} (W_1 + \theta, ..., W_n + \theta)$$

و در پی آن می‌توان نتیجه گرفت:

$$\begin{aligned} U(X_1, ..., X_n) &\stackrel{d}{=} U(W_1 + \theta, ..., W_n + \theta) \\ &= U(W_1, ..., W_n) \qquad (۱) \end{aligned}$$

که تساوی (۱) به‌دلیل ناوردای مکانی بودن آمارهٔ U است. با توجه به این‌که توزیع W_iها به پارامتر θ بستگی ندارد، در نتیجه توزیع آماره U وابسته به θ نیست و از این رو آمارهٔ کمکی است. □

به‌عنوان مثال اگر $X_1, ..., X_n$ نمونهٔ تصادفی از توزیع $U(\circ, \theta)$ باشد که $\theta > \circ$، از این‌که برای نمونهٔ تصادفی $W_1, ..., W_n$ از توزیع $U(\circ, 1)$، $\theta W_i \stackrel{d}{=} X_i$ نتیجه می‌شود که خانواده توزیع‌های $U(\circ, \theta)$ از یک مدل مقیاس هستند و چون برای آمارهٔ

$$U = \frac{X_1 + ... + X_n}{X_1}$$

نتیجه می‌شود

$$U(\lambda X_1,...,\lambda X_n)=\frac{\lambda X_1+...+\lambda X_n}{\lambda X_1}$$
$$=U(X_1,...,X_n)$$

آمارهٔ U ناوردای مقیاسی و طبق گزارهٔ ۲.۲ یک آمارهٔ کمکی است. [۱۱، ۱۰] نتایجی را برای توزیع نرمال و گاما به‌دست‌آمده است که از قضیهٔ ۱.۱ نتیجه می‌شوند:

نتیجه ۳.۲. [۱۱] اگر $X_1,...,X_n$ یک نمونهٔ تصادفی از توزیع نرمال باشد، $\bar{X}$ و $g(X_1,...,X_n)$ مستقل هستند، اگر و فقط اگر به‌ازای هر λ داشته باشیم:

$$g(X_1,...,X_n)\overset{d}{=}g(\lambda+X_1,...,\lambda+X_n).$$

نتیجه ۴.۲. [۱۰] اگر $X_1,...,X_n$ یک نمونهٔ تصادفی از توزیع گاما باشد، $\bar{X}$ و $g(X_1,...,X_n)$ مستقل هستند، اگر و فقط اگر به‌ازای هر λ داشته باشیم:

$$g(X_1,...,X_n)\overset{d}{=}g(\lambda X_1,...,\lambda X_n).$$

به راحتی با استفاده از قضیهٔ باسو و قضیهٔ ۱.۱ می‌توان نتایج اخیر را در حالت کلّی و در خانوادهٔ

$$P=\{N(\theta,\sigma^2,\gamma):\theta\in\mathbb{R},\sigma^2>0,\gamma\in\mathbb{R}^*\}$$

تعمیم داد که در قضیه زیر به آن اشاره شده است.

قضیهٔ ۵.۲. [۱۰] اگر $X_1,...,X_n$ یک نمونهٔ تصادفی از توزیع $N(\theta,\sigma^2,\gamma)$ باشد، $M_\gamma(X_1,...,X_n)$ از $g(X_1,...,X_n)$ مستقل است اگر و تنها اگر که به‌ازای هر λ داشته باشیم:

$$g(X_1,...,X_n)\overset{d}{=}g(\lambda+X_1,...,\lambda+X_n).$$

تذکر ۶.۲. همان‌طور که در قضیهٔ بالا مشاهده می‌شود برآوردگر هم‌وردای مکانی $M_\gamma(X_1,...,X_n)$ از آمارهٔ ناوردای $g(X_1,...,X_n)$ مستقل است. اثبات هم‌وردای مکانی بودن $M_\gamma(X_1,...,X_n)$ از تساوی زیر واضح است:

$$M_\gamma(X_1+c,...,X_n+c)=\begin{cases}\frac{1}{\gamma}\ln\left(\frac{1}{n}e^c\sum_{i=1}^n e^{\gamma x_i}\right) & \gamma\in\mathbb{R}-\{0\}\\ \bar{X}+c & \gamma=0\\ X_{(n)}+c & \gamma=+\infty\\ X_{(1)}+c & \gamma=-\infty\end{cases}$$
$$=M_\gamma(X_1,...,X_n)+c.$$

به‌طور مشابه $M_\gamma(X_1,...,X_n)$ فقط در حالت‌های $\gamma=0,\pm\infty$ برآوردگر هم‌وردای مقیاسی است.

نتایج ۳.۲ و ۴.۲، حالت خاص قضیهٔ ۵.۲ هستند که با قرار دادن $\gamma=0$ برای توزیع نرمال و $\gamma=1$ برای توزیع گاما حاصل می‌شوند.

گزارهٔ ۷.۲. [۵] اگر X و Y مستقل و برای برخی مقادیر حقیقی γ متغیرهای تصادفی $X-Y$ و $M_\gamma(X,Y)$ مستقل باشند، آن‌گاه X و Y هر دو دارای توزیع $N(\theta,\sigma^2,\gamma)$ هستند.

گزارهٔ ۷.۲ به‌ازای هر $\gamma\in\mathbb{R}$ صحیح است ولی برای $\gamma\to\pm\infty$ صدق نمی‌کند. با کمک قضیهٔ باسو مثال نقضی برای عدم برقراری این گزاره در حالت $\gamma\to\pm\infty$ به دست آورده شده است.

مثال ۸.۲. اگر X و Y دو متغیر تصادفی مستقل با توزیع هندسی $Ge(\theta)$ باشند، آن‌گاه تابع چگالی احتمال توأم X و Y به‌صورت زیر است:

$$P_\theta(X=x,Y=y)=(1-p)^2p^{x+y-2\theta}$$
$$I_{\{\theta,\theta+1,...\}\times\{\theta,\theta+1,...\}}(x,y).$$

اگر $\gamma\to-\infty$، آن‌گاه $M_{-\infty}(X,Y)=\min(X,Y)$ آمارهٔ بسندهٔ کامل برای θ است و از این رو طبق قضیهٔ باسو از آمارهٔ کمکی $X-Y$ مستقل است؛ در حالی که X و Y دارای توزیع $N(\theta,\sigma^2,-\infty)$ نیست. (ذکر این نکته ضروری است که [۶] نشان داد عکس مطلب نیز برقرار است، یعنی اگر $\min(X,Y)$ و $X-Y$ مستقل باشند، آن‌گاه X و Y دارای توزیع هندسی هستند.)

۳ نتیجه‌گیری

در این مقاله به‌کمک قضیهٔ باسو در یک خانواده نمایی عمومی یعنی توزیع نرمال تعمیم‌یافته با پارامتر مکان و مقیاس، استقلال آمارهٔ بسندهٔ هم‌وردای مکانی از یک آمارهٔ ناوردای مکانی مورد بررسی قرار گرفت. بررسی استقلال در حالت مقیاس و مکان-مقیاس و همچنین تحت گروه تبدیلات دیگر در این خانواده و همچنین اثبات وجود و یا عدم وجود یک آمارهٔ بسندهٔ کامل در این خانواده از مطالبی است که می‌تواند مسیر جدیدی را برای محققان علاقه‌مند باز کند.

مراجع

[1] Basu, D. (1955). On statistics independent of a complete sufficient statistic, *Sankhya*, **15**, 377-380.

[2] Basu, D. (1977). On the elimination of nuisance parameters, *Journal of the American Statististical Association*, **72**, 355-366.

[3] Duerinckx, M. and Ley, C. (2012). Maximum likelihood characterization of rotationally symmetric distributions on the sphere, *Sankhya A*, **74**(**2**), 249–262.

[4] Duerinckx, M. and Ley, C. and Swan, Y. (2014). Maximum likelihood characterization of distributions Bernoulli, *Euclid*, **20**(**2**), 775–802.

[5] Ferguson, T. S. (1962a). Location and scale parameters in exponential families of distributions, *Annals of Mathematical Statistics*, **33**, 986-1001.

[6] Ferguson, T. S. (1962b). A characterization of the geometric distribution, *Annals of Mathematical Statistics*, **33**, 1207.

[7] Hogg, R. V. and Craig, A. T. (1956). Sufficient statistics in elementary distribution theory, *Sankhya*, **17**, 209-216.

[8] Hürlimann, W. (2013). *A Characterization of the Compound Multiparameter Hermite Gamma Distribution via Gauss's Principle*, Hindawi Publishing Corporation The Scientific World Journal Volume.

[9] Kagan, A. M., Linnik, Y. V. and Rao, C. R. (1973). *Characterization Problems in Mathematical Statistics*, John Wiley Sons, New York-London-Sydney.

[10] Laha, R. G. (1954). On a characterisation of the gamma distribution, *Annals of Mathematical Statistics*, **25**, 784-787.

[11] Laha, R. G. (1956). On Some properties of the normal and gamma distributions, *Proceedings of American Mathematical Society*, **7**, 172-174.

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

[13] Marshall, A. W. and Olkin, I. (1993). Maximum likelihood characterizations of distributions, *Statistica Sinica*, **3**, 157–171.

[14] Puig, P. (2003). Characterizing additively closed discrete models by a property of their MLEs, with an application to generalized Hermite distributions, *Journal of the American Statististical Association*, **98**, 687-692.

[15] Puig, P. and Valero, J. (2006). Count data distributions: some characterizations with applications, *Journal of the American Statististical Association*, **101**, 332-340.

[16] Teicher, H. (1961). Maximum likelihood characterization of distributions, *Annals of Mathematical Statistics*, **32**, 1214–1222.