مجله علوم آماری، بهار ۱۳۸۶
جلد ۱، شماره ۱، ص ۷۳–۸۸

استنباط در توزیع نرمال براساس نمونه گیری وزنی

سید محمد رضا علوی، رحیم چینی‌پرداز
گروه آمار دانشگاه شهید چمران اهواز

تاریخ دریافت: ۱۳۸۵/۴/۵ تاریخ آخرین بازنگری: ۱۳۸۵/۱۱/۲۱

چکیده : روش‌های معمول، استنباط آماری مبتنی بر نمونه تصادفی است. اما در بسیاری از موارد مکانیسم نمونه‌گیری به گونه‌ای است که هر مشاهده با شانسی متناسب با تابعی نامنفی از آن ثبت می‌شود. چنین نمونه‌ای را نمونه وزنی گویند. بنابراین لازم است اصلاحات مناسبی که بستگی به تابع وزن دارد در روش های معمول استنباطی صورت گیرد. در این مقاله تحت چند وزن خاص چنین اصلاحاتی در روش های برآورد پارامترها در توزیع نرمال ارائه و در پایان توزیع نرمال دوقلو به عنوان یک توزیع نرمال وزنی معرفی شده است.

واژه‌های کلیدی : تابع وزن مستقل از پارامتر، تابع وزن وابسته به پارامتر، توزیع نرمال وزنی، توزیع نرمال دوقلو، متغیر تصادفی وزنی، نمونه‌گیری وزنی.

۱ مقدمه

دراستنباط آماری کلاسیک فرض می شود یک نمونه تصادفی از جامعه انتخاب و براساس آن درمورد جامعه استنباط‌های آماری مانند برآورد فاصله اطمینان و آزمون‌های فرض انجام می‌شود. بدیهی است اگر روش ثبت داده‌ها به گونه‌ای باشد

° آدرس الکترونیک مسئول مقاله: سید محمدرضا علوی، alavi_m@scu.ac.ir

[DOR: 20.1001.1.17358183.1386.1.1.13.0]

که مشاهدات شانس یکسان برای ثبت نداشته باشند، به نتایج حاصل از استنباط کلاسیک این داده‌ها نمی‌توان اعتماد کرد. به عنوان مثال فرض کنید روش صید ماهی‌ها به گونه‌ای باشد که ماهی‌های بزرگ شانس بیشتری برای صید داشته باشند (برای مثال با پرتاب نیزه به ماهی یا شلیک به آن)، یا در کنترل کیفیت و آنالیز داده‌های پرت مشاهدات ثبت شده فقط داده‌های خارج از کنترل یا داده‌های پرت باشند. واضح است که در چنین مواردی مشاهدات با وزن‌هایی خاص ثبت می‌شوند. در مثال اول (صید ماهی توسط پرتاب نیزه یا شلیک تصادفی به ماهی‌ها) توابع وزن زیر مناسب به نظر می‌رسند:

$$w(x,\alpha)=\exp\{\alpha x\},\qquad w(x,\alpha,\sigma^{2})=\exp\left\{\frac{\alpha x}{\sigma^{2}}\right\},\qquad w(x,\alpha)=x^{\alpha}$$

و در مثال های دیگر استفاده از تابع وزن زیر مفید خواهد بود:

$$w(x,\alpha)=(x-\mu)^{\alpha}$$

در هر کدام از این وزن‌ها اگر α به صفر میل کند نمونه تصادفی خواهد بود. در مثال‌های بالا و موارد مشابه که در عمل بسیار رخ می‌دهند (خصوصا وقتی طبیعت، نمونه را انتخاب می‌کند) استفاده از نمونه‌گیری وزنی ضروری است. ایده اصلی استفاده از توابع وزن ابتدا توسط فیشر (۱۹۳۴) بیان و سپس بوسیله رائو (۱۹۶۵) به طور رسمی به عنوان نمونه‌گیری وزنی تعریف شد. بعد از آن آمارشناسان زیادی جنبه‌های کاربردی و نظری آن را مطالعه کردند. به عنوان مثال در زمینه‌های کاربردی براون (۱۹۷۲) و اسمیت و پارنز (۱۹۹۴) در زمینه ترافیک، موریسون (۱۹۷۳) در بازاریابی و مطلوبیت منابع، در تحقیقات مربوط به تولیدات جنگلی وارن (۱۹۷۵)، در زمینه شیلات چن و کولینگ (۲۰۰۱)، از نمونه‌گیری وزنی استفاده کرده‌اند. در زمینه‌های نظری نیز می‌توان به گوپتا و آکمن (۱۹۹۵) و سنسگیری و آکمن (۲۰۰۱) در قابلیت اعتماد، کوک و مارتین (۱۹۷۴)، پاتیل و رائو (۱۹۷۸) در نمونه‌گیری ترابرشی خطی، پاتیل و ارد (۱۹۷۸) و آنی کریشنان و سوناج (۲۰۰۳) در پایایی در توزیعهای وزنی در حالت خاص، سلستین و مایزر (۲۰۰۳) درباره مشخصه‌های بیش پراکندگی و کم پراکندگی در توزیعهای وزنی پواسون، چاکرابرتی

[DOR: 20.1001.1.17358183.1386.1.1.13.0]

و داس (۲۰۰۶) راجع به مشخصه‌هایی از توزیعهای وزنی شبه دوجمله‌ای، گیولامن و همکاران (۱۹۹۸) در برآورد چگالی با استفاده از داده‌های وزنی به روش هسته‌ای، بایاری و دگروت (۱۹۸۷) راجع به تحلیل بیزی مدلهای انتخاب، بایاری و برگر (۱۹۹۸) در تحلیل بیزی استوار مدلهای انتخاب و جاکوبو (۲۰۰۳) در مورد نتایج بزرگ نمونه‌ای تحت نمونه‌گیری اریب در حضور متغیرهای وابسته اشاره کرد.

در این مقاله فرض شده است که نمونه‌گیری از توزیع نرمال تولید ولی مشاهدات تحت تابع وزن $w(x)$ ثبت می‌شوند و هدف اصلاح برآوردهای نقطه‌ای است وقتی که $w(x)$ وابسته و یا مستقل از پارامتر باشد. بدیهی است اگر توزیع اصلی نامعلوم باشد توزیع وزنی نیز نامعلوم خواهد بود. در بخش دوم مقاله توزیع وزنی و نمونه‌گیری وزنی مطرح می‌شوند و در بخش سوم برآورد پارامترهای جامعه نرمال تحت تابع وزن مستقل از پارامتر ارائه می‌شود. دربخش چهارم نیز فرض می‌شود که تابع وزن به پارامترهای جامعه وابسته است. روش های اصلاحی برآوردها و مقایسه آنها با بخش سوم به این بخش اختصاص یافته است. در بخش پنجم و نهایی تحت یک تابع وزن خاص یک توزیع جدید آماری معرفی و خواص آن مطرح خواهد شد.

۲ توزیع وزنی و نمونه گیری وزنی

۱.۲ توزیع وزنی

فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال (یا تابع جرم احتمال) $f(x;\theta)$ و θ بردارپارامتر جامعه باشد. فرض کنید نمونه‌گیری چنان باشد که مشاهده x شانسی متناسب با $w(x,\alpha) \geq 0$ برای ثبت شدن داشته باشد. α بردار پارامتر تابع وزن نامیده می‌شود که در این مقاله آن را معلوم فرض می‌کنیم. داریم:

$$\frac{P(\text{ثبت}\,|X=x)}{P(\text{ثبت}\,|X=y)} = \frac{w(x,\alpha)}{w(y,\alpha)} \qquad (۱)$$

[DOR: 20.1001.1.17358183.1386.1.1.13.0]

درچنین حالتی مشاهده ثبت شده برآمده ازجامعه اصلی X نیست بلکه از جامعه X^w با تابع چگالی احتمال زیر است:

$$f^w(x,\theta,\alpha) = \frac{w(x,\alpha)f(x,\theta)}{E[w(X,\alpha)]} \qquad (۲)$$

در اینجا X^w متغیر وزنی X است که توزیع آن را توزیع وزنی باتابع وزن $w(x,\alpha)$ و نمونه‌گیری متناظر آن را نمونه‌گیری وزنی می‌گویند. یک شرط ابتدایی ایجاب می‌کند که $E[w(X,\alpha)]$ موجود باشد. خاصیت چگالی بودن f^w نیز به سادگی قابل اثبات است. اگر $w(x,\alpha) = x$ باشد، X^w را متغیر اریب اندازه و نمونه‌گیری متناظر آن را نمونه‌گیری اریب اندازه گویند. اگر $w(x,\alpha) = c$ مقدار ثابتی باشد، $f^w(x;\theta) = f(x;\theta)$ و نمونه‌گیری، تصادفی است. بدیهی است اگر $f^w = f$ نباشد، احتمالا استنباط درباره پارامتر θ براساس مشاهدات وزنی به روش‌های معمول گمراه کننده خواهد بود. به همین دلیل بسیاری از استفاده کنندگان آمار به دلیل غیر تصادفی بودن، اعتقاد به دور انداختن مشاهدات دارند که مستلزم عدم آگاهی و از دست رفتن هزینه‌های نمونه‌گیری می‌شود. گروهی نیز در عمل با یکی دانستن f و f^w استنباط‌های نا دقیق و بعضا اشتباهی را مرتکب می‌شوند. راه منطقی حل مساله با آگاهی از عدم تصادفی بودن نمونه‌گیری، تلاش برای استفاده از داده‌ها و ارائه راهکاری مناسب خواهد بود. در این مقاله سعی شده است که راهکارهای‌مناسب برای برآورد پارامترهای جامعه نرمال ارائه شود.

۲.۲ نمونه‌گیری وزنی از توزیع نرمال

فرض کنید $X \sim N(\mu,\sigma^۲)$ و $x_۱, x_۲, \ldots, x_n$ مقادیر محقق شده یک نمونه تصادفی $X_۱^w, X_۲^w, \ldots, X_n^w$ از متغیر نرمال وزنی X^w باشد. در این صورت

$$f^w(x;\mu,\sigma,\alpha) = \frac{w(x;\alpha)}{\sigma\sqrt{۲\pi}E[w(X,\alpha)]} \exp\left\{-\frac{(x-\mu)^۲}{۲\sigma^۲}\right\} \qquad (۳)$$

$$\sigma > ۰, -\infty < \mu < \infty$$

خواهد بود. در اینجا اگر چه $N(\mu,\sigma^۲)$ عضو خانواده نمایی دو پارامتری است ولی f^w لزوما عضو چنین خانواده‌ای نخواهد بود. به عنوان مثال، اگر

[DOR: 20.1001.1.17358183.1386.1.1.13.0]

$w(x,\mu) = (x-\mu)^{۲}$ باشد نتایج مربوط به خانواده نمایی قابل استفاده نیست. می‌توان نشان داد اگر

$$w(x;\mu,\sigma) = \exp\{c(\mu,\sigma)b(x)\}$$

که در آن c تابعی فقط از (μ,σ) و b تابعی فقط از x باشد، f^w عضو خانواده نمایی و اگر $w(x,\alpha,\beta) = \exp\{\alpha x + \beta x^{۲}\}$ باشد، توزیع نرمال وزنی نیز یک توزیع نرمال است.

۳ روش های اصلاحی در برآورد پارامترها با تابع وزن مستقل از پارامترهای جامعه

در حالت کلی برآوردهای حداکثر درستنمایی پارامترها در رابطه (۳) از حل معادلات زیر در صورت وجود به دست می‌آیند

$$\begin{aligned} \frac{\partial \ell}{\partial \mu} &= \frac{\sum_{i=۱}^{n}(x_i-\hat{\mu})}{\sigma^{۲}} - \frac{n\partial \log E[w(X)]}{\partial \mu} = ۰ \\ \frac{\partial l}{\partial \sigma^{۲}} &= \frac{-\sum_{i=۱}^{n}(x_i-\hat{\mu})}{۲\hat{\sigma}^{۴}}\frac{n}{۲\hat{\sigma}^{۲}} - \frac{n\partial \log E[w(X)]}{\partial \sigma^{۲}} = ۰ \end{aligned} \qquad (۴)$$

که در آن ℓ لگاریتم تابع درستنمایی حاصل شده از (۳) است. معادلات بالا ممکن است جواب تحلیلی نداشته باشند و با استفاده از روش‌های عددی پارامترها برآورد شوند.

۱.۳ وزن نمایی درجه دوم

f^w تحت وزن $w(x,\alpha,\beta) = \exp\{\beta x^{۲} + \alpha x\}$ منجر به یک توزیع نرمال جدید با میانگین $\frac{\mu+\alpha\sigma^{۲}}{۱-۲\beta\sigma^{۲}}$ و واریانس $\frac{\sigma^{۲}}{۱-۲\beta\sigma^{۲}}$ خواهد شد (برای $\beta < (۲\sigma^{۲})^{-۱}$). در حالت خاص‌تر با $\beta = ۰$ و $w(x,\alpha) = \exp\{\alpha x\}$ توزیع وزنی، نرمال با میانگین $\mu + \alpha\sigma^{۲}$ و واریانس $\sigma^{۲}$ است. بنابراین توزیع میانگین نمونه وزنی $\bar{X}^w = \frac{۱}{n}\sum_{i=۱}^{n} X_i^w$ نرمال و

توزیع $\frac{\sum_{i=1}^{n}(X_i^w-\bar{X}^w)^2}{\sigma^2}$ خی‌دو با $n-1$ درجه آزادی خواهد بود. برآوردگرهای حداکثر درستنمایی عبارتند از:

$$\begin{aligned}\hat{\sigma}^2 &= \frac{\sum_{i=1}^{n}(X_i^w-\bar{X}^w)}{n}\\ \hat{\mu} &= \bar{X}^w-\alpha\hat{\sigma}^2.\end{aligned} \tag{5}$$

این برآوردگرها اریب اما برآوردگرهای زیر نااریب هستند:

$$\begin{aligned}S_w^2 &= \frac{\sum_{i=1}^{n}(X_i^w-\bar{X}^w)}{n-1}\\ \hat{\hat{\mu}} &= \bar{X}^w-\alpha S_w^2.\end{aligned} \tag{6}$$

برآوردگرهای فوق مستقل نیستند ولی از مستقل بودن $\bar{X}^w$ و S_w^2 در توزیع نرمال می‌توان توزیع $\hat{\hat{\mu}}$ را وقتی که n فرد باشد به صورت زیر به دست آورد:

$$\begin{aligned}f_{\hat{\hat{\mu}}}(x,\alpha;\mu,\sigma^2) &= K\exp\left\{\frac{(n-1)x}{2\alpha\sigma^2}\right\}\\ &\times \sum_{i=0}^{\frac{n-3}{2}}\binom{\frac{n-3}{2}}{i}\mu_i'\Big(\frac{\mu^*-x}{\sigma/\sqrt{n}}\Big)^{\frac{n-3}{2}-i}\end{aligned} \tag{7}$$

که در آن $\mu^*=\mu+\alpha\sigma^2-\frac{n-1}{2n\alpha}$ و

$$\begin{aligned}K &= \left[\frac{\sigma^2}{2(n-1)\alpha}\right]^{\frac{n-1}{2}}\left[\frac{\sigma}{\sqrt{n}}\right]^{\frac{n-3}{2}}\\ &\times \exp\left\{-\frac{n}{2\sigma^2}\left[-\frac{(n-1)^2}{4n^2\alpha^2}+\frac{n-1}{n\alpha}(1+\mu+\alpha\sigma^2)\right]\right\}\end{aligned}$$

μ_i' گشتاور مرتبه iام توزیع نرمال استاندارد است که برای i فرد صفر است. وقتی n زوج باشد توزیع شکل پیچیده‌ای خواهدداشت.

اگر $\sigma=\sigma_0$ معلوم باشد برآورد حداکثر درستنمایی μ برابر $\bar{X}^w-\alpha\sigma^2$ است که برآوردی نااریب و توزیع آن نرمال با میانگین μ و واریانس $\frac{\sigma_0^2}{n}$ خواهد بود.

اگر $\mu=\mu_0$ معلوم باشد، از معادله زیر:

$$n\alpha^2\hat{\sigma}^4+n\hat{\sigma}^2-\sum_{i=1}^{n}(X_i^w-\mu_0)^2=0$$

برآورد حداکثر درستنمایی $\sigma^۲$ به صورت زیر به دست می‌آید:

$$\hat{\sigma}^۲ = \left[۱ + \sqrt{۱ + ۴\alpha^۲ \frac{\sum_{i=۱}^{n}(X_i^w - \mu_\circ)^۲}{n}} \right] /۲$$

۲.۳ تابع وزن توانی

اگر تابع وزن $w(x,\alpha) = x^\alpha$ باشد در آن صورت $E[w(X,\alpha)] = E[X^\alpha] = \mu'_\alpha$ همیشه موجود است ($\alpha > \circ$) و برابر گشتاور غیرمرکزی مرتبه α است. بنابراین از رابطه (۲) داریم:

$$f^w(x,\mu,\sigma,\alpha) = \frac{x^\alpha}{\sigma\sqrt{۲\pi}\mu'_\alpha} \exp\left\{-\frac{(x-\mu)^۲}{۲\sigma^۲}\right\} \qquad (۸)$$

با جایگذاری در رابطه‌های (۳) می‌توان به سادگی برآوردهای حداکثر درستنمایی را نیز به دست آورد. درحالت خاص اریب اندازه ۱ $= \alpha$، تابع چگالی نرمال وزنی X^w، غیر نرمال و به صورت

$$f^w(x,\mu,\sigma) = \frac{x}{\sigma\mu\sqrt{۲\pi}} \exp\left\{-\frac{(x-\mu)^۲}{۲\sigma^۲}\right\}$$

با میانگین $E(X^w) = \mu + \frac{\sigma^۲}{\mu}$ و واریانس $\frac{\mu'_۳}{\mu} - [E(X^w)]^۲$ به دست می‌آید.

در اینجا اگر $\sigma = \sigma_\circ$ معلوم باشد برآورد حداکثر درستنمایی μ از رابطه

$$\bar{X}^w = \hat{\mu} + \frac{\sigma_\circ^۲}{\hat{\mu}}$$

و اگر σ نامعلوم باشد برآوردها از روش عددی با محاسبات عددی دو رابطه زیر به دست می‌آیند:

$$\bar{X}^w = \hat{\mu} + \frac{\hat{\sigma}^۲}{\hat{\mu}}, \qquad \hat{\sigma}^۲ = \frac{\sum_{i=۱}^{n}(X_i^w - \hat{\mu})^۲}{n}$$

[DOR: 20.1001.1.17358183.1386.1.1.13.0]

۴ روشهای اصلاحی در برآورد پارامترها با تابع وزن وابسته به پارامتر جامعه

در این بخش فرض می‌شود که تابع وزن به پارامترهای توزیع نرمال وابسته است.

۱.۴ تابع وزن نمایی

اگر تابع وزن به صورت:

$$w(x,\alpha,\sigma)=\exp\left\{\frac{\alpha x}{\sigma^{۲}}\right\}$$

باشد، که در آن تابع وزن به $\sigma^{۲}$ وابسته است، به سادگی می‌توان نشان داد X^w دارای توزیع نرمال با میانگین $\mu+\alpha$ و واریانس $\sigma^{۲}$ خواهد بود. لذا برآوردهای حداکثر درستنمایی پارامترهای μ و $\sigma^{۲}$ به صورت زیر به دست می‌آیند:

$$\hat{\mu}=\bar{X}^w-\alpha, \qquad \hat{\sigma^{۲}}=\frac{\sum_{i=۱}^{n}(X_i^w-\bar{X}^w)^{۲}}{n}$$

از اینجا می‌توان نتیجه گرفت که $\hat{\mu}$ و $S_w^{۲}=\frac{\sum_{i=۱}^{n}(X_i^w-\bar{X}^w)^{۲}}{n-۱}$ برآوردهای نااریب برای μ و $\sigma^{۲}$ هستند. نتایج بالا را می‌توان با تعریف تابع وزن

$$w(x,\alpha,\mu,\sigma)=\exp\left\{\beta\Big(\frac{x-\mu}{\sigma}\Big)^{۲}+\alpha\Big(\frac{x-\mu}{\sigma}\Big)\right\}$$

تکمیل کرد $(\beta<۱/۲)$ و نشان داد که برای X^w توزیع نرمال با میانگین $\mu+\frac{\alpha\sigma}{۱-۲\beta}$ و واریانس $\frac{\sigma^{۲}}{۱-۲\beta}$ به دست می‌آید. حالت های خاص آن با $\alpha=۰$ توزیع نرمال با میانگین μ و واریانس $\frac{\sigma^{۲}}{۱-۲\beta}$ و برای $\beta=۰$ توزیع نرمال با میانگین $\mu+\alpha\sigma$ و واریانس $\sigma^{۲}$ است.

۲.۴ تابع وزن توانی

اگر تابع وزن به صورت $w(x;\mu)=(x-\mu)^{\alpha}$ باشد، تابع چگالی احتمال X^w عبارت است از:

$$f^w(x,\mu,\sigma,\alpha)=\frac{(x-\mu)^{\alpha}}{\sigma\mu_{\alpha}\sqrt{۲\pi}}\exp\left\{-\frac{(x-\mu)^{۲}}{۲\sigma^{۲}}\right\} \qquad (۹)$$

که در آن μ_α گشتاور مرکزی مرتبه αی توزیع نرمال با میانگین μ و واریانس $\sigma^۲$ است. برای وقتی که α عدد طبیعی و زوج باشد، چگالی بالا یک خانواده از چگالی‌ها را معرفی می‌کند که دو مد متقارن نسبت به میانگین دارند. این خانواده را توزیع نرمال دوقلوی مرتبه α نامیده و با $DN(\alpha, \mu, \sigma)$ نمایش می‌دهیم. در بخش بعد بعضی از خواص توزیع نرمال دوقلوی مرتبه ۲ بحث می‌شود.

۵ نرمال دو قلوی مرتبهٔ دو

نرمال دوقلوی مرتبهٔ دو در حقیقت یک توزیع نرمال وزنی با وزن $w(x,\mu) = (x-\mu)^۲$ می‌باشد. این وزن زمانی اتفاق می‌افتد که داده‌ها بدون توجه به مقادیر آنها بلکه براساس مجذور انحراف از مرکز ثبت شوند. این توزیع یک توزیع دومدی متقارن با میانگین μ و واریانس $۳\sigma^۲$ با چگالی زیر است:

$$f^w(x,\mu,\sigma) = \frac{(x-\mu)^۲}{\sigma^۳\sqrt{۲\pi}} \exp\left\{-\frac{(x-\mu)^۲}{۲\sigma^۲}\right\}, \quad -\infty < x < \infty \qquad (۱۰)$$

و تا کنون به وسیله آمارشناسان پیشنهاد نشده است. در عمل ممکن است توزیع مشاهدات پرت و خارج از کنترل از این توزیع پیروی کنند. با تبدیل $Z^w = \frac{X^w-\mu}{\sigma}$ توزیع نرمال دو قلوی شبه استاندارد با چگالی زیر به دست می‌آید:

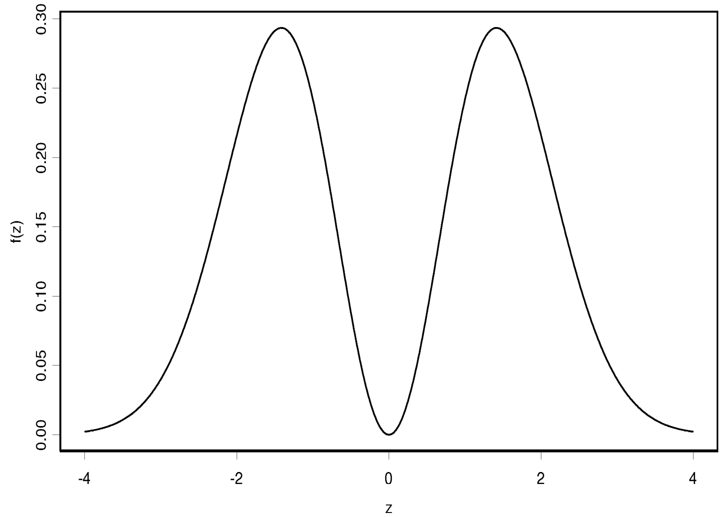
$$f^w(z) = \frac{z^۲}{\sqrt{۲\pi}} \exp\left\{-\frac{z^۲}{۲}\right\}, \quad -\infty < z < \infty \qquad (۱۱)$$

به سادگی می‌توان نشان داد که

$$Z^w = \frac{X^w-\mu}{\sigma} \sim DN(۲, ۰, ۱)$$

نمودار چگالی آن در شکل ۱ نشان داده شده است. همان طور که در شکل ۱ پیداست توزیع حول صفر متقارن و دارای دو مد است. بااندکی محاسبات ریاضی و استفاده از (۱۱) قضایای زیر قابل اثبات هستند.

[DOR: 20.1001.1.17358183.1386.1.1.13.0]



شکل ۱: نمودار تابع چگالی احتمال نرمال دوقلوی شبه استاندارد

قضیه ۱ : اگر Z^w دارای توزیع نرمال دوقلوی شبه استاندارد باشد در آن صورت تابع توزیع آن برابر است با:

$$F_{Z^w}(z) = \phi(z) - z\phi(z)$$

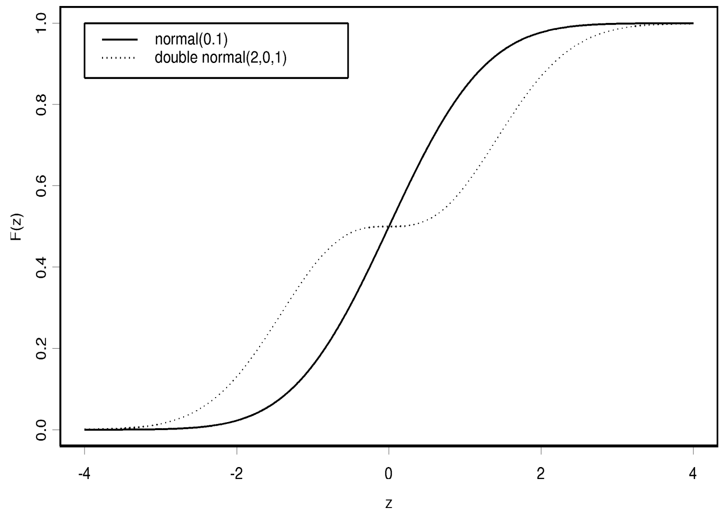
که در اینجا $\phi(z)$ و $\phi(z)$ به ترتیب تابع چگالی احتمال و تابع توزیع تجمعی نرمال استاندارد هستند.

قضیه ۲ : اگر Z^w دارای توزیع نرمال دوقلوی شبه استاندارد باشد برای $k = 0, 1, \ldots$ داریم:

$$E(X^w)^n = \begin{cases} \frac{\Gamma(k+1+\frac{1}{2})}{\Gamma(\frac{1}{2})} 2^{k+1} & n = 2k \\ 0 & n = 2k+1 \end{cases}$$

و تابع مولد گشتاور نیز موجود بوده عبارت است از

$$M_{Z^w}(t) = (t^2+1)\exp\{\frac{1}{2}t^2\}.$$



شکل ۲: نمودار تابع توزیع نرمال استاندارد و نرمال دوقلوی شبه استاندارد

تابع توزیع تجمعی نرمال استاندارد $\phi(z)$ و نرمال دوقلوی شبه استاندارد $F_{Z^w}(z)$ در شکل ۲ نشان داده شده است. همان گونه که از شکل ۲ پیداست برای مقادیر کمتر از صفر مقدار تابع توزیع نرمال دوقلوی شبه استاندارد (منحنی خط چین) و برای مقادیر بیشتر از صفر مقدار تابع توزیع نرمال استاندارد (منحنی با خط سیاه) بیشتر است. سطح زیر منحنی نرمال دوقلوی شبه استاندارد بین صفر تا z با استفاده از نرم افزار Splus تا چهار رقم اعشار محاسبه و در جدول ۱ نشان داده شده است.

قضیه ۳ : اگر Z^w دارای توزیع نرمال دوقلوی شبه استاندارد باشد، متغیر تصادفی $Y^w = {Z^w}^2$ دارای توزیع خی‌دو χ^2 با سه درجه آزادی است.

قضیه ۴ : اگر Z^w دارای نرمال دوقلوی شبه استاندارد باشد در آن صورت $Y^w = aZ^w + b$ برای $a \neq 0$ دارای توزیع نرمال دوقلوی مرتبه دو با پارامترهای b و $|a|$ می‌باشد.

فرع ۱ : هر ترکیب خطی از توزیع نرمال دوقلوی مرتبهٔ دو، یک نرمال دوقلوی

[DOR: 20.1001.1.17358183.1386.1.1.13.0]

مرتبهٔ دو می‌باشد.

فرع ۲ : اگر X^w دارای توزیع نرمال دوقلوی مرتبهٔ دو با پارامترهای μ و σ باشد، درآن صورت

$$Z^w = \frac{X^w - \mu}{\sigma} \sim DN(2, 0, 1)$$

فرع ۳ : اگر X^w دارای توزیع نرمال دوقلوی مرتبهٔ دو با پارامترهای μ و σ باشد، در آن صورت تابع مولد گشتاور و تابع توزیع X^w برابر است با

$$\begin{aligned} M_{X^w}(t) &= (t^2+1)\exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right) \\ F_{X^w}(x) &= \phi\Big(\frac{x-\mu}{\sigma}\Big) - \Big(\frac{x-\mu}{\sigma}\Big)\phi\Big(\frac{x-\mu}{\sigma}\Big). \end{aligned}$$

اگر $\mu = \mu_0$ معلوم باشد، برآورد حداکثر درستنمایی σ^2 برابر است با $\hat{\sigma}^2 = \frac{\sum_{i=1}^n (X_i{}^w - \mu_0)^2}{3n}$ ، توزیع $\frac{3n\hat{\sigma}^2}{\sigma^2}$ خی‌دو با $3n$ درجه آزادی است ولی اگر هر دو پارامتر نامعلوم باشند از معادلات زیر برآورد پارامترها با روشهای عددی به‌دست می‌آیند.

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (X_i{}^w - \hat{\mu})^2}{3n}$$

و

$$\sum_{i=1}^n \frac{1}{X_i{}^w - \hat{\mu}} + \frac{\sum_{i=1}^n (X_i{}^w - \hat{\mu})}{\hat{\sigma}^2} = 0$$

ماتریس اطلاع فیشر برای پارامترهای این توزیع برابراست با

$$I(\mu, \sigma) = \begin{pmatrix} \frac{3n}{\sigma^2} & 0 \\ 0 & \frac{6n}{\sigma^2} \end{pmatrix}$$

توزیع مجانبی توام برآوردهای حداکثر درستنمایی پارامترهای نرمال دوقلو، نرمال دو متغیره با بردار میانگین پارامترها و ماتریس کوواریانس معکوس اطلاع فیشر می‌باشد، که به‌صورت زیر است:

$$\begin{pmatrix} \hat{\mu} \\ \hat{\sigma} \end{pmatrix} = N_2\Big(\begin{pmatrix} \mu \\ \sigma \end{pmatrix}, \begin{pmatrix} \frac{\sigma^2}{3n} & 0 \\ 0 & \frac{\sigma^2}{6n} \end{pmatrix}\Big).$$

جدول ۱: سطح زیر منحنی نرمال دوقلوی شبه استاندارد بین صفر و z

z	۰/۰۰	۰/۰۱	۰/۰۲	۰/۰۳	۰/۰۴	۰/۰۵	۰/۰۶	۰/۰۷	۰/۰۸	۰/۰۹
۰/۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۱	۰/۰۰۰۱
۰/۱	۰/۰۰۰۱	۰/۰۰۰۲	۰/۰۰۰۲	۰/۰۰۰۳	۰/۰۰۰۴	۰/۰۰۰۴	۰/۰۰۰۵	۰/۰۰۰۶	۰/۰۰۰۸	۰/۰۰۰۹
۰/۲	۰/۰۰۱۱	۰/۰۰۱۲	۰/۰۰۱۴	۰/۰۰۱۶	۰/۰۰۱۸	۰/۰۰۲۰	۰/۰۰۲۳	۰/۰۰۲۶	۰/۰۰۲۹	۰/۰۰۳۲
۰/۳	۰/۰۰۳۵	۰/۰۰۳۸	۰/۰۰۴۲	۰/۰۰۴۶	۰/۰۰۵۰	۰/۰۰۵۵	۰/۰۰۶۰	۰/۰۰۶۵	۰/۰۰۷۰	۰/۰۰۷۵
۰/۴	۰/۰۰۸۱	۰/۰۰۸۷	۰/۰۰۹۳	۰/۰۱۰۰	۰/۰۱۰۷	۰/۰۱۱۴	۰/۰۱۲۲	۰/۰۱۲۹	۰/۰۱۳۷	۰/۰۱۴۶
۰/۵	۰/۰۱۵۴	۰/۰۱۶۳	۰/۰۱۷۳	۰/۰۱۸۲	۰/۰۱۹۲	۰/۰۲۰۲	۰/۰۲۱۳	۰/۰۲۲۴	۰/۰۲۳۵	۰/۰۲۴۶
۰/۶	۰/۰۲۵۸	۰/۰۲۷۰	۰/۰۲۸۳	۰/۰۲۹۶	۰/۰۳۰۹	۰/۰۳۲۲	۰/۰۳۳۶	۰/۰۳۵۰	۰/۰۳۶۵	۰/۰۳۷۹
۰/۷	۰/۰۳۹۵	۰/۰۴۱۰	۰/۰۴۲۶	۰/۰۴۴۲	۰/۰۴۵۸	۰/۰۴۷۵	۰/۰۴۹۲	۰/۰۵۱۰	۰/۰۵۲۷	۰/۰۵۴۶
۰/۸	۰/۰۵۶۴	۰/۰۵۸۳	۰/۰۶۰۲	۰/۰۶۲۱	۰/۰۶۴۱	۰/۰۶۶۱	۰/۰۶۸۱	۰/۰۷۰۱	۰/۰۷۲۲	۰/۰۷۴۳
۰/۹	۰/۰۷۶۵	۰/۰۷۸۶	۰/۰۸۰۸	۰/۰۸۳۱	۰/۰۸۵۳	۰/۰۸۷۶	۰/۰۸۹۹	۰/۰۹۲۲	۰/۰۹۴۶	۰/۰۹۷۰
۱/۰	۰/۰۹۹۴	۰/۱۰۱۸	۰/۱۰۴۳	۰/۱۰۶۷	۰/۱۰۹۲	۰/۱۱۱۸	۰/۱۱۴۳	۰/۱۱۶۹	۰/۱۱۹۵	۰/۱۲۲۱
۱/۱	۰/۱۲۴۷	۰/۱۲۷۳	۰/۱۳۰۰	۰/۱۳۲۷	۰/۱۳۵۴	۰/۱۳۸۱	۰/۱۴۰۸	۰/۱۴۳۶	۰/۱۴۶۳	۰/۱۴۹۱
۱/۲	۰/۱۵۱۹	۰/۱۵۴۷	۰/۱۵۷۵	۰/۱۶۰۴	۰/۱۶۳۲	۰/۱۶۶۰	۰/۱۶۸۹	۰/۱۷۱۸	۰/۱۷۴۶	۰/۱۷۷۵
۱/۳	۰/۱۸۰۴	۰/۱۸۳۳	۰/۱۸۶۲	۰/۱۸۹۱	۰/۱۹۲۱	۰/۱۹۵۰	۰/۱۹۷۹	۰/۲۰۰۸	۰/۲۰۳۸	۰/۲۰۶۷
۱/۴	۰/۲۰۹۶	۰/۲۱۲۶	۰/۲۱۵۵	۰/۲۱۸۴	۰/۲۲۱۴	۰/۲۲۴۳	۰/۲۲۷۲	۰/۲۳۰۲	۰/۲۳۳۱	۰/۲۳۶۰
۱/۵	۰/۲۳۸۹	۰/۲۴۱۸	۰/۲۴۴۷	۰/۲۴۷۶	۰/۲۵۰۵	۰/۲۵۳۴	۰/۲۵۶۳	۰/۲۵۹۲	۰/۲۶۲۰	۰/۲۶۴۹
۱/۶	۰/۲۶۷۷	۰/۲۷۰۶	۰/۲۷۳۴	۰/۲۷۶۲	۰/۲۷۹۰	۰/۲۸۱۸	۰/۲۸۴۶	۰/۲۸۷۳	۰/۲۹۰۱	۰/۲۹۲۸
۱/۷	۰/۲۹۵۶	۰/۲۹۸۳	۰/۳۰۱۰	۰/۳۰۳۶	۰/۳۰۶۳	۰/۳۰۹۰	۰/۳۱۱۶	۰/۳۱۴۲	۰/۳۱۶۸	۰/۳۱۹۴
۱/۸	۰/۳۲۲۰	۰/۳۲۴۵	۰/۳۲۷۰	۰/۳۲۹۶	۰/۳۳۲۰	۰/۳۳۴۵	۰/۳۳۷۰	۰/۳۳۹۴	۰/۳۴۱۸	۰/۳۴۴۲
۱/۹	۰/۳۴۶۶	۰/۳۴۹۰	۰/۳۵۱۳	۰/۳۵۳۶	۰/۳۵۵۹	۰/۳۵۸۲	۰/۳۶۰۵	۰/۳۶۲۷	۰/۳۶۴۹	۰/۳۶۷۱
۲/۰	۰/۳۶۹۳	۰/۳۷۱۴	۰/۳۷۳۵	۰/۳۷۵۶	۰/۳۷۷۷	۰/۳۷۹۸	۰/۳۸۱۸	۰/۳۸۳۹	۰/۳۸۵۸	۰/۳۸۷۸
۲/۱	۰/۳۸۹۸	۰/۳۹۱۷	۰/۳۹۳۶	۰/۳۹۵۵	۰/۳۹۷۴	۰/۳۹۹۲	۰/۴۰۱۰	۰/۴۰۲۸	۰/۴۰۴۶	۰/۴۰۶۳
۲/۲	۰/۴۰۸۱	۰/۴۰۹۸	۰/۴۱۱۴	۰/۴۱۳۱	۰/۴۱۴۷	۰/۴۱۶۴	۰/۴۱۸۰	۰/۴۱۹۵	۰/۴۲۱۱	۰/۴۲۲۶
۲/۳	۰/۴۲۴۱	۰/۴۲۵۶	۰/۴۲۷۱	۰/۴۲۸۵	۰/۴۲۹۹	۰/۴۳۱۴	۰/۴۳۲۷	۰/۴۳۴۱	۰/۴۳۵۴	۰/۴۳۶۸
۲/۴	۰/۴۳۸۱	۰/۴۳۹۳	۰/۴۴۰۶	۰/۴۴۱۸	۰/۴۴۳۱	۰/۴۴۴۳	۰/۴۴۵۴	۰/۴۴۶۶	۰/۴۴۷۷	۰/۴۴۸۹
۲/۵	۰/۴۵۰۰	۰/۴۵۱۱	۰/۴۵۲۱	۰/۴۵۳۲	۰/۴۵۴۲	۰/۴۵۵۲	۰/۴۵۶۲	۰/۴۵۷۲	۰/۴۵۸۲	۰/۴۵۹۱
۲/۶	۰/۴۶۰۰	۰/۴۶۰۹	۰/۴۶۱۸	۰/۴۶۲۷	۰/۴۶۳۶	۰/۴۶۴۴	۰/۴۶۵۲	۰/۴۶۶۰	۰/۴۶۶۸	۰/۴۶۷۶
۲/۷	۰/۴۶۸۴	۰/۴۶۹۱	۰/۴۶۹۹	۰/۴۷۰۶	۰/۴۷۱۳	۰/۴۷۲۰	۰/۴۷۲۷	۰/۴۷۳۴	۰/۴۷۴۰	۰/۴۷۴۷
۲/۸	۰/۴۷۵۳	۰/۴۷۵۹	۰/۴۷۶۵	۰/۴۷۷۱	۰/۴۷۷۷	۰/۴۷۸۲	۰/۴۷۸۸	۰/۴۷۹۳	۰/۴۷۹۸	۰/۴۸۰۴
۲/۹	۰/۴۸۰۹	۰/۴۸۱۴	۰/۴۸۱۹	۰/۴۸۲۳	۰/۴۸۲۸	۰/۴۸۳۲	۰/۴۸۳۷	۰/۴۸۴۱	۰/۴۸۴۵	۰/۴۸۵۰
۳/۰	۰/۴۸۵۴	۰/۴۸۵۷	۰/۴۸۶۱	۰/۴۸۶۵	۰/۴۸۶۹	۰/۴۸۷۲	۰/۴۸۷۶	۰/۴۸۷۹	۰/۴۸۸۳	۰/۴۸۸۶
۳/۱	۰/۴۸۸۹	۰/۴۸۹۲	۰/۴۸۹۵	۰/۴۸۹۸	۰/۴۹۰۱	۰/۴۹۰۴	۰/۴۹۰۷	۰/۴۹۰۹	۰/۴۹۱۲	۰/۴۹۱۴
۳/۲	۰/۴۹۱۷	۰/۴۹۱۹	۰/۴۹۲۲	۰/۴۹۲۴	۰/۴۹۲۶	۰/۴۹۲۸	۰/۴۹۳۰	۰/۴۹۳۲	۰/۴۹۳۴	۰/۴۹۳۶
۳/۳	۰/۴۹۳۸	۰/۴۹۴۰	۰/۴۹۴۲	۰/۴۹۴۴	۰/۴۹۴۵	۰/۴۹۴	۰/۴۹۴۹	۰/۴۹۵۰	۰/۴۹۵۲	۰/۴۹۵۳
۳/۴	۰/۴۹۵۵	۰/۴۹۵۶	۰/۴۹۵۸	۰/۴۹۵۹	۰/۴۹۶۰	۰/۴۹۶۱	۰/۴۹۶۳	۰/۴۹۶۴	۰/۴۹۶۵	۰/۴۹۶۶
۳/۵	۰/۴۹۶۷	۰/۴۹۶۸	۰/۴۹۶۹	۰/۴۹۷۰	۰/۴۹۷۱	۰/۴۹۷۲	۰/۴۹۷۳	۰/۴۹۷۴	۰/۴۹۷۵	۰/۴۹۷۶
۳/۶	۰/۴۹۷۶	۰/۴۹۷۷	۰/۴۹۷۸	۰/۴۹۷۹	۰/۴۹۷۹	۰/۴۹۸۰	۰/۴۹۸۱	۰/۴۹۸۱	۰/۴۹۸۲	۰/۴۹۸۳
۳/۷	۰/۴۹۸۳	۰/۴۹۸۴	۰/۴۹۸۴	۰/۴۹۸۵	۰/۴۹۸۵	۰/۴۹۸۶	۰/۴۹۸۶	۰/۴۹۸۷	۰/۴۹۸۷	۰/۴۹۸۸
۳/۸	۰/۴۹۸۸	۰/۴۹۸۹	۰/۴۹۸۹	۰/۴۹۸۹	۰/۴۹۹۰	۰/۴۹۹۰	۰/۴۹۹۰	۰/۴۹۹۱	۰/۴۹۹۱	۰/۴۹۹۱
۳/۹	۰/۴۹۹۲	۰/۴۹۹۲	۰/۴۹۹۲	۰/۴۹۹۳	۰/۴۹۹۳	۰/۴۹۹۳	۰/۴۹۹۳	۰/۴۹۹۴	۰/۴۹۹۴	۰/۴۹۹۴

۶ بحث و نتیجه‌گیری

در این مقاله استنباط در باره توزیع نرمال تحت نمونه گیری وزنی با چند وزن خاص بررسی گردید. برآوردهای گشتاوری و حداکثر درستنمایی تحت این وزنها به دست آمد. یک توزیع تحت نام توزیع نرمال دوقلو به عنوان یک توزیع وزنی نرمال تحت وزن خاصی معرفی گردید. خواص این توزیع بررسی و جدول آن با استفاده از برنامه‌های کامپیوتری به دست آمد. در ادامه این کار تحقیقاتی می‌توان وزن‌های ارائه شده در این مقاله و یا حتی وزن‌های مناسب دیگر را برای توزیع‌های مشابه دنبال کرد. واضح است که در صورت مواجهه با نمونه‌گیری وزنی استفاده از

چنین توزیع‌هایی می‌تواند سبب اصلاح روش‌های برآورد پارامترها شود.

مراجع

Bayarri, M. J. and Berger, J. O. (1998), *Robust Bayesian Analysis of Selection Models.*, Annals of Statistics, **26**, 645-659.

Bayarri, M. J. and DeGroot, M. H. (1987), *Bayesian Analysis of Selection Models,* The Statistician, **36**, 137-146.

Brown, M. (1972), *Low Density Traffic Stream.*, Advances in Applied Probability **4**, 177-192.

Celestin, C. and Mizere, D. (2005), *Overdispersion and Underdispersion Characterization of Weighted Poisson Distributions.*, LMA Technical Report No. 0523.

Chakraborty, S. and Das, K. (2006),*On Some Properties of a Class of Weighted Quasi-Binomial Distributions.*, Journal of Statistical Planing and Inference, **136**, 159-182.

Chen, S . X. and Cowling, A.(2001), *Measurment Error in Line Transect Surveys where Detectablity Varies with Distance and Size.*, Biometrics, **57**, 732-742.

Cook, D. R., Martin, F. B. (1974), *A Model for Quadrat Sampling with Visibility bais.*, Journal of the American Statistical Association, **69**, 345-349

Fisher, R. A. (1934), *The Effect of Models of Ascertainment upon the Estimation of Frequencies.*, Annal Eugen, **6**, 13-25.

Guillamon, A., Navarro, J. and Ruiz, J. M. (1998). *Kernel Density Estimation Using Weighted Data.*, Communication in Statistics -Theory and Methods, **27(9)**, 2123-2135.

Gupta, R. C. and Akman, H. O. (1995), *On the Reliability Studies of a Weighted Inverse Gaussian Model*, Journal of Statistical Planning and Inference, **48**, 69-83.

Jacobo, D. (2003), *Large Sample Results under Biased Sampling when Covariables are Present.*, Statistics and Probability Letters, **63**, 287-293.

Morrison, D. G. (1973), *Some Results for Waiting Times with an Application to Survey Data.*, American Statistician, **27**, 226-227.

Patil, G. P. and Ord, J. K. (1978), *On Size-Biased Sampling and Related Form-Invariant Weighted Distributions.*, Sankhya, **38**, 48-61.

Patil, G. D. and Rao, C. R. (1978). *Weighted Distributions and Size-Biased Sampling with Application to Wildlife Populations and Human Families,* Biometrics, **34**, 179-184.

Rao, C. R. (1965), *On Discrete Distributions Arising out of Methods of Ascertainment.*, Sankhya A., **27**, 311-324.

Sansgiry, P. S. and Akman, O. (2001), *Reliability Estimation Via Length-Biased Transformation.*, Communication in Statistics -Theory and Methods, **30(11)**, 2473-2479.

Smith, W. and Parnez, M. (1994), *Mean Street: The Median of a Size-Biased Sample and the Population Mean.*, Journal of the American Statistical Association, **48,** 106-110.

[DOR: 20.1001.1.17358183.1386.1.1.13.0]

Unnikrishnan Nair, N. and Sunoj, S. M. (2003), *Form-Invariant Bivariate Weighted Models.*, Statistics, **37(3)**, 259-269.

Warren, W. G. (1975), *Statistical Distributions in Forestry and Forest Products Research,* in Statistical Distributions in Scientific Work, Patil, G. P., Kotz, S., Ord, K. eds., **2**, 369-384.