دانشگاه پیام نور

پایان نامه برای دریافت درجه کارشناسی ارشد در رشته آمار ریاضی

دانشکده علوم پایه

گروه علمی آمار

عنوان پایان نامه:

آزمون های نیکویی برازش برای توزیع نرمال- چوله

استاد راهنما:

سرکار خانم دکتر نرگس عباسی

استاد مشاور:

آقای دکتر علیرضا عرب پور

نگارش:

پریا جنگی پور افشار

شهریور ۱۳۸۹

تقدیم به پدرم

چو کوهی استوار تکیه گاه زندگی من است

تقدیم به مادرم

که حرف به حرف محبت را برایم هجی کرد

تقدیم به همسرم

اگر محبت و حمایت او نبود هرگز به این مهم دست نمی یافتم

تکیه بر تقوی و دانش در طریقت کافریست

راهرو گر صد هنر دارد توکل بایدش

- سپاس و ستایش از آن پروردگار بزرگ و بلند مرتبه ای است که انسان را آفرید و خرد را در نهادش قرار داد تا او نیز بیافریند.

- سپاس فراوان از اساتید ارجمند جناب آقای دکتر علیرضا عرب پور و سرکار خانم دکتر نرگس عباسی که با راهنمایی ها و عنایاتشان، بنده را در طول این پایان نامه همراهی فرمودند.

چکیده

معمولاً تحلیل‌های آماری با فرض نرمال بودن داده‌ها صورت می‌گیرد، در حالی‌که در عمل در بسیاری از موارد داده‌ها نامتقارن بوده و دیگر نمی‌توان از توزیع نرمال برای تحلیل آنها استفاده کرد، در این موارد اغلب سعی می‌شود با استفاده از تبدیلی، توزیع داده‌ها را به نرمال نزدیک کرده و سپس تحلیل داده‌ها را انجام داد. اما این روش، خود با مسائل و دشواری‌هایی از جمله چگونگی انتخاب تبدیل مناسب و اریبی برآوردها مواجه است.

با توجه به این مسأله، به‌کارگیری توزیع‌های نامتقارنی که خواصی مشابه توزیع نرمال دارند، مورد توجه قرار گرفته است. از جمله مهم‌ترین این توزیع‌ها، توزیع نرمال-چوله است که در پایان‌نامه، ضمن معرفی و بررسی خواص این توزیع، آزمون‌های نیکویی برازش در این توزیع ارائه می‌شود. سپس در انتها معرفی توزیع چوله t وآزمون‌های نیکویی برازش برای این توزیع بیان می‌شود.

واژه‌های کلیدی: توزیع نرمال-چوله، آزمون‌های نیکویی برازش، توزیع t -چوله.

فهرست

پیشگفتار

توزیع نرمال–چوله یک متغیری با مقادیر حقیقی، همراه با یک پارامتر اضافی که میزان چولگی را در کنترل دارد، توسط آزالینی[۱] (1985,1986) و در کتاب آزالینی و کاپیتانیو[۲] (1999) و آزالینی و دالاواله[۳] (1996) مطرح و تعمیم چند متغیری آن را مورد بررسی قرار دادند. این مدل در بردارنده ویژگی‌های جبری مشخص و مشابه با ویژگی‌های جبری توزیع نرمال معمولی است (رجوع شود به آزالینی 2005) اما هنگامی که فرض نرمال بودن با توجه به چولگی داده‌ها غیر قابل توجیه است یک مدل مناسب‌تر محسوب می‌شود. توزیع نرمال–چوله و تعمیم آن در بسیاری از مباحث علمی کاربرد دارد کتاب جینتون[۴] (2004) و آزالینی (2005) جزئیات مشروح پیشرفت در این شاخه از تحقیق را همراه با مثال‌های تشریحی و کاربرد آن در داده‌های حقیقی بیان شده است.

جنبه‌های مشخص آزمون نیکویی برازش برای توزیع نرمال–چوله امروزه مورد توجه قرار گرفته است. سالوان[۵] (1986) مشکلات آزمون نرمال بودن را در مدل نرمال–چوله بررسی کرد و نشان داد که ضریب چولگی نیرومندترین آماره‌ی آزمون مکان و مقیاس پایا است.

پوسی[۶] (2000) از همین آماره استفاده کرد تا فاصله جبری را از توزیع حدی نیم نرمال را آزمون کند.
گوپتا و چن[۷] (2001) از آماره‌هایی براساس تابع توزیع تجربی (EDF) استفاده کردند تا فرضیه صفر یک توزیع نرمال–چوله را در موقعیتی آزمایش کنند که در آن پارامترهای توزیع کاملاً مشخص‌اند.
تاکنون هیچ مرجعی در کتاب‌ها برای مسئله آزمون نیکویی برازش یعنی برای آزمون فرضیه صفر که یک نمونه تصادفی از یک جمعیت نرمال–چوله است و پارامترهای آن از داده‌ها به دست آمده است مطرح نشده است. خاطر نشان می‌کنیم که در اثر منتشر شده دالاواله (تحت چاپ) اصطلاح آماره‌ی اندرسن – دارلینگ برای مبحث آزمون نیکویی برازش مطرح شده است.

---

[1]. Azzalini
[2]. Capitanio
[3]. Dalla Valle
[4]. Genton
[5]. Salvan
[6]. Pewsy
[7]. Gupta and Chen

فصل اول:

مروری بر مفاهیم پایه

۱-۱- مقدمه

آزمون نیکویی برازش برای توزیع نرمال-چوله بر پایه‌ی تابع توزیع تجربی بیان می‌شود. این آزمون می‌تواند به خوبی برای توزیع‌های یک متغیره و چند متغیره به کار برده شود.

در این فصل در بخش ۱-۲ تعریف آماره‌ی تابع توزیع تجربی و در بخش ۱-۳ تعریف آزمون نیکویی برازش بیان شده است و در بخش ۱-۴ آزمون‌های نیکویی برازش آورده شده است.

۱-۲- آماره‌ی تابع توزیع تجربی

آماره‌ی تابع توزیع تجربی (EDF)[1] یک تابع پله‌ای است که بر اساس نمونه محاسبه می‌شود به گونه‌ای که تابع توزیع جامعه را برآورد می‌کند. آماره‌ی EDF اختلاف بین تابع توزیع تجربی نمونه و تابع توزیع داده شده را اندازه‌گیری می‌کند. دو نوع آماره EDF وجود دارد: ۱- نوع کولموگروف-اسمیرنوف و ۲- نوع مربعی.

۱-۳-آزمون نیکویی-برازش خی-دو

این آزمون برای داده‌های گسسته و پیوسته، تک متغییره و چند متغییره قابل استفاده است. این آزمون قدیمی‌ترین آزمون نیکویی-برازش می‌باشد که اولین بار توسط کارل پیرسن[2] ارائه گردید. این آزمون فراوانی مشاهده شده و مورد انتظار (بر اساس توزیع فرض صفر) را برای گروه‌های مختلف محاسبه می‌کند که در آن m همان تعداد خانه‌های جدول می‌باشد.

$$\chi^2 = \sum_{i=1}^{m} \frac{(\text{مورد انتظار } i\text{ام}) - (\text{مشاهده } i\text{ام}))^2}{(\text{مقدار مورد انتظار } i\text{ام})}$$

یادآور می‌شویم که داده‌های مشاهده شده بر اساس طبقه بندی یا رده بندی داده‌ها جمع‌آوری می‌شوند، و این به خاطر وجود رابطه‌ی بین توزیع خی-دو و مجموع توان دوم مقادیر z مربوط به خانه‌های مختلف می‌باشد که به این صورت است.

$$Z = \frac{\text{مقدار مورد انتظار} - \text{مقدار مشاهده شده}}{\text{انحراف معیار مقدار مشاهده شده}}$$

[1]. Empirical Characteristic Function
[2]. Karl Pearson

(که مقادیر مورد انتظار همان میانگین نمونه هستند) که این رابطه را می‌توان به صورت زیر نوشت.

$$\chi^2 = \sum_{i=1}^{m} (\text{آماره } z \text{ سلول } i\text{ام})^2$$

این رابطه توسط هاسکینگ[3] برای محاسبه‌ی مقادیر احتمال گشتاورهای خطی مورد استفاده قرار گرفت از آنجا که آزمون خی-دو داده‌ها در قالب رده‌ها دسته‌بندی شده است از این رو قسمتی از اطلاعات داده‌ها را از دست می‌دهد. بنابر این موری پیشنهاد کرد که این آزمون برای توزیع های یک متغیره که هدف آن آزمون برازش یا آزمون بر پایه آماره تابع توزیع تجربی (EDF) است استفاده نشود.

۱-۴-آزمون‌های کلی نیکویی برازش براساس (EDF):

در این بخش آماره‌های نیکویی برازش مطرح شده براساس EDF را بررسی می‌کنیم که برای آزمون فرضیه صفر

$H_0 : F(0)$ نمونه‌ی تصادفی از توزیع $X_1, X_2, ..., X_n$

به‌کار می‌رود. آماره‌ی آزمون براساس EDF تفاوت بین تابع توزیع $F(0)$ که در فرضیه‌ی صفر مطرح شده و EDF (یک تابع پله‌ای که با $F_n(0)$ نشان داده شده است و به‌صورت زیر محاسبه می‌شود) را اندازه‌گیری می‌کنند.

$$F_n(x) = \begin{cases} 0 & x \langle X_{(1)} \\ i/n & X_{(i)} \le x < X_{(i+1)} \\ 1 & X_{(n)} < x \end{cases}$$

که در آن $X_{(1)} \langle X_{(2)} \langle ... \langle X_{(n)}$ آماره‌های ترتیبی هستند. برای مقایسه‌ی دو تابع توزیع چندین آماره مورد استفاده قرارمی‌گیرد. که استفان[4] (۱۹۸۶) آنها را به دو دسته تقسیم کرده است. دسته‌ی کرامر – ون میسز شامل آماره‌ی کرامر – ون میسز W^2، آماره‌ی U^2 واتسون و آماره‌ی اندرسون – دارلینگ A^2 است که به صورت زیر تعریف شده‌اند.

$$W^2 = n\int_{-\infty}^{\infty} \left(\left(F_n(x) - F(x) \right)^2 dF(x) \right)$$

$$U^2 = n\int_{-\infty}^{\infty} \left[F_n(x) - F(x) - \int_{-\infty}^{\infty} \left[F_n(t) - F(t) dF(t) \right] \right]^2 dF(x)$$

$$A^2 = n\int_{-\infty}^{\infty} \left(F_n(x) - F(x) \right)^2 \left[F(x)(1 - F(x)) \right]^{-1} dF(x)$$

---
[3]. Hosking
[4]. Stephens

دسته‌ی کولموگوروف- اسمیرنوف شامل آماره‌های D^-, D^+ ، آماره‌ی کولموگوروف –اسمیرنوف D و آماره‌ی کوپر V است که به صورت زیر تعریف شده‌اند.

$$D^- = \sup\left(F(x) - F_n(x)\right) \qquad D^+ = \sup\left(F_n(x) - F(x)\right) \tag{۱-۱}$$

$$D = \max\left(D^+, D^-\right) \qquad V = D^+ + D^- \tag{۲-۱}$$

استفان (۱۹۸۶) از فرمول‌های ساده زیر برای محاسبه این آماره‌ها استفاده کرد:

$$W^۲ = \sum_{i=۱}^{n}\left(p_{(i)} - \frac{۲i-۱}{۲n}\right)^۲ + \frac{۱}{۱۲n} \tag{۳-۱}$$

$$U^۲ = W^۲ - n\left(\bar{p} - \frac{۱}{۲}\right)^۲ \tag{۴-۱}$$

$$A^۲ = -n - \frac{۱}{n}\sum_{i=۱}^{n}(۲i-۱)\left[\log p_{(i)} + \log\left(۱ - p_{(n+۱-i)}\right)\right] \tag{۵-۱}$$

$$D^+ = \max_i\left(\frac{i}{n} - p_{(i)}\right) \tag{۶-۱}$$

$$D^- = \max_i\left(p_{(i)} - \frac{i-۱}{n}\right) \tag{۷-۱}$$

$$D = \max\left(D^+, D^-\right) \tag{۸-۱}$$

$$V = D^+ + D^- \tag{۹-۱}$$

که در آن $p_{(i)} = F\left(x_{(i)}\right)$ و $\bar{p} = \sum_{i=۱}^{n} p_{(i)} \Big/ n$ مقادیر بزرگ آماره‌ها، تفاوت بین توابع توزیع فرضی و توزیع تجربی را نشان می‌دهد و بنابر این فرضیه صفر رد می‌شود.

به طور کلی هنگامی که مقدار پارامتر توزیع فرضی کاملاً مشخص باشد توزیع نمونه‌ای هر کدام از شاخص‌های آماری EDF به صورت دقیق مشخص می‌شود و جدول‌های نقطه‌های درصدی در دسترس قرار می‌گیرند (رجوع کنید به استفان ۱۹۸۶[۵] جدول ۴,۲) اما هنگامی که مقدار به دست آمده از پارامترهای توزیع نامشخص باشد و باید آنها را از نمونه برآورد کرد توزیع نمونه‌ای هر آماره‌ی EDF بستگی به توزیع آزمون شده، اندازه‌ی نمونه، مقدار واقعی پارامترهای نامشخص و روش‌های به کار رفته برای تخمین پارامترها دارد.

در این شرایط باید از تکنیک‌های مونت کارلو بهره گرفت تا توزیع نمونه‌ای یک آماره آزمون تقریب زده شود. خاطر نشان می‌کنیم که یک تئوری کلی موجود است که می‌تواند برای به دست آوردن توزیع مجانبی آماره‌های $\mathrm{A}^۲$, $\mathrm{U}^۲$, $\mathrm{W}^۲$ مورد استفاده قرار گیرد.

---

۵. Stephens

هنگامی که پارامترهای نامشخص، پارامترهای مکان یا مقیاس باشند و با استفاده از برآوردکننده‌های مکانی و مقیاسی هم متغیر (یعنی برآوردکننده‌های حداکثر درست نمایی) تخمین زده می‌شود، توزیع نمونه‌ای آماره‌ی EDF وابسته به مقدار این پارامترها نیست. در مورد توزیع نرمال-چوله λ,η,ξ که به ترتیب پارامترهای مکان و مقیاس و شکل هستند. برآورد کننده‌های حداکثر درستنمایی $\hat{\lambda},\hat{\eta},\hat{\xi}$ دارای این ویژگی هستند که توزیع $\hat{\lambda},\left(\hat{\xi}-\xi\right)/\hat{\eta}$ فقط بستگی به اندازه نمونه n و مقدار واقعی λ دارد (رجوع کنید به ایستمان وبین ۱۹۷۳[۶]) بنابراین هنگامی که پارامترها برآورد می‌شوند توزیع نمونه‌ای آماره‌های EDF فقط بستگی به مقدار واقعی λ و اندازه نمونه n دارد.

همچنین می‌توان از تکنیک‌های مونت کارلو استفاده کرد تا این توزیع نمونه‌ای پارامترها را برآورد کرد.

۶. Eastman and Bain

فصل دوم:

توزیع نرمال-چوله و خواص آن

۲-۱ مقدمه

توزیع نرمال یکی از محدود توزیع‌های پیوسته است که در آمار از اهمیت زیادی برخوردار است. این توزیع خاصیت‌هایی دارد که باعث افزایش کاربردهای آن در عمل شده است معمولاً در تحلیل‌های آماری فرض نرمال بودن داده‌ها را می‌پذیرند سپس به استنباط آماری می‌پردازد. در بسیاری از موارد توزیع داده‌های واقعی مقداری چولگی دارد. بنابراین ممکن است در عمل فرض نرمال بودن داده‌ها رد شود. هنگامی که داده‌ها نا متقارن هستند. دیگر نمی‌توان ازتوزیع نرمال برای تحلیل آنها استفاده کرد در این موار اغلب سعی می‌شود با استفاده از تبدیلی مناسب توزیع داده‌هارا به نرمال نزدیک کرده و سپس تحلیل روی داده‌های تبدیل یافته انجام گیرد.

پیدا کردن تبدیل مناسب بعضا کاری دشوار و در عمل ممکن است چنین تبدیلی یافت نشود. اگر توزیع نامتقارنی وجود داشته باشد که دارای خواصی مشابه توزیع نرمال باشد این توزیع می‌تواند نقش اساسی در تحلیل داده‌های نامتقارن ایفا کند. یکی از توزیع‌ها که اخیرا معرفی شده و مورد توجه قرار گرفته است و خواصی مشابه توزیع نرمال دارد به توزیع نرمال–چوله معروف شده است. این توزیع نامتقارن دارای یک پارامتر برای تنظیم چولگی است. اگر مقدار این پارامتر صفر گردد توزیع نرمال–چوله به توزیع نرمال تبدیل می‌شود. توزیع نرمال–چوله یک متغیره نخستین بار توسط آزالینی (۱۹۸۵) ارائه شد.

۲-۲ تعاریف و مفاهیم اولیه

در این قسمت روش ساخت تابع توزیع چوله بر اساس یک تابع توزیع متقارن ارائه می‌گردد. برای این منظور ابتدا یک قضیه مهم بیان می‌شود. این قضیه چگونگی ساختن یک توزیع چوله بر اساس توزیعی متقارن را بیان می‌کند. سپس توزیع نرمال-چوله تعریف می‌شود.

قضیه ۲-۲-۱: با فرض اینکه $f_{\circ}$ یک تابع چگالی متقارن حول صفر، G توزیعی مطلقا پیوسته، به گونه‌ای که تابع چگالی متناظر با آن، یعنی G' حول صفر متقارن و W یک تابع فرد باشد آنگاه :

$$f(y) = ۲f_{\circ}(y)G(W(y)) \qquad (۱-۲)$$

یک تابع چگالی است($W,G,f_{\circ}$توابعی یک بعدی هستند.)

اثبات: چون $۲f_{\circ}(y)G(W(y))$ مثبت است کافی است نشان دهیم که $\int_{-\infty}^{\infty} f_{\circ}(y)G(W(y))dy = \frac{۱}{۲}$ است فرض کنید Y, X دو متغیر تصادفی مستقل با چگالی‌های $G', f_{\circ}$ باشند در این صورت

$$P(Y \le W(X)) = E\{I_{(Y \le W(X))}\}$$

$$= E\{E\{I_{(Y \le W(X))}|X\}\}$$

$$= E\{P(Y \le W(X))|X\}$$

$$= \int_{-\infty}^{\infty} P(Y \le W(X)|X = x) f_{\circ}(x)dx$$

$$= \int_{-\infty}^{\infty} P(Y \le W(x)) f_{\circ}(x)dx$$

$$= \int_{-\infty}^{\infty} G(w(x)) f_{\circ}(x)dx \qquad (۱)$$

$$= \int_{-\infty}^{\circ} G(w(x)) f_{\circ}(x)dx + \int_{\circ}^{\infty} G(w(x)) f_{\circ}(x)dx$$

با توجه به فرد بودن $W(x)$

$$= \int_{0}^{\infty} G(-w(x)) f_{0}(x) dx + \int_{0}^{\infty} G(w(x)) f_{0}(x) dx$$

$G(-u) = 1 - G(u)$

$$= \int_{0}^{\infty} \{G(-w(x)) + G(w(x))\} f_{0}(x) dx$$

$$= \int_{0}^{\infty} f_{0}(x) dx = \frac{1}{2} \qquad (2)$$

از تساوی (۱)و (۲)، رابطه (۲–۱) به سادگی نتیجه می‌شود. □

قضیه‌ی ۲–۲–۱ روشی ساده برای تولید توزیع‌های چوله فراهم می‌کند.

نتیجه ۲–۲–۱: با فرض آنکه f_{0} و G' چنان باشند که در قضیه‌ی۲–۲–۱ ذکر شده است، $Y \sim f_{0}\ ,\ X \sim G'$ متغیرهای تصادفی مستقل باشند آنگاه متغیر تصادفی Z، که به صورت

$$Z = \begin{cases} Y & W(Y) \geq X \\ -Y & X > W(Y) \end{cases}$$

تعریف می‌شود دارای تابع چگالی(۲–۱) است.

اثبات:

$$P(Z \leq z) = P(Z \leq z, X \leq W(Y)) + P(Z \leq z, X > W(Y))$$

$$= P(Y \leq z, X \leq W(Y)) + P(-Y \leq z, X > W(Y))$$

$$= \int_{-\infty}^{z} P(X \leq W(Y) | Y = y) f_{0}(y) dy + \int_{-z}^{\infty} P(X > W(Y) | Y = y) f_{0}(y) dy$$

$$= \int_{-\infty}^{z} P(X \leq W(y)) f_{0}(y) dy + \int_{-z}^{\infty} P(X > W(y)) f_{0}(y) dy$$

$$= \int_{-\infty}^{z} G(W(y)) f_{0}(y) dy + \int_{-z}^{\infty} \{1 - G(w(y))\} f_{0}(y) dy$$

$$\Rightarrow f_{Z}(z) = \frac{\partial}{\partial} P(Z \leq z)$$

$$= f_{0}(z) G(W(z)) + f_{0}(-z)\{1 - G(W(-z))\}$$

$$= f_{0}(z) G(W(z)) + f_{0}(z)\{1 - G(W(z))\}$$

$$= 2 f_{0}(z) G(W(z))$$

□

نتیجه ۲-۲-۲: اگر $Z \sim f$, $Y \sim f_{\circ}$ باشند آنگاه $|Y| \overset{d}{=} |Z|$ یعنی $|Y|,|Z|$ هم‌توزیع هستند.

اثبات:

$$h_{|z|}(z) = ۲f_{\circ}(z)G(w(z)) + ۲f_{\circ}(-z)G(W(-z))$$

$$= ۲f_{\circ}(z)\{G(w(z)) + G(-w(z))\}$$

$$= ۲f_{\circ}(z)$$

$$h_{|Y|}(z) = f_{\circ}(z) + f_{\circ}(-z) = ۲f_{\circ}(z)$$

$$\Rightarrow h_{|z|}(z) = h_{|Y|}(z)$$

□

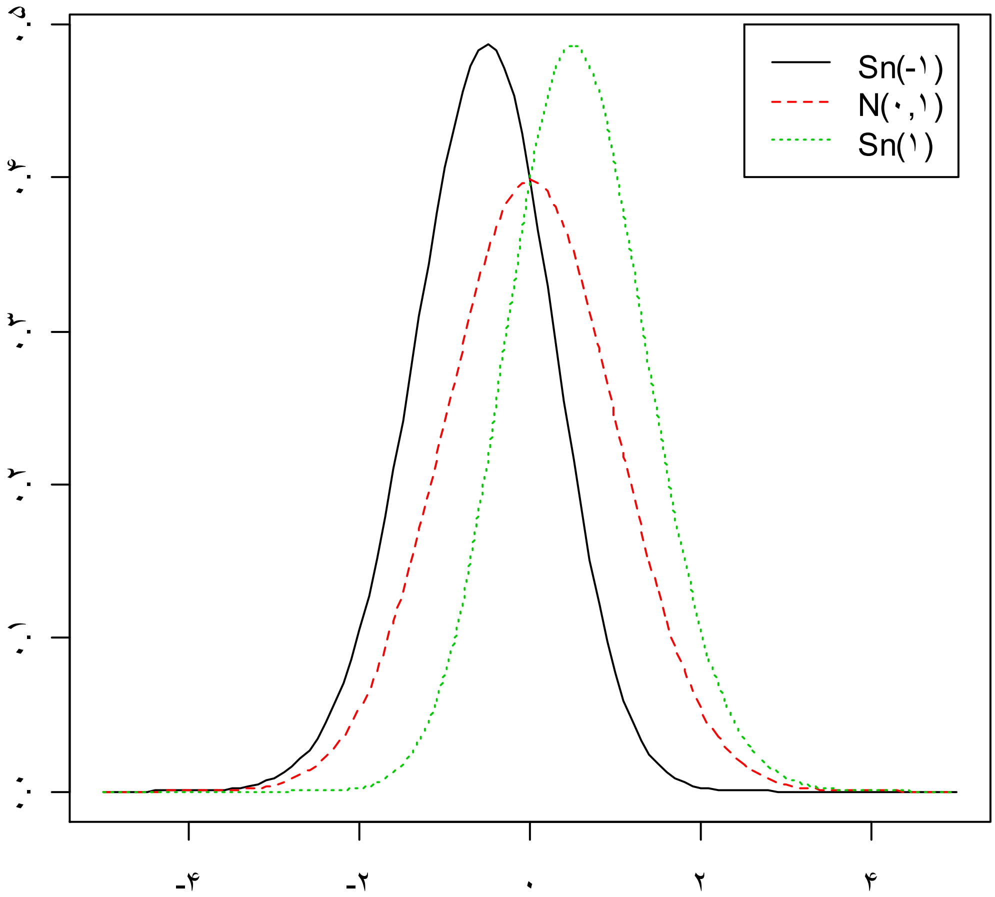
نتیجه ۲-۲-۳: اگر در قضیه‌ی (۲-۲-۱)، $W(y) = \lambda y$ باشد که در آن λ یک عدد حقیقی است آنگاه تابع $۲f_{\circ}(y)G(\lambda y)$، یک تابع چگالی است.

اگر تابع $f_{\circ}$ و تابع توزیع G ثابت باشند به ازای مقادیر مختلف $\lambda \in R$ خانواده‌ای از توزیع‌ها تولید می‌شود که ازای $\lambda = ۰$ متقارن و برابر $f_{\circ}$ است اما به ازای $\lambda \neq ۰$ چوله است. با استفاده از این نتیجه توزیع نرمال – چوله در سال ۱۹۸۵ توسط آزالینی تعریف گردید.

تعریف ۲-۲-۱: (توزیع نرمال – چوله): متغیر تصادفی Z نرمال چوله با پارامتر حقیقی λ نامیده می‌شود. هر گاه تابع چگالی آن به صورت

$$\phi(z;\lambda) = ۲\phi(z)\Phi(\lambda z) \qquad z,\lambda \in R$$

باشد که در آن Φ,ϕ به ترتیب تابع چگالی و تابع توزیع نرمال استاندارد هستند. این توزیع که به صورت $Z \sim SN(\lambda)$ نمایش داده می‌شود، توزیع نرمال – چوله استاندارد (نرمال – چوله آزالینی) با پارامتر شکل λ نامیده می‌شود. پارامتر شکل λ را می‌توان "پارامتر کنترل چولگی" نامیده، تابع چگالی فوق به ازای مقادیر مثبت λ چوله به راست به ازای مقادیر منفی چوله به چپ و به ازای $\lambda = ۰$ متقارن و به چگالی نرمال استاندارد تبدیل می‌شود. شکل (۲-۲-۱) تابع چگالی نرمال–چوله را به ازای چند مقدار از پارامتر آن نشان می‌دهد.



شکل (۲-۲-۱): توزیع نرمال – چوله در حالت SN(۱) , SN(۰), SN(-۱)

۲-۳ ویژگی‌هایی توزیع نرمال – چوله

در بخش قبل توزیع نرمال – چوله معرفی شد. در این بخش به بررسی بعضی از ویژگی‌های توزیعی نرمال – چوله شامل تابع چگالی، تابع توزیع، تابع مولد گشتاور و تابع مشخصه پرداخته می‌شود.

۲-۳-۱: چند خاصیت تابع چگالی نرمال – چوله

فرض کنید متغیر تصادفی Z دارای توزیع نرمال – چوله با پارامتر λ باشد در این صورت

۱) $-Z$ دارای چگالی نرمال – چوله با پارامتر $-\lambda$ است

$$f_{-Z}(z) = F_Z(-z) = ۲\phi(-z)\Phi(\lambda(-z)) = ۲\phi(z)\Phi((-\lambda)z) \qquad \square$$

۲) اگر $\lambda = ۰$ باشد آنگاه $Z \sim N(۰,۱)$ است.

۳) اگر $\lambda = ۱$ باشد آنگاه توزیع Z برابر توزیع آماره ماکسیم در یک نمونه تصادفی دوتایی از توزیع نرمال استاندارد است. اگر $\lambda = -۱$ باشد توزیع Z برابر توزیع آماره مینیمم در یک نمونه تصادفی دوتایی از توزیع نرمال استاندارد است.

۴) اگر $\lambda \to +\infty$ آنگاه توزیع Z به توزیع نیم نرمال میل می‌کند یعنی

$$\lim_{\lambda\to+\infty} ۲\phi(z)\Phi(\lambda z) = ۲\phi(z) \lim_{\lambda\to+\infty} \int_{-\infty}^{\lambda z} \phi(t)dt = \sqrt{\frac{۲}{\pi}} e^{-z^۲/۲} I_{(-\infty,۰)}(z)$$

این خاصیت یک نقطه ضعف برای توزیع نرمال – چوله محسوب می‌شود. زیرا وقتی λ به سمت $+\infty$ میل می‌کند، مقدار احتمالی که توزیع : نرمال – چوله به اعداد منفی نسبت می‌دهد به مقدار صفر کاهش می‌یابد.

به‌طور مشابه می‌توان نشان داد که اگر $\lambda \to -\infty$ آنگاه توزیع Z به توزیع نیم نرمال $\sqrt{\frac{۲}{\pi}} e^{-z^۲/۲} I_{(-\infty,۰)}(z)$ میل می‌کند.

۵) تابع توزیع نرمال – چوله به ازای یک λ ثابت قویا تک مدی است. (رامش گوپتا و نیکول براون[۱] ۲۰۰۱).

برای اینکه نشان دهیم تابع توزیع $f_z(z,\lambda)$ قویا تک مدی است کافی است نشان دهیم که $\frac{\partial^۲}{\partial z^۲} \ln f_z(z,\lambda)$ برای هر z تابعی منفی است یعنی $\ln f_z(z,\lambda)$ تابعی مقعر از z است.

---

[۱]. Ramesh Gupta and Nikol Brown

$$\frac{\partial^2}{\partial z^2} Ln\phi_Z(z,\lambda) = \frac{\partial^2}{\partial z^2}\{Ln2 + Ln\phi(z) + Ln\Phi(\lambda z)\}$$
$$= \frac{\partial^2}{\partial z^2} Ln\phi(z) + \frac{\partial^2}{\partial z^2} Ln\Phi(\lambda z)$$
$$= \frac{\partial}{\partial z}\left\{\frac{-z\phi(z)}{\phi(z)}\right\} + \frac{\partial}{\partial z}\left\{\frac{\lambda\phi(\lambda z)}{\Phi(\lambda z)}\right\}$$
$$= -1 + \lambda\left\{\frac{-z\lambda^2\phi(\lambda z)\Phi(\lambda z)}{\Phi^2(\lambda z)} - \frac{\lambda\Phi^2(\lambda z)}{\Phi^2(\lambda z)}\right\}$$
$$= -1 + \frac{-\lambda^2\phi(\lambda z)}{\Phi(\lambda z)}\left\{\frac{\phi(\lambda z)}{\Phi(\lambda z)} + \lambda z\right\}$$

برای اینکه نشان دهیم عبارت فوق منفی است کافی‌است نشان دهیم که عبارت $\frac{\phi(\lambda z)}{\Phi(\lambda z)} + \lambda z$ مثبت است ، به ازای هر z ، $\phi(\lambda z)$و $\Phi(\lambda z)$ مثبت‌اند.

الف)اگر $\lambda z \geq 0$ باشد آنگاه $\left\{\frac{\phi(\lambda z)}{\Phi(\lambda z)} + \lambda z\right\} \geq 0$ است.

ب)اگر $\lambda z < 0$ باشد تغییر متغیر $t = -\lambda Z$ را انجام می‌دهیم، بنابراین

$$\phi(\lambda z) = \phi(-\lambda z) = \phi(t)$$

$$\Phi(\lambda z) = 1 - \Phi(-\lambda z) = 1 - \Phi(t)$$

$$\frac{\phi(\lambda z)}{\Phi(\lambda z)} + \lambda z = \frac{\phi(t)}{1-\Phi(t)} t = h(t) - t$$

که در آن $h(t)$ تابع نرخ شکست در توزیع نرمال استاندارد است. می‌توان نشان داد که در توزیع

نرمال استاندارد به ازای هر t ، $h(t) > t$ است زیرا از نا مساوی میل[2] داریم

$$\frac{x}{x^2+1}\phi(x) \leq 1 - \Phi(x) \leq \frac{1}{x}\phi(x)$$

$$\Rightarrow x \leq \frac{\phi(x)}{1-\Phi(x)}$$

بنابراین حکم ثابت می‌شود. □

---

[2]. Mills Ration