مجله علوم آماری، پاییز و زمستان ۱۳۸۹
جلد ۴، شماره ۲، ص ۱۶۷–۱۸۲

یکپارچه‌سازی ویژگی‌های گوناگون توزیعی با استفاده از توزیع‌های موزون

حمیدرضا چاره، افشین فلاح
گروه آمار، دانشگاه بین‌المللی امام خمینی

تاریخ دریافت: ۱۳۸۹/۳/۱۰ **تاریخ آخرین بازنگری:** ۱۳۸۹/۹/۲۱

چکیده : این مقاله در راستای ایده یکپارچه‌سازی مباحث مربوط به ساختن توزیع‌های چوله متقارن (چوله‌نرمال) و توزیع‌های دو (چند) نمایی، بر اساس توزیع‌های متقارن و تک مدی، توزیع‌های موزون را مورد توجه قرار داده است. بحث شده است که توزیع چوله نرمال و تعمیم‌های آن، که در سال‌های اخیر بخش قابل توجهی از تحقیقات را به خود اختصاص داده‌اند، را می‌توان از دیدگاه توزیع‌های موزون به‌صورت کلی‌تر و با ویژگی‌های دیگری مانند چند نمایی بودن به‌صورت یکپارچه مورد مطالعه قرار داد. به کمک دو مورد از تحقیقات شاخص اخیر در زمینه توزیع‌های چوله نرمال، نشان داده شده است که توزیع‌های مورد نظر این تحقیقات و خواص مطلوب آن‌ها را می‌توان به‌صورت کلی‌تر به همراه برخی ویژگی‌های مطلوب دیگر، از دیدگاه توزیع‌های موزون و تنها به‌عنوان حالات خاص به‌دست آورد. تاکید شده است که چولگی تنها یکی از ویژگی‌هایی است که می‌توان با انتخاب وزن مناسب به هر توزیع متقارن افزود.

آدرس الکترونیک مسئول مقاله: افشین فلاح، fallah@ikiu.ac.ir
کد موضوع‌بندی ریاضی (۲۰۱۰): ۶۲E۱۵

واژه‌های کلیدی : چولگی، توزیع چوله نرمال، توزیع موزون، توزیع دونمایی.

۱ مقدمه

توزیع نرمال به دلیل ویژگی‌های خاص و گاه منحصر به فرد خود از جایگاه ویژه‌ای در آمار برخوردار است، تا حدی که بخش قابل توجهی از مباحث مدل‌سازی آماری به نوعی وابسته به توزیع نرمال است. با این وجود، موارد بسیاری نیز وجود دارند که این فرضیه به سادگی قابل پذیرش یا دستیابی نیست. این موارد امروزه با گسترش نیاز به کاربست روش‌های مختلف آماری خصوصا روش‌های مدل‌سازی در سایر شاخه‌های علوم، بیشتر از گذشته مشاهده می‌شود. در چنین شرایطی نظریه‌ها و روش‌های آماری مبتنی بر فرض نرمال بودن به‌صورت معمول قابل کاربرد نیستند. تبدیل داده‌ها و تغییر در روش مدل‌سازی به گونه‌ای که به فرض نرمال بودن وابسته نباشد یا دست کم وابستگی کمی داشته باشد، از جمله ایده‌هایی بوده که در این زمینه مطرح و منجر به شکل‌گیری روش‌های مختلف ناپارامتری و استوار[۱] شده است. گرچه هر یک از این روش‌ها ممکن است در کاربردهای خاصی راهگشا و قابل استفاده باشند، اما به دلیل ویژگی‌های خاص توزیع نرمال و گستردگی ادبیات و تحقیقات انجام شده پیرامون آن، ایده تعمیم این خانواده از توزیع‌ها و افزودن ویژگی‌هایی مانند قابلیت مدلبندی عدم تقارن، چند نمایی بودن و غیره، خصوصاً در بیش از دو دهه گذشته همواره مورد علاقه بوده است. به گونه‌ای که پس از ارائه تعمیمی از توزیع نرمال با قابلیت مدل‌بندی چولگی توسط آزالینی (۱۹۸۵)، ایده وی به سرعت مورد توجه محققان بسیاری قرار گرفت. آزالینی و واله (۱۹۹۶) شکل چند متغیره این توزیع را ارائه کردند. لیزئو و لوپرفیدو (۲۰۰۳) استنباط بیزی بر اساس توزیع چوله‌نرمال را مورد مطالعه قرار دادند. تعمیم‌های متعددی نیز بر پایه همین ایده توسط نویسندگان مختلفی مانند گنزالس و همکاران (۲۰۰۴)، آرلانو ـ واله و گنتون (۲۰۰۵)، یادگاری و همکاران (۲۰۰۷) و بهرامی و همکاران (۱۳۸۸) ارائه شده است. به‌علاوه خردمندی و سنجری (۱۳۸۸)

---
۱ Robust

و خردمندی و همکاران (۲۰۱۰) نیز توزیع‌های چوله t-نرمال را مطرح و تعمیم‌ها و نحوه کاربست آن‌ها در تحلیل داده‌های واقعی را بیان کردند. در این راستا و به دلیل تعدد تلاش‌هایی که در این خصوص صورت پذیرفته است، در سال‌های اخیر برخی محققان کوشیده‌اند مباحث مرتبط با توزیع چوله‌نرمال و تعمیم‌های آن را یکپارچه و در قالب یک توزیع جامع ارائه نمایند. از جمله این تلاش‌ها می‌توان به آرلانو ـ واله و آزالینی (۲۰۰۶) اشاره کرد، که در آن فرم یکپارچه‌ای از توزیع چوله‌نرمال و تعمیم‌های آن که بسیاری از توزیع‌های چوله‌نرمال را به عنوان حالت خاص در بر می‌گیرد، ارائه شده است.

از دیگر ایده‌های مطرح در این زمینه، استفاده از توزیع‌های موزون[۲] بوده است (فیشر، ۱۹۳۴). این توزیع‌ها با لحاظ کردن یک تابع وزنی، انعطاف پذیری بیشتری را به توزیع اصلی می‌بخشند، به‌گونه‌ای که با استفاده از توابع وزن متفاوت می‌توان برخی ویژگی‌های توزیع اصلی را تغییر داد. برخی از توابع وزن مهم به‌همراه ویژگی‌های شاخص توزیع‌های موزون توسط علوی و چینی‌پرداز (۱۳۸۶) و زادکرمی (۱۳۸۷) مورد بررسی قرار گرفته‌اند. این مقاله در راستای ایده یکپارچه‌سازی مباحث مربوط به ساختن توزیع‌های چوله متقارن (چوله‌نرمال) و توزیع‌های دو (چند) نمایی بر اساس توزیع‌های متقارن و تک مدی، توزیع‌های موزون را مورد توجه قرار داده است. نشان داده شده است که توزیع چوله‌نرمال و تعمیم‌های آن، که در سال‌های اخیر بخش قابل توجهی از ادبیات آماری را به خود اختصاص داده‌اند، را می‌توان از دیدگاه توزیع‌های موزون به‌صورت کلی‌تر و به ویژه با ویژگی‌های دیگری مانند چند نمایی بودن به‌صورت یکپارچه مورد مطالعه قرار داد.

در بخش دوم، به اجمال توزیع چوله‌نرمال و فرم یکپارچه آرلانو ـ واله و آزالینی (۲۰۰۶) برای این توزیع به همراه برخی ویژگی‌های آن شرح داده شده است. در بخش سوم، مفهوم توزیع‌های موزون به اختصار شرح داده شده است. سپس، در راستای ایده مطرح شده در مقدمه، دو مورد از تحقیقات شاخص سال‌های اخیر درباره توزیع چوله‌نرمال و تعمیم‌های آن، از دیدگاه توزیع‌های وزنی بررسی و نشان

---

[۲] Weighed Distributions

داده شده است، که توزیع‌های مورد بحث در این تحقیقات و ویژگی‌های آنها را می‌توان به عنوان حالات خاصی از دو توزیع وزنی به‌دست آورد.

۲ توزیع چوله‌نرمال و تعمیم‌های آن

متغیر تصادفی X دارای توزیع چوله‌نرمال استاندارد است ($Z \sim SN(\lambda)$)، هرگاه تابع چگالی آن به‌صورت $\phi(z;\lambda) = 2\phi(x)\Phi(\lambda x),\ \ x,\lambda \in \mathbb{R}$ باشد، که در آن $\phi(\cdot)$ و $\Phi(\cdot)$ به‌ترتیب تابع چگالی و تابع توزیع تجمعی توزیع نرمال استاندارد هستند. این توزیع به ازای $\lambda > 0$ ($\lambda < 0$) چوله به راست (چپ) است و به ازای $\lambda = 0$ به توزیع نرمال استاندارد تبدیل می‌شود (آزالینی، ۱۹۸۵). آرلانو ـ واله و آزالینی (۲۰۰۶) فرم یکپارچه‌ای از توزیع چوله‌نرمال را معرفی کردند که بسیاری از توزیع‌های چوله‌نرمال معرفی شده در ادبیات موضوع را به عنوان حالت خاص در بر می‌گیرد. اهمیت اصلی این توزیع یکپارچه، انعطاف پذیری مطلوب آن در برازش به مشاهدات گوناگون و فراهم آوردن شرایط یکپارچه نمودن مباحث مدل‌سازی مرتبط با توزیع چوله‌نرمال و تعمیم‌های آن است. فرض کنید $\mathbf{U}$ و $\mathbf{V}$ به‌ترتیب نشان دهنده دو بردار تصادفی m و n بعدی باشند، که از توزیع نرمال چند متغیره بلوکی به‌صورت

$$\begin{pmatrix} \boldsymbol{U} \\ \boldsymbol{V} \end{pmatrix} \sim N_{m+n}\left(\begin{pmatrix} \boldsymbol{\delta} \\ \boldsymbol{\xi} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Gamma} & \boldsymbol{\Delta}^T \\ \boldsymbol{\Delta} & \boldsymbol{\Omega} \end{pmatrix}\right), \tag{۱}$$

پیروی می‌کنند. در این صورت بردار تصادفی n ـ بعدی $\boldsymbol{X}$ که به‌صورت $\boldsymbol{X} \stackrel{d}{=} \boldsymbol{V}|\boldsymbol{U} > 0$ تعریف می‌شود، دارای توزیع چوله‌نرمال یکپارچه[۳] n متغیره با پارامتر $\boldsymbol{\theta} = (\boldsymbol{\xi}, \boldsymbol{\delta}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\Delta})$ است و با نماد $\boldsymbol{X} \sim SUN_{n,m}(\boldsymbol{\theta})$ نشان داده می‌شود. تابع چگالی و تابع مولد گشتاور این توزیع به‌ترتیب به‌صورت

$$\phi_{SUN_{n,m}}(\boldsymbol{x};\boldsymbol{\theta}) = \frac{\phi_n(\boldsymbol{x};\boldsymbol{\xi},\boldsymbol{\Omega})\Phi_m(\boldsymbol{\delta} + \boldsymbol{\Delta}^T\boldsymbol{\Omega}^{-1}(\boldsymbol{x}-\boldsymbol{\xi}), \boldsymbol{\Gamma} - \boldsymbol{\Delta}^T\boldsymbol{\Omega}^{-1}\boldsymbol{\Delta})}{\Phi_m(\boldsymbol{\delta};\boldsymbol{\Gamma})} \tag{۲}$$

$$M_{SUN_{n,m}}(\boldsymbol{s};\boldsymbol{\theta}) = \frac{\exp(\boldsymbol{\xi}^T\boldsymbol{s} + \frac{1}{2}\boldsymbol{s}^T\boldsymbol{\Omega}\boldsymbol{s})\Phi_m(\boldsymbol{\delta} + \boldsymbol{\Delta}^T\boldsymbol{s};\boldsymbol{\Gamma})}{\Phi_m(\boldsymbol{\delta};\boldsymbol{\Gamma})},$$

---

۳ Unified Skew-Normal Distribution

هستند، که در آن‌ها $\phi_n(.;\boldsymbol{\xi},\boldsymbol{\Omega})$ تابع چگالی توزیع نرمال چند متغیره $N_n(\boldsymbol{\xi},\boldsymbol{\Omega})$ و $\Phi_m(.;\boldsymbol{\Gamma})$ تابع توزیع تجمعی توزیع نرمال چند متغیره $N_m(۰,\boldsymbol{\Gamma})$ هستند. این توزیع به ازای $m=۱$، $n=۱$ و $\theta=(۰,۰,۱,۱,\lambda)$ به توزیع چوله‌نرمال آزالینی با پارامتر چولگی $\alpha=\frac{\lambda}{\sqrt{۱-\lambda^۲}}$ تبدیل می‌شود. به علاوه، با جایگزینی مقادیر $m=۱$ و $n=۱$ در (۲) توزیع چوله‌نرمال یکپارچه یک متغیره به‌صورت

$$\phi_{SN}(x;\theta)=\frac{۱}{\Phi(\frac{\delta}{\gamma})}\phi(x;\xi,\omega)\Phi(\frac{\delta+\frac{\lambda}{\omega}(x-\xi)}{\sqrt{\gamma^۲-\frac{\lambda^۲}{\omega}}})\quad -\infty<x<+\infty,$$

حاصل می‌شود، که در آن $\phi(.;\xi,\omega)$ تابع چگالی توزیع نرمال $N(\xi,\omega)$، $\Phi(.)$ تابع توزیع تجمعی نرمال استاندارد و $\theta=(\xi,\delta,\omega,\gamma^۲,\lambda)$ بردار پارامترهای توزیع است. تابع مولد گشتاورها و دو گشتاور اول این توزیع به‌ترتیب به‌صورت

$$\begin{aligned}
M_{SN}(t;\theta) &= \frac{۱}{\Phi(\frac{\delta}{\gamma})}\exp(\xi t+\frac{۱}{۲}wt^۲)\Phi(\frac{\delta+\lambda t}{\gamma})\\
E(X_\theta) &= \xi+\frac{\lambda}{\gamma}r\left(\frac{\delta}{\gamma}\right)\\
E(X_\theta^۲) &= \xi^۲+\omega+\frac{\lambda}{\gamma}r\left(\frac{\delta}{\gamma}\right)\{\xi-\frac{\lambda\delta}{\gamma^۲}\}
\end{aligned}$$

هستند، که در آن $r(t)=\frac{\phi(t)}{\Phi(t)}$ وارون تابع نرخ مخاطره توزیع نرمال استاندارد است. یکی از ویژگی‌های مهم و جالب توزیع چوله‌نرمال یکپارچه بسته بودن آن تحت حاشیه‌ای سازی و شرطی ساختن است. یعنی تمام توزیع‌های کناری و شرطی این توزیع نیز چوله‌نرمال یکپارچه هستند (آرلانو ـ واله و آزالینی، ۲۰۰۶). فرض کنید بردار تصادفی $\boldsymbol{X}\sim SUN_{n,m}(\boldsymbol{\theta})$ به‌صورت $(\boldsymbol{X}_۱,\boldsymbol{X}_۲)^t$ که در آن $\boldsymbol{X}_۱$ و $\boldsymbol{X}_۲$ به‌ترتیب بردارهای $n_۱$ و $n_۲=n-n_۱$ بعدی هستند، افراز شده است. اگر بردار $\boldsymbol{\xi}$ و ماتریس‌های $\boldsymbol{\Delta}$ و $\boldsymbol{\Omega}$ به‌ترتیب به‌صورت $(\boldsymbol{\xi}_۱,\boldsymbol{\xi}_۲)^t$، $(\boldsymbol{\Delta}_۱,\boldsymbol{\Delta}_۲)^t$ و

$$\begin{pmatrix}\boldsymbol{\Omega}_{۱۱} & \boldsymbol{\Omega}_{۱۲}\\ \boldsymbol{\Omega}_{۲۱} & \boldsymbol{\Omega}_{۲۲}\end{pmatrix},$$

افراز شوند، در این صورت X_i برای $i=۱,۲$ دارای توزیع چوله‌نرمال یکپارچه n_i–بعدی با بردار پارامترهای $\boldsymbol{\theta}_i=(\boldsymbol{\xi}_i,\boldsymbol{\delta},\boldsymbol{\Omega}_{ii},\boldsymbol{\Gamma},\boldsymbol{\Delta}_i)^t$ است. همچنین $X_۱|X_۲$

دارای توزیع $SUN_{n_1}(\boldsymbol{\xi}_{1.2}, \boldsymbol{\delta}_{1.2}, \boldsymbol{\Omega}_{11.2}, \boldsymbol{\Gamma}_{1.2}, \boldsymbol{\Delta}_{1.2})$ است، که در آن

$$\begin{aligned}
\boldsymbol{\xi}_{1.2} &= \boldsymbol{\xi}_1 + \boldsymbol{\Omega}_{12}^T \boldsymbol{\Omega}_{22}^{-1}(X_2 - \boldsymbol{\xi}_2) \quad , \quad \boldsymbol{\delta}_{1.2} = \boldsymbol{\delta} + \boldsymbol{\Delta}_2 \boldsymbol{\Omega}_{22}^{-1}(X_2 - \boldsymbol{\xi}_2) \\
\boldsymbol{\Omega}_{11.2} &= \boldsymbol{\Omega}_{11} - \boldsymbol{\Omega}_{12}^t \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\Omega}_{21} \quad , \quad \boldsymbol{\Gamma}_{1.2} = \boldsymbol{\Gamma} - \boldsymbol{\Delta}_2^t \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\Delta}_2 \\
\boldsymbol{\Delta}_{1.2} &= \boldsymbol{\Delta}_1 - \boldsymbol{\Omega}_{12}^t \boldsymbol{\Omega}_{22}^{-1} \boldsymbol{\Delta}_2.
\end{aligned}$$

برای جزئیات بیشتر در مورد توزیع چوله‌نرمال یکپارچه و ارتباط آن با سایر توزیع‌ها به آزالینی و واله (۲۰۰۶) مراجعه شود.

۳ دیدگاه توزیع‌های موزون

این بخش، به عنوان نمونه، دو مورد از تحقیقات سال‌های اخیر در زمینه توزیع چوله‌نرمال و تعمیم‌های آن را از دیدگاه توزیع‌های موزون مورد بررسی قرار می‌دهد. در نمونه اول، با در نظر گرفتن یک تابع وزن چند متغیره، شکل تعمیم یافته‌ای از توزیع نرمال وزنی معرفی و نشان داده شده است که توزیع چوله‌نرمال یکپارچه آرلانو ـ واله و آزالینی (۲۰۰۶) که خود خانواده گسترده‌ای از توزیع چوله‌نرمال و تعمیم‌های آن را در بر می‌گیرد، حالتی خاص از آن است. به عنوان نمونه دوم، توزیع کیم (۲۰۰۵) مورد توجه قرار گرفته است. این توزیع از این جهت قابل توجه است که دو ویژگی چولگی و دو نمایی بودن را توأما مد نظر قرار داده است. نشان داده شده است که از دیدگاه توزیع‌های موزون به سادگی می‌توان توزیع‌هایی مطلوب‌تر از این توزیع را با انتخاب وزن مناسب ارائه کرد، به گونه‌ای که این توزیع را به عنوان حالت خاص در برگرفته و افزون بر آن قابلیت‌های مدل‌بندی مطلوب‌تری را نیز دارا باشند. ایده اصلی توزیع‌های وزنی اولین بار توسط فیشر (۱۹۳۴) مطرح شده است. اگر متغیر تصادفی X دارای تابع چگالی احتمال $f(x;\theta)$ و $w(x;\alpha) \geq 0$ یک تابع وزن باشد، آنگاه

$$f_{X_w}(x;\theta,\alpha) = \frac{w(x;\alpha)}{E_X[w(X;\alpha)]} f(x;\theta), \tag{۳}$$

تابع چگالی احتمال وزن‌دار شده X است. توزیع‌های موزون که با در نظر گرفتن وزن‌های مختلف در (۳) به‌دست می‌آیند، از انعطاف پذیری مطلوبی در برازش به

مشاهدات مختلف برخوردارند. تعداد و تنوع توابع وزنی که می‌توان مد نظر قرار داد، بسیار زیاد است و محقق می‌تواند در هر شرایطی با توجه به ماهیت مشاهدات وزن مناسب را به گونه‌ای انتخاب نماید تا بهترین برازش ممکن حاصل شود. به عنوان مثال، اگر تابع وزن به‌صورت $w(x) = I_{(a,b)}(x)$ در نظر گرفته شود، که در آن a و b دو عدد ثابت و $I_A(\cdot)$ تابع نشانگر مجموعه A را نشان می‌دهد، توزیعی از دو سو بریده[۴] به‌دست می‌آید. به همین ترتیب اگر تابع وزن به‌صورت $w(x,\lambda) = \Phi(\lambda x)$ انتخاب شود، که در آن $\Phi(\cdot)$ تابع توزیع تجمعی نرمال استاندارد است، خانواده توزیع‌های چوله ـ متقارن با شکل کلی $f_{X_\lambda}(x) = ۲f_X(x)\Phi(\lambda x)$ حاصل می‌شود، که در آن $f_X(.)$ یک توزیع متقارن دلخواه می‌باشد. اگر توزیع $f_X(\cdot)$ نرمال استاندارد انتخاب شود، توزیع چوله‌نرمال آزالینی (۱۹۸۵) به‌دست می‌آید. در حالت چند متغیره، با فرض $w(\boldsymbol{x},\boldsymbol{\theta}) = \Phi_m(\boldsymbol{\delta} + \boldsymbol{\Delta}^t\boldsymbol{\Omega}^{-۱}(\boldsymbol{x}-\boldsymbol{\xi}), \boldsymbol{\Gamma} - \boldsymbol{\Delta}^t\boldsymbol{\Omega}^{-۱}\boldsymbol{\Delta})$ که در آن $\boldsymbol{\theta} = (\boldsymbol{\xi}, \boldsymbol{\delta}, \boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{\Delta})$ بردار پارامترها است، توزیع چوله‌نرمال یکپارچه (۲) به‌دست می‌آید.

اکنون فرض کنید $\boldsymbol{X} \sim SUN_{n,m}(\boldsymbol{\theta})$، $Var(\boldsymbol{X}) = \boldsymbol{\Pi}$ و $E(\boldsymbol{X}) = \boldsymbol{\Lambda}$. در این صورت، اگر تابع وزن به‌صورت $w(\boldsymbol{x};\boldsymbol{\Lambda}) = (\boldsymbol{x}-\boldsymbol{\Lambda})^{\mathbf{t}}(\boldsymbol{x}-\boldsymbol{\Lambda})$ در نظر گرفته شود، آنگاه تابع چگالی احتمال توزیع نرمال موزون تعمیم یافته‌ای به‌صورت

$$\begin{aligned}\phi_{WSUN_{n,m}}(\boldsymbol{x};\boldsymbol{\theta}) &= (\boldsymbol{x}-\boldsymbol{\Lambda})^T\boldsymbol{\Pi}^{-۱}(\boldsymbol{x}-\boldsymbol{\Lambda}) \\ &\quad\times\frac{\phi_n(\boldsymbol{x};\boldsymbol{\xi},\boldsymbol{\Omega})\Phi_m(\boldsymbol{\delta}+\boldsymbol{\Delta}^T\boldsymbol{\Omega}^{-۱}(\boldsymbol{x}-\boldsymbol{\xi});\boldsymbol{\Gamma}-\boldsymbol{\Delta}^t\boldsymbol{\Omega}^{-۱}\boldsymbol{\Delta})}{\Phi_m(\boldsymbol{\delta};\boldsymbol{\Gamma})}, \quad (۴)\end{aligned}$$

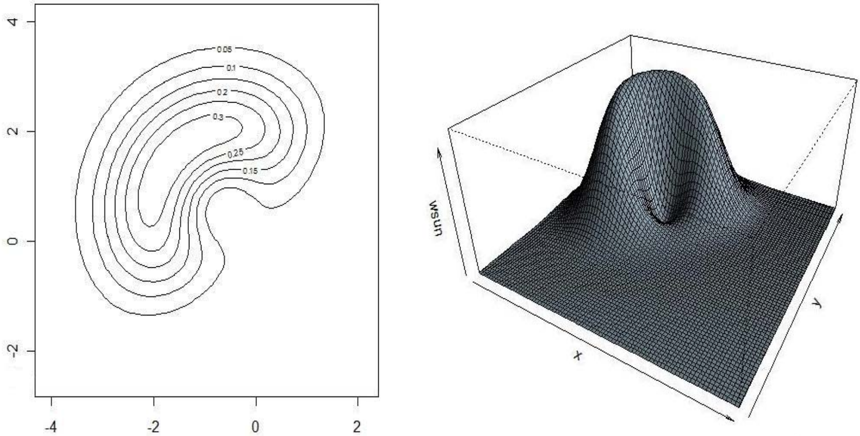
به‌دست می‌آید، که با $\boldsymbol{X} \sim WSUN_{n,m}(\boldsymbol{\theta})$ نشان داده می‌شود. به سادگی داریم

$$\begin{aligned}\int \phi_{WSUN_{n,m}}(\boldsymbol{x};\boldsymbol{\theta})d\boldsymbol{x} &= E[(\boldsymbol{X}-\boldsymbol{\Lambda})^t\boldsymbol{\Pi}^{-۱}(\boldsymbol{X}-\boldsymbol{\Lambda})] \\ &= E[(\boldsymbol{X}-\boldsymbol{\Lambda})^t(\boldsymbol{X}-\boldsymbol{\Lambda})]\boldsymbol{\Pi}^{-۱} \\ &= \boldsymbol{\Pi}\boldsymbol{\Pi}^{-۱} = ۱.\end{aligned}$$

گشتاورهای این توزیع با استفاده از لم زیر که اثبات ساده‌ای دارد، قابل محاسبه‌اند.

---

۴ Truncated



شکل ۱: توزیع چوله‌نرمال یکپارچه موزون به ازای مقادیر $\Lambda = (-1, 1)'$

لم ۱ : فرض کنید $\boldsymbol{X} \sim SUN(\boldsymbol{\theta})$ و $\boldsymbol{X}_w \sim WSUN_{m,n}(\boldsymbol{\theta})$، در این صورت

$$E(\boldsymbol{X}_w) = E(\boldsymbol{X}[(\boldsymbol{X} - \boldsymbol{\Lambda})^t \boldsymbol{\Pi}^{-1} (\boldsymbol{X} - \boldsymbol{\Lambda})]).$$

شکل ۱ نمودار تابع چگالی (۴) را به ازای $\boldsymbol{\Lambda} = (-1, 1)'$ نشان می‌دهد. به‌طور خاص، به ازای $m = 1$ و $n = 1$، توزیع نرمال موزون تعمیم‌یافته یک متغیره با تابع چگالی احتمال

$$\begin{aligned} f_{X_w}(x;\theta) &= \frac{\left[x - (\xi + \frac{\lambda}{\gamma} r\left(\frac{\delta}{\gamma}\right))\right]^2}{\left[\omega - \frac{\lambda^2}{\gamma^2} r\left(\frac{\delta}{\gamma}\right)\left\{\frac{\delta}{\gamma} + r\left(\frac{\delta}{\gamma}\right)\right\}\right]\Phi(\frac{\delta}{\gamma})} \phi(x;\xi,\omega) \\ &\times \Phi(\frac{\delta + \frac{\lambda}{\omega}(x-\xi)}{\sqrt{\gamma^2 - \frac{\lambda^2}{\omega}}}), \qquad -\infty < x < +\infty \end{aligned}$$

حاصل می‌شود، که در آن $r(t) = \frac{\phi(t)}{\Phi(t)}, t \in \mathbb{R}$ وارون تابع نرخ مخاطره توزیع نرمال استاندارد یک متغیره است. گشتاورهای این توزیع نیز به راحتی با استفاده از لم ۱ قابل محاسبه هستند، به عنوان مثال گشتاور اول آن عبارت است از

$$E(X_w) = \frac{1}{\left[\omega - \frac{\lambda^2}{\gamma^2} r\left(\frac{\delta}{\gamma}\right)\left\{\frac{\delta}{\gamma} + r\left(\frac{\delta}{\gamma}\right)\right\}\right]}$$

$$\begin{aligned}\times & \left[\xi\{\xi^۲ + \xi\omega + ۲\omega\} + \frac{\lambda}{\gamma} r\left(\frac{\delta}{\gamma}\right)\left\{\xi^۲ + ۲\omega - \frac{\lambda^۲}{\gamma^۲} - \frac{\delta}{\gamma}\right\}\right. \\ - & ۲\left(\xi + \frac{\lambda}{\gamma} r\left(\frac{\delta}{\gamma}\right)\right)\left(\xi^۲ + \omega + \frac{\lambda}{\gamma} r\left(\frac{\delta}{\gamma}\right)\left\{\xi - \frac{\lambda\delta}{\gamma^۲}\right\}\right) \\ & \left. + \left(\xi + \frac{\lambda}{\gamma} r\left(\frac{\delta}{\gamma}\right)\right)^۳\right].\end{aligned}$$

همچنین به ازای $n = ۱$، $m = ۱$ و $\theta = (۰, ۰, ۱, ۱, \lambda)$، توزیع چوله‌نرمال وزنی استاندارد با تابع چگالی

$$f_X(x;\lambda) = \frac{(x - \sqrt{\frac{۲}{\pi}}\lambda)^۲}{۱ - \frac{۲}{\pi}\lambda^۲} ۲\phi(x)\Phi(\frac{\lambda x}{\sqrt{۱ - \lambda^۲}}), \quad -\infty < x < +\infty$$

به‌دست می‌آید، که دو گشتاور اول آن عبارتند از

$$\begin{aligned} E(X) &= \frac{(\frac{۲}{\pi} - ۱)\sqrt{\frac{۲}{\pi}}\lambda^۳}{۱ - \frac{۲}{\pi}\lambda^۲} \\ E(X^۲) &= \frac{\lambda^۳(\frac{۴}{\pi} - ۱) + \lambda^۲(\frac{۲}{\pi} - \sqrt{\frac{۲}{\pi}}) + \lambda(۲\sqrt{\frac{۲}{\pi}} + ۱ - \frac{۸}{\pi})}{۱ - \frac{۲}{\pi}\lambda^۲}. \end{aligned}$$

از دیدگاه توزیع‌های موزون، توزیع $WSUN$ توزیعی موزون با یک تابع وزن دو تکه ضربی به‌صورت

$$w(\boldsymbol{x};\Lambda) = (\boldsymbol{x} - \Lambda)^t(\boldsymbol{x} - \Lambda)\Phi_m(\boldsymbol{\delta} + \boldsymbol{\Delta}^t\boldsymbol{\Omega}^{-۱}(\boldsymbol{x} - \boldsymbol{\xi}), \boldsymbol{\Gamma} - \boldsymbol{\Delta}^t\boldsymbol{\Omega}^{-۱}\boldsymbol{\Delta}),$$

است. مولفه‌های اول و دوم (از سمت راست) در این تابع وزنی به‌ترتیب قابلیت مدل‌بندی چولگی و دو مدی بودن را به توزیع نرمال چند متغیره القا می‌کنند. از طرف دیگر این توزیع به وضوح شکل موزون توزیع چوله‌نرمال یکپارچه آزالینی و واله (۲۰۰۶) است و از این‌رو، آن را به‌عنوان حالت خاص در برمی‌گیرد. چون سایر ویژگی‌های این توزیع را می‌توان از طریق رابطه بین توزیع‌های SUN و $WSUN$ به‌دست آورد، از توضیح بیشتر در این رابطه خودداری می‌شود. آرلانو ـ واله و آزالینی (۲۰۰۶) را ببینید.

یکی دیگر از ویژگی‌هایی که در بعضی کاربردها استفاده از توزیع نرمال را دشوار می‌سازد، وضعیتی است که در آن مشاهدات علاوه بر عدم تقارن، ساختاری با

دو نما را منعکس می‌سازند. کیم (۲۰۰۵) با مبنا قرار دادن توزیع نرمال و به روشی مشابه روش آزالینی (۱۹۸۵) توزیعی دو نمایی به‌صورت

$$f_X = c(\lambda)\phi(x)\Phi(\lambda|x|), \quad -\infty < x, \lambda < \infty,$$

را ارائه کرد، که در آن $c(\lambda)$ ثابت چگالی ساز است. این توزیع تنها برای مدل‌بندی مشاهداتی بکار می‌آید که علاوه بر ساختار دو نمایی و نامتقارن، ارتفاع یکسانی را در دو نما نمایش دهند. از دیدگاه توزیع‌های موزون می‌توان با انتخاب تابع وزن مناسب برای توزیع نرمال، توزیعی ساخت که همه ویژگی‌های توزیع کیم (۲۰۰۵) را داشته و نقطه ضعف این توزیع (ارتفاع لزوماً یکسان در دو نما) را نیز نداشته باشد.

قضیه ۱ : تابع وزن $w(x;\lambda_1,\lambda_2) = \Phi(\lambda_1 x + \lambda_2|x|)$ را برای توزیع نرمال در نظر بگیرید. در این صورت تابع چگالی توزیع نرمال موزون به‌صورت

$$\phi^*(x;\lambda_1,\lambda_2) = c(\lambda_1,\lambda_2)\phi(x)\Phi(\lambda_1 x + \lambda_2|x|), \qquad (5)$$

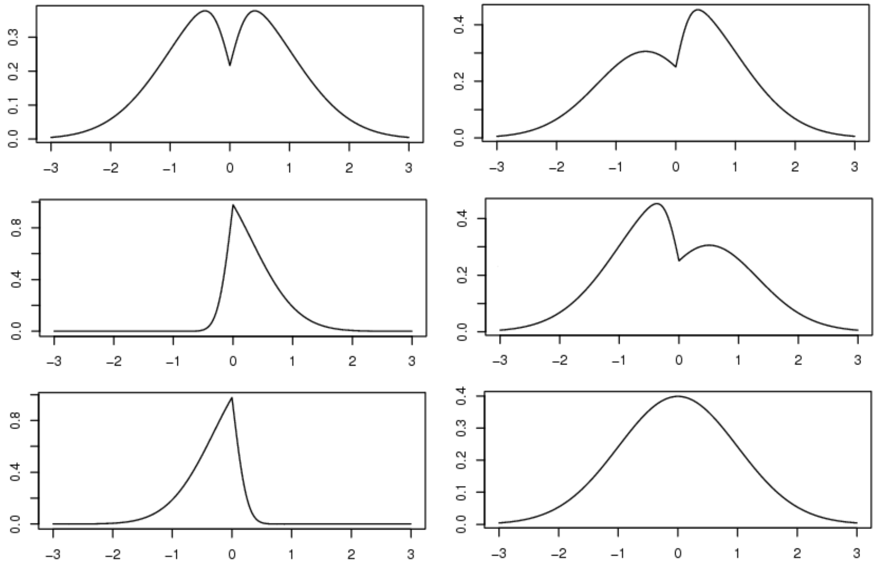
است، که در آن

$$c(\lambda_1,\lambda_2) = \frac{1}{\frac{1}{2} + \frac{1}{\pi}\tan^{-1}(\frac{\lambda_2}{\sqrt{1+\lambda_1^2}})}.$$

برهان : بر اساس تعریف توزیع‌های وزنی به سادگی می‌توان نوشت

$$\begin{aligned} \frac{1}{c(\lambda_1,\lambda_2)} &= \int_{-\infty}^{+\infty} \phi^*(x;\lambda_1,\lambda_2)dx \\ &= E[\Phi(\lambda_1 X + \lambda_2|X|)] \\ &= E[P(Y \leq \lambda_1 X + \lambda_2|X|)] \\ &= P(\frac{Y-\lambda_1 X}{|X|} \leq \lambda_2), \\ &= \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(\frac{\lambda_2}{\sqrt{1+\lambda_1^2}}). \end{aligned}$$

در روابط بالا برابری آخر به این دلیل است که متغیر تصادفی $Y - \lambda_1 X$ دارای توزیع نرمال $N(0, 1+\lambda_1^2)$ می‌باشد و به همین دلیل $\frac{Y-\lambda_1 X}{|X|}$ دارای توزیع کوشی



شکل ۲: توزیع چوله‌نرمال دو تکه برای مقادیر مختلف λ_1 و λ_2

$C(0, 1+\lambda_1^2)$ است. شکل ۲ تابع چگالی این توزیع را برای مقادیر مختلف λ_2 و λ_1 نشان می‌دهد. ملاحظه می‌شود که توزیع حاصل ساختاری دو نمایی دارد، در اطراف هر نما لزوماً متقارن نیست و قابلیت مدل‌بندی چولگی را نیز داراست. همچنین به ازای $\lambda_1 = 0$ و $\lambda_2 = 0$ به‌ترتیب به توزیع‌های چوله‌نرمال کیم (۲۰۰۵) و چوله نرمال آزالینی (۱۹۸۵) تبدیل می‌شود. از این رو به دلیل ویژگی‌های یاد شده در مقایسه با همتاهای خود به وضوح قابلیت برازش به مشاهدات متنوع‌تری را داراست. در ادامه برخی از ویژگی‌های توزیع (۵) که آن را توزیع چوله‌نرمال دو تکه[۵] نامیده و با نماد $TPSN(\lambda_1, \lambda_2)$ نمایش می‌دهیم، مورد بررسی قرار گرفته است.

قضیه ۲ : توزیع $TPSN(\lambda_1, \lambda_2)$ آمیخته‌ای از دو توزیع چوله‌نرمال بریده با پارامترهای $\lambda_2 + \lambda_1$ و $\lambda_2 - \lambda_1$ به‌ترتیب برای مقادیر مثبت و منفی متغیر تصادفی X است.

۵ Two Part Skew-Normal

برهان : فرض کنید

$$\begin{aligned}\phi_{X^+}(x;\lambda_1,\lambda_2) &= \frac{2\phi(x)\Phi((\lambda_1+\lambda_2)x)}{\frac{1}{2}+\frac{1}{\pi}\tan^{-1}(\frac{\lambda_2}{\sqrt{1+\lambda_1^2}})}I_{(0,+\infty)}(x)\\ \phi_{X^-}(x;\lambda_1,\lambda_2) &= \frac{2\phi(x)\Phi((\lambda_1-\lambda_2)x)}{\frac{1}{2}+\frac{1}{\pi}\tan^{-1}(\frac{\lambda_2}{\sqrt{1+\lambda_1^2}})}I_{(-\infty,0)}(x)\end{aligned}$$

در این صورت، اگر متغیر تصادفی X^* به‌صورت

$$X^* = \begin{cases} X^+ & \text{با احتمال } \frac{1}{2} \\ X^- & \text{با احتمال } \frac{1}{2} \end{cases} \qquad (6)$$

تعریف شود، به سادگی رابطه زیر را خوهیم داشت و حکم ثابت است،

$$\phi_{X^*}(x;\lambda_1,\lambda_2) = \frac{1}{2}\phi_{X^+}(x;\lambda_1,\lambda_2)I_{(0,+\infty)}(x) + \frac{1}{2}\phi_{X^-}(x;\lambda_1,\lambda_2)I_{(-\infty,0)}(x).$$

قضیه ۳ : تابع مولد گشتاورهای توزیع $TPSN(\lambda_1,\lambda_2)$ را می‌توان به‌صورت

$$\begin{aligned}M_X(t) &= \exp(\frac{t^2}{2})c(\lambda_1,\lambda_2)\\ &\quad [P(Y\le(\lambda_1-\lambda_2)X, X<0)+P(Y\le(\lambda_1+\lambda_2)X, X>0)],\end{aligned}$$

نوشت، که در آن $Y \sim N(0,1)$ و $X \sim N(t,1)$، از یکدیگر مستقل هستند.

برهان : بر اساس تعریف تابع مولد گشتاورها می‌توان نوشت

$$\begin{aligned}M_X(t) &= \int_{-\infty}^{\infty} e^{tz}\phi(z)\Phi(\lambda_1 z+\lambda_2|z|)dz\\ &= \int_{-\infty}^{0} e^{tz}\phi(z)\Phi((\lambda_1-\lambda_2)z)dz + \int_{0}^{\infty} e^{tz}\phi(z)\Phi((\lambda_1+\lambda_2)z)dz\\ &= e^{t^2/2}\left[\int_{-\infty}^{0}\phi(z-t)\Phi((\lambda_1-\lambda_2)z)dz\right.\\ &\quad \left. + \int_{0}^{\infty}\phi(z-t)\Phi((\lambda_1+\lambda_2)z)dz\right]\\ &= e^{t^2/2}\left[P(Y\le(\lambda_1-\lambda_2)X, X<0)+P((\lambda_1+\lambda_2)X, X>0)\right].\end{aligned}$$

شبیه‌سازی از توزیع $TPSN(\lambda_1,\lambda_2)$ نیز براساس لم زیر به‌سادگی ممکن است.

لم ۲ : فرض کنید $Z_۱$ و $Z_۲$ دو متغیر تصادفی مستقل از نرمال استاندارد باشند، آنگاه

$$Z_۱ \mid Z_۲ < \lambda_۱ Z_۱ + \lambda_۲ |Z_۱| \stackrel{d}{=} TPSN(\lambda_۱, \lambda_۲).$$

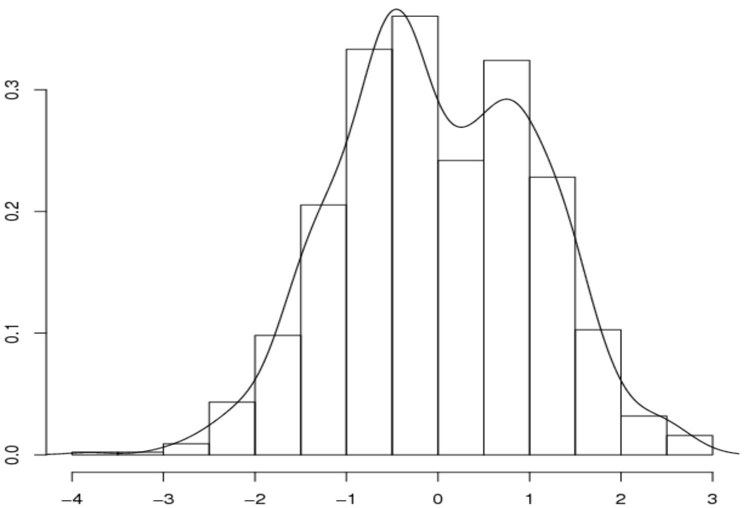
برهان :

$$\begin{aligned} P[Z_۱ \mid Z_۲ < \lambda_۱ Z_۱ + \lambda_۲ |Z_۱|] &= \frac{P_{Z_۱}(z) P[Z_۲ < \lambda_۱ Z_۱ + \lambda_۲ |Z_۱| \mid Z_۱ = z]}{P(Z_۲ < \lambda_۱ Z_۱ + \lambda_۲ |Z_۱|)} \\ &= \frac{\phi(z)\Phi(\lambda_۱ z + \lambda_۲ |z|)}{\frac{۱}{۲} + \frac{۱}{\pi}\tan^{-۱}(\frac{\lambda_۲}{\sqrt{۱+\lambda_۱^۲}})} \\ &= c(\lambda_۱, \lambda_۲)\phi(z)\Phi(\lambda_۱ z + \lambda_۲ |z|). \end{aligned}$$

شکل۳ بافت‌نگار و تابع چگالی یک نمونه تصادفی ۵۰۰۰ تایی از توزیع $TPSN(-۱, ۲)$ را که بر اساس لم ۲ شبیه‌سازی شده است، نشان می‌دهد. سایر ویژگی‌های این توزیع نیز با تکنیک‌هایی مشابه آنچه در موارد بالا مورد استفاده قرار گرفتند، قابل دستیابی هستند.

۴ نتیجه‌گیری

توزیع چوله‌نرمال و تعمیم‌های آن در دو دهه اخیر به‌صورت فزاینده‌ای مورد توجه محققان مختلف قرار گرفته و بخش قابل توجهی از ادبیات آماری را به خود اختصاص داده است. در سال‌های اخیر، برخی محققان کوشیده‌اند مباحث مرتبط با توزیع چوله‌نرمال و تعمیم‌های آن را یکپارچه و در قالب یک خانواده جامع‌تر ارائه کنند. این مقاله نشان می‌دهند که با مبنی قرار دادن توزیع‌های موزون، می‌توان نتایج بسیاری از تحقیقات سال‌های اخیر در زمینه توزیع چوله‌نرمال و تعمیم‌های آن و سایر توزیع‌های مشابه را در حالت کلی‌تر و به همراه برخی ویژگی‌های مطلوب دیگر به‌دست آورد. به عبارتی، بسیاری از نتایج مربوط به خانواده توزیع چوله‌نرمال و تعمیم‌های آن تنها حالات خاصی از نتایج حاصل از خانواده جامع‌تر توزیع‌های موزون هستند. توزیع‌های موزون در مقایسه با بسیاری خانواده‌های توزیعی دیگر



شکل ۳: بافت‌نگار و تابع چگالی یک نمونه تصادفی ۵۰۰۰ تایی شبیه‌سازی شده از توزیع $TPSN(-۱,۲)$

مانند خانواده توزیع‌های چوله‌نرمال و تعمیم‌های آن، از انعطاف پذیری به مراتب بیشتری در برازش به مشاهدات مختلف برخوردارند. تعداد و تنوع توابع وزنی که می‌توان مد نظر قرار داد، بسیار زیاد است و محقق می‌تواند در هر شرایطی با توجه به ماهیت مشاهدات وزن مناسب را به گونه‌ای انتخاب کند تا بهترین برازش ممکن حاصل شود. ویژگی‌های متفاوت و مورد علاقه‌ای مانند عدم تقارن، دو یا چند نمایی بودن و غیره، را می‌توان با در نظر گرفتن وزن مناسب با هر توزیع دلخواهی القا نمود.

تشکر و قدردانی

نویسندگان از پیشنهادهای ارزنده‌ی داوران محترم، در راستای افزایش سطح کیفی مقاله، کمال سپاسگزاری را دارند.

مراجع

علوی، م. ر. و چینی‌پرداز، ر. (۱۳۸۶). استنباط در توزیع نرمال با استفاده از نمونه‌گیری وزنی، مجله علوم آماری، ۱، **۱**، ۸۸-۷۳.

بهرامی، و.، آگاهی، ح. و رنگین، ح. (۱۳۸۸). توزیع چوله نرمال بالاکریشنان دوپارامتره، مجله پژوهش‌های آماری ایران، ۶، **۲**، ۲۴۲-۲۳۱.

خردمندی، آمنه و سنجری فارسی پور، ن. (۱۳۸۸). ویژگی‌های توزیع چوله t-نرمال و مدل‌بندی داده‌های آلودگی تالاب شادگان. مجله علوم آماری، ۳، **۱**، ۵۸-۴۷.

زادکرمی، م. (۱۳۸۷). شناسایی‌پذیری در توزیع‌های وزنی با استفاده از برآوردگر ماکسیمم درستنمایی تعمیم یافته، مجله پژوهش‌های آماری ایران، ۷، ۸۴-۷۳.

Arellano-Valle, R. B. and Azzalini, A.,(2006). On the Unification of Families of Skew-Normal Distribution, *Scandinavian Journal of Statistics*, **33**, 561-574.

Arellano-Valle, R. B. and Genton, M. G., (2005). On Fundamental Skew Distribution, *Journal of Multivariate Analysis*, **96**, 93-116.

Azzalini, A. and Valle, A. D., (1996). The Multivariate Skew-Normal Distribution, *Biometrika*, 83, 715-726.

Azzalini, A., (1985). A Class of Distributions which Includes the Normal Ones, *Scandinavian Journal of Statistics*, **12**, 171-178.

Fisher, R. A., (1934). The Effects of Methods of Ascertainment Upon The Estimation of Frequencies, *Ann. Eugenics*, **6**, 13-25.

Gonzalez, F. G., Dominguez, M., A. and Gupta, A. K., (2004). Additive Properties of Skew-Normal Vector, *Journal of Stattistics, Planning and Inference*, **126**, 521-534.

Kheradmandi, A. Mohammadzadeh, M., and Sanjari, F. N., (2010), Generalizations of the Skew t-Normal Distribution and their Properties, Journal of Sciences, Islamic Republic of Iran, Vol. 21, **4**, 343-351.

Kim, H. J., (2005). On a Class of Two-Piece Skew-Normal Distribution, *statistics*, **39**, 537-553.

Liseo, B., Loperfido, N., (2003). A Bayesian Interpertation of The Multivariate Skew-Normal Distribution, *Statistics and Probability letter*, **61**, 396-401.

Yaddegari, I., Gerami, A. and Khaledi, M. J. (2008). A generalization of the Balakrishnan Skew-Normal Distribution, *Statistica and Probability Letters*, **87**, 1165-1167.

273 Journal of Statistical Sciences, Autumn & Winter 2010
. Vol. 4, No. 2, pp 167-182.

Incorporating Various Distributional Properties Using Weigthed Distributions

Chareh, H. and Fallah, A.

Department of Statistics, Imam Khomeini International University, Ghazvin, Iran.

Abstract: This paper considers the weighted distributions in order to incorporating the topics related to construction of skew-symetric (skew-normal) and bimodal distributions. It discusses that many of skew-normal distributions disscussed in recent years researches can be studid in more general form along with some other interesting aspects in context of weigthed distributions. Two cosiderable case of the recent years reaserches have been disscussed. It is shown that the introduced distributions in these reseaches along with all of their interesting properties can be obtain from weigthed distributions perspective as only special cases.

Keywords: Skewness, Skew-Normal Distribution, Weighted Distribution, Bimodal Distribution.

Mathematics Subject Classification (2000): 62E15