دانشگاه اصفهان
دانشکده علوم
گروه آمار

پایان نامه ی دکتری رشته ی آمار

مطالعه‌ای در بعضی توزیع‌های چوله

استاد راهنما :
دکتر محمدحسین علامت‌ساز

پژوهشگر:
وحید نکوخو

مهرماه ۱۳۹۰

سپاس‌گزاری

حمد و سپاس بی‌کران خداوند مهربان را که نعمت سلامتی و توفیق کسب علم و دانش عطا فرمود.

در تدوین این رساله از مساعدت‌های عزیزانی بهره‌مند گردیده‌ام که بر خود لازم می‌دانم از آن‌ها تقدیر و تشکر به عمل آید.

در ابتدا از استاد عزیز و ارجمندم جناب آقای دکتر محمدحسین علامت‌ساز که راهنمایی این رساله را به عهده داشتند و در طول مدت تحصیل همواره مرا مورد حمایت و یاری قرار دادند و از راهنمایی‌های گهربارشان نهایت استفاده را به عمل آوردم و نیز در تصحیح رساله و رفع اشکال‌ها نهایت همکاری و لطف را مبذول فرمودند، صمیمانه تشکر و قدردانی می‌نمایم.

از اعضای محترم کمیته داوری رساله، جناب آقایان دکتر نادر نعمت‌الهی، دکتر مجید اسدی و دکتر ایرج کاظمی که با بررسی دقیق رساله و ارایه نظرات ارزشمندشان بر غنای آن افزودند تشکر و قدردانی می‌نمایم.

از دوستان عزیزم جناب آقایان دکتر حمید بیدرام و امیرحسین آقاجانی به جهت کمک‌های بی‌دریغشان تشکر می‌نمایم. از ریاست محترم دانشکده ریاضی و کامپیوتر خوانسار جناب آقای دکتر حشمت‌ا... یاوری جهت حمایت‌های بی‌دریغشان از اینجانب قدردانی می‌نمایم. همچنین از دوست بزرگوار و ارجمندم جناب آقای ایرج یادگاری که از دانسته‌ها و تجربه‌های ایشان بسیار آموختم قدردانی و سپاس‌گزاری می‌نمایم.

از اعضای محترم هیأت علمی گروه آمار دانشگاه اصفهان که در دوره‌های مختلف آموزشی زحمات فراوانی در جهت رشد و ارتقای علمی اینجانب متحمل شدند تشکر و قدردانی می‌نمایم.

در پایان، از تک تک اعضای خانواده‌ام به ویژه پدر بزرگوار و مادر عزیز و مهربانم که با تلاش فراوان در نهایت از خود گذشتگی از هیچ کمکی دریغ نکردند و شرایط مناسبی جهت تحصیل برایم فراهم نمودند سپاس‌گزاری می‌نمایم.

برای همه این عزیزان سلامتی و طول عمر همراه با عزت و سربلندی آرزومندم.

وحید نکوخو

مهرماه ۱۳۹۰

تقدیم به زیباترین واژگان حیات،

پدر بزرگوار

و

مادر عزیزم

چکیده

در متون آماری مدل‌های بسیاری جهت توصیف و تحلیل داده‌ها ارایه شده که هر یک از آن‌ها برای نوع خاصی از داده‌ها مناسب است. در این میان، توزیع نرمال یکی از مهم‌ترین توزیع‌های آماری است که در هر دوی آمار نظری و آمارکاربردی نقش کلیدی دارد. به ویژه اهمیت این توزیع در قضیه حد مرکزی آشکار می‌شود. پیش‌فرض بسیاری از روش‌های آماری نرمال بودن توزیع جامعه می‌باشد. اما، عملاً در بسیاری از مسائل روزمره و واقعی قبول چنین فرضی معقول به نظر نمی رسد زیرا امکان دارد بر خلاف توزیع نرمال، داده‌ها نوعاً نامتقارن و یا حتی دومدی باشند.

در دهه‌های اخیر بررسی توزیع واقعی چنین جوامعی مورد توجه آماردانان قرار گرفته است و پیامد آن معرفی خانواده توزیع‌هایی است که هم می‌توانند متقارن و هم نامتقارن باشند. در واقع، چنین توزیع‌هایی حاوی پارامترهایی هستند که به کمک آن‌ها می‌توان چولگی و برجستگی توزیع را کنترل کرد و در صورت لزوم با انتخاب مقادیر خاصی از این پارامترها، حالت متقارن توزیع را در نظر گرفت. توزیع چوله-نرمال یکی از معروف‌ترین این توزیع‌ها است. این توزیع می‌تواند به سمت چپ یا راست چوله باشد و به توزیع نرمال متقارن نیز تبدیل شود. اما برای استفاده از این توزیع در عمل با محدودیت‌هایی مواجه هستیم. از جمله این که توزیع چوله-نرمال تک‌مدی است و برای تحلیل داده‌هایی که از چنین جامعه‌ای پیروی می‌کنند مناسب است، در حالی که در کاربردهای عملی ممکن است تحلیل داده‌های دومدی یا چندمدی مد نظر باشد.

در این رساله پس از ارایه مقدمه و مفاهیم اولیه، توزیع چوله- نرمال را بررسی می‌کنیم. سپس تعمیم جدیدی از این توزیع ارایه می‌کنیم که هم برای تحلیل جوامع تک‌مدی و هم دومدی مناسب است. این موضوع به دلیل وجود اندک توزیع‌های دومدی رایج، ممکن است در تشریح جوامع دومدی عملاً بسیار سودمند باشد. به ویژه با ارایه مثالی از داده‌های واقعی ضرورت کاربردی آن را نیز تشریح خواهیم کرد.

به دلیل کاربردهای فراوان توزیع لاپلاس در بسیاری از زمینه‌ها از جمله پدیده‌های اقتصادی، در این رساله، ساختار نامتقارنی از این توزیع ارایه می‌کنیم که انعطاف‌پذیری بالاتری نسبت به توزیع لاپلاس متقارن دارد. در واقع، با معرفی خانواده جدیدی از توزیع‌های چوله-لاپلاس به بررسی خواص آن می‌پردازیم. همچنین ارتباطی بین توزیع‌های چوله-لاپلاس بریده شده و نمایی تعمیم‌یافته وجود دارد که به آن خواهیم پرداخت.

توزیع نمایی تعمیم‌یافته یکی از توزیع‌های چوله مهم در نظریه توزیع‌ها می باشد که کاربردهای فراوانی در تحلیل مدل‌های طول عمر دارد. این در حالی است که در بسیاری از مسائل عملی امکان دارد طول عمر یک وسیله بر مبنای یک مقیاس گسسته اندازه‌گیری شود. این موضوع نیاز به وجود توزیع‌های احتمال گسسته با خواص مطلوب، به خصوص در زمینه تحلیل داده‌ها، را خاطر نشان می‌کند. در این رساله، ساختارهای گسسته‌ای از توزیع نمایی تعمیم‌یافته ارایه می‌کنیم که برای تجزیه و تحلیل طیف گسترده‌ای از داده‌های گسسته نامنفی مناسب هستند. این ساختارها تعمیم‌هایی از توزیع‌های هندسی هستند و می‌توانند رقیب‌های مناسبی برای توزیع‌های گسسته کلاسیک باشند. مثال‌های تشریح شده در این رساله نیز حاکی از همین واقعیت هستند.

هدف این رساله، ارایه ساختارهای توزیعی و احتمالی گوناگون در موضوع مورد بررسی است به طوری که هر یک از این مدل‌ها پاسخگوی نیاز محقق باشد و وی با توجه به نیار خود بتواند از هر یک از این ساختارها استفاده کند.

کلید واژه: توزیع چوله-لاپلاس، توزیع چوله-نرمال، توزیع لاپلاس، توزیع نرمال، توزیع نمایی تعمیم‌یافته، توزیع هندسی

فهرست مطالب

عنوان **صفحه**

عنوان **صفحه**

فهرست شکل‌ها

عنوان **صفحه**

فهرست جدول‌ها

عنوان **صفحه**

فصل اول

مقدمه

خانواده توزیع‌های آماری از مهمترین ابزار اولیه‌ای هستند که در هر موضوع و مبحث آماری مطرح و مورد نیازند. یکی از مهم‌ترین اعضای این خانواده توزیع نرمال است. معمولاً تحلیل‌های آماری با فرض نرمال بودن داده‌ها انجام می‌گیرد. این در حالی است که در عمل در بسیاری از موارد توزیع داده ها نامتقارن بوده و استفاده از توزیع نرمال برای تحلیل آن‌ها مناسب نیست.

توزیع‌های چوله-متقارن[۱]، در بین توزیع‌های دیگر، مرکز توجه بسیاری از پژوهشگران آماری در دو دهه اخیر بوده است. محققین بسیاری از جمله آرنولد[۲] و لین[۳] (۲۰۰۴) و جنتون[۴] و لوپرفیدو[۵] (۲۰۰۵) ساختار کلی توزیع‌های چوله-متقارن را معرفی و خواص آن‌ها را بررسی کرده‌اند. همچنین نکوخو (۲۰۰۷) منبع جامع و مفیدی در این زمینه می‌باشد. قبل از معرفی این ساختار، ابتدا تابع چوله ساز را معرفی می‌کنیم.

تعریف ۱-۱ تابع $\pi : R \to [0,1]$ را تابع چوله ساز می‌نامیم، هر گاه به ازای هر $x \in R$، رابطه

$$\pi(x) + \pi(-x) = 1, \tag{۱-۱}$$

برقرار باشد.

---

[۱]-Skew-symmetric

[۲]-Arnold

[۳]-Lin

[۴]-Genton

[۵]-Loperfido

با

این تعریف، هر تابع توزیع متقارن حول صفر یک تابع چوله ساز است، در حالی که هر تابع چوله ساز یک تابع توزیع نیست زیرا ممکن است غیر نزولی نباشد.

حال می‌توان ساختار کلی توزیع‌های چوله-متقارن را تعریف نمود.

تعریف ۱-۲ گوییم متغیر تصادفی X دارای توزیع چوله-متقارن است، هر گاه تابع چگالی آن به صورت

$$h(x) = 2f(x)\pi(x), \quad x \in R, \tag{۱-۲}$$

باشد، که در آن π تابع چوله ساز و f هر تابع چگالی پیوسته‌ای است که حول صفر متقارن باشد.

تابع h در رابطه (۱-۲) یک تابع چگالی است، زیرا

$$\begin{aligned}
\int_{-\infty}^{\infty} h(x)\,dx &= 2\left[\int_{-\infty}^{0} f(x)\pi(x)\,dx + \int_{0}^{\infty} f(x)\pi(x)\,dx\right] \\
&= 2\left[\int_{0}^{\infty} f(x)\{1-\pi(x)\}\,dx + \int_{0}^{\infty} f(x)\pi(x)\,dx\right] \\
&= 2\int_{0}^{\infty} f(x)\,dx \\
&= 1.
\end{aligned}$$

اگر

$$\pi(x) = \frac{1}{2}, \quad x \in R,$$

آنگاه تابع چگالی رابطه (۱-۲) به تابع چگالی f که متقارن است تبدیل می‌شود. به بیان دیگر، تابع چگالی (۱-۲) تابع چگالی متقارن f را به عنوان عضوی خاص در بر می‌گیرد. اما به ازای هر تابع غیر ثابت π، تابع چگالی (۱-۲) متقارن نیست.

آرنولد و لین (۲۰۰۴) با جایگزین کردن تابع چگالی نرمال استاندارد به جای f در ساختار (۱-۲) تعمیم نامتقارنی از توزیع نرمال به دست آوردند و ویژگی‌های آن را بررسی نمودند.

فرض کنید π یک تابع چوله ساز پارامتری باشد. در این حالت می‌توان به توزیع‌های نامتقارنی دست یافت که حاوی پارامترهایی هستند که با استفاده از آن‌ها می‌توان چولگی و برجستگی این توزیع‌ها را کنترل کرد. به بیان دیگر، به توزیع‌هایی دسترسی پیدا می‌کنیم که انعطاف پذیری بهتری نسبت به توزیع‌های متقارن دارند. بنابراین انتظار داریم این گونه مدل‌ها در کاربردهای عملی مفید باشند.

یکی از مهم‌ترین اعضای خانواده توزیع‌های چوله-متقارن (۱-۲)، توزیع چوله-نرمال[۱] است که توسط آزالینی[۲] (۱۹۸۵) ارایه شد. تابع چگالی توزیع چوله-نرمال با انتخاب تابع چگالی نرمال استاندارد به جای f و

$$\pi(x) = \Phi(\lambda x), \quad x \in R,$$

که در آن Φ تابع توزیع تجمعی نرمال استاندارد و $\lambda \in R$ پارامتر ثابتی است، در رابطه (۱-۲) به دست می‌آید.

بخش قابل توجهی از متون آماری به مطالعه خانواده توزیع‌های چوله-متقارن اختصاص یافته است. آزالینی (۱۹۸۵) و (۱۹۸۶)، آرنولد و لین (۲۰۰۴) و ناداراجا[۳] و کاتس[۴] (۲۰۰۳) مراجع ارزشمندی در این زمینه هستند. بحث‌های تکمیلی و کاربردهای مهمی توسط ما[۵] و جنتون (۲۰۰۴)، آرلانو-واله[۶] و همکاران (۲۰۰۴)، اومباخ[۷] (۲۰۰۶) و گومز[۸] و همکاران (۲۰۰۷) انجام پذیرفته است. هنز[۹] (۱۹۸۶) گشتاورهای مهم توزیع چوله-نرمال از جمله گشتاورهای مرتبه زوج و فرد آن را محاسبه نمود. آزالینی و دالا واله[۱۰] (۱۹۹۶) تابع چگالی چوله-نرمال چند متغیره را معرفی کردند. آزالینی و کاپیتانیو[۱۱] (۱۹۹۹) کاربردهایی از توزیع چوله-نرمال چند متغیره ارایه نمودند. بهبودیان و همکاران (۲۰۰۶)، با استفاده از نسبت دو متغیر تصادفی مستقل چوله-نرمال و نرمال استاندارد، نوعی توزیع چوله-کوشی معرفی کردند. هوانگ[۱۲] و چن[۱۳] (۲۰۰۷) توزیع چوله-کوشی دیگری را به صورت نسبت دو متغیر تصادفی مستقل چوله-نرمال معرفی نمودند که این تعمیم، توزیع چوله-کوشی بهبودیان و همکاران را به عنوان حالتی خاص در بر می‌گیرد.

در این رساله، ضمن بررسی توزیع چوله-نرمال یک متغیره که عضو خاصی از خانواده توزیع‌های چوله-متقارن است به ارایه یک تعمیم جدید از این توزیع و بررسی خواص آن می‌پردازیم. این تعمیم جدید در نکوخو و همکاران (۲۰۱۱a) ارایه گردیده و برای تحلیل داده های تک مدی و دو مدی مناسب است.

۱-skew-normal
۲-Azzalini
۳-Nadarajah
۴-Kotz
۵-Ma
۶-Arellano-Valle
۷-Umbach
۸-Gómez
۹-Henze
۱۰-Dalla Valle
۱۱-Capitanio
۱۲-Huang
۱۳-Chen

یکی از توزیع‌های مهم آماری، توزیع لاپلاس[۱] است که از اختلاف دو متغیر تصادفی مستقل و هم‌توزیع نمایی به دست می‌آید. توزیع لاپلاس در بسیاری از زمینه‌ها از جمله پدیده‌های اقتصادی کاربرد دارد. از طرفی با جایگزین کردن این توزیع در ساختار تابع چگالی (۱-۲) و انتخاب تابع توزیع تجمعی متغیر تصادفی لاپلاس به جای π در این ساختار، می‌توان به تابع چگالی چوله-لاپلاس نامتقارنی دست یافت که انعطاف پذیری بهتری نسبت به توزیع لاپلاس متقارن دارد.

آریال[۲] و رائو[۳] (۲۰۰۵) برخی از خواص توزیع چوله-لاپلاس بریده شده را مطالعه کردند. کوزوبووسکی[۴] و نولان[۵] (۲۰۰۸) نشان دادند که تابع چگالی چوله-لاپلاس بینهایت بار تقسیم پذیر[۶] است. این در حالی است که در زمینه مطالعه توزیع‌های چوله-متقارن کمتر به توزیع چوله-لاپلاس پرداخته شده است. بر این اساس نکوخو و علامت‌ساز (۲۰۱۱) خانواده جدیدی از توزیع‌های چوله-لاپلاس را معرفی و خواص آن را بررسی کردند که به معرفی این خانواده نیز خواهیم پرداخت.

ترتیب‌بندی‌ها در شاخه‌های مختلف احتمال و آمار کاربرد دارند. هدف از ارایه یک مسأله ترتیب‌بندی مقایسه توزیع چند جامعه یا مقایسه توزیع چند آماره مبتنی بر یک جامعه می‌باشد. مثلاً به این روش می‌توان متغیرهای تصادفی مرتب شده از توزیع‌های متفاوت و یا چندین متغیر تصادفی مرتب شده از یک توزیع را با یکدیگر مقایسه نمود. این در حالی است که در یک خانواده چوله-متقارن این سؤال مطرح می‌شود که از لحاظ انواع ترتیب‌بندی‌های تصادفی چه روابطی بین توزیع‌های چوله و متقارن وجود دارد. نکوخو و علامت‌ساز (۲۰۱۱) به برخی از این سؤال ها پاسخ داده‌اند.

نکوخو و علامت‌ساز (۲۰۱۱) همچنین ارتباطی بین توزیع‌های چوله-لاپلاس بریده شده و نمایی تعمیم‌یافته به دست آورده‌اند. توزیع نمایی تعمیم‌یافته یکی از توزیع‌های چوله مهم در نظریه توزیع‌ها می‌باشد و در هر دوی آمار نظری و کاربردی مطرح است. این توزیع کاربردهای فراوانی به خصوص در تحلیل مدل‌های طول عمر و

---

[۱]-Laplace

[۲]-Aryal

[۳]-Rao

[۴]-Kozubowski

[۵]-Nolan

[۶]-Infinitely divisible

قابلیت اعتماد دارد. گوپتا[۱] و کوندو[۲] در سلسله مقالاتی که از مهمترین آن‌ها می‌توان به گوپتا و کوندو (۱۹۹۹) اشاره نمود، مطالب ارزشمندی درباره توزیع نمایی تعمیم‌یافته ارایه کردند. آن‌ها، در این مقاله، ویژگی‌های توزیع نمایی تعمیم‌یافته سه پارامتری را بررسی نموده و مقایسه‌هایی بین این توزیع و توزیع‌های گاما و وایبل انجام دادند. اما در بسیاری از مسائل امکان دارد طول عمر یک وسیله بر مبنای یک مقیاس گسسته اندازه‌گیری شود. در دو دهه اخیر توزیع‌های گسسته کلاسیک مانند هندسی و دوجمله‌ای منفی برای مدل‌سازی طول عمرهای گسسته به کار رفته است. این در حالی است که نیاز به وجود توزیع‌های احتمال گسسته بیشتر، با انعطاف پذیری بهتر و ویژگی‌های برتر، به خصوص در زمینه تحلیل داده‌ها احساس می‌شود. نکوخو و همکاران (۲۰۱۱b) و (۲۰۱۱c) ساختارهای گسسته ای از توزیع نمایی تعمیم‌یافته ارایه و معرفی نمودند که در این رساله با آن‌ها آشنا می‌شویم. همان‌گونه که توزیع نمایی تعمیم‌یافته، توزیع نمایی را به عنوان حالتی خاص در بر می‌گیرد، تعمیم‌های ارایه شده توسط نکوخو و همکاران نیز توزیع‌های هندسی را به عنوان حالت‌های خاص در خود دارند.

در زمینه گسسته‌سازی توزیع‌های پیوسته فعالیت‌های چشمگیری صورت گرفته است. برای مثال می‌توان به توزیع نرمال گسسته لیسمن[۳] و وان زویلن[۴] (۱۹۷۲) اشاره نمود که توسط کمپ[۵] (۱۹۹۷) مطالعه شد. در این زمینه زمینه داس‌گوپتا[۶] (۱۹۹۳) و زابلوسکی[۷] (۲۰۰۱) نیز بررسی‌هایی انجام دادند. اخیراً اینوسا[۸] و کوزوبووسکی (۲۰۰۶) و کوزوبووسکی و اینوسا (۲۰۰۶)، به ترتیب، توزیع‌های لاپلاس و چوله- لاپلاس را روی اعداد صحیح ایجاد و خواص آن‌ها را معرفی نمودند. همچنین تعمیم‌هایی از توزیع هندسی به دست آمده است که از مهم‌ترین آن‌ها تعمیم به دست آمده توسط گومز-دنیز[۹] (۲۰۱۰) می‌باشد.

همان‌گونه که بیان شد، در متون آماری مدل‌های بسیاری جهت توصیف و تحلیل داده‌ها ارایه شده که هر یک از آن‌ها برای نوع خاصی از داده‌ها مناسب است. از طرفی در بسیاری از مسائل عملی با یک جامعه متقارن و پیوسته مواجه نیستیم. بنا براین ارایه مدل‌هایی با انعطاف‌پذیری بالا و خواص مطلوب در چنین حالت‌هایی از

۱-Gupta
۲-Kundu
۳-Lisman
۴-van Zuylen
۵-Kemp
۶- Dasgupta
۷- Szablowski
۸-Inusah
۹-Gómez-Déniz

اهمیت خاصی برخوردار است. در این رساله، ساختارهای گوناگون و متفاوتی در موضوع مورد بررسی ارایه می‌گردد که هر یک از مدل‌ها پاسخگوی نیاز محقق و آماردان می باشد به طوری که محقق با توجه به نیاز خود می‌تواند از هر یک از این ساختارها استفاده کند. این رساله در پنج فصل تدوین شده است. در فصل دوم توزیع چوله-نرمال یک متغیره و ویژگی‌های مختلف آن را بررسی می‌کنیم. در فصل سوم یک تعمیم جدید از توزیع چوله-نرمال ارایه و ویژگی‌های مختلف آن را بررسی می کنیم. در فصل چهارم به معرفی خانواده جدیدی از توزیع‌های چوله-لاپلاس می‌پردازیم و خصوصیات مختلف آن را بیان می‌کنیم. در این فصل همچنین به بحث پیرامون برخی از انواع ترتیب‌بندی‌های تصادفی بین توزیع‌های متقارن و چوله می‌پردازیم. در فصل پنجم ساختارهای گسسته متفاوتی از توزیع نمایی تعمیم‌یافته ارایه و ویژگی‌های متفاوت آن‌ها را بررسی می کنیم.

فصل دوم

توزیع چوله- نرمال یک متغیره و ویژگی‌های آن

۲-۱ مقدمه

نظریه توزیع‌ها یکی از شاخه‌های مهم علم آمار است و توزیع نرمال از مهم‌ترین و اساسی‌ترین توزیع‌های تک‌مدی است. این توزیع ویژگی‌هایی دارد که باعث افزایش کاربردهای آن در عمل شده است. معمولاً در تحلیل‌های آماری فرض نرمال بودن داده‌ها را می‌پذیرند، سپس به استنباط آماری می‌پردازند. اما در بسیاری از مسائل عملی امکان دارد با یک جامعه تک‌مدی روبرو شویم که فرض نرمال بودن مجموعه داده‌ها رد شود. هنگامی که داده‌ها نامتقارن هستند، استفاده از توزیع نرمال برای تحلیل آن‌ها مناسب نیست. در این موارد اغلب سعی می‌شود با استفاده از تبدیلی مناسب، توزیع داده‌ها را به نرمال نزدیک کرده و سپس تحلیل داده‌های تبدیل‌یافته انجام گیرد. در این میان پیدا کردن تبدیل مناسب شاید کاری دشوار باشد و حتی در عمل ممکن است چنین تبدیلی یافت نشود. در این میان اگر توزیع نامتقارنی وجود داشته باشد که دارای خواصی مشابه توزیع نرمال باشد، این توزیع می‌تواند نقشی اساسی در تحلیل داده‌های نامتقارن ایفا کند. یکی از این توزیع‌ها که اخیراً معرفی شده و مورد توجه بسیاری از آماردانان و پژوهشگران قرار گرفته است و خصوصیاتی مشابه توزیع نرمال دارد، توزیع چوله-نرمال است. این توزیع نامتقارن دارای یک پارامتر برای تنظیم چولگی است و مقادیر مختلف آن شکل‌های متفاوتی به تابع چگالی خواهد داد. اگر مقدار این پارامتر صفر گردد توزیع چوله-نرمال به توزیع نرمال تبدیل می شود. در واقع توزیع چوله-نرمال، توزیع نرمال را به عنوان حالتی خاص در بر می‌گیرد. توزیع چوله-نرمال یک متغیره نخستین بار توسط آزالینی (۱۹۸۵) ارایه شد. آزالینی و دالاواله (۱۹۹۶) توزیع چوله-نرمال چند متغیره را تعریف کردند. آزالینی و کاپیتانیو (۱۹۹۹) ویژگی‌های توزیع چوله-نرمال چند متغیره را بررسی کردند.

هنز (۱۹۸۶) گشتاورهای مهم توزیع چوله- نرمال از جمله گشتاورهای مرتبه زوج و فرد آن را محاسبه نمود. همچنین تعمیم‌های مختلفی از این توزیع، برای مثال، توسط آرلانو-واله و همکاران (۲۰۰۴)، ما و جنتون (۲۰۰۴) و شرفی و بهبودیان (۲۰۰۸) ارایه شده است. اخیراً نکوخو و همکاران (۲۰۱۱a) تعمیم دیگری از این توزیع را ارایه کرده‌اند.

در این فصل توزیع چوله-نرمال یک متغیره معرفی می‌گردد. سپس مفاهیم و تعاریف مقدماتی آن، روش‌های مختلف شبیه سازی از آن، تابع توزیع و تابع مولد گشتاور آن ارایه می‌گردد. همچنین پیرامون توزیع ترکیب‌های خطی از متغیرهای چوله-نرمال نیز بحث می‌شود.

۲-۲ تعاریف و نمادها

در ابتدای این بخش یک قضیه مهم که توسط آزالینی و کاپیتانیو (۱۹۹۹) ارایه شده است را بیان می‌کنیم. بعد از آن به سراغ توزیع چوله-نرمال رفته و با معرفی آن به بررسی ویژگی‌های این توزیع می‌پردازیم.

قضیه ۲-۱ فرض کنید f یک تابع چگالی متقارن حول صفر، G تابع توزیعی مطلقاً پیوسته به گونه‌ای که مشتق آن حول صفر متقارن و w یک تابع فرد پیوسته باشد. در این صورت

$$h(x) = 2f(x)G\big(w(x)\big), \qquad (۲-۱)$$

یک تابع چگالی است.

اثبات. تابع h در رابطه (۲-۱) حالت خاصی از تابع چگالی (۱-۲) می باشد. به عبارت دیگر با انتخاب $\pi(x) = G\big(w(x)\big)$ در رابطه (۱-۲)، تابع چگالی رابطه (۲-۱) به دست می آید.

تابع چگالی (۲-۱) عضو خاصی از خانواده توزیع‌های چوله-متقارن معرفی شده در رابطه (۱-۲) می‌باشد. با این حال، تابع چگالی (۲-۱) را نیز تابع چگالی چوله-متقارن می‌نامیم.

نتیجه ۲-۱ در قضیه (۲-۱)، اگر $w(x) = \lambda x$، که در آن λ یک عدد حقیقی ثابت است، آنگاه تابع چگالی (۲-۱) به

$$h(x) = 2f(x)G(\lambda x), \qquad (۲-۲)$$

خلاصه می‌شود.

اگر در رابطه (۲-۲) تابع چگالی f و تابع توزیع G ثابت نگه داشته شوند، به ازای مقادیر مختلف λ، خانواده ای از توزیع ها تولید می‌شود که به ازای $\lambda = 0$ متقارن و برابر f بوده و به ازای $\lambda \neq 0$ چوله است.

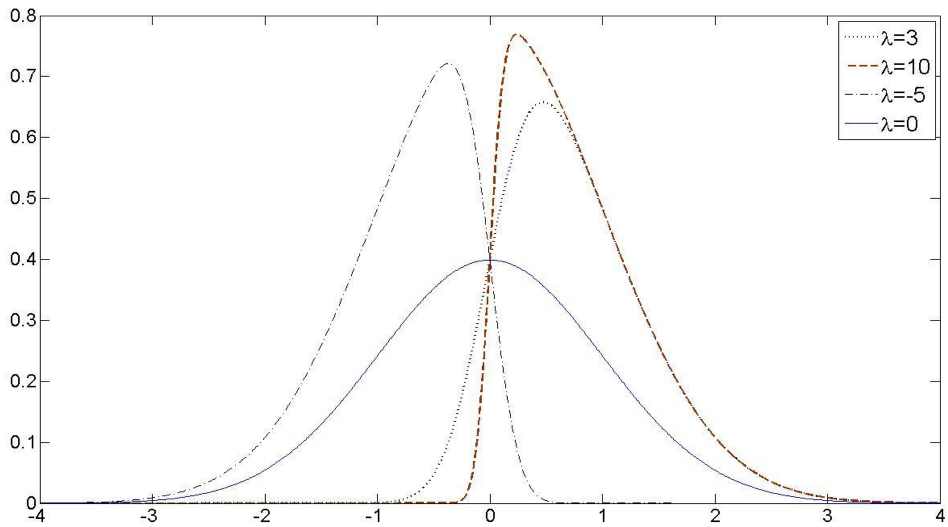
توزیع چوله-نرمال ارایه شده توسط آزالینی (۱۹۸۵) حالت خاصی از تابع چگالی (۲-۲) بوده و سرآغاز مطالعات انجام گرفته در زمینه توزیع‌های چوله-متقارن است.

تعریف ۲-۱ گوییم متغیر تصادفی Z_λ دارای توزیع چوله- نرمال با پارامتر حقیقی λ است اگر تابع چگالی آن به صورت

$$\phi(z;\lambda) = 2\phi(z)\,\Phi(\lambda z), \qquad z \in R, \tag{۲-۳}$$

باشد که در آن ϕ و Φ به ترتیب توابع چگالی و توزیع تجمعی نرمال استاندارد هستند. هنگامی که Z_λ دارای توزیع چوله-نرمال است به اختصار آن را با نماد $Z_\lambda \sim SN(\lambda)$ نمایش می‌دهیم.

پارامتر λ در این توزیع را می‌توان پارامتر کنترل چولگی نامید. تابع چگالی چوله-نرمال به ازای مقادیر مثبت λ چوله به راست، به ازای مقادیر منفی چوله به چپ و به ازای $\lambda = 0$ متقارن و به چگالی نرمال استاندارد تبدیل می‌شود. شکل (۲-۱) تابع چگالی چوله-نرمال را به ازای چند مقدار مختلف پارامتر آن نشان می‌دهد.



شکل ۲-۱ تابع چگالی چوله- نرمال به ازای چند مقدار مختلف پارامتر λ

۲-۳ ویژگی‌های توزیع چوله-نرمال

در این بخش بعضی از ویژگی‌های مهم توزیع چوله-نرمال آزالینی را بیان می‌کنیم. همچنین به بررسی تابع توزیع، تابع مولد گشتاور و گشتاورهای آن می پردازیم. سپس بعضی از روش‌های ساختن توزیع چوله-نرمال آزالینی را ارایه می‌کنیم. برای ملاحظه اثبات های تفصیلی نتایج ارایه شده در این فصل، که اثبات آن‌ها آورده نشده است، می توان به نکوخو (۲۰۰۷) مراجعه کرد.

۲-۳-۱ چند خاصیت تابع چگالی چوله-نرمال

فرض کنید متغیرهای تصادفی Z و Z_λ به ترتیب دارای توزیع نرمال استاندارد و توزیع چوله-نرمال باشند. در این صورت روابط و نتایج زیر را خواهیم داشت.

الف) $-Z_\lambda$ دارای توزیع چوله-نرمال با پارامتر $-\lambda$ است.

ب) اگر $\lambda = 0$، آنگاه $Z_\lambda \sim N(0,1)$. به عبارت دیگر در این حالت داریم $Z_\lambda \stackrel{d}{=} Z$.

پ) اگر $\lambda = 1$، آنگاه توزیع Z_λ برابر توزیع آماره ماکسیمم در یک نمونه تصادفی دو تایی از توزیع نرمال استاندارد است. اگر $\lambda = -1$، آنگاه توزیع Z_λ برابر توزیع آماره می‌نیمم در چنین حالتی است.

ت) اگر $\lambda \to +\infty$، آنگاه توزیع Z_λ به توزیع نیم نرمال[1] میل می‌کند. به عبارت دیگر در این حالت داریم $Z_\lambda \stackrel{d}{=} |Z|$. همچنین اگر $\lambda \to -\infty$، آنگاه خواهیم داشت $Z_\lambda \stackrel{d}{=} -|Z|$.

ث) تابع چگالی چوله-نرمال قویاً تک‌مدی[2] است (یعنی پیچش آن با هر توزیع تک‌مدی دیگری توزیعی تک‌مدی خواهد بود).

ج) اگر t تابعی زوج باشد، آنگاه رابطه هم توزیعی $t(Z_\lambda) \stackrel{d}{=} t(Z)$ برقرار است. برای مثال داریم $|Z_\lambda| \stackrel{d}{=} |Z|$ و $Z_\lambda^2 \stackrel{d}{=} Z^2$.

۲-۳-۲ تابع توزیع متغیر تصادفی چوله-نرمال

تابع توزیع متغیر تصادفی چوله-نرمال عبارت است از

$$\Phi(z;\lambda) = \int_{-\infty}^{z} 2\phi(t)\Phi(\lambda t)\, dt = 2\int_{-\infty}^{z}\int_{-\infty}^{\lambda t} \phi(t)\phi(u)\, du\, dt.$$

در لم زیر ارتباط تابع توزیع چوله-نرمال و تابع توزیع نرمال استاندارد ارایه می شود.

[1]-Half normal distribution

[2]-Strongly unimodal