دانشگاه صنعتی اصفهان

دانشکده علوم ریاضی

توزیع‌های متقارن چوله

پایان‌نامه کارشناسی ارشد آمار ریاضی

فریده علی‌محمدی

استاد راهنما

دکتر صفیه محمودی

دانشگاه صنعتی اصفهان

دانشکده علوم ریاضی

پایان‌نامه کارشناسی ارشد آمار ریاضی خانم فریده علی‌محمدی

تحت عنوان

توزیع‌های متقارن چوله

در تاریخ توسط کمیته تخصصی زیر مورد بررسی و تصویب نهائی قرار گرفت.

۱ــ استاد راهنمای پایان‌نامه دکتر صفیه محمودی

۲ــ استاد مشاور پایان‌نامه دکتر سعید پولادساز

۳ــ استاد داور ۱

()

۴ــ استاد داور ۲

سرپرست تحصیلات تکمیلی دانشکده دکتر رسول نصر اصفهانی

فهرست مطالب

چکیده:

در طول دهه‌ی گذشته علاقه‌ی فزاینده‌ای به ایجاد کلاس‌های پارامتری از توزیع‌ها برای نمایش مناسب داده‌ها با رد فرض‌های غیر واقعی وجود داشته است. انگیزه‌ی اصلی این امر از مدل‌بندی داده‌هایی ناشی شده است که اغلب بعضی فرض‌های استاندارد از جمله نرمال بودن را ندارند. کلاس توزیع‌های متقارن چوله یک مثال از این توزیع‌ها می‌باشد. در این پایان‌نامه، این کلاس از توزیع‌ها و ویژگی‌های اساسی آن‌ها مورد بررسی قرار گرفته است. توزیع نرمال چوله یکی از این توزیع‌ها است که با وجود چولگی بعضی از ویژگی‌های توزیع نرمال را دارد. در این تحقیق، مسأله‌ی برآورد پارامتر به روش گشتاوری، ماکزیمم درستنمایی و ماکزیمم درستنمایی اصلاح شده برای توزیع نرمال چوله‌ی یک پارامتری بررسی شده است. هم چنین دو روش معمول آزمون نیکویی برازش برای این توزیع مورد بحث قرار گرفته است.

فصل ۱

مقدمه

۱—۱ تقارن

از دیرباز تقارن مفهوم مهمی در هنر، ریاضیات و علوم دیگر بوده است. تقارن در زیباشناسی یک قانون، در ریاضیات محصول ساختار هندسی و در فلسفه، تعادل و هارمونی است. در علم آمار و در توزیع‌های احتمال تقارن دارای تعابیر مختلفی می‌باشد. می‌توان تقارن را به عنوان ویژگی ساختاری یک توزیع بر اساس تابع توزیع، تابع مشخصه و یا تابع چگالی آن توزیع تعریف کرد. معمولاً توزیع X حول θ را برای تعریف تقارن متغیر تصادفی X مورد استفاده قرار می‌دهند.

تعریف ۱.۱ در توزیع‌های یک متغیره اگر

$$X - \theta \stackrel{d}{=} \theta - X, \qquad (۱—۱)$$

آن‌گاه متغیر تصادفی X را حول θ متقارن گویند.

معمولاً منحنی فراوانی نرمال استاندارد به دلیل زنگی شکل و متقارن بودن مبنای قضاوت در مورد تقارن یک توزیع احتمال قرار می‌گیرد. هم چنین مفهوم تقارن در توزیع‌های چند متغیره نیز وجود دارد و دارای انواع مختلفی از جمله تقارن کروی [۱] ، تقارن بیضی [۲] ، تقارن مرکزی [۳] و تقارن زاویه‌ای [۴]

[۱] spherical symmetry
[۲] elliptical symmetry
[۳] central symmetry
[۴] angular symmetry

می‌باشد.

۱—۲ اندازه‌ی چولگی و عدم تقارن

وقتی که تقارن به عنوان یک ویژگی توزیع متغیر تصادفی در نظر گرفته می‌شود، می‌توان چولگی را به عنوان نوع یا جهت انحراف از تقارن تعریف کرد. به طور کلی واژه‌ی چولگی برای اشاره به هر چیزی که خارج از تقارن باشد یا به یک طرف کج شده باشد مورد استفاده قرار می‌گیرد. در آمار و تئوری احتمال چولگی اندازه‌ای از عدم تقارن توزیع متغیر تصادفی است. اندازه‌ی چولگی باید از پارامترهای مکان و مقیاس آزاد باشد و برای توزیع‌های متقارن به صفر کاهش یابد. علامت و اندازه‌ی چولگی به ترتیب جهت و مقدار عدم تقارن یک متغیر تصادفی را نشان می‌دهند. به طور کلی دو نوع چولگی تعریف شده است

۱) چولگی مثبت

در این نوع چولگی دم راست منحنی تابع چگالی طولانی‌تر است و بیشترین جرم توزیع در سمت چپ متمرکز شده است، چنین توزیعی را چوله به راست گویند.

۲) چولگی منفی

در این حالت دم چپ منحنی تابع چگالی طولانی‌تر است و بیشترین جرم توزیع در سمت راست متمرکز شده است، چنین توزیعی را چوله به چپ گویند.

از جمله توزیع‌های چوله‌ی مشهور می‌توان از کای‌دو و توزیع F نام برد. شاخص‌های متفاوتی برای اندازه‌گیری چولگی یک متغیر تصادفی معرفی شده‌اند که اغلب آن‌ها بر اساس گشتاورهای توزیع تعریف می‌شوند. از جمله

ضرایب چولگی پیرسن : کارل پیرسن در سال ۱۸۹۵ دو فرمول ساده را برای اندازه‌گیری چولگی مطرح کرد.

$$\frac{E(X)-M}{\sigma},$$
$$\frac{۳\Big(E(X)-m\Big)}{\sigma},$$

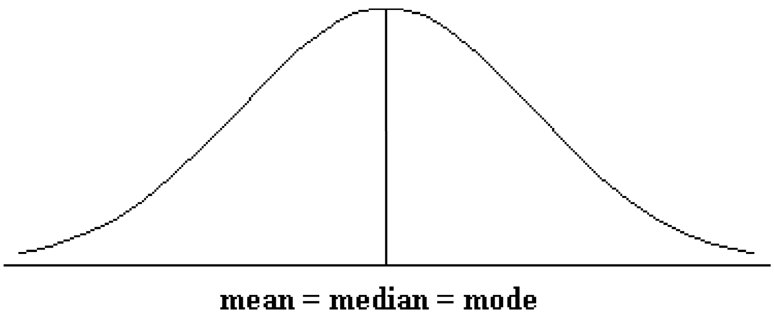
به طوری که M مد، σ انحراف معیار و m میانه‌ی متغیر تصادفی X است.

گشتاور استاندارد شده‌ی مرتبه‌ی سوم : این گشتاور که شاخص دیگری برای اندازه‌گیری چولگی است با γ_1 نشان داده می‌شود و به صورت

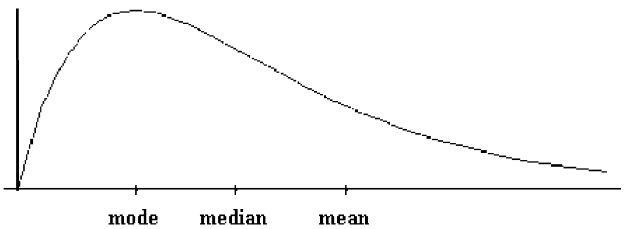
$$\gamma_1 = \frac{\mu_3}{\sigma^3},$$

است. که در آن μ_3 گشتاور مرتبه‌ی سوم حول میانگین، $\mu_3 = E(X-\mu)^3$، و σ انحراف معیار می‌باشد.

روابطی که بیان شد همگی کمیت‌های ساده و مفیدی برای سنجش میزان عدم تقارن یک توزیع هستند و به واحد اندازه‌گیری نیز بستگی ندارند. در صورتی که داده‌ها نسبت به میانگین متقارن باشند، ضرایب گفته شده برابر با صفر هستند. شکل‌های ۱—۱، ۱—۲ و ۱—۳ حالت‌های مختلف چولگی را نشان می‌دهند.

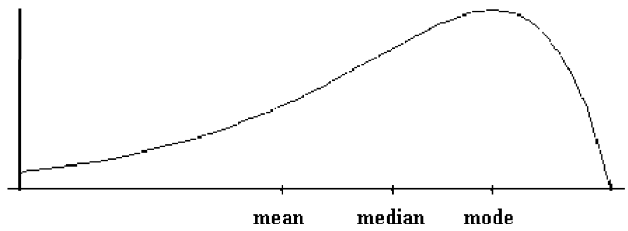


شکل ۱—۱: چولگی صفر.



شکل ۱—۲: چولگی مثبت.

همان طور که ملاحظه می‌شود مقایسه‌ی بین میانگین، میانه و مد در یک توزیع یک راه مناسب برای ارزیابی تقارن آن توزیع است. وجود فاصله بین این کمیت‌ها در یک توزیع حاکی از چولگی آن توزیع و علامت اختلاف بین آن‌ها جهت چولگی را نشان می‌دهد. برای یک توزیع متقارن میانگین، میانه و مد بر هم منطبق هستند.



شکل ۱—۳: چولگی منفی.

۱—۳ ساختن چگالی چوله

توزیع نرمال قرن‌هاست که شناخته شده است. عمومیت این توزیع به خاطر سادگی تحلیل و قضیه‌ی حد مرکزی (که در سال ۱۷۳۳ میلادی توسط دموار[۵] بیان شد) می‌باشد. هم چنین حالت چند متغیره‌ی آن فرم پیچیده‌ای ندارد و توزیع‌های شرطی و حاشیه‌ای آن نیز نرمال می‌باشند خاصیتی که در توزیع‌های چند متغیره‌ی دیگر کمتر دیده می‌شود. در سال ۱۸۰۹ کارل فریدریش گاوس [۶] دریافت که توزیع خطاها نرمال می‌باشد و نیز لاپلاس[۷] در سال ۱۸۱۰ نشان داد که میانگین نمونه‌ای دارای توزیع نرمال می‌باشد. توزیع نرمال در احتمال، فرایند تصادفی و آمار بیز کاربرد گسترده‌ای دارد.

اما شک و تردیدهایی در باره‌ی استفاده‌ی نامحدود از این توزیع وجود دارد. گری [۸] [۱۳] بیان کرده است:«نرمال بودن یک افسانه است؛ توزیع نرمال هیچ‌گاه وجود نداشته و وجود نخواهد داشت.» هیل [۹] و دیکسون [۱۰] [۱۹] نشان دادند که توزیع داده‌های واقعی در کاربرد عملی اغلب چوله می‌باشد اما در بسیاری از موارد این چولگی نادیده گرفته می‌شود و در عمل توزیع نرمال استفاده می‌شود.

بعضی از آمارادانان در جستجوی خانواده‌ی توزیع چوله‌ای بوده‌اند که شامل توزیع نرمال باشد و دارای پارامتر چولگی باشد که بتوان آن را با استفاده از داده‌ها برآورد کرد و هم چنین بسیاری از ویژگی‌های توزیع نرمال را در حالت تک متغیره و چند متغیره داشته باشد. برای این منظور توزیع چوله‌ای پیشنهاد شد که با ضرب کردن تابع چگالی نرمال در سری توانی به دست می‌آید و ضرایب آن را می‌توان به روش کمترین مربعات محاسبه کرد که منجر به سری‌های مشهور گرام—چارلیه [۱۱] و اجوورث [۱۲] می‌شود که از

[۵] De Moiver
[۶] Carl Friedrich Guass
[۷] Laplace
[۸] Geary
[۹] Hill
[۱۰] Dixon
[۱۱] Gram-Charlier series
[۱۲] Edgeworth

اولین شکل توزیع‌های چوله در آمار می‌باشد اما بسیاری از ویژگی‌های توزیع نرمال را ندارد. برای جزئیات بیشتر می‌توان به تایل [۱۳] [۳۴] مراجعه کرد. خانواده‌ی دیگری از توزیع‌ها که شامل توزیع نرمال می‌باشد خانواده‌ی توزیع توان نمایی [۱۴] است که تحت عنوان توزیع خطای تعمیم یافته نیز شناخته می‌شود، تابع چگالی این خانواده متناسب با

$$\exp\Big(-\frac{|x|^r}{r}\Big),$$

می‌باشد[۹]. که در اینجا پارامتر r کشیدگی تابع چگالی را تعیین می‌کند، ولی چولگی هم چنان صفر است. پرنتیک [۱۵] [۳۱] کلاس پارامتری وسیعی از توزیع‌ها را مطالعه کرد که تحت شرایط پارامتری مناسب شامل توزیع نرمال استاندارد می‌باشد. توزیع‌های g و h توکی [۱۶] (برای جزئیات بیشتر به هواگلین[۱۷] [۲۰] رجوع کنید) نوع دیگری از توزیع‌های چوله هستند ولی چگالی آن‌ها بسط ساده‌ای ندارد.

در طول دهه‌ی گذشته علاقه‌ی فزاینده‌ای به ایجاد کلاس پارامتری انعطاف‌پذیری از توزیع‌ها وجود داشته است که شامل چولگی و کشیدگی بوده و از توزیع نرمال متفاوت باشد که به وسیله‌ی آن بتوان داده‌ها را تا آنجایی که ممکن است به طور مناسب و با رد فرض‌های غیر واقعی نمایش داد. علت اصلی این امر وجود مجموعه داده‌هایی می‌باشد که فرض‌های استانداردی مانند نرمال بودن را برقرار نمی‌کنند.

ایده‌ی کلی ایجاد مدل‌هایی که از نرمال بودن انحراف داشته باشند تغییر تابع چگالی نرمال استاندارد می‌باشد. اگر چه این روش مدت‌ها در ادبیات آماری استفاده شده است اما این آزالینی [۱۸] [۳] بود که اولین بار با استفاده از آن یک تابع چگالی احتمال جدید معرفی کرد.

آزالینی در سال ۱۹۸۵ تابع چگالی احتمال جدیدی را معرفی کرد که از حاصل ضرب توابع چگالی و توزیع متغیرهای تصادفی متقارن به دست می‌آید. انگیزه‌ی اصلی او برای معرفی این خانواده از توزیع‌ها یافتن کلاس جدیدی از توزیع‌ها بود که

۱) شامل توزیع نرمال باشد.

۲) فرم تابع چگالی آن ساده باشد.

۳) دامنه‌ی وسیعی از چولگی و کشیدگی را داشته باشد.

کلاسی از توزیع‌های احتمال که شرایط بالا را برآورده کند پتانسیل بالایی در مدل‌سازی آماری دارد. کلاس توزیع‌های معرفی شده توسط آزالینی دارای ویژگی اول و دوم می‌باشد ولی ویژگی سوم را برآورده

[۱۳] Thiele
[۱۴] exponential power distribution
[۱۵] Prentice
[۱۶] Tukey's g- and- h distribution
[۱۷] Hoaglin
[۱۸] Azzalini

نمی‌کند و این یکی از محدودیت‌های این توزیع است.

آزالینی [۳] دریافت که اگر f یک تابع چگالی احتمال متقارن باشد یعنی همواره $f(x)=f(-x)$ و G تابع توزیع مطلقاً پیوسته‌ای باشد که مشتق آن، G'، حول نقطه‌ی صفر متقارن است، آن‌گاه به ازای هر مقدار حقیقی y و λ

$$2f(y)G(\lambda y), \qquad (1-2)$$

یک تابع چگالی احتمال می‌باشد. هم چنین او با در نظر گرفتن یک حالت خاص از (۱−۲)، هنگامی که f تابع چگالی نرمال استاندارد و G تابع توزیع نرمال استاندارد باشد، کلاس جدیدی از توزیع‌های احتمال را تحت عنوان نرمال چوله[۱۹] معرفی کرد و درباره‌ی ویژگی‌های آن به بحث پرداخت. گوپتا[۲۰] و چانگ[۲۱] [۱۴] حالت چند متغیره‌ی تابع چگالی احتمال (۱−۲) و ویژگی‌های آن را مطرح کردند. وانگ[۲۲] و همکاران [۳۷] درباره‌ی توزیع توابع زوج متغیرهای تصادفی متقارن چوله به بحث پرداختند. اومبچ[۲۳] [۳۵] درباره‌ی امید ریاضی و گشتاورهای توزیع‌های متقارن چوله بحث کرد. هم چنین هوانگ[۲۴] و چن[۲۵] [۲۱] فرم درجه دوم توزیع‌های متقارن چوله را در حالتی که تابع چگالی f در (۱−۲)، تابع چگالی نرمال استاندارد باشد مطرح کرده و نشان دادند که توزیع آن از پارامتر λ مستقل است. هنز[۲۶] [۱۸] فرم بسته‌ای را برای گشتاورهای توزیع نرمال چوله در حالت یک پارامتری به دست آورد. آزالینی و دالاویل[۲۷] [۵] حالت چند متغیره نرمال چوله را معرفی کردند. نویسندگان دیگر از جمله گوپتا و همکاران [۱٦] حالت‌های مختلف چند متغیره‌ی دیگری را برای نرمال چوله به دست آوردند. بعدها این ایده به سایر توزیع‌های متقارن دیگر نظیر توزیع t، کشی و... تعمیم داده شد. ساتوری[۲۸] [۳۳] نیز در باره‌ی برآورد پارامتر شکل توزیع نرمال چوله و مشکلات مربوط به برآورد این پارامتر به بحث پرداخت. گوپتا و چن [۱۵] نیز آزمون نیکویی برازش را برای توزیع نرمال چوله یک پارامتری مطرح کردند.

این پایان‌نامه شامل ۵ فصل است. در فصل دوم توزیع‌های متقارن چوله در حالت یک متغیره و چند متغیره و ویژگی‌های مهم آن در این دو حالت مطرح شده است. در فصل سوم توزیع نرمال چوله، ویژگی‌های آن، روش‌های مختلف تولید متغیرهای تصادفی از این توزیع، حالت چند متغیره و فرم‌های

[۱۹] skew-normal
[۲۰] Gupta
[۲۱] Chang
[۲۲] Wang
[۲۳] Umbach
[۲۴] Huang
[۲۵] Chen
[۲٦] Henze
[۲۷] Dalla Valle
[۲۸] Sartori

درجه دوم توزیع‌های متقارن چوله با هسته‌ی نرمال استاندارد مطرح شده است. در فصل چهارم مشکلات مربوط به برآورد پارامتر λ مطرح و سه روش برآورد گشتاوری، برآورد ماکزیمم درستنمایی و ماکزیمم درستنمایی اصلاح شده برای این پارامتر مورد بررسی قرار گرفته است. در فصل پنجم دو آزمون معروف کی‌دو و کلموگروف—اسمیرنوف برای این توزیع انجام گرفته است. در پیوست نیز بعضی تعاریف و قضایا و برنامه‌های کامپیوتری ارائه شده است.

فصل ۲

توزیع‌های متقارن چوله

۲–۱ مقدمه

خانواده‌ی توزیع‌های متقارن چوله، خانواده‌ای از توزیع‌های چوله حول یک توزیع متقارن می‌باشند. توزیع متقارن را به اصطلاح هسته [۱] می‌گویند که عضو مرکزی این خانواده می‌باشد. تابعی که برای ایجاد چولگی در هسته استفاده می‌شود و این خانواده از توزیع‌ها را به وجود می‌آورد تابع چولگی [۲] نامیده می‌شود. در این فصل به معرفی مختصر این توزیع‌ها پرداخته شده است. این فصل شامل پنج بخش است. در بخش دوم توزیع‌های متقارن چوله در حالت کلی و ویژگی‌های مهم آن‌ها بررسی شده است. حالت چند متغیره‌ی این توزیع در بخش سوم مطرح شده است. در بخش چهار با ذکر یک قضیه نشان داده شده است که می‌توان توزیع‌های متقارن را بر اساس یک حالت خاص از خانواده‌ی متقارن چوله‌ی وابسته به آن‌ها تجزیه و تحلیل کرد. و در بخش آخر مثال‌هایی از این نوع توزیع‌ها همراه با شکل آن‌ها آورده شده است. مطالب این فصل عمدتاً مبتنی بر مراجع [۱۴]، [۳۵]، [۳۶]، [۳۷] و [۳۸] می‌باشد.

---

۱ kernel

۲ skewing function

۲—۲ خانواده‌ی توزیع‌های متقارن چوله

نظریه‌ی توزیع‌ها یکی از شاخه‌های مهم علم آمار است و توزیع‌های احتمالی که نسبت به توزیع نرمال انعطاف‌پذیرتر باشند غالباً در مدل‌سازی آماری مورد نیاز هستند. در دو دهه‌ی اخیر خانواده‌ی توزیع‌های متقارن چوله معرفی شده‌اند که این خانواده شامل کلاس پارامتری انعطاف‌پذیری از توزیع‌ها می‌باشد که دارای چولگی هستند. در این کلاس چولگی را به وسیله‌ی یک پارامتر موسوم به پارامتر شکل یا پارامتر چولگی می‌توان کنترل کرد. دامنه‌ی تغییرات این پارامتر مجموعه‌ی اعداد حقیقی می‌باشد و به ازای هر یک از مقادیر این پارامتر یک توزیع با چولگی متفاوت را می‌توان به دست آورد. در لم زیر که توسط وانگ و همکاران [۳۸] در سال ۲۰۰۴ ارائه شده است شکل کلی این خانواده از توزیع‌ها معرفی شده است.

لم ۱.۲ اگر f یک تابع چگالی احتمال باشد که نسبت به مبدأ متقارن است یعنی، $f(y) = f(-y)$ و π تابعی با دامنه‌ی اعداد حقیقی باشد به طوری که به ازای هر مقدار حقیقی y، $0 \leq \pi(y) \leq 1$ و نیز $\pi(y) + \pi(-y) = 1$، آن‌گاه

$$2f(y)\pi(y), \qquad (۲—۱)$$

به ازای همه‌ی مقادیر حقیقی y یک تابع چگالی احتمال می‌باشد.

اثبات. چون به ازای هر مقدار حقیقی y، توابع $f(y)$ و $\pi(y)$ همواره غیر منفی هستند. برای اثبات تابع چگالی احتمال بودن این رابطه کافیست نشان دهیم که انتگرال (۲—۱) روی مجموعه‌ی اعداد حقیقی برابر یک است. برای این منظور داریم

$$\begin{aligned}\int_{\mathbb{R}} 2f(y)\pi(y)dy &= \int_{-\infty}^{0} 2f(y)\pi(y)dy + \int_{0}^{\infty} 2f(y)\pi(y)dy \\ &= 2\int_{0}^{\infty} \Big(f(y)\pi(-y) + f(y)\pi(y)\Big)dy \\ &= \int_{\mathbb{R}} f(y)dy \\ &= 1.\end{aligned}$$

تساوی دوم با تغییر متغیر y به $-y$ و خاصیت متقارن بودن تابع چگالی احتمال f به دست می‌آید. ■

رابطه‌ی (۲—۱) تابع چگالی احتمال توزیع متقارن چوله می‌باشد. در این رابطه تابع چگالی متقارن f را به اصطلاح هسته و π را تابع چولگی می‌گویند. تعداد بیشماری انتخاب برای تابع π وجود دارد که این امر موجب می‌شود کلاس توزیع‌های متقارن چوله بسیاری از توزیع‌ها را شامل شود و به اندازه‌ی کافی انعطاف‌پذیر باشد. از جمله π را می‌توان تابع توزیع یک متغیر تصادفی پیوسته‌ی متقارن نظیر t، کشی یا لاپلاس در نظر گرفت. مثلاً π می‌تواند F، تابع توزیع هسته‌ی متقارن f، باشد. هم چنین تابع چولگی می

تواند یک تابع ثابت باشد. به عنوان مثال به ازای هر مقدار حقیقی از y، $\pi(y) = \frac{1}{2}$ باشد. یا حتی می‌تواند یک تابع غیریکنوا نظیر $\pi(y) = \frac{1+\sin(y)}{2}$ باشد.

آزالینی و کاپیتانیو[۳] [۴] در سال ۲۰۰۳ تابع چگالی احتمالی به شکل

$$2f(y)G(w(y)), \tag{۲–۲}$$

به ازای همه‌ی مقادیر حقیقی y معرفی کردند به طوری که f یک تابع چگالی احتمال متقارن حول نقطه‌ی صفر، G تابع توزیع به طور مطلق پیوسته‌ای است که G'، مشتق تابع G، حول صفر متقارن می‌باشد و توزیع چولگی نامیده می‌شود هم چنین w یک تابع فرد حقیقی مقدار با دامنه‌ی اعداد حقیقی است. با توجه به لم ۱.۲ رابطه‌ی (۲–۲) یک تابع چگالی احتمال است. همان طور که در گزاره‌ی زیر نشان داده شده است کلاس توزیع‌های به وجود آمده به وسیله‌ی توابع چگالی احتمال (۲–۱) و (۲–۲) یکسان هستند. این گزاره در وانگ و همکاران [۳۸] آمده است.

گزاره **۲.۲** کلاس توزیع‌های متقارن چوله‌ی (۲–۱) مشابه کلاس توصیف شده در رابطه‌ی (۲–۲) می‌باشد.

اثبات. ابتدا نشان می‌دهیم کلاس توابع به شکل $G(w(\cdot))$ مشابه کلاس توابع چولگی است. با توجه به ویژگی‌های تابع توزیع G و تابع فرد w داریم

$$G(w(-s)) = G(-w(s)) = 1 - G(w(s)).$$

با توجه به رابطه‌ی بالا و با توجه به این که همواره $0 \leq G(w(\cdot)) \leq 1$، بنابراین $G(w(\cdot))$ یک تابع چولگی می‌باشد. از طرف دیگر فرض کنید H تابع توزیع تجمعی یک متغیر تصادفی متقارن حول نقطه‌ی صفر باشد که اکیداً صعودی است. برای هر تابع چولگی π می‌توان نوشت $\pi(s) = H(k(s))$ به طوری که $k(s) = H^{-1}(\pi(s))$. با توجه به ویژگی‌های توابع H و π می‌توان نوشت

$$H^{-1}(\pi(-s)) = H^{-1}\Big(1 - \pi(s)\Big) = -H^{-1}(\pi(s)).$$

بنابراین $k(s)$ یک تابع فرد است و این دو کلاس یکسان می‌باشند. ■

در این فصل در باره‌ی یک حالت خاص از توابع چگالی احتمال معرفی شده در رابطه‌ی (۲–۲) به بحث پرداخته شده است که شکل کلی این کلاس از توزیع‌ها در گزاره‌ی ۳.۲ آمده است. این کلاس اولین نوع از توزیع‌های متقارن چوله می‌باشد که آزالینی[۳] در سال ۱۹۸۵ آن را معرفی کرد.

[۳] Capitanio

گزاره **۳.۲** اگر f یک تابع چگالی احتمال متقارن حول صفر، $f(x) = f(-x)$، و G یک تابع توزیع مطلقاً پیوسته باشد به طوری که G' ، مشتق تابع G، حول نقطه‌ی صفر متقارن باشد، آن‌گاه به ازای هر مقدار حقیقی λ و y

$$f_\lambda(y) = ۲f(y)G(\lambda y), \qquad (۲-۳)$$

یک تابع چگالی احتمال است.

در تابع چگالی متقارن چوله‌ی (۲−۳) پارامتر λ انحراف از توزیع متقارن f را کنترل می‌کند در واقع میزان چولگی را نشان می‌دهد به همین دلیل آن را پارامتر شکل یا پارامتر چولگی می‌گویند. چوله به راست یا چپ بودن منحنی تابع چگالی احتمال (۲−۳) به علامت λ بستگی دارد به ازای λ های مثبت منحنی چوله به راست و به ازای λ های منفی چوله به چپ خواهد بود. هم چنین نمودار تابع چگالی احتمال (۲−۳) به ازای مقادیر یکسان λ با علامت‌های مختلف تقارن یکدیگر هستند. نمودار تابع چگالی متقارن چوله به ازای توابع چگالی احتمال متقارن f و مقادیر مختلف λ در بخش ۵ همین فصل آورده شده است.

تابع چگالی (۲−۳) دارای حالت‌های خاص زیر می‌باشد

۱) f_λ به ازای λ برابر صفر همان تابع چگالی احتمال متقارن f است.

بنابراین (۲−۳) تابع چگالی متقارن f را به عنوان عضو مرکزی خود در بر می‌گیرد. یعنی مدلی در اختیار داریم، f_λ، که دارای تغییرات در اطراف مدل متقارن f است.

۲) اگر λ به سمت $\pm\infty$ میل کند، آن‌گاه توزیع متغیر تصادفی Y به توزیع تاخورده‌ی[۴] X یعنی توزیع $\pm|X|$ وقتی X دارای توزیع متقارن f است، میل می‌کند.

بنابراین خانواده‌ی توزیع‌های $\{f_\lambda(y) : \lambda \in \mathbb{R}\}$ شامل مجموعه‌ای از توزیع‌هاست که این توزیع‌ها به ترتیب از توزیع تاخورده‌ی منفی f (توزیع $-|X|$ وقتی که X دارای توزیع متقارن f باشد) به ازای $\lambda \longrightarrow -\infty$ شروع شده تا توزیع متقارن f به ازای λ برابر صفر ادامه یافته و به توزیع تاخورده‌ی f (توزیع $|X|$ وقتی که X دارای توزیع متقارن f باشد) به ازای $\lambda \longrightarrow \infty$ ختم می‌شود.

یکی از ویژگی‌های جالب توزیع‌های متقارن چوله خاصیت تقارن آن‌ها نسبت به مقادیر λ با علامت‌های مختلف است. به این صورت که اگر متغیر تصادفی Y دارای توزیع متقارن چوله با پارامتر λ باشد، آن‌گاه $-Y$ دارای همین توزیع با پارامتر $-\lambda$ می‌باشد. این ویژگی در مسائل مربوط به استنباط آماری مفید است. با توجه به این خاصیت می‌توان در بعضی مسائل از جمله برآورد پارامتر فقط حالت $\lambda \geq ۰$ یا $\lambda \leq ۰$ را در نظر گرفت و نتایج را به کل مجموعه‌ی اعداد حقیقی تعمیم داد. این ویژگی به

[۴] folded distribution

راحتی و با استفاده از تعریف تابع چگالی احتمال متقارن چوله قابل اثبات است.

امید ریاضی توابع زوج و فرد متغیر تصادفی متقارن چوله دارای ویژگی‌هایی است که در اومبچ [۳۵] آمده است در دو قضیه‌ی ۴.۲ و ۶.۲ به این ویژگی‌ها اشاره شده است. در اینجا فرض می‌کنیم Y متغیر تصادفی متقارن چوله با تابع چگالی احتمال (۲−۳)، X یک متغیر تصادفی پیوسته با تابع چگالی احتمال متقارن f و Z برابر قدرمطلق X باشد.

قضیه **۴.۲** اگر h یک تابع زوج باشد $\big(h(x) = h(-x)\big)$ و امید ریاضی قدرمطلق $h(Z)$ متناهی باشد، آن‌گاه همواره رابطه‌ی

$$E(h(Y)) = E(h(X)) = E(h(Z)),$$

برقرار است.

اثبات. چون Z برابر قدرمطلق X است بدیهی است که امید ریاضی توابع $h(X)$ و $h(Z)$ همواره برابر هستند. هم چنین داریم

$$\begin{aligned}
E(h(Y)) &= ۲\int_{\mathbb{R}} h(y)f(y)G(\lambda y)dy \\
&= ۲\int_{-\infty}^{۰} h(y)f(y)G(\lambda y)dy + ۲\int_{۰}^{\infty} h(y)f(y)G(\lambda y)dy \\
&= ۲\int_{۰}^{\infty} h(y)f(y)\Big(۱ - G(\lambda y)\Big)dy + ۲\int_{۰}^{\infty} h(y)f(y)G(\lambda y)dy \\
&= ۲\int_{۰}^{\infty} h(y)f(y)dy \\
&= E(h(Z)).
\end{aligned}$$

تساوی سوم با توجه به تغییر متغیر $y \longrightarrow -y$ و خاصیت تقارن توابع f و g به دست می‌آید. ■

به ویژه برای گشتاورهای زوج متغیر تصادفی متقارن چوله نتیجه‌ی زیر را می‌توان از قضیه‌ی ۴.۲ به دست آورد.

نتیجه **۵.۲** گشتاورهای زوج متغیر تصادفی متقارن‌ـچوله‌ی Y، از پارامتر چولگی λ مستقل و به ازای n های متعلق به اعداد طبیعی همواره رابطه‌ی

$$E(Y^{۲n}) = E(X^{۲n}) = E(Z^{۲n}),$$

برقرار است.

قضیه ۶.۲ اگر h یک تابع فرد باشد $\big(h(x) = -h(-x)\big)$ و به ازای تمام مقادیر مثبت x، $h(x)$ همواره غیر منفی باشد و نیز امید ریاضی قدر مطلق $h(Z)$ متناهی باشد ، آن‌گاه رابطه‌ی

$$\circ = E(h(X)) \leq E(h(Y)) \leq E(h(Z)),$$

به ازای تمام مقادیر غیر منفی λ برقرار است.

اثبات. چون X یک متغیر تصادفی متقارن است در نتیجه بدیهی است که همواره امید ریاضی $h(X)$ برابر با صفر است. حال داریم

$$\begin{aligned} E(h(Y)) &= ۲\int_{-\infty}^{\circ} h(y)f(y)G(\lambda y)dy + ۴\int_{\circ}^{\infty} h(y)f(y)G(\lambda y)dy \\ &= -۲\int_{\circ}^{\infty} h(y)f(y)\Big(۱ - G(\lambda y)\Big)dy + ۲\int_{\circ}^{\infty} h(y)f(y)G(\lambda y)dy \\ &= -۲\int_{\circ}^{\infty} h(y)f(y)dy + ۴\int_{\circ}^{\infty} h(y)f(y)G(\lambda y)dy \qquad (۲-۴) \\ &= -E(h(Z)) + ۴\int_{\circ}^{\infty} h(y)f(y)G(\lambda y)dy \\ &\leq -E(h(Z)) + ۴\int_{\circ}^{\infty} h(y)f(y)dy \\ &= -E(h(Z)) + ۲E(h(Z)) \\ &= E(h(Z)). \end{aligned}$$

هم چنین با توجه به رابطه‌ی (۲-۴) داریم

$$E(h(Y)) = ۴\int_{\circ}^{\infty} h(y)f(y)\Big(G(\lambda y) - \frac{۱}{۲}\Big)dy.$$

چون به ازای مقادیر مثبت y همواره $G(\lambda y) - \frac{۱}{۲}$ غیر منفی است. بنابراین خواهیم داشت $E(h(Y)) \geq \circ = E(h(X))$. ■

به ویژه برای گشتاورهای فرد متغیر تصادفی متقارن چوله نتیجه‌ی زیر را می‌توان از قضیه‌ی ۶.۲ به دست آورد.

نتیجه ۷.۲ به ازای n های متعلق به اعداد طبیعی رابطه‌ی

$$\circ = E(X^{۲n-۱}) \leq E(Y^{۲n-۱}) \leq E(Z^{۲n-۱}),$$

برای متغیرهای تصادفی X، Y و Z همواره برقرار است.

بر اساس قضیه‌های ۴.۲ و ۶.۲ نتیجه‌ی زیر را خواهیم داشت.

نتیجه ۸.۲ اگر X، Y و Z متغیرهای تصادفی تعریف شده در قضیه‌ی قبل باشند، آن‌گاه رابطه‌ی زیر برای واریانس آن‌ها برقرار است

$$Var(X) \geq Var(Y) \geq Var(Z).$$

با استفاده از خاصیت تقارن توزیع متقارن چوله به ازای مقادیر یکسان λ با علامت‌های مختلف رابطه‌ی زیر برای گشتاورهای فرد به ازای λ های منفی و n های متعلق به اعداد طبیعی برقرار است

$$-E(Z^{2n-1}) \leq E(Y^{2n-1}) \leq E(X^{2n-1}) = 0.$$

ممکن است انتظار داشته باشیم چنین نتایجی برای چولگی نیز برقرار باشد. با توجه به مثال‌هایی که در اومبچ[۳۵] آمده است نمی‌توان یک حالت خاص را برای روابط بین چولگی متغیرهای تصادفی X، Y و Z به دست آورد. در مثال ۹.۲ حالتی را داریم که توزیع متغیر تصادفی Z چوله نیست ولی متغیر تصادفی Y دارای چولگی منفی است. در مثال ۱۰.۲، Z دارای چولگی مثبت است ولی چولگی Y منفی است. در مثال ۱۱.۲ چولگی Z منفی است، اما چولگی متغیر تصادفی Y با توجه به انتخاب تابع توزیع G می‌تواند مثبت یا منفی باشد. در اینجا برای محاسبه‌ی چولگی از گشتاور مرتبه‌ی سوم استفاده شده است.

مثال ۹.۲ فرض کنید متغیر تصادفی X دارای تابع چگالی احتمال $f(x) = 3|x|(1-|x|)I_{(-1,1)}(x)$ باشد در این صورت تابع چگالی احتمال متغیر تصادفی Z برابر $g(z) = 6z(1-z)I_{(0,1)}(z)$ خواهد بود. با توجه به این‌که X و Z هر دو متقارن هستند چولگی آن‌ها برابر صفر است. اگر تابع توزیع G را به صورت $G(y) = \frac{1}{2}(1+y)I_{[-1,1)}(y) + I_{[1,\infty)}(y)$ در نظر بگیریم، آن‌گاه امید ریاضی و گشتاور مرتبه‌ی سوم متغیر تصادفی Y را به صورت زیر خواهیم داشت

$$\begin{aligned} E(Y) &= \int_{-1}^{1} y f_\lambda(y)dy = 0.3, \\ E[(Y-0.3)^3] &= \int_{-1}^{1} (y-0.3)^3 f_\lambda(y)dy = -0.073. \end{aligned}$$

مثال ۱۰.۲ تابع چگالی احتمال متغیر تصادفی X را به صورت $f(x) = \frac{1}{10}(7-4|x|)I_{(-1,1)}(x)$ در نظر می‌گیریم در این صورت تابع چگالی Z به صورت $g(z) = \frac{1}{5}(7-4z)I_{(0,1)}(z)$ خواهد بود بنابراین داریم

$$E(Z) = 0.033, \qquad E[(Z-0.033)^3] = 0.006.$$

اگر G را تابع توزیع معرفی شده در مثال ۹.۲ در نظر بگیریم، خواهیم داشت

$$E(Y) = 0.267, \qquad E[(Y-0.267)^3] = -0.029.$$