دانشگاه سیستان و بلوچستان

پایان نامۀ کارشناسی ارشد آمار

عنوان

برآوردهای پارامتر شکل توزیع گاوسی تعمیم یافته

استاد راهنما

دکتر حسن زارعی

تحقیق و نگارش

محمد فضایلی

پاییز ۸۸

چکیده

توزیع گاوسی تعمیم یافته [1] به دلیل کاربرد در آمار و علوم مهندسی به ویژه در الگوسازی توابع چگالی احتمال سیگنال‌ها، مورد توجه بسیاری از محققان بوده است. به عنوان مثال، منحنی ضرایب موجی با حالت خاصی از تابع توزیع گاوسی تعمیم یافته الگوسازی می‌شود که در آن پارامترها توسط گشتاورهای دوم و چهارم داده‌ها برآورد می‌شوند. همچنین **مولر**، بهترین برازش را برای توزیع ضرایب تبدیل کسینوسی گسسته [2] به‌دست آورد و **گازور** و **ژانگ** [۷] با کمک توزیع گاوسی تعمیم یافته، توزیع نمونه های گفتاری [3] را تخمین زدند.

در این پایان نامه، توزیع گاوسی تعمیم یافته را معرفی کرده و توزیع‌های لاپلاس، نرمال، $\alpha-$پایدار [4] و گاما را به‌عنوان حالت‌های خاص آن مطالعه می‌کنیم. علاوه براین توزیع‌های زیر گاوسی و سوپر گاوسی که به‌ازای مقادیر مختلف پارامتر شکل حاصل می‌شود، را بررسی می‌نماییم.

به منظور برآورد پارامتر شکل توزیع با استفاده از روش‌های کلاسیک، تابع مولد گشتاور توزیع محاسبه می‌شود. پس از تعیین گشتاورها، برآوردکننده‌های حاصل از روش گشتاوری، درستنمایی ماکسیمم، ماکسیمم آنتروپی و نسبت کشیدگی برای پارامتر شکل تعیین می‌گردد. چون روش‌های درستنمایی ماکسیمم و ماکسیمم آنتروپی به‌طور تحلیلی قابل حل نمی‌باشند، بنابراین از روش‌های عددی و شبیه‌سازی، برآوردهای درستنمایی ماکسیمم و ماکسیمم آنتروپی را به‌دست می‌آوریم.

---

[1] Generalized Gaussian Distribution(GGD)

[2] Discrete Cosine Transform(DCT)

[3] Speech

[4] $\alpha - stable$

فهرست مندرجات

۱ تعاریف مقدماتی ۴

۱ـ۱ گشتاورها ۷

۱ـ۲ برآوردگر گشتاوری ۸

۱ـ۳ برآوردگر درستنمایی ماکسیمم ۱۰

۱ـ۴ کشیدگی ۱۱

۱ـ۵ آنتروپی ۱۵

۱ـ۶ توزیع پایدار ۱۷

۱ـ۷ توابع بسل ۱۹

۱—۸ **توزیع دیریکله** ۲۱

۱—۹ **توزیع لوی** ۲۲

۲ تابع چگالی احتمال توزیع گاوسی تعمیم یافته ۲۴

۲—۱ **حالت‌های خاص** ۲۶

۲—۱—۱ **توزیع لاپلاس** ۲۶

۲—۱—۲ **توزیع نرمال** ۲۷

۲—۱—۳ α — **پایدار** ۲۷

۲—۱—۴ **توزیع گاوسی معکوس تعمیم یافته** ۲۸

۲—۱—۵ **توزیع گاوسی تعمیم یافته مالات** ۲۹

۲—۲ **مشخصه های توزیع** ۳۰

۲—۲—۱ **گشتاورهای توزیع** ۳۰

۲—۲—۲ **تابع مولد گشتاور توزیع** ۳۳

۲—۲—۳ **نسبت مالات توزیع** ۳۵

۲—۲—۴ **نسبت کشیدگی توزیع** ۳۵

۲—۲—۵ **آنتروپی توزیع** ۳۹

۳ برآوردگرهای پارامتر شکل توزیع گاوسی تعمیم یافته ۴۳

۳–۱ برآورد گشتاوری پارامتر شکل توزیع گاوسی تعمیم یافته ۴۴

۳–۲ برآورد پارامتر شکل توزیع گاوسی تعمیم یافته با استفاده از نسبت مالات ۵۰

۳–۳ برآورد پارامتر شکل توزیع گاوسی تعمیم یافته با استفاده از نسبت کشیدگی ۵۱

۳–۴ برآورد پارامتر شکل توزیع گاوسی تعمیم یافته با استفاده از آنتروپی ۵۲

۳–۵ برآورد به روش درستنمایی ماکسیمم ۵۶

۴ برآورد وزنی تصادفی پارامترها در توزیع گاوسی تعمیم یافته ۶۱

۴–۱ تعاریف اولیه ۶۲

۴–۲ برآورد وزنی تصادفی پارامترهای توزیع گاوسی تعمیم یافته ۶۵

۴–۲–۱ گشتاور قدر مطلق مرتبه K برای پارامترهای توزیع گاوسی تعمیم یافته . . . ۶۵

۴–۲–۲ برآورد وزنی تصادفی پارامترهای α و β ۶۶

۴–۳ تحلیل همگرایی ۶۷

A واژه نامه ۷۷

B مراجع ۸۴

فصل ۱

تعاریف مقدماتی

مقدمه

توزیع گاوسی تعمیم یافته به‌طور بسیار گسترده‌ای در الگوسازی تابع چگالی احتمال [1] سیگنال‌ها به‌کار برده می‌شود. همچنین کاربرد آن‌ها در ضرایب تبدیل، از قبیل تبدیل کسینوسی گسسته [2] یا ضرایب موجی بسیار رایج است.

مولر[٦] آماره کای‌ـدو ضرایب تبدیل کسینوسی گسستهٔ جریان متناوب [3] را برای دو توزیع گاوسی تعمیم یافته و توزیع لاپلاس [4] از روی منحنی نرمال مقایسه کرد و بیان نمود که به‌ازای تمامی منحنی‌ها، توزیع گاوسی تعمیم یافته بهترین برازش را برای ضرایب تبدیل کسینوسی گسسته ایجاد می‌کند. وی پارامتر شکل توزیع گاوسی تعمیم یافته را به کمک روش درستنمایی ماکسیمم[5] برآورد نمود. لازم به ذکر است که فرمول های برآورد به روش درستنمایی ماکسیمم، توسط **دی — یو**[٨] طرح شده بود.

عیازی [٣] زیر گروهی از منحنی‌های لناو بابون را به کمک توزیع گاوسی تعمیم یافته الگوسازی نمود.

طنابی و **فروردین**[١١] زیر گروه منحنی‌ها و ضرایب تبدیل کسینوسی گسسته را با استفاده از توزیع گاوسی تعمیم یافته الگوسازی کردند.

هیستوگرام جزئیات یک تصویر را **مالات**[١٢] به کمک توزیع گاوسی تعمیم یافته الگوسازی کرد و برآورد پارامتر شکل توزیع گاوسی تعمیم یافته را بر مبنای گشتاورهای اول و دوم، بنا نهاد. این برآوردگر توسط **برنی**[١٣] مورد استفاده قرار گرفت.

جوشی و **فیشر**[١۴] برآوردگرهای پارامتر شکل توزیع گاوسی تعمیم یافته بیان شده توسط مالات و دی ـ یو را مقایسه کردند. آن‌ها اظهار نمودند که‌روش ارائه شده توسط مالات از نظر محاسباتی پیچیدگی بسیار کمتری دارد.

---

[1] Probability Density Function (PDF)

[2] Descrete Cosine Transform(DCT)

[3] $ACDCT$

[4] Laplasian Distribution

[5] Maximum Liklihood(ML)

کلارک[١٥] نشان داد که ضرایب تبدیل کسینوسی گسسته، تبدیل والش ـهادامارد[٦] و تبدیل سینوسی گسسته[٧] می‌تواند توسط توزیع گاوسی تعمیم یافته الگوسازی شود.

در این پایان نامه سعی شده است که تابع توزیع گاوسی تعمیم یافته، معرفی و ارتباط آن با برخی توزیع‌های آماری نشان داده شود و همچنین پارامترهای موجود در توزیع، به روش‌های متفاوت برآورد گردد. از جمله این روش ها روش گشتاوری، آنتروپی، مالات، درستنمایی ماکسیمم، و نسبت کشیدگی است.

در پایان، برآورد پارامتر شکل توزیع گاوسی تعمیم یافته را به روش عددی و با توجه به برآوردگرهای گفته شده بدست می آوریم (آن قسمت از مطالبی که در بین علائم ♣ تا ♠ قرار دارند نتایجی است که توسط اینجانب بدست آمده است).

[٦] Walsh-Hadamard Transform(WHT)

[٧] discrete sine transform(DST)

۱—۱ گشتاورها

گشتاورهای متغیر تصادفی یک توزیع، امید ریاضی توان‌های مختلف آن متغیر تصادفی هستند که دارای توزیع مشخصی می‌باشند.

تعریف ۱.۱.۱: اگر X یک متغیر تصادفی باشد، گشتاور مرتبه r اُم آن که معمولاً با μ_r نشان داده می‌شود، به‌صورت زیر تعریف می‌شود:

$$\mu_r = E[X^r] ,$$

به شرط این که امید ریاضی وجود داشته باشد.

تعریف ۲.۱.۱: اگر X یک متغیر تصادفی باشد، گشتاور مرتبه r اُم X حول a به‌صورت $E[(X-a)^r]$ تعریف می‌شود. اگر $a = \mu_x$ باشد، گشتاور مرکزی مرتبه r اُم X حول μ_x که با μ'_r نشان داده می‌شود، برابر است با:

$$\mu'_r = E[(X-\mu_x)^r] .$$

توجه کنید که، $\mu'_۱ = E[(X-\mu_x)] = ۰$ و $\mu'_۲ = E[(x-\mu_x)^۲]$ واریانس X است. هم‌چنین توجه کنید اگر تابع چگالی احتمال X حول μ_x متقارن باشد، همهٔ گشتاورهای فرد X حول μ_x صفر است، مشروط بر این که گشتاورها وجود داشته باشند.

در حالی که یک گشتاور ویژه یا چند گشتاور می‌توانند اطلاعات اندکی درباره توزیع فراهم کنند، مجموعه کامل گشتاورها یا $\mu_۱, \mu_۲, \mu_۳, \ldots$ ، معمولاً توزیع X را دقیقاً تعیین می‌کند. در آمار کاربردی، گشتاورهای مرتبه اول، دوم، سوم و چهارم از اهمیت زیادی برخوردارند، اما گشتاورهای مرتبه پنجم و بالاتر کمتر مورد استفاده قرار می‌گیرند.

اکنون تابعی را معرفی می‌کنیم که نمایشی از تمام گشتاورهاست. چنین تابعی، تابع مولّد گشتاور نامیده می‌شود.

تعریف ۳.۱.۱: اگر X یک متغیر تصادفی با تابع چگالی احتمال $f(x)$ باشد و $h > ۰$ وجود داشته باشد

به‌طوری که $\forall t \in (-h,h)$، $E(e^{tx})$ وجود داشته باشد، آن‌گاه تابع مولد گشتاور X که با نماد $M_X(t)$ نمایش داده می‌شود برابر است با:

$$M_X(t) = E(e^{tX}) = \begin{cases} \int_{-\infty}^{\infty} e^{tx} f_X(x)dx & ; \quad \text{if } X \text{ is continuous} \\ \sum_x e^{tx} f_X(x) & ; \quad \text{if } X \text{ is discrete} \end{cases}$$

باید توجه داشته باشیم که تابع مولد گشتاور بر حسب یک تابع چگالی تعریف می‌شود.

اگر از تابع مولد گشتاور، r مرتبه نسبت به t مشتق بگیریم داریم:

$$\frac{\partial^r}{\partial t^r} M_X(t) = \int_{-\infty}^{\infty} x^r e^{tx} f_X(x)dx \; .$$

و چنانچه $t \to 0$ ، داریم:

$$\frac{\partial^r}{\partial t^r} M_X(t)|_{t=0} = E[X^r] = \mu_r \; .$$

بنابراین گشتاورهای یک توزیع را می‌توان با مشتق‌گیری از تابع مولد گشتاور بدست آورد که این نام گذاری از اینجا نتیجه می‌شود.

۱—۲ برآوردگر گشتاوری

یکی از قدیمی‌ترین روش‌های برآوردیابی، روش برآورد گشتاوری است که در سال ۱۸۹۴ توسط آماردان مشهور **کارل پیرسن**[۸] معرفی شده است. این روش عبارت است از دستورالعملی برای به‌دست آوردن برآوردگری به نام «برآوردگر گشتاوری» که آن را به اختصار با MME نشان می‌دهیم.

برای روشن شدن موضوع، فرض کنید $X_1, X_2, \ldots, X_n$ یک نمونه تصادفی n تایی از توزیع $F_\theta(X)$ باشد به‌طوری‌که $\underline{\theta} = (\theta_1, \ldots, \theta_k) \in \Theta \subseteq \Re^k$. هم‌چنین فرض کنید k گشتاور اول این توزیع، که به‌صورت توابعی از θ هستند، وجود داشته باشند. می‌دانیم که گشتاور r اُم توزیع، در صورت وجود، به‌صورت زیر تعریف

[۸] Carl Pearson

می‌شود:

$$\mu_r = \mu_r(\theta) = E_\theta(X^r) \quad ; \quad r=۱,\ldots,k \; .$$

اگر M_r نمایانگر r اُمین گشتاور نمونه‌ای بر پایه نمونه تصادفی داده شده باشد که به‌صورت زیر تعریف می‌شود:

$$M_r = \frac{۱}{n}\sum_{i=۱}^{n} {X_i}^r \quad ; \quad r=۱,\ldots,k \; .$$

آن‌گاه برآورد گشتاوری پارامترهای مجهول $\theta_۱,\ldots,\theta_k$ ، بر اساس یک ایده ساده و از تشکیل و حل k معادله زیر حاصل خواهد شد:

$$\mu_r = M_r \quad ; \quad r=۱,\ldots,k \; .$$

تبصره ۱.۲.۱: برآورد گشتاوری روش منحصر به فردی ندارد و روش ارائه شده برای حل معادلات فوق مبتنی بر استفاده از k گشتاور اول است که می‌تواند با استفاده از سایر گشتاورهای توزیع یا گشتاورهای مرکزی نیز بدست آید.

تبصره ۲.۲.۱: اگر علاقه‌مند به برآورد توابعی از θ ، مثلاً $\gamma_۱(\theta),\ldots,\gamma_k(\theta)$، بر اساس روش گشتاوری باشیم، می‌توانیم به روش‌های مختلفی این کار را انجام دهیم. یک روش، می‌تواند ابتدا با محاسبه $\tilde{\theta}_۱,\ldots,\tilde{\theta}_k$ و سپس انتخاب $\gamma_r(\tilde{\theta}_۱,\ldots,\tilde{\theta}_k)$ به عنوان برآورد گشتاوری $\gamma_r(\theta_۱,\ldots,\theta_k)$ که $r=۱,\ldots,k$ بدست آید. روش دیگر می‌تواند از حل معادلات زیر حاصل شود:

$$\mu_r(\gamma_۱,\ldots,\gamma_k) = M_r \quad ; \quad r=۱,\ldots,k \; .$$

که $\mu_r(\gamma_۱,\ldots,\gamma_k)$ ، r اُمین گشتاور بر حسب $\gamma_۱(\theta),\ldots,\gamma_k(\theta)$ است.

۱–۳ برآوردگر درستنمایی ماکسیمم

اکنون به بررسی روش برآوردی که غالباً مورد استفاده قرار می‌گیرد، یعنی روش «درستنمایی ماکسیمم (ML)[۹]» می‌پردازیم. این روش یکی از قدیمی‌ترین و پراهمیت ترین روش‌ها در نظریه برآوردهاست.

روش درستنمایی ماکزیمم عبارت است از دستورالعملی برای به‌دست آوردن برآوردگری به‌نام « برآوردگر درستنمایی ماکزیمم » که از آن به اختصار MLE[۱۰] یاد خواهیم کرد و مبتنی بر یک تابع آماری مهم به نام « تابع درستنمایی » است. فرض کنید $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ ، برداری از n متغیر تصادفی با تابع چگالی احتمال توأم $f_\theta(x)$ که $\theta \in \Theta \subseteq \Re^k$ باشد.

تعریف ۱.۳.۱: برای هر مقدار داده شده $\mathbf{X} = \mathbf{x}$ ، **تابع درستنمایی** $\mathbf{X}$ را تابع چگالی احتمال توأم $X_1, X_2, \ldots, X_n$، یعنی $f_\theta(\mathbf{x})$ تعریف می‌کنیم که به‌صورت تابعی از θ در نظر گرفته می‌شود و آن را با نماد $L(\theta)$ نمایش می‌دهیم:

$$
\begin{aligned}
L(\theta) &= f_\theta(x) \\
&= L(\theta|x) \ .
\end{aligned}
$$

تبصره ۱.۳.۱: تابع درستنمایی یا $L(\theta)$ لزوماً نسبت به θ مشتق‌پذیر نیست.

تبصره ۲.۳.۱: تابع درستنمایی یا $L(\theta)$ بر حسب θ یک تابع چگالی احتمال نیست.

تبصره ۳.۳.۱: اگر $X_1, X_2, \ldots, X_n$ یک نمونه تصادفی از خانواده چگالی‌های $\{f_\theta(x); \theta \in \Theta\}$ باشد آن‌گاه:

$$
L(\theta) = \prod_{i=1}^{n} f_\theta(x_i) \ .
$$

---

[۹] Maximum Likelihood

[۱۰] Maximum Likelihood Estimator

تعریف ۲.۳.۱: اگر $\delta(X)$ برآوردگری برای θ باشد طوری که:

$$۱)\quad P_\theta(\delta(X) \in \Theta) = ۱ \qquad ; \qquad \forall \theta \in \Theta \,.$$

$$۲)\quad L(\delta(x)) \geq L(\theta) \qquad ; \qquad \forall \theta \in \Theta \,.$$

آن‌گاه $\delta(X)$ به عنوان **برآوردگر درستنمایی ماکسیمم** θ تعریف می‌شود.

معمولاً برآوردگر درستنمایی ماکزیمم θ را با $\hat{\theta}$ نمایش می‌دهند و لزومی ندارد که منحصر به فرد باشد.
هم‌چنین بر اساس تعریف $\hat{\theta}$ داریم:

$$L(\hat{\theta}) = sup_{\theta \in \Theta} L(\theta) \,.$$

یا اگر $l(\theta) = \ln L(\theta)$ تعریف شود، آن‌گاه داریم:

$$l(\hat{\theta}) = sup_{\theta \in \Theta} l(\theta) \,.$$

۱ـ۴ کشیدگی

در ابتدا به بررسی حالت کلی کومولانت[۱۱] می‌پردازیم. می‌دانیم تابع مشخصه X یعنی $\Psi(t)$ به‌صورت زیر حاصل می‌شود:

$$\Psi(t) \;=\; E(e^{itX}) \;=\; \int_{-\infty}^{\infty} e^{itx} f_X(x)\, dx \,.$$

که $i = \sqrt{-۱}$ است. هر توزیع احتمال، به طور یکتا از تابع مشخصه‌اش به‌دست می‌آید. حال با بسط دادن تابع مشخصه $\Psi(t)$ به سری تیلور آن داریم:

$$\Psi(t) \;=\; \int_{-\infty}^{\infty} \Big(\sum_{k=۰}^{\infty} \frac{x^k\, (it)^k}{k!}\Big)\, f_X(x)\, dx$$

---

[۱۱] Cumulant

$$= \sum_{k=0}^{\infty} E(X^k)\frac{(it)^k}{k!} \ .$$

بنابراین جمله ضرایب بسط فوق، گشتاور X ($E(X^k)$) است. (البته با این فرض که این گشتاورها موجود باشند). به این دلیل گاهی تابع مشخصه یا $\Psi(t)$ را تابع مولد گشتاور نیز می‌نامند.

اغلب از تابع مشخصه دوم X یعنی $\phi(t)$ یا تابع مولد کومولانت استفاده می‌شود.

تعریف ۱.۴.۱: تابع کومولانت با گرفتن لگاریتم طبیعی از تابع مشخصه اول به‌صورت زیر حاصل می‌شود:

$$\phi(t) = \ln(\Psi(t)) = \ln[E(e^{itX})] \ .$$

کومولانت K_k به روش مشابهی از گشتاورها، تحت ضرایب بسط سری تیلور از تابع مشخصه دوم، تعریف می‌شود:

$$\phi(t) = \sum_{k=0}^{\infty} K_k \frac{it^k}{k!} \ .$$

که k امین کومولانت با مشتق‌گیری به‌صورت زیر حاصل می‌شود:

$$K_k = (-i)^k \ \frac{d^k\phi(t)}{dt^k}|_{t=0} \ .$$

برای متغیر تصادفی X چهار کومولانت اول به‌صورت زیر هستند:

$$K_1 = E(X) \ .$$

$$K_2 = E(X^2) - [E(X)]^2 \ .$$

$$K_3 = E(X^3) - 3E(X^2)E(X) + 2(E(X))^3 \ .$$

$$K_4 = E(X^4) - 3[E(X^2)]^2 - 4E(X^3)E(X) + 12E(X^2)(E(X))^2 - 6(E(X))^4 \ .$$

این فرمول‌ها بعد از به‌کارگیری تابع مشخصه دوم به‌صورت بسیار دشواری حاصل می‌شود. بنابراین عبارات مراتب بالاتر کومولانت بسیار پیچیده است که در این پایان نامه مورد بررسی قرار نمی‌گیرد.

در حالت خاصی که میانگین متغیر تصادفی X صفر است، چهار کومولانت اول به‌صورت زیر بدست می‌آیند:

$$K_۱ = ۰ .$$

$$K_۲ = E(X^۲) .$$

$$K_۳ = E(X^۳) .$$

$$K_۴ = E(X^۴) - ۳(E(X^۲))^۲ .$$

بنابراین سه کومولانت اول، مرتبط با گشتاورها هستند و چهارمین کومولانت یا $K_۴$ تحت عنوان کشیدگی مورد بررسی قرار می‌گیرد.

اغلب به‌جای گشتاور مرکزی مرتبه چهارم یعنی $\mu'_۴ = E[(X - E(X))^۴]$، آماره مرتبه چهارم که کشیدگی نام دارد استفاده می‌شود. چون این آماره برخی خواص مفید دارد که گشتاور مرکزی مرتبه چهارم از آن بی‌بهره است. این آماره بخصوص در تحلیل مولفه مستقل[۱۲] و تفکیک منبع نور[۱۳] در مهندسی کاربرد مهمی دارد.

کشیدگی نرمال شده به‌صورت زیر محاسبه می‌شود:

$$\tilde{K}(X) = \frac{E(X^۴)}{(E(X^۲))^۲} - ۳ .$$

یک خاصیت مهم کشیدگی، خاصیت جمع‌پذیری آن است.

لم ۱.۴.۱: الف) اگر X و Y دو متغیر تصادفی مستقل با میانگین صفر باشند، داریم:

$$kurt(X + Y) = kurt(X) + kurt(Y) .$$

ب) به ازای هر اسکالر β داریم:

$$kurt(\beta X) = \beta^۴ kurt(X) .$$

اثبات: الف)

$$\clubsuit \ kurt(X + Y) = E[(X + Y)^۴] - ۳\{E[(X + Y)^۲]\}^۲$$

---

[۱۲] Independent Component Analysis(ICA)

[۱۳] Blind Source Separation

$$= E[X^4 + 4X^3Y + 6X^2Y^2 + 4XY^3 + Y^4] - 3[E(X^2 + 2XY + Y^2)]^2$$

$$= E(X^4) + 4E(X^3Y) + 6E(X^2Y^2) + 4E(XY^3) + E(Y^4)$$

$$-3[E^2(X^2) + E^2(Y^2) + 4E^2(XY) + 4E(X^2)E(XY)$$

$$+4E(Y^2)E(XY) + 2E(X^2)E(Y^2)]$$

$$= E(X^4) + 4E(X^3)E(Y) + 6E(X^2)E(Y^2) + 4E(X)E(Y^3) + E(Y^4)$$

$$-3E^2(X^2) - 3E^2(Y^2) - 12E^2(X)E^2(Y)$$

$$-12E(X^2)E(X)E(Y) - 12E(Y^2)E(X)E(Y) - 6E(X^2)E(Y^2)$$

حال با توجه به استقلال بین متغیرهای X و Y و هم‌چنین میانگین صفر متغیرها داریم:

$$kurt(X+Y) = E(X^4) - 3[E(X^2]^2 + E(Y^4) - 3[E(Y^2)]^2$$

$$= kurt(X) + kurt(Y)\ . \qquad \square$$

دقت کنید که این خاصیت برای گشتاور چهارم برقرار نیست که این یکی از برتری‌های کومولانت نسبت به گشتاورهاست.

ب)

$$kurt(\beta X) = E[(\beta X)^4] - 3\{E[(\beta X)^2]\}^2$$

$$= E(\beta^4 X^4) - 3\{E[(\beta^2 X^2)]\}^2$$

$$= \beta^4 E(X^4) - 3[\beta^2 E(X^2)]^2$$

$$= \beta^4 E(X^4) - 3\beta^4 [E(X^2)]^2$$

$$= \beta^{4}\{E(X^{4}) - 3[E(X^{2})]^{2}\} . \qquad \square \qquad \spadesuit$$

به این جهت کشیدگی خطی نیست ([۱۸]).

۱—۵ آنتروپی

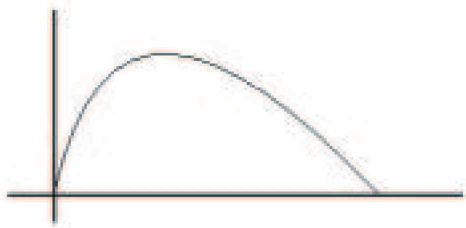
آنتروپی [۱۴] فرض اساسی تئوری اطلاع است. آنتروپی برای متغیر تصادفی گسسته X به‌صورت زیر تعریف می‌شود:

$$H(X) = -\sum_{i} P(X = a_i) \log P(X = a_i) \ .$$

که a_i ، مقادیر ممکن X است. بر اساس این‌که مبنای لگاریتم چه عددی باشد، واحدهای متفاوت آنتروپی حاصل می‌شود. معمولاً لگاریتم در مبنای ۲ مورد استفاده قرار می‌گیرد، که در این حالت هر واحد یک بیت نامیده می‌شود. به ازای $P = 0$ و $P = 1$ آنتروپی ($H(X)$) برابر با صفر است و برای سایر مقادیر بزرگتر از صفر می‌باشد. بر اساس شکل (۱—۵—۱) وقتی P نزدیک به صفر یا یک باشد آنتروپی کوچک و زمانی که بین این دو عدد است آنتروپی بزرگ می‌شود. اگر $H(P) = -\sum P \log P$ آن‌گاه نمودار $H(P)$ به‌صورت شکل (۱—۵—۱) است که در آن P احتمال است.

---

[۱۴] Entropy



شکل (۱—۵—۱): نمودار آنتروپی

آنتروپی برای حالتی که X متغیر تصادفی پیوسته باشد نیز قابل تعمیم است که در این حالت به آن آنتروپی مشتق گویند و به‌صورت زیر تعریف می‌شود:

$$H(X) = -\int_{-\infty}^{\infty} f_X(x) \log f_X(x) dx \ .$$

اگر متغیر تصادفی روی برخی بازه‌های کوچک متمرکز شده باشد، آنتروپی آن کوچک می‌شود. در حالت آنتروپی پیوسته، این عدد می‌تواند منفی هم شود، ولی در حالت گسسته این عدد منفی نمی‌شود، چون احتمال در بازه $[\circ, ۱]$ می‌باشد. ولی چگالی‌های احتمال پیوسته می‌توانند بزرگتر از یک شوند که در این حالت آنتروپی مقادیر منفی را در بر می‌گیرد.

اکنون به راحتی می‌توان مشاهده نمود کدام نوع از متغیرهای تصادفی، آنتروپی‌های کوچک می‌گیرند واضح است که چگالی‌هایی که احتمال آن‌ها مقادیر بزرگ و نزدیک‌تر به یک را می‌گیرند، چون سهم منفی زیادی را از انتگرال دارند آنتروپی‌های کوچک می‌گیرند.

مثال ۱.۵.۱ : توزیع یکنواخت X را در بازه $[\circ, a]$ در نظر بگیرید

$$f_X(x) = \begin{cases} ۱/a & \circ \leq x \leq a \\ \circ & o.w \end{cases}$$

آنتروپی آن به‌صورت زیر به‌دست می‌آید:

$$H(X) = -\int_{\circ}^{a} \frac{۱}{a} \log \frac{۱}{a} dx = \log a \ .$$

پس می‌بینیم که آنتروپی بزرگ می‌شود، اگر a بزرگ شود و بالعکس کوچک می‌شود، اگر a کوچک شود. در حالت حدی وقتی a به سمت صفر میل می‌کند، آنتروپی به $-\infty$ میل می‌کند ([۱۸]).

۱—۶ توزیع پایدار

در تئوری احتمال، گفته می‌شود یک متغیر تصادفی پایدار است (یا دارای توزیع پایداری است) اگر این خاصیت را داشته باشد که ترکیب خطی از دو جفت مستقل از متغیرها، نه تنها دارای توزیع یکسانی باشند حتی پارامترهای مکان و مقیاس یکسانی نیز داشته باشند.
توزیع نرمال یک خانواده از توزیع‌های پایدار می‌باشد. بر اساس قضیه حد مرکزی، مجموع نرم‌دارِ مجموعه‌ای از متغیرهای تصادفی با واریانس‌های متناهی زمانی که تعداد متغیرها افزایش می‌یابد به توزیع نرمال گرایش خواهد داشت. بدون فرض واریانس متناهی، ممکن است توزیع حدی، یک توزیع پایدار باشد. توزیع‌های پایداری که نرمال نیستند، اغلب توزیع‌های **پارِتویی پایدار**[۱۵] **نامیده می‌شود که برگرفته از نام‌ویلفردو پارِتو**[۱۶] **است. یعنی اگر X_1 و X_2 زوج‌های مستقل از متغیر تصادفی پایدار X باشند آن‌گاه به ازای مقادیر ثابت a و b، متغیر تصادفی $aX_1 + bX_2$ دارای توزیع یکسانی با متغیر تصادفی $cX + d$ (c و d مقادیر ثابتند) خواهد بود. توزیع اکیداً پایدار نامیده می‌شود هرگاه $d = 0$ باشد.**
چون توزیع‌های نرمال، کوشی و لوی[۱۷]، هر سه خصوصیت بالا را دارا می‌باشند، بنابراین می‌توانیم چنین نتیجه بگیریم که توزیع‌های فوق موارد خاصی از توزیع‌های پایدار می‌باشند. چنین توزیع‌هایی یک خانواده چهار پارامتری از توزیع‌های احتمالی پیوسته که پارامتربندی شده‌اند، تشکیل می‌دهند که این چهار پارامتر عبارتند از پارامترهای مکان و مقیاس μ و c و دو پارامتر شکلِ α و β که به ترتیب متناظر با میزان تقارن و میزان تمرکز توزیع می‌باشند.

---
[۱۵] Stable Paretian distributions
[۱۶] Vilfredo Pareto
[۱۷] Lévy

توزیع‌های پایدار سوپر گاوسی[۱۸] بوده و دارای دم‌های پهن[۱۹] می‌باشند. حال به تعریف دیگری از پایداری می‌پردازیم:

متغیر تصادفی X پایدار است اگر برای n مؤلفه مستقل X_i از X ثابتی مانند d وجود داشته باشد به طوری که

$$X_۱ + X_۲ + \ldots + X_n \stackrel{d}{=} n^{\frac{۱}{a}} X + d_n \ .$$

هم‌چنین متغیر تصادفی X پایدار است اگر تابع مشخصه آن به‌صورت زیر باشد:

$$\varphi(t) = exp\{i\lambda t - \gamma |t|^{\alpha} \ [۱ + i\beta sign(t) \ W(t,\alpha)]\} \ .$$

که در آن

$$W(t,\alpha) = \begin{cases} tan\frac{\alpha\pi}{۲} & \alpha \neq ۱ \\ \frac{۲}{\pi}\log|t| & \alpha = ۱ \end{cases} \qquad sign(t) = \begin{cases} ۱ & t > ۰ \\ ۰ & t = ۰ \\ -۱ & t < ۰ \end{cases}$$

پارامترهای تابع مشخصه عبارتند از: پارامتر مکان $(\infty > \lambda > -\infty)$، پارامتر پراکندگی $(۰ > \gamma)$، شاخص تقارن $(۱ \leq \beta \leq -۱)$ و نمای مشخصه$(۲ \leq \alpha < ۰)$. نمای مشخصه، ضخامت دم تابع چگالی احتمال را کنترل می‌کند. برای مقادیر کوچکتر α ، دم‌ها کشیده‌تر هستند.

ساختار توزیع پایدار: یک توزیع پایدار دارای چهار پارامتر می‌باشد. می‌توان نشان داد که هر توزیع پایدار دارای تابع چگالی پیوسته می‌باشد. اگر $f(x,a,\beta,c,\mu)$ تابع چگالی متغیر تصادفی X باشد و $Y = \sum_{i=۱}^{N} k_i(X_i - \mu)$ (مجموع مؤلفه‌های مستقل از X) تعریف شود آن‌گاه Y دارای تابع چگالی احتمال $s^{-۱} \ f(\frac{y}{s}, a, \beta, c, ۰)$ خواهد بود که در آن s برابر است با:

$$s = \left(\sum_{i=۱}^{N} |k_i|^a\right)^{\frac{۱}{a}} \qquad ۰ < a \leq ۲ \ .$$

و $f(x)$ توزیعی معادل توزیع زیر دارد که چنین تعریف می‌شود:

$$f(x) \sim \frac{ac^a(۱+\beta)\sin(\frac{\pi a}{۲})\frac{\Gamma(a)}{\pi}}{|x|^{۱+a}} \ . \qquad (۱ - ٦ - ۱)$$

[۱۸] Super-Gaussian (Leptokurtotic)

[۱۹] Heavy-tailed