تعیین کران‌های واریانس توزیع‌های تک‌مدی دم‌سنگین به کمک آنتروپی توانی

فاطمه آشتاب [۱] محمدحسین دهقان [۲] منیژه صانعی‌طبس [۳]

تاریخ دریافت: ۱۴۰۲/۰۷/۱۵
تاریخ پذیرش: ۱۴۰۲/۱۱/۱۵

چکیده:
واریانس و آنتروپی معیارهایی متمایز هستند که معمولاً برای اندازه‌گیری عدم قطعیت متغیرهای تصادفی استفاده می‌شوند. همان‌طور که واریانس هر متغیر تصادفی، پراکندگی آن حول میانگین را نشان می‌دهد، معیار آنتروپی عدم قطعیت یک رویکرد اطلاعاتی را اندازه‌گیری می‌کند به‌عبارت‌دیگر میانگین مقدار اطلاع یک متغیر تصادفی را اندازه‌گیری می‌کند. برای دو توزیع یکنواخت و نرمال واریانس نسبتی از آنتروپی توانی است. یافتن یک چنین رابطه یکنوا بین واریانس و آنتروپی برای یک رده بزرگ‌تر از این دو توزیع اهمیت و کاربرد زیادی در پردازش سیگنال، یادگیری ماشین، نظریه اطلاع و احتمال و آمار دارد. برای کم کردن خطاهای برآوردگرها مورداستفاده قرار می‌گیرد و یک راهبردی را انتخاب می‌کند که به‌طور متوسط بیشترین یا تقریباً بزرگ‌ترین کاهش در آنتروپی توزیع مکان هدف داشته باشد و اثربخشی این روش با استفاده از شبیه‌سازی‌ها با مدل‌های سنجش کاوی امتحان می‌شوند. در این مقاله کران بالای واریانس برای توزیع‌های تک مدی که دم آن‌ها سنگین‌تر از دم توزیع نمایی است به کمک آنتروپی توانی ایجاد می‌گردد.

واژه‌های کلیدی: آنتروپی توانی، کران‌های واریانس، توزیع‌های تک‌مدی، پیوستگی لیپ‌شیتز.

۱ مقدمه

آنتروپی یک مفهوم مهم در حوزه اطلاعات و تئوری اطلاعات است. در اصل، آنتروپی یک میزان از عدم قطعیت و تصادفی بودن یک مجموعه از رویدادها را نشان می‌دهد. آنتروپی به‌عنوان یک معیار اطلاعاتی استفاده می‌شود و میزان اطلاعات موجود در یک سیستم را مشخص می‌کند. مقدار آنتروپی بیشتر، نشان‌دهنده بیشترین عدم قطعیت و تصادفی بودن سیستم است [۲].

تعریف ۱.۱. *آنتروپی یک مفهوم کلیدی در حوزه‌های مختلف علمی و مهندسی، ازجمله فیزیک، اطلاعاتی، نظریه اطلاع، تکنولوژی اطلاعات، آمار و احتمالات و رمزگذاری است. این مفهوم ابتدا توسط رودولف کلوزیوس در قرن نوزدهم مطرح شد و سپس توسط کلود شانون در دهه ۱۹۴۰ و ۱۹۵۰ به‌صورت دقیق‌تر موردبررسی قرار گرفت. در فیزیک، آنتروپی به‌عنوان یک مفهوم مرتبط با نظریه ترمودینامیک استفاده می‌شود. آنتروپی معمولاً به‌عنوان یک معیار برای اندازه‌گیری ناهمگنی و بی‌نظمی یک سیستم تعریف می‌شود. در ترمودینامیک، آنتروپی نشان‌دهنده تعداد حالات میکروسکوپیک ممکن است که متناسب با یک حالت ماکروسکوپیک (وضعیت قابل‌مشاهده) است؛ به‌عبارت‌دیگر، آنتروپی میزان عدم قطعیت در یک سیستم را نشان می‌دهد و به میزان ترتیب یا بی‌نظمی در سیستم نسبت می‌دهد. در نظریه اطلاع، آنتروپی به‌عنوان یک مفهوم برای اندازه‌گیری میزان اطلاعات موجود در یک مجموعه از داده‌ها استفاده می‌شود. آنتروپی در این حالت نشان‌دهنده میزان عدم قطعیت در پیام یا داده‌ها است؛ به‌عبارت‌دیگر، آنتروپی بیشتر به معنی وجود اطلاعات بیشتر و عدم قطعیت کمتر است؛ که بعد از گفته شدن آنتروپی شانون، دانشمندان با تعمیم دادن آن معیارهای دیگر نیز برای آنتروپی ارائه دادند [۲].*

[۱] دانشجوی کارشناسی ارشد دانشگاه سیستان و بلوچستان
[۲] عضو هیئت‌علمی دانشگاه سیستان و بلوچستان
[۳] عضو هیئت‌علمی دانشگاه سیستان و بلوچستان (نویسنده مسئول: manijesanei@math.usb.ac.ir)

۱.۱ آنتروپی در حالت پیوسته

تعریف ۲.۱. *اگر X یک متغیر تصادفی در حالت پیوسته با تابع چگالی احتمال $f_X(x)$ باشد آنتروپی آن عبارت است از:*

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

۲.۱ آنتروپی در حالت گسسته

تعریف ۳.۱. *اگر X یک متغیر تصادفی گسسته با تابع احتمال $P_X(x)$ باشد، آنگاه*

$$H(X) = -\sum_{x} P_X(x)\log P_X(x).$$

آنتروپی متغیرهای تصادفی گسسته خواهد بود.

تعریف ۴.۱. *آنتروپی توانی مفهومی است که در حوزه تحلیل سیگنال و پردازش سیگنال مورداستفاده قرار می‌گیرد. آنتروپی توانی به‌صورت معمول در مطالعه توزیع‌های احتمالی پیوسته استفاده می‌شود. در واقع، آنتروپی توانی یک معیار از تنوع و پیچیدگی سیگنال به‌صورت زمانی است و مقدار آنتروپی توانی بیشتر، نشان‌دهنده وجود تغییرات پیچیده و تنوع بیشتر در سیگنال است [۲۹]. به‌طور خلاصه، آنتروپی مفهومی در حوزه اطلاعات و تئوری اطلاعات است که میزان عدم قطعیت و تصادفی بودن سیستم را اندازه‌گیری می‌کند، درحالی‌که آنتروپی توانی به‌عنوان یک مفهوم در حوزه پردازش سیگنال استفاده می‌شود و تنوع و پیچیدگی سیگنال را نشان می‌دهد.*

آنتروپی توانی یک مفهوم در نظریه اطلاع است که برای اندازه‌گیری میزان تنوع و پیچیدگی یک سیگنال استفاده می‌شود. آنتروپی توانی یک سیگنال، مقداری است که نشان می‌دهد چقدر انرژی در سیگنال توزیع‌شده است. معمولاً آنتروپی توانی برای سیگنال‌های پیوسته محاسبه می‌شود. برای محاسبه آنتروپی توانی یک سیگنال پیوسته، ابتدا توان سیگنال را در طول زمان یا فضا محاسبه می‌کنیم و سپس میانگین لگاریتم طبیعی این توان‌ها را محاسبه می‌کنیم. این میانگین توان را می‌توان با استفاده از انتگرال‌گیری نیز محاسبه کرد. آنتروپی توانی در برخی موارد می‌تواند به‌عنوان یک معیار برای تحلیل سیگنال‌های پیچیده و تصادفی مورداستفاده قرار گیرد. می‌توانیم یک رابطه مستقیم بین واریانس و آنتروپی تفاضلی برای توزیع‌های یکنواخت متقارن و نامتقارن داشته باشیم. هم‌چنین در نظریه اطلاع نابرابری آنتروپی توانی به‌اصطلاح به قدرت آنتروپی متغیر تصادفی مربوط می‌شود. این نشان می‌دهد، آنتروپی مفهومی است که در حوزه‌های مختلف برای اندازه‌گیری ناهمگنی، بی‌نظمی و عدم قطعیت استفاده می‌شود، اما اصول عمومی آنتروپی برای اندازه‌گیری وجود اطلاع و عدم قطعیت کاربرد دارد. تعیین کران‌های واریانس توزیع تک مدی با استفاده از آنتروپی توانی، در واقع یک روش برای تخمین زدن حداکثر و حداقل مقادیر واریانس در یک توزیع است. این روش به ما اطلاعاتی در مورد پراکندگی داده‌ها و نقطه تمرکز توزیع می‌دهد.

اهمیت تعیین کران‌های واریانس توسط آنتروپی توانی به دلایل زیر است [۲۷]:

۱. ارزیابی کیفیت و تنوع داده‌ها: واریانس یک معیار بسیار مهم برای اندازه‌گیری پراکندگی داده‌ها است. با تعیین کران‌های واریانس، می‌توانیم محدوده‌ای را مشخص کنیم که بیانگر پراکندگی مجموعه داده است. این محدوده ما را قادر می‌سازد تا به‌طور کیفی ارزیابی کنیم که آیا داده‌ها واحدی و یکنواخت هستند یا نه. اگر واریانس بسیار کم باشد، نشان می‌دهد که داده‌ها به‌طورکلی متمرکز و کم پراکنده هستند، درحالی‌که واریانس بزرگ نشانگر پراکندگی بیشتر و تنوع بیشتر است.

۲. تحلیل و پیش‌بینی رفتار توزیع: با تعیین کران‌های واریانس، می‌توانیم رفتار توزیع را بررسی کنیم و درک بهتری از شکل و خصوصیات آن پیدا کنیم. مثلاً، اگر واریانس بسیار کوچک باشد، نشان‌دهنده این است که توزیع تقریباً یکنواخت است و بیشتری تمرکز بر روی مقدار میانگین دارد؛ به‌عبارت‌دیگر، نقطه تمرکز توزیع به‌شدت تأثیرگذار است. از طرف دیگر، واریانس بزرگ نشان می‌دهد که داده‌ها پراکنده‌تر هستند و نقطه تمرکز توزیع کمتری اهمیت دارد.

۳. استفاده در تحلیل آماری: واریانس در بسیاری از روش‌ها و تحلیل‌های آماری به‌عنوان یک معیار مهم استفاده می‌شود. با تعیین کران‌های واریانس، می‌توانیم مطمئن شویم که تخمین‌های آماری ما دقیق و قابل‌اعتماد هستند.

بنابراین، تعیین کران‌های واریانس با استفاده از آنتروپی توانی، ابزاری قدرتمند برای ارزیابی و تحلیل داده‌هاست. این روش به ما کمک می‌کند تا به‌طور کمّی و کیفی بهتری از پراکندگی و تنوع داده‌ها دست پیدا کنیم و به‌صورت کمّی رفتار توزیع را تحلیل کنیم.

نابرابری آنتروپی توانی در سال ۱۹۴۸ توسط کلودشانون [۶] در مقاله اصلی خود او که به نام نظریه ارتباطات ریاضی بود، اثبات شد که نابرابری آنتروپی توانی را در ادامه بیان خواهیم کرد.

تعریف ۵.۱. *اگر X یک بردار تصادفی در $\mathbb{R}^n$ با چگالی f و آنتروپی $h(X)$ باشد، آنتروپی توانی آن به‌صورت زیر تعریف می‌گردد:*

$$N(X) = \frac{1}{2\pi e} e^{\frac{2}{n}h(X)} \qquad (1)$$

در حالت خاص وقتی X دارای توزیع نرمال با ماتریس کوواریانس K باشد، داریم:

$$N(X) = |K|^{\frac{1}{n}}$$

یک نابرابری مهم و پرکاربرد در رابطه با آنتروپی توانی معرفی‌شده است که به نابرابری آنتروپی توانی معروف است (EPI) در ادامه این نابرابری ارائه می‌گردد.

تعریف ۶.۱. *اگر X و Y دو بردار تصادفی مستقل در $\mathbb{R}^n$ با تابع چگالی مربوطه و گشتاور مرتبه دوم متناهی باشند، نابرابری زیر برقرار است[۶]:*

$$N(X+Y) \geq N(X) + N(Y) \qquad (۲)$$

برای مطالعه بیشتر و کامل‌تر در خصوص آنتروپی و سایر مفاهیم نظریه اطلاع به کتاب نظریه اطلاع تألیف کاور و توماس [۲]مراجعه شود.

۳.۱ کران‌دار کردن واریانس

کران بالای واریانس نشان‌دهنده یک هدف مهم است، زیرا در بسیاری از زمینه‌های آماری مانند تخمین واریانس، هم‌چنین به‌عنوان یک حداکثر برای فرآیند تصادفی، پرکاربرد است. اولین بار موضوع کران واریانس برای متغیر تصادفی که دارای کران بالا است به‌صورت زیر مطرح گردید: فرض کنید C یک عدد مثبت ثابت باشد. فرض کنید X یک متغیر تصادفی باشد که فقط مقادیری بین ۰ و C می‌گیرد. این دلالت بر $P(۰ \leq X \leq C) = ۱$ دارد. پس نابرابری زیر در مورد واریانس $V(X)$ ثابت می‌گردد:

$$V(X) \leq \frac{C^۲}{۴}$$

نابرابری‌های پوانکاره (هم‌سوپریمتری) که کران‌های بالایی را برای واریانس تابعی از یک متغیر تصادفی، به دست می‌دهند دارای تاریخچه طولانی و غنی است که با کار چرنوف (۱۹۸۱) [۴]شروع می‌شود. چرنوف با استفاده از چندجمله‌ای هرمیت، کران بالایی را برای واریانس تابعی از یک متغیر تصادفی معمولی به دست آورد. چن (۱۹۸۲) [۳]با استفاده از نابرابری کوشی- شوارتز، اثبات متفاوتی ارائه کرد و نابرابری را به حالت عادی چند متغیره تعمیم داد. کاکلوس (۱۹۸۲) [۵]نشان داده است که چگونه می‌توان کران‌های بالایی مشابه را برای توزیع‌های دیگر، ازجمله موارد گسسته، به دست آورد. علاوه بر این، با استفاده از نابرابری کرامر-رائو [۴] ، کران‌های پایینی مشابهی برای واریانس تابعی از یک متغیر تصادفی که شرایط نظم را برآورده می‌کند، ارائه داده است. همه این کران شامل دو گشتاور اول مشتقات یا تفاوت‌های تابع است. واریانس و آنتروپی معیارهایی متمایز هستند که معمولاً برای اندازه‌گیری عدم قطعیت متغیرهای تصادفی استفاده می‌شوند. درحالی‌که واریانس نشان می‌دهد که چگونه یک متغیر تصادفی بیشتر از حد انتظارش گسترش می‌یابد، معیار آنتروپی رویکرد متفاوتی دارد. آنتروپی عدم قطعیت یک رویکرد اطلاعاتی را اندازه‌گیری می‌کند به‌عبارت‌دیگر میانگین مقدار اطلاع یک متغیر تصادفی را اندازه‌گیری می‌کند. اگرچه واریانس و آنتروپی به‌طورکلی یکسان نیستند، یک مورد وجود دارد که واریانس و آنتروپی معیارهای معادل هستند. هنگامی‌که توزیع نرمال در نظر گرفته می‌شود، حداکثر آنتروپی برابر واریانس به‌دست‌آمده است. این نشان می‌دهد که واریانس و آنتروپی تا حدی مرتبط هستند هنگامی‌که توزیع احتمال اساساً چندوجهی باشد به‌گونه‌ای که برای هر آزمایش بیش از دو نتیجه، واریانس به‌عنوان یک معیار عدم قطعیت ناکارآمد می‌شود. به‌عنوان‌مثال، ترکیبی از دو توزیع نرمال تک متغیره هم‌وزن با واریانس واحد و میانگین‌های متفاوت را نظر بگیرید. هنگامی‌که اختلاف میانگین دو متغیر بینهایت شود واریانس کل توزیع بی‌نهایت می‌شود ولی آنتروپی در این موقعیت پایدار باقی می‌ماند؛ بنابراین، آنتروپی به‌عنوان معیار مناسب‌تری برای عدم قطعیت متغیرهای تصادفی چندوجهی در نظر گرفته می‌شود. این مطلب به کمک نابرابری معروفی بنام نابرابری فانو نشان داده می‌شود که به آن اشاره خواهیم کرد.

تعریف ۷.۱. *برای هر برآوردگر $\widehat{X}$ به‌طوری‌که $X \longrightarrow Y \longrightarrow \widehat{X}$ با $(X \neq \widehat{X})$ و با در نظرگرفتن احتمال خطا به شکل $P_e = P(X \neq \widehat{X})$ نابرابری فانو به‌صورت زیر بیان می‌شود [۲۱]:*

$$H(P_e) + P_e \log |X| \geq H(X|\widehat{X}) \geq H(X|Y) \qquad (۳)$$

توجه داشته باشید که در رابطه (۳) $P_e = ۰$ به این معنی است که $H(X|Y) = ۰$

هم‌چنین فرض کنید با شناخت یک متغیر تصادفی Y می‌خواهیم مقدار یک متغیر تصادفی همبسته X را حدس بزنیم. نابرابری فانو احتمال خطا در حدس زدن متغیر تصادفی X را به آنتروپی شرطی آن $H(X|Y)$ مرتبط می‌کند.

[4] Cramer-Rao

۲ کران‌های واریانس برای توزیع‌های تک‌مدی

تعریف ۱.۲. *در توزیع‌های تک‌مدی مقادیر در ابتدا افزایش می‌یابند و به یک قله منفرد می‌رسند و سپس از پرتکرارترین عدد در یک حالت کاهش می‌یابند. به‌طورکلی اگر در یک توزیع یک مد وجود داشته باشد به آن توزیع تک مدی گفته می‌شود و این حالت هم در توزیع‌های گسسته و هم‌پیوسته تعریف‌شده است که می‌توانیم به چند توزیع مانند توزیع گاما، برنولی، هندسی، دوجمله‌ای منفی، فوق هندسی، پواسون، نمایی، خی‌دو، نرمال و غیره اشاره‌کرد. تعریف رسمی خاصیت تک‌مدي به‌صورت زیر است [۲۲].*

تعریف ۲.۲. *یک متغیر تصادفی Z با تابع توزیع F را تک‌مدی حول (a) می‌گوییم، اگر F در فاصله‌ی $(-\infty, a)$ محدب (تقعر رو به بالا) و در فاصله‌ی (a, ∞) مقعر (تقعر رو به پایین) باشد.*

تعریف ۳.۲. ***توزیع تک‌مدی پیوسته:*** *اگر تابع چگالی یک متغیر تصادفی پیوسته مطلق روي فاصله $(-\infty, a)$ صعودی و در فاصله (a, ∞) نزولی باشد، آنگاه آن توزیع تک‌مدی پیوسته حول a خوانده می‌شود.*

تعریف ۴.۲. ***توزیع تک‌مدی گسسته:*** *یک توزیع گسسته با تابع جرم احتمال (p_n) را روی مجموعه‌ي اعداد صحیح، تک‌مدي حول a می‌نامند اگر در رابطه‌ی زیر صدق کند:*

$$\begin{cases} p_n \geq p_{n-1}, & n \leq a \\ p_n \geq p_{n+1}, & n \geq a. \end{cases}$$

برای توزیع‌های تک‌مدی گسسته توزیع‌هایی چون برنولی، دوجمله‌ای، هندسی، پواسون و ... را می‌توان نام برد.

لم ۵.۲. *پیوستگی لیپ‌شیتز: تحلیل ریاضی، پیوستگی لیپ‌شیتز که به نام ریاضیدان رودولف لیپ‌شیتز نام‌گذاری شده است، یک شکل قوی از پیوستگی یکنواخت برای توابع است. به‌طور شهودی یک تابع پیوسته لیپ‌شیتز سرعت تغییر آن محدود است. فرض کنیم $f : (X, d_1) \to (Y, d_2)$ یک تابع باشد. گوییم f در شرط لیپ‌شیتز با ثابت M صدق می‌کند هرگاه عددی ثابت مانند $M > 0$ موجود باشد به قسمی‌که به ازای هر $x, y \in X$ داشته باشیم*

$$d_2(f(x), f(y)) \leq M d_1(x, y).$$

چانگ ۲۰۱۵ [۱]

۱.۲ توزیع دم‌سنگین

توزیع دم‌سنگین [5] یک مفهوم در آمار و احتمال است که به توزیعی اشاره دارد که دارای دم‌های بزرگ و احتمالات نادر در این دم‌ها است؛ به‌عبارت‌دیگر، در یک توزیع دم‌سنگین، احتمال وقوع رویدادهای نادر و خارج از محدوده معمول بزرگ است.

این مفهوم معمولاً در زمینه تحلیل داده‌ها و مدل‌سازی استفاده می‌شود. توزیع دم‌سنگین می‌تواند در مواردی مانند توزیع درآمد، سایز فایل‌ها در سیستم‌عامل، تعداد دنبال کنندگان در شبکه‌های اجتماعی و غیره مشاهده شود.

توزیع دم‌سنگین معمولاً به معنای وجود انحرافات شدید در داده‌ها است که از سنت‌های نرمال یا یک توزیع دم متوسط که دارای دمی سنگین‌تر است، جدا می‌شود. این انحرافات می‌توانند نشان‌دهنده رویدادهای نادر یا غیرمعمول باشند که در تحلیل داده‌ها بسیار مهم هستند.

مهم‌ترین خصوصیت توزیع دم‌سنگین، وجود مقادیر نادر (اعداد بزرگ) است که می‌توانند تأثیر بسیار زیادی بر نتایج تحلیل داشته باشند. این موضوع ممکن است به تحلیل‌های آماری سنتی‌که بر اساس فرضیاتی مانند توزیع نرمال بر مبنای داده‌های معمول توسعه‌یافته‌اند، چالش بیاورد؛ بنابراین، در تجزیه‌وتحلیل داده‌ها و استنتاج‌های آماری بر مبنای توزیع دم‌سنگین، نیاز به روش‌های خاص و مناسبی وجود دارد [۲۸].

۲.۲ کران‌های واریانس چگالی‌های مخلوط تک‌مدی متقارن

ابتدا چگالی‌های مخلوط خطی تک‌مدی متقارن به فرم زیر را که مخلوطی از توزیع‌های نزولی نمایی، $p_i(x) \propto e^{-\beta_i |x-m|^{\theta_i}}$ برای هر $\beta_i > 0$، $\theta_i > 0$ به ازای $i = 1, 2, \ldots, n$ است را در نظر می‌گیریم [۱]:

$$p(x) = \sum_{i=1}^{n} \alpha_i p_i(x), \qquad \alpha_i > 0, \sum_{i=1}^{n} \alpha_i = 1. \qquad (۴)$$

که $p_i(x) \propto e^{-\beta_i |x-m|^{\theta_i}}$ آمیخته‌ای از توزیع‌های نزولی نمایی برای هر $\beta_i > 0$ و $\theta_i > 0$ است.

برای هر قضیه ۶.۲ کران بالای واریانس چگالی نزولی نمایی مخلوط خطی تک مدی را تحت این فرض که نسبت مابین واریانس‌های

[5] heavy-tailed distribution

ماکزیمم و مینیمم اجزای مخلوط $p_i(x)$ کران‌دار است را ثابت می‌کند.

قضیهٔ ۶.۲. *فرض کنید $p(x)$ چگالی مخلوط خطی تک‌مدی متقارن به فرم (۴) با مؤلفه $p_i(x) = \frac{1}{Z_i(\theta_i,\beta_i)}e^{-\beta_i|x-m|^{\theta_i}}$ باشد که $Z_i(\theta_i,\beta_i)$ ثابت نرمال شده است و $\theta_i,\beta_i > 0$. فرض کنید σ_i^2 نشان‌دهنده واریانس $p_i(x)$ و نسبت مؤلفه واریانس‌های σ_i^2/σ_j^2 برای همه $i \neq j$ کران‌دار باشد، به‌عبارت‌دیگر، فرض کنید $r = \max_{i,j\in\{1,2,\ldots,n\}}\left\{\frac{\sigma_i^2}{\sigma_j^2}\right\}$. در این صورت واریانس چگالی $p(x)$ در رابطه زیر صدق می‌کند[۱]:*

$$\frac{e^{2h(p)}}{2\pi e} \le Var(X) \le B(\boldsymbol{\theta},r)e^{2h(p)}. \qquad (۵)$$

در اینجا،

$$B(\boldsymbol{\theta},r) = M(r)\cdot\prod_{i=1}^{n}\left(\frac{1}{A(\theta_i)}\right)^{\alpha_i}, \qquad (۶)$$

برای $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_n)^T$، که

$$A(\theta) = 4\theta^{-2}\frac{(\Gamma(1/\theta))^3}{\Gamma(3/\theta)}e^{2/\theta}, \qquad (۷)$$

با تابع گاما $\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}$ که برای همه $t > 0$ تعریف می‌شود و

$$M(r) = \frac{(r-1)r^{\frac{1}{r-1}}}{e\log r}, \qquad r \ge 1. \qquad (۸)$$

برابری در واریانس کران پایین در رابطه (۵) به دست می‌آید اگر و تنها اگر $p(x)$ توزیع گاوسی باشد. برابری در واریانس کران پایین زمانی برآورده می‌شود که همه $p_i(x)$ها دارای یک توزیع باشند؛ به‌عبارت‌دیگر، برای هر i، $p(x) = p_i(x)$. در نتیجه، کران‌های بالا و پایین هم‌ارز هستند هرگاه همه $p_i(x)$ها دارای توزیع گاوسی باشند. برای یک توزیع مخلوط به فرم (۴)، احتمالاً با $p_i(x) \neq p_j(x)$ برای هر جفت (i,j)، اسکالر ثابت $B(\boldsymbol{\theta},r)$ در رابطه (۶) بزرگ‌تر یا مساوی $1/2\pi e$ است و ازاین‌رو میانگین هندسی $1/A(\theta_i) \ge 1/(2\pi e)$ برای هر θ_i، با وزن $\{\alpha_i\}$، به‌صورت

$$\prod_{i=1}^{n}(1/A(\theta_i))^{\alpha_i} \ge 1/(2\pi e),$$

داده می‌شود و $M(r) \ge 1$ با $\lim_{r\to 1} M(r) = 1$.

لم ۷.۲. *تابع $1/A(\theta)$ با $A(\theta)$ داده‌شده در (۷) روی $[0,2]$ نزولی و روی $[2,+\infty)$ صعودی است. مینیمم $1/A(\theta)$ در $\theta = 2$ با مقدار $1/A(2) = 1/(2\pi e)$ رخ می‌دهد. هرگاه $\theta \to 0$ باشد[۱]، $\lim_{\theta\to 0} 1/A(\theta) = \infty$ و هرگاه $\theta \to \infty$، $\lim_{\theta\to\infty} 1/A(\theta) = 1/12$.*

با استفاده از این لم ۷.۲ ، می‌توانیم کران بالای واریانس توزیع مخلوط را ساده کنیم هرگاه همه θ_iها مؤلفه مخلوط $p_i(x) = \frac{1}{Z_i(\theta_i,\beta_i)}e^{-\beta_i|x-m|^{\theta_i}}$ برای $i = 1,\ldots,n$ کران‌دار پایینی به ثابت مثبت باشند. برای مثال، هرگاه برای هر i، $\theta_i \ge 1$ باشد، بیشترین نسبت بین کران بالایی و کران پایینی برابر با $\frac{2\pi e}{12}M(r)$ می‌شود ازاین‌رو

$$\prod_{i=1}^{n}(1/A(\theta_i))^{\alpha_i} \le 1/12.$$

هرگاه برای هر i، $\theta_i \ge 1/2$ باشد، آنگاه $\prod_{i=1}^{n}(1/A(\theta_i))^{\alpha_i} \le (2e^4)/15$، بیشترین نسبت بین کران بالایی و کران پایینی برابر $\frac{4\pi e^5}{15}M(r)$ می‌شود. توجه داشته باشید هرگاه برای هر i، $\theta_i \ge 1/2$ باشد، مؤلفه‌های آمیخته $p_i(x)$ می‌توانند توزیع‌های دُم-سنگین باشند اما واریانس $p(x)$ هنوز از بالا به تابع یکنوا با آنتروپی توانی $e^{2h(p)}$ کران‌دار است[۱].

نتیجه ۸.۲. *چگالی تک‌مدی متقارن به فرم $p(x) = \sum_{i=1}^{n}\alpha_i p_i(x)$ برای $\alpha_i > 0$، در نظر بگیرید که $p_i(x) = \frac{1}{Z_i(\theta_i,\beta_i)}e^{-\beta_i|x-m|^{\theta_i}}$ با ثابت نرمال شده $Z_i(\theta_i,\beta_i)$ و $\beta_i > 0$ است. هرگاه برای هر i، مرتبه $\theta_i \ge 1$ باشد، واریانس به‌صورت زیر کران‌دار است[۱]:*

$$\frac{e^{2h(p)}}{2\pi e} \le Var(X) \le \frac{M(r)}{12}e^{2h(p)}. \qquad (۹)$$

هرگاه برای هر i، $\theta_i \ge 1/2$،

$$\frac{e^{2h(p)}}{2\pi e} \le Var(X) \le \frac{2e^4 M(r)}{15}e^{2h(p)}, \qquad (۱۰)$$

که $M(r)$ در رابطه (۸) داده‌شده است.

نتیجه ۹.۲. *فرض کنید که $p(x)$ چگالی تک‌مدی متقارن با تکیه‌گاه کران‌دار به فرم $p(x) = \sum_{i=1}^{n}\alpha_i p_i(x)$ باشد که $p_i(x) = \text{unif}\left(m-\frac{1}{2\epsilon_i}, m+\frac{1}{2\epsilon_i}\right)$ برای $\epsilon_i > 0$. همچنین، فرض کنید $r = \max_{i,j\in\{1,\ldots,n\}}\left\{\frac{\epsilon_i^2}{\epsilon_j^2}\right\}$ کران‌دار باشد. در این صورت واریانس به‌صورت زیر کران‌دار است[۱]:*

$$\frac{e^{2h(p)}}{2\pi e} \le Var(X) \le \frac{M(r)}{12}e^{2h(p)}, \qquad (۱۱)$$

که $M(r)$ در رابطه (۸) داده‌شده است.

در قضیه ۶.۲ و نتیجه ۸.۲ کران‌داری نسبت r بین واریانس ماکزیمم و واریانس مینیمم مؤلفه‌های مخلوط را فرض کردیم؛ یعنی، $\max_{i,j\in\{1,\ldots,n\}}\left\{\sigma_i^2/\sigma_j^2\right\}$ در قضیه ۶.۲ و $\max_{i,j\in\{1,\ldots,n\}}\left\{\epsilon_i^2/\epsilon_j^2\right\}$

در نتیجه ۹.۲ کران‌دار هستند. در اینجا، با یک مثال نقض نشان می‌دهیم که هرگاه این فرض نقض شود، واریانس چگالی مخلوط تک‌مدی متقارن لزوماً محدود در تابع یکنوای آنتروپی توانی نیست.

مثال ۱۰.۲. *توزیع تک‌مدی متقارن تشکیل‌شده از دو چگالی یکنواخت به فرم زیر را در نظر بگیرید:*

$$p(x)=\sum_{i=1}^{2}\alpha_i\cdot \mathrm{unif}\left(-\frac{1}{2\epsilon_i},\frac{1}{2\epsilon_i}\right) \qquad (12)$$

که $\alpha_i>0$ *و برای* $\epsilon_1>\epsilon_2>0$، $\sum_{i=1}^{2}\alpha_i=1$. *واریانس این توزیع برابر است با:*

$$Var(X)=\frac{1}{12}\left(\alpha_i\frac{1}{\epsilon_1^2}+\alpha_2\frac{1}{\epsilon_2^2}\right), \qquad (13)$$

و آنتروپی تفاضلی این توزیع به‌صورت زیر است:

$$\begin{aligned} h(p) &= -\frac{1}{\epsilon_1}(\alpha_1\epsilon_1+\alpha_2\epsilon_2)\log(\alpha_1\epsilon_1+\alpha_2\epsilon_2) \\ &\quad -\left(\frac{1}{\epsilon_2}-\frac{1}{\epsilon_1}\right)\alpha_2\epsilon_2\log(\alpha_2\epsilon_2). \end{aligned} \qquad (14)$$

هرگاه $\epsilon_1/\epsilon_2\to\infty$ *باشد، حد آنتروپی تفاضلی به‌صورت رابطه زیر درمی‌آید:*

$$\lim_{\epsilon_1/\epsilon_2\to\infty}h(p)=-\alpha_1\log\epsilon_1-\alpha_2\log\epsilon_2+H_B(\alpha_1) \qquad (15)$$

که $H_B(\alpha_1)=-\alpha_1\log\alpha_1-(1-\alpha_1)\log(1-\alpha_1)$. *در این صورت آنتروپی توانی به‌صورت زیر درمی‌آید:*

$$\lim_{\epsilon_1/\epsilon_2\to\infty}e^{2h(p)}=e^{H_B(\alpha_1)}\left(\frac{1}{\epsilon_1^2}\right)^{\alpha_1}\left(\frac{1}{\epsilon_2^2}\right)^{\alpha_2}. \qquad (16)$$

برای یافتن یک کران بالا روی واریانس (۱۳) با آنتروپی توانی (۱۶)، نیاز به یک کران بالا روی میانگین حسابی $\left(\alpha_1\frac{1}{\epsilon_1^2}+\alpha_2\frac{1}{\epsilon_2^2}\right)$ *در جملات میانگین هندسی* $\left(\frac{1}{\epsilon_1^2}\right)^{\alpha_1}$ *و* $\left(\frac{1}{\epsilon_2^2}\right)^{\alpha_2}$ *داریم. بااین‌وجود، اگر* $\epsilon_2\to 0$ *برای ثابت* ϵ_1*، بنا به این‌که میانگین حسابی روی مرتبه* $\Theta(1/\epsilon_2^2)$ *افزایش می‌یابد درحالی‌که میانگین هندسی روی مرتبه* $\Theta\left(1/\epsilon_2^{(2\alpha_2)}\right)$ *برای* $\alpha_2<1$ *افزایش واریانس سریع‌تر از* $e^{2h(p)}$ *است برای این‌که نمی‌تواند از بالا برای هر ثابت اسکالر آنتروپی توانی کران‌دار شود. از طرف دیگر، اگر* $\epsilon_1\to\infty$ *برای ثابت* ϵ_2*، میانگین حسابی که متناسب با واریانس است، به‌طور تقریبی* α_2/ϵ_2^2 *است که یک ثابت است اما میانگین هندسی که متناسب با* $e^{2h(p)}$ *است برای* $\alpha_1<1$ *به* 0 *از مرتبه* $\Theta\left(1/\epsilon_1^{(2\alpha_1)}\right)$ *میل می‌کند. در نتیجه، برای هر حالت که در ویژگی* $\epsilon_1/\epsilon_2\to\infty$ *صدق می‌کند، واریانس* $p(x)$ *نمی‌تواند از بالا به ثابت اسکالر* $e^{2h(p)}$ *کران‌دار شود. این مثال نشان می‌دهد که هرگاه نسبت بین واریانس‌های مؤلفه‌های مخلوط بی‌کران باشد، لزوماً کران بالا روی واریانس توزیع تک‌مدی متقارن وجود ندارد یعنی یک ثابت آنتروپی توانی مدرج شده است.*

۳.۲ کران‌های واریانس چگالی‌های تک‌مدی پیوسته-لیپ‌شیتز با تکیه‌گاه کران‌دار

با مؤلفه‌های مخلوط به‌اندازه کافی بزرگ، هر چگالی تک‌مدی متقارن می‌تواند به‌دلخواه همانند یک مخلوط خطی به فرم (۴) تقریب زده شود. این برای حالت خاص چگالی‌های یکنواخت $p_i(x)$ اثبات‌شده است [۲۳]. ما در اینجا، این نتیجه را تعمیم می‌دهیم. فرض کنید $p(x)$ چگالی تک‌مدی متقارن با تکیه‌گاه کران‌دار $[m-s,m+s]$ باشد. فرض کنید که $p(x)$ در شرط پیوستگی-لیپ‌شیتز با ثابت $c_s>0$ صدق کند، به‌عبارت‌دیگر

$$|p(x+y)-p(x)|\le c_s|y| \qquad (17)$$

برای هر $x,y\in[m-s,m+s]$. برای چنین $p(x)$ای، چگالی مخلوط خطی $\overline{p}_n(x)$ را می‌سازیم که تقریب‌های $p(x)$ به قسمی که تمایز بین توان‌های آنتروپی $p(x)$ و $\overline{p}_n(x)$ و واریانس‌های $p(x)$ و $\overline{p}_n(x)$ می‌توانند به شکل زیر کران‌دار شوند[۱]:

$$\begin{aligned} &e^{2h(\overline{p}_n)}\le e^{2h(p)}\left(1+c_1n^{-1}\log n\right), \\ &\left|\int_{m-s}^{m+s}(x-m)^2(p(x)-p_n(x))dx\right|\le c_2n^{-1} \end{aligned} \qquad (18)$$

همچنین، یک کران بالای واریانس از چنین چگالی مخلوط خطی $\overline{p}_n(x)$ را ثابت می‌کنیم و با نماد $Var(\overline{p}_n)$ نشان می‌دهیم به قسمی که

$$Var(\overline{p}_n)\le\frac{c_ss^2e^{c_ss^2}}{24}e^{2h(\overline{p}_n)}(1+c_2n^{-1}) \qquad (19)$$

با ترکیب روابط (۱۸) و (۱۹) و با فرض $n\to\infty$، کران بالای واریانس هر چگالی تک‌مدی متقارن پیوسته-لیپ‌شیتز با تکیه‌گاه کران‌دار را در آنتروپی توانی ثابت می‌کنیم.

قضیهٔ ۱۱.۲. *برای هر چگالی تک‌مدی متقارن* $p(x)$ *با تکیه‌گاه کران‌دار* $[m-s,m+s]$*، هرگاه* $p(x)$ *در شرط لیپ‌شیتز (۱۷) با ثابت* $c_s>0$ *صدق کند، واریانس* $Var(X)$ *می‌تواند از بالا و پایین به ثابت اسکال آنتروپی توانی به‌صورت زیر است:*

$$\frac{e^{2h(p)}}{2\pi e}\le Var(X)\le\frac{c_ss^2e^{c_ss^2}}{24}e^{2h(p)}. \qquad (20)$$

که c_s ثابت لیپ‌شیتز و s نصف اندازه تکیه‌گاه است. در ادامه، واریانس کران بالای را برای حالتی که چگالی‌ها تک‌مدی متقارن لیپ‌شیتز پیوسته باشند، تعمیم می‌دهیم. چگالی تک‌مدی $p(x)$ با تکیه‌گاه کران‌دار $[b-s_l, b+s_r]$ با $s_l, s_r > 0$ را در نظر بگیرید. تعریف کنید: $s = \max\{s_l, s_r\}$. فرض کنید حالت مُد این چگالی در $x = b$ رخ دهد. فرض کنید m نشان‌دهنده میانگین این چگالی باشد، $m = \int_{-\infty}^{+\infty} xp(x)dx$. فرض کنید $p(x)$ لیپ‌شیتز پیوسته باشد؛ به‌عبارت‌دیگر،

$$|p(x+y) - p(x)| \le c_s|y| \qquad (۲۱)$$

برای هر $x, y \in [b-s_l, b+s_r]$. در نتیجه، واریانس این چگالی $p(x)$ می‌تواند از بالا به منهای ثابت اسکال آنتروپی توانی $(m-b)^2$ کران‌دار شود که ثابت یک تابع با ثابت لیپ‌شیتز c_s و اندازه ماکزیمم $s = \max\{s_l, s_r\}$ از تکیه‌گاه یک‌بعدی از مد $x = b$ است[۱].

قضیهٔ ۱۲.۲. *برای هر چگالی تک‌مدی $p(x)$ روی تکیه‌گاه کران‌دار $[b-s_l, b+s_r]$ با مد $x = b$ و میانگین m که $p(x)$ در شرط لیپ‌شیتز (۲۱) صدق می‌کند، واریانس $Var(X)$ از بالا و پایین کران‌دار به جملات آنتروپی توانی به‌صورت زیر است:*

$$\frac{e^{2h(p)}}{2\pi e} \le Var(X) \le \frac{c_s s^2 e^{c_s s^2} M(128(c_s s^2)^4)}{6} e^{2h(p)} - (m-b)^2 \qquad (۲۲)$$

با $s = \max\{s_l, s_r\}$ و $M(r)$ که در رابطه (۸) داده‌شده است[۱].

۴.۲ کران بالای واریانس چگالی‌های مخلوط تک‌مدی متقارن

فرض کنید توزیع مخلوط $p(x)$ از تعداد متناهی توزیع‌های نزولی نمایی، $p_i(x) = \frac{1}{Z_i(\theta_i,\beta_i)} e^{-\beta_i|x-m|^{\theta_i}}$ با مرتبه $\theta_i > 0$ و $\beta_i > 0$ برای $i = 1, \dots, n$ با وزن‌های مخلوط α_i در نظر می‌گیریم؛ به‌عبارت‌دیگر

$$p(x) = \sum_{i=1}^{n} \left(\frac{1}{Z_i(\theta_i,\beta_i)} e^{-\beta_i|x-m|^{\theta_i}} \right) \qquad (۲۳)$$

که $\alpha_i > 0$ و $\sum_{i=1}^n \alpha_i = 1$. همه مؤلفه‌های مخلوط $p_i(x)$ دارای همان میانگین m با تضمین تک‌مدی $p(x)$ است. توجه داشته باشید برای هر مؤلفه مخلوط $p_i(x)$، ثابت نرمال شده برابر $Z_i(\theta_i,\beta_i) = 2\beta_i^{-1/\theta_i}\theta_i^{-1}\Gamma(1/\theta_i)$ و واریانس $\sigma_i^2 = \beta_i^{-2}\frac{\Gamma(3/\theta_i)}{\Gamma(1/\theta_i)}$ است. واریانس چگالی مخلوط تک‌مدی متقارن ۲۳ به‌صورت زیر است:

$$Var(X) = \int_{-\infty}^{\infty} (x-m)^2 p(x)dx = \sum_{i=1}^{n} \alpha_i \sigma_i^2. \qquad (۲۴)$$

برای چگالی مخلوط خطی $p(x)$ از (۲۳) کران بالا روی واریانس در جملات آنتروپی توانی همانند چیزی که در قضیه ۶.۲ بیان شد به دست می‌آید که در اینجا اثبات را ارائه می‌دهیم[۱].

با استفاده از مقعر بودن آنتروپی تفاضلی $h(p)$ در توزیع $p(x)$ داریم

$$h(p) \ge \sum_{i=1} \alpha_i h(p_i), \qquad (۲۵)$$

و در نتیجه

$$e^{2h(p)} \ge e^{\sum_{i=1}^n 2\alpha_i h(p_i)} = \prod_{i=1}^{n} \left(e^{2h(p_i)}\right)^{\alpha_i}. \qquad (۲۶)$$

برای $p_i(x) = \frac{1}{Z_i(\theta_i,\beta_i)} e^{-\beta_i|x-m|^{\theta_i}}$ داریم

$$\sigma_i^2 = \frac{1}{A(\theta_i)} \cdot e^{2h(p_i)} \qquad (۲۷)$$

و در نتیجه

$$e^{2h(p)} \ge \left(\prod_{j=1}^{n} A(\theta_j)^{\alpha_j}\right) \cdot \left(\prod_{i=1}^{n} (\sigma_i^2)^{\alpha_i}\right). \qquad (۲۸)$$

با استفاده از معکوس نامساوی میانگین توانی داده‌شده در [۲۴، ۲۵] یک کران بالا روی میانگین هندسی $\{\sigma_i^2\}$ با مرتبه‌های $\{\alpha_i\}$ در جملات میانگین حسابی $\{\sigma_i^2\}$ با مرتبه‌های $\{\alpha_i\}$ به‌صورت زیر داده می‌شود:

$$\sum_{i=1}^{n} \alpha_i \sigma_i^2 \le M(r) \prod_{i=1}^{n} (\sigma_i^2)^{\alpha_i} \qquad (۲۹)$$

که $M(r)$ در رابطه (۸) تعریف‌شده است و $r = \max_{i,j\in\{1,\dots,n\}} \left\{ \frac{\sigma_i^2}{\sigma_j^2} \right\}$.

بنا به این‌که واریانس توزیع مخلوط $p(x) = \sum_{i=1}^n \alpha_i p_i(x)$ برابر $Var(X) = \sum_{i=1}^n \alpha_i \sigma_i^2$ است، با ترکیب روابط (۲۸) و (۲۹) داریم:

$$\begin{aligned} Var(X) &= \sum_{i=1}^{n} \alpha_i \sigma_i^2 \\ &\le M(r) \prod_{i=1}^{n} (\sigma_i^2)^{\alpha_i} \\ &\le M(r) \left(\prod_{i=1}^{n} \left(\frac{1}{A(\theta_i)}\right)^{\alpha_i}\right) e^{2h(p)}. \end{aligned} \qquad (۳۰)$$

۳ مثال‌ها

۱. نمودار کران واریانس توزیع نمایی:

توزیع نمایی تنها توزیع پیوسته‌ای است که خاصیت بی‌حافظگی دارد و ازاین‌رو بیشتر در حل مسائل احتمال و نظریه صف به کار گرفته می‌شود. هم‌چنین از این توزیع برای مدل‌سازی کردن و آسان ساختن شیوه‌ی حل مسائل واقعی استفاده می‌کنند. این ویژگی تابع را می‌توان این‌طور تفسیر کرد که رویدادهایی را که در گذشته اتفاق افتاده می‌توانیم در نظر نگیریم و از زمان حال

به بعد را مبدأ زمان قرار بدهیم. مثلاً لامپی که طول عمرش ۱۰ ساعت است و تا ساعت ۶ هنوز نسوخته است را می‌توان مثل یک لامپ نو به‌حساب آورد.

$$f(x) = \lambda e^{-\lambda x}, \qquad \forall x \geq 0.$$

سپس

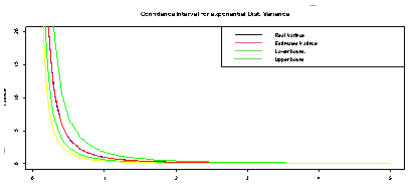
$$\begin{aligned} h(p) &= -\int_0^\infty \lambda e^{-\lambda x} \log(\lambda e^{-\lambda x}) dx h_e(X) \\ &= -\left(\int_0^\infty (\log \lambda) \lambda e^{-\lambda x} dx + \int_0^\infty (-\lambda x) \lambda e^{-\lambda x} dx \right) \\ &= -\log \lambda \int_0^\infty f(x) dx + \lambda E[X] \\ &= -\log \lambda + 1. \end{aligned}$$

به‌سادگی می‌توان نتیجه گرفت:

$$h(p) = -\log \lambda + 1 \Rightarrow \lambda = e^{1-h(p)}$$

واریانس توزیع نمایی:

$$\begin{aligned} v(x) &= \frac{1}{\lambda^2} = e^{2(h(p)-1)}, \\ \frac{e^{2h(p)}}{2\pi e} &\leq e^{2(h(p)-1)} = v(x) \end{aligned}$$



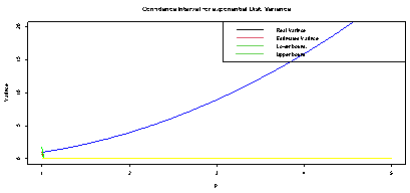
شکل ۱: نمودار کران واریانس توزیع نمایی.

۲. نمودار کران واریانس توزیع نرمال:

$$\begin{aligned} h(p) &= \int_{-\infty}^{\infty} f(x) \log \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) dx \\ &= \int_{-\infty}^{\infty} f(x) \log \frac{1}{\sqrt{2\pi\sigma^2}} dx + \\ &\quad \log(e) \int_{-\infty}^{\infty} f(x) \left(-\frac{(x-\mu)^2}{2\sigma^2} \right) dx \\ &= -\frac{1}{2} \log(2\pi\sigma^2) - \log(e) \frac{\sigma^2}{2\sigma^2} \\ &= -\frac{1}{2} \left(\log(2\pi\sigma^2) + \log(e) \right) \\ &= -\frac{1}{2} \log(2\pi e\sigma^2). \end{aligned}$$

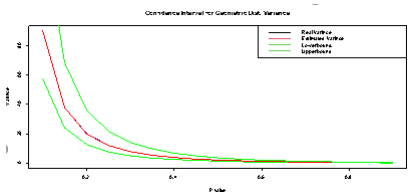
از رابطه بالا واریانس توزیع نرمال نتیجه می‌شود که:

$$h(p) = \frac{1}{2} \log(2\pi e\sigma^2) \Rightarrow \sigma^2 = \frac{e^{2h(p)}}{2\pi e}.$$



شکل ۲: نمودار کران واریانس توزیع نرمال.

۳. نمودار کران واریانس توزیع هندسی:



شکل ۳: نمودار کران واریانس توزیع هندسی.

۴ بحث و نتیجه‌گیری

در این مقاله کران‌های بالا روی واریانس از زیررده‌های توزیع‌های تک‌مدی در جملات آنتروپی توانی را ثابت کردیم. ابتدا چگالی‌های مخلوط تک‌مدی متقارن از توزیع‌های نزولی نمایی و توزیع‌های یکنواخت را در نظر گرفتیم. تنگی کران‌های بالا روی واریانس روی نسبت بین ماکزیمم و مینیمم واریانس‌ها از مؤلفه‌های مخلوط مستقل است. با ارائه یک مثال نقض، نشان دادیم که هرگاه نسبت بین واریانس‌های مؤلفه‌های مخلوط بی‌کران است و لزوماً کران بالا روی واریانس چگالی‌های مخلوط تک‌مدی یکنواخت وجود ندارد که ثابت مدرج شده آنتروپی توانی باشد. همچنین نشان دادیم که واریانس هر چگالی تک‌مدی لیپ‌شیتز-پیوسته با تکیه‌گاه کران‌دار می‌تواند به جملات آنتروپی توانی کران‌دار شود. همه کران‌های بالا روی واریانس چگالی‌های تک‌مدی ارائه‌شده در اینجا مدرج-پایا هستند.

در پردازش سیگنال، سنجش تطبیقی و یادگیری ماشین، جانشین‌های نظری اطلاعات مانند واگرایی کولبک لیبلر، آنتروپی و اطلاعات فیشر به‌طور گسترده به‌جای توابع هزینه‌های خاص کار مانند میانگین مجذور خطا یا احتمال خطای طبقه‌بندی مورداستفاده قرارگرفته‌اند. ازآنجایی‌که چنین توابع هزینه‌های خاص کار اغلب غیرقابل‌حل هستند، جانشین‌های اطلاعاتی به‌عنوان اهداف طبیعی برای توسعه شکل‌های موج یا استراتژی‌های انتخاب حسگر برای جمع‌آوری و یا فیلتر کردن اطلاعات استفاده می‌شوند. نتایج گزارش‌شده در اینجا می‌تواند برای توجیه استفاده از آنتروپی دیفرانسیل به‌عنوان

جایگزینی برای میانگین مربعات خطا در چنین کاربردهایی با ضمانت عملکرد قابل‌اثبات، زمانی که توزیع خلفی لیپ‌شیتز و تک‌وجهی با پشتیبانی محدود است یا زمانی که می‌تواند با مخلوطی خطی از چگالی‌های به‌طور نمایی کاهشی و چگالی‌های یکنواخت تقریبی شده است، مورداستفاده قرار گیرد.

یکی از مسیرهای جالب تحقیقات آینده، گسترش این کار به یک محیط چند متغیره است. نامساوی‌های مرتبط برای توابع تعداد زیادی از متغیرهای تصادفی یک موضوع تحقیقاتی فعال با کاربردهای متنوع در ابعاد بالا بوده است. آمار، یادگیری ماشین و نظریه اطلاعات. پیشرفت‌های اخیر در نابرابری‌های تمرکز و کاربردهای آن در نظریه اطلاعات را می‌توان در یک مونوگراف عالی یافت [۱۹]. در میان بسیاری از جنبه‌های جالب، ما به‌ویژه به یافتن شرایطی در توزیع چند متغیره برای استخراج نابرابری غلظت از نظر توان آنتروپی علاقه‌مندیم. هرگاه متغیرهای تصادفی با بردار تصادفی $k-$بعدی $\boldsymbol{X} = [X_1, X_2, \ldots, X_k]$ دوبه‌دو مجزا ترکیب و توزیع‌های حاشیه‌ای p_i از هر بردار تصادفی X_i تک‌مدی باشد و در شرایط بحث شده در اینجا صدق کند، واریانس $Var(X_i)$ هر متغیر تصادفی X_i از بالا به $Var(X_i) \leq c_i e^{2h(p_i)}$ برای ثابت $c_i > 0$ کران‌دار و تعیین‌کننده ماتریس کوواریانس $k \times k$، Σ برابر $|\Sigma| = \prod_{i=1}^{k} Var(X_i)$ است. به‌علاوه، آنتروپی توانی $e^{2h(X)}$ برابر $e^{2h(X)} = \prod_{i=1}^{k} e^{2h(p_i)}$ است. با استفاده از این حقایق، می‌توانیم داشته باشیم

$$|\Sigma| \leq c e^{2h(X)} \qquad (۳۱)$$

برای ثابت مثبت $c > 0$. بنابراین برای این حالت توزیع چندمتغیره با متغیرهای تصادفی مستقل در شرط لیپ‌شیتز-پیوسته صدق می‌کند و تک‌مدی است همان‌گونه که آنتروپی $h(\boldsymbol{X})$ از بردار تصادفی $\boldsymbol{X}$ به $-\infty$ میل می‌کند، تعیین‌کننده ماتریس کوواریانس است که از واریانس‌های توزیع‌های حاشیه‌ای تعیین می‌شود همگرا به 0 است و توزیع‌های چندمتغیره تصادفی با کران از ماتریس کوواریانس تعیین می‌شوند؛ اما توجه به این نکته مهم است که تعیین‌کننده ماتریس کوواریانس همیشه مرتبط با توزیع $k-$بعدی نیست هرگاه متغیرهای تصادفی دوبه‌دو مستقل نباشند. برای مثال، هرگاه $X_1 = X_2 = \ldots = X_k$ بیان‌کننده واریانس X_i باشد، تعیین‌کننده ماتریس کوواریانس برابر 0 است. بنابراین، یک سؤال باز جالب پیدا کردن یک متریک مناسب است که یک توزیع چند متغیره را نشان دهد و یک نابرابری مرتبط برای متریک برحسب توان آنتروپی برای یک رده مناسب از توزیع‌ها استخراج شود.

مراجع

[1] Chung, H. W., Sadler, B. M., Hero, A. O. (2015, September). Bounds on variance for symmetric unimodal distributions. *In 2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, (pp. 1235-1240). IEEE.

[2] Cover, T. M. (1999).*Elements of information theory*. John Wiley & Sons.

[3] Chen, L. H. (1982). An inequality for the multivariate normal distribution. *Journal of Multivariate Analysis*, **12(2)**, .306-315

[4] Chernoff, H. (1981). A note on an inequality involving the normal distribution. *The Annals of Probability*, **9(3)**, .533-535

[5] Cacoullos, T. (1982). On upper and lower bounds for the variance of a function of a random variable. *The Annals of Probability*, **10(3)**, .799-809

[6] Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal*, **27(3)**, .379-423

[7] Jedynak, B., Frazier, P. I., Sznitman, R. (2012). Twenty questions with noise: Bayes optimal policies for entropy loss. *Journal of Applied Probability*, **49(1)**, .114-136

[8] Wang, H., Yao, K., Pottie, G., Estrin, D. (2004, April). Entropy-based sensor selection heuristic for target localization. *In Proceedings of the 3rd international symposium on Information processing in sensor networks*, (pp. .(36-45

[9] Bakry, D., Barthe, F., Cattiaux, P., and Guillin, A. (2008). A simple proof of the Poincaré inequality for a large class of probability measures. *Electron. Commun. Probab.*, **13**, .60-66

[10] Bobkov, S. G. (1999). Isoperimetric and analytic inequalities for log-concave probability measures. *The Annals of Probability*, **27(4)**, .1903-1921

[11] Madiman, M., Melbourne, J., Xu, P. (2017). Forward and reverse entropy power inequalities in convex geometry. *In Convexity and concentration*, (pp. 427-485). New York, NY: Springer New York.

[12] Bobkov, S., Madiman, M. (2011). The entropy per coordinate of a random vector is highly constrained under convexity conditions. *IEEE Transactions on Information Theory*, **57(8)**, .4940-4954

[13] Hensley, D. (1980). Slicing convex bodies—bounds for slice area in terms of the body's covariance. *Proceedings of the American Mathematical Society*, **79(4)**, .619-625

[14] Webb, S. P. (1996). *Central slices of the regular simplex*. University of London, University College London (United Kingdom).

[15] Marsiglietti, A., Kostina, V. (2017, June). A lower bound on the differential entropy for log-concave random variables with applications to rate-distortion theory.*In 2017 IEEE International Symposium on Information Theory (ISIT)*, (pp. 46-50). IEEE.

[16] Walther, G. (2009). Inference and modeling with log-concave distributions. *Statistical Science*, **24(3)**, .319-327

[17] Bagnoli, M., Bergstrom, T. (2006). Log-concave probability and its applications. *In Rationality and Equilibrium: A Symposium in Honor of Marcel K. Richter*, (pp. 217-241). Springer Berlin Heidelberg.

[18] Saumard, A., Wellner, J. A. (2014). Log-concavity and strong log-concavity: a review. *Statistics surveys*, **8**, .45

[19] Raginsky, M., Sason, I. (2013). Concentration of measure inequalities in information theory, communications, and coding. *Foundations and Trends in Communications and Information Theory*, **10(1-2)**, .1-246

[20] Principe, J. C. (2010). *Information theoretic learning: Renyi's entropy and kernel perspectives*. Springer Science Business Media.

[21] Fano, R. M. (2008). Fano inequality. *Scholarpedia*, **3(10)**, .6648

[22] Alimohammadi, M., Alamatsaz, M. H. (2011). Properties of Unimodal Distributions and Some of Their Applications. *Andisheye Amari*, **16(1)**, .47-62

[23] Feller, W. (1971). *An introduction to probability theory and its applications*. John Wiley & Sons.

[24] Specht, W. (1960). Zur theorie der elementaren mittel. *Mathematische Zeitschrift*, **74(1)**, .91-98

[25] Mitrinovic, D. S., Vasic, P. M. (1970). *Analytic inequalities (Vol. 1)*. Berlin: Springer-verlag.

[26] Tsiligkaridis, T., Sadler, B. M., Hero, A. O. (2014). Collaborative 20 questions for target localization. *IEEE Transactions on Information Theory*, 60(4), 2233-2252.

[27] Travers, N. F. (2013). *Bounds on Convergence of Entropy Rate Approximations in Hidden Markov Processes*. University of California, Davis.

[28] Resnick, S. I. (2007). *Heavy-tail phenomena: probabilistic and statistical modeling*. Springer Science Business Media.

[29] Kwok, T. (2005). Power entropy and its applications. *IEEE Signal Processing Magazine*, **22(5)**, 151-15.

Determining the variance boundaries of single-mode distributions using power entropy

Fateme Ashtab[1] Dr. Mohammad Hossein Dehghan[2] Dr. Manizheh Sanei Tabas[3]

Abstract:

Variance and entropy are distinct metrics commonly used to measure the uncertainty of random variables. While variance shows how a random variable spreads more than expected, entropy measures the uncertainty of an information approach; in other words, it measures the average amount of information of a random variable. For both uniform and normal distributions, variance is a measure of power entropy. Finding such a monotonic relationship between variance and entropy for a larger class of these two distributions is very important and useful in signal processing, machine learning, information theory, probability, and statistics. For example, it is used to reduce the errors of estimators and choose a strategy that gives, on average, the greatest or nearly greatest reduction in the entropy of the distribution of the target location. The effectiveness of this method is tested using simulations with mining assay models. In this article, the upper bound of the variance for single-mode distributions, whose tails are heavier than the tails of exponential distributions, is created with the help of power entropy.

Keywords: Power entropy, Vriance bounds, unimodal distributions, Lipshitz continuity.

[1] **Master student, University of Sistan and Baluchestan**

[2] **Assistant Professor, Dept.of statistics, Faculty of Math, University of Sistan and Baluchestan, Daneshgah Ave., Zahedan, Iran.**

[3] **Assistant Professor, Dept.of statistics, Faculty of Math, University of Sistan and Baluchestan, Daneshgah Ave., Zahedan, Iran.**