مجله مدل‌سازی پیشرفته ریاضی
سال ۱۴۰۰، دوره ۱۱، شماره ۴، ص ۶۹۹-۷۱۱

شاپا چاپی: ۲۲۵۱-۸۰۸۸
DOI: 10.22055/JAMM.2021.38241.1954

اندازه‌های اطلاع جنسن-فیشر و جنسن-χ^2 برای توزیع‌های آمیخته متناهی

امید خوارزمی [۱] * ، مراد علیزاده [۲]

[۱] گروه آمار، دانشکده علوم ریاضی، دانشگاه ولی عصر (عج) رفسنجان، رفسنجان، ایران
[۲] گروه آمار، دانشکده مهندسی سیستم‌های هوشمند و علوم داده، دانشگاه خلیج فارس، بوشهر، ایران

دبیر مسئول: غلامعلی پرهام

تاریخ دریافت: ۱۴۰۰/۵/۲۲ تاریخ پذیرش: ۱۴۰۰/۹/۴

چکیده: در این مقاله، ابتدا با در نظر گرفتن اطلاع فیشر نوع پارامتری، یک اندازه اطلاع جدید براساس نامساوی جنسن معرفی می‌کنیم. سپس ماتریس اطلاع فیشر و اطلاع جنسن-فیشر را برای یک توزیع آمیخته متناهی از توابع چگالی احتمال مطالعه می‌شود. همچنین یک معیار اطلاع دیگر تحت عنوان جنسن- χ^2 بر اساس یک توزیع آمیخته از توابع چگالی احتمال معرفی می‌شود. تعمیم‌هایی از دو اندازه اطلاع جنسن-فیشر و جنسن-χ^2 براساس m تابع چگالی احتمال ارائه و ارتباط بین این دو معیار اطلاع جدید و همچنین ارتباط بین اطلاع جنسن-فیشر با برخی از معیارهای اطلاع شناخته شده از قبیل اطلاع جنسن-شانون و اطلاع جفریز مطالعه می‌شود.

واژه‌های کلیدی: اطلاع فیشر، تابع محدب، نامساوی جنسن، ماتریس اطلاع فیشر، اندازه اطلاع کای-دو.

رده‌بندی ریاضی: 62B10, 94A16

* نویسنده مسئول مقاله
رایانامه: omid.kharazmi@vru.ac.ir (O. Kharazmi),

۱ مقدمه

مباحث نظریه اطلاع دارای منشاءهای مختلفی از جمله نظریه ترمودینامیک و ارتباطات می‌باشد. امروزه این نظریه دارای کاربردهای فراوانی در زمینه‌های مختلف مانند فیزیک، مهندسی، زیست‌شناسی، پزشکی، اقتصاد، زبان‌شناسی و آمار است. در حقیقت نظریه‌ی اطلاع یکی از شاخه‌های نوین علمی است که در آن اطلاع از نقطه نظر ریاضی کمّی شده و مورد مطالعه قرار می‌گیرد. به‌طورکلی دو نوع مهم و پرکاربرد از معیارهای اطلاع آنتروپی شانون و اطلاع فیشر می‌باشد. به دلیل تنوع کاربرد این معیارهای اطلاع، در سه دهه‌ی اخیر محققان تعمیم‌های مختلفی از این معیارها ارائه کرده‌اند. شانون با انتشار مقاله‌ای در سال ۱۹۴۸ تحت عنوان نظریه‌ی ریاضی ارتباطات، با استفاده از مفهوم بی‌نظمی (آنتروپی) در مبحث ترمودینامیک و مکانیک آماری، اطلاع را به‌صورت کمّی تعریف کرد. شانون با بکارگیری مفاهیم احتمال بسیاری از سوالات مطرح شده در نظریه‌ی ارتباطات را پاسخ داد. امروزه از ایشان به عنوان پدر عصر رقمی (دیجیتال) یاد می‌شود (به مرجع [۱] مراجعه شود).

یکی از اندازه‌های اطلاع معروف و مرتبط با رفتارهای لحظه‌ای سیستم‌ها (ساختار درونی)، اطلاع فیشر است که به‌طورگسترده در علوم مختلف و بویژه فیزیک، آمار و مخابرات مورد توجه قرار گرفته است. این اندازه در سال ۱۹۲۹ توسط فیشر معرفی شده است. برای مطالعه‌ی جزئیات بیشتر به مراجع [۲] و [۵] مراجعه شود.

نامساوی‌های ریاضی نقش مهمی در رشته آمار و نظریه اطلاع دارند. برای مثال اندازه‌های اطلاع از نوع جنسن، نوعی از معیارهای اطلاع هستند که بر اساس ویژگی محدب (مقعر) یک اندازه اطلاع و با بکارگیری نامساوی جنسن ارائه شده‌اند. رائو [۳] یک رهیافت کلی برای معرفی اندازه‌های تعمیم یافته از آنتروپی شانون برای اندازه‌گیری فاصله بین دو مدل احتمال بر اساس نامساوی جنسن ارائه کرد. اخیراً خوارزمی و بالاکریشنان ([۴]، [۵]، [۶]) اندازه‌های اطلاع جدیدی از نوع جنسن در دو حالت مدل‌های احتمال گسسته و پیوسته ارائه کرده‌اند.

دو شکل متداول اطلاع فیشر عبارتند از اطلاع فیشر برای پارامتر مدل (اطلاع فیشر نوع پارامتری) و اطلاع فیشر برای تابع چگالی احتمال (اطلاع فیشر نوع تابع چگالی). اطلاع فیشر برای پارامتر مدل به‌صورت زیر تعریف می‌شود.

فرض کنید متغیر تصادفی X دارای مدل احتمال $f^\theta(x) = f(x|\theta)$ که در آن θ پارامتر مدل می‌باشد و $x \in \mathbb{R}$، $\theta \in \Theta \subseteq \mathbb{R}$. تحت فرض اینکه $f^\theta(x)$ نسبت به پارامتر θ مشتق‌پذیر باشد و به شرط برقراری شرایط نظم، میزان اطلاع فیشر در مورد پارامتر θ برای یک مشاهده تصادفی عبارت است از

$$\mathcal{I}(\theta) = \mathcal{I}\left(f^\theta\right) = \int_{\mathbb{R}} \left[\frac{\partial}{\partial\theta} \log f^\theta(x)\right]^{۲} f^\theta(x)dx. \qquad (۱.۱)$$

این کمیت علاوه بر کاربردهای گسترده‌اش در نظریه برآوردیابی آماری مانند استفاده در نامساوی کرامر–رائو، واریانس توزیع‌های مجانبی و حد اطلاع کولبک، به‌طورگسترده در دیگر شاخه‌های علم بویژه فیزیک مورد توجه قرار گرفته است [۷]. فرض کنید X یک متغیر تصادفی با تابع چگالی پیوسته f بر روی $\mathbb{R}$ باشد، اطلاع فیشر مربوط به تابع چگالی f عبارت است از

$$\mathcal{I}(f) = \int_{\mathbb{R}} \rho^{۲}(x)f(x)dx \qquad (۲.۱)$$

که در آن $\rho(x) = \frac{\partial}{\partial x}\log f(x)$، تابع امتیاز مربوط به چگالی f می‌باشد. کمیت $\mathcal{I}(f)$ همواره نامنفی است و نسبت به f محدب می‌باشد. با توجه به خاصیت محدب بودن اندازه اطلاع فیشر در رابطه (۲.۱) نسبت به تابع چگالی احتمال f، سانچز و همکاران [۸] اندازه اطلاع جنسن–فیشر دو تابع چگالی احتمال $f_۱$ و $f_۲$ (اطلاع جنسن–فیشر نوع تابع چگالی) برای $۰ \leq \alpha \leq ۱$ به‌صورت

$$\mathcal{JF}(f_۱, f_۲; \alpha) = \alpha\mathcal{I}(f_۱) + (۱-\alpha)\mathcal{I}(f_۲) - \mathcal{I}(\alpha f_۱ + (۱-\alpha)f_۲) \qquad (۳.۱)$$

تعریف کردند. هدف اصلی در این مقاله، معرفی دو اندازه اطلاع جدید در کلاس اندازه‌های اطلاع نوع جنسن و مطالعه‌ی نتایجی مرتبط با آن‌ها می‌باشد. ابتدا با در نظر گرفتن اطلاع فیشر نوع پارامتری در رابطه (۱.۱)، اندازه اطلاع جدید

جنسن-فیشر (JF) را معرفی می‌کنیم. یکی از ویژگی‌های مهم این معیار جدید اطلاع این است که اندازه اطلاع جنسن-فیشر نوع تابع چگالی در رابطه (۳.۱) را به عنوان حالت خاص وقتی که پارامتر θ به‌صورت پارامترمکان باشد؛ یعنی $f^{\theta}(x) = f(x-\theta)$، شامل می شود. یکی از اهداف مهم مرتبط با این معیار اطلاع بررسی ماتریس اطلاع فیشر و اندازه اطلاع جنسن-فیشر برای یک ترکیب محدب متناهی از توابع چگالی احتمال آمیخته می باشد. اندازه اطلاع جنسن-$\chi^{۲}$ (JC) یکی دیگر از دو اندازه اطلاع جدید است که در این مقاله معرفی می‌شود. همچنین علاقه‌مند به تعمیم این معیارهای اطلاع در حالت m تابع چگالی هستیم. مطالعه‌ی ارتباط بین دو معیار اطلاع JF و JC و همچنین ارتباط بین اطلاع JF با دیگر اندازه های اطلاع از قبیل جنسن-شانون و جفریز از دیگر اهداف این پژوهش می‌باشد. در ادامه معیارهای اطلاع مورد نیاز در این مقاله را به اختصار معرفی می‌کنیم.

فرض کنید X یک متغیر تصادفی پیوسته با تابع چگالی f بر روی $\mathbb{R}$ باشد. آنتروپی تفاضلی (آنتروپی شانون) f به‌صورت زیر تعریف می‌شود

$$H(f) = -\int_{\mathbb{R}} f(x)\log f(x)dx. \tag{۴.۱}$$

این معیار برای اندازه‌گیری میزان عدم حتمیت X بکار می‌رود. به عبارتی دیگر، $H(f)$ میزان تخت بودن و یا قله بودن f را اندازه‌گیری می‌کند. علی‌رغم کاربردهای فراوان آن، از نواقص این معیار آن است که مقدار آن ممکن است منفی باشد و در حالت کلی تعریف آن به وجود f بستگی دارد.

فرض کنید X و Y دو متغیر تصادفی با توابع چگالی پیوسته به‌ترتیب f و g بر روی $\mathbb{R}$ باشند، آنگاه اندازه اطلاع کولبک-لیبلر بین f و g عبارت است از

$$K(f,g) = \int_{\mathbb{R}} f(x)\log\frac{f(x)}{g(x)}dx. \tag{۵.۱}$$

از مهمترین ویژگی‌های اطلاع کولبک-لیبلر این است که همواره نامنفی است و با استفاده از آن می‌توان میزان مشابهت بین دو مدل f و g را بررسی کرد.

یکی از مهمترین اندازه‌های اطلاع، اطلاع جنسن-شانون است و برای دو تابع چگالی $f_{۱}$ و $f_{۲}$ برای $۰ \leq \alpha \leq ۱$ به‌صورت زیر تعریف می‌شود

$$JS(f_{۱}, f_{۲}; \alpha) = H(\alpha f_{۱} + (۱-\alpha)f_{۲}) - \{\alpha H(f_{۱}) + (۱-\alpha)H(f_{۲})\} \tag{۶.۱}$$

که در آن H آنتروپی شانون می‌باشد (به [۹] مراجعه شود). در حقیقت این اندازه اطلاع بر اساس ویژگی مقعر بودن آنتروپی شانون نسبت به تابع چگالی تعریف می‌شود. از مهمترین ویژگی‌های اطلاع جنسن-شانون این است که متریک می‌باشد و همیشه متناهی است، در صورتی که معیار اطلاع کولبک-لیبلر نامتقارن است و متریک نمی‌باشد. علاوه براین با توجه به رابطه (۶.۱) هرچه توابع چگالی مؤلفه‌ها از میانگین دو تابع چگالی دورتر باشند، مشابهت بین دو مدل کمتر می‌شود و برعکس. اندازه اطلاع جنسن-شانون از مهمترین معیارها برای بررسی مشابهت بین دو مدل احتمال است.

فرض کنید X و Y دو متغیر تصادفی با توابع چگالی پیوسته به‌ترتیب f و g باشند، آنگاه فاصله اطلاع کای-دو بین f و g عبارت است از

$$\chi^{۲}(f,g) = \int_{\mathbb{R}} \frac{[f(x)-g(x)]^{۲}}{f(x)}dx. \tag{۷.۱}$$

اندازه اطلاع کای-دو یکی از معیارهای مهم در شاخه آمار و نظریه اطلاع است و دارای کاربردهای زیادی در زمینه آزمون فرضیه‌های آمار پارامتری و ناپارامتری است.

بر این اساس ساختار مقاله به صورت زیر خواهد بود: ابتدا در بخش ۲ با در نظر گرفتن اطلاع فیشر نوع پارامتری، و بر اساس خاصیت محدب این اندازه اطلاع، اطلاع جدید جنسن-فیشر نوع پارامتری با استفاده از نامساوی جنسن معرفی می‌شود. سپس نشان می‌دهیم که اطلاع جدید معرفی شده را می‌توان به‌صورت یک ترکیب محدب از فاصله‌های اطلاع فیشر ارائه کرد. در بخش ۳، اندازه اطلاع جنسن-فیشر و ماتریس اطلاع فیشر را برای یک ترکیب محدب از توابع

چگالی احتمال محاسبه می کنیم و نشان می‌دهیم که ماتریس اطلاع فیشر برای بردار پارامتر یک توزیع آمیخته را می‌توان بر اساس اندازه کای-دو بیان کرد. علاوه براین در این بخش اندازه اطلاع جدید جنسن-$\chi^۲$ بین دوتابع چگالی آمیخته را معرفی می‌کنیم. تعمیم دو اطلاع جنسن-فیشر و جنسن-$\chi^۲$ و ارتباط بین آن‌ها در بخش ۴ ارائه می‌شود. همچنین در این بخش نشان می‌دهیم که اطلاع جنسن-فیشر بیزی تحت توزیع پیشین یکنواخت دارای ارتباط مستقیم با برخی از اندازه‌های اطلاع از قبیل جنسن-شانون و جفریز می‌باشد. در پایان، مقاله در بخش ۵ نتیجه‌گیری می‌شود.

۲ اندازه اطلاع جنسن-فیشر نوع پارامتری

در این بخش، ابتدا با در نظر گرفتن اطلاع فیشر نوع پارامتری، در رابطه (۱.۱)، اندازه اطلاع جنسن-فیشر مرتبط با آن معرفی می‌شود. سپس نشان می دهیم که این اندازه اطلاع جدید را می توان به‌صورت یک ترکیب محدب از فواصل اطلاع فیشر بیان کرد.

تعریف ۱.۲. فرض کنید $X_۱$،...،X_n متغیرهای تصادفی پیوسته و دارای توابع چگالی به ترتیب $f_۱^\theta$،...،f_n^θ باشند و همچنین $\alpha_۱, ..., \alpha_n$ اعداد حقیقی مقدار باشند بطوری که $\sum_{i=۱}^n \alpha_i = ۱$. آنگاه اندازه اطلاع جنسن-فیشر نوع پارامتری به صورت

$$JFI(f_۱^\theta, ..., f_n^\theta, \boldsymbol{\alpha}) \;=\; \sum_{i=۱}^n \alpha_i \mathcal{I}(f_i^\theta) - \mathcal{I}\left(\sum_{i=۱}^n \alpha_i f_i^\theta\right) \qquad (۱.۲)$$

تعریف می شود.

قضیه ۲.۲. فرض کنید متغیر های تصادفی $X_۱, ..., X_n$ دارای توابع چگالی احتمال به ترتیب $f_۱^\theta, ..., f_n^\theta$ باشند، آنگاه اطلاع جنسن-فیشر را می‌توان به‌صورت یک ترکیب آمیخته‌ای از فاصله‌های اطلاع فیشر به‌صورت زیر بیان کرد:

$$JFI(f_۱^\theta, ..., f_n^\theta, \boldsymbol{\alpha}) = \sum_{i=۱}^n \alpha_i D(f_i^\theta, f_T^\theta),$$

که در آن $f_T^\theta = \sum_{i=۱}^n \alpha_i f_i^\theta$ تابع چگالی احتمال آمیخته می باشد، همچنین $D(f_i^\theta, f_T^\theta)$ فاصله اطلاع فیشر بین دو تابع چگالی f_i^θ و f_T^θ است و به‌صورت

$$D(f_i^\theta, f_T^\theta) \;=\; \int_۰^\infty \left\{\frac{\partial \log f_i^\theta(x)}{\partial\theta} - \frac{\partial \log \sum_{i=۱}^n \alpha_i f_i^\theta(x)}{\partial\theta}\right\}^۲ f_i^\theta(x)dx$$

تعریف می‌شود.

اثبات. با توجه به رابطه (۱.۲)، خواهیم داشت

$$\begin{aligned} JFI(f_۱^\theta, ..., f_n^\theta, \boldsymbol{\alpha}) \;&=\; \sum_{i=۱}^n \alpha_i \mathcal{I}(f_i^\theta) - \mathcal{I}\left(\sum_{i=۱}^n \alpha_i f_i^\theta\right) \\ &=\; \sum_{i=۱}^n \alpha_i \int_۰^\infty \left\{\frac{\partial \log f_i^\theta(x)}{\partial\theta}\right\}^۲ f_i^\theta(x)dx \\ &\quad - \int_۰^\infty \left\{\frac{\partial \log \sum_{i=۱}^n \alpha_i f_i^\theta(x)}{\partial\theta}\right\}^۲ \sum_{i=۱}^n \alpha_i f_i^\theta(x)dx. \end{aligned}$$

از طرف دیگر داریم

$$\begin{aligned}\sum_{i=۱}^{n}\alpha_i D(f_i^\theta, f_T^\theta) &= \sum_{i=۱}^{n}\alpha_i\int_{۰}^{\infty}\left\{\frac{\partial \log f_i^\theta(x)}{\partial\theta}-\frac{\partial\log\sum_{i=۱}^{n}\alpha_i f_i^\theta(x)}{\partial\theta}\right\}^{۲} f_i^\theta(x)dx\\ &= \sum_{i=۱}^{n}\alpha_i\int_{۰}^{\infty}\left\{\left(\frac{\partial \log f_i^\theta(x)}{\partial\theta}\right)^{۲}-۲\left(\frac{\partial \log f_i^\theta(x)}{\partial\theta}\right)\frac{\partial\log\sum_{i=۱}^{n}\alpha_i f_i^\theta(x)}{\partial\theta}\right.\\ &\quad\left.+\left(\frac{\partial\log\sum_{i=۱}^{n}\alpha_i f_i^\theta(x)}{\partial\theta}\right)^{۲}\right\}f_i^\theta(x)dx\\ &= \sum_{i=۱}^{n}\alpha_i\int_{۰}^{\infty}\left\{\frac{\partial \log f_i^\theta(x)}{\partial\theta}\right\}^{۲} f_i^\theta(x)dx\\ &\quad-\int_{۰}^{\infty}\left\{\frac{\partial\log\sum_{i=۱}^{n}\alpha_i f_i^\theta(x)}{\partial\theta}\right\}^{۲}\sum_{i=۱}^{n}\alpha_i f_i^\theta(x)dx,\end{aligned}$$

با توجه به برابری دو رابطه بالا اثبات کامل می‌شود.

$\square$

۳ ماتریس اطلاع فیشر و اندازه اطلاع جنسن-فیشر برای یک توزیع آمیخته متناهی

در این بخش، ماتریس اطلاع فیشر برای بردار پارامتر آمیخته یک تابع چگالی آمیخته و اندازه اطلاع جنسن-فیشر بین دوتابع چگالی و همچنین تعریف اندازه اطلاع جنسن-$\chi^{۲}$ بین دو تابع چگالی آمیخته ارائه می‌شود.

فرض کنید $f_۰$، $f_۱$، $g_۰$ و $g_۱$ توابع چگالی احتمال پیوسته دلخواه باشند. تابع چگالی آمیخته f_θ، با بردار پارامتر $\theta=(\alpha,\Lambda)$، $۰\le\alpha\le۱$، $۰\le\Lambda\le۱$، بر پایه ترکیبات خطی از توابع چگالی $f_۰$، $f_۱$، $g_۰$ و $g_۱$ به‌صورت زیر در نظر بگیرید

$$f_\theta(x)=\alpha f_\Lambda(x)+(۱-\alpha)g_\Lambda(x), \tag{۱.۳}$$

که در آن f^Λ و g^Λ توابع چگالی آمیخته و به‌صورت

$$f^\Lambda(x)=\Lambda f_۰(x)+(۱-\Lambda)f_۱(x), \tag{۲.۳}$$

$$g^\Lambda(x)=\Lambda g_۰(x)+(۱-\Lambda)g_۱(x), \tag{۳.۳}$$

تعریف شده اند.

۱.۳ ماتریس اطلاع فیشر برای پارامتر آمیخته یک توزیع آمیخته متناهی

در قضیه زیر ماتریس اطلاع فیشر برای بردار پارامتر $\theta=(\alpha,\Lambda)$ ارائه شده است.

قضیه ۱.۳. فرض کنید X یک متغیر تصادفی پیوسته و دارای تابع چگالی احتمال به صورت (۱.۳) باشد، آنگاه ماتریس اطلاع فیشر برای بردار پارامتر $\theta=(\alpha,\Lambda)$ به‌صورت زیر است

$$\begin{aligned}\mathcal{I} &= \begin{bmatrix}\mathcal{I}(\alpha) & \mathcal{I}(\alpha,\Lambda)\\ \mathcal{I}(\Lambda,\alpha) & \mathcal{I}(\alpha)\end{bmatrix}\\ &= \begin{bmatrix}\frac{۱}{\alpha^{۲}}\chi^{۲}(f_{\boldsymbol\theta},g^\Lambda) & E_{f_۰}\left[\frac{g_۱(X)}{f_\theta(X)}\right]-E_{f_۱}\left[\frac{g_۰(X)}{f_\theta(X)}\right]\\ E_{f_۰}\left[\frac{g_۱(X)}{f_\theta(X)}\right]-E_{f_۱}\left[\frac{g_۰(X)}{f_\theta(X)}\right] & \frac{۱}{\Lambda^{۲}}\chi^{۲}(f_{\boldsymbol\theta},m_\alpha)\end{bmatrix},\end{aligned}$$

که در آن $m_\alpha(x)=\alpha f_1(x)+(1-\alpha)g_1$،

$$\mathcal{I}(\alpha)=\int\left(\frac{\partial}{\partial\alpha}\log f_\theta(x)\right)^2 f_\theta(x)dx,$$

$$\mathcal{I}(\Lambda)=\int\left(\frac{\partial}{\partial\Lambda}\log f_\theta(x)\right)^2 f_\theta(x)dx$$

و

$$\mathcal{I}(\alpha,\Lambda)=\mathcal{I}(\Lambda,\alpha)=\int\left(\frac{\partial^2}{\partial\Lambda\partial\alpha}\log f_\theta(x)\right)f_\theta(x)dx.$$

اثبات. طبق تعریف اطلاع فیشر و در نظر گرفتن تابع چگالی احتمال (۱.۳) و برابری

$$f_\theta(x)-g^\Lambda(x)=\alpha\Big(f^\Lambda(x)-g^\Lambda(x)\Big)$$

داریم

$$\begin{aligned}
\mathcal{I}_{11} &= \mathcal{I}(\alpha)\\
&= \int\left(\frac{\partial}{\partial\alpha}\log f_\theta(x)\right)^2 f_\theta(x)dx\\
&= \int\frac{\big(f^\Lambda(x)-g^\Lambda(x)\big)^2}{f_\theta(x)}dx\\
&= \frac{1}{\alpha^2}\int\frac{\big(f_\theta(x)-g^\Lambda(x)\big)^2}{f_\theta(x)}dx\\
&= \frac{1}{\alpha^2}\chi^2\big(f_{\boldsymbol\theta},g^\Lambda\big).
\end{aligned}$$

به‌طور مشابه و طبق فرض $m_\alpha(x)=\alpha f_1(x)+(1-\alpha)g_1$، اطلاع فیشر پارامتر Λ را می‌توان به‌صورت زیر محاسبه کرد

$$\begin{aligned}
\mathcal{I}_{22} &= \mathcal{I}(\Lambda)\\
&= \int\left(\frac{\partial}{\partial\Lambda}\log f_\theta(x)\right)^2 f_\theta(x)dx\\
&= \int\frac{\big\{\alpha(f_0(x)-f_1(x))+(1-\alpha)(g_0(x)-g_1(x))\big\}^2}{f_\theta(x)}dx\\
&= \frac{1}{\Lambda^2}\int\frac{\big(f_\theta(x)-m_\alpha(x)\big)^2}{f_\theta(x)}dx\\
&= \frac{1}{\Lambda^2}\chi^2\big(f_{\boldsymbol\theta},m_\alpha\big).
\end{aligned}$$

با توجه به تعریف اطلاع فیشر توأم برای پارامترهای α و Λ و برابری $\mathcal{I}_{12}=\mathcal{I}_{21}=\mathcal{I}(\alpha,\Lambda)$، کافی است تا نشان دهیم که

$$\mathcal{I}(\alpha,\Lambda)=E_{f_0}\Big[\frac{g_1(X)}{f_\theta(X)}\Big]-E_{f_1}\Big[\frac{g_0(X)}{f_\theta(X)}\Big].$$

بنابراین داریم

$$\begin{aligned}
\mathcal{I}_{۱۲} &= \mathcal{I}(\alpha,\Lambda)\\
&= \int \left(\frac{\partial^{۲}}{\partial\Lambda\partial\alpha}\log f_{\theta}(x)\right) f_{\theta}(x)dx\\
&= \int \left\{\frac{\partial}{\partial\Lambda}\frac{f^{\Lambda}(x)-g^{\Lambda}(x)}{f_{\theta}(x)}\right\} f_{\theta}(x)dx\\
&= \int \frac{f_{۰}(x)g_{۱}(x)-f_{۱}(x)g_{۰}(x)}{f_{\theta}(x)}dx\\
&= \int \frac{g_{۱}(x)}{f_{\theta}(x)}f_{۰}(x)dx - \int \frac{g_{۰}(x)}{f_{\theta}(x)}f_{۱}(x)dx\\
&= E_{f_{۰}}[\frac{g_{۱}(X)}{f_{\theta}(X)}] - E_{f_{۱}}[\frac{g_{۰}(X)}{f_{\theta}(X)}],
\end{aligned}$$

اثبات قضیه کامل است.

□

۲.۳ اندازه‌های اطلاع جنسن-فیشر و جنسن-$\chi^{۲}$ بین دو توزیع آمیخته متناهی

فرض کنید f^{Λ} و g^{Λ} به‌ترتیب توابع چگالی احتمال به‌صورت (۲.۳) و (۳.۳) باشند. آنگاه اطلاع جنسن-فیشر (JF) مرتبط با پارامتر آمیخته Λ، برای مقادیر $۰ \leq \alpha \leq ۱$، به‌صورت

$$JF(f^{\Lambda},g^{\Lambda};\alpha)=\alpha\mathcal{I}(f^{\Lambda})+(۱-\alpha)\mathcal{I}(g^{\Lambda})-\mathcal{I}(h^{\Lambda}_{\alpha}), \tag{۴.۳}$$

تعریف می شود، که در آن

$$\mathcal{I}(f^{\Lambda})=\int\left(\frac{\partial}{\partial\Lambda}\log f^{\Lambda}(x)\right)^{۲} f^{\Lambda}(x)dx,\ \mathcal{I}(g^{\Lambda})=\int\left(\frac{\partial}{\partial\Lambda}\log g^{\Lambda}(x)\right)^{۲} g^{\Lambda}(x)dx,$$

$$\mathcal{I}(h^{\Lambda}_{\alpha})=\int\left(\frac{\partial}{\partial\Lambda}\log h^{\Lambda}_{\alpha}(x)\right)^{۲} h^{\Lambda}_{\alpha}(x)dx$$

و

$$h^{\Lambda}_{\alpha}(x)=\alpha f^{\Lambda}(x)+(۱-\alpha)g^{\Lambda}(x). \tag{۵.۳}$$

تعریف ۲.۳. توابع چگالی احتمال $f_{۰}, f_{۱}, g_{۰}$ و $g_{۱}$ را در نظر بگیرید، برای $\Lambda\in(۰,۱)$، فرض کنید f_{Λ} و g_{Λ} به ترتیب به‌صورت روابط (۲.۳) و (۳.۳) تعریف شده باشند. آنگاه اندازه اطلاع جنسن-$\chi^{۲}$ (JC) بین دو تابع چگالی احتمال f^{Λ} و g^{Λ} برای $\alpha\in(۰,۱)$ به‌صورت زیر تعریف می‌شود

$$JC(f^{\Lambda},g^{\Lambda};\alpha)=\alpha\chi^{۲}(f^{\Lambda},f_{۱})+(۱-\alpha)\chi^{۲}(g^{\Lambda},g_{۱})-\chi^{۲}(h^{\Lambda}_{\alpha},\alpha f_{۱}+(۱-\alpha)g_{۱}), \tag{۶.۳}$$

که در آن تابع چگالی h^{Λ}_{α} به‌صورت رابطه (۵.۳) تعریف شده است.

۴ تعمیم اندازه های اطلاع JF و JC و ارتباط آن‌ها با برخی از معیارهای شناخته شده اطلاع

در این بخش، ابتدا تعمیم‌هایی از تعاریف اندازه‌های اطلاع جنسن-فیشر در رابطه (۴.۳) و جنسن-$\chi^۲$ در رابطه (۶.۳) در حالت m تابع چگالی احتمال ارائه می‌شود. سپس ارتباط بین دو اندازه اطلاع جدید، بررسی اطلاع جنسن-فیشر بیزی تحت توزیع پیشین یکنواخت و ارتباط آن با اندازه‌های اطلاع جفریز و جنسن-شانون مطالعه می‌شود.

فرض کنید $F=\{f_{ij}\}_{i=۱,...,m,\ j=۱,...,n}$ یک ماتریس $m\times n$ از تابع چگالی احتمال پیوسته باشد و f_i^{Λ} و $f_i^{\Lambda-j}$ توابع چگالی آمیخته از درایه‌های ماتریس F برای مقادیر $n\geq ۲$، به‌صورت

$$f_i^{\Lambda}(x)=\frac{۱}{n-۱}\sum_{j=۱}^{n-۱}\Lambda_k f_{ik}(x)+\left(۱-\frac{\sum_{k=۱}^{n-۱}\Lambda_k}{n-۱}\right)f_{in}(x),\ i=۱,...,m, \tag{۱.۴}$$

و

$$f_i^{\Lambda-j}(x)=\frac{n-۲}{n-۱}f_{ij}+\frac{۱}{n-۱}\sum_{k=۱,k\neq j}^{n-۱}\Lambda_k f_{ik}(x)+\frac{۱}{n-۱}\left(۱-\sum_{k=۱,k\neq j}^{n-۱}\Lambda_k\right)f_{in}(x), \tag{۲.۴}$$

تعریف شده باشند که در آن $۰\leq \Lambda_i\leq ۱,\ i=۱,...,n-۱$ و $\sum_{i=۱}^{n-۱}\Lambda_i\leq ۱$. آنگاه اندازه اطلاع جنسن-$\chi^۲$ بر اساس m توابع چگالی احتمال $f_i^{\Lambda}\ (i=۱,...,m)$ به‌صورت

$$JCD(f_۱^{\Lambda},...,f_m^{\Lambda},\boldsymbol{\alpha})=\sum_{i=۱}^{m}\alpha_i\,\chi^۲(f_i^{\Lambda},f_i^{\Lambda-j})-\chi^۲\left(\sum_{i=۱}^{m}\alpha_i f_i^{\Lambda},\sum_{i=۱}^{m}\alpha_i f_i^{\Lambda-j}\right) \tag{۳.۴}$$

تعریف می‌شود که در آن $\alpha_۱,...,\alpha_n$ مقادیر حقیقی و نامنفی هستند و $\sum_{i=۱}^{n}\alpha_i=۱$. علاوه بر این اندازه اطلاع جنسن-فیشر بر اساس m تابع چگالی احتمال $f_i^{\Lambda}\ (i=۱,...,m)$ برای هر $j=۱,...,n-۱$ به‌صورت زیر تعریف می‌شود

$$JF_{\Lambda_j}(f_۱^{\Lambda},...,f_m^{\Lambda}:\boldsymbol{\alpha})\ =\ \sum_{i=۱}^{m}\alpha_i\mathcal{I}_{\Lambda_j}(f_i^{\Lambda})-\mathcal{I}_{\Lambda_j}\big(\sum_{i=۱}^{m}\alpha_i f_i^{\Lambda}\big). \tag{۴.۴}$$

قضیه ۱.۴ فرض کنید $f_۱^{\Lambda}$، ...، f_m^{Λ} توابع چگالی احتمال به‌صورت رابطه (۱.۴) باشند، آنگاه رابطه‌ی بین دو اندازه اطلاع جنسن-$\chi^۲$ در (۳.۴) و جنسن-فیشر در (۴.۴)، برای مقادیر $j=۱,...,n-۱$ به‌صورت

$$JF_{\Lambda_j}(f_۱^{\Lambda},...,f_m^{\Lambda}:\boldsymbol{\alpha})\ =\ \frac{۱}{(\Lambda_j-(n-۱))^۲}JC_{\Lambda_j}(f_۱^{\Lambda},...,f_m^{\Lambda}:\boldsymbol{\alpha}) \tag{۵.۴}$$

می‌باشد.

اثبات. طبق تعریف اطلاع جنسن-فیشر برای $j=۱,...,n-۱$ و با توجه به این نکته که

$$f_i^{\Lambda}(x)-f_i^{\Lambda-j}(x)=\frac{n-۲-\Lambda_j}{n-۱}\Big(f_{ij}(x)-f_{in}(x)\Big),$$

داریم

$$\begin{aligned}
JF_{\Lambda_j}(f_{۱}^{\Lambda},...,f_m^{\Lambda}:\boldsymbol{\alpha}) &= \sum_{i=۱}^{m}\alpha_i\mathcal{I}_{\Lambda_j}(f_i^{\Lambda})-\mathcal{I}_{\Lambda_j}\big(\sum_{i=۱}^{m}\alpha_i f_i^{\Lambda}\big)\\
&= \sum_{i=۱}^{m}\alpha_i\int\left(\frac{\partial}{\partial\Lambda_j}\log f_j^{\Lambda}(x)\right)^{۲}f_j^{\Lambda}(x)dx\\
&\quad -\int\left(\frac{\partial}{\partial\Lambda_j}\log\sum_{i=۱}^{m}\alpha_i f_i^{\Lambda}(x)\right)^{۲}\sum_{i=۱}^{m}\alpha_i f_i^{\Lambda}(x)dx\\
&= \sum_{i=۱}^{m}\alpha_i\int\frac{۱}{(n-۱)^{۲}}\frac{\big(f_{ij}(x)-f_{in}(x)\big)^{۲}}{f_i^{\Lambda}(x)}dx\\
&\quad -\int\frac{۱}{(n-۱)^{۲}}\frac{\left(\sum_{i=۱}^{m}\alpha_i\big(f_{ij}(x)-f_{in}(x)\big)\right)^{۲}}{\sum_{i=۱}^{m}\alpha_i f_i^{\Lambda}(x)}dx\\
&= \sum_{i=۱}^{m}\alpha_i\int\frac{۱}{(n-۲-\Lambda_j)^{۲}}\frac{\big(f_i^{\Lambda}(x)-f_i^{\Lambda-j}(x)\big)^{۲}}{f_i^{\Lambda}(x)}dx\\
&\quad -\int\frac{۱}{(n-۲-\Lambda_j)^{۲}}\frac{\left(\sum_{i=۱}^{m}\alpha_i\big(f_i^{\Lambda}(x)-f_i^{\Lambda-j}(x)\big)\right)^{۲}}{\sum_{i=۱}^{m}\alpha_i f_i^{\Lambda}(x)}dx\\
&= \frac{۱}{(n-۲-\Lambda_j)^{۲}}\sum_{i=۱}^{m}\alpha_i\,\chi^{۲}(f_i^{\Lambda},f_i^{\Lambda-j})\\
&\quad -\frac{۱}{(n-۲-\Lambda_j)^{۲}}\chi^{۲}\left(\sum_{i=۱}^{m}\alpha_i f_i^{\Lambda},\sum_{i=۱}^{m}\alpha_i f_i^{\Lambda-j}\right)\\
&= \frac{۱}{(\Lambda_j-(n-۱))^{۲}}JC_{\Lambda_j}(f_{۱}^{\Lambda},...,f_m^{\Lambda}:\boldsymbol{\alpha}),
\end{aligned}$$

اثبات کامل است. □

در ادامه این بخش، با توجه به مفهوم اطلاع فیشر بیزی (مرجع [۶] را ببینید)، اطلاع جنسن-فیشر بیزی برای اندازه اطلاع رابطه (۴.۳)، تحت تابع پیشین یکنواخت برای پارامتر Λ محاسبه و ارتباط آن با اندازه‌های اطلاع جفریز و جنسن-شانون نشان داده می‌شود.

قضیه ۲.۴. ارتباط بین اطلاع جنسن-فیشر بیزی و اندازه اطلاع جفریز، تحت توزیع پیشین $\pi(\Lambda)=\frac{۱}{s-r}$ برای مقادیر حقیقی s و r که در آن $۰\leq r<s\leq ۱$، به‌صورت

$$\int_r^s JF(f^{\Lambda},g^{\Lambda},\alpha)\pi(\Lambda)d\Lambda = \frac{۱}{(s-r)^{۲}}\Big\{\alpha J(f^s,f^r)+(۱-\alpha)J(g^s,g^r)-J(h_{\alpha}^s,h_{\alpha}^r)\Big\}, \tag{۶.۴}$$

می‌باشد که در آن

$$h_{\alpha}^s(x)=\alpha f^s(x)+(۱-\alpha)g^s(x),\quad h_{\alpha}^r(x)=\alpha f^r(x)+(۱-\alpha)g^r(x),$$

و $f^s(f^r)$ و $g^s(g^r)$ توابع چگالی به‌ترتیب به شکل‌های (۲.۳) و (۳.۳) می‌باشند. همچنین J اندازه اطلاع جفریز بین دو تابع چگالی احتمال است و به‌صورت $J(f,g)=KL(f,g)+KL(g,f)$ تعریف می‌شود.

اثبات. ابتدا با مقداری محاسبات جبری می‌توان نشان داد که

$$(s-r)\int_r^s \mathcal{I}(f^\Lambda)d\Lambda = J(f^s, f^r), \quad (s-r)\int_r^s \mathcal{I}(g^\Lambda)d\Lambda = J(g^s, g^r).$$

حال با استفاده از تعریف اندازه اطلاع جنسن-فیشر (۴.۳) و در نظرگرفتن $K=(s-r)\int_s^r JF(f^\Lambda, g^\Lambda;\alpha)d\Lambda$، داریم

$$\begin{aligned} K &= (s-r)\Big\{\alpha\int_r^s \mathcal{I}(f^\Lambda)d\Lambda + (۱-\alpha)\int_r^s \mathcal{I}(g^\Lambda)d\Lambda\Big\} - (s-r)\int_r^s I(h_\alpha^\Lambda)d\Lambda \\ &= \alpha J(f^s, f^r) + (۱-\alpha)J(g^s, g^r) - (s-r)\int_r^s I(h_\alpha^\Lambda)d\Lambda. \end{aligned}$$

اکنون برای کامل شدن اثبات کافی است تا نشان دهیم که

$$(s-r)\int_r^s \mathcal{I}(h_\alpha^\Lambda)d\lambda = J(h_\alpha^s, h_\alpha^r).$$

از طرفی با توجه به تعریف تابع چگالی h_α^Λ داریم

$$h_\alpha^\Lambda(x) = \alpha\Big(\Lambda f_۰(x) + (۱-\Lambda)f_۱(x)\Big) + (۱-\alpha)\Big(\Lambda g_۰(x) + (۱-\Lambda)g_۱(x)\Big).$$

قرار دهید

$$Z_\alpha(x) = \alpha\Big(f_۰(x) - f_۱(x)\Big) + (۱-\alpha)\Big(g_۰(x) - g_۱(x)\Big)$$

بنابراین نتیجه می‌شود که $\frac{\partial}{\partial\Lambda}\log(h_\alpha^\Lambda) = Z_\alpha(x)$. همچنین بدست می‌آوریم که

$$\begin{aligned} (s-r)Z_\alpha(x) &= (s-r)\Big\{\alpha\big(f_۰(x) - f_۱(x)\big) + (۱-\alpha)\big(g_۰(x) - g_۱(x)\big)\Big\} \\ &= \alpha\Big(sf_۰(x) + (۱-s)f_۱(x)\Big) + (۱-\alpha)\Big(sg_۰(x) + (۱-s)g_۱(x)\Big) \\ &- \alpha\Big(rf_۰(x) + (۱-r)f_۱(x)\Big) + (۱-\alpha)\Big(rg_۰(x) + (۱-r)g_۱(x)\Big) \\ &= \alpha f_s(x) + (۱-\alpha)g_s(x) - \Big(\alpha f_r(x) + (۱-\alpha)g_r(x)\Big) \\ &= h_\alpha^s(x) - h_\alpha^r(x). \end{aligned}$$

اکنون با توجه به نتایج بالا و استفاده از قضیه فوبینی داریم

$$(s-r)\int_r^s \mathcal{I}(h_\alpha^\Lambda)d\Lambda = (s-r)\int Z_\alpha(x)\Big\{\int_r^s \frac{Z_\alpha(x)}{h_\alpha^\Lambda(x)}d\Lambda\Big\}dx. \quad (۷.۴)$$

انتگرال داخلی رابطه (۷.۴) نتیجه می‌دهد که

$$\begin{aligned} \int_r^s \frac{Z_\alpha(x)}{h_\alpha^\Lambda(x)}d\Lambda dx &= \int_r^s \frac{\alpha(f_۰(x) - f_۱(x)) + (۱-\alpha)(g_۰(x) - g_۱(x))}{h_\alpha^\Lambda(x)}d\Lambda \\ &= \log\left\{\frac{\alpha f_s(x) + (۱-\alpha)g_s(x)}{\alpha f_r(x) + (۱-\alpha)g_r(x)}\right\} \\ &= \log\left\{\frac{h_\alpha^s(x)}{h_\alpha^r(x)}\right\}. \end{aligned}$$

(۸.۴)

در نهایت با جایگذاری رابطه (۸.۴) در رابطه (۷.۴)، بدست می‌آوریم که

$$\begin{aligned}
(s-r)\int_r^s \mathcal{I}(h_\alpha^\Lambda)d\Lambda &= (s-r)\int Z_\alpha(x)\log\left\{\frac{h_\alpha^s(x)}{h_\alpha^r(x)}\right\}dx \\
&= \int (s-r)Z_\alpha(x)\log\left\{\frac{h_\alpha^s(x)}{h_\alpha^r(x)}\right\}dx \\
&= \int \left(h_\alpha^s(x)-h_\alpha^r(x)\right)\log\left\{\frac{h_\alpha^s(x)}{h_\alpha^r(x)}\right\}dx \\
&= KL(h_\alpha^s,h_\alpha^r)+KL(h_\alpha^r,h_\alpha^s) \\
&= J(h_\alpha^s,h_\alpha^r),
\end{aligned}$$

اثبات کامل است. □

نتیجه ۳.۴. با توجه به قضیه ۲.۴ برای حالت خاص وقتی که $s=1$ و $r=0$، داریم

$$\int_0^1 JF(f^\Lambda,g^\Lambda;\alpha)d\Lambda = \alpha J(f_0,f_1)+(1-\alpha)J(g_0,g_1)-J(m_\alpha,n_\alpha),$$

که در آن $m_\alpha(x)=\alpha f_0(x)+(1-\alpha)g_0(x)$ و $n_\alpha(x)=\alpha f_1(x)+(1-\alpha)g_1(x)$.

با توجه به نتیجه ۳.۴ یک نمایش دیگر برای اندازه اطلاع جنسن-فیشر بیزی (۴.۳)، بر اساس اطلاع جنسن-شانون و اندازه اطلاع نادرستی کریج به‌صورت

$$\begin{aligned}
\int_0^1 JF(f_\Lambda,g_\Lambda,\alpha)d\Lambda &= JS(f_0,g_0;\alpha)+JS(f_1,g_1;\alpha) \\
&+ \alpha(K(f_0,f_1)+K(f_1,f_0))+(1-\alpha)(K(g_0,g_1)+K(g_1,g_0)) \\
&- [K(m_\alpha,n_\alpha)+K(n_\alpha,m_\alpha)],
\end{aligned}$$

ارائه می‌شود که در آن $K(f,g)$ اندازه اطلاع نادرستی کریج است و به‌صورت $K(f,g)=H(f)+KL(f,g)$ تعریف می‌شود.

ملاحظه ۴.۴. با توجه به نتیجه ۳.۴ به راحتی می‌توان نتیجه گرفت که اندازه اطلاع جفریز نسبت به بردارهای توابع چگالی (f_0,f_1) و (g_0,g_1) توأماً محدب می باشد.

۵ نتیجه‌گیری

در این مقاله دو اندازه اطلاع جدید بر اساس نامساوی جنسن، تحت عناوین جنسن-فیشر پارامتری و جنسن-χ^2 معرفی شده است. یکی از نتایج مرتبط با اطلاع جنسن-فیشر پارامتری این است که این اندازه اطلاع جدید را می‌توان به‌صورت آمیخته‌ای از فواصل اطلاع فیشر بیان کرد. همچنین ماتریس اطلاع فیشر برای یک توزیع آمیخته از توابع چگالی احتمال محاسبه شده است. تعمیم‌هایی از دو اندازه اطلاع جدید بر اساس m تابع چگالی احتمال آمیخته ارائه شده است. یکی از نتایج بدست آمده این است که دو اندازه اطلاع معرفی شده با هم در ارتباط می‌باشند. همچنین نشان دادیم که اطلاع جنسن-فیشر بیزی تحت توزیع پیشین یکنواخت برای پارامتر آمیخته، با اطلاع جفریز و جنسن-شانون مرتبط است. یکی دیگر از نتایج مرتبط با اطلاع جنسن-فیشر بیزی این است که اطلاع جفریز به‌صورت یک تابع توأماً محدب نسبت به توابع چگالی احتمال می‌باشد.

فهرست منابع

[1] C. E. Shannon, A mathematical theory of communication, *Bell System Technical Journal*, **27** (1948), 379-423.

[2] R. A. Fisher, Tests of significance in harmonic analysis,*Proceedings of the Royal Society of London, Series A*, **125** (1929), 54-59.

[3] C. R. Rao, Diversity and dissimilarity coefficients: a unified approach, *Theoretical population biology*, **21** (1982), 24–43.

[4] O. Kharazmi & N. Balakrishnan, Jensen-information generating function and its connections to some well-known information measures, *Statistics & Probability Letters*, **170** (2021), .108995

[5] O. Kharazmi & N. Balakrishnan, Discrete Versions of Jensen–Fisher, Fisher and Bayes–Fisher Information Measures of Finite Mixture Distributions, *Entropy*, **23** (2021), .363

[6] O. Kharazmi & N. Balakrishnan, Cumulative residual and relative cumulative residual Fisher information and their properties, *IEEE Transactions on Information Theory*,(2021), DOI: 10.1109/TIT.2021.3073789.

[7] B. R. Frieden, Science from Fisher Information: A Unification, *Cambridge: Cambridge University Press*, (2004), MR2069674.

[8] P. Sánchez-Moreno, A. Zarzo, and J. S. Dehesa, Jensen divergence based on Fisher’s information, *Journal of Physics A: Mathematical and Theoretical*, **45** (2012), 125305.

[9] J. Lin, Divergence measures based on the Shannon entropy,*IEEE Transactions on Information Theory*, **37** (1991), 145–151.

Jensen-Fisher and Jensen-χ^2 information measures for finite mixture distributions

Omid Kharazmi[1, †] **, Morad Alizadeh**[2]

(1) Department of Statistics, Faculty of Mathematical Sciences, Vali-e-Asr University of Rafsanjan, Iran
(2) Department of Statistics, Faculty of Intelligent Systems Engineering and Data Science, Persian Gulf University, Bushehr, Iran

Communicated by: Gholam Ali Parham

Received: 2021/8/13 Accepted: 2021/11/25

Abstract:
In this paper, first, considering Fisher information of parametric type, we introduce a new information measure based on Jensen inequality. Then, the Fisher information matrix and Jensen-Fisher are studied for a finite mixture distribution of probability density functions. Further, another information criterion is introduced as Jensen-χ^2 based on a mixture of probability density functions. Generalizations of Jensen-Fisher and Jensen-χ^2 information measures are presented based on m probability density functions and the relationship between these two new information criteria as well as the relationship between Jensen-Fisher information and some known information criteria such as Jensen-Shannon and Jeffreys information measures are studied.

Keywords: Fisher information, Convex function, Jensen inequality, Fisher information matrix, Chi-square divergence.

†Corresponding author.
E-mail addresses: omid.kharazmi@vru.ac.ir (Omid Kharazmi), m.alizadeh@pgu.ac.ir (Morad Alizadeh).