Kharazmi University

Mathematical Researches

Print ISSN: 2588-2546
Online ISSN: 2588-2554

Homepage: https://jemr.khu.ac.ir/

New Approach to Fisher Information

M. Shams[1]

1. Corresponding Author, Department of Statistics, Faculty of Mathematical Sciences, University of Kashan, Kashan, Iran, E-mail: mehdishams@kashanu.ac.ir

Article Info

Article type:
Research Article

Article history:
Received: 7 October 2023
Received in revised form: 8 February 2024
Accepted: 17 August 2024
Published online: 10 November 2024

Keywords:
Shannon differential entropy,
Fisher information,
Markov chains,
Cramer-Rao inequality,
Sufficiency.

ABSTRACT

Introduction

Fisher information is a measure of the information inside a random variable about an unknown parameter. Mutual information shows the dependence between two variables and relative entropy shows the difference between the two probability distributions.

Material and Methods

In this paper, Fisher information is generalized for mutual information and relative entropy and the monotonicity properties of Fisher information are examined. Then concepts such as information correlation and information correlation coefficient are introduced.

Results and discussion

It is shown that Shannon differential entropy, which measures the behavior of a random variable, and conditional Fisher information are used to determine the probability of estimation error.

Conclusion

In this paper, with the help of Fisher's information structure, some concepts of information theory were generalized. For a Markov chain, it was shown that when the statistic moves away from the estimated parameter, the mutual Fisher information decreases, and when the Fisher information is conditioned on closer values, it is higher than when it is conditioned on more distant values. Two open problems are proposed for researchers who are interested in this topic. In cases where a topological group acts on random variables, the criteria mentioned in this paper should be analyzed. Also, with the help of the convolution function, the criteria mentioned in the paper can be examined for linear combinations of random variables.

How to cite: Shams, M., (2024). New Approach to Fisher Information. *Mathematical Researches*, **10** (3), 72 – 99.

 Publisher: Kharazmi University

شاپا چاپی: ۲۰۴۶-۲۵۸۸
شاپا الکترونیکی: ۲۵۵۴-۲۵۸۸

پژوهش‌های ریاضی

Homepage: https://mmr.khu.ac.ir/

رویکرد جدید برای اطلاع فیشر

مهدی شمس[1] ✉

۱. نویسنده مسئول، گروه آمار، دانشکده علوم ریاضی، دانشگاه کاشان، کاشان، ایران. ✉رایانامه: mehdishams@kashanu.ac.ir

اطلاعات مقاله

نوع مقاله: مقاله پژوهشی

تاریخ دریافت: ۱۴۰۲/۷/۱۵
تاریخ بازنگری: ۱۴۰۲/۱۱/۱۹
تاریخ پذیرش: ۱۴۰۳/۵/۲۷
تاریخ انتشار: ۱۴۰۳/۸/۲۰

چکیده

اطلاع فیشر معیاری برای اندازه‌گیری اطلاعات نهفته در متغیر تصادفی راجع به پارامتر مجهول است. اطلاع متقابل وابستگی بین دو متغیر و آنتروپی نسبی تفاوت بین دو توزیع احتمال را نشان می‌دهد. در این مقاله اطلاع متقابل و آنتروپی نسبی برای اطلاع فیشر تعمیم داده می‌شود و خواص یکنوایی اطلاع فیشر مورد بررسی قرار می‌گیرد. سپس مفاهیمی چون همبستگی اطلاع و ضریب همبستگی اطلاع معرفی می‌شوند. در پایان نشان داده می‌شود که آنتروپی دیفرانسیل شانون که متغیر تصادفی را به صورت کمی اندازه‌گیری می‌کند و اطلاع فیشر شرطی برای تعیین احتمال خطای برآورد مورد استفاده قرار می‌گیرند.

واژه‌های کلیدی:
آنتروپی دیفرانسیل شانون،
اطلاع فیشر،
زنجیر مارکف،
نامساوی کرامر-رائو،
بسندگی.

استناد: شمس، مهدی (۱۴۰۳). رویکرد جدید برای اطلاع فیشر. *پژوهش‌های ریاضی*، **۱۰** (۳)، ۷۲ –۹۹.

ناشر: دانشگاه خوارزمی

پیش‌گفتار

تولد شاخه نظریه اطلاع به تحقیقات شانون [۳۱] بر پایه مطالعه رفتار سیستم‌هایی که توسط توابع چگالی توصیف می شدند، بر می‌گردد. قبل از آن فیشر [۹] کمیتی به نام اطلاع فیشر ارائه کرد که علاقه‌مندان به کاربردهای اطلاع فیشر می توانند مقاله مروری [۲۵] را مطالعه کنند.

یکی از روابط بین اطلاع فیشر و آنتروپی دیفرانسیل شانون، توسط کولبک [۱۹] بیان شد که نشان داد مشتق‌های مرتبه دوم فاصله کولبک-لایبلر نسبت به پارامترهای توابع چگالی، مؤلفه‌های ماتریس اطلاع فیشر را تولید می‌کنند. تحقیقات مشابه در این زمینه در [۸،۱۰] انجام شد. بعد از آن، این نتایج توسط [۱۵] برای اغتشاشات غیرگاوسی نیز تعمیم داده شد که منجر به نامساوی پیچش برای اطلاع فیشر شد [۸]. در [۲۱،۲۳] آنتروپی ماکزیمم با اطلاع فیشر مینیمم آمیخته می‌شود. به عبارت دیگر ترکیب آنتروپی دیفرانسیل شانون، اطلاع فیشر و قضیه حد مرکزی برای نمونه بزرگ، امکان جست‌وجوی مقدار مینیمم برای فاصله کولبک-لایبلر را فراهم می‌کند [۳۵]. در مقالاتی نظیر [۱۰،۱۱،۱۲،۱۴،۱۶،۱۷،۲۹،۳۳]، ارتباط اطلاع فیشر با مفاهیم فیزیکی به خصوص نظریه ترمودینامیک بیان شده‌اند. در [۲۸] کاربردهای اطلاع فیشر در شبکه‌های کامپیوتری و یادگیری ماشین بیان شد.

در این مقاله پس از مرور تحقیقات قبلی در مورد خواص اطلاع فیشر [۸،۱۰]، دیدگاه جدیدی در مورد اطلاع فیشر مطرح می‌شود. در حقیقت ترکیبی از دو دیدگاه که سیستم را به طور کامل مشخص می‌کنند، یعنی رفتار (شانون) و ساختار (فیشر) ارائه می‌شود. به عبارت دیگر به کمک ساختار اطلاع فیشر که از مفاهیم استنباط آماری است، برخی مفاهیم نظریه اطلاع نظیر اطلاع متقابل و فاصله کولبک-لایبلر تعمیم داده می‌شود که به ترتیب اطلاع فیشر متقابل و اطلاع فیشر نسبی نامگذاری می‌شوند.

مفاهیم اطلاع فیشر، نامساوی کرامر-رائو، اطلاع فیشر شرطی و قاعده زنجیره‌ای برای اطلاع فیشر در بخش دوم معرفی می شوند.

در بخش دوم بعد از تعمیم اطلاع فیشر نسبی برای توابع محدب با مقادیر حقیقی و دو تعمیم دیگر از این معیار، مفهوم همبستگی اطلاع و خواص آن مورد بررسی قرار می‌گیرد. سپس معیار جدیدی به نام ضریب همبستگی اطلاع معرفی می‌گردد و در نهایت یک شرط کافی برای صفر بودن این معیار بیان می‌شود.

در بخش سوم به کمک اطلاع فیشر، اطلاع متقابل تعمیم داده می‌شود که از آن به‌عنوان اطلاع فیشر متقابل یاد می‌شود. به کمک دو تعمیم از اطلاع فیشر، دو صورت کلی از معیار اطلاع فیشر متقابل ارائه می‌شود. به کمک سه تعمیم ذکر شده برای اطلاع فیشر نسبی در بخش سوم و ارتباط این معیار با اطلاع فیشر متقابل، سه تعمیم دیگر برای اطلاع فیشر متقابل معرفی می‌شود. سپس مفاهیمی چون همبستگی اطلاع شرطی و اطلاع فیشر متقابل شرطی معرفی خواهند شد.

در بخش چهارم ثابت می‌شود شرطی کردن، اطلاع فیشر را افزایش می‌دهد و در زنجیرهای مارکف، اطلاع فیشر متقابل زمانی که متغیرهای تصادفی از پارامترهای برآوردشده دور شوند، کاهش می‌یابد. در پایان یک کران بالا برای احتمال خطای

برآورد بر حسب اطلاع فیشر شرطی معرفی می‌شود.

۱. اطلاع فیشر

در این بخش مفاهیم اطلاع فیشر، نامساوی کرامر-رائو، اطلاع فیشر شرطی و قاعده زنجیره‌ای برای اطلاع فیشر یادآوری می‌شود.

فرض کنید متغیر تصادفی $\mathbf{X}$ دارای تابع چگالی $f_{\mathbf{X};\boldsymbol{\theta}}$ باشد و $\mathcal{X}$ تکیه‌گاه متغیر تصادفی و Θ فضای پارامتر باشد. اطلاع فیشر متناظر با θ_k یعنی مؤلفه kام بردار $\boldsymbol{\theta}\in\Theta$ به صورت زیر تعریف می‌شود:

$$I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)=E_{\mathbf{X};\boldsymbol{\theta}}\left(\tfrac{\partial}{\partial\theta_k}\ln f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})\right)^2 =\int_{\mathcal{X}}\left[\tfrac{\partial}{\partial\theta_k}\ln f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})\right]^2 f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})d\mathbf{x}\geq 0$$

که تساوی برقرار است اگر توزیع X به θ_k بستگی نداشته باشد. به عنوان مثال به راحتی می‌توان نشان داد اگر X دارای توزیع $N(\mu,\sigma^2)$ باشد، $I_{X;\mu,\sigma^2}(\mu)=\frac{1}{\sigma^2}$ [۲۲].

در [۱۸] نشان داده شد، اگر شرایط نظم برقرار باشد، آن‌گاه

$$I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)=-E_{\mathbf{X};\boldsymbol{\theta}}\left(\tfrac{\partial^2}{\partial\theta_k^2}\ln f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})\right)=-\int_{\mathcal{X}}\left(\tfrac{\partial^2}{\partial\theta_k^2}\ln f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})\right)f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})d\mathbf{x}.$$

به راحتی می‌توان نشان داد که اطلاع فیشر یک متغیر تصادفی X با تکیه‌گاه $\mathcal{X}=\mathbb{R}$ متناظر با پارامتر θ_k تحت تبدیلات مکان، ناوردا است، یعنی $I_{X+a;\boldsymbol{\theta}}(\theta_k)=I_{X;\boldsymbol{\theta}}(\theta_k)$.

قضیه ۱.۱. (نامساوی کرامر-رائو، [۲۲]). تحت شرایط نظم و وجود برآوردگر نااریب $\hat{\theta}_k(\mathbf{X})$ برای θ_k داریم $\sigma^2_{\hat{\theta}_k}\geq\frac{1}{I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)}$ که در آن $\sigma^2_{\hat{\theta}_k}=\mathrm{Var}_{\mathbf{X};\boldsymbol{\theta}}\left(\hat{\theta}_k(\mathbf{X})\right)=\int_{\mathcal{X}}\left(\hat{\theta}_k(\mathbf{x})-\theta_k\right)^2 f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})d\mathbf{x}$. ■

کران پایین نامساوی کرامر-رائو، یعنی عکس اطلاع فیشر یک کران پایین برای واریانس برآوردگر نااریب ارائه می‌دهد. هر برآوردگری که واریانس آن برابر با این کران پایین است را کارا می‌نامند [۱۸]. از این به بعد تا انتهای مقاله، برای راحتی انتگرال‌های معین به صورت انتگرال نامعین نوشته خواهند شد.

فرض کنید $(\mathbf{X},\mathbf{Y})$ دارای تابع چگالی توأم $f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})$ باشد. اطلاع فیشر متناظر با θ_k به صورت زیر تعریف می‌شود [۲۲]:

$$I_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)=E_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\left[\tfrac{\partial}{\partial\theta_k}\ln f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})\right]^2=\iint\left[\tfrac{\partial}{\partial\theta_k}\ln f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})\right]^2 f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y}.$$

در [۱۸] ثابت شد اگر شرایط نظم برقرار باشد، آن‌گاه

$$I_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)=-E_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\left(\tfrac{\partial^2}{\partial\theta_k^2}\ln f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})\right)=-\iint\left(\tfrac{\partial^2}{\partial\theta_k^2}\ln f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})\right)f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y}.$$

به طور مشابه اطلاع فیشر شرطی نیز به صورت زیر تعریف می‌شود [۲۲]:

$$I_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\theta_k)=E_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\left[\tfrac{\partial}{\partial\theta_k}\ln f_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\mathbf{Y}|\mathbf{X})\right]^2=\iint\left[\tfrac{\partial}{\partial\theta_k}\ln f_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\mathbf{y}|\mathbf{x})\right]^2 f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y}.$$

زمیر [۳۴] نشان داد که اگر شرایط نظم برقرار باشد، آن‌گاه قاعده زنجیره‌ای برای اطلاع فیشر برقرار است، یعنی

$$I_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)=I_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\theta_k)+I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)=I_{\mathbf{X}|\mathbf{Y};\boldsymbol{\theta}}(\theta_k)+I_{\mathbf{Y};\boldsymbol{\theta}}(\theta_k).$$

برای اثبات این برابری به صورت زیر عمل می‌شود:

$$\begin{aligned}
I_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)&=\iint f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})\Big(\frac{\partial\ln f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}{\partial\theta_k}\Big)^2 d\mathbf{x}d\mathbf{y}\\
&=\iint f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})\Big(\frac{\partial\ln(f_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\mathbf{y}\,|\,\mathbf{x})f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x}))}{\partial\theta_k}\Big)^2 d\mathbf{x}d\mathbf{y}\\
&=\iint f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})\Big(\frac{\partial\ln f_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\mathbf{y}\,|\,\mathbf{x})}{\partial\theta_k}+\frac{\partial\ln f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})}{\partial\theta_k}\Big)^2 d\mathbf{x}d\mathbf{y}\\
&=I_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\theta_k)+I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)+2\iint f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})\frac{\partial\ln f_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\mathbf{y}\,|\,\mathbf{x})}{\partial\theta_k}\frac{\partial\ln f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})}{\partial\theta_k}d\mathbf{x}d\mathbf{y}\\
&=I_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\theta_k)+I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)+\iint\frac{\partial f_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\mathbf{y}\,|\,\mathbf{x})}{\partial\theta_k}\frac{\partial f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})}{\partial\theta_k}d\mathbf{x}d\mathbf{y}\\
&=I_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\theta_k)+I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)+\int\Big(\frac{\partial f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})}{\partial\theta_k}\int\frac{\partial f_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\mathbf{y}\,|\,\mathbf{x})}{\partial\theta_k}d\mathbf{y}\Big)d\mathbf{x}\\
&=I_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\theta_k)+I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)+\int\Big(\frac{\partial f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})}{\partial\theta_k}\frac{\partial}{\partial\theta_k}\int f_{\mathbf{Y}|X\,;\boldsymbol{\theta}}(\mathbf{y}\,|\,\mathbf{x})d\mathbf{y}\Big)d\mathbf{x}\\
&=I_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\theta_k)+I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)
\end{aligned}$$

که در تساوی آخر با توجه به شرایط نظم (برقراری شرط جابه‌جایی انتگرال با مشتق) عبارت سوم برابر با صفر است

$$\int\frac{\partial f_{\mathbf{Y}|X\,;\boldsymbol{\theta}}(\mathbf{y}\,|\,\mathbf{x})}{\partial\theta_k}d\mathbf{y}=\frac{\partial}{\partial\theta_k}\int f_{\mathbf{Y}|X\,;\boldsymbol{\theta}}(\mathbf{y}\,|\,\mathbf{x})d\mathbf{y}=0$$

و اثبات کامل می‌شود. در حالت چند متغیره داریم $I_{\mathbf{X}_1,\ldots,\mathbf{X}_n;\boldsymbol{\theta}}(\theta_k)=\sum_{k=1}^{n}I_{\mathbf{X}_k|\mathbf{X}_{k-1},\ldots,\mathbf{X}_1;\boldsymbol{\theta}}(\theta_k)$.

فرض کنید $T(\mathbf{X})$ یک آماره دلخواه باشد. به راحتی ثابت می‌شود $I_{T(\mathbf{X});\boldsymbol{\theta}}(\theta_k)\le I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)$ و شرط لازم و کافی برای این که تساوی برقرار باشد این است که $T(\mathbf{X})$ یک آماره بسنده باشد [۲۲]. این حقیقت نشان می‌دهد اطلاعی که

آماره نسبت به پارامتر می‌دهد از اطلاعی که خود نمونه اصلی نسبت به پارامتر می‌دهد کمتر است و زمانی دقیقاً این دو معیار با هم برابر هستند که آماره مورد نظر به گونه‌ای فشرده شده باشد که هیچ اطلاعی راجع به پارامتر را از دست ندهد و آماره بسنده این خصوصیت را دارد.

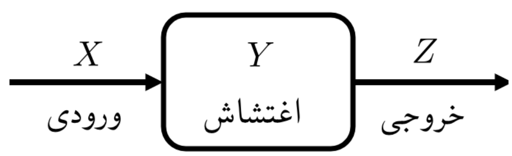
همچنین با توجه به قاعده زنجیره‌ای برای اطلاع فیشر، به ازای متغیرهای تصادفی مستقل $\mathbf{X}$ و $\mathbf{Y}$ داریم

$$I_{\mathbf{X},\,\mathbf{Y};\boldsymbol{\theta}}(\theta_k)=I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)+I_{\mathbf{Y};\boldsymbol{\theta}}(\theta_k)$$

که به عنوان نمونه با توجه به قضیه باسو [۱] برای آماره بسنده کامل $T(\mathbf{X})$ و آماره کمکی $U(\mathbf{X})$ داریم $I_{T(\mathbf{X}),U(\mathbf{X});\boldsymbol{\theta}}(\theta_k)=I_{T(\mathbf{X});\boldsymbol{\theta}}(\theta_k)+I_{U(\mathbf{X});\boldsymbol{\theta}}(\theta_k)$ و در حالت خاص که آماره $U(\mathbf{X})$ برای پارامتر θ_k کمکی است، به سادگی می‌توان نوشت $I_{T(\mathbf{X}),U(\mathbf{X});\boldsymbol{\theta}}(\theta_k)=I_{T(\mathbf{X});\boldsymbol{\theta}}(\theta_k)$ و این تساوی بیان کننده این حقیقت است که اطلاع فیشر توزیع توأم یک آماره بسنده کامل با یک آماره کمکی برابر با اطلاع فیشر آماره بسنده کامل است و در حقیقت آماره کمکی نمی‌تواند اطلاعاتی به یک آماره بسنده کامل راجع به پارامتر مجهولی که توزیع آماره کمکی به آن پارامتر وابسته نیست اضافه کند.

مثال ۱.۱. فرض کنید X و Y دو متغیر تصادفی مستقل به ترتیب دارای توزیع $N(\mu,\sigma^2)$ و $N(0,\eta^2)$ باشند و توزیع (X,Y) در دسترس باشد و $Z=X+Y$. از این‌رو:

$$\begin{aligned}I_{Z,X;\mu,\sigma^2,\eta^2}(\mu)&=I_{Z|X;\mu,\sigma^2,\eta^2}(\mu)+I_{X;\mu,\sigma^2}(\mu)\\&=I_{Y;\eta^2}(\mu)+I_{X;\mu,\sigma^2}(\mu)\\&=I_{X;\mu,\sigma^2}(\mu)=\tfrac{1}{\sigma^2}.\ \blacksquare\end{aligned}$$



شکل ۱. یک سامانه خطی با ورودی X، خروجی Z و اغتشاش جمعی Y

مثال ۱.۱ نشان می‌دهد، اگر تابع چگالی توأم خروجی Z و ورودی X در دسترس باشد، اغتشاش Y در فرایند برآورد بی‌تأثیر است (شکل ۱). نظر به این که تمام اطلاعات نهفته در X از طریق تابع چگالی توأم در دسترس است، اطلاع فیشر تابع چگالی توأم متناظر با اطلاع فیشر توزیع کناری $f_{X;\mu,\sigma^2}$ است. از این که Z دارای توزیع $N(\mu,\sigma^2+\eta^2)$ است، $I_{Z;\mu,\sigma^2,\eta^2}(\mu)=\frac{1}{\sigma^2+\eta 2}$ و با استفاده از قاعده زنجیره‌ای برای اطلاع فیشر داریم:

$$I_{Z,X;\mu,\sigma^2,\eta^2}(\mu)=I_{X|Z;\mu,\sigma^2,\eta^2}(\mu)+I_{Z;\mu,\sigma^2,\eta^2}(\mu)$$

و لذا $I_{X|Z;\mu,\sigma^2,\eta^2}(\mu)=\frac{1}{\sigma^2}-\frac{1}{\sigma^2+\eta^2}=\frac{\eta^2}{\sigma^2\left(\sigma^2+\eta^2\right)}$.

۲. رویکرد جدید آنتروپی نسبی برای اطلاع فیشر و همبستگی اطلاع

در ابتدا تعمیم اطلاع فیشر نسبی برای توابع محدب با مقادیر حقیقی و دو تعمیم دیگر ارائه می‌شود. سپس مفهوم همبستگی اطلاع و خواص آن مورد بررسی قرار می‌گیرد. در ادامه معیار جدیدی به نام ضریب همبستگی اطلاع معرفی می‌گردد که شبیه ضریب همبستگی بین ۱- و ۱ کران‌دار بوده و در نهایت نشان داده می‌شود یک شرط کافی برای صفر بودن ضریب همبستگی اطلاع این است که دو متغیر تصادفی مستقل باشند یا حداقل توزیع یکی از آن‌ها به پارامتر مجهول وابسته نباشد.

تعریف اطلاع فیشر نسبی برای اولین بار در [۲۶] معرفی شد و سپس در [۳۳،۳۰،۱۷،۳] کاربردهایی از آن ارائه شد. در نظریه اطلاع معیاری به نام فاصله کولبک-لایبلر یا آنتروپی نسبی یا واگرایی بین دو تابع چگالی $f_{\mathbf{X};\boldsymbol{\theta}}$ و $f_{\mathbf{Y};\boldsymbol{\theta}}$ تعریف می‌شود:

$$D(f_{\mathbf{X};\boldsymbol{\theta}} \,||\, f_{\mathbf{Y};\boldsymbol{\theta}}) = E_{\mathbf{X};\boldsymbol{\theta}}\left(\ln \frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})}{f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X})}\right) = \int \frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})}{f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x})} f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x}) d\mathbf{x}$$

که این معیار توسط کولبک و لایبلر [۲۰] برای تشخیص میزان اختلاف توابع چگالی معرفی شد. اطلاع فیشر نسبی زمانی که در هر کدام از فرمول‌های اطلاع فیشر به جای تابع چگالی، نسبت دو تابع چگالی جایگزین شود به دست می‌آید که این نامگذاری به واسطه شباهت آن با فرمول فاصله کولبک-لایبلر است. از این‌رو معیار آنتروپی نسبی با اطلاع فیشر آمیخته شده و معیار جدیدی به نام اطلاع فیشر نسبی تعریف می‌شود. از این‌رو اطلاع فیشر نسبی نوع I به صورت زیر تعریف می‌شود:

$$D^{I}\left(f_{\mathbf{X};\boldsymbol{\theta}} \,||\, f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k) = E_{\mathbf{X};\boldsymbol{\theta}}\left(\tfrac{\partial}{\partial\theta_k} \ln \tfrac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})}{f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X})}\right)^2 = \int \left(\tfrac{\partial}{\partial\theta_k} \ln \tfrac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})}{f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x})}\right)^2 f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x}) d\mathbf{x}.$$

به طور مشابه با قبل می‌توان عبارت بالا را بر حسب مشتق مرتبه دوم نیز نوشت که به آن اطلاع فیشر نسبی نوع II گویند:

$$D^{II}\left(f_{\mathbf{X};\boldsymbol{\theta}} \,||\, f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k) = -E_{\mathbf{X};\boldsymbol{\theta}}\left(\tfrac{\partial^2}{\partial\theta_k^2} \ln \tfrac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})}{f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X})}\right) = -\int \tfrac{\partial^2}{\partial\theta_k^2} \ln \tfrac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})}{f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x})} f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x}) d\mathbf{x}$$

که لزوماً D^{I} و D^{II} مساوی نیستند، زیرا در اطلاع فیشر هر دو فرمول از روی تابع چگالی ساخته شده بودند در صورتی که برای اطلاع فیشر نسبی دو فرمول از روی نسبت توابع چگالی ساخته می‌شوند که خود تابع چگالی نیست و از این‌رو $D^{I} \neq D^{II}$.

اطلاع فیشر نسبی نوع I و II شامل برخی خواص مشابه با آنتروپی نسبی هست. به عنوان نمونه شرط لازم و کافی برای این که $D^{i}\left(f_{\mathbf{X};\boldsymbol{\theta}} \,||\, f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k) = 0$، $i = I, II$ آن است که تقریباً همه جا $f_{\mathbf{X};\boldsymbol{\theta}} = f_{\mathbf{Y};\boldsymbol{\theta}}$. همچنین اطلاع فیشر نسبی یک معیار متقارن نیست. برای این منظور سانچز-مورنو و همکاران [۳۰]، معیار متقارنی به نام واگرایی فیشر به صورت زیر معرفی کردند:

$$FD^I\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)=D^I\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)+D^I\left(f_{\mathbf{Y};\boldsymbol{\theta}}\,||\,f_{\mathbf{X};\boldsymbol{\theta}}\right)(\theta_k)$$
$$=E_{\mathbf{X};\boldsymbol{\theta}}\left(\tfrac{\partial}{\partial\theta_k}\ln\tfrac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})}{f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X})}\right)^2+E_{\mathbf{Y};\boldsymbol{\theta}}\left(\tfrac{\partial}{\partial\theta_k}\ln\tfrac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})}{f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X})}\right)^2$$
$$=\int\left(\tfrac{\partial}{\partial\theta_k}\ln\tfrac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})}{f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x})}\right)^2\left(f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})+f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x})\right)d\mathbf{x}$$

که در فیزیک مورد استفاده قرار می‌گیرد و علاوه بر آن معیار واگرایی ینسن-فیشر نیز به صورت زیر معرفی کردند:

$$JFD^I\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)=D^I\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,\frac{f_{\mathbf{X};\boldsymbol{\theta}}+f_{\mathbf{Y};\boldsymbol{\theta}}}{2}\right)(\theta_k)+D^I\left(f_{\mathbf{Y};\boldsymbol{\theta}}\,||\,\frac{f_{\mathbf{X};\boldsymbol{\theta}}+f_{\mathbf{Y};\boldsymbol{\theta}}}{2}\right)(\theta_k)$$

که می‌تواند به عنوان یک متر در فضای توزیع‌های احتمال مورد استفاده قرار گیرد. واگرایی ینسن-فیشر در مکانیک کوانتوم، بیوانفورماتیک و فیزیک اتمی کاربرد دارد [۳۰].

در [۳۰] ثابت می‌شود واگرایی ینسن-فیشر را می‌توان بر حسب اطلاع فیشر نیز نوشت:

$$JFD^I\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)=\frac{I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)+I_{\mathbf{Y};\boldsymbol{\theta}}(\theta_k)}{2}-I_{\mathbf{Z};\boldsymbol{\theta}}(\theta_k)$$

که در آن متغیر تصادفی $\mathbf{Z}$ داری تابع چگالی آمیخته $\frac{f_{\mathbf{X};\boldsymbol{\theta}}+f_{\mathbf{Y};\boldsymbol{\theta}}}{2}$ است.

نامنفی بودن $JFD^I\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)$ نشان می‌دهد $\frac{I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)+I_{\mathbf{Y};\boldsymbol{\theta}}(\theta_k)}{2}\geq I_{\mathbf{Z};\boldsymbol{\theta}}(\theta_k)$ و این نامساوی ثابت می‌کند، اطلاع فیشر $I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)$ نسبت به $f_{\mathbf{X};\boldsymbol{\theta}}$ محدب است [۷].

بنابراین شرط لازم و کافی برای این که $JFD^I\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)=0$ برقرار باشد آن است که تقریباً همه جا $f_{\mathbf{X};\boldsymbol{\theta}}=f_{\mathbf{Y};\boldsymbol{\theta}}$.

به کمک تعمیم‌های آنتروپی نسبی معرفی‌شده در [۳۲،۲۷،۵]، می‌توان اطلاع فیشر نسبی را تعمیم داد که در ادامه به آن اشاره می‌شود.

در [۵] به کمک تابع محدب با مقادیر حقیقی φ، معیار φ- واگرایی بین توابع چگالی $f_{\mathbf{X};\boldsymbol{\theta}}$ و $f_{\mathbf{Y};\boldsymbol{\theta}}$ با ضابطه

$$D_\varphi(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}})=E_{\mathbf{X};\boldsymbol{\theta}}\Big[\varphi\Big(\frac{f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})}\Big)\Big]=\int\varphi\Big(\frac{f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})}\Big)f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})d\mathbf{x}$$

تعریف شد که در حالت خاص برای تابع محدب $\varphi_0(x)=-\ln(x)$ به راحتی می‌توان نوشت

$$D_{\varphi_0}(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}})=D(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}).$$

اکنون می‌خواهیم معیار φ- واگرایی را با رویکرد اطلاع فیشر تعمیم دهیم. برای این منظور برای تابع محدب با مقادیر حقیقی φ ، معیار φ- اطلاع فیشر نسبی نوع I بین توابع چگالی $f_{\mathbf{X};\boldsymbol{\theta}}$ و $f_{\mathbf{Y};\boldsymbol{\theta}}$ را به صورت زیر تعریف می‌کنیم:

$$D_{\varphi}^{I}\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)=E_{\mathbf{X};\boldsymbol{\theta}}\left[\left(\frac{\partial}{\partial\theta_k}\varphi\left(\frac{f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})}\right)\right)^2\right]=\int\left(\frac{\partial}{\partial\theta_k}\varphi\left(\frac{f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})}\right)\right)^2 f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})d\mathbf{x}. \quad (۱)$$

عبارت بالا را می‌توان بر حسب مشتق مرتبه دوم نیز نوشت که آن را φ-اطلاع فیشر نسبی نوع II نامیم:

$$D_{\varphi}^{II}\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)=-E_{\mathbf{X};\boldsymbol{\theta}}\left[\frac{\partial^2}{\partial\theta_k^2}\varphi\left(\frac{f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})}\right)\right]=-\int\frac{\partial^2}{\partial\theta_k^2}\varphi\left(\frac{f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})}\right)f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})d\mathbf{x}. \quad (۲)$$

در حالت خاص برای تابع محدب $\varphi_0(x)=-\ln(x)$ به راحتی می‌توان نوشت

$$D_{\varphi_0}^{i}\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)=D^{i}\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)\ \ ,i=I,II.$$

به کمک معیار متقارن واگرایی فیشر معرفی شده در [۳۰]، معیار مشابهی برای متقارن کردن φ- اطلاع فیشر نسبی نوع I به صورت زیر ارائه می‌شود:

$$\begin{aligned}FD_{\varphi}^{I}\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)&=D_{\varphi}^{I}\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)+D_{\varphi}^{I}\left(f_{\mathbf{Y};\boldsymbol{\theta}}\,||\,f_{\mathbf{X};\boldsymbol{\theta}}\right)(\theta_k)\\&=E_{\mathbf{X};\boldsymbol{\theta}}\left[\left(\frac{\partial}{\partial\theta_k}\varphi\left(\frac{f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})}\right)\right)^2\right]+E_{\mathbf{Y};\boldsymbol{\theta}}\left[\left(\frac{\partial}{\partial\theta_k}\varphi\left(\frac{f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})}\right)\right)^2\right]\\&=\int\left(\frac{\partial}{\partial\theta_k}\varphi\left(\frac{f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})}\right)\right)^2\left(f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})+f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x})\right)d\mathbf{x}.\end{aligned}$$

علاوه بر آن، معیار واگرایی ینسن-فیشر ارائه شده در [۳۰] نیز برای φ- اطلاع فیشر نسبی نوع I به صورت زیر تعریف می‌شود:

$$JFD_{\varphi}^{I}\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)=D_{\varphi}^{I}\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,\frac{f_{\mathbf{X};\boldsymbol{\theta}}+f_{\mathbf{Y};\boldsymbol{\theta}}}{2}\right)(\theta_k)+D_{\varphi}^{I}\left(f_{\mathbf{Y};\boldsymbol{\theta}}\,||\,\frac{f_{\mathbf{X};\boldsymbol{\theta}}+f_{\mathbf{Y};\boldsymbol{\theta}}}{2}\right)(\theta_k).$$

معیارهای φ- اطلاع فیشر نسبی نوع I و معیار واگرایی ینسن-فیشر یعنی $FD_{\varphi}^{I}\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)$ و $JFD_{\varphi}^{I}\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)$ هر دو به ترتیب تعمیم‌های $FD^{I}\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)$ و $JFD^{I}\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)$ هستند، به طوری که

$$\begin{aligned}FD_{-\ln}^{I}\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)&=FD^{I}\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k),\\JFD_{-\ln}^{I}\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)&=JFD^{I}\left(f_{\mathbf{X};\boldsymbol{\theta}}\,||\,f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k).\end{aligned}$$

سالیس [۳۲]، تعمیمی برای آنتروپی نسبی بین دو تابع چگالی $f_{\mathbf{X};\boldsymbol{\theta}}$ و $f_{\mathbf{Y};\boldsymbol{\theta}}$ به صورت زیر ارائه کرد:

$$D^{(q)}\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)=E_{\mathbf{X};\boldsymbol{\theta}}\Big[\frac{\left[f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})/f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X})\right]^{q-1}-1}{q-1}\Big]=\int\frac{\left[f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})/f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x})\right]^{q-1}-1}{q-1}f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})d\mathbf{x}$$

و در حالت خاص $\lim_{q\to 1} D^{(q)}\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)=D\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)$.

در این قسمت، با آمیختن $D^{(q)}\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)$ با اطلاع فیشر، معیارهای جدید زیر برای اطلاع فیشر نسبی نوع $i=I,II$ تعریف می‌شود:

$$D^{I,(q)}\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)=E_{\mathbf{X};\boldsymbol{\theta}}\Big[\tfrac{\partial}{\partial\theta_k}\Big(\frac{\left[f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})/f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X})\right]^{q-1}-1}{q-1}\Big)\Big]^2$$
$$=\int\tfrac{\partial}{\partial\theta_k}\Big(\frac{\left[f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})/f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x})\right]^{q-1}-1}{q-1}\Big)^2 f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})d\mathbf{x},$$

$$D^{II,(q)}\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)=-E_{\mathbf{X};\boldsymbol{\theta}}\Big[\tfrac{\partial^2}{\partial\theta_k^2}\Big(\frac{\left[f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})/f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X})\right]^{q-1}-1}{q-1}\Big)\Big]$$
$$=-\int\tfrac{\partial^2}{\partial\theta_k^2}\Big(\frac{\left[f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})/f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x})\right]^{q-1}-1}{q-1}\Big)f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})d\mathbf{x}.$$

به سادگی مشاهده می‌شود $\lim_{q\to 1} D^{i,(q)}\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)=D^{i}\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)$، $i=I,II$.

نکته قابل توجه‌ای با مقایسه دو معیار ارائه شده در مقالات [۵] و [۳۲] مشاهده می‌شود. برای این منظور تابع محدب

$$\varphi_q(x)=\frac{x^{q-1}-1}{q-1};\ x>0,q>1 \qquad (۳)$$

را در نظر بگیرید. با توجه به این که $\lim_{q\to 1}\varphi_q(x)=ln(x)$ به راحتی می‌توان نتیجه گرفت که

$$\lim_{q\to 1} D^{I}_{\varphi_q}\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)=\lim_{q\to 1} D^{I,(q)}\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)=D^{I}\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k),$$
$$-\lim_{q\to 1} D^{II,(q)}_{\varphi_q}\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)=\lim_{q\to 1} D^{II,(q)}\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)=D^{II}\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k).$$

از این رو برای تابع محدب (۳)، معیار φ_q- اطلاع فیشر نسبی نوع I بین توابع چگالی $f_{\mathbf{X};\boldsymbol{\theta}}$ و $f_{\mathbf{Y};\boldsymbol{\theta}}$، مقدار حدی $D^{I,(q)}\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)$ وقتی $q\to 1$ و همچنین اطلاع فیشر نسبی نوع I همگی با هم برابر خواهند بود. اما اطلاع فیشر نسبی نوع II بین توابع چگالی $f_{\mathbf{X};\boldsymbol{\theta}}$ و $f_{\mathbf{Y};\boldsymbol{\theta}}$ و مقدار حدی $D^{II,(q)}\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)$ وقتی $q\to 1$، با قرینه معیار φ_q- اطلاع فیشر نسبی نوع II برابر هستند و این حقیقت به خاطر مقعر بودن $\lim_{q\to 1}\varphi_q(x)=ln(x)$ است.

پارک و باسو [۲۷]، τ- آنتروپی نسبی تعمیم‌یافته $f_{\mathbf{X};\boldsymbol{\theta}}$ و $f_{\mathbf{Y};\boldsymbol{\theta}}$ را به صورت زیر تعریف کردند:

$$D_\tau\left(f_{\mathbf{X};\theta} \,||\, f_{\mathbf{Y};\theta}\right) = \int\Big[\frac{f_{\mathbf{Y};\theta}(\mathbf{x})}{\bar{\tau}} \ \ln\left(\frac{f_{\mathbf{Y};\theta}(\mathbf{x})}{f_{\mathbf{X};\theta}(\mathbf{x})}\right) - \left(\frac{f_{\mathbf{Y};\theta}(\mathbf{x})}{\bar{\tau}} + \frac{f_{\mathbf{X};\theta}(\mathbf{x})}{\tau}\right)\ln\left(\tau\frac{f_{\mathbf{Y};\theta}(\mathbf{x})}{f_{\mathbf{X};\theta}(\mathbf{x})} + \bar{\tau}\right)\Big]\, d\mathbf{x}$$

به طوری که $\tau \in [0,1]$ و $\bar{\tau} = 1-\tau$.

در حالت خاص به سادگی می‌توان نوشت $\lim_{\tau\to 1} D_\tau\left(f_{\mathbf{X};\theta} \,||\, f_{\mathbf{Y};\theta}\right) = D\left(f_{\mathbf{X};\theta} \,||\, f_{\mathbf{Y};\theta}\right)$.

به منظور استفاده از این معیار و آمیختن آن با اطلاع فیشر، به صورت کنونی نمی‌توان ایده اطلاع فیشر نسبی را پیاده کرد و برای رسیدن به این هدف ابتدا $D_\tau\left(f_{\mathbf{X};\theta} \,||\, f_{\mathbf{Y};\theta}\right)$ را به صورت زیر بازنویسی می‌کنیم:

$$\begin{aligned}
D_\tau\left(f_{\mathbf{X};\theta} \,||\, f_{\mathbf{Y};\theta}\right) &= \int\Big[\frac{f_{\mathbf{Y};\theta}(\mathbf{x})}{\bar{\tau} f_{\mathbf{X};\theta}(\mathbf{x})}\ln\left(\frac{f_{\mathbf{Y};\theta}(\mathbf{x})}{f_{\mathbf{X};\theta}(\mathbf{x})}\right) - \left(\frac{f_{\mathbf{Y};\theta}(\mathbf{x})}{\bar{\tau} f_{\mathbf{X};\theta}(\mathbf{x})} + \frac{1}{\tau}\right)\ln\left(\tau\frac{f_{\mathbf{Y};\theta}(\mathbf{x})}{f_{\mathbf{X};\theta}(\mathbf{x})} + \bar{\tau}\right)\Big]\, f_{\mathbf{X};\theta}(\mathbf{x})d\mathbf{x} \\
&= E_{\mathbf{X};\theta}\Big[\frac{f_{\mathbf{Y};\theta}(\mathbf{X})}{\bar{\tau} f_{\mathbf{X};\theta}(\mathbf{X})}\ln\left(\frac{f_{\mathbf{Y};\theta}(\mathbf{X})}{f_{\mathbf{X};\theta}(\mathbf{x})}\right) - \left(\frac{f_{\mathbf{Y};\theta}(\mathbf{X})}{\bar{\tau} f_{\mathbf{X};\theta}(\mathbf{X})} + \frac{1}{\tau}\right)\ln\left(\tau\frac{f_{\mathbf{Y};\theta}(\mathbf{X})}{f_{\mathbf{X};\theta}(\mathbf{X})} + \bar{\tau}\right)\Big].
\end{aligned}$$

اکنون - τاطلاع فیشر نسبی نوع $i = I, II$ به صورت زیر تعریف می‌شود:

$$D_\tau^{I}\left(f_{\mathbf{X};\theta} \,||\, f_{\mathbf{Y};\theta}\right)(\theta_k) = E_{\mathbf{X};\theta}\Big[\tfrac{\partial}{\partial\theta_k}\Big(\frac{f_{\mathbf{Y};\theta}(\mathbf{X})}{\bar{\tau} f_{\mathbf{X};\theta}(\mathbf{X})}\ln\Big(\frac{f_{\mathbf{Y};\theta}(\mathbf{X})}{f_{\mathbf{X};\theta}(\mathbf{X})}\Big) - \Big(\frac{f_{\mathbf{Y};\theta}(\mathbf{X})}{\bar{\tau} f_{\mathbf{X};\theta}(\mathbf{X})} + \frac{1}{\tau}\Big)\ln\Big(\tau\frac{f_{\mathbf{Y};\theta}(\mathbf{X})}{f_{\mathbf{X};\theta}(\mathbf{X})} + \bar{\tau}\Big)\Big)\Big]^2,$$

$$D_\tau^{II}\left(f_{\mathbf{X};\theta} \,||\, f_{\mathbf{Y};\theta}\right)(\theta_k) = -E_{\mathbf{X};\theta}\Big[\tfrac{\partial^2}{\partial\theta_k^2}\Big(\frac{f_{\mathbf{Y};\theta}(\mathbf{X})}{\bar{\tau} f_{\mathbf{X};\theta}(\mathbf{X})}\ln\Big(\frac{f_{\mathbf{Y};\theta}(\mathbf{X})}{f_{\mathbf{X};\theta}(\mathbf{x})}\Big) - \Big(\frac{f_{\mathbf{Y};\theta}(\mathbf{X})}{\bar{\tau} f_{\mathbf{X};\theta}(\mathbf{X})} + \frac{1}{\tau}\Big)\ln\Big(\tau\frac{f_{\mathbf{Y};\theta}(\mathbf{X})}{f_{\mathbf{X};\theta}(\mathbf{X})} + \bar{\tau}\Big)\Big)\Big].$$

به راحتی مشاهده می‌شود $\lim_{\tau\to 1} D_\tau^{i}\left(f_{\mathbf{X};\theta} \,||\, f_{\mathbf{Y};\theta}\right)(\theta_k) = D^{i}\left(f_{\mathbf{X};\theta} \,||\, f_{\mathbf{Y};\theta}\right)(\theta_k)$، $i = I, II$.

با مقایسه دو معیار معرفی‌شده در [۵] و [۲۷] و با در نظر گرفتن تابع

$$\varphi_\tau(x) = \frac{x}{\bar{\tau}}\ln(x) - \left(\frac{x}{\bar{\tau}} + \frac{1}{\tau}\right)\ln\left(\tau x + \bar{\tau}\right), \ \ x > 0, \tag{۴}$$

و این که $\varphi_\tau''(x) = \frac{1}{x(\tau x + \bar{\tau})} > 0$، مشاهده می‌شود تابع $\varphi_\tau(x)$ محدب است. با توجه به این که $\lim_{\tau\to 1}\varphi_\tau(x) = -ln(x)$، به راحتی می‌توان نتیجه گرفت که

$$\lim_{\tau\to 1} D_i^{\tau}\left(f_{\mathbf{X};\theta} \,||\, f_{\mathbf{Y};\theta}\right)(\theta_k) = \lim_{\tau\to 1} D_{\varphi_\tau}^{i}\left(f_{\mathbf{X};\theta} \,||\, f_{\mathbf{Y};\theta}\right)(\theta_k) = D^{i}\left(f_{\mathbf{X};\theta} \,||\, f_{\mathbf{Y};\theta}\right)(\theta_k), \ \ i = I, II.$$

از این رو برای تابع محدب (۴)، معیار - φ_τاطلاع فیشر نسبی نوع $i = I, II$ بین توابع چگالی $f_{\mathbf{X};\theta}$ و $f_{\mathbf{Y};\theta}$، مقدار حدی نوع $i = I, II$ برای معیار $D_\tau^{i}\left(f_{\mathbf{X};\theta} \,||\, f_{\mathbf{Y};\theta}\right)(\theta_k)$ وقتی $\tau \to 1$ و همچنین اطلاع فیشر نسبی نوع $i = I, II$ همگی با هم برابر خواهند بود.

در قضیه زیر تعمیم‌های ارائه‌شده برای اطلاع فیشر نسبی جمع‌بندی می‌شوند.

قضیه ۱.۲. هر کدام از حالت‌های زیر یک تعمیم برای اطلاع فیشر نسبی $D^i\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)$، $i=I,II$ است:

الف) φ- اطلاع فیشر نسبی نوع $i=I,II$ برای تابع محدب φ، $D_{\varphi}^i\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)$:

$$D_{-\ln}^i\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)=D^i\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k),\quad i=I,II.$$

ب) اطلاع فیشر نسبی نوع $i=I,II$ از مرتبه q، $D^{i,(q)}\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)$:

$$D^{i,(q)}\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)=D^i\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k),\quad i=I,II.$$

ج) τ- اطلاع فیشر نسبی نوع $i=I,II$ برای $\tau\in[0,1]$، $D_{\tau}^i\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)$:

$$\lim_{\tau\to 1} D_{\tau}^i\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)=D^i\left(f_{\mathbf{X};\boldsymbol{\theta}} \| f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k),\quad i=I,II. \blacksquare$$

در ادامه این بخش معیار همبستگی اطلاع را به شرح ذیل تعریف می‌شود.

تعریف ۱.۲. همبستگی اطلاع بین متغیرهای تصادفی $\mathbf{X}$ و $\mathbf{Y}$ نسبت به θ_k به صورت زیر تعریف می‌شود:

$$\begin{aligned}IC_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)&=E_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\left[\left(\tfrac{\partial}{\partial\theta_k}\ln f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})\right)\left(\tfrac{\partial}{\partial\theta_k}\ln f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})\right)\right]\\&=\iint\left(\tfrac{\partial}{\partial\theta_k}\ln f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})\right)\left(\tfrac{\partial}{\partial\theta_k}\ln f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})\right)f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y}. \blacksquare\end{aligned}$$

ذکر این نکته ضروری است که این تعریف با تعریف ماتریس اطلاع فیشر متفاوت است [۳۳].

با توجه به تعریف ۱.۲، به سادگی مشاهده می‌شود $IC_{\mathbf{X},\mathbf{X};\boldsymbol{\theta}}(\theta_k)=I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)$ و $IC_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)=IC_{\mathbf{Y},\mathbf{X};\boldsymbol{\theta}}(\theta_k)$. از این‌رو همبستگی اطلاع یک تعمیم از اطلاع فیشر است به گونه‌ای که همبستگی اطلاع بین یک متغیر با خودش، همان اطلاع فیشر است و همین‌طور همبستگی اطلاع خاصیت جابه‌جایی دارد.

مثال ۱.۲. فرض کنید X و Z متغیرهای تصادفی معرفی‌شده در مثال ۱.۱ باشد. در این صورت همبستگی اطلاع بین Z و X به صورت $IC_{Z,X;\mu,\sigma^2,\eta^2}(\mu)=\frac{1}{\sigma^2+\eta^2}$ به دست می‌آید.

ابتدا مشتق مرتبه اول توابع چگالی Z و X به صورت زیر محاسبه می‌شود:

$$\begin{aligned}&\frac{d\ln f_{Z;\mu,\sigma^2,\eta^2}(z)}{d\mu}=\frac{d}{d\mu}\left(\ln\frac{1}{\sqrt{2\pi}\sqrt{\sigma^2+\eta^2}}-\frac{(z-\mu)^2}{2(\sigma^2+\eta^2)}\right),\\&\frac{d\ln f_{X;\mu,\sigma^2}(z)}{d\mu}=\frac{x-\mu}{\sigma^2}.\end{aligned}\tag{۵}$$

با استفاده از رابطه (۵)، همبستگی اطلاع بین Z و X عبارت است از:

$$IC_{Z,X;\mu,\sigma^2,\eta^2}(\mu)=\iint f_{Z,X;\mu,\sigma^2,\eta^2}(z,x)\frac{d\ln f_{Z;\mu,\sigma^2,\eta^2}(z)}{d\mu}\frac{d\ln f_{X;\mu,\sigma^2}(x)}{d\mu}dzdx$$
$$=\iint f_{Z|X;\mu,\sigma^2,\eta^2}(z|x)f_{X;\mu,\sigma^2}(x)\frac{z-\mu}{\sigma^2+\eta^2}\frac{x-\mu}{\sigma^2}dzdx$$
$$=\frac{1}{\sigma^2(\sigma^2+\eta^2)}\iint f_{Y;\eta^2}(z-x)f_{X;\mu,\sigma^2}(x)(z-\mu)(x-\mu)dzdx$$
$$=\frac{1}{\sigma^2(\sigma^2+\eta^2)}\iint\frac{1}{\sqrt{2\pi}\eta}e^{-\frac{(z-x)^2}{2\eta^2}}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}(z-\mu)(x-\mu)dzdx$$
$$=\frac{1}{2\pi\eta\sigma(\sigma^2+\eta^2)\sigma^2}\iint(z-\mu)(x-\mu)e^{-\frac{1}{2}\left(\frac{(z-x)^2}{\eta^2}+\frac{(x-\mu)^2}{\sigma^2}\right)}dzdx$$
$$=\frac{1}{2\pi\eta\sigma(\sigma^2+\eta^2)\sigma^2}\int(x-\mu)e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}\left(\int(z-\mu)e^{-\frac{1}{2}\frac{(z-x)^2}{\eta^2}}dz\right)dx=\frac{1}{\sigma^2+\eta^2}.\ \blacksquare$$

قضیه ۲.۲. همبستگی اطلاع کران‌دار است، یعنی $\left(IC_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)\right)^2\le I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)I_{\mathbf{Y};\boldsymbol{\theta}}(\theta_k)$.

اثبات. می‌دانیم برای هر $a\in\mathbb{R}$:

$$0\le E_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\left(a\frac{\partial\ln f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})}{\partial\theta_k}+\frac{\partial\ln f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})}{\partial\theta_k}\right)^2=a^2I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)+2aIC_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)+I_{\mathbf{Y};\boldsymbol{\theta}}(\theta_k).$$

شرط برقراری رابط اخیر به ازای هر $a\in\mathbb{R}$ به صورت $\Delta'=\left(IC_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)\right)^2-I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)I_{\mathbf{Y};\boldsymbol{\theta}}(\theta_k)\le 0$ است که منجر به حکم می‌شود. ■

قضیه ۲.۲ بیان می‌کند که همبستگی اطلاع بین دو متغیر تصادفی (مشابه کوواریانس) به مجذور حاصل‌ضرب اطلاع فیشر آن دو متغیر تصادفی، کران‌دار است که نامساوی کوشی-شوارتز [۱۸] نیز بر همین مهم تأکید دارد.

در تعریف بعدی همبستگی اطلاع توسط اطلاع فیشر دو متغیر تصادفی نرمال‌سازی می‌شود و معیار جدیدی به دست می‌آید.

تعریف ۲.۲. فرض کنید $\mathbf{X}$ و $\mathbf{Y}$ دو متغیر تصادفی دلخواه باشند. در این صورت ضریب همبستگی اطلاع بین متغیرهای تصادفی $\mathbf{X}$ و $\mathbf{Y}$ نسبت به θ_k به صورت $ICC_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)=\frac{IC_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)}{\sqrt{I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)I_{\mathbf{Y};\boldsymbol{\theta}}(\theta_k)}}$ تعریف می‌شود. ■

به کمک قضیه ۲.۲، ثابت می‌شود $-1\le ICC_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)\le 1$.

مثال ۲.۲. ضریب همبستگی اطلاع بین متغیرهای تصادفی Z و X تعریف شده در مثال ۱.۱ برابر است با $ICC_{Z,X;\mu,\sigma^2,\eta^2}(\mu)=\frac{(\sigma^2+\eta^2)^{-1}}{\sqrt{(\sigma^2)^{-1}(\sigma^2+\eta^2)^{-1}}}=\frac{\sigma}{\sqrt{\sigma^2+\eta^2}}$ که به وضوح کوچک‌تر از ۱ است. ■

قضیه ۳.۲. مقدار همبستگی اطلاع بین متغیرهای تصادفی $\mathbf{X}$ و $\mathbf{Y}$ نسبت به θ_k برابر با صفر است، یعنی $IC_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)=0$، اگر حداقل یکی از دو شرط زیر برقرار باشد:

الف) $\mathbf{X}$ و $\mathbf{Y}$ مستقل باشند،

ب) $f_{\mathbf{X};\boldsymbol{\theta}}$ یا $f_{\mathbf{Y};\boldsymbol{\theta}}$ به θ_k بستگی نداشته باشند. ■

یک نتیجه مستقیم از قضیه ۳.۲ این است که مقدار همبستگی اطلاع بین آماره‌های $T(\mathbf{X})$ و $U(\mathbf{X})$ برابر با صفر است، یعنی $IC_{T(\mathbf{X}),U(\mathbf{X});\boldsymbol{\theta}}(\theta_k)=0$ هرگاه یکی از حالت‌های زیر برقرار باشد:

الف) دو آماره $T(\mathbf{X})$ و $U(\mathbf{X})$ مستقل باشند.

ب) حداقل یکی از آماره‌های $T(\mathbf{X})$ و $U(\mathbf{X})$ نسبت به پارامتر مجهول θ_k کمکی باشند.

۳. رویکرد جدید اطلاع متقابل برای اطلاع فیشر

در این بخش اطلاع فیشر متقابل در دو نوع مختلف تعریف می‌شود و ثابت می‌شود شرط لازم و کافی برای استقلال دو متغیر تصادفی، صفر بودن این معیار است.

سپس همبستگی اطلاع شرطی و در پی آن اطلاع فیشر متقابل شرطی نیز در دو نوع مختلف معرفی خواهند شد. برای هر یک از این مفاهیم خواص و روابط بین این مقادیر با اطلاع فیشر و اطلاع فیشر شرطی و همبستگی اطلاع بیان می‌شود. در پایان بخش به کمک تعمیم‌های اطلاع فیشر، دو تعمیم برای اطلاع فیشر متقابل ارائه می‌شود و با استفاده از تعمیم‌های اطلاع فیشر نسبی، سه تعمیم دیگر برای اطلاع فیشر متقابل معرفی می‌شود.

اطلاع فیشر متقابل نوع I به صورت زیر تعریف می‌شود:

$$M^{I}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)=E_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\left[\tfrac{\partial}{\partial\theta_k}\ln\frac{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})}\right]^2$$
$$=\iint\left[\tfrac{\partial}{\partial\theta_k}\ln\frac{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})}\right]^2 f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y}\geq 0.$$

قضیه ۱.۳. فرض کنید $\mathbf{X}$ و $\mathbf{Y}$ دو متغیر تصادفی دلخواه باشند. تحت شرایط نظم،

$$M^{I}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)=I_{\mathbf{X}|\mathbf{Y};\boldsymbol{\theta}}(\theta_k)-I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)+2IC_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)=I_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\theta_k)-I_{\mathbf{Y};\boldsymbol{\theta}}(\theta_k)+2IC_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k). \quad \blacksquare$$

مثال ۱.۳. در ادامه مثال ۱.۱ اطلاع فیشر متقابل نوع I بین Z و X را به صورت زیر می‌توان نوشت:

$$\begin{aligned}M^{I}_{Z,X;\mu,\sigma^2,\eta^2}(\mu)&=I_{Z|X}(\mu)-I_{Z}(\mu)+2IC_{Z,X;\mu,\sigma^2,\eta^2}(\mu)\\&=I_{Y}(\mu)-I_{Z}(\mu)+2IC_{Z,X;\mu,\sigma^2,\eta^2}(\mu)\\&=\frac{1}{\eta^2}-\frac{1}{\sigma^2+\eta^2}+\frac{2}{\sigma^2+\eta^2}=\frac{1}{\sigma^2+\eta^2}+\frac{1}{\eta^2}. \quad \blacksquare\end{aligned}$$

همبستگی اطلاع شرطی نسبت به θ_k برای متغیرهای تصادفی $\mathbf{X}$ و $\mathbf{Y}$ به شرط $\mathbf{Z}$ به صورت زیر تعریف می‌شود:

$$IC_{\mathbf{X},\mathbf{Y}|\mathbf{Z};\boldsymbol{\theta}}(\theta_k)=\iiint \tfrac{\partial}{\partial\theta_k}\ln f_{\mathbf{X}|\mathbf{Z};\boldsymbol{\theta}}(\mathbf{x}|\mathbf{z})\tfrac{\partial}{\partial\theta_k}\ln f_{\mathbf{Y}|\mathbf{Z};\boldsymbol{\theta}}(\mathbf{y}|\mathbf{z})f_{\mathbf{X},\mathbf{Y},\mathbf{Z};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y},\mathbf{z})d\mathbf{x}d\mathbf{y}d\mathbf{z}.$$

همچنین اطلاع فیشر متقابل شرطی نوع I متغیرهای تصادفی $\mathbf{X}$ و $\mathbf{Y}$ به شرط $\mathbf{Z}$ مطابق با فرمول زیر تعریف می شود:

$$M^{I}_{\mathbf{X},\mathbf{Y}|\mathbf{Z};\boldsymbol{\theta}}(\theta_k)=\iiint\left[\tfrac{\partial}{\partial\theta_k}\ln\frac{f_{\mathbf{X},\mathbf{Y}|\mathbf{Z};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y}|\mathbf{z})}{f_{\mathbf{X}|\mathbf{Z};\boldsymbol{\theta}}(\mathbf{x}|\mathbf{z})f_{\mathbf{Y}|\mathbf{Z};\boldsymbol{\theta}}(\mathbf{y}|\mathbf{z})}\right]^2 f_{\mathbf{X},\mathbf{Y},\mathbf{Z};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y},\mathbf{z})d\mathbf{x}d\mathbf{y}d\mathbf{z}.$$

لذا تحت شرایط نظم و مشابه با قضیه ۱.۳ داریم:

$$\begin{aligned}M^{I}_{\mathbf{X},\mathbf{Y}|\mathbf{Z};\boldsymbol{\theta}}(\theta_k)&=I_{\mathbf{X}|\mathbf{Y},\mathbf{Z};\boldsymbol{\theta}}(\theta_k)-I_{\mathbf{X}|\mathbf{Z};\boldsymbol{\theta}}(\theta_k)+2IC_{\mathbf{X},\mathbf{Y}|\mathbf{Z};\boldsymbol{\theta}}(\theta_k)\\&=I_{\mathbf{Y}|\mathbf{X},\mathbf{Z};\boldsymbol{\theta}}(\theta_k)-I_{\mathbf{Y}|\mathbf{Z};\boldsymbol{\theta}}(\theta_k)+2IC_{\mathbf{X},\mathbf{Y}|\mathbf{Z};\boldsymbol{\theta}}(\theta_k).\end{aligned}$$

به طور مشابه اطلاع فیشر متقابل نوع II نیز به صورت زیر تعریف می‌شود:

$$M^{II}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)=-E_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\left[\tfrac{\partial^2}{\partial\theta_k^2}\ln\frac{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})}\right]=-\iint\tfrac{\partial^2}{\partial\theta_k^2}\ln\frac{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})}f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y}$$

و با استفاده از قاعده زنجیره‌ای برای اطلاع فیشر ثابت می‌شود:

$$M^{II}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)=I_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)-I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)-I_{\mathbf{Y};\boldsymbol{\theta}}(\theta_k)=I_{\mathbf{X}|\mathbf{Y};\boldsymbol{\theta}}(\theta_k)-I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)=I_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\theta_k)-I_{\mathbf{Y};\boldsymbol{\theta}}(\theta_k)$$

که در آن تساوی اول از رابطه زیر نتیجه شده است:

$$\begin{aligned} M^{II}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k) &= -\iint f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})\frac{\partial^2}{\partial\theta_k^2}\ln\Big(\frac{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{y})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})}\Big)d\mathbf{x}d\mathbf{y} \\ &= -\iint f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})\frac{\partial^2 \ln f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}{\partial\theta_k^2}d\mathbf{x}d\mathbf{y} + \iint f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})\frac{\partial^2 \ln f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})}{\partial\theta_k^2}d\mathbf{x}d\mathbf{y} \\ &\quad + \iint f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})\frac{\partial^2 \ln f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})}{\partial\theta_k^2}d\mathbf{x}d\mathbf{y} \\ &= I_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k) - I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k) - I_{\mathbf{Y};\boldsymbol{\theta}}(\theta_k). \end{aligned}$$

با توجه به تعریف اطلاع فیشر متقابل و مقایسه آن با تعریف اطلاع فیشر نسبی به سادگی می‌توان نتیجه گرفت:

$$M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k) = D^i\left(f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}} \,||\, f_{\mathbf{X};\boldsymbol{\theta}}f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k),\ \ i = I, II.$$

از این‌رو با توجه به خواص اطلاع فیشر نسبی، شرط لازم و کافی برای استقلال متغیرهای تصادفی $\mathbf{X}$ و $\mathbf{Y}$ آن است که $M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k) = 0$، $i = I, II$. این ادعا را به روشی دیگر نیز می‌توان اثبات کرد. به عنوان نمونه فرض کنید $\mathbf{X}$ و $\mathbf{Y}$ متغیرهای تصادفی مستقل باشند. با استفاده از قضیه‌های ۳.۲ و ۱.۳ داریم:

$$M^{I}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k) = I_{\mathbf{X}|\mathbf{Y};\boldsymbol{\theta}}(\theta_k) - I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k) + 2IC_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k) = I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k) - I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k) = 0.$$

همچنین با توجه به این که در اطلاع فیشر شرطی برای متغیرهای تصادفی مستقل، شرط حذف می‌شود، می‌توان نتیجه گرفت $M^{II}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k) = I_{\mathbf{X}|\mathbf{Y};\boldsymbol{\theta}}(\theta_k) - I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k) = I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k) - I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k) + 0 = 0$ و از این رو برهان کامل می‌شود.

مثال ۲.۳. در ادامه مثال ۱.۱، اطلاع فیشر متقابل نوع II بین Z و X برابر است با:

$$M^{II}_{Z,X;\mu,\sigma^2,\eta^2}(\mu) = I_{Z|X}(\mu) - I_Z(\mu) = I_Y(\mu) - I_Z(\mu) = \frac{1}{\eta^2} - \frac{1}{\sigma^2+\eta^2} = \frac{\sigma^2}{\eta^2(\sigma^2+\eta^2)}. \ \blacksquare$$

در گزاره زیر ارتباط بین اطلاع فیشر متقابل نوع I و II بیان می‌شود.

گزاره ۱.۳. فرض کنید $\mathbf{X}$ و $\mathbf{Y}$ متغیرهای تصادفی دلخواه باشند. در این صورت:

$$M^{I}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k) = M^{II}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k) + 2IC_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k). \ \blacksquare$$

با توجه به گزاره ۱.۳ و قضیه ۳.۲، به سادگی می‌توان نتیجه گرفت، $M^{I}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k) = M^{II}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)$ اگر حداقل یکی از شرایط زیر برقرار باشد:

الف) $\mathbf{X}$ و $\mathbf{Y}$ مستقل باشند،

ب) $f_{\mathbf{X};\boldsymbol{\theta}}$ یا $f_{\mathbf{Y};\boldsymbol{\theta}}$ به θ_k بستگی نداشته باشند،

که نشان می‌دهد با برقراری هر یک از حالت‌های (الف) یا (ب)، هر دو معیار اطلاع فیشر متقابل نوع I و II با هم برابر هستند. قابل ذکر است که تحت شرط (الف)، $M^{I}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)=M^{II}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)=0$، در صورتی که تحت شرط (ب) دو معیار اطلاع فیشر متقابل نوع I و II مساوی هستند ولی لزوماً مقدار آنها برابر با صفر نیست.

مثال ۳.۳. با توجه به گزاره ۱.۳، می‌توان اطلاع فیشر متقابل II بین Z و X مطرح شده در مثال ۲.۳ را به صورت زیر نیز محاسبه کرد:

$$\begin{aligned} M^{II}_{Z,X;\mu,\sigma^2,\eta^2}(\mu) &= M^{I}_{Z,X;\mu,{}^2,\eta^2}(\mu)-2IC_{Z,X;\mu,\sigma^2,\eta^2}(\mu)\\ &=\frac{1}{\sigma^2+\eta^2}+\frac{1}{\eta^2}-2\frac{1}{\sigma^2+\eta^2}=\frac{\sigma^2}{\eta^2(\sigma^2+\eta^2)}. \ \blacksquare \end{aligned}$$

یادآوری می‌شود، نظر به این که $M^{I}\geq 0$، بسته به مقدار IC، علامت M^{II} می‌تواند مثبت یا منفی باشد.

در [۲۴،۶،۴،۲]، تعمیم‌های از اطلاع فیشر معرفی شده است که به کمک آنها می‌توان اطلاع فیشر متقابل را تعمیم داد که به عنوان نمونه، در ادامه برخی تعمیم‌های اطلاع فیشر متقابل ارائه می‌شود.

در [۴] مفهوم اطلاع فیشر نوع $i=I,II$ از مرتبه $s\geq 1$ یعنی

$$I^{I,(s)}_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)=\left[E_{\mathbf{X};\boldsymbol{\theta}}\left|\frac{\partial}{\partial\theta_k}\ln f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})\right|^{\frac{s}{s-1}}\right]^{s-1}=\left[\int\left|\frac{\partial}{\partial\theta_k}\ln f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})\right|^{\frac{s}{s-1}}f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})\mathrm{d}\mathbf{x}\right]^{s-1},$$

$$I^{II,(s)}_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)=-\left[E_{\mathbf{X};\boldsymbol{\theta}}\left(\frac{\partial^s}{\partial\theta_k^s}\ln f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})\right)^{\frac{1}{s-1}}\right]^{s-1}=-\left[\int\left(\frac{\partial^s}{\partial\theta_k^s}\ln f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})\right)^{\frac{1}{s-1}}f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})\mathrm{d}\mathbf{x}\right]^{s-1}$$

ارائه شد که در حالت $s=2$منجر به تعریف کلاسیک اطلاع فیشر یعنی $I^{I,(2)}_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)=I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)$ می‌شود. بوکِی [۴]، ثابت کرد برای هر $s\geq 2$، $I^{I,(s)}_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)$ یک تابع محدب نسبت به $f_{\mathbf{X};\boldsymbol{\theta}}$ است و برای $s\geq 1$ و آماره دلخواه $T(\mathbf{X})$داریم $I^{I,(s)}_{T(\mathbf{X});\boldsymbol{\theta}}(\theta_k)\leq I^{I,(s)}_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)$ و شرط لازم و کافی برای برقراری تساوی این است که $T(\mathbf{X})$ یک آماره بسنده باشد.

اکنون می‌توان از تعمیم [۴] برای اطلاع فیشر ایده گرفت و اطلاع فیشر متقابل نوع $i=I,II$ از مرتبه$s\geq 1$ بین متغیرهای تصادفی $\mathbf{X}$ و $\mathbf{Y}$ را تعریف کرد:

$$\begin{aligned} M^{I,(s)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)&=\left[E_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\left(\tfrac{\partial}{\partial\theta_k}\ln\frac{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})}\right)^{\frac{s}{s-1}}\right]^{s-1}\\ &=\left[\iint\left(\frac{\partial}{\partial\theta_k}\ln\frac{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})}\right)^{\frac{s}{s-1}}f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y}\right]^{s-1}, \end{aligned}$$

$$M^{II,(s)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k) = -\Big[E_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\Big(\tfrac{\partial^s}{\partial\theta_k^s}\ln\frac{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})}\Big)^{\frac{1}{s-1}}\Big]^{s-1}$$

$$= -\Big[\iint\Big(\frac{\partial^s}{\partial\theta_k^s}\ln\frac{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})}\Big)^{\frac{1}{s-1}}f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y}\Big]^{s-1}$$

که در حالت خاص $M^{i,(2)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k) = M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)$، $i=I,II$.

در [۲]، مفهوم جدیدی به نام اطلاع فیشر تعمیم‌یافته مرتبه (β,q) یعنی

$$I^{(\beta,q)}_{\mathbf{X};\boldsymbol{\theta}}(\theta_k) = E_{\mathbf{X};\boldsymbol{\theta}}\Big[\big|\tfrac{\partial}{\partial\theta_k}\ln f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})\big|^{\beta} f^{\beta(q-1)}_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})\Big] = \int\left|\frac{\frac{\partial}{\partial\theta_k}f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})}\right|^{\beta} f^{\beta(q-1)+1}_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})\mathrm{d}\mathbf{x}$$

معرفی شد که در حالت خاص $(\beta,q)=(2,1)$ برابر همان اطلاع فیشر است. اکنون می‌توان از تعمیم [۲] برای اطلاع فیشر ایده گرفت و اطلاع فیشر متقابل نوع $i=I,II$ از مرتبه (β,q) بین متغیرهای تصادفی $\mathbf{X}$ و $\mathbf{Y}$ را تعریف کرد:

$$M^{I}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}},(\beta,q)(\theta_k) = E_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\left[\left(\tfrac{\partial}{\partial\theta_k}\ln\frac{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})}\right)^{\beta} f^{\beta(q-1)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})\right]$$

$$= \iint\left(\tfrac{\partial}{\partial\theta_k}\ln\frac{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})\,f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})}\right)^{\beta} f^{\beta(q-1)+1}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y}\ ,$$

$$M^{II,(\beta,q)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k) = -E_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\Big[\tfrac{\partial^\beta}{\partial\theta_k^\beta}\Big(\ln\frac{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})}\Big)f^{\beta(q-1)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})\Big]$$

$$= -\iint \tfrac{\partial^\beta}{\partial\theta_k^\beta}\Big(\ln\frac{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})f_{\mathbf{y};\boldsymbol{\theta}}(\mathbf{y})}\Big)f^{\beta(q-1)+1}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y}$$

و به راحتی مشاهده می‌شود $M^{i,(2,1)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k) = M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)$، $i=I,II$.

در بخش دوم تعمیم‌هایی برای اطلاع فیشر نسبی ارائه شد که به طور مشابه می‌توان از آنها بهره جست و اطلاع فیشر متقابل را تعمیم داد که در ذیل به ذکر این موارد اشاره می‌شود.

به کمک ایده [۵] که توسط آن φ-اطلاع فیشر نسبی تعمیم داده شد، می‌توان تعمیم دیگری برای اطلاع فیشر متقابل در نظر گرفت. برای رسیدن به این منظور به ازای تابع محدب با مقادیر حقیقی φ، معیار φ-اطلاع فیشر متقابل نوع $i=I,II$ بین متغیرهای تصادفی $\mathbf{X}$ و $\mathbf{Y}$ به صورت زیر تعریف می‌شود:

$$M^{I}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta},\varphi}(\theta_k)=E_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\left[\tfrac{\partial}{\partial\theta_k}\varphi\Big(\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})}{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}\Big)\right]^2$$
$$=\iint\left[\tfrac{\partial}{\partial\theta_k}\varphi\Big(\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})}{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}\Big)\right]^2 f_{\mathbf{X},\mathbf{y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y},$$

$$M^{II}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta},\varphi}(\theta_k)=-E_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\left[\tfrac{\partial^2}{\partial\theta_k^2}\varphi\Big(\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})}{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}\Big)\right]$$
$$=-\int\tfrac{\partial^2}{\partial\theta_k^2}\varphi\Big(\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})}{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}\Big)f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y}$$

که در حالت خاص برای تابع محدب $\varphi_0(x)=-\ln(x)$، به راحتی می‌توان نوشت $M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta},\varphi_0}(\theta_k)=M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)$ ، $i=I,II$.

اکنون به کمک روابط (۱) و (۲)، ارتباط بین دو معیار φ-اطلاع فیشر نسبی و φ-اطلاع فیشر متقابل توسط رابطه زیر بیان می‌شود:

$$M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta},\varphi}(\theta_k)=D^{i}_{\varphi}\left(f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\,||\,f_{\mathbf{X};\boldsymbol{\theta}}f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k),\ \ i=I,II. \tag{۶}$$

به کمک معادله (۶) و با نظر گرفتن توابع محدب (۳) و (۴)، به راحتی می‌توان تعمیم‌های جدیدی برای اطلاع فیشر متقابل به شرح ذیل تعریف کرد. اطلاع فیشر متقابل نوع $i=I,II$ از مرتبه q به صورت زیر تعریف می‌شود:

$$M^{I',(q)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)=D^{I,(q)}\left(f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\,||\,f_{\mathbf{X};\boldsymbol{\theta}}f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)$$
$$=E_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\Big[\tfrac{\partial}{\partial\theta_k}\Big(\frac{\left[f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})/f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})\right]^{q-1}-1}{q-1}\Big)\Big]^2$$
$$=\int\Big[\tfrac{\partial}{\partial\theta_k}\Big(\frac{\left[f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})/f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})\right]^{q-1}-1}{q-1}\Big)\Big]^2 f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y},$$

$$M^{II',(q)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)=D^{II,(q)}\left(f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\,||\,f_{\mathbf{X};\boldsymbol{\theta}}f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)$$
$$=-E_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\Big[\tfrac{\partial^2}{\partial\theta_k^2}\Big(\frac{\left[f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})/f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})\right]^{q-1}-1}{q-1}\Big)\Big]$$
$$=-\int\tfrac{\partial^2}{\partial\theta_k^2}\Big(\frac{\left[f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})/f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})\right]^{q-1}-1}{q-1}\Big)f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y}.$$

با در نظر گرفتن تابع محدب $\varphi_q(x)$ در (۳) و به کار گیری (۶) داریم:

$$\begin{aligned}\lim_{q\to 1} M^{I',(q)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k) &= \lim_{q\to 1} D^{I,(q)}\left(f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}} \,||\, f_{\mathbf{X};\boldsymbol{\theta}} f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)\\ &= \lim_{q\to 1} D^{I}_{\varphi_q}\left(f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}} \,||\, f_{\mathbf{X};\boldsymbol{\theta}} f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)\\ &= D^{I}\left(f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}} \,||\, f_{\mathbf{X};\boldsymbol{\theta}} f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)\\ &= M^{I}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k),\end{aligned}$$

$$\begin{aligned}-\lim_{q\to 1} M^{II',(q)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k) &= -\lim_{q\to 1} D^{II,(q)}\left(f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}} \,||\, f_{\mathbf{X};\boldsymbol{\theta}} f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)\\ &= \lim_{q\to 1} D^{II}_{\varphi_q}\left(f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}} \,||\, f_{\mathbf{X};\boldsymbol{\theta}} f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)\\ &= D^{II}\left(f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}} \,||\, f_{\mathbf{X};\boldsymbol{\theta}} f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)\\ &= M^{II}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)\end{aligned}$$

و لذا مقدار حدی دو معیار $M^{I',(q)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)$ و $-M^{II',(q)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)$ وقتی $q\to 1$، به ترتیب برابر با اطلاع فیشر متقابل نوع $i=I,II$ است.

در انتها با استفاده از تابع محدب معرفی شده در (۴)، معیار τ-اطلاع فیشر متقابل نوع $i=I,II$ به صورت زیر تعریف می‌شود:

$$\begin{aligned}M^{I}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta};\varphi_\tau}(\theta_k) &= D_\tau^{I}\left(f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}} \,||\, f_{\mathbf{X};\boldsymbol{\theta}} f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)\\ &= E_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\Big[\tfrac{\partial}{\partial\theta_k}\Big(\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X}) f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})}{\bar{\tau} f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}\ln\Big(\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X}) f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})}{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}\Big)\\ &\qquad -\Big(\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X}) f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})}{\bar{\tau} f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}\Big)\ln\Big(\tau\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X}) f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})}{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}+\bar{\tau}\Big)\Big)\Big]^2,\\ &= \int\Big[\tfrac{\partial}{\partial\theta_k}\Big(\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x}) f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})}{\bar{\tau} f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}\ln\Big(\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x}) f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})}{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}\Big)\\ &\qquad -\Big(\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x}) f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})}{\bar{\tau} f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}\Big)\ln\Big(\tau\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x}) f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})}{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}+\bar{\tau}\Big)\Big)\Big]^2 f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})\,d\mathbf{x}\,d\mathbf{y},\end{aligned}$$

$$\begin{aligned}M^{II}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta};\varphi_\tau}(\theta_k) &= D^{II}_{\tau}\left(f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}} \,\|\, f_{\mathbf{X};\boldsymbol{\theta}} f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)\\
&= -E_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}\Big[\tfrac{\partial^2}{\partial\theta_k^2}\Big(\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})}{\bar{\tau}f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}\ln\Big(\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})}{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}\Big)\\
&\qquad -\Big(\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})}{\bar{\tau}f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}\Big)\ln\Big(\tau\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{X})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{Y})}{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{X},\mathbf{Y})}+\bar{\tau}\Big)\Big)\Big]\\
&= -\int \tfrac{\partial^2}{\partial\theta_k^2}\Big[\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})}{\bar{\tau}f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}\ln\Big(\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})}{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}\Big)\\
&\qquad -\Big(\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})}{\bar{\tau}f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}\Big)\ln\Big(\tau\frac{f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})f_{\mathbf{Y};\boldsymbol{\theta}}(\mathbf{y})}{f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})}+\bar{\tau}\Big)\Big]f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\mathbf{x},\mathbf{y})d\mathbf{x}d\mathbf{y}.\end{aligned}$$

با در نظر گرفتن تابع محدب φ_τ در (۴) و به کار گیری (۶) داریم:

$$\begin{aligned}\lim_{\tau\to 1} M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta};\varphi_\tau}(\theta_k) &= \lim_{\tau\to 1} D^{I}_{\tau}\left(f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}} \,\|\, f_{\mathbf{X};\boldsymbol{\theta}} f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)\\
&= \lim_{\tau\to 1} D^{I}_{\varphi_\tau}\left(f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}} \,\|\, f_{\mathbf{X};\boldsymbol{\theta}} f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)\\
&= D^{i}\left(f_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}} \,\|\, f_{\mathbf{X};\boldsymbol{\theta}} f_{\mathbf{Y};\boldsymbol{\theta}}\right)(\theta_k)\\
&= M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k), \quad i = I, II\end{aligned}$$

و لذا مقدار حدی دو معیار اطلاع فیشر متقابل تعمیم‌یافته $M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta};\varphi_\tau}(\theta_k)$، $i = I, II$ وقتی $\tau \to 1$ ، به ترتیب برابر با اطلاع فیشر متقابل نوع $i = I, II$ است.

به منظور جمع‌بندی تعمیم‌های ارائه‌شده برای اطلاع فیشر متقابل، مطالب بیان شده در قضیه زیر خلاصه می‌شود.

قضیه ۲.۳. هر کدام از حالت‌های زیر یک تعمیم برای اطلاع فیشر متقابل $M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)$، $i = I, II$ است:

الف) اطلاع فیشر متقابل نوع $i = I, II$ از مرتبه $s \geq 1$، $M^{i,(s)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)$:

$$M^{i,(2)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k) M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k), \;\; i = I, II.$$

ب) اطلاع فیشر متقابل نوع $i = I, II$ از مرتبه (β, q)، $M^{i,(\beta,q)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)$:

$$M^{i,(2,1)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k) = M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k), \;\; i = I, II.$$

ج) φ-اطلاع فیشر متقابل نوع $i = I, II$ برای تابع محدب φ، $M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta},\varphi}(\theta_k)$:

$$M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta},-\ln}(\theta_k) = M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k), i = I, II.$$

د) φ-اطلاع فیشر متقابل نوع $i = I', II'$ از مرتبه q، $M^{I',(q)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)$:

$$\lim_{q\to 1}\left| M^{i',(q)}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)\right| = M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k), \; i = I, II.$$

ه) τ-اطلاع فیشر متقابل نوع $i=I,II$ برای $\tau\in[0,1]$، $M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta};\varphi_\tau}(\theta_k)$:

$$\lim_{\tau\to 1} M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta};\varphi_\tau}(\theta_k)=M^{i}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k),\ i=I,II. \ \blacksquare$$

۴. خواص دیگر در زنجیرهای مارکف و یافتن کران بالا برای احتمال خطای برآورد

ابتدا در این بخش بررسی می‌شود که شرطی کردن باعث افزایش اطلاع فیشر می‌شود. سپس در حالتی که سه متغیر از یک زنجیر مارکف بیایند، نشان داده می‌شود اطلاع فیشر متقابل (از هر دو نوع) برای گذشته و آینده فرایند از مقادیر متناظر این کمیت برای حال و آینده فرایند کمتر است. همچنین در این حالت اطلاع فیشر شرطی آینده به شرط حال فرایند از اطلاع فیشر شرطی آینده به شرط گذشته فرایند بیشتر است. در پایان به کمک نامساوی فانو در نظریه اطلاع، یک کران بالا برای احتمال خطای برآورد بر حسب اطلاع فیشر شرطی بیان می‌شود.

فرض کنید بردار تصادفی $\mathbf{X}$ دارای امید ریاضی $\mu(\boldsymbol{\theta})=\int_{\mathcal{X}}\mathbf{x}f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})d\mathbf{x}$ و ماتریس واریانس-کوواریانس $\Sigma(\boldsymbol{\theta})=\int_{\mathcal{X}}\left(\mathbf{x}-\mu(\boldsymbol{\theta})\right)\left(\mathbf{x}-\mu(\boldsymbol{\theta})\right)^T f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})d\mathbf{x}$ باشد. اگر ماتریس اطلاع فیشر به صورت زیر تعریف شده باشد:

$$I_{\mathbf{X};\boldsymbol{\theta}}=\int_{\mathcal{X}}\left(\frac{\partial \ln f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})}{\partial\boldsymbol{\theta}}\right)^T\left(\frac{\partial \ln f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})}{\partial\boldsymbol{\theta}}\right)f_{\mathbf{X};\boldsymbol{\theta}}(\mathbf{x})d\mathbf{x}$$

در این صورت به شرط وجود $\frac{\partial\mu(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}$، داریم $I_{\mathbf{X};\boldsymbol{\theta}}\geq\left(\frac{\partial\mu(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)^T\Sigma^{-1}(\boldsymbol{\theta})\left(\frac{\partial\mu(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)$ و به طور مشابه:

$$\mathbf{x}^T\left(I_{\mathbf{X};\boldsymbol{\theta}}-\left(\frac{\partial\mu(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)^T\Sigma^{-1}(\boldsymbol{\theta})\left(\frac{\partial\mu(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right)\right)\mathbf{x}\geq 0$$

که نشان می‌دهد ماتریس داخل پرانتز، نیمه معین مثبت است و لذا اعضای قطری آن نامنفی‌اند [۸].

همچنین $I_{\mathbf{X};\boldsymbol{\theta}}(\theta_k)\geq\sum_{i=1}^{m}\sum_{j=1}^{m}\frac{\partial\mu_i}{\partial\theta_k}c_{ij}^{-1}\frac{\partial\mu_j}{\partial\theta_k}$ که در آن $\Sigma^{-1}(\boldsymbol{\theta})=\left(c_{ij}^{-1}\right)$.

زمیر [۳۴] نشان داد شرطی کردن باعث می‌شود که اطلاع فیشر افزایش یابد.

قضیه ۱.۴. اگر $f_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}$ به θ_k بستگی داشته باشد و $f_{\mathbf{X};\boldsymbol{\theta}}$ یا $f_{\mathbf{Y};\boldsymbol{\theta}}$ به θ_k وابسته نباشد، آن‌گاه $I_{\mathbf{Y};\boldsymbol{\theta}}(\theta_k)\leq I_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\theta_k)$.

برهان. با در نظر گرفتن قضیه ۳.۲، می‌توان نتیجه گرفت که $IC_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}=0$. لذا با توجه به قضیه ۱.۴ و همچنین نامنفی بودن اطلاع فیشر متقابل نوع I به سادگی می‌توان نوشت:

$$0\leq M^{I}_{\mathbf{X},\mathbf{Y};\boldsymbol{\theta}}(\theta_k)=I_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\theta_k)-I_{\mathbf{Y};\boldsymbol{\theta}}(\theta_k)$$

و از این رو برهان کامل می‌شود. ■

از قضیه ۱.۴ چند نکته را می‌توان نتیجه گرفت که به طور خلاصه به آنها اشاره می‌شود:

۱) فرض کنید $f_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}$ به θ_k بستگی نداشته باشد. در این حالت $I_{\mathbf{Y}|\mathbf{X};\boldsymbol{\theta}}(\theta_k)=0$ و در پی آن نامساوی ذکرشده در قضیه ۱.۴ به تساوی تبدیل می‌شود. در حالت اخیر این مفهوم استنباط می‌شود که اگر اطلاع فیشر یک توزیع شرطی $\mathbf{Y}|\mathbf{X}$ به پارامتر مجهول بستگی نداشته باشد و حداقل یکی از متغیرهای $\mathbf{X}$ و $\mathbf{Y}$ به پارامتر مجهول وابسته نباشند، آن‌گاه اطلاع فیشر توزیع غیرشرطی $\mathbf{Y}$ نیز برابر با صفر است.

۲) فرض کنید $T(\mathbf{X})$ یک آماره دلخواه و $U(\mathbf{X})$ یک آماره کمکی برای θ_k باشد. دو حالت زیر قابل بررسی است:

الف) فرض کنید تابع چگالی شرطی $f_{U(\mathbf{X})|T(\mathbf{X});\boldsymbol{\theta}}$ به پارامتر θ_k بستگی داشته باشد. در این صورت طبق قضیه ۱.۴، نامساوی بدیهی $0=I_{U(\mathbf{X});\boldsymbol{\theta}}(\theta_k)\leq I_{U(\mathbf{X})|T(\mathbf{X});\boldsymbol{\theta}}(\theta_k)$ همواره برقرار است و در حالت خاص که آماره $T(\mathbf{X})$ بسنده است نامساوی به تساوی تبدیل می‌شود.

این نامساوی نشان می‌دهد اطلاع فیشر آماره کمکی که روی یک آماره دلخواه دیگر شرطی شده باشد از اطلاع فیشر آماره کمکی غیر شرطی بیشتر است و در همین شرطی کردن اطلاعاتی که راجع به پارامتر مجهول θ_k نهفته هستند بازیابی می شوند.

ب) فرض کنید تابع چگالی شرطی $f_{T(\mathbf{X})|U(\mathbf{X});\boldsymbol{\theta}}$ به پارامتر θ_k بستگی داشته باشد. در این صورت طبق قضیه ۱.۴، نامساوی $I_{T(\mathbf{X});\boldsymbol{\theta}}(\theta_k)\leq I_{T(\mathbf{X})|U(\mathbf{X});\boldsymbol{\theta}}(\theta_k)$ برقرار است و این نامساوی نشان می‌دهد که اطلاع فیشر آماره دلخواه مشروط روی یک آماره کمکی، از اطلاع فیشر آماره دلخواه که شرطی نشده بیشتر است و از این رو شرطی کردن حتی روی یک آماره کمکی (که به ظاهر اطلاعاتی راجع به پارامتر مجهول نمی‌دهد)، می‌تواند باعث افزایش اطلاع فیشر شود.

برای سه متغیر تصادفی دلخواه X، Y و Z که زنجیر مارکف را نشان می‌دهند، یعنی $X\to Y\to Z$ داریم $f_{X,Z|Y;\boldsymbol{\theta}}=f_{X|Y;\boldsymbol{\theta}}f_{Z|Y;\boldsymbol{\theta}}$ [۸]. حال اگر فقط $f_{Z|Y;\boldsymbol{\theta}}$ به θ_k بستگی داشته باشد، آن‌گاه

$$M^{i}_{X,Z;\boldsymbol{\theta}}(\theta_k)\leq M^{i}_{Y,Z;\boldsymbol{\theta}}(\theta_k),\ i=I,II. \qquad (۷)$$

نامساوی (۷) نشان می‌دهد تحت شرایط ذکرشده، اطلاع فیشر متقابل (از هر دو نوع) برای گذشته و آینده فرایند از اطلاع فیشر متقابل برای حال و آینده فرایند کمتر است. به عبارت دیگر در یک زنجیر مارکف، اطلاع فیشر متقابل مراحل دورتر از یک فرایند تصادفی نسبت به مراحل نزدیک‌تر، کمتر بوده و این حقیقت به یک نوعی نشان دهنده کمتر بودن وابستگی بین مراحل دورتر در یک زنجیر مارکف است. برای اثبات نامساوی (۷)، مثلاً در حالت $i=I$ به صورت زیر عمل می‌کنیم:

با در نظر گرفتن قضیه ۱.۳، قاعده زنجیره‌ای برای اطلاع فیشر و همچنین تعریف اطلاع فیشر متقابل نوع I می‌توان

نتیجه گرفت که:

$$
\begin{aligned}
M^{I}_{(X,Y),Z;\boldsymbol{\theta}}(\theta_k) &= I_{X,Y|Z;\boldsymbol{\theta}}(\theta_k) - I_{X,Y;\boldsymbol{\theta}}(\theta_k) + 2IC_{(X,Y),Z;\boldsymbol{\theta}}(\theta_k) \\
&= I_{X|Y,Z;\boldsymbol{\theta}}(\theta_k) + I_{Y|Z;\boldsymbol{\theta}}(\theta_k) - I_{X|Y;\boldsymbol{\theta}}(\theta_k) - I_{Y;\boldsymbol{\theta}}(\theta_k) + 2IC_{(X,Y),Z;\boldsymbol{\theta}}(\theta_k) \\
&= \Big(I_{X|Y,Z;\boldsymbol{\theta}}(\theta_k) - I_{X|Y;\boldsymbol{\theta}}(\theta_k) + 2IC_{X,Z|Y;\boldsymbol{\theta}}(\theta_k)\Big) - 2IC_{X,Z|Y;\boldsymbol{\theta}}(\theta_k)\Big) \\
&\quad + \Big(I_{Y|Z;\boldsymbol{\theta}}(\theta_k) - I_{Y;\boldsymbol{\theta}}(\theta_k) + 2IC_{Y,Z;\boldsymbol{\theta}}(\theta_k)\Big) - 2IC_{Y,Z;\boldsymbol{\theta}}(\theta_k) + 2IC_{(X,Y),Z;\boldsymbol{\theta}}(\theta_k) \\
&= M^{I}_{(X,Z)|Y;\boldsymbol{\theta}}(\theta_k) + M^{I}_{Y,Z;\boldsymbol{\theta}}(\theta_k) - 2IC_{X,Z|Y;\boldsymbol{\theta}} - 2IC_{Y,Z;\boldsymbol{\theta}}(\theta_k) + 2IC_{(X,Y),Z;\boldsymbol{\theta}}(\theta_k).
\end{aligned}
$$

به طور مشابه به سادگی می‌توان نوشت:

$$
\begin{aligned}
M^{I}_{(X,Y),Z;\boldsymbol{\theta}}(\theta_k) &= M^{I}_{(Y,Z)|X;\boldsymbol{\theta}}(\theta_k) + M^{I}_{X,Z;\boldsymbol{\theta}}(\theta_k) \\
&\quad - 2IC_{Y,Z|X;\boldsymbol{\theta}}(\theta_k) - 2IC_{X,Z;\boldsymbol{\theta}}(\theta_k) + 2IC_{(X,Y),Z;\boldsymbol{\theta}}(\theta_k).
\end{aligned}
$$

در معادله بالا عبارت‌های همبستگی اطلاع صفر هستند، زیرا همگی شامل مشتق‌های توابع چگالی بوده و $f_{Z|Y;\boldsymbol{\theta}}$ فقط به θ_k بستگی دارد. لذا

$$
M^{I}_{X,Z|Y;\boldsymbol{\theta}}(\theta_k) + M^{I}_{Y,Z;\boldsymbol{\theta}}(\theta_k) = M^{I}_{Y,Z|X;\boldsymbol{\theta}}(\theta_k) + M^{I}_{X,Z;\boldsymbol{\theta}}(\theta_k).
$$

با توجه به این که $M^{I}_{X,Z|Y;\boldsymbol{\theta}}(\theta_k) = 0$ (زیرا X و Z به شرط Y مستقل هستند) و همچنین $M^{I}_{Y,Z|X;\boldsymbol{\theta}}(\theta_k) \geq 0$، می‌توان نوشت $M^{I}_{X,Z;\boldsymbol{\theta}}(\theta_k) \leq M^{I}_{Y,Z;\boldsymbol{\theta}}(\theta_k)$ و برهان کامل می‌شود.

در قضیه زیر نشان داده می‌شود که در یک زنجیر مارکف، اطلاع فیشر شرطی آینده به شرط حال فرایند از اطلاع فیشر شرطی آینده به شرط گذشته فرایند بیشتر است.

قضیه ۲.۴. اگر $X \to Y \to Z$ یک زنجیر مارکف باشد، آن‌گاه $I_{Z|X;\boldsymbol{\theta}}(\theta_k) \leq I_{Z|Y;\boldsymbol{\theta}}(\theta_k)$.

برهان. طبق (۷) برای $i = II$ می‌توان نوشت $I_{Z|X;\boldsymbol{\theta}}(\theta_k) - I_{Z;\boldsymbol{\theta}}(\theta_k) \leq I_{Z|Y;\boldsymbol{\theta}}(\theta_k) - I_{Z;\boldsymbol{\theta}}(\theta_k)$ و در پی آن برهان کامل می‌شود. ■

در [۸] این حقیقت آمده است که تابع چگالی نرمال، آنتروپی دیفرانسیل تمام توابع چگالی با واریانس ثابت σ^2 را ماکزیمم می‌کند. بنابراین برای یک تابع چگالی دلخواه $f_{\mathbf{X};\boldsymbol{\theta}}$، برخی اطلاعات Y و برآوردگر $\hat{X}(Y)$، طبق نامساوی فانو [۸] داریم $E_X\left(X - \hat{X}(Y)\right)^2 \geq \frac{1}{2\pi e} e^{2h(X|Y)}$ که در آن $h(X|Y) = -\int_S f(x,y)\log f(x|y)\,dx\,dy$ آنتروپی دیفرانسیل شرطی X به شرط Y است. به منظور استفاده از این حقیقت برای اطلاع فیشر، متغیر تصادفی داده‌شده R، متغیر تصادفی مشاهده شده Y و برآوردگر $\hat{R} = g(Y)$ را در نظر می‌گیریم. هدف یافتن کران برای

$P((R-\hat{R})^2>\varepsilon)$ است. توجه کنید که $R\to Y\to\hat{R}$ یک زنجیر مارکف است و $\hat{R}$ بستگی به $\boldsymbol{\theta}$ دارد. اگر خطای برآورد را با $E=(R-\hat{R})^2$ نشان دهیم، می‌توان نشان داد برای هر ε و برخی مقادیر $\xi\in[\varepsilon,\infty)$ داریم:

$$P\left(E>\varepsilon\right)\le\frac{I_{R|\hat{R};\boldsymbol{\theta}}(\theta_k)}{I_{R|\hat{R},E=\xi;\boldsymbol{\theta}}(\theta_k)}.$$

برای اثبات نامساوی بالا به صورت زیر عمل می‌کنیم.

با استفاده از قاعده زنجیره‌ای برای اطلاع فیشر داریم:

$$\$I_{R,E|\hat{R};\boldsymbol{\theta}}(\theta_k)=I_{R|\hat{R},\,E;\boldsymbol{\theta}}(\theta_k)+I_{E|\hat{R};\boldsymbol{\theta}}(\theta_k)=I_{E|R,\,\hat{R};\boldsymbol{\theta}}(\theta_k)+I_{R|\hat{R};\boldsymbol{\theta}}(\theta_k).$$

با توجه به این حقیقت که $E\,|\,R,\hat{R};\boldsymbol{\theta}$ به پارامتر مجهول θ_k بستگی ندارد، به سادگی می‌توان نوشت $I_{E|R,\,\hat{R};\boldsymbol{\theta}}(\theta_k)=0$ و از این‌رو با توجه به نامنفی بودن اطلاع فیشر می‌توان نتیجه گرفت که:

$$I_{R|\hat{R},E;\boldsymbol{\theta}}(\theta_k)\le I_{R|\hat{R},E;\boldsymbol{\theta}}(\theta_k)+I_{E|\hat{R};\boldsymbol{\theta}}(\theta_k)=I_{R|\hat{R};\boldsymbol{\theta}}(\theta_k).$$

از طرف دیگر،

$$\begin{aligned}
I_{R|\hat{R},E;\boldsymbol{\theta}}(\theta_k)&=\int_{e\in E}^{\infty}\iint_{r\in R,\hat{r}\in\hat{R}|(r-\hat{r})^2=e}f_{R,\hat{R},E;\boldsymbol{\theta}}(r,\hat{r},e)\Big(\frac{\partial\ln I_{R|\hat{R},E;\boldsymbol{\theta}}(r\,|\,\hat{r},e)}{\partial\theta_k}\Big)^2drd\hat{r}de\\
&=\int_{e\in E|e\le\varepsilon}^{\infty}\iint_{r\in R,\hat{r}\in\hat{R}|(r-\hat{r})^2=e}f_{R,\hat{R},E;\boldsymbol{\theta}}(r,\hat{r},e)\Big(\frac{\partial\ln I_{R|\hat{R},E;\boldsymbol{\theta}}(r\,|\,\hat{r},e)}{\partial\theta_k}\Big)^2drd\hat{r}de\\
&\quad+\int_{e\in E|e>\varepsilon}^{\infty}\iint_{r\in R,\hat{r}\in\hat{R}|(r-\hat{r})^2=e}f_{R,\hat{R},E;\boldsymbol{\theta}}(r,\hat{r},e)\Big(\frac{\partial\ln I_{R|\hat{R},E;\boldsymbol{\theta}}(r\,|\,\hat{r},e)}{\partial\theta_k}\Big)^2drd\hat{r}de\\
&\ge\int_{e\in E|e>\varepsilon}^{\infty}\iint_{r\in R,\hat{r}\in\hat{R}|(r-\hat{r})^2=e}f_{R,\hat{R},E;\boldsymbol{\theta}}(r,\hat{r},e)\Big(\frac{\partial\ln I_{R|\hat{R},E;\boldsymbol{\theta}}(r\,|\,\hat{r},e)}{\partial\theta_k}\Big)^2drd\hat{r}de\\
&=\int_{e\in E|e>\varepsilon}^{\infty}\iint_{r\in R,\hat{r}\in\hat{R}|(r-\hat{r})^2=e}f_{R,\hat{R}|E;\boldsymbol{\theta}}(r,\hat{r}\,|\,e)f_{E;\boldsymbol{\theta}}(e)\Big(\frac{\partial\ln I_{R|\hat{R},E;\boldsymbol{\theta}}(r\,|\,\hat{r},e)}{\partial\theta_k}\Big)^2drd\hat{r}de\\
&=\int_{e\in E|e>\varepsilon}^{\infty}f_{E;\boldsymbol{\theta}}(e)\iint_{r\in R,\hat{r}\in\hat{R}|(r-\hat{r})^2=e}f_{R,\hat{R}|E;\boldsymbol{\theta}}(r,\hat{r}\,|\,e)\Big(\frac{\partial\ln I_{R|\hat{R},E;\boldsymbol{\theta}}(r\,|\,\hat{r},e)}{\partial\theta_k}\Big)^2drd\hat{r}de\\
&=\int_{e\in E|e>\varepsilon}^{\infty}f_{E;\boldsymbol{\theta}}(e)I_{R|\hat{R},\,E=e;\boldsymbol{\theta}}(\theta_k)de.
\end{aligned}$$

با استفاده از قضیه مقدار میانی، یک $\xi\in[\varepsilon,\infty)$ وجود دارد به طوری که

$$I_{R|\hat{R},\,E=\xi;\boldsymbol{\theta}}(\theta_k)\int_{e\in E|e>\varepsilon}^{\infty}f_{E;\boldsymbol{\theta}}(e)de=I_{R|\hat{R},\,E=\xi;\boldsymbol{\theta}}(\theta_k).\,P\{E>\varepsilon\}\le I_{R|\hat{R},E;\boldsymbol{\theta}}(\theta_k)\le I_{R|\hat{R};\boldsymbol{\theta}}(\theta_k),$$

بنابراین

$$P\{E > \varepsilon\} \leq \frac{I_{R|\hat{R};\boldsymbol{\theta}}(\theta_k)}{I_{R|\hat{R},E=\xi;\boldsymbol{\theta}}(\theta_k)}.$$

۵. بحث و نتیجه‌گیری

در این مقاله به کمک ساختار اطلاع فیشر، برخی مفاهیم نظریه اطلاع نظیر اطلاع متقابل و فاصله کولبک-لایبلر تعمیم داده شد. برای یک زنجیر مارکف نشان داده شد وقتی آماره از پارامتر برآوردشده دور شود، اطلاع فیشر متقابل کاهش پیدا می‌کند، همچنین وقتی اطلاع فیشر روی مقادیر نزدیک‌تر شرطی شود نسبت به حالتی که روی مقادیر دورتر شرطی شود بیشتر است.

بالاخره به عنوان یک کاربرد از قضیه فانو در نظریه اطلاع، یک کران بالا برای احتمال خطای برآورد بر حسب اطلاع فیشر شرطی معرفی شد.

در پایان دو مسأله باز برای محققانی که به این موضوع علاقه‌مند هستند پیشنهاد می‌شود. در حالت‌هایی که یک گروه توپولوژیک روی متغیرهای تصادفی عمل می‌کند، معیارهای اطلاع فیشر متقابل، اطلاع فیشر نسبی، همبستگی اطلاع، ضریب همبستگی اطلاع بین متغیرهای تصادفی تبدیل‌یافته و نسخه‌های شرطی آن‌ها مورد تحلیل قرار گیرد. همچنین به کمک تابع پیچش می‌توان معیارهای مذکور را برای ترکیب‌های خطی متغیرهای تصادفی نیز مورد بررسی قرار داد.

References

1. D. Basu, On Statistics Independent of a Complete Sufficient Statistic, Sankhya, **15** (1955), 377-380.
2. J. F. Bercher, Some properties of generalized Fisher information in the context of nonextensive thermostatistics, Physica A: Statistical Mechanics and its Applications, **392** (2013), 3140 -3154.
3. S. G. Bobkov, G. P. Chistyakov, and F. Gotze, Fisher information and the central limit theorem, Probab.Theory Relat. Fields, **159** (2014), 1-59.
4. D. E. Boekee, Generalized Fisher information with application to estimation problems, IFAC Proceedings, **10** (1974), 75-82.
5. B. Bouchon, and F. Pessoa, Logarithmic information of degree q linked with an extension of Fisher's information, Kybernetika, **21** (1985), 346-359.
6. A. Cianchi, E. Lutwak, D. Yang, and G. Zhang, A unified approach to Cramér–Rao inequalities, IEEE Trans. Inf. Theory, **60** (2013), 643-650.
7. M. Cohen, The Fisher Information and Convexity, IEEE Trans. Inf. Theory, **14** (1968), 591-592.
8. T. Cover, and J. Thomas, Elements of Information Theory, John Wiley and

Sons, Inc.: Hoboken, NJ, USA, 2006.

9. R. Fisher, Theory of Statistical Estimation, Proc. Camb. Philos. Soc. **22** (1925), 700-725.

10. B. R. Frieden, Science from Fisher Information: A Unification, Cambridge University Press: Cambridge, UK, 2004.

11. B. R. Frieden, and R. A. Gatenby, Principle of maximum Fisher information from Hardy's axioms applied to statistical systems, Phys. Rev. E, **88** (2013), 1-6.

12. B. R. Frieden and M. Petri, Motion-dependent levels of order in a relativistic universe, Phys. Rev. E, **86** (2012), 1-5.

13. B. R. Frieden, A. Plastino, A. R. Plastino, and B. H. Soffer, Fisher-Based Thermodynamics: Its Legendre Transform and Concavity Properties, Phys. Rev. E, **60** (1999), 48-53.

14. B. R. Frieden, A. Plastino, A. R. Plastino, and B. H. Soffer, Non-equilibrium thermodynamics and Fisher information: An illustrative example, Phys. Lett. A, **304** (2002), 73-78.

15. D. Guo, Relative Entropy and Score Function: New Information-Estimation Relationships through Arbitrary Additive Perturbation, In Proceedings of the IEEE International Symposium on Information Theory, Seoul, Korea, 28 June–3 July; (2009), 814-818.

16. S. Flego, F. Olivares, A. Plastino, and M. Casas, Extreme Fisher Information, Non-Equilibrium Thermodynamics and Reciprocity Relations, Entropy, **13** (2011), 184-194.

17. M. Hirata, A. Nemoto, and H. Yoshida, An integral representation of the relative entropy, Entropy, **14** (2012), 1469-1477.

18. R. V. Hogg, and A. T. Craig, Introduction to Mathematical Statistics, Prentice Hall: Upper Saddle River, NJ, USA, 1995.

19. S. Kullback, Information Theory and Statistics, Dover Publications Inc.: Mineola, NY, USA, 1968.

20. S. Kullback, and R. A. Leibler, On Information and Sufficiency, Ann. Math. Stat. **22** (1951), 79-86.

21. R. S. Langley, Probability Functionals for Self-Consistent and Invariant Inference: Entropy and Fisher Information, IEEE Trans. Inf. Theory, **59** (2013), 4397-4407.

22. E. L. Lehmann, and G. Casella, Theory of Point Estimation, 2nd Edition, Springer-Verlage, New York, 1998.

23. S. Luo, Maximum Shannon entropy, minimum Fisher information, and an elementary game, Found. Phys, **32** (2002), 1757-1772.

24. S. Lv, General Fisher information matrices of a random vector, Advances in Applied Mathematics, **89** (2017), 18-40.

25. A. Ly, M. Marsman, J. Verhagen, R. P. Grasman, and E. J. Wagenmakers, A tutorial on Fisher information, Journal of Mathematical Psychology, **80** (2002), 40-55.

26. F. Otto, and C. Villani, Generalization of an Inequality by Talagrand and Links with the Logarithmic Sobolev Inequality, J. Funct. Anal., **173** (2000), 361-400.

27. C. Park, and A. Basu, The generalized Kullback-Leibler divergence and robust inference, Journal of Statistical Computation and Simulation, **73** (2003), 311-332.

28. R. Pascanu, and Y. Bengio, Revisiting Natural Gradient for Deep Networks, Cornell University Library: Ithaca, NY, USA, (2014), 1-18.

29. A. R. Plastino, and A. Plastino, Symmetries of the Fokker-Planck equation and the Fisher-Frieden

arrow of time, Phys. Rev. E, **54** (1996), 4423-4426.

30. P. Sánchez-Moreno, A. Zarzo, and J. S. Dehesa, Jensen divergence based on Fisher's information, J. Phys. A: Math. Theor, **45** (2012), 125-305.
31. C. A. Shannon, Mathematical Theory of Communication, Bell Syst. Tech. J. , **27** (1948), 379-423.
32. Tsallis, C. Generalized entropy-based criterion for consistent testing, Phys. Rev., **58** (1998), 1442-1445.
33. R. C. Venkatesan, and A. Plastino, Legendre transform structure and extremal properties of the relative Fisher information, Phys. Lett. A, **378** (2014), 1341–1345.
34. R. A. Zamir, Proof of the Fisher Information Inequality Via a Data Processing Argument, IEEE Trans. Inf. Theory, **44** (1998), 1246-1250.
35. P. Zegers, A. Fuentes, and C. Alarcon, Relative Entropy Derivative Bounds, Entropy, **15** (2013), 2861-2873.