بسم الله الرحمن الرحیم

دانشکده علوم ریاضی

رشته آمار، گرایش احتمال

رساله دکتری

معیارهای عدم دقت در برخی از متغیرهای تصادفی ترتیبی

نگارنده: صفیه دانشی

استادان راهنما

دکتر احمد نزاکتی
دکتر سعید طهماسبی

شهریور ۱۳۹۸

شماره:

تاریخ:

ویرایش:

باسمه تعالی

مدیریت تحصیلات تکمیلی

فرم شماره ۱۲: صورت جلسه نهایی دفاع از رساله دکتری (Ph.D)

(ویژه دانشجویان ورودی های ۹۴ و ما قبل)

بدینوسیله گواهی می شود **خانم صفیه دانشی** دانشجوی دکتری رشته **آمار - احتمال** به شماره دانشجویی **۹۲۴۶۲۱۵** ورودی بهمن ماه سال ۹۲ در تاریخ **۹۸/۶/۱۱** از رساله **نظری** ■ / **عملی** ☐ خود با عنوان : **معیارهای عدم دقت در برخی از متغیرهای تصادفی ترتیبی**

دفاع و با اخذ نمره ......۱۹/۷۵...... به درجه : ......عالی...... نائل گردید.

الف) درجه عالی: نمره ۲۰-۱۹ ☑	ب) درجه بسیار خوب: نمره ۱۸/۹۹ – ۱۷ ☐
ج) درجه خوب: نمره۱۶/۹۹– ۱۵ ☐	د) غیر قابل قبول و نیاز به دفاع مجدد دارد ☐
ه) رساله نیاز به اصلاحات دارد ☐	

ردیف	هیئت داوران	نام و نام خانوادگی	مرتبه علمی	امضاء
۱	دکتر احمد نزاکتی	استاد راهنما	دانشیار	
۲	دکتر سعید طهماسبی	استاد راهنما	دانشیار	
۳	دکتر سیمیندخت براتپور	استاد مدعو خارجی	دانشیار	
۴	دکتر محمد آرشی	استاد مدعو داخلی	دانشیار	
۵	دکتر حسین باغیشنی	استاد مدعو داخلی	استادیار	
۶	دکتر ابراهیم هاشمی	سرپرست (نماینده) تحصیلات تکمیلی دانشکده	استاد	

مدیر محترم تحصیلات تکمیلی دانشگاه:

ضمن تأیید مراتب فوق مقرر فرمائید اقدامات لازم در خصوص انجام مراحل دانش آموختگی **خانم صفیه دانشی** بعمل آید.

نام و نام خانوادگی رئیس دانشکده : دکتر ابراهیم هاشمی

تاریخ و امضاء و مهر دانشکده:

تقدیم به:

روح پاک پدرم که عالمانه به من آموخت تا چگونه در عرصه زندگی، ایستادگی را تجربه نمایم.

به مادرم، دریای بی کران فداکاری و عشق که وجودم برایش همه رنج بود و وجودش برایم همه مهر.

به همسرم، اسطوره زندگیم، پناه خستگیم و امید بودنم.

و به پسرم، که وجودش شادی بخش و صفایش مایه آرامش من است.

سپاس‌گزاری

شکر و سپاس خدا را که بزرگترین امید و یاور در لحظه لحظه زندگیست.
بسی شایسته است از اساتید فرهیخته و فرزانه‌ام جناب آقای دکتر احمد نزاکتی و جناب آقای دکتر سعید طهماسبی، صمیمانه تشکر و قدردانی کنم که قطعاً بدون راهنمایی‌های ارزنده ایشان، این مجموعه به اتمام نمی‌رسید.
از جناب آقای دکتر آرشی، جناب آقای دکتر باغیشنی و سرکار خانم دکتر براتپور که زحمت داوری این پایان‌نامه را تقبل فرمودند سپاسگذارم.

صفیه دانشی
شهریور ۱۳۹۸

تعهد نامه

اینجانب **صفیه دانشی** دانشجوی دکتری رشته **آمار علوم ریاضی** دانشگاه صنعتی شاهرود، نویسنده پایان‌نامه با عنوان **معیارهای عدم دقت در برخی از متغیرهای تصادفی ترتیبی** ، تحت راهنمایی **دکتر احمد نزاکتی** و **دکتر سعید طهماسبی** متعهد می‌شوم:

- تحقیقات در این پایان‌نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های دیگر پژوهش‌گران، به مرجع مورد استفاده استناد شده است.
- مطالب این پایان‌نامه، تا کنون توسط خود، یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ‌جا ارایه نشده است.
- حقوق معنوی این اثر، به دانشگاه صنعتی شاهرود تعلق دارد، و مقالات مستخرج با نام ‘‘ دانشگاه صنعتی شاهرود ’’ یا “ Shahrood University of Technology “ به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به‌دست آوردن نتایج اصلی پایان‌نامه تاثیرگذار بوده‌اند، در مقالات مستخرج از پایان‌نامه رعایت می‌گردد.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که از موجود زنده (یا بافت‌های آنها) استفاده شده است، ضوابط و اصول اخلاقی رعایت شده است.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته (یا استفاده شده است)، اصل رازداری و اصول اخلاق انسانی رعایت شده است.

صفیه دانشی
شهریور ۱۳۹۸

مالکیت نتایج و حق نشر

- تمام حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم‌افزارها و تجهیزات ساخته شده) متعلق به دانشگاه صنعتی شاهرود می‌باشد. این مطلب باید به نحو مقتضی، در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در این پایان‌نامه بدون ذکر منبع مجاز نمی‌باشد.

چکیده

نظریه اطلاع یکی از موضوعات مهم در علوم است. امروزه اندازه‌های اطلاع و عدم دقت جایگاه ویژه‌ای در قابلیت اطمینان دارند. در این رساله اندازه‌های اطلاع و متغیرهای تصادفی ترتیبی را معرفی می‌کنیم و اندازه عدم دقت را بین توزیع nــامین رکورد بالا (پایین) و متغیر تصادفی اصلی در نظر می‌گیریم. همچنین اندازه عدم دقت بین توزیع iــامین آماره ترتیبی و متغیر تصادفی اصلی را بیان می‌کنیم. اندازه عدم دقت تجمعی گذشته (باقیمانده) وزنی را در مقادیر رکورد و آماره‌های ترتیبی مطرح می‌کنیم و عدم دقت را بین rــامین همراه آماره ترتیبی تعمیم‌یافته و متغیر تصادفی اصلی در خانواده‌ی مورنگسترن به‌دست می‌آوریم. برای این مفاهیم برخی ویژگی‌ها و نتایج مشخصه‌سازی مانند ارتباط با توابع قابلیت اطمینان دیگر، کران‌ها، ترتیب تصادفی و تاثیر تبدیل خطی را به‌دست می‌آوریم و نسخه‌ی پویای این اندازه‌ها را در نظر می گیریم. همچنین موضوع برآورد اندازه عدم دقت را با استفاده از میانگین عدم دقت تجمعی تجربی و عدم دقت گذشته (باقیمانده) تجمعی وزنی در رکوردهای پایین بررسی می‌کنیم.

کلمات کلیدی: اندازه عدم دقت، عدم دقت تجمعی، آنتروپی تجمعی، مقادیر رکورد، آماره‌های ترتیبی، عدم دقت وزنی، رویکرد تجربی.

فهرست مقالات مستخرج از پایان‌نامه

1. Daneshi S., Nezakati A. and Tahmasebi S. (2019), ”On weighted cumulative past (residual) inaccuracy for record values”, **Journal of Inequalities and Applications**, no. 1, pp 134.
2. Tahmasebi S. and Daneshi S. (2018), ”Measures of inaccuracy in record values”, **Communications in Statistics-Theory and Methods**, 47(24), pp. 6002-6018.
3. Tahmasebi S., Nezakati A. and Daneshi S. (2018), ”Results on Cumulative Measure of Inaccuracy in Record Values”, **Journal of Statistical Theory and Applications**, 17(1), pp. 15-28.
4. Daneshi S., Nezakati A. and Tahmasebi S. (2019), ”Weighted Cumulative Residual (Past) Inaccuracy For Minimum (Maximum) of Order Statistics”, accepted to **Statistics, Optimization and Information Computing**.
5. Daneshi S., Nezakati A., Tahmasebi S. and Longobardi M. (2019), ”Inaccuracy Measures for Concomitants of Generalized Order Statistics in Morgenstern Family”, submitted to **Note di Mathematica**.
6. Tahmasebi S., Daneshi S., and Salehi M. (2017), ”Cumulative residual inaccuracy in upper record values”, 3rd Seminar on Reliability Theory and its Applications, p 303, Mashhad, Iran.
7. Daneshi S., Nezakati A., Tahmasebi S. (2018), ”Cumulative residual inaccuracy for minimum and maximum of order statistics”, 4th Seminar on Reliability Theory and its Applications, p 41, Shiraz, Iran.
8. Daneshi S., Nezakati A., Tahmasebi S. (2018), ”Measures of Inaccuracy for Concomitants of Order Statistics and Record Values in Morgenstern Family”, 14th Iranian Statistics Conference, p 154, Shahrood, Iran.
9. Daneshi S., Nezakati A., Tahmasebi S. (2019), ”Measures of Inaccuracy for Concomitants of Generalized Order Statistics”, 5th Seminar on Reliability Theory and its Applications, p 75, Yazd, Iran.

۱۰. منصوری ف، طهماسبی س، علیزاده م، فرخ سرشت ف، و دانشی ص. (۱۳۹۸)، پنجمین سمینار تخصصی نظریه قابلیت اعتماد و کاربردهای آن، ”کاربردی از اطلاع کولبک- لیبلر تجمعی در پردازش تصاویر“، ص ۱۳۰، یزد.

پیشگفتار

نظریه‌ی اطلاع، مدلی ریاضی از شرایط و عوامل مؤثر در پردازش و انتقال داده‌ها و اطلاعات فراهم می‌آورد. این نظریه با آنتروپی[۱]، فشرده‌سازی اطلاعات و سایر موضوعات وابسته ارتباط دارد. آنتروپی در نظریه‌ی اطلاع معیاری عددی است که میزان اطلاعات یا میزان عدم حتمیت یک متغیر تصادفی را بیان می‌کند. هرچه آنتروپی یک متغیر تصادفی بیشتر باشد اطلاعات حاصل از مشاهده قطعی آن بیشتر خواهد بود. شانون[۲] در سال ۱۹۴۸ اندازه‌ای از عدم حتمیت را برای متغیرهای تصادفی گسسته معرفی کرد که همواره مثبت است. اما آنتروپی شانون در حالت پیوسته دارای چالش‌هایی از جمله امکان منفی شدن بود. برای برطرف کردن این چالش‌ها، رائو[۳] و همکارانش (۲۰۰۴) معیار جدیدی به نام آنتروپی باقیمانده‌ی تجمعی ارائه دادند که با جایگزینی تابع بقا به جای تابع چگالی، در دستور آنتروپی شانون به‌دست می‌آید. همچنین دی‌کرشنزو و لونگوبردی[۴] (۲۰۰۹) با جایگزینی تابع توزیع به جای تابع چگالی در دستور آنتروپی شانون، آنتروپی گذشته تجمعی را معرفی کردند.

در مسائل قابلیت اطمینان عموماً متغیر تصادفی بیانگر طول عمر سیستم است. بنابراین با داشتن اطلاع در مورد طول عمر سیستم، این اندازه‌ها کارآمد نیستند. برای این منظور ابراهیمی[۵] (۱۹۹۶) آنتروپی متغیر تصادفی طول عمر باقیمانده را برای سیستم‌هایی که حداقل تا زمان t عمر می‌کنند، معرفی کرد. سپس اسدی و زهره‌وند[۶] (۲۰۰۷) حالت پویای آنتروپی باقیمانده را ارائه دادند.

کریج[۷] در سال ۱۹۶۱ یک اندازه را پیشنهاد داد که با عنوان اندازه عدم دقت شناخته می‌شود. این اندازه می‌تواند به‌عنوان تعمیمی از آنتروپی شانون در نظر گرفته شود. این اندازه علاوه بر عدم حتمیت برآمد X، عدم حتمیت مربوط به انتخاب مدل را نیز در نظر می‌گیرد. مانند آنتروپی شانون، اندازه عدم دقت نیز به اندازه عدم دقت باقیمانده و گذشته تجمعی تعمیم داده شد، سپس حالت پویای آن مورد بررسی قرار گرفت. برخی نویسندگان بر روی مقدار عدم

[۱] Entropy
[۲] Shannon
[۳] Rao
[۴] Di Cresenzo and Longobardi
[۵] Ebrahimi
[۶] Asadi and Zohrevand
[۷] Kerridge

دقت در آماره‌های ترتیبی مطالعه کردند. تاپلیال و تانجا[۸] (۲۰۱۳) عدم دقت را برای آماره‌های ترتیبی معرفی و حالت پویای آن را بررسی کردند. همچنین تاپلیال و همکاران (۲۰۱۵) نشان دادند که اندازه عدم دقت پویا در آماره‌های ترتیبی، تابع توزیع را به‌طور یکتا مشخص می‌کند.
این رساله به پنج فصل تقسیم شده است. در فصل اول، به بررسی برخی مفاهیم اولیه و نکاتی می‌پردازیم که در ادامه از آن‌ها استفاده می‌نماییم و کاربردی از اطلاع کولبک‌ـ لایبلر تجمعی را در پردازش تصاویر ارائه می‌کنیم. در فصل دوم پس از بیان مقدمات، اندازه عدم دقت را برای رکوردهای بالا و پایین با ذکر چند مثال معرفی می‌کنیم و برخی از ویژگی‌ها و نتایج مشخصه‌سازی را برای حالت‌های تجمعی و تجمعی پویا بیان می‌کنیم. همچنین اندازه عدم دقت تجربی را برای رکوردهای پایین به‌دست می‌آوریم. در فصل سوم، بعد از معرفی و مرور کارهای قبلی در خصوص اندازه عدم دقت در آماره‌های ترتیبی، عدم دقت تجمعی وزنی را برای آماره ترتیبی مینیمم و ماکزیمم به‌دست می‌آوریم و بررسی می‌کنیم. در فصل چهارم، عدم دقت گذشته‌ی تجمعی وزنی و تجمعی وزنی پویا را برای nـ امین رکورد پایین بررسی می‌کنیم و همچنین عدم دقت باقیمانده‌ی تجمعی وزنی را برای nـ امین رکورد بالا مطرح می‌کنیم. در فصل پنجم، معیارهای عدم دقت و خصوصیات آن‌ها را در متغیرهای همراه آماره‌های ترتیبی تعمیم‌یافته مطرح می‌کنیم. همچنین عدم دقت گذشته‌ی تجمعی را برای همراه GOS به‌دست می‌آوریم و برخی رفتارهای حدی آن را بررسی می‌کنیم.

[۸] Thapliyal and Taneja

فهرست مطالب

فهرست تصاویر ع

فهرست جداول ف

۱ مقدمات و مفاهیم اولیه ۱

۱.۱ مقدمه ۱
۲.۱ آماره‌های ترتیبی ۳
۱.۲.۱ کاربرد آماره‌های ترتیبی ۳
۲.۲.۱ توزیع آماره‌های ترتیبی ۴
۳.۲.۱ آماره‌های رکوردی ۵
۴.۲.۱ کاربرد رکوردها ۵
۵.۲.۱ توزیع رکوردها ۶
۶.۲.۱ سانسور ۷
۷.۲.۱ سانسور از راست ۷
۸.۲.۱ سانسور از چپ ۸
۹.۲.۱ سانسور تصادفی ۸
۱۰.۲.۱ آماره‌های ترتیبی تعمیم‌یافته ۸
۳.۱ مفاهیم کلی عدم دقت ۹
۱.۳.۱ خواص عدم دقت ۱۰
۲.۳.۱ آنتروپی شانون ۱۲
۳.۳.۱ تابع تشخیص کولبک_لایبلر ۱۳
۴.۳.۱ مقدار عدم دقت بین دو توزیع طول عمر باقیمانده و گذشته ۱۵
۴.۱ مروری مختصر بر کارهای انجام‌شده توسط دیگران ۱۶

۲ عدم دقت تجمعی در مقادیر رکورد بالا و پایین ۱۹

۱.۲ مقدمه ۱۹

۲.۲ اندازه عدم دقت در رکوردهای بالا ۲۱
۳.۲ اندازه عدم دقت باقیمانده برای R_n ۲۴
۴.۲ عدم دقت باقیمانده‌ی تجمعی برای R_n ۲۷
۵.۲ اندازه عدم دقت تجمعی در رکوردهای پایین ۳۶
۶.۲ اندازه عدم دقت تجمعی پویا برای L_n ۴۴
۷.۲ برآورد عدم دقت تجمعی برای L_n ۴۵

۳ معیارهای عدم دقت در آماره‌های ترتیبی ۴۹
۱.۳ مقدمه ۴۹
۲.۳ اندازه‌ی عدم دقت برای $X_{i:n}$ ۵۱
۳.۳ خصوصیات اندازه عدم دقت برای $X_{i:n}$ ۵۲
۴.۳ خصوصیات اندازه عدم دقت باقیمانده برای $X_{i:n}$ ۵۵
۵.۳ نتایج مشخصه‌سازی ۵۶
۶.۳ عدم دقت باقیمانده‌ی تجمعی وزنی برای $X_{(1:n)}$ ۵۸
۷.۳ عدم دقت گذشته‌ی تجمعی وزنی برای $X_{(n:n)}$ ۶۵

۴ معیارهای عدم دقت گذشته (باقیمانده) تجمعی وزنی در مقادیر رکورد ۷۵
۱.۴ مقدمه ۷۵
۲.۴ عدم دقت گذشته‌ی تجمعی وزنی برای L_n ۷۶
۳.۴ عدم دقت گذشته‌ی تجمعی وزنی پویا برای L_n ۸۱
۴.۴ برآورد عدم دقت گذشته‌ی تجمعی وزنی برای L_n ۸۲
۵.۴ عدم دقت باقیمانده‌ی تجمعی وزنی برای R_n ۸۵

۵ معیارهای عدم دقت در متغیرهای همراه از آماره‌های ترتیبی تعمیم‌یافته ۹۵
۱.۵ مقدمه ۹۵
۲.۵ اندازه عدم دقت برای همراه GOS ۹۷
۳.۵ CPI بین $Y_{[r,n,m,k]}$ و Y ۱۰۲
۴.۵ CPI تجربی برای همراه GOS ۱۰۴

مراجع ۱۱۳

آ تعاریف ۱۱۹
۱.آ قضیه همگرایی مغلوب ۱۱۹
۲.آ قضیه گلیونکو_کانتلی ۱۱۹
۳.آ شاخص جینی ۱۲۰

آ.۴ قضیه فوبینی ۱۲۰

فهرست تصاویر

۱.۲ تابع $I(f_{R_n},f)$ بر حسب β برای ۵ $= n$ و ۲ $= \lambda$ ۲۲
۲.۲ تابع $I(f_{R_n},f)$ بر حسب λ برای ۵ $= n$ و ۲ $= \beta$ ۲۲
۳.۲ تابع $I(f_{R_n},f)$ برای ۳ $< n$ در توزیع نرمال استاندارد ۲۳
۴.۲ تابع $I(F_{L_n},F)$ بر حسب n در توزیع وایبل معکوس برای ۲ $= \beta = \alpha$ (راست)، توزیع یکنواخت بر روی $(۰,۱)$ (وسط) و توزیع نمایی با ۲ $= \lambda$ (چپ) ۳۷

۱.۵ تصاویر اسکن مغز ۱۰۸
۲.۵ میزان سطح خاکستری مرتب‌شده در شکل ۱.۵ . . . . ۱۰۹
۳.۵ تصاویر منتخب ۱۱۰
۴.۵ سطح خاکستری مرتب‌شده برای تصاویر شکل ۳.۵ . . . . ۱۱۰

فهرست جداول

۱.۲ مقادیر عددی $E[\hat{I}(F_{L_n},F)]$ و $Var[\hat{I}(F_{L_n},F)]$ برای توزیع نمایی. ۴۷
۲.۲ مقادیر عددی $E[\hat{I}(F_{L_n},F)]$ و $Var[\hat{I}(F_{L_n},F)]$ برای توزیع یکنواخت. . . ۴۷

۱.۳ مقادیر عددی $\mathbb{E}[\hat{\tilde{I}}^w(F_{X_{(n:n)}},F)]$ و $Var[\hat{\tilde{I}}^w(F_{X_{(n:n)}},F)]$ برای توزیع وایبل ۷۲
۲.۳ مقادیر عددی $\mathbb{E}[\hat{\tilde{I}}^w(F_{X_{(n:n)}},F)]$ و $Var[\hat{\tilde{I}}^w(F_{X_{(n:n)}},F)]$ برای توزیع بتا . . ۷۲

۱.۴ مقادیر عددی $\mathbb{E}[\hat{I}^w(F_{L_n},F)]$ و $Var[\hat{I}^w(F_{L_n},F)]$ برای توزیع وایبل . . . ۸۴
۲.۴ مقادیر عددی $\mathbb{E}[\hat{I}^w(F_{L_n},F)]$ و $Var[\hat{I}^w(F_{L_n},F)]$ برای توزیع بتا ۸۴

۱.۵ مقادیر عددی $\mathbb{E}[\hat{I}(F_{R_{[r]}},F_Y)]$ و $Var[\hat{I}(F_{R_{[r]}},F_Y)]$ برای MTBGED با
$\lambda = ۱$ ۱۰۵
۲.۵ مقادیر عددی $\mathbb{E}[\hat{I}(F_{R_{[r]}},F_Y)]$ و $Var[\hat{I}(F_{R_{[r]}},F_Y)]$ برای MTBUD با . . ۱۰۶
۳.۵ میانگین سطح خاکستری برای تصاویر شکل ۱.۵ . . . . ۱۰۹
۴.۵ اندازه کولبک_لایبلر تجمعی تجربی برای تصاویر شکل ۱.۵ . . . . ۱۰۹
۵.۵ میانگین سطح خاکستری برای تصاویر شکل ۳.۵ . . . . ۱۱۰
۶.۵ اندازه‌های کولبک_لایبلر تجمعی تجربی برای تصاویر شکل ۳.۵ . . . . . ۱۱۱

فهرست نمادها

تابع جرم احتمال	$P(.)$
تابع چگالی احتمال	$f(.)$
تابع توزیع تجمعی	$F(.)$
تابع بقا	$\bar{F}(.)$
عدم دقت	$I(f,g)$
آنتروپی شانون	$H(.)$
اطلاع کولبک_لایبلر	$K(f\|g)$
میانگین زمان غیرفعال	$\mu(t)$
n_امین رکورد بالا	R_n
n_امین رکورد پایین	L_n
شاخص جینی	$gini[.]$
آنتروپی تجمعی	$\mathcal{CE}(.)$
آنتروپی تجمعی تعمیم‌یافته	$\mathcal{CE}_{j+1}(.)$
آنتروپی باقیمانده‌ی تجمعی	$\mathcal{E}(.)$
آنتروپی باقیمانده‌ی تجمعی تعمیم‌یافته	$\mathcal{E}_{j+1}(.)$
عدم دقت وزنی	$I^w(f,g)$
آنتروپی تجمعی وزنی	$\mathcal{CE}^w(.)$
آنتروپی باقیمانده‌ی تجمعی وزنی	$\mathcal{E}^w(.)$
میانگین وزنی طول عمر باقیمانده	$M^w(.)$
میانگین زمان غیرفعال وزنی	$\mu^w(t)$
عدم دقت تعمیم یافته از مرتبه m	$I_m(f,g)$
عدم دقت تعمیم یافته‌ی وزنی از مرتبه m	$I_m^w(f,g)$
میانگین ترکیبی عمر باقیمانده	$m^c(t)$

فصل ۱

مقدمات و مفاهیم اولیه

۱.۱ مقدمه

امروزه آماره‌های مرتب در اکثر شاخه‌های آمار و احتمال مورد توجه پژوهشگان قرار گرفته‌اند. انتشار کتاب ساراهان و گرینبرگ[۱] (۱۹۶۲) که تحقیقات انجام گرفته از آغاز تا اوایل سال ۱۹۶۰ را در برداشت و همچنین کتاب دیوید[۲] (۱۹۷۰) و تجدید نظر بعدی وی در سال ۱۹۸۱، نقطه‌ی شروع انجام تحقیقات بیشتر در زمینه‌ی آماره‌های مرتب بود. محققینی همچون هارتر[۳] (۱۹۷۰)، گالامبوس[۴] (۱۹۷۸)، کاستیلو[۵] (۱۹۸۱) و بارنت و لویس[۶] (۱۹۸۴) نیز در این زمینه آثاری را به چاپ رسانده‌اند. کتاب‌های ''آماره‌های ترتیبی'' دیوید و ناگاراجا[۷] (۲۰۰۳) و همچنین''نخستین درس در آماره‌های ترتیبی'' آرنولد و همکاران[۸] (۱۹۹۲)، از جمله کتاب‌های معروفی هستند که به اهمیت آماره‌های ترتیبی در مسائل مربوط به چندک‌ها

---

[۱] Sarahan and Greenberg
[۲] David
[۳] Harter
[۴] Galambos
[۵] Castillo
[۶] Barnett and Lewis
[۷] David and Nagaraja
[۸] Arnold et al.

پرداخته‌اند. همچنین پارک[۹] (۱۹۹۶) به محاسبه ی اطلاع فیشر در آماره‌های مرتب پرداخته است.
محاسبه‌ی میزان آنتروپی در آماره‌های مرتب نیز توجه محققین زیادی را به خود جلب کرده است. پارک (۲۰۰۳) به تجزیه‌ی آنتروپی آماره‌های مرتب پرداخته است و همچنین ابراهیمی و همکاران (۲۰۰۴) ویژگی‌های آنتروپی در آماره‌های ترتیبی را بررسی کرده‌اند.

اما مواردی وجود دارند که همه‌ی داده‌ها ثبت نمی‌شوند و فقط کوچکترین و بزرگترین آنها نسبت به اهمیتشان ثبت می‌شوند. این نوع داده‌ها در متون آماری، به داده‌های رکوردی معروفند. مقادیر مربوط به رکوردها و زمان‌های بین این مقادیر همیشه مورد توجه بوده است. یک هواشناس می خواهد بداند که آیا بارش بی‌سابقه‌ی برف در یک منطقه‌ی خاص دوباره تکرار خواهد شد. کسانی که در مورد زلزله تحقیق می‌کنند همیشه از خود می‌پرسند که آیا زلزله‌ای بزرگتر و مخرب‌تر دوباره به وقوع خواهد پیوست. البته مقادیر رکوردی در میادین ورزشی ملموس‌ترند و ورزش‌کاران و ورزش‌دوستان با مفهوم رکورد به خوبی آشنا هستند. تنش‌های گوناگون چالش بزرگی است که پیش روی مهندسان عمران قرار دارد و به دنبال ترکیبی هستند که با بیشترین نیرو کمترین گسیختگی را داشته باشند. در تمام این موارد و موردهای مشابه با یک دنباله مشاهده متوالی مواجه‌ایم که مقادیر کوچک یا مقادیر بزرگ این مشاهدات به گونه‌ای از اهمیت خاصی برخوردار هستند. این مقادیر در جنبه‌های گوناگون زندگی ظاهر می شوند. اگرچه اغلب افراد از وجود قوانین احتمالی حاکم بر این مقادیر بی‌اطلاع هستند، اما گاهی اوقات از روی تجربه و مشاهده در مورد رکوردها، نه تنها اظهار نظر می‌کنند بلکه پیش‌بینی نیز می‌نمایند. یک احساس درونی به ما می‌گوید که رکوردها عاقبت شکسته خواهند شد. شاید درجه ی حرارت از گرم‌ترین روز دنیا تجاوز کند. شاید زلزله‌ای بیاید که رکورد شدت زلزله‌ی ۳۰ سال اخیر ایران شکسته شود و شاید رقیبی برای قهرمان اتومبیل رانی پیدا شود و ...
شکسته شدن رکوردها، چه رکوردهای بالا و چه رکوردهای پایین همیشه با نوعی تردید و ابهام همراه هستند. وجود این تردیدها و ابهامات باعث شد تا افرادی مثل چاندلر[۱۰] (۱۹۵۲) مباحث مربوط به رکوردها را مورد مطالعه قرار دهند و به تکوین نظریه‌ای بپردازند که به دنباله‌ای از متغیرهای تصادفی وابسته است. تلاش‌های چاندلر باعث شد تا یک الگوی احتمالی مناسب برای این پدیده پیدا شود. پس از معرفی موضوع رکوردها، طیف وسیعی از کارهای پژوهشی در این زمینه انجام گرفت و نتایج فراوانی به دست آمد. تعدادی از آماردانان بر روی مسائلی کار کردند که مربوط به رکوردها بود. ویلکس[۱۱] (۱۹۵۹) این سوال را مطرح کرد که چند مشاهده ی اضافی نیاز است تا باعث افزایش بیشینه ی n مشاهده گردد. رنی[۱۲] (۱۹۶۲)

[۹] Park
[۱۰] Chandler
[۱۱] Wilks
[۱۲] Renyi

نشان داد که زمان‌های رکوردها، متغیرهای تصادفی مستقل هستند. آرنولد[۱۳] (۱۹۹۸) به برآورد راستنمایی پارامترها، بر اساس رکوردها پرداخت. بازاک و بالاکریشنان[۱۴] (۲۰۰۳) بحث پیشگویی رکوردهای آینده را مطرح کردند.
در این فصل تلاش شده است تا مفاهیم اولیه، تعاریف و توزیع احتمال آماره‌های ترتیبی و رکوردی بیان شوند.

۲.۱ آماره‌های ترتیبی

متغیرهای تصادفی $X_1, \dots, X_n$ را در نظر بگیرید. اگر آن‌ها را به صورت صعودی مرتب کنیم به rـامین آن‌ها، rـامین آماره‌ی ترتیبی گفته می‌شود و آن را با نماد $X_{r:n}$ نشان می‌دهیم. به‌طوری که نامساوی $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ برای آن‌ها برقرار است.

۱.۲.۱ کاربرد آماره‌های ترتیبی

در قسمت‌های مختلف علم آمار، اعم از آمار توصیفی و استنباطی می‌توان نقش آماره‌های ترتیبی را به وضوح دید. باید توجه داشت در قسمت آمار توصیفی یکی از مهمترین شاخص‌های تمرکز در یک مجموعه از داده‌های کمی، میانگین است. اما در صورت وجود یک یا چند داده‌ی پرت در مجموعه‌ی داده‌ها، میانگین‌های اصلاح‌شده یا میانه می‌توانند معیارهای مناسبی برای تعیین تمرکز باشند که در محاسبه ی آن‌ها آماره‌های ترتیبی نقش عمده‌ای دارند. همچنین در صورت کیفی بودن داده‌ها، میانه بهترین شاخص تمرکزیابی است. در آمار استنباطی نیز مثال‌های زیادی از کاربرد آماره‌های ترتیبی وجود دارند. آرنولد و همکاران (۱۹۹۲) و دیوید و ناگاراجا (۲۰۰۳) مواردی از کاربردهای آماره‌های ترتیبی را بیان می‌کنند. در فهرست زیر به برخی از آن‌ها اشاره می‌کنیم.

۱. آماره‌ی ترتیبی kـام طول عمر مولفه‌های یک سیستم k از n، همان طول عمر سیستم می‌باشد. این یکی از کاربردهای آماره‌های ترتیبی در قابلیت اعتماد است. همچنین در تحلیل بقا به ویژه زمانی که با داده‌های سانسورشده مواجه می‌شویم نیز به آماره‌های ترتیبی نیاز داریم.

۲. گاهی اوقات بخصوص در استنباط‌های ناپارامتری، آماره‌های ترتیبی به عنوان آماره بسنده معرفی می شوند و برآوردگرهای نااریب با کمترین واریانس، فواصل اطمینان، تواناترین آزمون برای پارامترهای مجهول جامعه و همچنین Uـآماره‌ها بر پایه‌ی آن‌ها ساخته می‌شوند.

[۱۳] Arnold
[۱۴] Basak and Balakrishnan

۳. در موضوعات مربوط به کنترل کیفیت، برای بررسی در کنترل بودن تولید، اغلب از نمودار میانگین یا نمودار میانه و دامنه تغییرات آن‌ها استفاده می‌شود، که برای محاسبه میانه و دامنه تغییرات ملزم به استفاده از آماره‌های ترتیبی هستیم.

۴. برای محاسبه تابع توزیع تجربی و همچنین در آزمون‌های نیکویی برازش به منظور بررسی تفاوت توزیع تجربی و تابع توزیع فرضی از آماره‌های ترتیبی استفاده می‌کنیم.

۵. برای محاسبه و بررسی ویژگی‌های مختلف در داده‌های سانسورشده، از آماره‌های ترتیبی و خواص مربوط به آن‌ها استفاده می‌کنیم.

۶. از آماره‌های ترتیبی در بحث مشخصه‌سازی توزیع‌ها نیز استفاده می‌شود.

۲.۲.۱ توزیع آماره‌های ترتیبی

فرض کنید $X_1, \ldots, X_n$ متغیرهای تصادفی مستقل و هم‌توزیع از توزیع F باشند. تابع توزیع تجمعی آماره‌ی ترتیبی nــام و آماره‌ی نخست به ترتیب به‌صورت زیر هستند:

$$F_{n:n}(x) = [F(x)]^n,$$

و

$$F_{1:n}(x) = 1 - [1 - F(x)]^n.$$

به طور کلی تابع توزیع تجمعی $X_{r:n}$ به‌صورت زیر می‌باشد:

$$F_{r:n}(x) = \sum_{i=r}^{n} \binom{n}{i} F(x)^i [1 - F(x)]^{n-i}.$$

با فرض پیوسته بودن متغیرهای تصادفی $X_1, \ldots, X_n$، تابع چگالی حاشیه‌ای آماره‌ی ترتیبی nــام، آماره‌ی نخست و آماره‌ی ترتیبی rــام به ترتیب به‌صورت زیر هستند:

$$f_{n:n}(x) = nF^{n-1}(x) f(x) \quad x \in \mathcal{R},$$

$$f_{1:n}(x) = n[1 - F(x)]^{n-1} f(x) \quad x \in \mathcal{R},$$

و

$$f_{r:n}(x) = \frac{n!}{(r-1)!(n-r)!} F(x)^{i-1} [1 - F(x)]^{n-i} f(x) \quad x \in \mathcal{R}.$$

برای $x_1 \leq x_2$، تابع چگالی احتمال توام آماره‌های ترتیبی rــام و sــام عبارتست از:

$$f_{r,s:n}(x_1, x_2) = \frac{n!}{(r-1)!(s-r-1)!(n-s)!} F(x_1)^{r-1} [F(x_2) - F(x_1)]^{s-r-1} [1 - F(x_2)]^{n-s} f(x_1) f(x_2).$$

و نیز تابع چگالی احتمال توام $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ به‌صورت زیر است:

$$f_{1,\ldots,n:n}(x_1, \ldots, x_n) = n! \prod_{i=1}^{n} f(x_i) \quad x_1 \leq x_2 \leq \ldots \leq x_n.$$

۳.۲.۱ آماره‌های رکوردی

فرض کنید $X_۱, X_۲, ...$ یک دنباله از متغیرهای تصادفی پیوسته، مستقل و هم‌توزیع با تابع توزیع مشترک $F(x)$ و تابع چگالی $f(x)$ باشد.

تعریف ۱.۲.۱. مشاهده‌ی X_j را یک رکورد بالا می‌گوییم اگر مقدار آن از مقدار تمام مشاهدات قبلی بیشتر باشد به عبارت دیگر $\forall i < j$ ، $X_i < X_j$. مشابه تعریف رکوردهای بالا، مشاهده‌ی X_j را یک رکورد پایین می‌گوییم اگر مقدار آن از مقدار تمام مشاهدات قبلی خود کمتر باشد، یعنی $\forall i < j$ ، $X_i > X_j$.

تعریف ۲.۲.۱. شماره اندیس‌هایی را که در آن‌ها رکوردهای بالا و پایین ظاهر می‌شوند، زمان رکوردها می‌نامند و با U_j و L_j نشان داده و به‌صورت زیر تعریف می‌شوند:

$$U_۱ = ۱,\ U_j = \min\{k | k > U_{j-۱}, X_k > X_{U_{j-۱}}\}, j \geq ۲$$

$$L_۱ = ۱,\ L_j = \min\{k | k > L_{j-۱}, X_k < X_{L_{j-۱}}\}, j \geq ۲$$

تعریف ۳.۲.۱. فرض کنید N_n تعداد رکوردهای یک نمونه‌ی nتایی و Y_i یک متغیر تصادفی نشانگر باشد. فرآیند تعداد رکوردها به صورت زیر معرفی می‌شود:

$$N_n = \sum_{i=۱}^{n} Y_i,\ Y_i = \begin{cases} ۱ & \text{هرگاه } X_i \text{ رکورد باشد} \\ ۰ & \text{هرگاه } X_i \text{ رکورد نباشد} \end{cases} \qquad (۱.۱)$$

تعریف ۴.۲.۱. زمان بین رکوردها آماره‌ای است که با نماد Δ_r نشان داده می‌شود و تعریف آن عبارت است از:

$$\Delta_r = U_{r+۱} - U_r,\ r = ۱, ۲, ۳, \ldots \qquad \text{برای رکوردهای بالا}$$

$$\Delta_r = L_{r+۱} - L_r,\ r = ۱, ۲, ۳, \ldots \qquad \text{برای رکوردهای پایین}$$

تعریف ۵.۲.۱. دنباله‌ی مقادیر رکوردها با R_j نشان داده و به‌صورت زیر تعریف می‌شود:

$$R_j = X_{U_j} \qquad \text{مقادیر رکوردهای بالا}$$

$$R_j = X_{L_j} \qquad \text{مقادیر رکوردهای پایین}$$

۴.۲.۱ کاربرد رکوردها

نظریه رکوردها کاربردهای وسیع در زندگی واقعی دارند. رکوردها در مسائلی از قبیل آزمون طول عمر قطعات الکترونیکی گران قیمت یا آزمون استقامت الوارهای چوب، داده‌های مربوط به مسابقات ورزشی، هواشناسی، ژئوفیزیک، زلزله نگاری، تعمیر مینیمال در مسایل قابلیت اعتماد، فرآیند پواسن ناهمگن و مانند آن نقش شایان توجهی ایفا می‌کنند. در اینجا دو مثال در این زمینه مطرح می‌کنیم.

مثال ۱.۲.۱. دستگاه سنگ خردکنی را در نظر بگیرید که در هنگام کار نیاز است در هر دفعه‌ای که اندازه‌ی سنگ دریافتی از اندازه‌ی سنگ‌های خردشده‌ی قبلی بزرگتر باشد، تنظیم مجدد گردد. در چنین دستگاهی اندازه‌هایی از سنگ‌ها برای کاربر اهمیت پیدا می‌کند که از مقادیر قبلی خود بزرگتر باشند. مشخص است که تعداد این مقادیر نشان می‌دهد که دستگاه چند بار تنظیم شده است.

مثال ۲.۲.۱. بسیاری از محصولات در اثر تنش از کار می‌افتند. مثلا محیط گرم باعث تنزل کارایی یک مولفه‌ی الکتریکی می‌شود یا یک باطری با گذشت زمان خالی می‌شود. مقدار تنش یک محصول تا نقطه‌ی کارافت برای دو محصول کاملا مشابه، متفاوت بوده و یک متغیر تصادفی می‌باشد. فرض کنید که نقطه‌ی کارافت یک محصول را بتوان در آزمایشگاه با افزایش تدریجی تنش (مثلاً نیرو، درجه‌ی حرارت و فشار) به دست آورد. اگر چنین آزمونی را برای صد محصول مشابه انجام دهیم، حاصل صد مقدار $X_۱, X_۲, ..., X_{۱۰۰}$، متناظر با نقاط کارافت محصولات خواهد بود.

فرض کنید هدف از انجام این آزمایش یافتن ضعیف‌ترین محصول در نمونه باشد. یک روش در این آزمایش این است که تک تک نمونه‌ها را تحت تنش قرار دهیم تا نقطه‌ی کارافت هر کدام پیدا شود. بی‌تردید با اعمال این روش تعدادی از محصولات و در مواردی خاص تمام محصولات را از دست خواهیم داد. روش دیگر، روش راهبرد دنباله‌ای است. به این صورت که ابتدا محصول اول را آزمایش می‌کنیم و مقدار $X_۱$ را یادداشت می‌کنیم. اگر نمونه‌ی دوم این مقدار را تحمل کرد و از کار نیفتاد، آزمایش را متوقف کرده سراغ نمونه ی سوم می‌رویم. با این روش $X_۲$ در صورتی دقیقاً مشخص می‌شود که $X_۲ < X_۱$ در غیر این صورت ما فقط این اطلاع را به دست خواهیم آورد که $X_۲ > X_۱$ و اطلاع دیگری در مورد $X_۲$ نداریم. اگر این روند را برای تمام محصولات ادامه دهیم در این صورت نتیجه‌ی آزمایش دنباله‌ای به صورت زیر خواهد بود و ضعیف‌ترین محصول مربوط به نقطه‌ی کارافت X_k می‌باشد.

$$X_۱ < X_۲ < ... < X_k.$$

در این مثال دنباله‌ای از مشاهدات برای ما اهمیت پیدا می‌کنند که هر کدام از مقدار قبلی خود کوچکتر است. این مثال حالتی از طرح آزمایش سانسور نوع دوم می‌باشد.

۵.۲.۱ توزیع رکوردها

فرض کنید $X_n, n \geq ۱$ دنباله‌ای از متغیرهای تصادفی پیوسته، مستقل و هم‌توزیع از جامعه‌ای با تابع توزیع احتمال $F(x)$ و تابع چگالی $f(x)$ باشد. همچنین فرض کنید $\lambda(x) = \frac{f(x)}{۱-F(x)}$ تابع نرخ خطر و $R(x) = -\log(۱ - F(x))$، در این صورت چگالی توام اولین n رکورد بالا به‌صورت زیر است:

$$f_{۱,\ldots,n}(x_۱, x_۲, \ldots, x_n) = \prod_{i=۱}^{n-۱} \lambda(x_i) f(x_n),\ -\infty < x_۱ < x_۲ < \ldots < x_n < \infty. \qquad (۲.۱)$$

و تابع چگالی حاشیه‌ای nـ امین رکورد بالا برابر است با:

$$f_n(x) = f(x)\frac{(R(x))^{n-1}}{(n-1)!} \tag{۳.۱}$$

همچنین چگالی توام X_{U_i} و X_{U_j} $(i<j)$ از رابطه‌ی زیر محاسبه می‌شود:

$$f_{i,j}(x_i,x_j) = \lambda(x_i)f(x_j)\frac{[R(x_i)]^{i-1}[R(x_j)-R(x_i)]^{j-i-1}}{(i-1)!(j-i-1)!},\ -\infty < x_i < x_j < \infty. \tag{۴.۱}$$

با استفاده از روابط قبلی چگالی شرطی $X_{U_j}|X_{U_i}=x_i$ برابر است با:

$$f_{j|i}(x_j|X_{U_i}=x_i) = \frac{[R(x_j)-R(x_i)]^{j-i-1}f(x_j)}{(j-i-1)!(1-F(x_i))}\,,\ x_i < x_j. \tag{۵.۱}$$

مشابه روابط بالا، فرض کنید $\tilde{\lambda}(x)=\frac{f(x)}{F(x)}$ تابع نرخ خطر معکوس و $H(x)=-\log(F(x))$. چگالی توام اولین n رکورد پایین برابر است با:

$$f_{1,\dots,n}(x_i,x_2,\dots,x_n) = \prod_{i=1}^{n-1}\tilde{\lambda}(x_i)f(x_n)\,,\ -\infty < x_n < x_{n-1} < \cdots < x_1 < \infty. \tag{۶.۱}$$

و تابع چگالی حاشیه‌ای nـ امین رکورد پایین برابر است با:

$$f_n(x) = f(x)\frac{(H(x))^{n-1}}{(n-1)!}. \tag{۷.۱}$$

۶.۲.۱ سانسور

در طبیعت گاه نمی‌توانیم همه داده‌های موجود را (مثلاً برای بررسی طول عمر موجودات و حتی لوازم و اشیاء از قبیل وسایل برقی) مورد بررسی قرار دهیم. گاه با این مسئله مواجه می‌شویم که به عنوان مثال، می‌خواهیم طول عمر لامپ‌ها را بررسی کنیم اما زمان لازم برای این کار را نداریم. تا ۴۰ روز منتظر می‌مانیم. تعداد لامپ‌هایی که می‌سوزند اندک هستند ولی ما به دلیل کم بودن وقت بقیه لامپ‌ها را مورد آزمایش قرار نمی‌دهیم و استنباط‌های خود را براساس همین لامپ‌های موجود انجام می‌دهیم. به این داده‌ها، داده‌های سانسورشده می‌گویند. چند نوع سانسور عبارتند از:

۷.۲.۱ سانسور از راست

تحت این نوع سانسور، مطالعه تا زمان t_c که از قبل توسط آزمایشگر مشخص شده است، ادامه می‌یابد. تحت طرح سانسور از راست، مشاهدات عبارتند از

$$Y_i = \begin{cases} X_i & X_i \le t_c \\ t_c & X_i > t_c \end{cases}$$

بنابراین نمونه شامل n مشاهده است که m تا از آنها کامل اندازه‌گیری شده و R تای بقیه $(R=n-m)$ سانسور شده‌اند.

۸.۲.۱ سانسور از چپ

تحت این نوع از سانسور مشاهداتی که از مقدار ثابت و از قبل تعیین شده ی t_c کمتر هستند را سانسور شده از چپ نامیده و مقدار ثبت شده برای آن‌ها t_c است و به ازای سایر مشاهدات مقدار دقیق X ثبت می‌شود. تحت طرح سانسور از چپ، مشاهدات عبارتند از

$$Y_i = \begin{cases} X_i & X_i \geq t_c \\ t_c & X_i < t_c \end{cases}$$

۹.۲.۱ سانسور تصادفی

اگر C_i به ازای $i = ۱, ۲, \ldots, n$ زمان سانسور متغیر تصادفی X_i باشد و C_iها مستقل و هم‌توزیع باشند، دراین‌صورت تحت این نوع از سانسور مشاهدات عبارتند از $(Y_۱, \delta_۱), (Y_۲, \delta_۲), \ldots, (Y_n, \delta_n)$ که $Y_i = \min(X_i, C_i)$ و

$$\delta_i = \begin{cases} ۱ & X_i \leq C_i \\ ۰ & X_i > C_i \end{cases}$$

در این حالت $Y_۱, Y_۲, \ldots, Y_n$ مستقل و هم‌توزیع هستند و δ_i مشخص می‌کند که سانسور رخ داده است $(\delta = ۰)$ یا زمان رخداد پیشامد مورد نظر به‌طور دقیق ثبت شده است $(\delta = ۱)$. اگر به ازای $i = ۱, ۲, \ldots, n$، $C_i = t_c$ و زمان سانسور کلیه واحدها یکسان باشد، طرح سانسور تصافی تبدیل به طرح سانسور از راست می‌شود. سانسور تصادفی در مطالعات پزشکی، همه‌گیرشناسی و مهندسی کاربرد دارد. به عنوان مثال، بیماران معمولاً در زمان‌های متفاوتی از دوره بیماری وارد مطالعه می‌شوند و در نتیجه هر یک از بیماران می‌بایست متناسب با زمان ورود خود، تحت درمان قرار گیرند.

۱۰.۲.۱ آماره‌های ترتیبی تعمیم‌یافته

آماره‌های ترتیبی تعمیم‌یافته[۱۵] (GOS) در ده سال اخیر مورد توجه قرار گرفته‌اند زیرا آنها در قابلیت اطمینان، مدل‌سازی آماری و استنباط انعطاف‌پذیری بیشتری را دارند. آماره‌های ترتیبی تعمیم‌یافته به عنوان یک توزیع یکپارچه که به‌صورت نظری تنظیم شده معرفی شده است و شامل مدل‌های مختلفی از متغیرهای تصادفی می‌شود. کمپس[۱۶] (۱۹۹۵) مفهوم GOS را بیان کرد و نشان داد که آماره‌های ترتیبی، مقادیر رکورد، و تعداد دیگری از متغیرهای تصادفی ترتیبی می توانند به عنوان حالت خاصی از آماره‌های ترتیبی تعمیم‌یافته در نظر گرفته شوند.

[۱۵] Generalized order statistics

[۱۶] Kamps

آماره ترتیبی تعمیم‌یافته بر اساس تابع توزیع دلخواه F به صورت میانگین تابع معکوس F بیان می‌شود.

تعریف ۶.۲.۱. متغیرهای تصادفی $X(1,n,\tilde{m},k),\ldots,X(n,n,\tilde{m},k)$ متغیرهای ترتیبی تعمیم‌یافته بر اساس تابع توزیع تجمعی $F(x)$ نامیده می‌شوند، اگر تابع احتمال پیوسته‌ی آن‌ها به‌صورت زیر داده شده باشد

(۸.۱)

$$f_{1,\ldots,n}(x_1,\ldots,x_n)=\begin{cases} k\left(\prod_{j=1}^{n-1}\gamma_j\right)\left[\prod_{j=1}^{n-1}\left(1-F(x_i)\right)^{m_i}f(x_i)\right].\left(1-F(x_n)\right)^{k-1}f(x_n), \\ for\, F^{-1}(0)<x_1<\ldots<x_n<F^{-1}(1) \\ 0, \qquad \text{در غیر این صورت} \end{cases}$$

با متغیرهای $n\in\mathcal{N}$، $n\geq 2$، $k>0$، $\tilde{m}=(m_1,m_2,\ldots,m_{n-1})$، $M_r=\sum_{j=r}^{n-1}m_j$، به‌طوری که $\gamma_r=k+n-r+M_r>0$ برای تمام مقادیر $r\in\{1,\ldots,n-1\}$، قرار دهید $c_{r-1}=\prod_{j=1}^{r}\gamma_j$، $r=1,2,\ldots,n-1$ و $\gamma_n=k$.

تابع توزیع تجمعی (cdf) متغیر ترتیبی تعمیم‌یافته $X(r,n,m,k)$ به‌صورت زیر است

$$F_r(X)=1-c_{r-1}\sum_{j=1}^{r}\frac{a_j(r)}{\gamma_j}\left(1-F(X)\right)^{\gamma_j}, \qquad (9.1)$$

به‌طوری‌که $a_j(r)=\prod_{i=1,i\neq j}^{r}\left(\frac{1}{\gamma_i-\gamma_j}\right)$ یا $1\leq j\leq r\leq n$.

حال اگر $m=0$ و $k=1$، تابع چگالی توام n آماره ترتیبی معمولی $X_{1,n},X_{2,n},\ldots,X_{n,n}$ به دست می‌آید. به‌علاوه اگر $m=-1$ و $k=1$، تابع چگالی اولین n رکورد بالا از متغیرهای تصادفی مستقل و هم‌توزیع به‌دست می‌آید.

۳.۱ مفاهیم کلی عدم دقت

فرض کنید که یک آزمایشگر احتمال مشاهده(رخداد) پیشامدهای مختلف یک آزمایش را بیان می‌کند. به دو طریق اظهاراتش می‌تواند فاقد دقت باشد: اطلاعات کافی نداشته و گزارشاتش مبهم باشد، یا مقداری از اطلاعاتش ناصحیح باشند. اغلب برآوردهای آماری و مسایل استنباطی در ارتباط با ساخت گزارشاتی هستند که ممکن است به یکی از این طریق‌ها یا هر دوی آن‌ها نادرست باشند. نظریه ارتباط شانون و ویور[17] (۱۹۴۹) یک نظریه کلی از عدم حتمیت را ارائه می‌دهد که ما را برای رویارویی با جنبه‌ی ابهام از عدم دقت قادر می‌سازد. سودمندی این نظریه در مسایل آماری به وسیله‌ی نویسندگان متعددی نشان داده شده است،

[17] Shannon and Weaver

مانند لیندلی و کولبک[۱۸] (۱۹۵۷) و لایبلر[۱۹] (۱۹۵۱). با این حال این نظریه تا کنون قادر به ارتباط با عدم دقت در مفهوم گسترده‌تر آن نشده و بنابراین کاربردش محدود شده است. برای غلبه بر این محدودیت‌ها در نظریه ارتباط، کریج (۱۹۶۱) یک اندازه را پیشنهاد داد که با عنوان ''اندازه عدم دقت''[۲۰] شناخته می‌شود. این اندازه می‌تواند به‌عنوان تعمیمی از آنتروپی شانون در نظر گرفته شود.

فرض کنید آزمایشگر ادعا می‌کند که احتمال رخداد iـامین پیشامد ممکن ϕ_i است و احتمال واقعی آن θ_i باشد و $\sum\theta_i = \sum\phi_i = ۱$. نشان داده می‌شود که عدم دقت گفته‌ی آزمایشگر می‌تواند با $I = I(\theta_۱, \theta_۲, \ldots, \phi_۱, \phi_۲, \ldots) = -\sum\theta_i \log\phi_i$ اندازه‌گیری شود و این تنها تعریفی است که مجموعه‌ای از فرضیات شهودی را پوشش می‌دهد.

طبیعتاً فرضیات عدم دقت با فرضیات معادله‌ی شانون $-\sum\theta_i \log\theta_i$ که به‌عنوان اندازه‌ی عدم حتمیت به‌کار رفته است تقریباً یکسان هستند. از آنجا که وقتی تمام اطلاعات ارائه شده درست باشند، عدم دقت به عدم حتمیت میل می‌کند.

۱.۳.۱ خواص عدم دقت

فرض کنید $I = I(\theta_۱, \theta_۲, \ldots; \phi_۱, \phi_۲, \ldots)$ مقدار عدم دقت باشد. کریج (۱۹۶۱) فرضیات زیر را مطرح ساخت:

۱. تابع I روی θها و ϕها پیوسته است.

۲. وقتی آزمایشگر N پیشامد را هم‌شانس و با احتمال $۱/N$ بیان می‌کند، عدم دقت تابعی یکنوا صعودی از N می‌باشد.

۳. اندازه عدم دقت به‌صورت یک ترکیب وزنی معنی‌دار از زیرفرض‌های عدم دقت اصلی است. به عنوان مثال داریم

$$I(\alpha, \beta, \gamma; \theta, \phi, \psi) \equiv I(\alpha, ۱-\alpha; \theta, ۱-\theta) + (۱-\alpha) I\left(\frac{\beta}{۱-\alpha}, \frac{\gamma}{۱-\alpha}; \frac{\phi}{۱-\theta}, \frac{\psi}{۱-\theta}\right).$$

با توجه به اینکه $\alpha + \beta + \gamma = ۱$، رابطه‌ی بالا را به صورت زیر اثبات می‌کنیم:

$$\begin{aligned}
& I(\alpha, ۱-\alpha; \theta, ۱-\theta) + (۱-\alpha) I\left(\frac{\beta}{۱-\alpha}, \frac{\gamma}{۱-\alpha}; \frac{\phi}{۱-\theta}, \frac{\psi}{۱-\theta}\right) \\
&= -\alpha\log\theta - (۱-\alpha)\log(۱-\theta) - (۱-\alpha)\left\{\frac{\beta}{۱-\alpha}\log(\frac{\phi}{۱-\theta}) + \frac{\gamma}{۱-\alpha}\log(\frac{\psi}{۱-\theta})\right\} \\
&= -\alpha\log\theta - (۱-\alpha)\log(۱-\theta) - \beta\log\phi + \beta\log(۱-\theta) - \gamma\log\psi + \gamma\log(۱-\theta) \\
&= -\alpha\log\theta - \beta\log\phi - \gamma\log\psi \\
&= I(\alpha, \beta, \gamma; \theta, \phi, \psi).
\end{aligned}$$

---

[۱۸] Lindley and Kullback

[۱۹] Leibler

[۲۰] Inaccuracy

۴. اگر دو گزینه که در مورد یک ادعای مشابه ساخته شده‌اند ترکیب شوند، مقدار عدم دقت تغییری نمی‌کند. به‌عنوان مثال،

$$\begin{aligned} I(\alpha,\beta,\gamma;\theta,\phi,\phi) &= -\alpha\log\theta-\beta\log\phi-\gamma\log\phi \\ &= -\alpha\log\theta-(\beta+\gamma)\log\phi \\ &= I(\alpha,\beta+\gamma;\theta,\phi). \end{aligned}$$

۵. فرض کنید $\phi_i = n_i/N$ که n_iها مقادیر صحیح هستند و $N=\sum_i n_i$. آن‌گاه داریم:

$$\begin{aligned} I\left(\theta_۱,\theta_۲,\ldots;\frac{n_۱}{N},\frac{n_۲}{N},\ldots\right) &= -\sum_i \theta_i \log\frac{n_i}{N} \\ &= -\sum_i \theta_i\log n_i-\sum_i\theta_i\log N, \end{aligned} \qquad (۱۰.۱)$$

و

$$\begin{aligned} \sum_i \theta_i I\left(۱;\frac{۱}{n_i}\right) &= -\sum_i\theta_i\log\frac{۱}{n_i} \\ &= \sum_i\theta_i\log n_i. \end{aligned} \qquad (۱۱.۱)$$

اگر رابطه‌ی (۱۰.۱) و (۱۱.۱) را جمع کنیم، خواهیم داشت

$$(۱۰.۱)+(۱۱.۱) = -\sum_i\theta_i\log N = -\log N = I\left(۱;\frac{۱}{N}\right).$$

بنابراین رابطه‌ی زیر به‌دست می‌آید:

$$I\left(\theta_۱,\theta_۲,\ldots;\frac{n_۱}{N},\frac{n_۲}{N},\ldots\right)+\sum_i\theta_i I(۱;\frac{۱}{n_i}) \equiv I\left(۱;\frac{۱}{N}\right)$$

از این فرضیات، ۱ و ۲ و ۳ مشابه فرضیات شانون است. تنها فرضیه ی ۴ جدید است و به طور شهودی قابل قبول به نظر می‌رسد.

در حالت پیوسته، اگر $f(x)$ تابع چگالی واقعی مشاهدات باشد و $g(x)$ تابع چگالی باشد که آزمایشگر در نظر می‌گیرد، اندازه عدم دقت به‌صورت زیر تعریف می‌شود

$$I(f,g) = -\int f(x)\log g(x)dx.$$

در ادامه‌ی فصل ابتدا به تعاریف اولیه مقدار کولبک‌ـ‌لایبلر و آنتروپی شانون پرداخته و سپس میزان عدم دقت بین توزیع‌های طول عمر باقیمانده از سیستم را بیان می‌کنیم.

۲.۳.۱ آنتروپی شانون

امروزه آنتروپی از جایگاه ویژه‌ای در علوم مختلف از قبیل فیزیک، آمار و احتمال، نظریه‌ی ارتباطات، نظریه‌ی مجموعه‌های فازی و اقتصاد برخوردار است. آنتروپی شانون در زمینه نظریه‌ی اطلاع نقش بسیار مهمی دارد. آنتروپی (اطلاع شانون) یک پیشامد، میزان عدم حتمیت آن پیشامد را اندازه می‌گیرد. هر چه پیشامدی نادرتر باشد، رخ دادن آن متضمن اطلاع بیشتری است. مفهوم آنتروپی اولین بار در علم فیزیک معرفی شد و برای کمی‌سازی مفاهیم عدم حتمیت و بی‌نظمی بکار برده شده است. بعدها هارتلی[۲۱] (۱۹۲۸) و نیکویست[۲۲] (۱۹۳۲)، اولین افرادی بودند که تعاریفی از اطلاع را در مهندسی ارتباطات ارائه دادند. شانون (۱۹۱۶-۲۰۰۱) برای اولین بار با تعریف ریاضی آنتروپی ایده‌های قدیمی نیکویست و هارتلی را مطرح ساخت و با در نظر گرفتن ارتباط به عنوان یک مسئله ریاضی، به مهندسین راه تعیین ظرفیت کانال‌های ارتباطی برای انتقال اطلاعات که به صورت ارقام در پایه ۲ کدگذاری شده بودند را نشان داد. به‌طور کلی این باور وجود دارد که نظریه اطلاع با انتشار مقاله شانون (۱۹۴۸) تحت عنوان نظریه‌ی ریاضی ارتباطات نظم جدیدی یافت.

شانون برای استخراج مفهوم آنتروپی یک توزیع احتمال مانند $P = (p_1, \ldots, p_n)$، در جستجوی تابعی بود که در اصول زیر صدق کند:

۱. $f(1) = 0$

۲. $f(p)$ یک تابع پیوسته از p باشد.

۳. اگر $p < q$ آن‌گاه $f(p) > f(q)$. یعنی $f(p)$ یک تابع اکیداً نزولی از p باشد.

۴. $f(p * q) = f(p) + f(q),\ \forall p, q \in [0, 1]$.

وی نشان داد که تنها تابعی که در این اصول صدق می کند، $f(p) = -\log_b(p),\ b \in \mathcal{R}^+$ است. مبنای b با توجه به واحدی که قرار است استفاده شود، تعیین می‌گردد. بنابراین آنتروپی شانون متغیر تصادفی X با تابع توزیع گسسته‌ی F، به‌صورت زیر تعریف می‌شود

$$H(X) = H(F) = -\sum_i p_i \log p_i,$$

که در آن p_iها، مقادیر تابع جرم احتمال مربوط به توزیع F هستند. فرض کنید X یک متغیر تصادفی پیوسته با تابع چگالی احتمال $f(x)$ باشد، آن‌گاه آنتروپی شانون آن به‌صورت

$$H(X) = H(f) = -\int f(x) \log f(x) dx = -E_f[\log f(X)] = E_f\left[\log \frac{1}{f(x)}\right]$$

تعریف شده که میزان بی‌نظمی و اغتشاش تابع چگالی $f(x)$ را اندازه می‌گیرد، به این معنی که هر چقدر تابع چگالی $f(x)$ به تابع چگالی یکنواخت نزدیکتر باشد یا به عبارتی هر چه چگالی

[۲۱] Hartley

[۲۲] Nyquist

منظم‌تر باشد آنتروپی شانون نیز کمتر می‌شود و تصمیم‌گیری بر اساس آن مشکل‌تر خواهد شد. بنابراین تصمیم‌گیری بر اساس تابع چگالی $f(x)$ زمانی قابل درک است که اندازه آنتروپی شانون آن زیاد باشد. بنابراین مهم است بدانیم کدام چگالی برای متغیرهای تصادفی پیوسته به ماکزیمم اطلاع منجر می‌شود. اصل ماکزیمم آنتروپی نیز به برآورد توزیع متغیرهای تصادفی می‌پردازد و توزیعی به عنوان یک مدل مناسب انتخاب می‌شود که عدم حتمیت بیشتری داشته باشد.

ویژگی‌هایی که باعث می شود آنتروپی را یک معیار عدم حتمیت بدانیم عبارتند از

الف) آنتروپی برای متغیرهای تصادفی گسسته همواره نامنفی است و صفر می شود، اگر و تنها اگر حتمیت کامل وجود داشته باشد، یعنی همه‌ی احتمال دستگاه به یک پیشامد معین اختصاص یافته باشد. به عبارت دیگر $H(۱, ۰, \ldots, ۰) = ۰$.

ب) در آنتروپی گسسته، همواره $H(F) \leq \log n$ و تساوی وقتی رخ می‌دهد که
$$p_۱ = p_۲ = \ldots = p_n = \frac{۱}{n}.$$
در واقع وقتی همه‌ی پیشامدها هم‌احتمال هستند، بیشترین عدم حتمیت از این نظر که کدام پیشامد رخ خواهد داد، پیش می‌آید.

پ) در حالت پیوسته، لازم است که محدودیت‌هایی برای متغیر تصادفی در نظر گرفت. به عنوان مثال، برای کلیه متغیرهای پیوسته‌ای که تکیه‌گاه آن‌ها خط حقیقی و واریانس آن‌ها مقدار $\sigma^۲$ است، $H(X)$ ماکزیمم است، اگر و تنها اگر متغیر تصادفی X دارای توزیع $N(۰, \sigma^۲)$ باشد.

برای توضیحات بیشتر به کاور و توماس[۲۳] (۱۹۹۱) مراجعه شود.

۳.۳.۱ تابع تشخیص کولبک_ لایبلر

تابع تشخیص کولبک_لایبلر که اولین بار توسط کولبک و لایبلر (۱۹۵۱) معرفی شد، یکی از اندازه‌های اطلاع برای مقایسه دو توزیع و یکی از معیارهای اندازه‌گیری فاصله بین دو توزیع احتمال است. برای دو متغیر تصادفی پیوسته X و Y به ترتیب با توابع چگالی احتمال f و g، اطلاع کولبک_لایبلر به عنوان میزان تفاوت توزیع‌های دو متغیر به‌صورت
$$K(f|g) = \int f(x) \log\left(\frac{f(x)}{g(x)}\right) dx = I(f, g) - H(f)$$
تعریف می‌شود، که در آن g توزیع مرجع است. برای دو متغیر تصادفی گسسته X و Y با توابع چگالی احتمال $p(x)$ و $q(x)$ و با تکیه‌گاه مشترک χ، این فاصله عبارت است از
$$K(p|q) = \sum_{x \in \chi} p(x) \log \frac{p(x)}{q(x)}.$$

[۲۳] Cover and Thomas

از اندازه اطلاع کولبک ـ لایبلر برای تقریب توابع توزیع نیز استفاده می‌شود، به این معنا که اگر یکی از دو توزیع دارای تابع چگالی معلوم باشد، $K(f|g)$ قابلیت پیش‌بینی توزیع نامعلوم با استفاده از توزیع معلوم را نشان می‌دهد. برای توضیحات بیشتر به کاور و توماس (۱۹۹۱) مراجعه شود.

چند ویژگی قابل توجه از این فاصله عبارتند از:

۱. **نامنفی بودن؛** $K(f|g) \geq ۰$ و برابری رخ می‌دهد، اگر و تنها اگر به ازای تمام مقادیر x، $f(x) = g(x)$. این ویژگی مستقیماً از نامساوی جنسن یا نامساوی لگ مجموع اثبات می‌شود.

۲. **تحدب؛** فاصله‌ی کولبک ـ لایبلر نسبت به زوج (p, q) محدب است. یعنی اگر $(p_۱, q_۱)$ و $(p_۲, q_۲)$ دو زوج از توابع چگالی احتمال باشند، آن‌گاه

$$K(\lambda p_۱ + (۱-\lambda)p_۲|\lambda q_۱ + (۱-\lambda)q_۲) \leq \lambda K(p_۱|q_۱) + (۱-\lambda)K(p_۲|q_۲), \quad \forall \lambda \in [۰, ۱].$$

۳. **جمع پذیری؛** برای مشاهدات مستقل x و y، فاصله‌ی کولبک ـ لایبلر بین توابع چگالی توام $f(x, y)$ و $g(x, y)$ برابر است با مجموع فاصله‌ی کولبک ـ لایبلر چگالی حاشیه‌ای آن‌ها. به‌عبارت دیگر داریم

$$K(f(x, y)|g(x, y)) = K(f(x)|g(x)) + K(f(y)|g(y)).$$

۴. **نامتقارن بودن؛** می‌توان به سادگی نشان داد که فاصله ی کولبک ـ لایبلر متقارن نیست، یعنی

$$K(f|g) \neq K(g|f),$$

اما می‌توان نوع متقارن آن را به شکل زیر در نظر گرفت:

$$K(f|g) + K(g|f).$$

تعریف ۱.۳.۱. فرض کنید X و Y متغیرهای تصادفی مثبت با توابع توزیع $F(X)$ و $G(Y)$ و میانگینهای متناهی $E(X)$ و $E(Y)$ باشند. آن‌گاه اطلاع کولبک ـ لایبلر تجمعی X از Y به‌صورت زیر تعریف می‌شود:

$$\begin{aligned} K(F|G) &= \int_۰^\infty F(x) \log\left(\frac{F(x)}{G(x)}\right) dx + E(X) - E(Y) \\ &= \int_۰^\infty F(x) \log F(x) dx - \int_۰^\infty F(x) \log G(x) dx + E(X) - E(Y) \\ &= -\mathcal{CE}(X) + I(F, G) + E(X) - E(Y). \end{aligned} \qquad (۱۲.۱)$$

۴.۳.۱ مقدار عدم دقت بین دو توزیع طول عمر باقیمانده و گذشته

فرض کنید X و Y دو متغیر تصادفی غیر منفی با توابع چگالی f و g باشند که زمان شکست دو سیستم را نشان می‌دهند. فرض کنید $F(x) = P(X \leq x)$ و $G(y) = P(Y \leq y)$ توابع توزیع‌های شکست، $\tilde{\lambda}_F(x) = \frac{f(x)}{F(x)}$ و $\tilde{\lambda}_G(x) = \frac{g(x)}{G(x)}$ نرخ‌های بقای معکوس، و $\overline{F}(x) = 1 - F(x)$ و $\overline{G}(x) = 1 - G(x)$ به ترتیب توابع بقای X و Y باشند.

اندازه‌ی شانون (۱۹۴۸) عدم حتمیت مربوط به متغیر تصادفی X و اندازه‌ی کولبک (۱۹۵۹) تمایز X نسبت به Y به ترتیب به‌صورت زیر می‌باشند:

$$H(X) = H(f) = -\int_0^\infty f(x) \log f(x) dx, \tag{۱۳.۱}$$

$$K(f|g) = \int_0^\infty f(x) \log \frac{f(x)}{g(x)} dx. \tag{۱۴.۱}$$

با جمع کردن (۱۳.۱) و (۱۴.۱) داریم

$$I(f,g) = H(f) + K(f|g) = -\int_0^\infty f(x) \log g(x) dx, \tag{۱۵.۱}$$

که اندازه‌ی کریج (۱۹۶۱) از عدم دقت مربوط به متغیرهای X و Y می باشد. اگر فرض کنیم F تابع توزیع واقعی باشد، آن‌گاه G به‌عنوان تابع توزیع مرجع تفسیر می‌شود.
در صورتی که سیستم تا زمان t فعال باشد، این اندازه‌ها برای باقیمانده ی طول عمر مناسب نیستند. ابراهیمی (۱۹۹۶) فرض کرد که X_t نشان دهنده‌ی باقیمانده‌ی طول عمر سیستم باشد به شرط آنکه تا زمان t فعال بوده، آن‌گاه

$$X_t \overset{d}{=} (X - t)|X > t, \ t \geq 0,$$

که در آن $\overset{d}{=}$ نماد معادل بودن در توزیع است. اندازه‌های پویا متناظر با عدم حتمیت، تمایز X نسبت به Y، و عدم دقت به ترتیب به‌صورت زیر داده شده‌اند:

$$H(f;t) = -\int_t^\infty \frac{f(x)}{\overline{F}(t)} \log \left(\frac{f(x)}{\overline{F}(t)} \right) dx, \tag{۱۶.۱}$$

$$K(f|g;t) = \int_t^\infty \frac{f(x)}{\overline{F}(t)} \log \left(\frac{f(x)/\overline{F}(t)}{g(x)/\overline{G}(t)} \right) dx \tag{۱۷.۱}$$

و

$$I(f,g;t) = -\int_t^\infty \frac{f(x)}{\overline{F}(t)} \log \left(\frac{g(x)}{\overline{G}(t)} \right) dx. \tag{۱۸.۱}$$

وقتی که $t = 0$، آن‌گاه (۱۶.۱)، (۱۷.۱) و (۱۸.۱) به ترتیب به (۱۳.۱)، (۱۴.۱) و (۱۵.۱) ساده می‌شوند.

در بسیاری از موارد، عدم حتمیت لزوماً مربوط به آینده نیست بلکه می‌تواند به گذشته معطوف باشد. به عنون مثال اگر یک سیستم که تنها در برخی زمان‌های از قبل مشخص شده بازرسی می‌شود، در زمان t از کار افتاده مشاهده شود، آن‌گاه عدم حتمیت طول عمر سیستم بر گذشته تکیه دارد یعنی در لحظه‌ای بین $(\circ, t)$ سیستم از کار افتاده است.

متغیر تصادفی $X_t \stackrel{d}{=} (t-X)|X \leq t$ به عنوان زمان غیرفعال شناخته می‌شود. به این خاطر که یک بار در زمان X سیستم از کار می‌افتد و در زمان t نیز در حالت از کار افتاده مشاهده می‌شود، زمان تصادفی از کار افتادن سیستم X_t است. بر این اساس اندازه‌های عدم حتمیت، تمایز X نسبت به Y و عدم دقت بر اساس آنتروپی گذشته روی $(\circ, t)$ به ترتیب به‌صورت زیر است

$$\tilde{H}(f;t) = -\int_\circ^t \frac{f(x)}{F(t)} \log\left(\frac{f(x)}{F(t)}\right) dx \qquad (۱۹.۱)$$

$$\tilde{K}(f|g;t) = \int_\circ^t \frac{f(x)}{F(t)} \log\left(\frac{f(x)/F(t)}{g(x)/G(t)}\right) dx \qquad (۲۰.۱)$$

و

$$\tilde{I}(f,g;t) = -\int_\circ^t \frac{f(x)}{F(t)} \log\left(\frac{g(x)}{G(t)}\right) dx. \qquad (۲۱.۱)$$

۴.۱ مروری مختصر بر کارهای انجام‌شده توسط دیگران

ابراهیمی و همکاران (۲۰۰۴) برخی ویژگی‌های نظری اندازه‌های اطلاع را برای داده‌های ترتیبی با استفاده از تبدیل انتگرالی احتمال بررسی کردند و اندازه‌های آنتروپی شانون و اطلاعات نسبی کولبک را تعریف کردند، که به ترتیب به‌صورت زیر تعریف می‌شوند:

$$H_n(X_{i:n}) = H_n(f_{i:n}) = -\int_\circ^\infty f_{i:n}(x)\log f_{i:n}(x)dx = H_n(W_{i:n}) - E_{g_i}\left[\log f(F^{-۱}(W_i))\right],$$

و

$$K_n(f_{i:n}|f) = \int_\circ^\infty f_{i:n}(x)\log\left(\frac{f_{i:n}(x)}{f(x)}\right)dx = -H_n(W_{i:n}),$$

که در آن $W_{i:n}$، iـامین آماره ترتیبی از توزیع یکنواخت $U(\circ,۱)$ با تابع چگالی زیر است

$$g_i(w) = \frac{۱}{B(i,n-i+۱)} w^{i-۱}(۱-w)^i, \qquad \circ \leq w \leq ۱.$$

در اینجا $B(a,b)$ تابع بتا با پارامترهای a و b است.

برخی نویسندگان بر روی مقدار عدم دقت در آماره‌های ترتیبی مطالعه کرده‌اند. تاپلیال و تانجا (۲۰۱۳) عدم دقت بین تابع چگالی iـامین آماره‌ی ترتیبی و تابع چگالی f را به صورت زیر تعریف کردند

$$I_n(f_{i:n}, f) = -\int_\circ^\infty f_{i:n}(x)\log(f(x))\,dx. \qquad (۲۲.۱)$$

مشابه (۱۸.۱) معیار عدم دقت پویا بین $f_{i:n}$ و f به صورت زیر تعریف می‌شود

$$I_n(f_{i:n}, f; t) = -\int_t^\infty \frac{f_{i:n}(x)}{\bar{F}_{i:n}(t)} \log\left(\frac{f(x)}{\bar{F}(t)}\right) dx. \qquad (۲۳.۱)$$

این معیار به‌عنوان مقدار عدم دقت باقیمانده‌ی پویا مربوط به دو توزیع طول عمر باقیمانده‌ی $F_{i:n}$ و F بیان می‌شود. توجه داشته باشید که $\bar{F}_{i:n}(t) = ۱ - F_{i:n}(t)$ تابع بقا مربوط به $X_{i:n}$ می‌باشد و به‌صورت زیر است

$$\bar{F}_{i:n}(t) = \frac{\bar{B}_{F(t)}(i, n-i+۱)}{B(i, n-i+۱)},$$

که در آن

$$\bar{B}_x(a, b) = \int_x^۱ u^{a-۱}(۱-u)^{b-۱} du, \quad ۰ < x < ۱,\ b > ۰$$

تابع بتای ناقص می‌باشد. دقت کنید هرگاه $t = ۰$، (۲۳.۱) به اندازه عدم دقت (۲۲.۱) تبدیل می‌شود.

براتپور و همکاران (۲۰۰۷و۲۰۰۸) نشان دادند که آنتروپی شانون و آنتروپی رنی iـامین آماره‌ی ترتیبی تابع توزیع را بطور یکتا مشخص می‌کند. براتپور (۲۰۱۰) همچنین اثبات کرد آنتروپی باقیمانده‌ی تجمعی نخستین آماره ترتیبی، تابع توزیع مرجع را بطور یکتا مشخص می‌کند. گوپتا[۲۴] و همکاران (۲۰۱۴) ثابت کردند که آنتروپی پویای iـامین آماره ترتیبی تابع توزیع را به‌طور یکتا مشخص می‌نماید. تاپلیال و همکاران (۲۰۱۵) نشان دادند که اندازه‌ی عدم دقت باقیمانده‌ی پویا بین iـامین آماره‌ی ترتیبی و متغیر تصادفی مرجع، تابع توزیع را به‌طور یکتا تعیین می‌کند.

[۲۴] Gupta

فصل ۲

عدم دقت تجمعی در مقادیر رکورد بالا و پایین

۱.۲ مقدمه

فرض کنید X و Y دو متغیر تصادفی نامنفی باشند که به ترتیب دارای تابع توزیع $F(x)$ و $G(x)$ و تابع بقای $\bar{F}(x)$ و $\bar{G}(x)$ هستند. در مقایسه با اندازه عدم دقت کریج که در (۱۴.۱) تعریف شد، تاپلیال و تانجا (۲۰۱۵a) اندازه عدم دقت گذشته‌ی تجمعی[۱] (CPI) را به‌صورت زیر تعریف کردند

$$I(F,G) = -\int_{0}^{+\infty} F(x)\log G(x)dx. \tag{۱.۲}$$

کومار[۲] و تانجا (۲۰۱۵) عدم دقت باقیمانده‌ی تجمعی[۳] (CRI) را بر اساس $\bar{F}(x)$ و $\bar{G}(x)$ به‌صورت زیر به‌دست آوردند

$$I(\bar{F},\bar{G}) = -\int_{0}^{+\infty} \bar{F}(x)\log \bar{G}(x)dx. \tag{۲.۲}$$

[۱] Cumulative past inaccuracy
[۲] Kumar
[۳] Cumulative residual inaccuracy

تابع میانگین زمان غیرفعال[۴] (MIT) در موضوعات متعددی مانند قابلیت اطمینان، تحلیل بقا، مطالعات آماری و غیره کاربرد دارد. برای متغیر تصادفی X MIT به‌صورت زیر تعریف می‌شود

$$\mu_۱(t) = E(t - X \mid X < t).$$

گاش و کندو[۵] (۲۰۱۷) یک ارتباط بین $\mu_۱(t)$ و $I(F,G)$ را به‌صورت زیر به‌دست آوردند

$$I(F,G) = E\left(\frac{\mu_۱(Y)F(Y)}{G(Y)}\right).$$

حال فرض کنید $\{X_m, m \geq ۱\}$ یک دنباله از متغیرهای تصادفی مستقل و همتوزیع با تابع توزیع F و تابع چگالی f باشد. فرض کنید R_n نشان‌دهنده‌ی nـامین رکورد بالا و L_n، nـامین رکورد پایین از $\{X_m, m \geq ۱\}$ باشند. آن‌گاه تابع چگالی و تابع بقای R_n و تابع چگالی و تابع توزیع L_n به ترتیب، به‌صورت زیر است

$$f_{R_n}(x) = \frac{[-\log \bar{F}(x)]^{n-۱}}{(n-۱)!} f(x), \quad x > ۰, \quad n \geq ۱, \tag{۳.۲}$$

$$\bar{F}_{R_n}(x) = \sum_{j=۰}^{n-۱} \frac{[-\log \bar{F}(x)]^j}{j!} \bar{F}(x) = \frac{\Gamma(n; -\log \bar{F}(x))}{\Gamma(n)}, \tag{۴.۲}$$

$$f_{L_n}(x) = \frac{[-\log F(x)]^{n-۱}}{(n-۱)!} f(x), \tag{۵.۲}$$

$$F_{L_n}(x) = \sum_{j=۰}^{n-۱} \frac{[-\log F(x)]^j}{j!} F(x), \tag{۶.۲}$$

به‌طوری‌که $\Gamma(a;x)$ تابع گامای ناقص است که به‌صورت زیر تعریف می‌شود

$$\Gamma(a;x) = \int_x^{+\infty} u^{a-۱} e^{-u} du,$$

و $\Gamma(n) = \Gamma(n;۰)$.

[۴] Mean inactivity time

[۵] Ghosh and Kundu

۲.۲ اندازه عدم دقت در رکوردهای بالا

اندازه عدم دقت مربوط به توزیع nــامین مقدار رکورد R_n و تابع توزیع اولیه $f(x)$ به‌صورت زیر به‌دست می‌آید

$$\begin{aligned} I(f_{R_n}, f) &= -\int_{0}^{+\infty} f_{R_n}(x)\log f(x)dx \\ &= -\int_{0}^{+\infty} \frac{[-\log \bar{F}(x)]^{n-1}}{(n-1)!} f(x)\log f(x)dx \\ &= -\int_{0}^{+\infty} \frac{(w_n)^{n-1}}{(n-1)!} e^{-w_n}\log(f(F^{-1}(1-e^{-w_n})))dw_n \\ &= -E_{W_n}[\log(f(F^{-1}(1-e^{-W_n})))], \end{aligned}$$

که W_n دارای توزیع گاما با پارامترهای شکل n و مقیاس ۱ است (به براتپور و همکاران ۲۰۰۷ مراجعه شود). در مثال‌های زیر $I(f_{R_n}, f)$ را برای برخی توزیع‌های طول عمر به‌دست می‌آوریم.

مثال ۱.۲.۲. فرض کنید X یک متغیر تصادفی از توزیع نمایی با میانگین $\frac{1}{\theta}$ باشد، آن‌گاه $I(f_{R_n}, f) = n - \log\theta$. دقت کنید که برای یک مقدار ثابت n، $I(f_{R_n}, f)$ تابعی نزولی بر حسب θ است. به‌طور مشابه، برای مقدار ثابتی از θ عدم دقت بر اساس n صعودی است.

مثال ۲.۲.۲. فرض کنید X یک متغیر تصادفی از توزیع پارتو با تابع چگالی $f(x) = \theta x^{-(\theta+1)}$ و $\theta > 0$ باشد، آن‌گاه

$$I(f_{R_n}, f) = -\log\theta + \left(\frac{\theta+1}{\theta}\right)\Gamma(n+1).$$

دقت کنید که برای یک مقدار ثابت n، $I(f_{R_n}, f)$ تابعی نزولی بر حسب θ است. به‌طور مشابه، برای مقدار ثابتی از θ عدم دقت بر اساس n صعودی است.

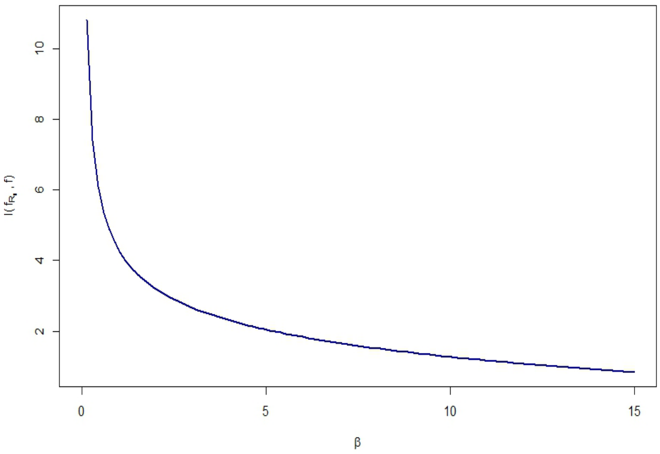
مثال ۳.۲.۲. فرض کنید X یک متغیر تصادفی از توزیع وایبل با تابع چگالی $f(x) = [\lambda\beta x^{(\beta-1)}]e^{-\lambda x^{\beta}} I_{(0,\infty)}(x)$ باشد، آن‌گاه

$$I(f_{R_n}, f) = -\log\beta - \frac{\log\lambda}{\beta} - \left(\frac{\beta-1}{\beta}\right)\psi(n) + n,$$

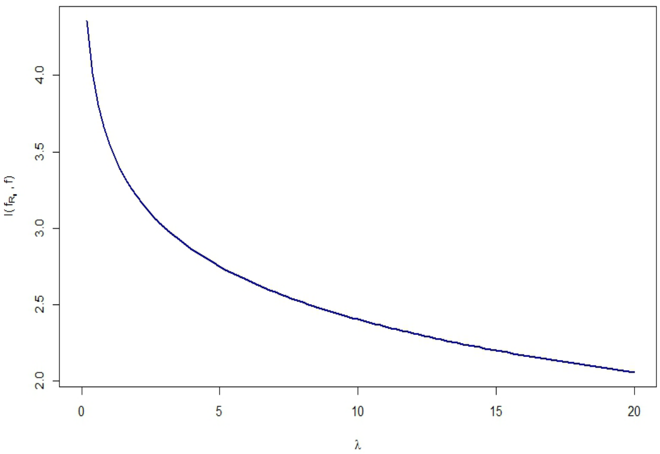
که $\psi(n)$ توزیع دی‌گاما است. برای یک مقدار ثابت $n \geq 2$، اگر $\lambda > 1$ باشد، آن‌گاه $I(f_{R_n}, f)$ تابعی نزولی از β می‌باشد. همچنین برای مقدار ثابتی از $\beta > 0$، $I(f_{R_n}, f)$ تابعی نزولی از λ است. شکل ۱.۲ تابع $I(f_{R_n}, f)$ را برای $n = 5$ و $\lambda = 2$ نشان می‌دهد. عدم دقت برای مقادیر مختلف β نزولی است. به‌طور مشابه، شکل ۲.۲ تابع $I(f_{R_n}, f)$ را برای $n = 5$ و $\beta = 2$ نشان می‌دهد. عدم دقت بر حسب $\lambda > 0$ نیر نزولی می‌باشد.

مثال ۴.۲.۲. فرض کنید $X \sim N(0,1)$، آن‌گاه

$$I(f_{R_n}, f) = \frac{1}{2}\log(2\pi) + \int_{0}^{+\infty} \frac{x^2 e^{\frac{-x^2}{2}}[-\log(1-\Phi(x))]^{n-1}}{2\Gamma(n)\sqrt{2\pi}}dx.$$



شکل ۱.۲: تابع $I(f_{R_n},f)$ بر حسب β برای $n=5$ و $\lambda=2$



شکل ۲.۲: تابع $I(f_{R_n},f)$ بر حسب λ برای $n=5$ و $\beta=2$

توجه کنید که $I(f_{R_n},f)$ تابعی نزولی بر حسب n است. شکل ۳.۲ نزولی بودن عدم دقت را برای مقادیر مختلف $n>3$ نشان می‌دهد.

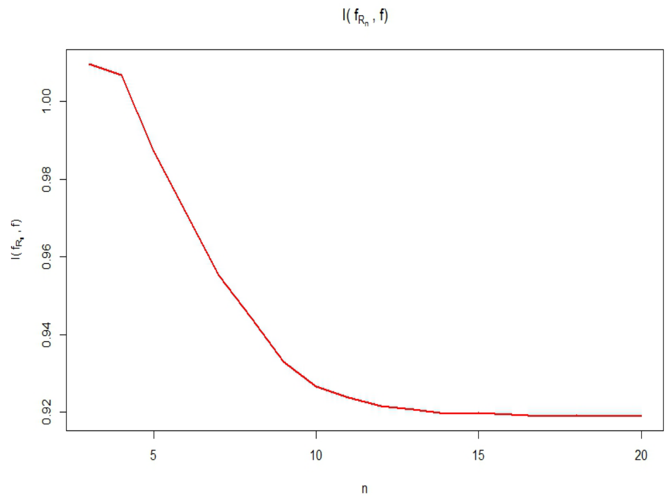
مثال ۵.۲.۲. فرض کنید X یک متغیر تصادفی با تابع چگالی زیر باشد

$$f(x)=\frac{\alpha\lambda e^{\lambda x}}{[e^{\lambda x}-(1-\alpha)]^2}I_{(0,\infty)}(x),\ \alpha>0,\ \lambda>0,$$

آن‌گاه

$$I(f_{R_n},f)=n-\log\lambda+\sum_{i=1}^{\infty}\frac{(\frac{\alpha-1}{\alpha})^i}{i(i+1)^n}.$$

حال با استفاده از برخی ویژگی‌های ترتیب تصادفی می‌توانیم خصوصیات مهمی را برای اندازه عدم دقت اثبات کنیم. ابتدا تعاریف زیر را بیان می‌کنیم:



شکل ۳.۲: تابع $I(f_{R_n}, f)$ برای $n > ۳$ در توزیع نرمال استاندارد

۱. می‌گوییم متغیر تصادفی X در ترتیب تصادفی از متغیر تصادفی Y کوچکتر است $(X \leq^{st} Y)$ اگر $P(X \geq x) \leq P(Y \geq x)$ به ازای تمام مقادیر x. می‌دانیم که به ازای تمامی توابع صعودی ϕ داریم

$$X \leq^{st} Y \Leftrightarrow E(\phi(X)) \leq E(\phi(Y)).$$

۲. متغیر تصادفی X با تابع چگالی احتمال $f(x)$ را در ترتیب نسبت درستنمایی کوچکتر از Y با تابع چگالی احتمال $g(y)$ می‌گوییم $(X \leq^{lr} Y)$ اگر $\frac{g(x)}{f(x)}$ تابعی صعودی از x باشد.

۳. متغیر تصادفی X با تابع بقای $\bar{F(x)}$ در ترتیب نرخ خطر از متغیر تصادفی Y با تابع بقای $\bar{G(y)}$ کوچکتر است $(X \leq^{hr} Y)$ اگر $\frac{\bar{G}(x)}{\bar{F}(x)}$ تابعی صعودی از x باشد.

۴. متغیر تصادفی X را در ترتیب محدب صعودی کوچکتر از Y می‌گوییم $(X \leq^{icx} Y)$ اگر برای تمامی توابع محدب صعودی ϕ که امیدشان وجود دارد، $E(\phi(X)) \leq E(\phi(Y))$.

۵. متغیر تصادفی X را در ترتیب محدب نزولی کوچکتر از Y می‌گوییم $(X \leq^{dcx} Y)$ اگر برای تمامی توابع محدب نزولی ϕ که امیدشان وجود دارد، $E(\phi(X)) \leq E(\phi(Y))$.

۶. می‌گوییم متغیر تصادفی نامنفی X با تابع توزیع $F(x)$ دارای نرخ خطر صعودی (نزولی) است، که با IFR (DFR) نشان داده می‌شود، اگر $\lambda_F(x) = \frac{f(x)}{\bar{F}(x)}$ تابعی صعودی (نزولی) از x باشد.

۷. می‌گوییم متغیر تصادفی نامنفی X دارای نرخ خطر معکوس نزولی است، که با DRHR نشان داده می‌شود، اگر $\tilde{\lambda}_F(x) = \frac{f(x)}{F(x)}$ تابعی نزولی از x باشد.

۸. می‌گوییم متغیر تصادفی نامنفی X با تابع توزیع F دارای میانگین نرخ شکست صعودی (نزولی) است، که با IFRA (DFRA) نشان داده می‌شود، اگر $\frac{\lambda_F(x)}{x}$ به‌ازای $x > ۰$ تابعی

صعودی (نزولی) باشد. توجه کنید که رده توزیع‌های IFR و DFR به‌ترتیب شامل رده توزیع‌های IFRA و DFRA نیز هستند.

قضیه ۱.۲.۲. اگر $f(x)$ در x صعودی باشد، آن‌گاه $I(f_{R_n}, f)$ در n نزولی است.

برهان. از آنجا که نسبت $\frac{f_{W_{n+1}}(x)}{f_{W_n}(x)} = \frac{x}{n}$ تابعی صعودی از x است، با توجه به شیکد و شانتکی کومار (۲۰۰۷) می‌دانیم $W_n \leqslant^{lr} W_{n+1}$، که نتیجه می‌دهد $W_n \leq^{st} W_{n+1}$. همچنین با توجه به فرض می‌دانیم که $f(F^{-1}(1-e^{-x}))$ در x صعودی است، بنابراین

$$-E[\log(f(F^{-1}(1-e^{-W_{n+1}})))] \leq -E[\log(f(F^{-1}(1-e^{-W_n})))].$$

با این رابطه نتیجه کامل می‌شود. □

براتپور و همکاران (۲۰۰۷a) یک کران بالا و پایین را برای $I(f_{R_n}, f)$ به‌صورت زیر به‌دست آوردند که به تابع نرخ خطر $\lambda_F(y) = \frac{f(y)}{\bar{F}(y)}$ وابسته است:

$$I(f_{R_n}, f) \leq -B_n \int_A \lambda_F(y) \log f(y) dy, \tag{۷.۲}$$

و

$$I(f_{R_n}, f) \geq -B_n \int_{A^c} \lambda_F(y) \log f(y) dy, \tag{۸.۲}$$

که در آن $B_n = \frac{(n-1)^{n-1}}{(n-1)!} e^{-(n-1)}$ و $A = \{y | f(y) \leq 1\}$.

۳.۲ اندازه عدم دقت باقیمانده برای R_n

در این بخش اندازه عدم دقت باقیمانده‌ی پویا را بین f_{R_n} و f به‌صورت زیر معرفی می‌کنیم:

$$\begin{aligned} I(f_{R_n}, f; t) &= -\int_t^{+\infty} \frac{f_{R_n}(x)}{\bar{F}_{R_n}(t)} \log\left(\frac{f(x)}{\bar{F}(t)}\right) dx \\ &= \log \bar{F}(t) - \frac{1}{\bar{F}_{R_n}(t)} \int_t^{+\infty} f_{R_n}(x) \log f(x) dx. \end{aligned} \tag{۹.۲}$$

توجه داشته باشید که $\lim_{t\to 0} I(f_{R_n}, f; t) = I(f_{R_n}, f)$. از آنجا که برای $t \geq 0$ به‌دست می‌آید $\log \bar{F}(t) \leq 0$، داریم

$$\begin{aligned} I(f_{R_n}, f; t) &\leq -\frac{1}{\bar{F}_{R_n}(t)} \int_t^{+\infty} f_{R_n}(x) \log f(x) dx. \\ &\leq -\frac{1}{\bar{F}_{R_n}(t)} \int_0^{+\infty} f_{R_n}(x) \log f(x) dx = \frac{I(f_{R_n}, f)}{\bar{F}_{R_n}(t)}. \end{aligned}$$

گزاره ۱.۳.۲. فرض کنید $m = \sup\{x; f(x) \leq M\}$ مقدار مد تابع توزیع است. آن‌گاه

$$I(f_{R_n}, f; t) \leq \log\left(\frac{\bar{F}(t)}{M}\right).$$

برهان. فرض کنید m مد توزیع باشد، بنابراین $\log f(x) \leq \log M$. با استفاده از این رابطه اثبات کامل می‌شود. □

براتپور و همکاران (۲۰۰۷b) مشخصه‌سازی را بر اساس آنتروپی شانون برای آماره‌های ترتیبی و مقادیر رکورد بررسی کردند. تاپلیال و تانجا (۲۰۱۵) بر روی نتایج مشخصه‌سازی بر اساس عدم دقت باقیمانده‌ی پویا برای آماره‌های ترتیبی تمرکز کردند و اثبات کردند که به‌طور یکتا تابع توزیع را مشخص می‌کند.

لم ۱.۳.۲. فرض کنید f تابعی پیوسته در دامنه‌ی $D \subset \mathbb{R}^۲$ باشد، و f در شرط لیپ‌شیتس (با توجه به y) در D صدق می‌کند، یعنی برای هر نقطه‌ی $(x,y_۱)$ و $(x,y_۲)$ در D

$$| f(x,y_۱) - f(x,y_۲) | \leq k \mid y_۱ - y_۲ \mid, \quad k > ۰. \tag{۱۰.۲}$$

آن‌گاه تابع $y = f(x)$ که از مسئله‌ی مقدار اولیه به‌دست می‌آید $y = f(x,y)$ و $\phi(x_۰) = y_۰$، $x \in I$ منحصر به‌فرد است.

برهان. به گوپتا و کرمانی[۶] (۲۰۰۸) مراجعه شود. □

اکنون برای هر تابع دومتغیره $f(x,y)$ که در $D \subset \mathbb{R}^۲$ تعریف شده باشد، یک شرط کافی را معرفی می‌کنیم که تضمین می‌کند شرط لیپ‌شیتس در D برقرار شود.

لم ۲.۳.۲. فرض کنید f تابعی پیوسته در فضای محدب $D \subset \mathbb{R}^۲$ باشد. همچنین فرض کنید که $\frac{\partial f}{\partial y}$ موجود و در D پیوسته باشد. آنگاه تابع f در D در شرط لیپ‌شیتس صدق می‌کند.

برهان. به گوپتا و کرمانی (۲۰۰۸) مراجعه شود. □

حال با استفاده از دو لم ذکر شده در بالا، قضیه‌ی زیر را اثبات می‌کنیم.

قضیه ۱.۳.۲. فرض کنید X یک متغیر تصادفی پیوسته و نامنفی با تابع بقای $\bar{F}(.)$ باشد. فرض کنید عدم دقت باقیمانده‌ی پویا مربوط به nـامین مقدار رکورد را با $I(f_{R_n}, f; t) < \infty$ و $t \geq ۰$، نشان دهیم. آن‌گاه $I(f_{R_n}, f; t)$ به‌طور یکتا تابع بقای $\bar{F}(.)$ مربوط به متغیر تصادفی X را مشخص می‌کند.

برهان. با توجه به (۹.۲) داریم

$$I(f_{R_n}, f; t) = \log \bar{F}(t) - \frac{۱}{\bar{F}_{R_n}(t)} \int_t^{+\infty} f_{R_n}(x) \log f(x) dx. \tag{۱۱.۲}$$

با مشتق‌گیری از دو طرف رابطه‌ی (۱۱.۲) نسبت به t، به‌دست می‌آوریم:

$$\begin{aligned} \acute{I} = \frac{d}{dt}[I(f_{R_n}, f; t)] &= -\lambda_F(t) + \lambda_{F_{R_n}}(t) \left[I(f_{R_n}, f; t) + \log\left(\lambda_F(t)\right)\right] \\ &= -\lambda_F(t) + c(t)\lambda_F(t) \left[I(f_{R_n}, f; t) + \log\left(\lambda_F(t)\right)\right] \\ &= \lambda_F(t) \left(-۱ + c(t) \left[I(f_{R_n}, f; t) + \log\left(\lambda_F(t)\right)\right]\right), \end{aligned} \tag{۱۲.۲}$$

[۶] Gupta and Kirmani

که در آن $\lambda_{F_{R_n}}(t)$ نرخ خطر(نرخ شکست) R_n بوده و

$$c(t) = \frac{[-\log \bar{F}(x)]^{n-1}}{(n-1)! \sum_{j=0}^{n-1} \frac{[-\log \bar{F}(x)]^j}{j!}}.$$

با مشتق‌گیری دوباره نسبت به t به‌دست می‌آوریم

$$\acute{\lambda}_F(t) = \frac{\lambda_F(t)\left[I''(f_{R_n}, f; t)c(t) - \left(\acute{c}(t)\lambda_F(t) + \acute{I}(f_{R_n}, f; t)\acute{c}(t) + \acute{I}(f_{R_n}, f; t)c^2(t)\lambda_F(t)\right)\right]}{c(t)\left[c(t)\lambda_F(t) + \acute{I}(f_{R_n}, f; t)\right]}. \tag{۱۳.۲}$$

فرض کنید دو تابع F و F^* وجود دارند، به‌طوری‌که

$$I(f_{R_n}, f; t) = I(f^*_{R_n}, f; t) = h(t).$$

بنابراین برای هر $t \geq 0$، با توجه به (۱۳.۲) داریم

$$\acute{\lambda}_F(t) = \psi(t, \lambda_F(t)), \quad \acute{\lambda}_{F^*}(t) = \psi(t, \lambda_{F^*}(t)),$$

که در آن

$$\psi(t, y) = \frac{y\left[h''(t)c(t) - \left(\acute{c}(t)y + \acute{h}(t)\acute{c}(t) + \acute{h}(t)c^2(t)y\right)\right]}{c(t)\left[c(t)y + \acute{h}(t)\right]}.$$

با استفاده از لم‌های ۱.۳.۲ و ۲.۳.۲ برای هر t داریم، $\lambda_F(t) = \lambda_{F^*}(t)$. با توجه به اینکه تابع نرخ خطر به‌طور یکتا تابع توزیع را مشخص می‌کند، اثبات کامل می‌شود. □

مشابه اندازه‌ی (۹.۲)، اندازه عدم دقت گذشته‌ی پویا بین f_{R_n} و f به‌صورت زیر به‌دست می‌آید

$$\begin{aligned}
\tilde{I}(f_{R_n}, f; t) &= -\int_0^t \frac{f_{R_n}(x)}{F_{R_n}(t)} \log\left(\frac{f(x)}{F(t)}\right) dx \\
&= \log F(t) - \frac{1}{F_{R_n}(t)} \int_0^t f_{R_n}(x) \log f(x) dx,
\end{aligned}$$

به سادگی می‌توان نشان داد $\lim_{t\to\infty} \tilde{I}(f_{R_n}, f; t) = \tilde{I}(f_{R_n}, f)$. توجه کنید که

$$\tilde{I}(f_{R_n}, f; t) \leq \frac{I(f_{R_n}, f)}{F_{R_n}(t)},$$

و به‌طور مشابه $\tilde{I}(f_{R_n}, f; t) \geq \log\left(\frac{F(t)}{M}\right)$.

قضیه ۲.۳.۲. فرض کنید X یک متغیر تصادفی پیوسته و نامنفی با تابع توزیع $F(.)$ باشد. فرض کنید عدم دقت گذشته‌ی پویا مربوط به nـامین مقدار رکورد برای هر $t \geq 0$ با $\tilde{I}(f_{R_n}, f; t) < \infty$ نشان داده شود. آن‌گاه $\tilde{I}(f_{R_n}, f; t)$ به‌طور یکتا تابع توزیع $F(\cdot)$ را مشخص می‌کند.

برهان. مشابه اثبات قضیه‌ی ۱.۳.۲ می‌باشد. □

۴.۲ عدم دقت باقیمانده‌ی تجمعی برای R_n

اندازه عدم دقت باقیمانده‌ی تجمعی بین $\bar{F}_{R_n}$ و $\bar{F}$ را به صورت زیر بیان می کنیم:

$$\begin{aligned} I(\bar{F}_{R_n}, \bar{F}) &= -\int_{0}^{+\infty} \bar{F}_{R_n}(x) \log\left(\bar{F}(x)\right) dx \\ &= -\int_{0}^{+\infty} \sum_{j=0}^{n-1} \frac{[-\log \bar{F}(x)]^{j}}{j!} \bar{F}(x) \log\left(\bar{F}(x)\right) dx \\ &= \sum_{j=0}^{n-1} \int_{0}^{+\infty} \frac{[-\log \bar{F}(x)]^{j+1}}{j!} \bar{F}(x) dx \\ &= \sum_{j=0}^{n-1} \int_{0}^{+\infty} \frac{[-\log \bar{F}(x)]^{j+1}}{j!} \times \frac{f(x)}{\frac{f(x)}{\bar{F}(x)}} dx \\ &= \sum_{j=0}^{n-1} (j+1) E_{R_{j+2}}\left(\frac{1}{\lambda_F(X)}\right), \end{aligned} \tag{۱۴.۲}$$

که R_{j+2} یک متغیر تصادفی با قابلیت اطمینان $\bar{F}_{R_{j+2}}$ است. در مثال زیر ما $I(\bar{F}_{R_n}, \bar{F})$ را برای برخی توزیع‌های طول عمر خاص که کاربرد وسیعی در نظریه قابلیت اطمینان و آزمون بقا دارند، بدست آورده‌ایم.

مثال ۱.۴.۲. (۱) اگر X دارای توزیع یکنواخت در $[0, \theta]$ باشد، آن‌گاه به راحتی می‌توان دید که برای هر عدد صحیح $n \geq 1$، $I(\bar{F}_{R_n}, \bar{F}) = \theta \sum_{j=0}^{n-1} \frac{j!(j+1)^2}{2^{j+2}}$.

(۲) اگر X دارای توزیع وایبل با تابع بقای $\bar{F}(x) = e^{-\alpha x^{\beta}}$ ، $x \geq 0$ ، $\alpha, \beta > 0$، باشد، آن‌گاه برای تمام مقادیر صحیح $n \geq 1$، $I(\bar{F}_{R_n}, \bar{F}) = \frac{1}{\beta \alpha^{\frac{1}{\alpha}}} \sum_{j=0}^{n-1} \frac{(\frac{1}{\beta}+j)!}{j!}$.

(۳) اگر X دارای توزیع پارتو با تابع بقای $\bar{F}(x) = \left(\frac{\lambda}{x+\lambda}\right)^{\gamma}$، $x \geq 0$، $\gamma > 1$، $\lambda > 0$، باشد، آن‌گاه برای تمام مقادیر صحیح $n \geq 1$، $I(\bar{F}_{R_n}, \bar{F}) = \frac{\lambda}{\gamma-1} \sum_{j=0}^{n-1} (j+1) \left(\frac{\gamma}{\gamma-1}\right)^{j+1}$.

(۴) فرض کنید X دارای توزیع نمایی با میانگین $\frac{1}{\lambda}$ باشد، آن‌گاه $I(\bar{F}_{R_n}, \bar{F}) = \frac{n(n+1)}{2\lambda}$.

(۵) فرض کنید X یک متغیر تصادفی نامنفی باشد که دارای توزیع نمایی‌ـ گاوسی معکوس با تابع بقای $\bar{F}(x) = e^{\frac{\alpha}{\beta}[1-\sqrt{1+2\beta x}]},\ x \geq 0,\ \alpha, \beta > 0$ است، آن‌گاه برای تمام مقادیر صحیح $n \geq 1$، $I(\bar{F}_{R_n}, \bar{F}) = \frac{1}{\alpha^2} \sum_{j=0}^{n-1} j!(j+1)^2[\alpha + (j+2)\beta]$.

گزاره ۱.۴.۲. فرض کنید X یک متغیر تصادفی نامنفی و مطلقاً پیوسته با تابع بقای $\bar{F}$ باشد. در این‌صورت

$$I(\bar{F}_{R_n}, \bar{F}) = \sum_{j=0}^{n-1} (j+1)[\mu_{j+2} - \mu_{j+1}],$$

که در آن $\mu_n = \int_0^{+\infty} \bar{F}_{R_n}(x) dx$.

برهان. با توجه به (۱۴.۲) داریم

$$\begin{aligned} I(\bar F_{R_n}, \bar F) &= \sum_{j=0}^{n-1} \int_0^{+\infty} \frac{[-\log \bar F(x)]^{j+1}}{j!} \bar F(x) dx \\ &= \sum_{j=0}^{n-1} (j+1) \int_0^{+\infty} [\bar F_{R_{j+2}}(x) - \bar F_{R_{j+1}}(x)] dx \\ &= \sum_{j=0}^{n-1} (j+1) [\mu_{j+2} - \mu_{j+1}]. \end{aligned}$$

□

گزاره ۲.۴.۲. فرض کنید $a, b > 0$. در این‌صورت برای $n = 1, 2, \ldots$

$$I(\bar F_{aR_n+b}, \bar F_{aX+b}) = a I(\bar F_{R_n}, \bar F).$$

برهان. با استفاده از (۱۴.۲) و توجه به این‌که $\bar F_{aX+b}(x) = \bar F(\frac{x-b}{a})$، داریم

$$I(\bar F_{aR_n+b}, \bar F_{aX+b}) = -\int_0^{+\infty} \sum_{j=0}^{n-1} \frac{[-\log \bar F_{aX+b}(x)]^{j+1}}{j!} \bar F_{aX+b}(x) dx = a I(\bar F_{R_n}, \bar F). \quad (۱۵.۲)$$

□

گزاره ۳.۴.۲. فرض کنید X یک متغیر تصادفی نامنفی و مطلقاً پیوسته با تابع بقای $\bar F$ باشد. در این‌صورت

$$I(\bar F_{R_n}, \bar F) = \sum_{j=0}^{n-1} \frac{1}{j!} \int_0^{\infty} \lambda_F(z) \left[\int_z^{\infty} [-\log \bar F(x)]^j \bar F(x) dx \right] dz.$$

برهان. با استفاده از (۱۴.۲) و این حقیقت که $-\log \bar F(x) = \int_0^x \lambda_F(z) dz$، داریم

$$\begin{aligned} I(\bar F_{R_n}, \bar F) &= \sum_{j=0}^{n-1} \int_0^{+\infty} \frac{[-\log \bar F(x)]^{j+1}}{j!} \bar F(x) dx \\ &= \sum_{j=0}^{n-1} \int_0^{+\infty} \left[\int_0^x \lambda_F(z) dz \right] \frac{[-\log \bar F(x)]^j}{j!} \bar F(x) dx \\ &= \sum_{j=0}^{n-1} \frac{1}{j!} \int_0^{+\infty} \lambda_F(z) \left[\int_z^{\infty} [-\log \bar F(x)]^j \bar F(x) dx \right] dz. \end{aligned}$$

□

گزاره ۴.۴.۲. فرض کنید X یک متغیر تصادفی نامنفی با تابع بقای $\bar F$ باشد. در این‌صورت

$$I(\bar F_{R_n}, \bar F) = \sum_{j=0}^{n-1} \frac{1}{j!} E[X(-\log \bar F(X))^{j+1}] - \sum_{j=0}^{n-1} \frac{j+1}{j!} E[X(-\log \bar F(X))^j].$$

برهان. با توجه به (۱۴.۲) و استفاده از قضیه‌ی فوبینی (پیوست آ.۴) داریم

$$\begin{aligned} I(\bar{F}_{R_n}, \bar{F}) &= \sum_{j=0}^{n-1} \int_0^{+\infty} \frac{[-\log \bar{F}(x)]^{j+1}}{j!} \bar{F}(x) dx \\ &= \sum_{j=0}^{n-1} \int_0^{+\infty} \frac{[-\log \bar{F}(x)]^{j+1}}{j!} \left[\int_x^{+\infty} f(z) dz\right] dx \\ &= \sum_{j=0}^{n-1} \int_0^{+\infty} \frac{f(z)}{j!} \left[\int_0^{z} [-\log \bar{F}(x)]^{j+1} dx\right] dz. \end{aligned}$$

با انتگرال‌گیری جزء به جزء نتیجه به راحتی به‌دست می‌آید. □

گزاره ۵.۴.۲. فرض کنید X یک متغیر تصادفی نامنفی و مطلقاً پیوسته با $I(\bar{F}_{R_n}, \bar{F}) < \infty$ برای $n \geq 1$ باشد. آن‌گاه داریم

$$I(\bar{F}_{R_n}, \bar{F}) = \sum_{j=0}^{n-1} \frac{1}{j!} E\left[h_{j+1}(X)\right], \tag{۱۶.۲}$$

که در آن

$$h_{j+1}(x) = \int_0^x \left[-\log \bar{F}(z)\right]^{j+1} dz, \quad x \geq 0.$$

برهان. با توجه به رابطه‌ی (۱۴.۲) و قضیه فوبینی به‌دست می‌آوریم

$$\begin{aligned} I(\bar{F}_{R_n}, \bar{F}) &= \sum_{j=0}^{n-1} \int_0^{\infty} \frac{[-\log \bar{F}(z)]^{j+1}}{j!} \bar{F}(z) dz \\ &= \sum_{j=0}^{n-1} \frac{1}{j!} \int_0^{\infty} \left[\int_z^{\infty} f(x) dx\right] [-\log \bar{F}(z)]^{j+1} dz \\ &= \sum_{j=0}^{n-1} \frac{1}{j!} \int_0^{\infty} f(x) \left[\int_0^{x} [-\log \bar{F}(z)]^{j+1} dz\right] dx = \sum_{j=0}^{n-1} \frac{1}{j!} E\left[h_{j+1}(X)\right]. \end{aligned}$$

□

در گزاره‌های زیر برخی کران‌های بالا و پایین را برای $I(\bar{F}_{R_n}, \bar{F})$ به‌دست می‌آوریم.

گزاره ۶.۴.۲. برای یک متغیر تصادفی نامنفی X و $n \geq 1$، رابطه‌ی زیر برقرار می‌شود

$$I(\bar{F}_{R_n}, \bar{F}) \geq \sum_{j=0}^{n-1} \frac{[\mathcal{E}(X)]^{j+1}}{j!}, \tag{۱۷.۲}$$

که در آن

$$\mathcal{E}(X) = -\int_0^{+\infty} \bar{F}(x) \log \bar{F}(x) dx, \tag{۱۸.۲}$$

آنتروپی باقیمانده‌ی تجمعی[۷] (CRE) است که توسط رائو و همکاران (۲۰۰۴) معرفی شده است.

[۷] Cumulative residual entropy

برهان. از آنجا که برای $n \geq ۱$، $\bar{F}(x) \geq (\bar{F}(x))^n$، داریم

$$I(\bar{F}_{R_n}, \bar{F}) = \sum_{j=۰}^{n-۱} \frac{۱}{j!} \int_۰^\infty \bar{F}(x)[-\log \bar{F}(x)]^{j+۱} dx \geq \sum_{j=۰}^{n-۱} \frac{۱}{j!} \int_۰^\infty [-\bar{F}(x) \log \bar{F}(x)]^{j+۱} dx.$$

با توجه به این‌که برای $g(x) = x^n, n \geq ۱$ تابعی محدب است، نامساوی جنسن نتیجه می‌دهد

$$I(\bar{F}_{R_n}, \bar{F}) = \sum_{j=۰}^{n-۱} \frac{۱}{j!} \int_۰^\infty \bar{F}(x)[-\log \bar{F}(x)]^{j+۱} dx \geq \sum_{j=۰}^{n-۱} \frac{۱}{j!} \left(-\int_۰^\infty \bar{F}(x) \log \bar{F}(x) dx \right)^{j+۱},$$

که با توجه به رابطه‌ی (۱۸.۲) اثبات به‌دست می‌آید. □

نتیجه ۱.۴.۲. فرض کنید X یک متغیر تصادفی نامنفی و مطلقاً پیوسته با تابع توزیع F باشد. آن‌گاه داریم

$$I(\bar{F}_{R_n}, \bar{F}) \geq \sum_{j=۰}^{n-۱} \frac{۱}{j!} \left[\int_۰^\infty F(x) \bar{F}(x) dx \right]^{j+۱}.$$

برهان. با توجه به گزاره ۱ در رائو (۲۰۰۵)، یک کران پایین برای CRE به‌صورت زیر است

$$\mathcal{E}(X) \geq \int_۰^{+\infty} F(x) \bar{F}(x) dx.$$

بنابراین، با استفاده از گزاره ۶.۴.۲ اثبات کامل می‌شود. □

نتیجه ۲.۴.۲. فرض کنید X یک متغیر تصادفی نامنفی پیوسته با تابع بقای $\bar{F}$ و امید متناهی باشد. آن‌گاه داریم

$$I(\bar{F}_{R_n}, \bar{F}) \geq \sum_{j=۰}^{n-۱} \frac{۱}{j!} [E(X) gini[X]]^{j+۱},$$

که در آن $gini[.]$ شاخص جینی (پیوست آ.۳) است.

برهان. با استفاده از گزاره‌ی ۵.۱ از وانگ[۸] (۱۹۹۸) داریم

$$\int_۰^{+\infty} F(x) \bar{F}(x) dx = \frac{۱}{۲} E[\mid X - Y \mid] = E(X) gini[X],$$

که X و Y مستقل و دارای توزیع یکسان هستند. بنابراین، نتیجه‌ی ۱.۴.۲ اثبات را کامل می‌کند. □

نتیجه ۳.۴.۲. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته باشد. در این‌صورت

$$I(\bar{F}_{R_n}, \bar{F}) \geq \sum_{j=۰}^{n-۱} \frac{C^{j+۱}}{j!} \left[e^{(j+۱)H(X)} \right],$$

که در آن $C = \exp\left(\int_۰^۱ \log(x \mid \log x \mid) dx \right) = ۰/۲۰۶۵$ و $H(X) = -\int_۰^{+\infty} f(x) \log f(x) dx$ آنتروپی شانون X است.

[۸] Wang

برهان. با توجه به رابطه‌ی (۱۷.۲) و گزاره‌ی ۲.۴ از دی کرشنزو و لونگوبردی[۹] (۲۰۰۹) اثبات به‌دست می‌آید. $\square$

گزاره ۷.۴.۲. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته با امید متناهی $\mu = E(X)$ باشد. در این‌صورت

$$I(\bar{F}_{R_n}, \bar{F}) \geq \sum_{j=0}^{n-1} \frac{1}{j!} \left[h_{j+1}(\mu)\right],$$

که $h_{j+1}(\mu)$ در گزاره‌ی ۵.۴.۲ تعریف شده است.

برهان. با توجه به اینکه $h_{j+1}(.)$ تابعی محدب است، با به‌کار بردن نامساوی جنسن در (۱۶.۲) اثبات کامل می‌شود. $\square$

گزاره ۸.۴.۲. فرض کنید X یک متغیر تصادفی نامنفی با تابع توزیع تجمعی کاملاً پیوسته $F(x)$ باشد. آن‌گاه برای $n = 1, 2, \ldots$ داریم

$$I(\bar{F}_{R_n}, \bar{F}) \leq \sum_{j=0}^{n-1} \frac{1}{j!} \int_0^\infty [-\log \bar{F}(x)]^{j+1} dx.$$

برهان. با استفاده از (۱۴.۲) و این واقعیت که برای هر x، $\bar{F}(x) \leq 1$، اثبات به آسانی به‌دست می‌آید. $\square$

گزاره ۹.۴.۲. فرض کنید X یک متغیر تصادفی نامنفی با تابع بقای $\bar{F}(x)$ باشد. آن‌گاه برای $n = 1, 2, \ldots$ داریم

$$I(\bar{F}_{R_n}, \bar{F}) \geq \sum_{j=0}^{n-1} \sum_{i=0}^{j+1} \frac{(-1)^i (j+1)}{i!(j+1-i)!} \int_0^\infty [\bar{F}(x)]^{i+1} dx.$$

برهان. از آنجا که $-\log \bar{F}(x) \geq 1 - \bar{F}(x)$، با یادآوری (۱۴.۲) اثبات نتیجه می‌شود. $\square$

در ادامه برخی نتایج را برای $I(\bar{F}_{R_n}, \bar{F})$ به‌دست می‌آوریم و ارتباط آن را با مفهوم قابلیت اطمینان بیان می‌کنیم.

گزاره ۱۰.۴.۲. اگر X DFR باشد، آن‌گاه برای $n = 1, 2, \ldots$ داریم

$$I(\bar{F}_{R_{n+1}}, \bar{F}) - I(\bar{F}_{R_n}, \bar{F}) \geq \sum_{i=1}^{n+1} E_{R_i}\left[\frac{1}{\lambda_F(X)}\right]. \qquad (۱۹.۲)$$

برهان. فرض کنید f_{R_n} تابع چگالی nــامین مقدار رکورد R_n باشد. چون نسبت

$$\frac{f_{R_{n+1}}(t)}{f_{R_n}(t)} = \frac{-\log \bar{F}(t)}{n}$$

[۹] Di Cresenzo and Longobardi

نسبت به t صعودی است، بنابراین $R_n \leq^{lr} R_{n+1}$ که نتیجه می‌دهد $R_n \leq^{st} R_{n+1}$، به این معنی که $\bar{F}_{R_n} \leq \bar{F}_{R_{n+1}}$ (برای جزئیات بیشتر شیکد و شانتی‌کومار[10] ۲۰۰۷، فصل ۱ را مطالعه کنید). بنابراین، اگر X DFR و $\lambda_F(x)$ نرخ مخاطره باشد، آن‌گاه $\frac{1}{\lambda_F(x)}$ تابعی صعودی از x است. بنابراین با استفاده از (۱۴.۲) و نامساوی (۱.A.۷) در شیکد و شانتی کومار (۲۰۰۷) داریم

$$\begin{aligned} I(\bar{F}_{R_{n+1}}, \bar{F}) &= \sum_{j=0}^{n}(j+1)E_{R_{j+2}}\left[\frac{1}{\lambda_F(X)}\right] \\ &\geq \sum_{j=0}^{n}(j+1)E_{R_{j+1}}\left[\frac{1}{\lambda_F(X)}\right] \\ &= \sum_{i=-1}^{n-1}(i+2)E_{R_{i+2}}\left[\frac{1}{\lambda_F(X)}\right] \\ &= \sum_{i=0}^{n-1}(i+2)E_{R_{i+2}}\left[\frac{1}{\lambda_F(X)}\right] + E_{R_1}\left[\frac{1}{\lambda_F(X)}\right] \\ &= I(\bar{F}_{R_n}, \bar{F}) + \sum_{i=1}^{n+1}E_{R_i}\left[\frac{1}{\lambda_F(X)}\right]. \end{aligned} \tag{۲۰.۲}$$

اثبات کامل است. □

گزاره ۱۱.۴.۲. اگر X دارای توزیع نمایی با میانگین $\mu = \frac{1}{\theta}$ باشد، از آنجاکه نرخ خطر مقداری ثابت است، رابطه‌ی $I(\bar{F}_{R_n}, \bar{F}) = \frac{n(n+1)}{2}\mu$ را به‌دست می‌آوریم که تابعی صعودی از n است.

گزاره ۱۲.۴.۲. فرض کنید X و Y دو متغیر تصادفی نامنفی با توابع بقای به‌ترتیب $\bar{F}(x)$ و $\bar{G}(x)$ باشند. اگر $X \leq^{hr} Y$، X DFR و $\tilde{R}_n$ رکورد بالای مربوط به Y باشد، آن‌گاه برای $n = 1, 2, \ldots$

$$I(\bar{F}_{R_n}, \bar{F}) \leq I(\bar{G}_{\tilde{R}_n}, \bar{G}). \tag{۲۱.۲}$$

برهان. می‌دانیم که $X \leq^{hr} Y$ نتیجه می‌دهد $X \leq^{st} Y$ (شیکد و شانتی کومار ۲۰۰۷ را ببینید). بنابراین، داریم

$$\bar{F}_{R_{j+2}} \leq \bar{G}_{\tilde{R}_{j+2}},$$

که $\bar{G}_{\tilde{R}_{j+2}}$ تابع بقای $\tilde{R}_{j+2}$ می‌باشد. در نتیجه رابطه‌ی $R_{j+2} \leq^{st} \tilde{R}_{j+2}$ برقرار می‌شود. این معادل است با این‌که برای تمام توابع صعودی ϕ داشته باشیم

$$E(\phi(R_{j+2})) \leq E(\phi(\tilde{R}_{j+2})),$$

در صورتی که امیدها موجود باشند (شیکد و شانتی کومار ۲۰۰۷، ص۴ را ببینید). بنابراین اگر فرض کنیم X DFR است و $\lambda_F(x)$ نرخ شکست است، آن‌گاه $\frac{1}{\lambda_F(x)}$ صعودی است و داریم

$$I(\bar{F}_{R_n}, \bar{F}) = \sum_{j=0}^{n-1}(j+1)E_{R_{j+2}}\left(\frac{1}{\lambda_F(X)}\right) \leq \sum_{j=0}^{n-1}(j+1)E_{\tilde{R}_{j+2}}\left(\frac{1}{\lambda_F(X)}\right).$$

[10] Shaked and Shanthikumar

از طرف دیگر، $X \leq^{hr} Y$ نتیجه می‌دهد که توابع نرخ خطر مربوطه در رابطه‌ی $\lambda_F(x) \geq \lambda_G(y)$ صدق می‌کنند. پس داریم

$$\sum_{j=0}^{n-1}(j+1)E_{\tilde{R}_{j+2}}\left(\frac{1}{\lambda_F(X)}\right) \leq \sum_{j=0}^{n-1}(j+1)E_{\tilde{R}_{j+2}}\left(\frac{1}{\lambda_G(Y)}\right) = I(\bar{G}_{\tilde{R}_n}, \bar{G}).$$

بنابراین، با استفاده از این دو عبارت به‌دست می‌آوریم $I(\bar{F}_{R_n}, \bar{F}) \leq I(\bar{G}_{\tilde{R}_n}, \bar{G})$. □

گزاره ۱۳.۴.۲. فرض کنید X و Y دو متغیر تصادفی نامنفی با توابع بقای $\bar{F}(x)$ و $\bar{G}(x)$ باشند. اگر $X \leq^{icx} Y$، آن‌گاه

$$I(\bar{F}_{R_n}, \bar{F}) \leq I(\bar{G}_{\tilde{R}_n}, \bar{G}).$$

برهان. از آنجا که $g_{j+1}(.)$ برای $j \geq 0$، یک تابع محدب صعودی است، با توجه به شیکد و شانتی کومار (۲۰۰۷) می‌دانیم که $X \leq^{icx} Y$ نتیجه می‌دهد $h_{j+1}(X) \leq^{icx} h_{j+1}(Y)$. با یادآوری تعریف ترتیب محدب صعودی و گزاره‌ی ۵.۴.۲ اثبات کامل می‌شود. □

گزاره ۱۴.۴.۲. اگر X IFRA (DFRA) باشد، آن‌گاه برای $n = 1, 2, \ldots$ داریم

$$I(\bar{F}_{R_n}, \bar{F}) \leq (\geq) \sum_{j=0}^{n-1} \frac{1}{j!} E\left[X\left(-\log \bar{F}(X)\right)^j\right]. \tag{۲۲.۲}$$

برهان. از رابطه‌ی (۱۴.۲) داریم

$$I(\bar{F}_{R_n}, \bar{F}) = \sum_{j=0}^{n-1} \int_0^{+\infty} \frac{[-\log \bar{F}(x)]^j}{j!}[-\log \bar{F}(x)]\bar{F}(x)dx. \tag{۲۳.۲}$$

حال از آنجا که X IFRA (DFRA) می‌باشد، $\frac{-\log \bar{F}(x)}{x}$ نسبت به $x > 0$ صعودی (نزولی) است که نتیجه می‌دهد

$$-\bar{F}(x)\log \bar{F}(x) \leq (\geq) x f(x), \quad x > 0. \tag{۲۴.۲}$$

بنابراین با توجه به (۲۳.۲) و (۲۴.۲) اثبات کامل می‌شود. □

گزاره ۱۵.۴.۲. فرض کنید X و Y دو متغیر تصادفی نامنفی با توابع بقای $\bar{F}(x)$ و $\bar{G}(x)$ باشند. اگر $X \leq^{hr} Y$، آن‌گاه برای $n = 1, 2, \ldots$

$$\frac{I(\bar{F}_{R_n}, \bar{F})}{E(X)} \leq \frac{I(\bar{G}_{\tilde{R}_n}, \bar{G})}{E(Y)}.$$

برهان. با توجه به اینکه $h_{j+1}(x) = \int_0^x [-\log \bar{F}(z)]^{j+1} dz$ یک تابع محدب صعودی است، تحت فرض $X \leq^{hr} Y$، طبق شیکد و شانتی کومار (۲۰۰۷) داریم

$$\sum_{j=0}^{n-1} \frac{1}{j!}\left[\frac{E\left[h_{j+1}(X)\right]}{E(X)}\right] \leq \sum_{j=0}^{n-1} \frac{1}{j!}\left[\frac{E\left[h_{j+1}(Y)\right]}{E(Y)}\right].$$

بنابراین، با یادآوری (۱۶.۲) اثبات کامل می‌شود. □

فرض کنید X_θ^* یک متغیر تصادفی کاملاً پیوسته و نامنفی را نشان می‌دهد که دارای تابع بقای $\bar{H}_\theta(x) = [\bar{F}(x)]^\theta, x \geq 0$ می‌باشد. حال اندازه عدم دقت باقیمانده‌ی تجمعی را بین $\bar{H}_{R_n}$ و $\bar{H}_\theta$ به‌صورت زیر به‌دست می‌آوریم:

$$\begin{aligned} I(\bar{H}_{R_n}, \bar{H}_\theta) &= -\int_0^{+\infty} \bar{H}_{R_n}(x) \log\left(\bar{H}_\theta(x)\right) dx \\ &= \sum_{j=0}^{n-1} \theta^{j+1} \int_0^{+\infty} \frac{[-\log \bar{F}(x)]^{j+1}}{j!} [\bar{F}(x)]^\theta dx. \end{aligned} \tag{۲۵.۲}$$

گزاره ۱۶.۴.۲. اگر $\theta \geq (\leq) 1$، آن‌گاه برای هر $n \geq 1$ داریم

$$I(\bar{H}_{R_n}, \bar{H}_\theta) \leq (\geq) \sum_{j=0}^{n-1} (j+1)\theta^{j+1} \mathcal{E}_{j+1}(X),$$

که $\mathcal{E}_{j+1}(X)$ آنتروپی باقیمانده‌ی تجمعی تعمیم یافته‌ی X است که توسط ساراکوس و ناوارو[11] (۲۰۱۳) به‌صورت زیر معرفی شده است

$$\mathcal{E}_{j+1}(X) = \int_0^{+\infty} \frac{\bar{F}(x)[-\log \bar{F}(x)]^{j+1}}{(j+1)!}.$$

برهان. فرض کنید $\theta \geq (\leq) 1$، آن‌گاه واضح است که $[\bar{F}(x)]^\theta \leq (\geq) \bar{F}(x)$، و بنابراین از (۲۴.۲) نتیجه می‌شود

$$I(\bar{H}_{R_n}, \bar{H}_\theta) \leq (\geq) \sum_{j=0}^{n-1} (j+1)\theta^{j+1} \mathcal{E}_{j+1}(X).$$

□

در ادامه‌ی این فصل، نسخه‌ی پویای $I(\bar{F}_{R_n}, \bar{F})$ را بررسی می‌کنیم. فرض کنید X طول عمر یک سیستم باشد به‌طوری که سیستم تا سن t سالم مانده باشد. به‌طور مشابه می‌توانیم نسخه‌ی پویای $I(\bar{F}_{R_n}, \bar{F})$ را به‌صورت زیر مطرح کنیم

$$\begin{aligned} I(\bar{F}_{R_n}, \bar{F}; t) &= -\int_t^{+\infty} \frac{\bar{F}_{R_n}(x)}{\bar{F}_{R_n}(t)} \log\left(\frac{\bar{F}(x)}{\bar{F}(t)}\right) dx \\ &= \log \bar{F}(t) \tilde{\mu}_n(t) - \int_t^{+\infty} \frac{\bar{F}_{R_n}(x)}{\bar{F}_{R_n}(t)} \log\left(\bar{F}(x)\right) dx \\ &= \log \bar{F}(t) \tilde{\mu}_n(t) + \frac{1}{\bar{F}_{R_n}(t)} \sum_{j=0}^{n-1} \int_t^{+\infty} \frac{[-\log \bar{F}(x)]^{j+1}}{j!} \bar{F}(x) dx \end{aligned} \tag{۲۶.۲}$$

که در آن $\tilde{\mu}_n(t) = \int_t^{+\infty} \frac{\bar{F}_{R_n}(x)}{\bar{F}_{R_n}(t)} dx$. توجه کنید که $\lim_{t\to 0} I(\bar{F}_{R_n}, \bar{F}; t) = I(\bar{F}_{R_n}, \bar{F})$. از آنجاکه برای $t \geq 0$، $\log \bar{F}(t) \leq 0$، داریم

$$\begin{aligned} I(\bar{F}_{R_n}, \bar{F}; t) &\leq \frac{1}{\bar{F}_{R_n}(t)} \sum_{j=0}^{n-1} \int_t^{+\infty} \frac{[-\log \bar{F}(x)]^{j+1}}{j!} \bar{F}(x) dx \\ &\leq \frac{1}{\bar{F}_{R_n}(t)} \sum_{j=0}^{n-1} \int_0^{+\infty} \frac{[-\log \bar{F}(x)]^{j+1}}{j!} \bar{F}(x) dx = \frac{I(\bar{F}_{R_n}, \bar{F})}{\bar{F}_{R_n}(t)}. \end{aligned}$$

[11] Psarrakos and Navarro

قضیه ۲.۴.۱. فرض کنید X یک متغیر تصادفی پیوسته و نامنفی با تابع توزیع $F(.)$ باشد. فرض کنید عدم دقت تجمعی پویا برای nـامین رکورد با $I(\bar{F}_{R_n}, \bar{F}; t) < \infty$ و $t \geq 0$ نشان داده شود. آنگاه $I(\bar{F}_{R_n}, \bar{F}; t)$ تابع توزیع را مشخصه‌سازی می‌کند.

برهان. با توجه به (۲.۲۶) داریم

$$I(\bar{F}_{R_n}, \bar{F}; t) = \log \bar{F}(t)\tilde{\mu}_n(t) + \frac{1}{\bar{F}_{R_n}(t)} \sum_{j=0}^{n-1} \int_t^{+\infty} \frac{[-\log \bar{F}(x)]^{j+1}}{j!} \bar{F}(x)dx. \quad (۲.۲۷)$$

با مشتق‌گیری از دو طرف (۲.۲۷) نسبت به t به‌دست می‌آوریم:

$$\begin{aligned} \frac{d}{dt}[I(\bar{F}_{R_n}, \bar{F}; t)] &= -\lambda_F(t)\tilde{\mu}_n(t) + \lambda_{F_{R_n}}(t) I(\bar{F}_{R_n}, \bar{F}; t) \\ &= -\lambda_F(t)\tilde{\mu}_n(t) + c(t)\lambda_F(t) I(\bar{F}_{R_n}, \bar{F}; t) \\ &= \lambda_F(t)\left[-\tilde{\mu}_n(t) + c(t) I(\bar{F}_{R_n}, \bar{F}; t)\right]. \end{aligned}$$

با مشتق‌گیری مجدد نسبت به t خواهیم داشت

$$\acute{\lambda}_F(t) = \frac{(\lambda_F(t))^2 \left[-\acute{c}(t) I(\bar{F}_{R_n}, \bar{F}; t) - c(t)\acute{I}(\bar{F}_{R_n}, \bar{F}; t) + 1 - c(t)\lambda_F(t)\tilde{\mu}_n(t)\right]}{\acute{I}(\bar{F}_{R_n}, \bar{F}; t)}. \quad (۲.۲۸)$$

فرض کنید دو تابع F و F^* موجود باشند به‌طوری‌که

$$I(\bar{F}_{R_n}, \bar{F}; t) = I(\bar{F}^*_{R_n}, \bar{F}^*; t) = z(t).$$

آنگاه برای تمام مقادیر t، طبق (۲.۲۸) داریم

$$\acute{\lambda}_F(t) = \varphi(t, \lambda_F(t)), \quad \acute{\lambda}_{F^*}(t) = \varphi(t, \lambda_{F^*}(t)),$$

که

$$\varphi(t, y) = \frac{y^2 \left[-\acute{c}(t)z(t) - c(t)\acute{z}(t) + 1 - c(t)yw(t)\right]}{\acute{z}(t)},$$

و $w(t) = \log \bar{F}(t)$. با استفاده از لم ۲.۳.۱ و ۲.۳.۲ برای هر مقدار t داریم $\lambda_F(t) = \lambda_{F^*}(t)$. از آنجا که تابع نرخ خطر تابع توزیع را به‌طور یکتا مشخصه‌سازی می‌کند، اثبات کامل می شود. □

۵.۲ اندازه عدم دقت تجمعی در رکوردهای پایین

اندازه عدم دقت تجمعی بین F_{L_n} (تابع توزیع nــامین رکورد پایین) و F به‌صورت زیر به‌دست می‌آید

$$\begin{aligned} I(F_{L_n}, F) &= -\int_0^\infty F_{L_n}(x)\log F(x)dx \\ &= -\int_0^\infty F(x)\sum_{j=0}^{n-1}\frac{[-\log F(x)]^j}{j!}\log F(x)dx \\ &= \sum_{j=0}^{n-1}\int_0^\infty \frac{[-\log F(x)]^{j+1}}{j!}F(x)dx \\ &= \sum_{j=0}^{n-1}(j+1)\int_0^\infty \frac{[-\log F(x)]^{j+1}}{(j+1)!}f(x)\frac{1}{\tilde{\lambda}(x)}dx \\ &= \sum_{j=0}^{n-1}(j+1)E_{L_{j+2}}\left[\frac{1}{\tilde{\lambda}(X)}\right], \end{aligned} \quad (۲۹.۲)$$

که در آن $\tilde{\lambda}(x) = \frac{f(x)}{F(x)}$ نرخ خطر معکوس و L_{j+2} یک متغیر تصادفی با تابع چگالی $f_{L_{j+2}}(x) = \frac{[-\log F(x)]^{j+1}f(x)}{(j+1)!}$ است.

در زیر چند مثال و برخی خواص $I(F_{L_n}, F)$ را مطرح می‌کنیم.

مثال ۱.۵.۲. ۱. اگر X دارای توزیع وایبل معکوس با تابع توزیع $F(x) = \exp(-(\frac{\alpha}{x})^\beta),\ x > 0$ باشد، آن‌گاه

$$I(F_{L_n}, F) = \frac{\alpha}{\beta}\sum_{j=0}^{n-1}\frac{\Gamma\left(\frac{(j+1)\beta-1}{\beta}\right)}{j!}.$$

شکل ۴.۲ تابع $I(F_{L_n}, F)$ را برای $\alpha = \beta = 2$ نشان می‌دهد. این تابع بر حسب n صعودی است.

۲. اگر X دارای توزیع یکنواخت در $[0, \theta]$ باشد، به‌دست می‌آوریم

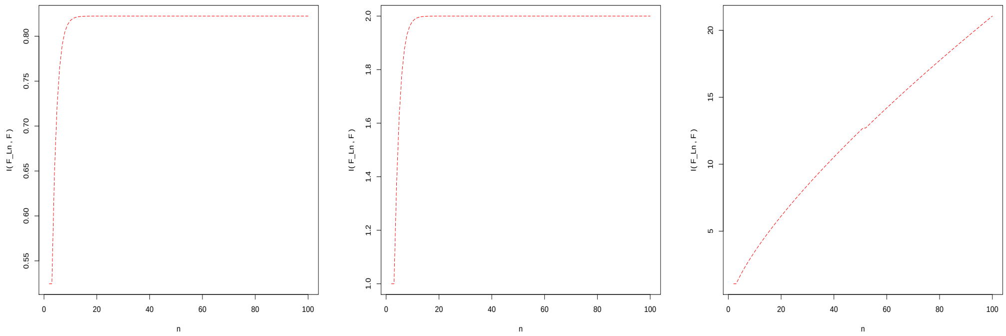
$$I(F_{L_n}, F) = \theta\sum_{j=0}^{n-1}(j+1)\left(\frac{1}{2}\right)^{j+1}.$$

از شکل ۴.۲ مشخص است که $I(F_{L_n}, F)$ برای توزیع یکنواخت استاندارد تابعی صعودی از n است و $\lim_{n\to\infty} I(F_{L_n}, F) = 2\theta$.

۳. اگر X دارای توزیع نمایی با میانگین $\frac{1}{\lambda}$ باشد، آن‌گاه

$$I(F_{L_n}, F) = \frac{1}{\lambda}\sum_{j=0}^{n-1}\sum_{k=0}^{\infty}(j+1)\left[\frac{1}{k+2}\right]^{j+2}.$$

از شکل ۴.۲ واضح است که $I(F_{L_n}, F)$ برای توزیع نمایی با میانگین $\frac{1}{2}$ تابعی صعودی از n است و $\lim_{n\to\infty} I(F_{L_n}, F) = \frac{1/644}{\lambda}$.



شکل ۴.۲: تابع $I(F_{L_n},F)$ بر حسب n در توزیع وایبل معکوس برای $\alpha=\beta=۲$ (راست)، توزیع یکنواخت بر روی $(۰,۱)$ (وسط) و توزیع نمایی با $\lambda=۲$ (چپ)

گزاره ۱.۵.۲. فرض کنید X یک متغیر تصادفی نامنفی با تابع توزیع F باشد. آن‌گاه

$$I(F_{L_n},F)=\int_۰^\infty \sum_{j=۰}^{n-۱}(j+۱)[F_{L_{j+۲}}(x)-F_{L_{j+۱}}(x)]dx.$$

برهان. با توجه به (۲۹.۲) اثبات به‌دست می‌آید. □

گزاره ۲.۵.۲. فرض کنید X یک متغیر تصادفی نامنفی با تابع توزیع F باشد. آن‌گاه

$$I(F_{L_n},F)=\sum_{j=۰}^{n-۱}\frac{۱}{j!}\int_۰^\infty \tilde{\lambda}(z)\left[\int_۰^z [-\log F(x)]^j\, F(x)dx\right]dz. \qquad (۳۰.۲)$$

برهان. با توجه به (۲۹.۲) و رابطه‌ی $-\log F(x)=\int_x^\infty \tilde{\lambda}(z)dz$ داریم

$$\begin{aligned}
I(F_{L_n},F)&=\sum_{j=۰}^{n-۱}\int_۰^\infty \frac{[-\log F(x)]^{j+۱}}{j!}F(x)dx\\
&=\sum_{j=۰}^{n-۱}\int_۰^\infty \left[\int_x^\infty \tilde{\lambda}(z)dz\right]\frac{[-\log F(x)]^j}{j!}F(x)dx\\
&=\sum_{j=۰}^{n-۱}\frac{۱}{j!}\int_۰^\infty \tilde{\lambda}(z)\left[\int_۰^z [-\log F(x)]^j\, F(x)dx\right]dz.
\end{aligned}$$

بنابراین اثبات کامل می‌شود. □

گزاره ۳.۵.۲. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته برای تمام $n\geq ۱$، با $I(F_{L_n},F)<\infty$ باشد. آن‌گاه

$$I(F_{L_n},F)=\sum_{j=۰}^{n-۱}\frac{۱}{j!}E\left(\tilde{h}_{j+۱}(T)\right), \qquad (۳۱.۲)$$

که در آن $\tilde{h}_{j+۱}(t)=\int_t^\infty [-\log F(x)]^{j+۱}\,dx$.

برهان. با توجه به (۲۹.۲) و قضیه‌ی فوبینی به‌دست می‌آوریم

$$\begin{aligned} I(F_{L_n}, F) &= \sum_{j=0}^{n-1} \int_0^\infty \frac{[-\log F(x)]^{j+1}}{j!} \left[\int_0^x f(t)dt\right] dx \\ &= \sum_{j=0}^{n-1} \int_0^\infty \frac{f(t)}{j!} \left[\int_t^\infty [-\log F(x)]^{j+1}\, dx\right] dt \\ &= \sum_{j=0}^{n-1} \frac{1}{j!} E\left[\tilde{h}_{j+1}(T)\right]. \end{aligned}$$

□

ملاحظه ۱.۵.۲. فرض کنید X یک متغیر تصادفی متقارن نسبت به امید متناهی $\mu = E(X)$ باشد، به این معنی که برای تمام $x \in \mathbb{R}$، $F(x+\mu) = 1 - F(\mu - x)$. آن‌گاه

$$I(F_{L_n}, F) = I(\bar{F}_{R_n}, \bar{F}),$$

که $I(\bar{F}_{R_n}, \bar{F})$ اندازه عدم دقت باقیمانده‌ی تجمعی بین $\bar{F}_{R_n}$ (تابع بقای nـامین رکورد بالا R_n) و $\bar{F}$ است.

کایال (۲۰۱۶) میانگین زمان غیرفعال را برای رکورد پایین به‌صورت زیر تعریف کرد

$$\mu_n(t) = E[t - L_n \mid L_n \le t] = \frac{\int_0^t F_{L_n}(x)dx}{F_{L_n}(t)} = \frac{\sum_{j=0}^{n-1} \int_0^t \frac{F(x)[-\log F(x)]^j}{j!} dx}{\sum_{j=0}^{n-1} \frac{F(t)[-\log F(t)]^j}{j!}}.$$

توجه کنید که $\mu_1(t) = \frac{\int_0^t F(x)dx}{F(t)}$. حال ما یک ارتباط را بین $\mu_n(t)$ و $I(F_{L_n}, F)$ در نظر می‌گیریم.

گزاره ۴.۵.۲. فرض کنید X یک متغیر تصادفی نامنفی با تابع توزیع F باشد. آن‌گاه

$$I(F_{L_n}, F) = \sum_{j=0}^{n-1} E[\mu_n(X_{j+1})].$$

برهان. با استفاده از (۳۰.۲) داریم

$$\begin{aligned} I(F_{L_n}, F) &= \sum_{j=0}^{n-1} \frac{1}{j!} \int_0^\infty \tilde{\lambda}(z) \left[\int_0^z [-\log F(x)]^j\, F(x)dx\right] dz \\ &= \int_0^\infty \tilde{\lambda}(z) \sum_{j=0}^{n-1} \frac{1}{j!} \left[\int_0^z [-\log F(x)]^j\, F(x)dx\right] dz \\ &= \int_0^\infty \tilde{\lambda}(z)\mu_n(z) \left[\sum_{j=0}^{n-1} \frac{1}{j!} [-\log F(z)]^j\, F(z)\right] dz \\ &= \int_0^\infty \mu_n(z) \left[\sum_{j=0}^{n-1} \frac{1}{j!} f(z)\, [-\log F(z)]^j\right] dz \\ &= \sum_{j=0}^{n-1} \int_0^\infty \mu_n(z) f_{L_{j+1}}(z) = \sum_{j=0}^{n-1} E[\mu_n(X_{j+1})]. \end{aligned}$$

بنابراین نتیجه‌ی مورد نظر به‌دست می‌آید. □

گزاره ۵.۵.۲. فرض کنید X یک متغیر تصادفی نامنفی با تابع توزیع F باشد. آن‌گاه

$$I(F_{L_n},F)=\sum_{j=0}^{n-1}\frac{1}{j!}\left[\sum_{i=0}^{j}\frac{1}{i!}E\left[(-\log F(X))^i\mu_{j+1}(X)\right]-\sum_{i=0}^{j-1}\frac{1}{i!}E\left[(-\log F(z))^i\mu_j(X)\right]\right].$$

برهان. با استفاده از (۶.۲) و (۳۰.۲) به‌دست می‌آوریم

$$\begin{aligned}
I(F_{L_n},F) &= \sum_{j=0}^{n-1}\frac{1}{j!}\int_0^\infty \tilde{\lambda}(z)\left[\int_0^z [F_{L_{j+1}}(x)-F_{L_j}(x)]dx\right]dz\\
&= \sum_{j=0}^{n-1}\frac{1}{j!}\int_0^\infty \tilde{\lambda}(z)\left[\mu_{j+1}(z)F_{L_{j+1}}(z)-\mu_j(z)F_{L_j}(z)\right]dz\\
&= \sum_{j=0}^{n-1}\frac{1}{j!}(\sum_{i=0}^{j}\frac{1}{i!}\int_0^\infty f(z)[-\log F(z)]^i\mu_{j+1}(z)dz\\
&- \sum_{i=0}^{j-1}\frac{1}{i!}\int_0^\infty f(z)[-\log F(z)]^i\mu_j(z)dz)\\
&= \sum_{j=0}^{n-1}\frac{1}{j!}\left[\sum_{i=0}^{j}\frac{1}{i!}E\left[(-\log F(X))^i\mu_{j+1}(X)\right]-\sum_{i=0}^{j-1}\frac{1}{i!}E\left[(-\log F(z))^i\mu_j(X)\right]\right].
\end{aligned}$$

درنتیجه اثبات کامل می‌شود. □

ملاحظه ۲.۵.۲. فرض کنید X یک متغیر تصادفی نامنفی با تابع توزیع F و X_{i+1}، $(i+1)$-امین رکورد پایین با تابع چگالی $f_{L_{i+1}}(x)$ باشد. آن‌گاه برای $n\geq 1$ داریم

$$I(F_{L_n},F)=\sum_{j=0}^{n-1}\sum_{i=0}^{j}\frac{1}{j!}\left[E\left[\mu_{j+1}(X_{i+1})\right]-E\left[\mu_j(X_{i+1})\right]\right].$$

برهان. از گزاره‌ی ۵.۵.۲ اثبات به‌دست می‌آید. □

ملاحظه ۳.۵.۲. مشابه (۲۹.۲)، اندازه عدم دقت گذشته‌ی تجمعی مرتبط با F و F_{L_n} به‌صورت زیر داده می‌شود

$$I(F,F_{L_n})=\mathcal{CE}(X)-E\left[U\log\left(\sum_{j=0}^{n-1}\frac{(-\log U)^j}{j!}\right)f(F^{-1}(U))\right],$$

که در آن $\mathcal{CE}(X)=-\int_0^\infty F(x)\log F(x)dx$ آنتروپی تجمعی است (دی کرشنزو و لونگوبردی (۲۰۰۹) را ببینید).

گزاره ۶.۵.۲. فرض کنید X یک متغیر تصادفی نامنفی با تابع توزیع F باشد، آن‌گاه یک بیان تحلیلی دیگر برای $I(F_{L_n},F)$ به‌صورت زیر به‌دست می‌آید

$$I(F_{L_n},F)=\sum_{j=0}^{n-1}\int_0^\infty\frac{[-\log F(x)]^{j+1}}{j!}F(x)dx=\sum_{j=0}^{n-1}(j+1)\mathcal{CE}_{j+1}(X),\qquad (۳۲.۲)$$

که

$$\mathcal{CE}_{j+1}(X)=\int_{0}^{\infty}\frac{[-\log F(x)]^{j+1}}{(j+1)!}F(x)dx, \tag{۳۳.۲}$$

آنتروپی تجمعی تعمیم‌یافته است (کایال[۱۲] (۲۰۱۶) را ببینید).

گزاره ۷.۵.۲. فرض کنید $a,b>0$. آن‌گاه برای $n=1,2,\ldots$ داریم

$$I(F_{aL_n+b},F_{aX+b})=aI(F_{L_n},F). \tag{۳۴.۲}$$

برهان. باتوجه به (۳۲.۲) داریم

$$I(F_{aL_n+b},F_{aX+b})=\sum_{j=0}^{n-1}(j+1)\mathcal{CE}_{j+1}(aX+b)=a\sum_{j=0}^{n-1}(j+1)\mathcal{CE}_{j+1}(X)=aI(F_{L_n},F). \tag{۳۵.۲}$$

اثبات کامل است. □

در ادامه کران بالا را برای $I(F,F_{L_n})$ به‌دست می‌آوریم.

گزاره ۸.۵.۲. فرض کنید X یک متغیر تصادفی نامنفی باشد که در $[0,a]$ مقدار می‌گیرد. آن‌گاه

$$I(F,F_{L_n})\leq[a-E(X)]\left|\log\left(1-\frac{E(L_n)}{a}\right)\right|$$

برهان. اثبات با استفاده از گزاره‌ی ۱.۹ در گاش و کوندو (۲۰۱۷) و با کمک نامساوی جمع لگاریتم به‌دست می‌آید. □

در گزاره‌های بعد برخی کران‌های پایین را برای $I(F_{L_n},F)$ نشان می‌دهیم.

گزاره ۹.۵.۲. اگر X یک متغیر تصادفی نامنفی و کاملاً پیوسته با میانگین $\mu=EX<\infty$ باشد، آن‌گاه برای $n\geq 1$ داریم

$$I(F_{L_n},F)\geq\sum_{j=0}^{n-1}\frac{\tilde{h}_{j+1}(\mu)}{j!}, \tag{۳۶.۲}$$

که تابع $\tilde{h}_{j+1}(.)$ در گزاره‌ی ۳.۵.۲ تعریف شده است.

برهان. با توجه به (۳۲.۲) داریم

$$\begin{aligned}I(F_{L_n},F)&=\sum_{j=0}^{n-1}(j+1)\mathcal{CE}_{j+1}(X)\\&=\sum_{j=0}^{n-1}\frac{E\left(\tilde{h}_{j+1}(X)\right)}{j!}.\end{aligned}$$

---
[۱۲] Kayal

از آنجا که $\tilde{h}_{j+1}(X)$ یک تابع محدب است، با استفاده از نامسای جنسن به‌دست می‌آوریم

$$I(F_{L_n}, F) \geq \sum_{j=0}^{n-1} \frac{\tilde{h}_{j+1}(\mu)}{j!}.$$

□

گزاره ۱۰.۵.۲. فرض کنید X یک متغیر تصادفی نامنفی با تابع توزیع F باشد. آن‌گاه

$$I(F_{L_n}, F) = \sum_{j=0}^{n-1} (j+1)\mathcal{CE}_{j+1}(X) \geq \sum_{j=0}^{n-1} \frac{[\mathcal{CE}(X)]^{j+1}}{j!}, \tag{۳۷.۲}$$

که $\mathcal{CE}(X)$ در ملاحظه‌ی ۳.۵.۲ داده شده است (برای جزییات بیشتر دی کرشنزو و لونگوبردی، ۲۰۰۹ را ببینید).

برهان. از آنجا که $(F(x))^n \leq F(x)$، برای تمام $n = 1, 2, \ldots$، داریم

$$\begin{aligned} I(F_{L_n}, F) &= \sum_{j=0}^{n-1} \int_0^\infty \frac{(-\log F(x))^{j+1}}{j!} F(x) dx \\ &\geq \sum_{j=0}^{n-1} \int_0^\infty \frac{(-\log F(x))^{j+1}}{j!} (F(x))^{j+1} \\ &= \sum_{j=0}^{n-1} \int_0^\infty \frac{[(-\log F(x))\, F(x)]^{j+1}}{j!} dx \\ &\geq \sum_{j=0}^{n-1} \frac{1}{j!} \left[\int_0^\infty (-\log F(x))\, F(x) dx \right]^{j+1}, \end{aligned}$$

سریعاً (۳۷.۲) حاصل می‌شود. □

ملاحظه ۴.۵.۲. فرض کنید X یک متغیر تصادفی نامنفی با تابع توزیع F باشد، آن‌گاه برای $n = 1, 2, \ldots$ داریم

$$I(F_{L_n}, F) \geq \sum_{j=0}^{n-1} \frac{1}{j!} \left[\int_0^\infty F(x)\bar{F}(x) dx \right]^{j+1}. \tag{۳۸.۲}$$

برهان. با استفاده از گزاره‌ی ۴.۳ از دی کرشنزو و لونگوبردی (۲۰۰۹) یک کران پایین برای $\mathcal{CE}(X)$ به‌صورت زیر است

$$\mathcal{CE}(X) \geq \int_0^\infty F(x)\bar{F}(x) dx.$$

حال، با استفاده از گزاره‌ی ۱۰.۵.۲ اثبات کامل می‌شود. □

گزاره ۱۱.۵.۲. برای یک متغیر تصادفی نامنفی و $n = 1, 2, \ldots$ داریم

$$I(F_{L_n}, F) \geq \sum_{j=0}^{n-1} \frac{1}{j!} [\mu.gini(X)]^{j+1}, \tag{۳۹.۲}$$

که $\mu = E(X)$ و $gini[.]$ شاخص جینی است.

برهان. با توجه به گزاره‌ی ۱.۵ از ونگ (۱۹۹۸) داریم

$$\begin{aligned}\int_{۰}^{\infty} F(x)\bar{F}(x)dx &= \frac{۱}{۲}E\left(|X-Y|\right)\\ &= E(X).gini[X],\end{aligned}$$

که X و Y مستقل و هم‌توزیع هستند. بنابراین، با توجه به رابطه‌ی (۳۸.۲) اثبات کامل می‌شود. □

نتیجه ۱.۵.۲. فرض کنید X یک متغیر تصادفی نامنفی با تابع بقای $\bar{F}(x)$ باشد. آن‌گاه داریم

$$I(F_{L_n}, F) \geq \sum_{j=۰}^{n-۱} \frac{۱}{j!} c^{j+۱} e^{(j+۱)H(X)}, \qquad (۴۰.۲)$$

که در آن $c = exp\left(\int_{۰}^{۱} \log(x|\log x|)dx\right) = ۰/۲۰۶۵$ و $H(X) = -\int_{۰}^{\infty} f(x)\log f(x)dx$ آنتروپی شانون X است.

برهان. با یادآوری (۳۷.۲) و گزاره‌ی ۲.۴ از دی کرشنزو و لونگوبردی (۲۰۰۹)، اثبات به‌دست می‌آید. □

حال با استفاده از برخی ویژگی‌های ترتیب تصادفی می‌توانیم خصوصیات مهمی را برای اندازه عدم دقت اثبات کنیم.

قضیه ۱.۵.۲. فرض کنید متغیر تصادفی نامنفی X DRHR است. آن‌گاه

$$I(F_{L_{n+۱}}, F) - I(F_{L_n}, F) \leq \sum_{i=۱}^{n+۱} E_{L_i}\left[\frac{۱}{\tilde{\lambda}(X)}\right]. \qquad (۴۱.۲)$$

برهان. فرض کنید $f_{L_n}(x)$ تابع چگالی nـامین رکورد پایین X_{L_n} باشد. آن‌گاه نسبت $\frac{f_{L_n}(x)}{f_{L_{n+۱}}(x)} = \frac{-n}{\log F(x)}$ در x صعودی است. بنابراین $X_{n+۱} \leq^{lr} X_n$، و این نتیجه می‌دهد که $X_{n+۱} \leq^{st} X_n$، یعنی $\bar{F}_{n+۱}(x) \leq \bar{F}_n(x)$ (برای جزییات بیشتر شیکد و شانتی کومار (۲۰۰۷، فصل ۱) را ببینید). این معادل است (شیکد و شانتی کومار (۲۰۰۷، صفحه ۴) را ببینید) با اینکه برای تمام توابع صعودی ϕ داشته باشیم

$$E(\phi(X_{n+۱})) \leq E(\phi(X_n)),$$

به شرطی که این امیدها موجود باشند. بنابراین، اگر X DRHR باشد و $\tilde{\lambda}(x)$ نرخ خطر

معکوس باشد، آن‌گاه $\frac{1}{\tilde{\lambda}(x)}$ در x صعودی است. به‌عنوان یک نتیجه، از (۲.۲۹) داریم

$$\begin{aligned} I(F_{L_{n+1}},F) &= \sum_{j=0}^{n}(j+1)E_{L_{j+2}}\left[\frac{1}{\tilde{\lambda}(X)}\right] \\ &\leq \sum_{j=0}^{n}(j+1)E_{L_{j+1}}\left[\frac{1}{\tilde{\lambda}(X)}\right] \\ &= \sum_{i=-1}^{n-1}(i+2)E_{L_{i+2}}\left[\frac{1}{\tilde{\lambda}(X)}\right] \\ &= \sum_{i=0}^{n-1}(i+2)E_{L_{i+2}}\left[\frac{1}{\tilde{\lambda}(X)}\right]+E_{L_1}\left[\frac{1}{\tilde{\lambda}(X)}\right] \\ &= I(F_{L_n},F)+\sum_{i=1}^{n+1}E_{L_i}\left[\frac{1}{\tilde{\lambda}(X)}\right]. \end{aligned} \quad (۲.۴۲)$$

بنابراین اثبات کامل می‌شود. □

قضیه ۲.۵.۲. فرض کنید X و Y دو متغیر تصادفی نامنفی باشند به‌طوری که $X \leq_{dcx} Y$. آن‌گاه داریم

$$I(F_{L_n},F) \leq I(G_{L_n},G).$$

برهان. از آنجا که $\tilde{h}_{j+1}(x)$ یک تابع محدب نزولی در x است، اثبات بلافاصله از (۲.۳۱) نتیجه می‌شود. □

گزاره ۲.۵.۱۲. فرض کنید X یک متغیر تصادفی نامنفی با تابع توزیع کاملاً پیوسته‌ی $F(x)$ باشد. آنگاه برای $n=1,2,\ldots$ داریم

$$I(F_{L_n},F) \geq \sum_{j=0}^{n-1}\sum_{i=0}^{j+1}\frac{(-1)^i(j+1)}{i!(j+1-i)!}\int_0^{\infty}[F(x)]^{i+1}dx.$$

برهان. از آنجا که $-\log F(x) \geq 1-F(x)$، با توجه به (۲.۲۹) اثبات به‌دست می‌آید. □

گزاره ۲.۵.۱۳. فرض کنید X یک متغیر تصادفی نامنفی با تابع توزیع کاملاً پیوسته‌ی $F(x)$ باشد. آنگاه برای $n=1,2,\ldots$ داریم

$$I(F_{L_n},F) \leq \sum_{j=0}^{n-1}\frac{1}{j!}\int_0^{\infty}[-\log F(x)]^{j+1}dx.$$

فرض کنید $\tilde{X}_\theta$ یک متغیر تصادفی کاملاً پیوسته و نامنفی با تابع توزیع $H_\theta(x)=[F(x)]^\theta$ و $x \geq 0$ را نشان دهد. حال اندازه عدم دقت تجمعی بین H_{L_n} و H_θ را به‌صورت زیر به‌دست می‌آوریم:

$$\begin{aligned} I(H_{L_n},H_\theta) &= -\int_0^{+\infty}H_{L_n}(x)\log\left(H_\theta(x)\right)dx \\ &= \sum_{j=0}^{n-1}\theta^{j+1}\int_0^{+\infty}\frac{[-\log F(x)]^{j+1}}{j!}[F(x)]^\theta dx. \end{aligned} \quad (۲.۴۳)$$

گزاره ۱۴.۵.۲. اگر $\theta \geq (\leq) ۱$، آن‌گاه برای هر $n = ۱, ۲, \ldots$ داریم

$$I(H_{L_n}, H_\theta) = \sum_{j=۰}^{n-۱}(j+۱)\mathcal{CE}_{j+۱}(\tilde{X}_\theta) \leq (\geq) \sum_{j=۰}^{n-۱}\theta^{j+۱}(j+۱)\mathcal{CE}_{j+۱}(X). \qquad (۴۴.۲)$$

برهان. فرض کنید که $\theta \geq (\leq) ۱$. آن‌گاه واضح است که $[F(x)]^\theta \leq (\geq) F(x)$ و بنابراین داریم

$$I(H_{L_n}, H_\theta) = \sum_{j=۰}^{n-۱}(j+۱)\mathcal{CE}_{j+۱}(\tilde{X}_\theta) \leq (\geq) \sum_{j=۰}^{n-۱}\theta^{j+۱}(j+۱)\mathcal{CE}_{j+۱}(X).$$

□

۶.۲ اندازه عدم دقت تجمعی پویا برای L_n

در تئوری قابلیت اطمینان اندازه‌های پویا برای توصیف محتوای اطلاعاتی که توسط طول عمرهای تصادفی مانند تغییرات سن بیان می‌شوند، کاربرد دارند. در این بخش ما نسخه‌ی پویای $I(F_{L_n}, F)$ را بررسی می‌کنیم. اگر یک سیستم که در زمان ۰ شروع به‌کار کرده فقط در زمان‌های بازرسی معین شده مشاهده گردد، و در زمان t از کار افتاده دیده شود، آن‌گاه اندازه عدم دقت پویا را به‌صورت زیر در نظر می‌گیریم

$$\begin{aligned} I(F_{L_n}, F; t) &= -\int_۰^t \frac{F_{L_n}(x)}{F_{L_n}(t)} \log\left(\frac{F(x)}{F(t)}\right) dx \\ &= \log F(t)\mu_n(t) - \int_۰^t \frac{F_{L_n}(x)}{F_{L_n}(t)} \log(F(x))\, dx \\ &= \log F(t)\mu_n(t) + \frac{۱}{F_{L_n}(t)} \sum_{j=۰}^{n-۱} \int_۰^t \frac{[-\log F(x)]^{j+۱}}{j!} F(x)dx. \qquad (۴۵.۲) \end{aligned}$$

دقت کنید که $\lim_{t\to\infty} I(F_{L_n}, F; t) = I(F_{L_n}, F)$. از آنجا که $\log F(t) \leq ۰$ برای $t \geq ۰$، داریم

$$\begin{aligned} I(F_{L_n}, F; t) &\leq \frac{۱}{F_{L_n}(t)} \sum_{j=۰}^{n-۱} \int_۰^t \frac{[-\log F(x)]^{j+۱}}{j!} F(x)dx \\ &\leq \frac{۱}{F_{L_n}(t)} \sum_{j=۰}^{n-۱} \int_۰^{+\infty} \frac{[-\log F(x)]^{j+۱}}{j!} F(x)dx = \frac{I(F_{L_n}, F)}{F_{L_n}(t)}. \end{aligned}$$

در قضیه‌ی بعد اثبات می‌کنیم که $I(F_{L_n}, F; t)$ به‌طور یکتا تابع توزیع را مشخصه‌سازی می‌کند.

قضیه ۱.۶.۲. فرض کنید X یک متغیر تصادفی پیوسته و نامنفی با تابع توزیع $F(.)$ است. فرض کنید عدم دقت تجمعی پویا برای nـامین رکورد با $I(F_{L_n}, F; t) < \infty$ و $t \geq ۰$ نشان داده شود. آن‌گاه $I(F_{L_n}, F; t)$ تابع توزیع را مشخصه‌سازی می‌کند.

برهان. از (۴۵.۲) داریم

$$I(F_{L_n},F;t)=\log F(t)\mu_n(t)+\frac{1}{F_{L_n}(t)}\sum_{j=0}^{n-1}\int_0^t\frac{[-\log F(x)]^{j+1}}{j!}F(x)dx. \qquad (۴۶.۲)$$

با مشتق‌گیری از دو طرف رابطه‌ی (۴۶.۲) نسبت به t به‌دست می‌آوریم

$$\begin{aligned}\frac{d}{dt}[I(F_{L_n},F;t)] &= \tilde{\lambda}_F(t)\mu_n(t)-\tilde{\lambda}_{F_{L_n}}(t)I(F_{L_n},F;t)\\ &= \tilde{\lambda}_F(t)\mu_n(t)-c(t)\tilde{\lambda}_F(t)I(F_{L_n},F;t)\\ &= \tilde{\lambda}_F(t)\left[\mu_n(t)-c(t)I(F_{R_n},F;t)\right].\end{aligned}$$

با مشتق‌گیری مجدد نسبت به t داریم

$$\acute{\tilde{\lambda}}_F(t)=\frac{(\tilde{\lambda}_F(t))^2\left[\acute{c}(t)I(F_{L_n},F;t)+c(t)\acute{I}(F_{L_n},F;t)-1+c(t)\tilde{\lambda}_F(t)\mu_n(t)\right]}{\acute{I}(F_{L_n},F;t)}. \qquad (۴۷.۲)$$

فرض کنید دو تابع F و F^* وجود دارند به‌طوری‌که

$$I(F_{L_n},F;t)=I(F^*_{L_n},F^*;t)=z(t).$$

آن‌گاه برای تمام مقادیر t، با توجه به (۴۷.۲) به‌دست می‌آوریم

$$\acute{\tilde{\lambda}}_F(t)=\varphi(t,\tilde{\lambda}_F(t)),\quad \acute{\tilde{\lambda}}_{F^*}(t)=\varphi(t,\tilde{\lambda}_{F^*}(t)),$$

که

$$\varphi(t,y)=\frac{y^2\left[\acute{c}(t)z(t)-c(t)\acute{z}(t)-1+c(t)yw(t)\right]}{\acute{z}(t)},$$

و $w(t)=\mu_n(t)$. با استفاده از قضیه‌ی ۱.۲ و لم ۲.۲ از گوپتا و کرمانی (۲۰۰۸) داریم $\tilde{\lambda}_F(t)=\tilde{\lambda}_{F^*}(t)$، برای هر t. از آنجا که نرخ خطر معکوس به‌طور یکتا تابع توزیع را مشخصه‌سازی می‌کند، اثبات کامل می‌شود. □

۷.۲ برآورد عدم دقت تجمعی برای L_n

در این بخش، موضوع برآورد کردن اندازه عدم دقت تجمعی را با استفاده از میانگین عدم دقت تجمعی تجربی در رکوردهای پایین بررسی می‌کنیم. فرض کنید $X_1,X_2,\dots,X_m$ یک نمونه‌ی تصادفی با حجم m از تابع توزیع تجمعی و مطلقاً پیوسته‌ی $F(x)$ باشد. آن‌گاه طبق رابطه‌ی (۳۲.۲)، عدم دقت تجمعی تجربی به‌صورت زیر تعریف می‌شود

$$\hat{I}(F_{L_n},F)=\sum_{j=0}^{n-1}\int_0^\infty\frac{\left[-\log\hat{F}_m(x)\right]^{j+1}}{j!}\hat{F}_m(x)dx=\sum_{j=0}^{n-1}(j+1)\mathcal{CE}_{j+1}(\hat{F}_m), \qquad (۴۸.۲)$$

که در آن

$$\hat{F}_m(x) = \frac{1}{m}\sum_{i=1}^{m} I_{(X_i \le x)}, \; x \in \mathbb{R}.$$

توزیع تجربی نمونه و I تابع نشانگر است. اگر $X_{(1)} \le X_{(2)} \le \ldots \le X_{(m)}$ آماره‌های ترتیبی نمونه باشند، آن‌گاه (۴۸.۲) را می‌توان به‌صورت زیر نوشت

$$\hat{I}(F_{L_n}, F) = \sum_{j=0}^{n-1}\sum_{k=1}^{m-1} \int_{X_{(k)}}^{X_{(k+1)}} \frac{\left[-\log \hat{F}_m(x)\right]^{j+1}}{j!} \hat{F}_m(x) dx. \qquad (۴۹.۲)$$

علاوه بر این،

$$\hat{F}_m(x) = \begin{cases} 0, & x < X_{(1)}, \\ \frac{k}{m}, & X_{(k)} \le x \le X_{(k+1)}, \quad k = 1, 2, \ldots, m-1 \\ 1, & x > X_{(k+1)}. \end{cases}$$

بنابراین، (۴۹.۲) می‌تواند به‌صورت زیر نوشته شود

$$\hat{I}(F_{L_n}, F) = \sum_{j=0}^{n-1}\sum_{k=1}^{m-1} \frac{1}{j!} U_{k+1} \frac{k}{m}\left(-\ln \frac{k}{m}\right)^{j+1}, \qquad (۵۰.۲)$$

که در آن $U_{k+1} = X_{(k+1)} - X_{(k)}, k = 1, 2, \ldots, m-1$.

مثال ۱.۷.۲. فرض کنید $X_1, X_2, \ldots, X_m$ یک نمونه‌ی تصادفی باشد که از توزیع نمایی با میانگین $\frac{1}{\lambda}$ گرفته شده باشد. آن‌گاه فاصله‌های نمونه‌ای U_{k+1} مستقل و دارای توزیع نمایی با میانگین $\frac{1}{\lambda(m-k)}$ هستند. حال از (۵۰.۲) به‌دست می‌آوریم

$$E[\hat{I}(F_{L_n}, F)] = \frac{1}{\lambda}\sum_{j=0}^{n-1}\sum_{k=1}^{m-1} \frac{1}{j!(m-k)}\left(\frac{k}{m}\right)\left(-\ln \frac{k}{m}\right)^{j+1} \qquad (۵۱.۲)$$

و

$$Var[\hat{I}(F_{L_n}, F)] = \frac{1}{\lambda^2}\sum_{j=0}^{n-1}\sum_{k=1}^{m-1} \frac{1}{(j!)^2(m-k)^2}\left(\frac{k}{m}\right)^2\left(-\ln \frac{k}{m}\right)^{2(j+1)}. \qquad (۵۲.۲)$$

ما مقادیر عددی $E[\hat{I}(F_{L_n}, F)]$ و $Var[\hat{I}(F_{L_n}, F)]$ را برای حجم‌های نمونه‌ی $m = 10, 15, 20$، $\lambda = 0/5, 1, 2$ و $n = 2, 3, 4, 5$ در جدول ۱.۲ حساب کردیم. به آسانی مشاهده می‌شود که $E[\hat{I}(F_{L_n}, F)]$ در m صعودی است. همچنین، $Var[\hat{I}(F_{L_n}, F)]$ در m نزولی می‌باشد.

مثال ۲.۷.۲. فرض کنید $X_1, X_2, \ldots, X_m$ یک نمونه‌ی تصادفی از جمعیتی با توزیع یکنواخت در $(0, 1)$ باشد. آن‌گاه متغیرهای U_{k+1} مستقل و دارای توزیع بتا با پارامتر ۱ و m می‌باشند (برای جزییات بیشتر پیک (۱۹۶۵) را ببینید). اکنون از (۵۰.۲) به‌دست می‌آوریم

$$E[\hat{I}(F_{L_n}, F)] = \sum_{j=0}^{n-1}\sum_{k=1}^{m-1} \frac{1}{j!(m+1)}\left(\frac{k}{m}\right)\left(-\ln \frac{k}{m}\right)^{j+1}, \qquad (۵۳.۲)$$

و

$$Var[\hat{I}(F_{L_n},F)] = \sum_{j=۰}^{n-۱}\sum_{k=۱}^{m-۱} \frac{۱}{(j!)^{۲}(m)(m+۲)} \left(\frac{k}{m}\right)^{۲} \left(-\ln\frac{k}{m}\right)^{۲(j+۱)}. \qquad (۵۴.۲)$$

ما مقادیر عددی $E[\hat{I}(F_{L_n},F)]$ و $Var[\hat{I}(F_{L_n},F)]$ را برای حجم‌های نمونه‌ی $m = ۱۰, ۱۵, ۲۰$ و $n = ۲, ۳, ۴, ۵$ در جدول ۲.۲ حساب کردیم. به آسانی می‌توان دید که $E[\hat{I}(F_{L_n},F)]$ در m و n صعودی است. همچنین، ملاحظه می‌شود که $\lim_{m\to\infty} Var[\hat{I}(F_{L_n},F)] = ۰$.

جدول ۱.۲: مقادیر عددی $E[\hat{I}(F_{L_n},F)]$ و $Var[\hat{I}(F_{L_n},F)]$ برای توزیع نمایی.

	$E[\hat{I}(F_{L_n},F)]$											
α	۰.۵	۱	۲	۰.۵	۱	۲	۰.۵	۱	۲	۰.۵	۱	۲
m	$n=۲$			$n=۳$			$n=۴$			$n=۵$		
۱۰	۱.۹۶	۰.۹۸۰	۰.۴۹۰	۲.۳۹۵	۱.۱۹۷	۰.۵۹۸	۲.۶۱۴	۱.۳۰۷	۰.۶۵۳	۲.۷۱۱	۱.۳۵۵	۰.۶۷۷
۱۵	۲.۰۱۱	۱.۰۰۵	۰.۵۰۲	۲.۴۷۱	۱.۲۳۵	۰.۶۱۷	۲.۷۱۶	۱.۳۵۸	۰.۶۷۹	۲.۸۳۴	۱.۴۱۷	۰.۷۰۸
۲۰	۲.۰۳۵	۱.۰۱۷	۰.۵۰۸	۲.۵۰۶	۱.۲۵۳	۰.۶۲۶	۲.۷۶۵	۱.۳۸۲	۰.۶۹۱	۲.۸۹۶	۱.۴۴۸	۰.۷۲۴

	$Var[\hat{I}(F_{L_n},F)]$											
α	۰.۵	۱	۲	۰.۵	۱	۲	۰.۵	۱	۲	۰.۵	۱	۲
m	$n=۲$			$n=۳$			$n=۴$			$n=۵$		
۱۰	۰.۲۵۲	۰.۰۶۳	۰.۰۱۵	۰.۲۹۱	۰.۰۷۲	۰.۰۱۸	۰.۳۰۶	۰.۰۷۶	۰.۰۱۹	۰.۳۱۰	۰.۰۷۷	۰.۰۱۹
۱۵	۰.۱۷۳	۰.۰۴۳	۰.۰۱۰	۰.۲۰۱	۰.۰۵۰	۰.۰۱۲	۰.۲۱۴	۰.۰۵۳	۰.۰۱۳	۰.۲۱۹	۰.۰۵۴	۰.۰۱۳
۲۰	۰.۱۳۱	۰.۰۳۲	۰.۰۰۸	۰.۱۵۳	۰.۰۳۸	۰.۰۰۹	۰.۱۶۴	۰.۰۴۱	۰.۰۱۰	۰.۱۶۸	۰.۰۴۲	۰.۰۱۰

جدول ۲.۲: مقادیر عددی $E[\hat{I}(F_{L_n},F)]$ و $Var[\hat{I}(F_{L_n},F)]$ برای توزیع یکنواخت.

	$E[\hat{I}(F_{L_n},F)]$				$Var[\hat{I}(F_{L_n},F)]$			
m	n=۲	n=۳	n=۴	n=۵	n=۲	n=۳	n=۴	n=۵
۱۰	۰.۴۳۷	۰.۵۸۱	۰.۶۶۰	۰.۶۹۷	۰.۰۱۴	۰.۰۱۹	۰.۰۲۱	۰.۰۲۲
۱۵	۰.۴۵۹	۰.۶۱۹	۰.۷۱۳	۰.۷۶۰	۰.۰۱۰	۰.۰۱۴	۰.۰۱۶	۰.۰۱۷
۲۰	۰.۴۷۰	۰.۶۳۷	۰.۷۳۹	۰.۷۹۴	۰.۰۰۷	۰.۰۱۱	۰.۰۱۳	۰.۰۱۴

فصل ۳

معیارهای عدم دقت در آمارههای ترتیبی

۱.۳ مقدمه

فرض کنید $X_1, X_2, ..., X_n$ مشاهدات مستقل و همتوزیع از تابع توزیع $F(X)$ و تابع چگالی $f(x)$ باشند. اگر مشاهدات $X_1, X_2, ..., X_n$ را از کوچکتر به بزرگتر مرتب کنیم آمارههای ترتیبی نمونه نامیده و با $X_{1:n}, X_{2:n}, ..., X_{n:n}$ نشان داده می‌شوند. این آماره‌ها در موضوعات مختلفی مانند تشخیص داده‌های پرت، مشخصه‌سازی تابع توزیع احتمال، کنترل کیفیت و آزمون مقاومت مواد کاربرد دارد. در مبحث قابلیت اطمینان آماره‌های ترتیبی برای مدل‌سازی آماری بکار برده می‌شوند. kـامین آماره ترتیبی از نمونه‌ای به حجم n طول عمر یک سیستم $(n-k+1)$ از n را نشان می‌دهد.

نویسندگان متعددی در مورد خواص نظری نظریه اطلاعات یک داده‌ی ترتیبی مطالعه کرده‌اند. وانگ و چن[1] (۱۹۹۰) نشان دادند که تفاوت بین میانگین آنتروپی آماره‌های ترتیبی و آنتروپی توزیع اولیه یک مقدار ثابت است. پارک (۱۹۹۵) برخی روابط بازگشتی را برای آنتروپی آماره‌های ترتیبی بدست آورد. ابراهیمی و همکاران (۲۰۰۴) برخی ویژگی‌های آنتروپی شانون

[1] Wong and Chen

آماره‌های ترتیبی را بدست آوردند و نشان دادند که توابع اطلاع کولبک_لایبلر مربوط به آماره‌های ترتیبی آزاد توزیع می‌باشند. در این فصل اندازه عدم دقت را برای آماره‌های ترتیبی بیان می‌کنیم و برخی ویژگی‌های آن را بررسی می‌نماییم.

اندازه عدم حتمیت شانون مربوط به i–امین آماره ترتیبی X_i به‌صورت زیر است

$$H(X_{i:n}) = -\int_{۰}^{\infty} f_{i:n}(x)\log f_{i:n}(x)dx, \tag{۱.۳}$$

که در آن

$$f_{i:n}(x) = \frac{۱}{B(i,n-i+۱)}(F(X))^{i-۱}(۱-F(X))^{n-i}f(x) \tag{۲.۳}$$

تابع چگالی i–امین آماره ترتیبی برای $i = ۱,۲,...,n$ است. در اینجا

$$B(a,b) = \int_{۰}^{۱} x^{a-۱}(۱-x)^{b-۱}dx, \quad a>۰,\ b>۰, \tag{۳.۳}$$

تابع بتا با پارامترهای a و b می‌باشد.

توجه کنید که به ازای $n=۱$ رابطه‌ی (۱.۳) به آنتروپی شانون X تغییر می‌کند. با استفاده از تبدیل $U=F(X)$، بطوریکه U دارای توزیع یکنواخت استاندارد است، آنتروپی i–امین آماره‌ی ترتیبی به‌صورت زیر است

$$H(X_{i:n}) = H_n(W_i) - E_{g_i}[\log(f(F^{-۱}(W_i)))], \tag{۴.۳}$$

که در آن

$$H_n(W_i) = B(i,n-i+۱) - (i-۱)[\psi(i)-\psi(n+۱)] - (n-i)[\psi(n-i+۱)-\psi(n+۱)], \tag{۵.۳}$$

آنتروپی i–امین آماره ترتیبی از توزیع یکنواخت استاندارد را نشان می‌دهد که تابع چگالی آن به‌صورت زیر است

$$g_i(w) = \frac{۱}{B(i,n-i+۱)}w^{i-۱}(۱-w)^{n-i}, \quad ۰<w<۱, \tag{۶.۳}$$

و $\psi(z) = \frac{d\log\Gamma(z)}{dz}$ تابع دی‌گاما است.

دی‌کرشنزو و لونگوبردی (۲۰۰۶) آنتروپی وزنی را معرفی کردند که به‌صورت زیر است

$$H^w(X) = -\int_{۰}^{+\infty} xf(x)\log f(x)dx. \tag{۷.۳}$$

اخیراً اندازه اطلاع جدیدی مطرح شده است که با جایگذاری تابع بقا $\bar{F}=۱-F$ به جای تابع چگالی در آنتروپی شانون به‌دست می‌آید. آنتروپی باقیمانده‌ی تجمعی توسط رائو و همکاران (۲۰۰۴) به‌صورت زیر معرفی شده است

$$\mathcal{E}(X) = \int_{۰}^{+\infty} \bar{F}(x)\Lambda(x)dx,$$

که در آن $\Lambda(x) = -\log \bar{F}(x)$. اندازه اطلاع جدیدی مشابه CRE توسط دی‌کرشنزو و لونگوبردی (۲۰۰۹) مطرح شده است که به‌صورت زیر است

$$\mathcal{CE}(X) = \int_{0}^{+\infty} F(x)\tilde{\Lambda}(x)dx, \tag{۸.۳}$$

که $\tilde{\Lambda}(x) = -\log F(x)$. در مقایسه با (۷.۳) میثاق و همکاران[۲] (۲۰۱۱) آنتروپی باقیمانده‌ی تجمعی وزنی (WCRE)[۳] را به‌صورت زیر معرفی کردند

$$\mathcal{E}^{w}(X) = \int_{0}^{+\infty} x\bar{F}(x)\Lambda(x)dx. \tag{۹.۳}$$

به‌طور مشابه، میثاق و همکاران (۲۰۱۱) آنتروپی تجمعی وزنی[۴] (WCE) را به‌صورت زیر مطرح کردند

$$\mathcal{CE}^{w}(X) = \int_{0}^{+\infty} xF(x)\tilde{\Lambda}(x)dx. \tag{۱۰.۳}$$

۲.۳ اندازه‌ی عدم دقت برای $X_{i:n}$

اندازه‌ی کولبک_لایبلر بین توزیع i_امین آماره‌ی ترتیبی و توزیع داده‌ها به‌صورت زیر است

$$K_n(f_{i:n}, f_X) = \int_{0}^{\infty} f_{i:n}(y)\log\left(\frac{f_{i:n}(y)}{f_X(y)}\right)dy \tag{۱۱.۳}$$

با استفاده از تبدیل انتگرالی احتمال $U = F(X)$، این رابطه به‌صورت زیر نوشته می‌شود

$$K_n(f_{i:n}, f_X) = K_n(\mathrm{g}_i, U) = \int_{0}^{\infty} \mathrm{g}_i(w)\log \mathrm{g}_i(w)dw = -H_n(W_i), \tag{۱۲.۳}$$

که $f_x(y)$ تابع چگالی متغیر اولیه X، $f_{i:n}$ تابع چگالی i_امین آماره‌ی ترتیبی، g_i تابع چگالی احتمال بتا (۶.۳) و U متغیر تصادفی یکنواخت است.

با جمع روابط (۱.۳) و (۱۱.۳) خواهیم داشت

$$\begin{aligned} H(X_{i:n}) + K_n(f_{i:n}, f_X) &= -\int_{0}^{\infty} f_{i:n}(y)\log f_{i:n}(y) + \int_{0}^{\infty} f_{i:n}(y)\log\left(\frac{f_{i:n}(y)}{f_X(y)}\right)dy \\ &= -\int_{0}^{\infty} f_{i:n}(y)\log f_X(y)dy. \end{aligned} \tag{۱۳.۳}$$

با استفاده از تبدیل $U = F(X)$، رابطه (۱۳.۳) به $-E_{\mathrm{g}_i}[\log(f(F^{-1}(W_i)))]$ تبدیل می‌شود. به‌علاوه با جمع روابط (۴.۳) و (۱۲.۳) به‌دست می‌آید

$$H(X_{i:n}) + K_n(f_{i:n}, f_X) = -E_{\mathrm{g}_i}[\log(f(F^{-1}(W_i)))], \tag{۱۴.۳}$$

[۲] Misagh et al.
[۳] Weighted cumulative residual entropy
[۴] Weighted cumulative entropy

که تاییدی بر نتایجی است که قبلاً به‌دست آمده‌اند.
مشابه اندازه‌ی عدم دقت کریج بین دو تابع چگالی f و g که در (۱۴.۱) تعریف شد، اندازه‌ی

$$I_n(f_{i:n}, f) = -\int_0^\infty f_{i:n}(x)\log f(x)dx = -E_{g_i}[\log(f(F^{-1}(W_i)))] \qquad (۱۵.۳)$$

به‌عنوان اندازه‌ی عدم دقت مرتبط با توزیع i–امین آماره‌ی ترتیبی و تابع توزیع اولیه‌ی $f(x)$ تعریف می‌شود.

۳.۳ خصوصیات اندازه عدم دقت برای $X_{i:n}$

در این بخش کران‌هایی را برای اندازه عدم دقت در آماره‌های ترتیبی بر حسب آنتروپی به‌دست می‌آوریم. همچنین مقدار میانگین این اندازه را نشان می‌دهیم.

قضیه ۱.۳.۳. برای هر متغیر تصادفی X با آنتروپی $H(X) < \infty$

(i) اگر B_i، i_امین جمله از $B(n-1, p_i)$ و $p_i = \frac{i-1}{n-1}$ باشد، آن‌گاه

$$nB_i(H(X)+I(A)) \le I_n(f_{i:n}, f) \le nB_i[H(X)+I(\bar{A})] \qquad (۱۶.۳)$$

که $I(A) = \int_A f(x)\log f(x)dx$ و $A = \{x; f(x) \le 1\}$ و $\bar{A} = \{x; f(x) > 1\}$.

(ii) اگر $M = f(m) < \infty$ که m مد توزیع می‌باشد، آن‌گاه

$$-\log M \le I_n(f_{i:n}, f) \le nB_i[H(X)+\log M] - \log M. \qquad (۱۷.۳)$$

برهان. ابراهیمی و همکاران (۲۰۰۴) برای $H(X_{i:n})$ کران‌ها زیر را بدست آوردند

$$H_n(W_i) + nB_i(H(X)+I(A)) \le H(X_{i:n}) \le H_n(W_i) + nB_i[H(X)+I(\bar{A})] \qquad (۱۸.۳)$$

که $H_n(W_i)$ در رابطه (۵.۳) تعریف شده است. با جمع روابط (۱۲.۳) و (۱۸.۳) رابطه‌ی (۱۶.۳) به‌دست می‌آید.
برای اثبات (ii) از نتیجه‌ی ارایه شده در ابراهیمی و همکاران (۲۰۰۴) استفاده می‌کنیم

$$H_n(W_i) - \log M \le H(X_{i:n}) \le H_n(W_i) - \log M + nB_i[H(X)+\log M]. \qquad (۱۹.۳)$$

با جمع (۱۲.۳) و (۱۹.۳)، (۱۷.۳) به‌دست خواهد آمد. □

مثال ۱.۳.۳. فرض کنید X یک متغیر تصادفی از توزیع نمایی با تابع چگالی $f(x) = \theta e^{-\theta x}$ و $\theta > 0, x \ge 0$ باشد. آن‌گاه $F(x) = 1 - e^{-\theta x}$.
برای $i = 1$، که حالت مینیمم نمونه است، داریم

$$I_n(f_{1:n}, f) = -E_{g_1}[\log(f(F^{-1}(W_1)))] = \frac{1}{n} - \log\theta. \qquad (۲۰.۳)$$

توجه کنید که

(i) برای یک مقدار ثابت n، با افزایش مقدار θ عدم دقت مینیمم نمونه برای توزیع نمایی کاهش می‌یابد.

(ii) بطور مشابه، اگر مقدار θ را ثابت در نظر بگیریم، با افزایش حجم نمونه عدم دقت کاهش می‌یابد.

برای $i=n$ که حالت ماکزیمم نمونه است

$$I_n(f_{n:n},f)=-E_{g_n}[\log(f(F^{-1}(W_n)))]=\gamma+\psi(n)-\log\theta+\frac{1}{n}. \tag{۲۱.۳}$$

که $\psi(1)=-\gamma=0/5772$ ثابت اویلر است و $\psi(n+1)=\psi(n)+\frac{1}{n}$.
در این حالت

(i) برای مقدار ثابت n، با افزایش مقدار پارامتر θ عدم دقت ماکزیمم نمونه کاهش می‌یابد.

(ii) $I_n(f_{n:n},f)-I_n(f_{1:n},f)=\gamma+\psi(n)\geq 0$، تساوی زمانی برقرار می‌شود که $n=1$. بنابراین نتیجه می‌گیریم برای توزیع نمایی عدم دقت مربوط به ماکزیمم همیشه بیشتر از مینیمم است.

ملاحظه ۱.۳.۳. برای توزیع نمایی با میانگین $\frac{1}{\theta}$ داریم $M=\theta$ و $H(X)=1-\log\theta$. با استفاده از (۱۷.۳) داریم

$$-\log\theta\leq I_n(f_{i:n},f)\leq nB_i-\log\theta. \tag{۲۲.۳}$$

برای $i=1$، رابطه‌ی (۲۲.۳) تبدیل می‌شود به

$$-\log\theta\leq I_n(f_{1:n},f)\leq n-\log\theta. \tag{۲۳.۳}$$

در حالی‌که

$$I_n(f_{1:n},f)=\frac{1}{n}-\log\theta. \tag{۲۴.۳}$$

اختلاف بین مقدار واقعی $I_n(f_{1:n},f)$ و کران پایین در رابطه‌ی (۲۳.۳) برابر است با $\frac{1}{n}$ که وقتی n به سمت بینهایت میل کند به سمت صفر می‌رود. بنابراین برای توزیع نمایی وقتی حجم نمونه بزرگ باشد، کران پایین سودمند است.

قضیه ۲.۳.۳. میانگین مقدار اندازه عدم دقت بین دو تابع چگالی $f_{i:n}$ و f برابر است با آنتروپی متغیر تصادفی اولیه X، به این معنا که

$$\frac{1}{n}\sum_{i=1}^{n}I_n(f_{i:n},f)=H(X). \tag{۲۵.۳}$$

برهان. در نظر بگیرید

$$\begin{aligned}-\sum_{i=1}^{n}\int f_{i:n}(y)\log f(y)dy &= -\sum_{i=1}^{n}\int \frac{1}{B(i,n-i+1)}(F(y))^{n-1}(1-F(y))^{n-i}f(y)\log f(y)dy\\ &= -\sum_{i=1}^{n}\int g_i(F(y))f(y)\log f(y)dy\\ &= -\int\sum_{i=1}^{n} nq_{i-1}f(y)\log f(y)dy\\ &= nH(X),\end{aligned}$$

که

$$g_i(w)=\frac{1}{B(i,n-i+1)}w^{i-1}(1-w)^{n-i},\quad 0\le w\le 1,$$

تابع چگالی iـامین آماره‌ی ترتیبی از توزیع یکنواخت استاندارد است و q_{i-1} نشان‌دهنده‌ی $(i-1)$ـامین جمله از $B(n-1,p)$ است، که $p=F(x)$ و $\sum_{i=1}^{n} q_{i-1}=1$. بنابراین نتیجه‌ی مورد نظر در (۲۵.۳) به‌دست می‌آید. □

مثال ۲.۳.۳. فرض کنید X یک متغیر تصادفی از توزیع نمایی با تابع چگالی $f(x)=\theta e^{-\theta x}$ و $x\ge 0, \theta>0$ باشد. آن‌گاه

$$f_{i;n}(y)=\frac{1}{B(i,n-i+1)}F(y)^{i-1}(1-F(y))^{n-i}f(y). \tag{۲۶.۳}$$

برای $i=1,2$ و $n=2$، با استفاده از (۱۵.۳) داریم

$$I_2(f_{1:2},f)=-\log\theta-\frac{1}{2}$$

و

$$I_2(f_{2:2},f)=-\log\theta+\frac{3}{2}.$$

بنابراین

$$\frac{1}{2}(I_2(f_{1:2},f)+I_2(f_{2:2},f))=1-\log\theta. \tag{۲۷.۳}$$

همچنین با استفاده از (۱۱.۲) داریم

$$H(X)=1-\log\theta$$

که برابر است با میانگین اندازه عدم دقت که در (۲۷.۳) محاسبه شد.

۴.۳ خصوصیات اندازه عدم دقت باقیمانده برای $X_{i:n}$

گزاره ۱.۴.۳. فرض کنید $m = \sup\{x; f(x) \leq M\}$ مد توزیع است به‌طوری که $M < \infty$. آن‌گاه

$$I_n(f_{i:n}, f; t) \geq \log\left(\bar{F}(t)\right) - \log M = \log\left(\frac{\bar{F}(t)}{M}\right).$$

برهان. از (۲۳.۱) داریم

$$\begin{aligned} I_n(f_{i:n}, f; t) &= -\int_t^{\infty} \frac{f_{i:n}(x)}{\bar{F}_{i:n}(t)} \log\left(\frac{f(x)}{\bar{F}(t)}\right) dx \\ &= \log\left(\bar{F}(t)\right) - \frac{۱}{\bar{F}_{i:n}(t)} \int_t^{\infty} f_{i:n}(x) \log f(x) dx. \end{aligned}$$

از آنجا که m مد توزیع است، بنابراین $\log f(x) \leq \log M$. با استفاده از این حقیقت در رابطه‌ی بالا داریم

$$I_n(f_{i:n}, f; t) \geq \log\left(\bar{F}(t)\right) - \log M.$$

□

قضیه ۱.۴.۳. فرض کنید $X_۱, X_۲, ..., X_n$ متغیرهای تصادفی مستقل و همتوزیع باشند که برای $i = ۱$ طول عمر یک سیستم سری را نشان می‌دهند. فرض کنید F و f به‌ترتیب تابع توزیع و تابع چگالی را نشان دهند. اگر f در دامنه‌ی خود نزولی باشد آن‌گاه عدم دقت مربوط به آن تابعی نزولی از n است.

برهان. می‌دانیم که متغیر تصادفی $\{X_{i:n} | X_{i:n} > t\}$ دارای تابع چگالی زیر است

$$g_i(y) = \frac{۱}{\bar{B}_{F(t)}(i, n)} y^{i-۱}(۱-y)^{n-۱}, \ F(t) \leq y \leq ۱,$$

که $\bar{B}_{F(t)}(a, b) = \int_{F(t)}^{۱} x^{a-۱}(۱-x)^{b-۱} dx$ تابع بتای ناقص است.

از آنجا که f برای یک سیستم سری (وقتی که $i = ۱$) در دامنه‌اش نزولی است، بنابراین

$$\frac{g_{n+۱}(x)}{g_n(x)} = \frac{\bar{B}_{F(t)}(۱, n) y}{\bar{B}_{F(t)}(۱, n+۱)}, \ F(t) \leq y \leq ۱$$

تابعی نزولی است. این به این معنی است که $X_{n+۱} \leq^{lr} X_n$ که نتیجه می‌دهد $X_{n+۱} \leq^{st} X_n$.
همچنین می‌دانیم که $f(F^{-۱}(x))$ تابعی نزولی از x است. بنابراین، برای $i = ۱$ داریم

$$E_{g_۱}\left[\log\left(f\left(F^{-۱}(U_n)\right)\right)\right] \leq E_{g_۱}\left[\log\left(f\left(F^{-۱}(U_{n+۱})\right)\right)\right].$$

همچنین از (۲۳.۱) عدم دقت باقیمانده‌ی iـامین آماره ترتیبی با استفاده از تبدیل $U = F(X)$، برابر می‌شود با

$$\begin{aligned} I_n(f_{i:n}, f; t) &= \log\left(\bar{F}(t)\right) - \frac{1}{\bar{F}_{i:n}(t)} \int_t^\infty f_{i:n}(x) \log f(x) dx \\ &= \log\left(\bar{F}(t)\right) - \int_{F(t)}^1 \frac{u^{i-1}(1-u)^{n-i} \log(f(F^{-1}(u)))du}{\bar{B}_{F(t)}(1, n)} \\ &= -E_{g_i}\left[\log\left(f\left(F^{-1}(U_n)\right)\right)\right] + \log\left(\bar{F}(t)\right). \end{aligned}$$

بنابراین برای $i = 1$ و $n \geq 1$ داریم

$$\begin{aligned} I_n(f_{1:n}, f; t) - I_{n+1}(f_{1:n+1}, f; t) &= -E_{g_1}\left[\log\left(f\left(F^{-1}(U_n)\right)\right)\right] + \log\left(\bar{F}(t)\right) \\ &\quad + E_{g_1}\left[\log\left(f\left(F^{-1}(U_{n+1})\right)\right)\right] - \log\left(\bar{F}(t)\right) \geq 0. \end{aligned}$$

اثبات کامل می‌شود. □

۵.۳ نتایج مشخصه‌سازی

قضیه ۱.۵.۳. فرض کنید X یک متغیر تصادفی پیوسته و نامنفی با تابع توزیع $F(.)$ باشد. همچنین فرض کنید عدم دقت باقیمانده‌ی پویا مربوط به iـامین آماره ترتیبی بر اساس یک نمونه‌ی تصادفی به حجم n را با $I_n(f_{i:n}, f; t) < \infty$ و $t \geq 0$ نشان دهیم. آن‌گاه $I_n(f_{i:n}, f; t)$ تابع توزیع را مشخصه‌سازی می‌کند.

برهان. می‌دانیم که

$$\begin{aligned} I_n(f_{i:n}, f; t) &= -\int_t^\infty \frac{f_{i:n}(x)}{\bar{F}_{i:n}(t)} \log\left(\frac{f(x)}{\bar{F}(t)}\right) dx \\ &= \log\left(\bar{F}(t)\right) - \frac{1}{\bar{F}_{i:n}(t)} \int_t^\infty f_{i:n}(x) \log f(x) dx. \end{aligned}$$

با مشتق‌گیری از هر دو طرف نسبت به t داریم

$$\frac{d}{dt}\left[I_n(f_{i:n}, f; t)\right] = -\lambda_F(t) + \lambda_{F_{i:n}}(t)\left(I_n(f_{i:n}, f; t) + \log\left(\lambda_F(t)\right)\right),$$

که $\lambda_F(t)$ و $\lambda_{F_{i:n}}(t)$ به‌ترتیب نرخ خطر X و $X_{i:n}$ هستند.

با مشتق‌گیری دوباره نسبت به t و استفاده از رابطه‌ی

$$\lambda_{F_{i:n}}(t) = c(t)\lambda_F(t)$$

که

$$c(t) = \left[\frac{(F(t))^{i-1}(1-F(t))^{n-i+1}}{\bar{B}_{F(t)}(i, n-i+1)}\right] \lambda_F(t),$$

به‌دست می‌آوریم

(۲۸.۳)
$$\acute{\lambda}_F(t) = \left(\frac{\lambda_F(t)\left[c(t)\acute{I}_n(f_{i:n}, f; t)\right] - \left[\acute{c}(t)\acute{I}_n(f_{i:n}, f; t) + \acute{c}(t)\lambda_F(t) + c^{2}(t)\lambda_F(t)\acute{I}_n(f_{i:n}, f; t)\right]}{c(t)\left[\lambda_F(t) + \acute{I}_n(f_{i:n}, f; t)\right]} \right).$$

فرض کنید که دو تابع F و F^* وجود دارند به‌طوری‌که

$$I_n(f_{i:n}, f; t) = I_n(f^*_{i:n}, f; t) = h(t).$$

آن‌گاه برای هر t، از (۲۸.۳) به‌دست می‌آوریم

$$\acute{\lambda}_F(t) = \psi(t, \lambda_F(t)), \quad \acute{\lambda}^*_F(t) = \psi(t, \lambda^*_F(t)),$$

که

$$\psi(t, y) = \left(\frac{y\left[c(t)h''(t)\right] - \left[\acute{c}(t)\acute{h}(t) + \acute{c}(t)y + c''(t)y\acute{h}(t)\right]}{c(t)\left[yc(t) + \acute{h}(t)\right]} \right).$$

با استفاده از لم‌های ۱.۳.۲ و ۲.۳.۲ داریم $\lambda_F(t) = \lambda^*_F(t)$ برای هر t. با دانستن این که تابع نرخ خطر تابع توزیع را به‌طور یکتا مشخصه‌سازی می‌کند، نتیجه‌ی مورد نظر حاصل می‌شود. □

در ادامه با در نظر گرفتن یک رابطه بین عدم دقت باقیمانده‌ی پویای اولین آماره ترتیبی و تابع نرخ خطر، برخی توابع طول عمر خاص را مشخصه‌سازی می‌کنیم.

قضیه ۲.۵.۳. فرض کنید X یک متغیر تصادفی پیوسته و نامنفی با تابع توزیع $F(.)$ باشد. عدم دقت باقیمانده‌ی پویا مربوط به آماره ترتیبی اول بر اساس یک نمونه تصادفی به حجم n را با $I_n(f_{1:n}, f; t) < \infty,\ t \geq 0$ نشان دهیم. فرض کنید $\lambda_F(t)$ تابع نرخ خطر X باشد و

$$I_n(f_{1:n}, f; t) = c - \log \lambda_F(t), \qquad (۲۹.۳)$$

که c مقداری ثابت است. آن‌گاه X دارای

۱. توزیع نمایی است اگر و تنها اگر $c = \frac{1}{n}$،

۲. توزیع پارتو است اگر و تنها اگر $c > \frac{1}{n}$،

۳. یک توزیع با دامنه متناهی است اگر و تنها اگر $c < \frac{1}{n}$.

برهان. فرض می‌کنیم که

$$I_n(f_{1:n}, f; t) = c - \log \lambda_F(t).$$

با مشتق‌گیری نسبت به t از دو طرف رابطه‌ی بالا داریم

$$\frac{d}{dt}[I_n(f_{1:n}, f; t)] = -\lambda_F(t) + \lambda_{F_{1:n}}(t)\left(I_n(f_{1:n}, f; t) + \log \lambda_F(t)\right), \qquad (۳۰.۳)$$

که در آن $\lambda_F(t)$ و $\lambda_{F_{1:n}}(t)$ به‌ترتیب نرخ خطر X و $X_{1:n}$ هستند. به آسانی می‌توان دید که $\lambda_{F_{1:n}}(t) = n\lambda_F(t)$. با استفاده از $I_n(f_{1:n}, f; t) = c - \log \lambda_F(t)$ و قرار دادن مقدار $\lambda_{F_{1:n}}(t)$ در (۳۰.۳) داریم

$$-\acute{\lambda}_F(t) = (nc-1)\lambda_F^2(t).$$

جواب این معادله‌ی دیفرانسیلی برابر می‌شود با

$$\lambda_F(t) = \frac{1}{at+b}, \tag{۳۱.۳}$$

که $a = (nc-1)$ و $b = \lambda_F 0$.

۱. اگر $c = \frac{1}{n}$، آن‌گاه $a = 0$ و از (۳۱.۳) معلوم می‌شود که $\lambda_F(t)$ مقداری ثابت است، که این در صورتی ممکن است که X دارای توزیع نمایی باشد.

۲. اگر $c > \frac{1}{n}$، آن‌گاه $a > 0$ و (۳۱.۳) تابع نرخ خطر توزیع پارتو می‌شود.

۳. اگر $c < \frac{1}{n}$، آن‌گاه $a < 0$ و (۳۱.۳) تابع نرخ خطر توزیع با دامنه متناهی می‌شود.

تنها اگر یکی از این حالت‌ها اثبات شود مابقی آسان است. □

۶.۳ عدم دقت باقیمانده‌ی تجمعی وزنی برای $X_{(1:n)}$

فرض کنید X و Y دو متغیر تصادفی نامنفی با توابع توزیع F و G باشند. آن‌گاه عدم دقت باقیمانده‌ی تجمعی وزنی[۵] (WCRI) بین $\bar{F}$ و $\bar{G}$ به‌صورت زیر است

$$I^w(\bar{F}, \bar{G}) = -\int_0^\infty x\bar{F}(x)\log \bar{G}(x)dx.$$

در این بخش ما عدم دقت باقیمانده‌ی تجمعی وزنی بین $\bar{F}_{X_{(1:n)}}$ و $\bar{F}$ را به‌صورت زیر معرفی می‌کنیم:

$$I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) = -\int_0^{+\infty} x\bar{F}_{X_{(1:n)}}(x)\Lambda(x)dx = \frac{1}{n}\mathcal{E}^w(X_{(1:n)}), \tag{۳۲.۳}$$

که

$$\mathcal{E}^w(X_{(1:n)}) = n\int_0^{+\infty} x[\bar{F}(x)]^n\Lambda(x)dx.$$

در ادامه برخی ویژگی‌های $I^w(\bar{F}_{X_{(1:n)}}, \bar{F})$ را بیان می‌کنیم.

گزاره ۱.۶.۳. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته با $I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) < \infty$ برای $n \geq 1$ باشد. آن‌گاه داریم

$$\begin{aligned} I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) &= \int_0^{+\infty} x[\bar{F}(x)]^n \left(\int_0^x \frac{f(z)}{\bar{F}(z)}dz\right) dx \\ &= \int_0^{+\infty}\int_z^{+\infty} \lambda(z)x[\bar{F}(x)]^n dxdz, \end{aligned} \tag{۳۳.۳}$$

[۵] Weighted cumulative residual inaccuracy

که $\lambda(.)$ تابع نرخ خطر است.

گزاره ۲.۶.۳. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته با $I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) < \infty$ برای $n \geq 1$ باشد. آن‌گاه داریم

$$I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) = E\left(M^w_{(1:n)}(Z)(\bar{F}(Z))^{n-1}\right), \tag{۳۴.۳}$$

که

$$M^w_{(1:n)}(z) = \frac{1}{(\bar{F}(z))^n}\int_z^{+\infty} x(\bar{F}(x))^n dx$$

میانگین وزنی طول عمر باقیمانده[۶] (WMRL) $X_{(1:n)}$ است.

برهان. از (۳۳.۳) داریم

$$\begin{aligned} I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) &= \int_0^{+\infty}\int_z^{+\infty} \lambda(z)x[\bar{F}(x)]^n dxdz \\ &= \int_0^{+\infty} \frac{f(z)}{\bar{F}(z)}dz\left[\int_z^{+\infty} x[\bar{F}(x)]^n dx\right] \\ &= \int_0^{+\infty} \frac{f(z)}{\bar{F}(z)}[\bar{F}(z)]^n M^w_{(1:n)}(z)dz = \int_0^{+\infty} f(z)[\bar{F}(z)]^{n-1}M^w_{(1:n)}(z)dz. \end{aligned} \tag{۳۵.۳}$$

بنابراین اثبات کامل است. □

گزاره ۳.۶.۳. فرض کنید $a, b > 0$. برای $n = 1, 2, \ldots$

$$I^w(\bar{F}_{aX_{(1:n)}+b}, \bar{F}_{aX+b}) = aI^w(\bar{F}_{X_{(1:n)}}, \bar{F}).$$

در قضایای بعد برخی کران‌های پایین و بالا را برای $I^w(\bar{F}_{X_{(1:n)}}, \bar{F})$ نشان می‌دهیم.

گزاره ۴.۶.۳. برای متغیر تصادفی نامنفی X و $n \geq 1$، رابطه‌ی زیر برقرار می‌شود

$$I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) \geq M^w_{(1:n)}(t)|\log \bar{F}(t)|[\bar{F}(t)]^n, \tag{۳۶.۳}$$

که $M^w_{(1:n)}(t)$ WMRL مربوط به $X_{(1:n)}$ است.

برهان. از براتپور (۲۰۱۰) اثبات به‌دست می‌آید. □

گزاره ۵.۶.۳. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته با $I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) < \infty$ و $n \geq 1$ باشد. آن‌گاه داریم

$$I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) \geq \int_0^{+\infty} x(\bar{F}(x))^n F(x)dx. \tag{۳۷.۳}$$

[۶] Weighted mean residual life

برهان. با یادآوری اینکه $-\log \bar{F}(x) \geq F(x)$ اثبات به‌دست می‌آید. □

گزاره ۳.۶.۶. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته با $I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) < \infty$ برای $n \geq 1$ باشد. آن‌گاه داریم

$$I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) \leq \mathcal{E}^w(X). \tag{۳.۳۸}$$

برهان. از آنجا که $\bar{F}(x) \geq [\bar{F}(x)]^n$ و $x \geq 0$، وقتی $n \geq 1$، اثبات به‌دست می‌آید. □

گزاره ۳.۶.۷. اگر X IFRA (DFRA) باشد، آن‌گاه

$$I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) \leq (\geq) \mathcal{E}\left(X^2 (\bar{F}(X))^{n-1}\right). \tag{۳.۳۹}$$

برهان. از آنجا که X IFRA (DFRA) است، $\frac{\Lambda(x)}{x}$ نسبت به $x > 0$ صعودی (نزولی) است که نتیجه می‌دهد

$$\bar{F}(x)\Lambda(x) \leq (\geq) x f(x), \quad x > 0. \tag{۳.۴۰}$$

با ضرب $x[\bar{F}(x)]^{n-1} \geq 0$ در (۳.۴۰) و سپس انتگرال‌گیری، نتیجه حاصل می‌شود. □

گزاره ۳.۶.۸. اگر X یک متغیر تصادفی نامنفی و کاملاً پیوسته با $I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) < \infty$ برای $n \geq 1$ باشد. آن‌گاه داریم

$$I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) = \mathbb{E}\left[h^w(X)\right], \tag{۳.۴۱}$$

که

$$h^w(x) = \int_0^x z\left[-\log \bar{F}(z)\right] [\bar{F}(z)]^{n-1} dz, \quad x \geq 0.$$

برهان. از (۳.۳۲) و با استفاده از قضیه فوبینی به‌دست می‌آوریم

$$\begin{aligned}
I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) &= \int_0^\infty z[-\log \bar{F}(z)][\bar{F}(z)]^{n-1}\bar{F}(z)dz \\
&= \int_0^\infty \left[\int_z^\infty f(x)dx\right] z[\bar{F}(z)]^{n-1}[-\log \bar{F}(z)]dz \\
&= \int_0^\infty f(x)\left[\int_0^x z[\bar{F}(z)]^{n-1}[-\log \bar{F}(z)]dz\right] dx = \mathbb{E}\left[h^w(X)\right].
\end{aligned}$$

□

گزاره ۳.۶.۹. فرض کنید X و Y دو متغیر تصادفی نامنفی با تابع قابلیت اعتماد $\bar{F}(x)$ و $\bar{G}(x)$ باشند. اگر $X \leq^{icx} Y$ آن‌گاه

$$I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) \leq I^w(\bar{G}_{Y_{(1:n)}}, \bar{G}).$$

برهان. از آنجا که $h^w(.)$ یک تابع محدب صعودی برای $n \geq 1$ است، از شیکد و شانتی‌کومار (۲۰۰۷) نتیجه می‌شود که $X \leq^{icx} Y$ معادل است با $h^w(X) \leq^{icx} h^w(Y)$. با یادآوری تعریف ترتیب محدب صعودی و گزاره‌ی ۸.۶.۳ اثبات کامل می‌شود. □

گزاره ۱۰.۶.۳. فرض کنید X و Y دو متغیر تصادفی نامنفی با تابع قابلیت اعتماد $\bar{F}(x)$ و $\bar{G}(x)$ باشند. اگر $X \leq^{hr} Y$، آن‌گاه برای $n = 1, 2, \ldots$ برقرار می‌شود

$$\frac{I^w(\bar{F}_{X_{(1:n)}}, \bar{F})}{E(X)} \leq \frac{I^w(\bar{G}_{Y_{(1:n)}}, \bar{G})}{E(Y)}.$$

برهان. با توجه به اینکه تابع $h^w(x) = \int_0^x z[\bar{F}(z)]^{n-1}[-\log \bar{F}(z)]dz$ یک تابع محدب صعودی است، تحت فرض $X \leq^{hr} Y$، از شیکد و شانتی‌کومار (۲۰۰۷) به‌دست می‌آید

$$\frac{E\left[h^w(X)\right]}{E(X)} \leq \frac{E\left[h^w(Y)\right]}{E(Y)}.$$

از این‌رو، با یادآوری (۴۱.۳) اثبات کامل می‌شود. □

گزاره ۱۱.۶.۳. (۱) فرض کنید X یک متغیر تصادفی پیوسته با تابع بقای $\bar{F}(.)$ باشد که در $[0, b]$ مقدار می‌گیرد و b متناهی است. آن‌گاه

$$I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) \leq bI(\bar{F}_{X_{(1:n)}}, \bar{F}).$$

(۲) فرض کنید X یک متغیر تصادفی پیوسته‌ی نامنفی با تابع بقای $\bar{F}(.)$ باشد که در $[a, \infty)$ مقدار می‌گیرد و $a > 0$ متناهی است. آن‌گاه

$$I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) \geq aI(\bar{F}_{X_{(1:n)}}, \bar{F}).$$

فرض کنید که X_θ^* یک متغیر تصادفی کاملاً پیوسته و نامنفی با تابع بقای $\bar{H}_\theta(x) = [\bar{F}(x)]^\theta$ و $x \geq 0$ را نشان دهد. این مدل به‌عنوان مدل نرخ خطر نسبی شناخته می‌شود. حال ما اندازه عدم دقت باقیمانده‌ی تجمعی وزنی را بین $\bar{H}_{X_{(1:n)}}$ و $\bar{H}$ به‌صورت زیر به‌دست می‌آوریم:

$$\begin{aligned} I^w(\bar{H}_{X_{(1:n)}}, \bar{H}) &= -\int_0^{+\infty} x\bar{H}_{X_{(1:n)}}(x) \log\left(\bar{H}(x)\right) dx \\ &= -\theta \int_0^{+\infty} x(\bar{F}(x))^{n\theta} \log \bar{F}(x) dx. \end{aligned} \qquad (۴۲.۳)$$

گزاره ۱۲.۶.۳. اگر $\theta \geq (\leq) 1$، آن‌گاه برای هر $n \geq 1$ داریم

$$I^w(\bar{H}_{X_{(1:n)}}, \bar{H}) \leq (\geq) \theta I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) \leq (\geq) \theta \mathcal{E}^w(X).$$

برهان. فرض کنید که $\theta \geq (\leq) 1$، آن‌گاه واضح است که $[\bar{F}(x)]^\theta \leq (\geq) \bar{F}(x)$ و بنابراین از (۴۲.۳) نتیجه می‌شود

$$I^w(\bar{H}_{X_{(1:n)}}, \bar{H}) \leq (\geq) \theta \mathcal{E}^w(X).$$

□

قضیه ۱.۶.۳. $I^w(\bar{F}_{X_{(1:n)}},\bar{F})=0$ اگر و تنها اگر X تباهیده باشد.

برهان. فرض کنید X در نقطه‌ی a تباهیده باشد. با توجه به تعریف تابع تباهیده و تعریف $I^w(\bar{F}_{X_{(1:n)}},\bar{F})$، واضح است که داریم $I^w(\bar{F}_{X_{(1:n)}},\bar{F})=0$.
حال فرض کنید که $I^w(\bar{F}_{X_{(1:n)}},\bar{F})=0$، یعنی

$$-\int_0^{\infty} x[\bar{F}(x)]^n \log \bar{F}(x)dx=0. \tag{۴۳.۳}$$

آن‌گاه با توجه به این‌که انتگرال (۴۳.۳) نامنفی است، نتیجه می‌گیریم $-x[\bar{F}(x)]^n \log \bar{F}(x)=0$ برای تقریباً تمام $x\in\mathbb{R}^+$. بنابراین، $\bar{F}(x)=0 \ or\ 1$ برای تقریباً تمام $x\in\mathbb{R}^+$. □

اخیراً کالی[۷] و همکاران (۲۰۱۸) CPI تعمیم یافته‌ی مرتبه m را معرفی کرده است که به‌صورت زیر تعریف می‌شود

$$I_m(F,G)=\frac{1}{m!}\int_0^{+\infty} F(x)[-\log G(x)]^m dx. \tag{۴۴.۳}$$

ما اکنون CRI تعمیم یافته‌ی وزنی (WGCRI) از مرتبه m را در مقایسه با اندازه‌ی تعریف شده در (۴۴.۳) به‌صورت زیر تعریف می‌کنیم

$$I^w_m(\bar{F},\bar{G})=\frac{1}{m!}\int_0^{+\infty} x\bar{F}(x)[-\log \bar{G}(x)]^m dx. \tag{۴۵.۳}$$

ملاحظه ۱.۶.۳. فرض کنید X یک متغیر تصادفی کاملاً پیوسته و نامنفی با تابع توزیع F باشد. آن‌گاه، WGCRI مرتبه m بین $\bar{F}_{X_{(1:n)}}$ و F برابر است با

$$\begin{aligned} I^w_m(\bar{F}_{X_{(1:n)}},\bar{F}) &= \frac{1}{m!}\int_0^{\infty} x[\bar{F}(x)]^n[-\log \bar{F}(x)]^m dx \\ &= \frac{1}{n^m}\mathcal{E}^w_m(X_{(1:n)}), \end{aligned} \tag{۴۶.۳}$$

که

$$\mathcal{E}^w_m(X)=\int_0^{\infty} x\frac{[-\log \bar{F}(x)]^m}{m!}\bar{F}(x)dx,$$

آنتروپی باقیمانده‌ی تجمعی تعمیم یافته‌ی وزنی (WGCRE) است که توسط کایال[۸] (۲۰۱۸) معرفی شده است.

ملاحظه ۲.۶.۳. در مقایسه با (۳۲.۳)، اندازه WCRI مربوط به $\bar{F}$ و $\bar{F}_{X_{(1:n)}}$ به‌دست می‌آید

$$I^w(\bar{F},\bar{F}_{X_{(1:n)}})=-\int_0^{+\infty} x\bar{F}(x)\log\left(\bar{F}_{X_{(1:n)}}(x)\right)dx=n\mathcal{E}^w(X). \tag{۴۷.۳}$$

[۷] Cali
[۸] Kayal

در ادامه‌ی این بخش، نسخه‌ی پویای $I^w(\bar{F}_{X_{(1:n)}}, \bar{F})$ را بررسی می‌کنیم. فرض کنید X طول عمر یک سیستم باشد تحت این فرض که سیستم تا سن t سالم مانده باشد. به‌طور مشابه، همچنین می‌توانیم نسخه‌ی پویای $I^w(\bar{F}_{X_{(1:n)}}, \bar{F})$ را به‌صورت زیر در نظر بگیریم

$$\begin{aligned} I^w(\bar{F}_{X_{(1:n)}}, \bar{F}; t) &= -\int_t^{+\infty} x \frac{\bar{F}_{X_{(1:n)}}(x)}{\bar{F}_{X_{(1:n)}}(t)} \log\left(\frac{\bar{F}(x)}{\bar{F}(t)}\right) dx \\ &= \log \bar{F}(t) M^w_{(1:n)}(t) - \int_t^{+\infty} x \frac{\bar{F}_{X_{(1:n)}}(x)}{\bar{F}_{X_{(1:n)}}(t)} \log\left(\bar{F}(x)\right) dx \\ &= \log \bar{F}(t) M^w_{(1:n)}(t) - \frac{1}{(\bar{F}(t))^n} \int_t^{+\infty} x(\bar{F}(x))^n \log \bar{F}(x) dx \quad (۴۸.۳) \end{aligned}$$

توجه کنید که $I^w(\bar{F}_{X_{(1:n)}}, \bar{F}) = \lim_{t\to 0} I^w(\bar{F}_{X_{(1:n)}}, \bar{F}; t)$. از آنجا که $\log \bar{F}(t) \le 0$ برای $t \ge 0$، داریم

$$\begin{aligned} I^w(\bar{F}_{X_{(1:n)}}, \bar{F}; t) &\le -\frac{1}{(\bar{F}(t))^n} \int_t^{+\infty} x(\bar{F}(x))^n \log \bar{F}(x) dx \\ &\le -\frac{1}{(\bar{F}(t))^n} \int_0^{+\infty} x(\bar{F}(x))^n \log \bar{F}(x) dx = \frac{I^w(\bar{F}_{X_{(1:n)}}, \bar{F})}{(\bar{F}(t))^n}. \end{aligned}$$

قضیه ۲.۶.۳. فرض کنید X یک متغیر تصادفی پیوسته و نامنفی با تابع توزیع $F(.)$ باشد. فرض کنید عدم دقت تجمعی پویای وزنی مربوط به اولین آماره ترتیبی را با $I^w(\bar{F}_{X_{(1:n)}}, \bar{F}; t) < \infty$، $t \ge 0$ نشان دهیم. آن‌گاه $I^w(\bar{F}_{X_{(1:n)}}, \bar{F}; t)$ تابع توزیع را مشخصه‌سازی می‌کند.

برهان. از (۴۸.۳) داریم

$$I^w(\bar{F}_{X_{(1:n)}}, \bar{F}; t) = \log \bar{F}(t) M^w_{(1:n)}(t) - \frac{1}{(\bar{F}(t))^n} \int_t^{+\infty} x(\bar{F}(x))^n \log \bar{F}(x) dx. \quad (۴۹.۳)$$

با مشتق‌گیری از دو طرف (۴۹.۳) نسبت به t به‌دست می‌آوریم:

$$\begin{aligned} \frac{\partial}{\partial t}[I^w(\bar{F}_{X_{(1:n)}}, \bar{F}; t)] &= -\lambda_F(t) M^w_{(1:n)}(t) + n\lambda_F(t) I^w(\bar{F}_{X_{(1:n)}}, \bar{F}; t) \\ &= \lambda_F(t)\left[n I^w(\bar{F}_{X_{(1:n)}}, \bar{F}; t) - M^w_{(1:n)}(t)\right]. \end{aligned}$$

با مشتق‌گیری مجدد نسبت به t به‌دست می‌آوریم

$$\acute{\lambda}_F(t) = \frac{(\lambda_F(t))^2 \left(n\lambda_F(t) M^w_{(1:n)}(t) + n\frac{\partial}{\partial t} I^w(\bar{F}_{X_{(1:n)}}, \bar{F}; t) - t\right)}{\frac{\partial}{\partial t} I^w(\bar{F}_{X_{(1:n)}}, \bar{F}; t)}. \quad (۵۰.۳)$$

فرض کنید که دو تابع F و F^* موجود باشند به‌طوری‌که

$$I^w(\bar{F}_{X_{(1:n)}}, \bar{F}; t) = I^w(\bar{F}^*_{X_{(1:n)}}, \bar{F}^*; t) = z(t).$$

آن‌گاه برای تمام مقادیر t از (۵۰.۳) داریم

$$\acute{\lambda}_F(t) = \varphi(t, \lambda_F(t)), \quad \acute{\lambda}_{F^*}(t) = \varphi(t, \lambda_{F^*}(t)),$$

که

$$\varphi(t,y) = \frac{y^{۲}\left[nys(t) + n\acute{z}(t) - t\right]}{\acute{z}(t)},$$

و $s(t) = M^{w}_{(۱:n)}(t)$. با استفاده از قضیه ۲.۳ و لم ۳.۳ از گوپتا و کرمانی (۲۰۰۸) داریم $\lambda_F(t) = \lambda_{F^*}(t)$، برای تمام مقادیر t. از آنجا که تابع نرخ خطر تابع توزیع را به‌طور یکتا مشخصه‌سازی می‌کند، اثبات کامل می‌شود. $\square$

گزاره ۱۳.۶.۳. اگر $X_{(۱)} \leq X_{(۲)} \leq \ldots \leq X_{(n)}$ آماره‌های ترتیبی نمونه‌ی $X_۱, X_۲, \ldots, X_n$ را نشان دهد آن‌گاه، اندازه‌ی تجربی $I^w(\bar{F}_{X_{(۱:n)}}, \bar{F})$ به‌صورت زیر به‌دست می‌آید

$$\begin{aligned}\hat{I}^w(\bar{F}_{X_{(۱:n)}}, \bar{F}) &= -\int_۰^{+\infty} x[\hat{\bar{F}}_n(x)]^n \log \hat{\bar{F}}_n(x)dx \\ &= -\sum_{k=۱}^{n-۱}\int_{X_{(k)}}^{X_{(k+۱)}} x\left(۱-\frac{k}{n}\right)^n \log\left(۱-\frac{k}{n}\right)dx \\ &= -\sum_{k=۱}^{n-۱} U_k\left(۱-\frac{k}{n}\right)^n \log\left(۱-\frac{k}{n}\right), \qquad (۵۱.۳)\end{aligned}$$

که $U_k = \frac{X^{۲}_{(k+۱)} - X^{۲}_{(k)}}{۲}, \quad k = ۱, ۲, \ldots, n-۱$.

قضیه ۳.۶.۳. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته با $I^w(\bar{F}_{X_{(۱:n)}}, \bar{F}) < \infty$ برای تمام $n \geq ۱$ باشد. آن‌گاه داریم

$$\hat{I}^w(\bar{F}_{X_{(۱:n)}}, \bar{F}) \longrightarrow I^w(\bar{F}_{X_{(۱:n)}}, \bar{F}) \quad a.s.$$

برهان. از (۵۱.۳) داریم

$$\begin{aligned}\hat{I}^w(\bar{F}_{X_{(۱:n)}}, \bar{F}) &= \int_۰^{\infty} x(-\log \hat{\bar{F}}_n(x))(\hat{\bar{F}}_n(x))^n dx \\ &= \int_۰^{۱} x(-\log \hat{\bar{F}}_n(x))(\hat{\bar{F}}_n(x))^n dx + \int_۱^{\infty} x(-\log \hat{\bar{F}}_n(x))(\hat{\bar{F}}_n(x))^n dx \\ &=: W_۱ + W_۲, \qquad (۵۲.۳)\end{aligned}$$

که

$$\begin{aligned} W_۱ &= \int_۰^{۱} x(-\log \hat{\bar{F}}_n(x))(\hat{\bar{F}}_n(x))^n dx, \\ W_۲ &= \int_۱^{\infty} x(-\log \hat{\bar{F}}_n(x))(\hat{\bar{F}}_n(x))^n dx.\end{aligned}$$

با استفاده از قضیه همگرایی مغلوب[۹] (پیوست آ.۱) و گلیونکو‌ـ‌کانتلی[۱۰] (پیوست آ.۲) داریم

$$\int_۰^{۱} x(-\log \hat{\bar{F}}_n(x))(\hat{\bar{F}}_n(x))^n dx \longrightarrow \int_۰^{۱} x(-\log \bar{F}(x))(\bar{F}(x))^n dx \quad as\ n \to \infty. \qquad (۵۳.۳)$$

[۹] Dominated Convergence Theorem (DCT)

[۱۰] Glivenko–Cantelli

با توجه به تعریف $\hat{\bar{F}}_n(x)$ به‌دست می‌آید

$$x^p \hat{\bar{F}}_n(x) \leq \frac{1}{n}\sum_{i=1}^{n} X_i^p.$$

علاوه بر این، با استفاده از قانون قوی اعداد بزرگ[11] داریم $\frac{1}{n}\sum_{i=1}^{n} X_i^p \longrightarrow \mathbb{E}(X^p)$ و $\sup_n(\frac{1}{n}\sum_{i=1}^{n} X_i^p) < \infty$، آن‌گاه $\hat{\bar{F}}_n(x) \leq x^{-p}\left(\sup_n(\frac{1}{n}\sum_{i=1}^{n} X_i^p)\right) = Cx^{-p}$. حال با استفاده از DCT داریم

$$\lim_{n\to\infty} W_2 = \int_1^{\infty} x(-\log \bar{F}(x))(\bar{F}(x))^n dx. \tag{۵۴.۳}$$

در آخر با استفاده از (۵۲.۳) نتیجه حاصل می‌شود. □

۷.۳ عدم دقت گذشته‌ی تجمعی وزنی برای $X_{(n:n)}$

عدم دقت گذشته‌ی تجمعی وزنی[12] (WCPI) بین $F_{X_{(n:n)}}$ و F را به‌صورت زیر در نظر می‌گیریم:

$$\tilde{I}^w(F_{X_{(n:n)}}, F) = -\int_0^{+\infty} xF_{X_{(n:n)}}(x)\log\left(F(x)\right)dx = \frac{1}{n}\mathcal{CE}^w(X_{(n:n)}), \tag{۵۵.۳}$$

که

$$\mathcal{CE}^w(X_{(n:n)}) = n\int_0^{+\infty} x[F(x)]^n \tilde{\lambda}(x)dx.$$

در ادامه برخی ویژگی‌های $\tilde{I}^w(F_{X_{(n:n)}}, F)$ را مطرح می‌کنیم.

گزاره ۱.۷.۳. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته با $\tilde{I}^w(F_{X_{(n:n)}}, F) < \infty$ برای $n \geq 1$ باشد. آن‌گاه داریم

$$\begin{aligned}\tilde{I}^w(F_{X_{(n:n)}}, F) &= \int_0^{+\infty} x[F(x)]^n \left(\int_x^{\infty} \frac{f(z)}{F(z)}dz\right) xdx \\ &= \int_0^{+\infty}\int_0^{z} \tilde{\lambda}(z)x[F(x)]^n dxdz,\end{aligned} \tag{۵۶.۳}$$

که $\tilde{\lambda}(.)$ تابع نرخ خطر معکوس است.

گزاره ۲.۷.۳. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته با $\tilde{I}^w(F_{X_{(n:n)}}, F) < \infty$ برای $n \geq 1$ باشد. آن‌گاه داریم

$$\tilde{I}^w(F_{X_{(n:n)}}, F) = E\left(\tilde{M}^w_{(n:n)}(Z)(F(Z))^{n-1}\right), \tag{۵۷.۳}$$

که

$$\tilde{M}^w_{(n:n)}(z) = \frac{1}{(F(z))^n}\int_0^{z} x(F(x))^n dx$$

میانگین وزنی زمان غیرفعال[13] (WMIT) $X_{(n:n)}$ است.

[11] SLLN
[12] Weighted cumulative past inaccuracy
[13] Weighted mean inactivity time

برهان. با توجه به (۵۶.۳) به‌دست می‌آوریم

$$\begin{aligned}\tilde{I}^w(F_{X_{(n:n)}}, F) &= \int_{0}^{+\infty}\int_{0}^{z}\tilde{\lambda}(z)x[F(x)]^n dxdz\\ &= \int_{0}^{+\infty}\frac{f(z)}{F(z)}dz\left[\int_{0}^{z}x[F(x)]^n dx\right]\\ &= \int_{0}^{+\infty}\frac{f(z)}{F(z)}[F(z)]^n\tilde{M}^w_{(n:n)}(z)dz = \int_{0}^{+\infty}f(z)[F(z)]^{n-1}\tilde{M}^w_{(n:n)}(z)dz.\end{aligned} \tag{۵۸.۳}$$

بنابراین اثبات کامل است. □

گزاره ۳.۷.۳. اگر $a, b > 0$، برای $n = 1, 2, \ldots$

$$\tilde{I}^w(F_{aX_{(n:n)}+b}, F_{aX+b}) = a\tilde{I}^w(F_{X_{(n:n)}}, F).$$

در قضایای بعدی برخی کران‌های بالا و پایین را برای $\tilde{I}(F_{X_{(n:n)}}, F)$ به دست می‌آوریم.

گزاره ۴.۷.۳. برای متغیر تصادفی نامنفی X و $n \geq 1$، برقرار می‌شود

$$\tilde{I}^w(F_{X_{(n:n)}}, F) \geq \tilde{M}^w_{(n:n)}(t)|\log F(t)|[F(t)]^n, \tag{۵۹.۳}$$

که $\tilde{M}_{(n:n)}(t)$ WMIT مربوط به $X_{(n:n)}$ است.

برهان. از براتپور (۲۰۱۰) اثبات به‌دست می‌آید. □

گزاره ۵.۷.۳. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته با $\tilde{I}^w(F_{X_{(n:n)}}, F) < \infty$ برای $n \geq 1$ باشد. آن‌گاه داریم

$$\tilde{I}^w(F_{X_{(n:n)}}, F) \geq \int_{0}^{+\infty}x(F(x))^n\bar{F}(x)dx. \tag{۶۰.۳}$$

برهان. با یادآوری اینکه $-\log F(x) \geq \bar{F}(x)$ اثبات به‌دست می‌آید. □

گزاره ۶.۷.۳. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته با $\tilde{I}^w(F_{X_{(n:n)}}, F) < \infty$ برای $n \geq 1$ باشد. آن‌گاه داریم

$$\tilde{I}^w(F_{X_{(n:n)}}, F) \leq \mathcal{CE}^w(X). \tag{۶۱.۳}$$

برهان. از آنجا که $F(x) \geq [F(x)]^n,\ x \geq 0$ وقتی که $n \geq 1$، اثبات به‌دست می‌آید. □

گزاره ۷.۷.۳. اگر X DRFRA باشد، آن‌گاه

$$\tilde{I}^w(F_{X_{(n:n)}}, F) \leq \mathbb{E}\left(X^2(F(X))^{n-1}\right). \tag{۶۲.۳}$$

برهان. از آنجا که X DRFRA است، $\frac{\tilde{\Lambda}(x)}{x}$ نسبت به $x>0$ نزولی است که نتیجه می‌دهد

$$F(x)\tilde{\Lambda}(x) \le xf(x), \quad x>0. \tag{۶۳.۳}$$

با ضرب $x[F(x)]^{n-1} \ge 0$ در (۶۳.۳) و انتگرال گیری از آن، نتیجه به‌دست می‌آید. □

گزاره ۸.۷.۳. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته با $\tilde{I}^w(F_{X_{(n:n)}},F)<\infty$ برای $n\ge 1$ باشد. آن‌گاه داریم

$$\tilde{I}^w(F_{X_{(n:n)}},F)=\mathbb{E}\left[\tilde{h}^w(X)\right], \tag{۶۴.۳}$$

که

$$\tilde{h}^w(x)=\int_x^\infty z\left[-\log F(z)\right]\left[F(z)\right]^{n-1}dz, \quad x\ge 0.$$

برهان. از (۵۵.۳) و با استفاده از قضیه فوبینی به‌دست می‌آوریم

$$\begin{aligned}
\tilde{I}^w(F_{X_{(n:n)}},F) &= \int_0^\infty z[-\log F(z)][F(z)]^{n-1}F(z)dz\\
&= \int_0^\infty \left[\int_0^z f(x)dx\right] z[F(z)]^{n-1}[-\log F(z)]dz\\
&= \int_0^\infty f(x)\left[\int_x^\infty z[F(z)]^{n-1}[-\log F(z)]dz\right]dx=\mathbb{E}\left[\tilde{h}^w(X)\right].
\end{aligned}$$

□

گزاره ۹.۷.۳. فرض کنید X و Y دو متغیر تصادفی نامنفی با توابع بقای $\bar{F}(x)$ و $\bar{G}(x)$ باشند. اگر $X\le^{icx} Y$ آن‌گاه

$$\tilde{I}^w(F_{X_{(n:n)}},F)\le \tilde{I}^w(G_{Y_{(n:n)}},G).$$

برهان. از آنجا که $\tilde{h}^w(.)$ یک تابع محدب صعودی برای $n\ge 1$ است، از شیکد و شانتی‌کومار (۲۰۰۷) نتیجه می‌شود که $X\le^{icx} Y$ معادل است با $\tilde{h}^w(X)\le^{icx}\tilde{h}^w(Y)$. با یادآوری تعریف ترتیب محدب صعودی و گزاره‌ی (۸.۷.۳) اثبات کامل می‌شود. □

گزاره ۱۰.۷.۳. فرض کنید X و Y دو متغیر تصادفی نامنفی با توابع بقای $\bar{F}(x)$ و $\bar{G}(x)$ باشند. اگر $X\le^{hr} Y$، آن‌گاه برای $n=1,2,\ldots$ برقرار می‌شود

$$\frac{\tilde{I}^w(F_{X_{(n:n)}},F)}{E(X)}\le\frac{I^w(\tilde{I}^w(G_{Y_{(n:n)}},G)}{E(Y)}.$$

برهان. با توجه به اینکه تابع $\tilde{h}^w(x)=\int_x^\infty z[F(z)]^{n-1}[-\log F(z)]dz$ یک تابع محدب صعودی است، تحت فرض $X\le^{hr} Y$، از شیکد و شانتی کومار (۲۰۰۷) نتیجه می‌شود

$$\frac{E\left[\tilde{h}^w(X)\right]}{E(X)}\le\frac{E\left[\tilde{h}^w(Y)\right]}{E(Y)}.$$

بنابراین، با یادآوری (۶۴.۳) اثبات کامل می‌شود. □

گزاره ۱۱.۷.۳. (۱) فرض کنید X یک متغیر تصادفی پیوسته با تابع بقای $\bar{F}(.)$ باشد که در $[0, b]$ مقدار می‌گیرد و b متناهی است. آن‌گاه

$$\tilde{I}^w(F_{X_{(n:n)}}, F) \leq b\tilde{I}(F_{X_{(n:n)}}, F).$$

(۲) فرض کنید X یک متغیر تصادفی پیوسته و نامنفی با تابع بقای $\bar{F}(.)$ باشد که در $[a, \infty)$ مقدار می‌گیرد و $a > 0$ متناهی است. آن‌گاه

$$\tilde{I}^w(F_{X_{(n:n)}}, F) \geq a\tilde{I}(F_{X_{(n:n)}}, F).$$

فرض کنید که X^*_θ نشان‌دهنده‌ی یک متغیر تصادفی کاملاً پیوسته و نامنفی با تابع توزیع $H_\theta(x) = [F(x)]^\theta, x \geq 0$ باشد. این مدل به‌عنوان مدل نرخ خطر نسبی شناخته می‌شود. حال، اندازه عدم دقت گذشته‌ی تجمعی وزنی بین $H_{X_{(n:n)}}$ و H را به‌صورت زیر به‌دست می‌آوریم:

$$\begin{aligned} \tilde{I}^w(H_{X_{(n:n)}}, H) &= -\int_0^{+\infty} xH_{X_{(n:n)}}(x)\log\left(H(x)\right)dx \\ &= -\theta\int_0^{+\infty} x(F(x))^{n\theta}\log F(x)dx. \end{aligned} \tag{۶۵.۳}$$

گزاره ۱۲.۷.۳. اگر $\theta \geq (\leq) 1$، آن‌گاه برای هر $n \geq 1$ داریم

$$\tilde{I}^w(H_{X_{(n:n)}}, H) \leq (\geq)\theta\tilde{I}^w(F_{X_{(n:n)}}, F) \leq (\geq)\theta\mathcal{CE}^w(X).$$

برهان. فرض کنید $\theta \geq (\leq) 1$. آن‌گاه واضح است که $[F(x)]^\theta \leq (\geq) F(x)$ و بنابراین از (۶۵.۳) نتیجه می‌شود

$$\tilde{I}^w(H_{X_{(n:n)}}, H) \leq (\geq)\theta\mathcal{CE}^w(X).$$

□

قضیه ۱.۷.۳. $\tilde{I}^w(F_{X_{(n:n)}}, F) = 0$، اگر و تنها اگر، X تباهیده باشد.

برهان. فرض کنید X در نقطه‌ی a تباهیده باشد. با توجه به تعریف تابع تباهیده و تعریف $\tilde{I}^w(F_{X_{(n:n)}}, F)$ واضح است که $\tilde{I}^w(F_{X_{(n:n)}}, F) = 0$.
حال فرض کنید $\tilde{I}^w(F_{X_{(n:n)}}, F) = 0$، یعنی

$$-\int_0^{\infty} x[F(x)]^n\log F(x)dx = 0. \tag{۶۶.۳}$$

آن‌گاه، با توجه به این‌که انتگرال (۶۶.۳) نامنفی است نتیجه می‌گیریم $-x[F(x)]^n\log F(x) = 0$، برای تقریباً تمام $x \in \mathbb{R}^+$. بنابراین، $F(x) = 0$ or 1 برای تقریباً تمام $x \in \mathbb{R}^+$. □

گزاره ۱۳.۷.۳. فرض کنید X یک متغیر تصادفی کاملاً پیوسته با $\tilde{I}^w(F_{X_{(n:n)}}, F) < \infty$ برای $n \geq 1$ باشد. آن‌گاه داریم

$$\tilde{I}^w(F_{X_{(n:n)}}, F) = \int_0^1 F_X^{-1}(u)\frac{\psi_n(u)}{f_X\left(F_X^{-1}(u)\right)}du,$$

که $\psi_n(u) = -u^n\log u$ برای $0 < u < 1$. توجه کنید که $\psi_n(0) = \psi_n(1) = 0$.

اخیراً کالی و همکاران (۲۰۱۸) CPI تعمیم یافته مرتبه m را به‌صورت زیر تعریف کرده‌اند

$$I_m(F,G)=\frac{1}{m!}\int_{0}^{+\infty}F(x)[-\log G(x)]^m dx. \tag{۶۷.۳}$$

اکنون CPI تعمیم یافته‌ی وزنی (WGCPI) مرتبه‌ی m را در مقایسه با اندازه‌ی تعریف شده در (۶۷.۳) به‌صورت زیر معرفی می‌کنیم

$$I_m^w(F,G)=\frac{1}{m!}\int_{0}^{+\infty}xF(x)[-\log G(x)]^m dx. \tag{۶۸.۳}$$

ملاحظه ۱.۷.۳. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته با تابع توزیع F باشد. آن‌گاه، WGCPI مرتبه‌ی m بین $F_{X_{(n:n)}}$ و F برابر است با

$$\begin{aligned}\tilde{I}_m^w(F_{X_{(n:n)}},F) &= \frac{1}{m!}\int_{0}^{\infty}x[F(x)]^n[-\log F(x)]^m dx\\ &= \frac{1}{n^m}\mathcal{CE}_m^w(X_{(n:n)}),\end{aligned} \tag{۶۹.۳}$$

که

$$\mathcal{CE}_m^w(X)=\int_{0}^{\infty}x\frac{[-\log F(x)]^m}{m!}F(x)dx,$$

آنتروپی تجمعی تعمیم‌یافته‌ی وزنی (WGCE) است که توسط کایال و مهارانا (۲۰۱۷) معرفی شده است.

ملاحظه ۲.۷.۳. در مقایسه با (۵۵.۳)، اندازه WCPI مرتبط با F و $F_{X_{(n:n)}}$ به‌صورت زیر است

$$\tilde{I}^w(F,F_{X_{(n:n)}})=-\int_{0}^{+\infty}xF(x)\log\left(F_{X_{(n:n)}}(x)\right)dx=n\mathcal{CE}^w(X). \tag{۷۰.۳}$$

در ادامه‌ی این بخش، نسخه‌ی پویای $\tilde{I}^w(F_{X_{(n:n)}},F)$ را بررسی می‌کنیم. اگر یک سیستم که در زمان ۰ شروع به کار کرده است و فقط در زمان‌های بازرسی معین مشاهده می‌شود، در زمان t از کار افتاده مشاهده شود، آن‌گاه نسخه‌ی پویای $\tilde{I}^w(F_{X_{(n:n)}},F)$ را به‌صورت زیر در نظر می‌گیریم

$$\begin{aligned}\tilde{I}^w(F_{X_{(n:n)}},F;t) &= -\int_{0}^{t}x\frac{F_{X_{(n:n)}}(x)}{F_{X_{(n:n)}}(t)}\log\left(\frac{F(x)}{F(t)}\right)dx\\ &= \log F(t)\tilde{M}_{(n:n)}^w(t)-\int_{0}^{t}x\frac{F_{X_{(n:n)}}(x)}{F_{X_{(n:n)}}(t)}\log\left(F(x)\right)dx\\ &= \log F(t)\tilde{M}_{(n:n)}^w(t)-\frac{1}{(F(t))^n}\int_{0}^{t}x(F(x))^n\log F(x)dx.\end{aligned} \tag{۷۱.۳}$$

دقت کنید که $\lim_{t\to 0}\tilde{I}^w(F_{X_{(n:n)}},F;t)=\tilde{I}^w(F_{X_{(n:n)}},F)$. از آنجا که $\log F(t)\leq 0$ برای $t\geq 0$، داریم

$$\begin{aligned}\tilde{I}^w(F_{X_{(n:n)}},F;t) &\leq -\frac{1}{(F(t))^n}\int_{0}^{t}x(F(x))^n\log F(x)dx\\ &\leq -\frac{1}{(F(t))^n}\int_{0}^{+\infty}x(F(x))^n\log F(x)dx=\frac{\tilde{I}^w(F_{X_{(n:n)}},F)}{(F(t))^n}.\end{aligned}$$

قضیه ۲.۷.۳. فرض کنید X یک متغیر تصادفی نامنفی با تابع توزیع $F(.)$ باشد. فرض کنید عدم دقت تجمعی پویای وزنی مربوط به nـامین رکورد را با $\tilde{I}^w(F_{X_{(n:n)}}, F; t) < \infty$ و $t \geq 0$ نشان دهیم. آنگاه $\tilde{I}^w(F_{X_{(n:n)}}, F; t)$ تابع توزیع را مشخصه‌سازی میکند.

برهان. از (۷۱.۳) داریم

$$\tilde{I}^w(F_{X_{(n:n)}}, F; t) = \log F(t)\tilde{M}^w_{(n:n)}(t) - \frac{1}{(F(t))^n}\int_0^t x(F(x))^n \log F(x). \tag{۷۲.۳}$$

با مشتقگیری از هر دو طرف رابطهی (۷۲.۳) نسبت به t بهدست میآوریم:

$$\begin{aligned}\frac{d}{dt}[\tilde{I}^w(F_{X_{(n:n)}}, F; t)] &= -\tilde{\lambda}_F(t)\tilde{M}^w_{(n:n)}(t) - n\tilde{\lambda}_F(t)\tilde{I}^w(F_{X_{(n:n)}}, F; t)\\ &= -\tilde{\lambda}_F(t)\left[\tilde{M}^w_{(n:n)}(t) - n\tilde{I}^w(F_{X_{(n:n)}}, F; t)\right]\end{aligned}$$

با مشتقگیری دوباره نسبت به t بهدست میآوریم

$$\acute{\tilde{\lambda}}_F(t) = \frac{\left(\tilde{\lambda}_F(t)\right)^2\left(n\tilde{\lambda}_F(t)\tilde{M}^w_{(n:n)}(t) + n\frac{\partial}{\partial t}\tilde{I}^w(F_{X_{(n:n)}}, F; t) - t\right)}{\frac{\partial}{\partial t}\tilde{I}^w(F_{X_{(n:n)}}, F; t)}. \tag{۷۳.۳}$$

فرض کنید دو تابع F و F^* وجود دارند بهطوریکه

$$\tilde{I}^w(F_{X_{(n:n)}}, F; t) = \tilde{I}^w(F^*_{X_{(n:n)}}, F^*; t) = z(t).$$

آنگاه برای تمام مقادیر t، از (۷۳.۳) داریم

$$\acute{\tilde{\lambda}}_F(t) = \varphi(t, \lambda_F(t)), \quad \acute{\tilde{\lambda}}_{F^*}(t) = \varphi(t, \lambda_{F^*}(t)),$$

که

$$\varphi(t, y) = \frac{y^2\left[nys(t) + n\acute{z}(t) - t\right]}{\acute{z}(t)},$$

و $\tilde{s}(t) = \tilde{M}^w_{(n:n)}(t)$. با استفاده از قضیه ۲.۳ و لم ۳.۳ از گوپتا و کرمانی (۲۰۰۸) داریم $\lambda_F(t) = \lambda_{F^*}(t)$ برای تمام مقادیر t. از آنجا که تابع نرخ خطر تابع توزیع را بهطور یکتا مشخصهسازی میکند، اثبات کامل میشود. □

گزاره ۱۴.۷.۳. اگر $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$ آمارههای ترتیبی از نمونهی $X_1, X_2, \ldots, X_n$ باشند آنگاه، اندازهی تجربی $\tilde{I}^w(F_{X_{(n:n)}}, F)$ بهصورت زیر داده میشود

$$\begin{aligned}\hat{\tilde{I}}^w(F_{X_{(n:n)}}, F) &= -\int_0^{+\infty} x[\hat{F}_n(x)]^n \log \hat{F}_n(x)dx\\ &= -\sum_{k=1}^{n-1}\int_{X_{(k)}}^{X_{(k+1)}} x\left(\frac{k}{n}\right)^n \log\left(\frac{k}{n}\right)dx\\ &= \frac{-1}{n^n}\sum_{k=1}^{n-1} k^n U_k \log\left(\frac{k}{n}\right),\end{aligned} \tag{۷۴.۳}$$

که $k = 1, 2, \ldots, n-1$، $U_k = \frac{X^2_{(k+1)} - X^2_{(k)}}{2}$.

مثال ۱.۷.۳. نمونه‌ی تصادفی $X_1, X_2, ..., X_n$ را از توزیع وایبل با تابع چگالی زیر در نظر بگیرید

$$f(x) = 2\lambda \exp(-\lambda x^2).$$

آن‌گاه $Y_k = X_k^2$ دارای توزیع نمایی با میانگین $\frac{1}{\lambda}$ است. در این حالت، $2U_k = X_{(k+1)}^2 - X_{(k)}^2$ مستقل و دارای توزیع نمایی با میانگین $\frac{1}{\lambda(n-k)}$ است (برای جزییات بیشتر پیک، ۱۹۶۵ را ببینید). اکنون از (۷۴.۳) به‌دست می‌آوریم

$$\mathbb{E}[\hat{\tilde{I}}^w(F_{X_{(n:n)}}, F)] = \frac{-1}{n^n} \sum_{k=1}^{n-1} \frac{k^n}{2\lambda(n-k)} \log\frac{k}{n}, \tag{۷۵.۳}$$

و

$$Var[\hat{\tilde{I}}^w(F_{X_{(n:n)}}, F)] = \frac{1}{n^{2n}} \sum_{k=1}^{n-1} \frac{k^{2n}}{4\lambda^2(n-k)^2} \left(\log\frac{k}{n}\right)^2. \tag{۷۶.۳}$$

ما مقادیر $\mathbb{E}[\hat{\tilde{I}}^w(F_{X_{(n:n)}}, F)]$ و $Var[\hat{\tilde{I}}^w(F_{X_{(n:n)}}, F)]$ را برای حجم‌های نمونه‌ی $n = 10, 15, 20$ و $\lambda = 0/5, 1, 2$ در جدول ۱.۳ محاسبه کرده‌ایم. به‌راحتی می‌توان مشاهده کرد که $\mathbb{E}[\hat{\tilde{I}}^w(F_{X_{(n:n)}}, F)]$ در n نزولی است. همچنین ملاحظه می کنیم که $\lim_{n\to\infty} Var[\hat{\tilde{I}}^w(F_{X_{(n:n)}}, F)] = 0$.

مثال ۲.۷.۳. فرض کنید $X_1, X_2, ..., X_n$ یک نمونه‌ی تصادفی از جمعیتی با تابع چگالی $f(x) = 2x$ و $0 < x < 1$ باشد. آن‌گاه $2U_k$ مستقل و دارای توزیع بتا با پارامتر ۱ و k است. حال از (۷۴.۳) به‌دست می‌آوریم

$$\mathbb{E}[\hat{\tilde{I}}^w(F_{X_{(n:n)}}, F)] = \frac{-1}{n^n} \sum_{k=1}^{n-1} \frac{k^n}{2(n+1)} \log\frac{k}{n}, \tag{۷۷.۳}$$

و

$$Var[\hat{\tilde{I}}^w(F_{X_{(n:n)}}, F)] = \frac{1}{n^{2n-1}} \sum_{k=1}^{n-1} \frac{k^n}{4(n+1)^2(n+2)} \left(\log\frac{k}{n}\right)^2. \tag{۷۸.۳}$$

ما مقادیر $\mathbb{E}[\hat{\tilde{I}}^w(F_{X_{(n:n)}}, F)]$ و $Var[\hat{\tilde{I}}^w(F_{X_{(n:n)}}, F)]$ را برای حجم‌های نمونه‌ی $n = 10, 15, 20$ در جدول ۲.۳ محاسبه کرده‌ایم. به‌راحتی می‌توان مشاهده کرد که $\mathbb{E}[\hat{\tilde{I}}^w(F_{X_{(n:n)}}, F)]$ در n نزولی است. همچنین، مشاهده می‌شود که $\lim_{n\to\infty} Var[\hat{\tilde{I}}^w(F_{X_{(n:n)}}, F)] = 0$.

قضیه ۳.۷.۳. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته با $\tilde{I}^w(F_{X_{(n:n)}}, F) < \infty$، برای تمام مقادیر $n \geq 1$ باشد. آن‌گاه داریم

$$\hat{\tilde{I}}^w(F_{X_{(n:n)}}, F) \longrightarrow \tilde{I}^w(F_{X_{(n:n)}}, F) \quad a.s.$$

جدول ۱.۳: مقادیر عددی $\mathbb{E}[\hat{\hat{I}}^w(F_{X_{(n:n)}}, F)]$ و $Var[\hat{\hat{I}}^w(F_{X_{(n:n)}}, F)]$ برای توزیع وایبل

	$\mathbb{E}[\hat{\hat{I}}^w(F_{X_{(n:n)}}, F)]$			$Var[\hat{\hat{I}}^w(F_{X_{(n:n)}}, F)]$		
n	$\lambda = ۰/۵$	$\lambda = ۱$	$\lambda = ۲$	$\lambda = ۰/۵$	$\lambda = ۱$	$\lambda = ۲$
۱۰	۰.۰۵۳۰	۰.۰۲۶۵	۰.۰۱۳۲	۰.۰۰۱۵۰	۰.۰۰۰۳۷	۰.۰۰۰۰۹
۱۵	۰.۰۳۶۵	۰.۰۱۸۲	۰.۰۰۹۱	۰.۰۰۰۶۸	۰.۰۰۰۱۷	۰.۰۰۰۰۴
۲۰	۰.۰۲۷۸	۰.۰۱۳۹	۰.۰۰۶۹	۰.۰۰۰۳۸	۰.۰۰۰۰۹	۰.۰۰۰۰۲

جدول ۲.۳: مقادیر عددی $\mathbb{E}[\hat{\hat{I}}^w(F_{X_{(n:n)}}, F)]$ و $Var[\hat{\hat{I}}^w(F_{X_{(n:n)}}, F)]$ برای توزیع بتا

$\mathbb{E}[\hat{\hat{I}}^w(F_{X_{(n:n)}}, F)]$			$Var[\hat{\hat{I}}^w(F_{X_{(n:n)}}, F)]$		
n=۱۰	n=۱۵	n=۲۰	n=۱۰	n=۱۵	n=۲۰
۰.۰۰۳۳۹	۰.۰۰۱۶۶	۰.۰۰۰۹۸	۲.۵۷e-۱۵	۱.۴۳e-۲۳	۲.۱۱e-۳۲

برهان. از (۷۴.۳) داریم

$$\begin{aligned}\hat{\hat{I}}^w(F_{X_{(n:n)}}, F) &= \int_۰^\infty x(-\log \hat{F}_n(x))(\hat{F}_n(x))^n dx \\ &= \int_۰^۱ x(-\log \hat{F}_n(x))(\hat{F}_n(x))^n dx + \int_۱^\infty x(-\log \hat{F}_n(x))(\hat{F}_n(x))^n dx \\ &=: R_۱ + R_۲, \end{aligned} \tag{۷۹.۳}$$

که

$$R_۱ = \int_۰^۱ x(-\log \hat{F}_n(x))(\hat{F}_n(x))^n dx,$$
$$R_۲ = \int_۱^\infty x(-\log \hat{F}_n(x))(\hat{F}_n(x))^n dx.$$

با استفاده از قضیه همگرایی مغلوب و گلیونکو_ کانتلی داریم

$$\int_۰^۱ x(-\log \hat{F}_n(x))(\hat{F}_n(x))^n dx \longrightarrow \int_۰^۱ x(-\log F(x))(F(x))^n dx \quad as\ m \to \infty. \tag{۸۰.۳}$$

هم‌اکنون با استفاده از تعریف تابع توزیع تجربی داریم

$$x^p \hat{F}_n(x) \leq \frac{۱}{n}\sum_{i=۱}^n X_i^p.$$

علاوه بر این، با استفاده از SLLN داریم $\frac{۱}{n}\sum_{i=۱}^n X_i^p \longrightarrow \mathbb{E}(X^p)$ و $\sup_n(\frac{۱}{n}\sum_{i=۱}^n X_i^p) < \infty$، آن‌گاه $\hat{F}_n(x) \leq x^{-p}\left(\sup_n(\frac{۱}{n}\sum_{i=۱}^n X_i^p)\right) = Cx^{-p}$. حال با استفاده از قضیه همگرایی مغلوب داریم

$$\lim_{n\to\infty} R_۲ = \int_۱^\infty x(-\log F(x))(F(x))^n dx. \tag{۸۱.۳}$$

در آخر با استفاده از (۷۹.۳) نتیجه حاصل می‌شود. □

در مثال‌های زیر $I^w(\bar{F}_{X_{(1:n)}},\bar{F})$ و $\tilde{I}^w(F_{X_{(n:n)}},F)$ را برای برخی توزیع‌های طول عمر خاص که در قابلیت اعتماد و آزمون بقا کاربرد دارند، حساب کرده‌ایم.

مثال ۳.۷.۳. (۱) اگر X دارای توزیع یکنواخت در $[0,\theta]$ باشد، آن‌گاه به‌راحتی می‌توان دید که $\tilde{I}^w(F_{X_{(n:n)}},F)=\frac{\theta^2}{(n+2)^2}$ و $I^w(\bar{F}_{X_{(1:n)}},\bar{F})=\frac{(2n+3)\theta^2}{(n^2+3n+2)^2}$، برای تمام مقادیر صحیح $n\geq 1$. توجه کنید که $\tilde{I}^w(F_{X_{(n:n)}},F)$ و $I^w(\bar{F}_{X_{(1:n)}},\bar{F})$ توابعی نزولی از n هستند.

(۲) اگر X دارای توزیع وایبل با تابع بقای $\bar{F}(x)=e^{-\lambda^q x^q},\ x>0,\ \lambda,q>0$ باشد، آن‌گاه برای تمام مقادیر صحیح $n\geq 1$ به‌دست می‌آوریم $I^w(\bar{F}_{X_{(1:n)}},\bar{F})=\frac{2}{(\lambda q)^2 n^{\frac{q+2}{q}}}\Gamma(\frac{2}{q})$.

(۳) اگر X دارای توزیع پارتو با تابع چگالی $f(x)=\frac{\alpha\beta^\alpha}{x^{\alpha+1}},\ x\geq\beta,\ \beta>0,\ \alpha>0$ باشد، آن‌گاه $I^w(\bar{F}_{X_{(1:n)}},\bar{F})=\frac{\alpha\beta^2}{(n\alpha-2)^2}$ برای تمام مقادیر صحیح $n>\frac{2}{\alpha}$. دقت کنید که $I^w(\bar{F}_{X_{(1:n)}},\bar{F})$ تابعی نزولی از n برای تمام مقادیر $\alpha>\frac{2}{n}$ است.

(۴) فرض کنید X دارای توزیع نمایی با میانگین $\frac{1}{\lambda}$ باشد، آن‌گاه $I^w(\bar{F}_{X_{(1:n)}},\bar{F})=\frac{2}{n^3\lambda^2}$. توجه کنید که $I(\bar{F}_{X_{(1:n)}},\bar{F})$ تابعی نزولی از n است.

(۵) فرض کنید X یک متغیر تصادفی نامنفی باشد که دارای توزیع وایبل با تابع توزیع $F(x)=\exp(-(\frac{\alpha}{x})^\beta)$ و $x>0$ است. آن‌گاه برای تمام مقادیر صحیح $n\geq 1$ به‌دست می‌آوریم $\tilde{I}^w(F_{X_{(n:n)}},F)=\frac{\alpha^2 n^{\frac{2-\beta}{\beta}}}{\beta}\Gamma(\frac{\beta-2}{\beta})$.

فصل ۴

معیارهای عدم دقت گذشته (باقیمانده) تجمعی وزنی در مقادیر رکورد

۱.۴ مقدمه

فرض کنید X و Y دو متغیر تصادفی نامنفی با تابع توزیع $F(x)$ و $G(x)$ باشند. اخیراً کوندو (۲۰۱۷) اندازه عدم دقت وزنی را به‌صورت زیر مطرح کرده است

$$I^w(f,g) = -\int_0^{+\infty} xf(x)\log g(x)dx. \tag{۱.۴}$$

مشابه (۱.۲)، ما عدم دقت گذشته‌ی تجمعی وزنی را به‌صورت زیر تعریف می‌کنیم

$$I^w(F,G) = -\int_0^{+\infty} xF(x)\log G(x)dx. \tag{۲.۴}$$

همچنین مشابه (۲.۲)، عدم دقت باقیمانده‌ی تجمعی وزنی را به‌صورت زیر تعریف می‌کنیم

$$\bar{I}^w(\bar{F},\bar{G}) = -\int_0^{+\infty} x\bar{F}(x)\log \bar{G}(x)dx. \tag{۳.۴}$$

۲.۴ عدم دقت گذشته‌ی تجمعی وزنی برای L_n

در این بخش ما اندازه‌ی وزنی را برای CPI بین F_{L_n} و F مطرح می‌کنیم. برای این منظور، تعدادی از ویژگی‌ها و نتایج مشخصه‌سازی را تحت برخی فرضیات بررسی می‌کنیم.

تعریف ۱.۲.۴. فرض کنید X یک متغیر تصادفی کاملاً پیوسته و نامنفی با تابع توزیع F باشد. آن‌گاه، WCPI بین F_{L_n} (تابع توزیع nـامین رکورد پایین L_n) و F را به‌صورت زیرتعریف می‌کنیم

$$\begin{aligned} I^w(F_{L_n},F) &= -\int_0^\infty xF_{L_n}(x)\log F(x)dx \\ &= \int_0^\infty \sum_{j=0}^{n-1}(j+1)x[F_{L_{j+2}}(x)-F_{L_{j+1}}(x)]dx \\ &= \sum_{j=0}^{n-1}(j+1)\mathbb{E}_{L_{j+2}}\left[\frac{X}{\tilde{\lambda}(X)}\right], \end{aligned} \tag{۴.۴}$$

در صورتی که $\tilde{\lambda}(x)=\frac{f(x)}{F(x)}$ تابع نرخ خطر معکوس و L_{j+2} یک متغیر تصادفی با تابع چگالی $f_{L_{j+2}}(x)=\frac{[-\log F(x)]^{j+1}f(x)}{(j+1)!}$ است.

در ادامه، چند مثال و برخی ویژگی‌های $I^w(F_{L_n},F)$ را بیان می کنیم.

مثال ۱.۲.۴. ۱. اگر X دارای توزیع وایبل معکوس با تابع توزیع $F(x)=\exp(-(\frac{\alpha}{x})^\beta),\ x>0$ باشد، آن‌گاه داریم

$$I^w(F_{L_n},F)=\frac{\alpha^2}{\beta}\sum_{j=0}^{n-1}\frac{\Gamma\left(\frac{(j+1)\beta-2}{\beta}\right)}{j!}.$$

۲. اگر X دارای توزیع یکنواخت در $[0,\theta]$ باشد، به‌دست می‌آوریم

$$I^w(F_{L_n},F)=\theta^2\sum_{j=0}^{n-1}(j+1)\left(\frac{1}{3}\right)^{j+2}.$$

۳. اگر X دارای توزیع توانی با تابع توزیع $F(x)=[\frac{x}{\alpha}]^\beta,\ 0<x<\alpha,\ \beta>0$ باشد، به‌دست می‌آوریم

$$I^w(F_{L_n},F)=\alpha^2\sum_{j=0}^{n-1}(j+1)\frac{\beta^{j+1}}{(2+\beta)^{j+2}}.$$

گزاره ۱.۲.۴. فرض کنید X یک متغیر تصادفی نامنفی با تابع توزیع F باشد. در این‌صورت

$$I^w(F_{L_n},F)=\sum_{j=0}^{n-1}\frac{1}{j!}\int_0^\infty\tilde{\lambda}(z)\left[\int_0^z x\,[-\log F(x)]^j\,F(x)dx\right]dz. \tag{۵.۴}$$

برهان. با توجه به (۴.۴) و رابطه‌ی $-\log F(x) = \int_x^\infty \tilde{\lambda}(z)dz$ داریم

$$I^w(F_{L_n}, F) = \sum_{j=0}^{n-1} \int_0^\infty x \frac{[-\log F(x)]^{j+1}}{j!} F(x)dx$$
$$= \sum_{j=0}^{n-1} \int_0^\infty \left[\int_x^\infty \tilde{\lambda}(z)dz\right] x \frac{[-\log F(x)]^j}{j!} F(x)dx$$
$$= \sum_{j=0}^{n-1} \frac{1}{j!} \int_0^\infty \tilde{\lambda}(z) \left[\int_0^z x\,[-\log F(x)]^j\, F(x)dx\right] dz.$$

بنابراین اثبات کامل می‌شود. □

تابع میانگین زمان غیرفعال وزنی یک متغیر تصادفی نامنفی به‌صورت زیر است

$$\mu^w(t) = \frac{\int_0^t xF(x)dx}{tF(t)}, \quad t > 0.$$

حال، WMIT L_n برابر است با

$$\mu_n^w(t) = \frac{\sum_{j=0}^{n-1} \frac{1}{j!} \int_0^t xF(x)[-\log F(x)]^j dx}{t\sum_{j=0}^{n-1} \frac{1}{j!} F(t)[-\log F(t)]^j}. \tag{۶.۴}$$

توجه کنید که $\mu_n^w(t)$ مشابه میانگین زمان انتظار باقیمانده است که در قابلیت اعتماد و تحلیل بقا به‌کار می‌رود (برای جزییات بیشتر بدیر و راکاب[1]، ۲۰۱۲ را ببینید).

گزاره ۲.۲.۴. فرض کنید X یک متغیر تصادفی نامنفی با تابع توزیع F باشد. آن‌گاه داریم

$$I^w(F_{L_n}, F) = \sum_{j=0}^{n-1} \mathbb{E}_{L_{j+1}}\left[X\mu_n^w(X)\right].$$

برهان. از رابطه‌ی (۶.۴) و گزاره‌ی ۱.۲.۴ به‌دست می‌آوریم

$$\begin{aligned} I^w(F_{L_n}, F) &= \int_0^\infty \tilde{\lambda}(z)\left[\sum_{j=0}^{n-1}\frac{1}{j!}\left[\int_0^z x[-\log F(x)]^j F(x)dx\right]\right]dz \\ &= \int_0^{+\infty}\sum_{j=0}^{n-1} z\mu_n^w(z) f_{L_{j+1}}(z)dz \\ &= \sum_{j=0}^{n-1}\int_0^{+\infty} z\mu_n^w(z) f_{L_{j+1}}(z)dz \\ &= \sum_{j=0}^{n-1} \mathbb{E}_{L_{j+1}}\left[X\mu_n^w(X)\right]. \end{aligned}$$

نتیجه به‌راحتی به‌دست می‌آید. □

[1] Bdair and Raqab

گزاره ۳.۲.۴. فرض کنید X یک متغیر تصادفی نامنفی و پیوسته باشد با $I^w(F_{L_n}, F) < \infty$، برای هر $n \geq ۱$. آن‌گاه داریم

$$I^w(F_{L_n}, F) = \sum_{j=۰}^{n-۱} \frac{۱}{j!} \mathbb{E}\left(\tilde{h}^w_{j+۱}(T)\right), \tag{۷.۴}$$

که $\tilde{h}^w_{j+۱}(t) = \int_t^\infty x\left[-\log F(x)\right]^{j+۱} dx$.

برهان. با استفاده از (۴.۴) و قضیه‌ی فوبینی به‌دست می‌آید

$$\begin{aligned} I^w(F_{L_n}, F) &= \sum_{j=۰}^{n-۱} \int_۰^\infty x \frac{\left[-\log F(x)\right]^{j+۱}}{j!} \left[\int_۰^x f(t)dt\right] dx \\ &= \sum_{j=۰}^{n-۱} \int_۰^\infty \frac{f(t)}{j!} \left[\int_t^\infty x\left[-\log F(x)\right]^{j+۱} dx\right] dt \\ &= \sum_{j=۰}^{n-۱} \frac{۱}{j!} \mathbb{E}\left[\tilde{h}^w_{j+۱}(T)\right]. \end{aligned}$$

□

ملاحظه ۱.۲.۴. فرض کنید X یک متغیر تصادفی متقارن نسبت به امید متناهی $\mu = E(X)$ باشد. آن‌گاه

$$I^w(F_{L_n}, F) = \bar{I}^w(\bar{F}_{R_n}, \bar{F}) - ۲\mu \bar{I}(\bar{F}_{R_n}, \bar{F}),$$

که $\bar{I}^w(\bar{F}_{R_n}, \bar{F})$ عدم دقت باقیمانده‌ی تجمعی وزنی بین $\bar{F}_{R_n}$ (تابع بقای nـامین رکورد بالا R_n) و $\bar{F}$ است.

اکنون می‌توانیم ویژگی مهمی از اندازه عدم دقت را با استفاده از دو مفهوم زیر اثبات کنیم.

۱. گفته می‌شود متغیر تصادفی نامنفی X دارای نرخ خطر معکوس نزولی در میانگین (DRHRA) است اگر $\frac{\tilde{\lambda}(x)}{x}$ در x نزولی باشد.

۲. گفته می‌شود متغیر تصادفی نامنفی X دارای نرخ خطر نزولی در میانگین (DHRA) است اگر $\frac{\lambda(x)}{x}$ در x نزولی باشد.

قضیه ۱.۲.۴. فرض کنید که متغیر تصادفی نامنفی X DRHRA است، آن‌گاه

$$I^w(F_{L_{n+۱}}, F) - I^w(F_{L_n}, F) \leq \sum_{i=۱}^{n+۱} \mathbb{E}_{L_i}\left[\frac{X}{\tilde{\lambda}(x)}\right]. \tag{۸.۴}$$

برهان. فرض کنید $f_{L_n}(x)$ تابع چگالی nـامین رکورد پایین باشد. آن‌گاه، نسبت $\frac{f_{Ln}(x)}{f_{L_{n+۱}}(x)} = \frac{-n}{\log F(x)}$ در x صعودی است. بنابراین $X_{n+۱} \leq^{lr} X_n$ که نتیجه می‌دهد $X_{n+۱} \leq^{st} X_n$، یعنی $\bar{F}_{n+۱}(x) \leq \bar{F}_n(x)$ (برای جزییات بیشتر شیکد و شانتی‌کومار، ۲۰۰۷ فصل ۱ را ببینید). این

معادل است (شیکد و شانتی‌کومار، ۲۰۰۷ صفحه ۴ را ببینید) با اینکه برای تمام توابع صعودی ϕ داشته باشیم

$$\mathbb{E}(\phi(X_{n+1})) \leq \mathbb{E}(\phi(X_n)),$$

به‌طوری که امیدها موجود باشند. بنابراین اگر X DRHRA و $\tilde{\lambda}(x)$ نرخ خطر معکوس باشد، آن‌گاه $\frac{x}{\tilde{\lambda}(x)}$ در x صعودی است. از (۴.۴) داریم

$$\begin{aligned}
I^w(F_{L_{n+1}}, F) &= \sum_{j=0}^{n}(j+1)\mathbb{E}_{L_{j+2}}\left[\frac{X}{\tilde{\lambda}(X)}\right] \\
&\leq \sum_{j=0}^{n}(j+1)\mathbb{E}_{L_{j+1}}\left[\frac{X}{\tilde{\lambda}(X)}\right] \\
&= \sum_{i=-1}^{n-1}(i+2)\mathbb{E}_{L_{i+2}}\left[\frac{X}{\tilde{\lambda}(X)}\right] \\
&= \sum_{i=0}^{n-1}(i+2)\mathbb{E}_{L_{i+2}}\left[\frac{X}{\tilde{\lambda}(X)}\right] + \mathbb{E}_{L_1}\left[\frac{X}{\tilde{\lambda}(X)}\right] \\
&= I^w(F_{L_n}, F) + \sum_{i=1}^{n+1}\mathbb{E}_{L_i}\left[\frac{X}{\tilde{\lambda}(X)}\right].
\end{aligned}$$

بنابراین اثبات کامل شده است. □

گزاره ۴.۲.۴. فرض کنید X یک متغیر تصادفی نامنفی با تابع توزیع کاملاً پیوسته‌ی $F(x)$ باشد. در این‌صورت برای $n = 1, 2, \dots$ داریم

$$I^w(F_{L_n}, F) \geq \sum_{j=0}^{n-1}\sum_{i=0}^{j+1}\frac{(-1)^i(j+1)}{i!(j+1-i)!}\int_0^{\infty} x[F(x)]^{i+1}dx.$$

برهان. از آنجا که $-\log F(x) \geq 1 - F(x)$، با یادآوری (۴.۴) اثبات به‌دست می‌آید. □

گزاره ۵.۲.۴. فرض کنید X یک متغیر تصادفی نامنفی با تابع توزیع کاملاً پیوسته‌ی $F(x)$ باشد. آن‌گاه برای $n = 1, 2, \dots$ داریم

$$I^w(F_{L_n}, F) \leq \sum_{j=0}^{n-1}\frac{1}{j!}\int_0^{\infty} x[-\log F(x)]^{j+1}dx.$$

فرض کنید که $\tilde{X}_\theta$ یک متغیر تصادفی نامنفی و کاملاً پیوسته با تابع توزیع $H_\theta(x) = [F(x)]^\theta$ و $x \geq 0$ را نشان دهد. اکنون عدم دقت تجمعی بین H_{L_n} و H را به‌صورت زیر به‌دست می‌آوریم:

$$\begin{aligned}
I^w(H_{L_n}, H_\theta) &= -\int_0^{+\infty} xH_{L_n}(x)\log\left(H_\theta(x)\right)dx \\
&= \sum_{j=0}^{n-1}\theta^{j+1}\int_0^{+\infty} x\frac{[-\log F(x)]^{j+1}}{j!}[F(x)]^\theta dx. \qquad (۹.۴)
\end{aligned}$$

گزاره ۶.۲.۴. اگر $\theta \geq ۱$، آن‌گاه برای هر $n = ۱, ۲, \ldots$ داریم

$$I^w(H_{L_n}, H_\theta) = \sum_{j=۰}^{n-۱}(j+۱)\mathcal{CE}^w_{j+۱}(\tilde{X}_\theta) \leq \sum_{j=۰}^{n-۱}\theta^{j+۱}(j+۱)\mathcal{CE}^w_{j+۱}(X), \tag{۱۰.۴}$$

که

$$\mathcal{CE}^w_{j+۱}(X) = \int_۰^\infty x\frac{[-\log F(x)]^{j+۱}}{(j+۱)!}F(x)dx, \tag{۱۱.۴}$$

آنتروپی تجمعی تعمیم‌یافته‌ی وزنی (WGCE) است که توسط کایال و مهارانا[۲] (۲۰۱۸) معرفی شده است.

برهان. فرض کنید که $\theta \geq ۱$، آن‌گاه واضح است که $[F(x)]^\theta \leq F(x)$ و بنابراین داریم

$$I^w(H_{L_n}, H_\theta) = \sum_{j=۰}^{n-۱}(j+۱)\mathcal{CE}^w_{j+۱}(\tilde{X}_\theta) \leq \sum_{j=۰}^{n-۱}\theta^{j+۱}(j+۱)\mathcal{CE}^w_{j+۱}(X).$$

□

گزاره ۷.۲.۴. فرض کنید X یک متغیر تصادفی نامنفی با تابع توزیع F باشد. آن‌گاه یک عبارت تحلیلی برای $I^w(F_{L_n}, F)$ به‌صورت زیر داده می‌شود

$$I^w(F_{L_n}, F) = \sum_{j=۰}^{n-۱}\int_۰^\infty x\frac{[-\log F(x)]^{j+۱}}{j!}F(x)dx = \sum_{j=۰}^{n-۱}(j+۱)\mathcal{CE}^w_{j+۱}(X). \tag{۱۲.۴}$$

گزاره ۸.۲.۴. فرض کنید $a, b > ۰$. برای $n = ۱, ۲, \ldots$ رابطه‌ی زیر برقرار می‌شود

$$I^w(F_{aL_n+b}, F_{aX+b}) = a^۲ I^w(F_{L_n}, F) + abI(F_{L_n}, F). \tag{۱۳.۴}$$

برهان. از (۱۲.۴) داریم

$$\begin{aligned} I^w(F_{aL_n+b}, F_{aX+b}) &= \sum_{j=۰}^{n-۱}(j+۱)\mathcal{CE}^w_{j+۱}(aX+b) \\ &= a^۲\sum_{j=۰}^{n-۱}(j+۱)\mathcal{CE}^w_{j+۱}(X) + ab\sum_{j=۰}^{n-۱}(j+۱)\mathcal{CE}_{j+۱}(X) \\ &= a^۲ I^w(F_{L_n}, F) + abI(F_{L_n}, F). \end{aligned}$$

اثبات کامل است. □

اخیراً کالی و همکاران (۲۰۱۸) CPI تعمیم یافته از مرتبه‌ی m را به‌صورت زیر تعریف کردند

$$I_m(F, G) = \frac{۱}{m!}\int_۰^{+\infty} F(x)[-\log G(x)]^m dx. \tag{۱۴.۴}$$

[۲] Kayal and Moharana

در مقایسه با اندازه‌ی تعریف شده در رابطه‌ی (۱۴.۴)، اکنون ما CPI تعمیم یافته‌ی وزنی از مرتبه‌ی m را به‌صورت زیر معرفی می‌کنیم

$$I_m^w(F,G)=\frac{1}{m!}\int_0^{+\infty}xF(x)[-\log G(x)]^m dx. \qquad (15.4)$$

ملاحظه ۲.۲.۴. فرض کنید X یک متغیر تصادفی نامنفی وکاملاً پیوسته با تابع توزیع F باشد. آن‌گاه WGCPI از مرتبه‌ی m بین F_{L_n} و F برابر است با

$$\begin{aligned} I_m^w(F_{L_n},F) &= \frac{1}{m!}\int_0^{\infty}xF_{L_n}(x)[-\log F(x)]^m dx \\ &= \sum_{j=0}^{n-1}\binom{m+j}{j}\mathcal{CE}_{m+j}^w(X). \end{aligned} \qquad (16.4)$$

۳.۴ عدم دقت گذشته‌ی تجمعی وزنی پویا برای L_n

در این بخش ما نسخه‌ی پویای $I^w(F_{L_n},F)$ را بررسی می‌کنیم. اگر یک سیستم که در زمان ۰ شروع به‌کار کرده است فقط در زمان‌های بازرسی تعیین شده مشاهده شود، و در زمان t از کار افتاده پیدا شود، آن‌گاه اندازه عدم دقت تجمعی پویا را به‌صورت زیر مطرح می‌کنیم

$$\begin{aligned} I^w(F_{L_n},F;t) &= -\int_0^t x\frac{F_{L_n}(x)}{F_{L_n}(t)}\log\left(\frac{F(x)}{F(t)}\right)dx \\ &= \log F(t)\mu_n^w(t)-\int_0^t x\frac{F_{L_n}(x)}{F_{L_n}(t)}\log\left(F(x)\right)dx \\ &= \log F(t)\mu_n^w(t)+\frac{1}{F_{L_n}(t)}\sum_{j=0}^{n-1}\int_0^t x\frac{[-\log F(x)]^{j+1}}{j!}F(x)dx. \end{aligned} \qquad (17.4)$$

توجه داشته باشید که $\lim_{t\to\infty}I^w(F_{L_n},F;t)=I^w(F_{L_n},F)$. از آنجا که $\log F(t)\leq 0$ برای $t\geq 0$، داریم

$$\begin{aligned} I^w(F_{L_n},F;t) &\leq \frac{1}{F_{L_n}(t)}\sum_{j=0}^{n-1}\int_0^t x\frac{[-\log F(x)]^{j+1}}{j!}F(x)dx \\ &\leq \frac{1}{F_{L_n}(t)}\sum_{j=0}^{n-1}\int_0^{+\infty} x\frac{[-\log F(x)]^{j+1}}{j!}F(x)dx=\frac{I^w(F_{L_n},F)}{F_{L_n}(t)}. \end{aligned}$$

در قضیه‌ی بعد ثابت می‌کنیم که $I^w(F_{L_n},F;t)$ تابع توزیع را به‌طور یکتا مشخص می‌کند.

قضیه ۱.۳.۴. فرض کنید X یک متغیر تصادفی پیوسته و نامنفی با تابع توزیع $F(.)$ باشد. فرض کنید عدم دقت تجمعی وزنی پویا مربوط به nـامین رکورد پایین با $I^w(F_{L_n},F;t)<\infty$ و $t\geq 0$ نشان داده شود. آن‌گاه $I^w(F_{L_n},F;t)$ تابع توزیع را مشخصه‌سازی می‌کند.

برهان. با توجه به (۱۷.۴) داریم

$$I^w(F_{L_n},F;t)=\log F(t)\mu_n^w(t)+\frac{1}{F_{L_n}(t)}\sum_{j=0}^{n-1}\int_0^t x\frac{[-\log F(x)]^{j+1}}{j!}F(x)dx. \qquad (۱۸.۴)$$

با مشتق‌گیری از دو طرف رابطه‌ی (۱۸.۴) نسبت به t به‌دست می‌آوریم :

$$\begin{aligned}\frac{\partial}{\partial t}[I^w(F_{L_n},F;t)] &= \tilde{\lambda}_F(t)\mu_n^w(t)-\tilde{\lambda}_{F_{L_n}}(t)I^w(F_{L_n},F;t)\\ &= \tilde{\lambda}_F(t)\mu_n^w(t)-c(t)\tilde{\lambda}_F(t)I^w(F_{L_n},F;t)\\ &= \tilde{\lambda}_F(t)\left[\mu_n^w(t)-c(t)I^w(F_{L_n},F;t)\right].\end{aligned}$$

مشتق‌گیری مجدد نسبت به t نتیجه می‌دهد

$$\acute{\tilde{\lambda}}_F(t)=\frac{(\tilde{\lambda}_F(t))^2\left(\acute{c}(t)I^w(F_{L_n},F;t)+c(t)\acute{I}^w(F_{L_n},F;t)-t+c(t)\tilde{\lambda}_F(t)\mu_n^w(t)\right)}{\acute{I}^w(F_{L_n},F;t)}. \qquad (۱۹.۴)$$

فرض کنید دو تابع F و F^* وجود دارند به‌طوری‌که

$$I^w(F_{L_n},F;t)=I^w(F^*_{L_n},F^*;t)=z(t).$$

آن‌گاه از (۱۹.۴) برای تمام مقادیر t به‌دست می‌آوریم

$$\acute{\tilde{\lambda}}_F(t)=\varphi(t,\tilde{\lambda}_F(t)),\quad \acute{\tilde{\lambda}}_{F^*}(t)=\varphi(t,\tilde{\lambda}_{F^*}(t)),$$

که

$$\varphi(t,y)=\frac{y^2\left[\acute{c}(t)z(t)+c(t)\left(\acute{z}(t)+ys(t)\right)-t\right]}{\acute{z}(t)},$$

و $s(t)=\mu_n^w(t)$. با استفاده از قضیه‌ی ۱.۲ و لم ۲.۲ از گوپتا و کرمانی (۲۰۰۸) داریم $\tilde{\lambda}_F(t)=\tilde{\lambda}_{F^*}(t)$، برای هر t. از آنجا که تابع نرخ خطر معکوس تابع توزیع را به‌طور یکتا مشخصه‌سازی می‌کند، اثبات کامل می‌شود. □

۴.۴ برآورد عدم دقت گذشته‌ی تجمعی وزنی برای L_n

در این بخش مسئله‌ی برآورد اندازه عدم دقت تجمعی وزنی را با استفاده از میانگین عدم دقت تجمعی وزنی تجربی عنوان می‌کنیم. فرض کنید $X_1,X_2,...,X_m$ یک نمونه‌ی تصادفی به حجم m از تابع توزیع تجمعی پیوسته‌ی $F(x)$ باشد. آن‌گاه طبق (۱۲.۴) اندازه عدم دقت تجمعی تجربی به‌صورت زیر تعریف می‌شود

$$\hat{I}^w(F_{L_n},F)=\sum_{j=0}^{n-1}\int_0^\infty x\frac{\left[-\log\hat{F}_m(x)\right]^{j+1}}{j!}\hat{F}_m(x)dx=\sum_{j=0}^{n-1}(j+1)\mathcal{CE}^w_{j+1}(\hat{F}_m). \qquad (۲۰.۴)$$

اگر $X_{(۱)} \leq X_{(۲)} \leq ... \leq X_{(m)}$ آماره‌های ترتیبی نمونه را نشان دهد، آن‌گاه (۲۰.۴) می‌تواند به‌صورت زیر نوشته شود

$$\hat{I}^w(F_{L_n}, F) = \sum_{j=۰}^{n-۱}\sum_{k=۱}^{m-۱}\int_{X_{(k)}}^{X_{(k+۱)}} x\frac{\left[-\log \hat{F}_m(x)\right]^{j+۱}}{j!}\hat{F}_m(x)dx. \quad (۲۱.۴)$$

علاوه بر این،

$$\hat{F}_m(x) = \begin{cases} ۰, & x < X_{(۱)}, \\ \frac{k}{m}, & X_{(k)} \leq x \leq X_{(k+۱)}, \quad k = ۱, ۲, ..., j \\ ۱, & x > X_{(k+۱)}. \end{cases}$$

بنابراین (۲۱.۴) می‌تواند به‌صورت زیر نوشته شود

$$\hat{I}^w(F_{L_n}, F) = \sum_{j=۰}^{n-۱}\sum_{k=۱}^{m-۱}\frac{۱}{j!}U_k\frac{k}{m}\left(-\log\frac{k}{m}\right)^{j+۱}, \quad (۲۲.۴)$$

که $U_k = \frac{X^۲_{(k+۱)} - X^۲_{(k)}}{۲}, k = ۱, ۲, ..., m-۱$.

مثال ۱.۴.۴. نمونه‌ی تصادفی $X_۱, X_۲, ..., X_m$ را از توزیع وایبل با تابع چگالی زیر در نظر بگیرید

$$f(x) = ۲\lambda\exp(-\lambda x^۲).$$

آن‌گاه $Y_k = X_k^۲$ دارای توزیع نمایی با میانگین $\frac{۱}{\lambda}$ است. در این حالت، متغیر تصادفی $۲U_k = X^۲_{(k+۱)} - X^۲_{(k)}$ مستقل و دارای توزیع نمایی با میانگین $\frac{۱}{\lambda(m-k)}$ است (برای جزئیات بیشتر پیک، ۱۹۶۵ را ببینید). حال از (۲۲.۴) به‌دست می‌آوریم

$$\mathbb{E}[\hat{I}^w(F_{L_n}, F)] = \sum_{j=۰}^{n-۱}\sum_{k=۱}^{m-۱}\frac{k}{۲\lambda j!(m-k)m}\left(-\log\frac{k}{m}\right)^{j+۱}, \quad (۲۳.۴)$$

و

$$Var[\hat{I}^w(F_{L_n}, F)] = \sum_{j=۰}^{n-۱}\sum_{k=۱}^{m-۱}\frac{k^۲}{۴\lambda^۲(j!)^۲(m-k)^۲m^۲}\left(-\log\frac{k}{m}\right)^{۲(j+۱)}. \quad (۲۴.۴)$$

مقادیر $\mathbb{E}[\hat{I}^w(F_{L_n}, F)]$ و $Var[\hat{I}^w(F_{L_n}, F)]$ را برای حجم نمونه‌ی $m = ۱۰, ۱۵, ۲۰$، $\lambda = ۰/۵, ۱, ۲$ و $n = ۲, ۳, ۴, ۵$ در جدول ۱.۴ محاسبه کردیم. به آسانی می‌توان دید که $\mathbb{E}[\hat{I}^w(F_{L_n}, F)]$ و $Var[\hat{I}^w(F_{L_n}, F)]$ در m نزولی است. همچنین $\lim_{m\to\infty} Var[\hat{I}^w(F_{L_n}, F)] = ۰$.

مثال ۲.۴.۴. فرض کنید $X_۱, X_۲, ..., X_m$ یک نمونه‌ی تصادفی از جمعیتی با تابع چگالی $f(x) = ۲x$ و $۰ < x < ۱$ باشد. در این‌صورت $۲U_k$ مستقل و دارای توزیع بتا با پارامتر ۱ و m است (برای جزئیات بیشتر پیک، ۱۹۶۵ را ببینید). حال از (۲۲.۴) به‌دست می‌آوریم

$$\mathbb{E}[\hat{I}^w(F_{L_n}, F)] = \sum_{j=۰}^{n-۱}\sum_{k=۱}^{m-۱}\frac{k}{۲j!(m+۱)m}\left(-\log\frac{k}{m}\right)^{j+۱}, \quad (۲۵.۴)$$

و

$$Var[\hat{I}^w(F_{L_n},F)] = \sum_{j=0}^{n-1}\sum_{k=1}^{m-1} \frac{k^2}{4(j!)^2(m+1)^2(m+2)m}\left(-\log\frac{k}{m}\right)^{2(j+1)}. \qquad (۲۶.۴)$$

ما مقادیر $\mathbb{E}[\hat{I}^w(F_{L_n},F)]$ و $Var[\hat{I}^w(F_{L_n},F)]$ را برای حجم‌های نمونه‌ی $m = 10, 15, 20$ و $n = 2, 3, 4, 5$ در جدول ۲.۴ محاسبه کرده‌ایم. به آسانی می‌توان دید که $Var[\hat{I}^w(F_{L_n},F)]$ در m نزولی است و $\lim_{m\to\infty} Var[\hat{I}^w(F_{L_n},F)] = 0$.

جدول ۱.۴: مقادیر عددی $\mathbb{E}[\hat{I}^w(F_{L_n},F)]$ و $Var[\hat{I}^w(F_{L_n},F)]$ برای توزیع وایبل

	$\mathbb{E}[\hat{I}^w(F_{L_n},F)]$											
λ	۰.۵	۱	۲	۰.۵	۱	۲	۰.۵	۱	۲	۰.۵	۱	۲
m	$n=2$			$n=3$			$n=4$			$n=5$		
۱۰	۰.۱۹۶	۰.۰۹۸	۰.۰۴۹	۰.۲۳۹	۰.۱۲۰	۰.۰۶۰	۰.۲۶۱	۰.۱۳۱	۰.۰۶۵	۰.۲۷۱	۰.۱۳۵	۰.۰۶۸
۱۵	۰.۱۳۴	۰.۰۶۷	۰.۰۳۳	۰.۱۶۵	۰.۰۸۲	۰.۰۴۱	۰.۱۸۱	۰.۰۹۰	۰.۰۴۵	۰.۱۸۹	۰.۰۹۴	۰.۰۴۷
۲۰	۰.۱۰۲	۰.۰۵۱	۰.۰۲۵	۰.۱۲۵	۰.۰۶۳	۰.۰۳۱	۰.۱۳۸	۰.۰۶۹	۰.۰۳۴	۰.۱۴۵	۰.۰۷۲	۰.۰۳۶

	$Var[\hat{I}^w(F_{L_n},F)]$											
λ	۰.۵	۱	۲	۰.۵	۱	۲	۰.۵	۱	۲	۰.۵	۱	۲
m	$n=2$			$n=3$			$n=4$			$n=5$		
۱۰	۰.۰۶۳	۰.۰۱۶	۰.۰۰۴	۰.۰۷۲	۰.۰۱۸	۰.۰۰۴	۰.۰۷۶	۰.۰۱۹	۰.۰۰۵	۰.۰۷۷	۰.۰۱۹	۰.۰۰۵
۱۵	۰.۰۴۳	۰.۰۱۱	۰.۰۰۲	۰.۰۵۰	۰.۰۱۲	۰.۰۰۳	۰.۰۵۳	۰.۰۱۳	۰.۰۰۳	۰.۰۵۴	۰.۰۱۴	۰.۰۰۳
۲۰	۰.۰۳۳	۰.۰۰۸	۰.۰۰۲	۰.۰۳۸	۰.۰۰۹	۰.۰۰۲	۰.۰۴۱	۰.۰۱۰	۰.۰۰۲	۰.۰۴۲	۰.۰۱۰	۰.۰۰۳

جدول ۲.۴: مقادیر عددی $\mathbb{E}[\hat{I}^w(F_{L_n},F)]$ و $Var[\hat{I}^w(F_{L_n},F)]$ برای توزیع بتا

	$\mathbb{E}[\hat{I}^w(F_{L_n},F)]$				$Var[\hat{I}^w(F_{L_n},F)]$			
m	n=۲	n=۳	n=۴	n=۵	n=۲	n=۳	n=۴	n=۵
۱۰	۰.۲۱۹	۰.۲۹۱	۰.۳۳۰	۰.۳۴۹	۰.۰۰۳	۰.۰۰۴	۰.۰۰۴	۰.۰۰۵
۱۵	۰.۲۳۰	۰.۳۰۹	۰.۳۵۶	۰.۳۸۰	۰.۰۰۲	۰.۰۰۳	۰.۰۰۴	۰.۰۰۴
۲۰	۰.۲۳۵	۰.۳۱۹	۰.۳۷۰	۰.۳۹۷	۰.۰۰۱	۰.۰۰۲	۰.۰۰۳	۰.۰۰۳

قضیه ۱.۴.۴. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته با $I^w(F_{L_n},F) < \infty$ برای هر $n \geq 1$ باشد. آن‌گاه داریم

$$\hat{I}^w(F_{L_n},F) \longrightarrow I^w(F_{L_n},F) \quad a.s.$$

برهان. از (۱۲.۴) داریم

$$\hat{I}^w(F_{L_n},F) = \sum_{j=0}^{n-1}(j+1)\mathcal{CE}^w_{j+1}(\hat{F}_m), \qquad (۲۷.۴)$$

که

$$\mathcal{CE}^w_{j+1}(\hat{F}_m) = \int_0^\infty x\frac{(-\log\hat{F}_m(x))^{j+1}}{(j+1)!}\hat{F}_m(x)dx$$

اکنون می‌توان به‌دست آورد

$$\begin{aligned}\frac{(j+1)!\mathcal{CE}^w_{j+1}(\hat{F}_m)}{(-1)^{j+1}} &= \int_0^\infty x(log\hat{F}_m(x))^{j+1}\hat{F}_m(x)dx \\ &= \int_0^1 x(log\hat{F}_m(x))^{j+1}\hat{F}_m(x)dx + \int_1^\infty x(log\hat{F}_m(x))^{j+1}\hat{F}_m(x)dx \\ &=: W_1 + W_2, \end{aligned} \tag{۲۸.۴}$$

که

$$W_1 = \int_0^1 x(\log\hat{F}_m(x))^{j+1}\hat{F}_m(x)dx,$$
$$W_2 = \int_1^\infty x(\log\hat{F}_m(x))^{j+1}\hat{F}_m(x)dx.$$

با استفاده از قضیه همگرایی مغلوب و گلیونکو ـ کانتلی داریم

$$\int_0^1 x(log\hat{F}_m(x))^{j+1}\hat{F}_m(x)dx \longrightarrow \int_0^1 x(logF(x))^{j+1}F(x)dx \quad as\ m\to\infty. \tag{۲۹.۴}$$

با توجه به تعریف $\hat{F}_m(x)$ به‌دست می‌آید

$$x^p\hat{F}_m(x) \leq \frac{1}{m}\sum_{i=1}^m X_i^p.$$

علاوه بر این با استفاده از قانون قوی اعداد بزرگ داریم $\frac{1}{m}\sum_{i=1}^m X_i^p \longrightarrow \mathbb{E}(X^p)$ و $\sup_m(\frac{1}{m}\sum_{i=1}^m X_i^p) < \infty$، آنگاه $\hat{F}_m(x) \leq x^{-p}\left(\sup_m(\frac{1}{m}\sum_{i=1}^m X_i^p)\right) = Cx^{-p}$. حال با به‌کار بردن DCT داریم

$$\lim_{m\to\infty} W_2 = \int_1^\infty xF(x)\left(\log F(x)\right)^{j+1}dx. \tag{۳۰.۴}$$

در انتها، با استفاده از (۲۷.۴) و (۲۸.۴) نتیجه به‌دست می‌آید. □

۵.۴ عدم دقت باقیمانده‌ی تجمعی وزنی برای R_n

در این بخش ما WCRI بین $\bar{F}_{R_n}$ و $\bar{F}$ را مطرح می‌کنیم. برخی ویژگی‌های WCRI مانند اثر تبدیل خطی، روابط با دیگر توابع قابلیت اطمینان، کران‌ها و ترتیب‌های تصادفی را بررسی می‌کنیم.

تعریف ۱.۵.۴. فرض کنید X یک متغیر تصادفی کاملاً پیوسته و نامنفی با تابع بقای $\bar{F}$ باشد.

آنگاه، WCRI بین $\bar{F}_{R_n}$ و $\bar{F}$ را به‌صورت زیر تعریف می‌کنیم :

$$\begin{aligned} \bar{I}^w(\bar{F}_{R_n}, \bar{F}) &= -\int_0^{+\infty} x\bar{F}_{R_n}(x)\log\left(\bar{F}(x)\right)dx \\ &= \sum_{j=0}^{n-1}\int_0^{+\infty} x\frac{[-\log\bar{F}(x)]^{j+1}}{j!}\bar{F}(x)dx \\ &= \sum_{j=0}^{n-1}(j+1)\mathbb{E}_{R_{j+2}}\left(\frac{X}{\lambda(X)}\right), \end{aligned} \tag{۳۱.۴}$$

که $\lambda(x)=\frac{f(x)}{\bar{F}(x)}$ تابع نرخ خطر و R_{j+2} یک متغیر تصادفی با قابلیت اطمینان $\bar{F}_{R_{j+2}}$ است.

در مثال زیر، $\bar{I}^w(\bar{F}_{R_n}, \bar{F})$ را برای برخی توزیع‌های طول عمر خاص که در قابلیت اطمینان و آزمون بقا بسیار کاربرد دارند حساب می‌کنیم.

مثال ۱.۵.۴. (۱) اگر X دارای توزیع یکنواخت در $[0,\theta]$ باشد، آنگاه به آسانی می‌توان دید که $\bar{I}^w(\bar{F}_{R_n}, \bar{F})=\theta^2\sum_{j=0}^{n-1}\frac{3^{j+2}-2^{j+2}}{6^{j+2}}(j+1)$، برای تمام اعداد صحیح $n\geq 1$.

(۲) اگر X دارای توزیع وایبل با تابع بقای $\bar{F}(x)=e^{-\alpha x^\beta},\ x\geq 0,\ \alpha,\beta>0$ باشد، آنگاه برای تمام اعداد صحیح $n\geq 1$ داریم $\bar{I}^w(\bar{F}_{R_n}, \bar{F})=\frac{1}{\beta}\sum_{j=0}^{n-1}\frac{\alpha^{2(1+j-\frac{1}{\beta})}(j+\frac{2}{\beta})!}{j!}$.

(۳) اگر X دارای توزیع نمایی با میانگین $\frac{1}{\lambda}$ باشد، آنگاه $\bar{I}^w(\bar{F}_{R_n}, \bar{F})=\frac{n(n+1)(n+2)}{3\lambda^2}$.

گزاره ۱.۵.۴. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته با تابع بقای $\bar{F}$ باشد. آن‌گاه داریم

$$\bar{I}^w(\bar{F}_{R_n}, \bar{F})=\sum_{j=0}^{n-1}(j+1)x[\mu_{j+2}-\mu_{j+1}],$$

که در آن $\mu_n=\int_0^{+\infty}\bar{F}_{R_n}(x)dx$.

برهان. از (۳۱.۴) داریم

$$\begin{aligned} \bar{I}^w(\bar{F}_{R_n}, \bar{F}) &= \sum_{j=0}^{n-1}\int_0^{+\infty} x\frac{[-\log\bar{F}(x)]^{j+1}}{j!}\bar{F}(x)dx \\ &= \sum_{j=0}^{n-1}(j+1)\int_0^{+\infty} x[\bar{F}_{R_{j+2}}(x)-\bar{F}_{R_{j+1}}(x)]dx \\ &= \sum_{j=0}^{n-1}(j+1)x[\mu_{j+2}-\mu_{j+1}]. \end{aligned}$$

□

گزاره ۲.۵.۴. فرض کنید $a,b>0$. برای $n=1,2,\ldots$ رابطه‌ی زیر برقرار می‌شود

$$\bar{I}^w(\bar{F}_{aR_n+b}, \bar{F}_{aX+b})=a^2\bar{I}^w(\bar{F}_{R_n}, \bar{F})+ab\bar{I}(\bar{F}_{R_n}, \bar{F}).$$

برهان. از (۳۱.۴) و با توجه به اینکه $\bar{F}_{aX+b}(x)=\bar{F}(\frac{x-b}{a})$، داریم

$$\begin{aligned}\bar{I}^w(\bar{F}_{aR_n+b},\bar{F}_{aX+b}) &= -\int_0^{+\infty}\sum_{j=0}^{n-1} x\frac{[-\log\bar{F}_{aX+b}(x)]^{j+1}}{j!}\bar{F}_{aX+b}(x)dx\\ &= -\int_0^{+\infty}\sum_{j=0}^{n-1} x\frac{[-\log\bar{F}_X(\frac{x-b}{a})]^{j+1}}{j!}\bar{F}_X(\frac{x-b}{a})dx\\ &= -\int_0^{+\infty}\sum_{j=0}^{n-1} a(ay+b)\frac{[-\log\bar{F}_X(y)]^{j+1}}{j!}\bar{F}_X(y)dy\\ &= a^2\bar{I}^w(\bar{F}_{R_n},\bar{F})+ab\bar{I}(\bar{F}_{R_n},\bar{F}). \qquad (۳۲.۴)\end{aligned}$$

□

کید و همکاران[۳] (۲۰۱۶) تابع میانگین ترکیبی عمر باقیمانده[۴] (CMRL) از X را به‌صورت زیر مطرح کردند

$$m^c(t)=\frac{\int_t^{+\infty}x\bar{F}(x)dx}{t\bar{F}(t)},\quad t>0.$$

حال CMRL مربوط به R_n به‌دست می‌آید

$$m_n^c(t)=\frac{\sum_{j=0}^{n-1}\frac{1}{j!}\int_t^{+\infty}x\bar{F}(x)[-\log\bar{F}(x)]^j dx}{t\sum_{j=0}^{n-1}\frac{1}{j!}\bar{F}(t)[-\log\bar{F}(t)]^j}. \qquad (۳۳.۴)$$

گزاره ۳.۵.۴. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته با تابع بقای $\bar{F}$ باشد. آنگاه داریم

$$\bar{I}^w(\bar{F}_{R_n},\bar{F})=\sum_{j=0}^{n-1}\frac{1}{j!}\int_0^{\infty}\lambda(z)\left[\int_z^{\infty}x[-\log\bar{F}(x)]^j\bar{F}(x)dx\right]dz.$$

برهان. با توجه به (۳۱.۴) و این حقیقت که $-\log\bar{F}(x)=\int_0^x\lambda(z)dz$، داریم

$$\begin{aligned}\bar{I}^w(\bar{F}_{R_n},\bar{F}) &= \sum_{j=0}^{n-1}\int_0^{+\infty}x\frac{[-\log\bar{F}(x)]^{j+1}}{j!}\bar{F}(x)dx\\ &= \sum_{j=0}^{n-1}\int_0^{+\infty}\left[\int_0^x\lambda(z)dz\right]x\frac{[-\log\bar{F}(x)]^j}{j!}\bar{F}(x)dx\\ &= \sum_{j=0}^{n-1}\frac{1}{j!}\int_0^{+\infty}\lambda(z)\left[\int_z^{\infty}x[-\log\bar{F}(x)]^j\bar{F}(x)dx\right]dz.\end{aligned}$$

□

[۳] Kayid et al.

[۴] Combination mean residual life

گزاره ۴.۵.۴. فرض کنید X یک متغیر تصادفی نامنفی با تابع بقای $\bar{F}$ باشد. آنگاه داریم

$$\bar{I}^w(\bar{F}_{R_n}, \bar{F}) = \sum_{j=0}^{n-1} \mathbb{E}_{R_{j+1}}\left[X m_n^c(X)\right].$$

برهان. از (۳۳.۴) و با استفاده از گزاره‌ی (۳.۵.۴) به‌دست می‌آوریم

$$\begin{aligned}
\bar{I}^w(\bar{F}_{R_n}, \bar{F}) &= \int_0^\infty \lambda(z)\left[\sum_{j=0}^{n-1}\frac{1}{j!}\left[\int_z^\infty x[-\log\bar{F}(x)]^j\bar{F}(x)dx\right]\right]dz \\
&= \int_0^{+\infty}\sum_{j=0}^{n-1} z m_n^c(z) f_{R_{j+1}}(z)dz \\
&= \sum_{j=0}^{n-1}\int_0^{+\infty} z m_n^c(z) f_{R_{j+1}}(z)dz \\
&= \sum_{j=0}^{n-1}\mathbb{E}_{R_{j+1}}\left[X m_n^c(X)\right].
\end{aligned}$$

اثبات کامل می‌شود. □

گزاره ۵.۵.۴. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته با $\bar{I}^w(\bar{F}_{R_n}, \bar{F}) < \infty$، برای $n \geq 1$، باشد. آنگاه داریم

$$\bar{I}^w(\bar{F}_{R_n}, \bar{F}) = \sum_{j=0}^{n-1}\frac{1}{j!}\mathbb{E}\left[h_{j+1}^w(X)\right], \tag{۳۴.۴}$$

که

$$h_{j+1}^w(x) = \int_0^x z\left[-\log\bar{F}(z)\right]^{j+1}dz, \quad x \geq 0.$$

برهان. از (۳۱.۴) و با استفاده از قضیه فوبینی به‌دست می‌آید

$$\begin{aligned}
\bar{I}^w(\bar{F}_{R_n}, \bar{F}) &= \sum_{j=0}^{n-1}\int_0^\infty z\frac{[-\log\bar{F}(z)]^{j+1}}{j!}\bar{F}(z)dz \\
&= \sum_{j=0}^{n-1}\frac{1}{j!}\int_0^\infty\left[\int_z^\infty f(x)dx\right]z[-\log\bar{F}(z)]^{j+1}dz \\
&= \sum_{j=0}^{n-1}\frac{1}{j!}\int_0^\infty f(x)\left[\int_0^x z[-\log\bar{F}(z)]^{j+1}dz\right]dx = \sum_{j=0}^{n-1}\frac{1}{j!}\mathbb{E}\left[h_{j+1}^w(X)\right].
\end{aligned}$$

□

گزاره ۶.۵.۴. فرض کنید X یک متغیر تصادفی نامنفی و کاملاً پیوسته با تابع توزیع F باشد. آنگاه

$$\bar{I}^w(\bar{F}_{R_n}, \bar{F}) \geq \sum_{j=0}^{n-1}\frac{1}{j!}[\mathcal{E}^{w_{j+1}}(X)]^{j+1},$$

که $\mathcal{E}^{w_{j+1}}(X) = -\int_0^\infty x^{\left(\frac{1}{j+1}\right)}\bar{F}(x)\log\bar{F}(x)dx$.

برهان. از (۳۱.۴) داریم

$$\begin{aligned}\bar{I}^w(\bar{F}_{R_n},\bar{F}) &= \sum_{j=0}^{n-1}\frac{1}{j!}\int_0^{+\infty} x\bar{F}(x)[-\log\bar{F}(x)]^{j+1}dx\\ &\geq \sum_{j=0}^{n-1}\frac{1}{j!}\int_0^{+\infty}\left[x^{\left(\frac{1}{j+1}\right)}\bar{F}(x)[-\log\bar{F}(x)]\right]^{j+1}dx\\ &\geq \sum_{j=0}^{n-1}\frac{1}{j!}\left[-\int_0^{+\infty}x^{\left(\frac{1}{j+1}\right)}\bar{F}(x)\log\bar{F}(x)dx\right]^{j+1}.\end{aligned}$$

اثبات کامل می‌شود. □

در گزاره‌های بعدی برخی کران‌های بالا و پایین را برای $\bar{I}^w(\bar{F}_{R_n},\bar{F})$ نشان می‌دهیم.

گزاره ۷.۵.۴. فرض کنید X یک متغیر تصادفی نامنفی با تابع توزیع کاملاً پیوسته‌ی $F(x)$ باشد. آنگاه برای $n=1,2,\ldots$ داریم

$$\bar{I}^w(\bar{F}_{R_n},\bar{F}) \leq \sum_{j=0}^{n-1}\frac{1}{j!}\int_0^{\infty}x[-\log\bar{F}(x)]^{j+1}dx.$$

برهان. با استفاده از (۳۱.۴) به راحتی اثبات می‌شود. □

گزاره ۸.۵.۴. فرض کنید X یک متغیر تصادفی نامنفی با تابع بقای $\bar{F}(x)$ باشد. آنگاه برای $n=1,2,\ldots$ داریم

$$\bar{I}^w(\bar{F}_{R_n},\bar{F}) \geq \sum_{j=0}^{n-1}\sum_{i=0}^{j+1}\frac{(-1)^i(j+1)}{i!(j+1-i)!}\int_0^{\infty}x[\bar{F}(x)]^{i+1}dx.$$

برهان. از آنجا که $-\log\bar{F}(x)\geq 1-\bar{F}(x)$، با یادآوری رابطه‌ی (۳۱.۴) اثبات به‌دست می‌آید.
□

در ادامه، برخی نتایج برای $I^w(\bar{F}_{R_n},\bar{F})$ و ارتباط آن با مفهوم قابلیت اطمینان را به‌دست می‌آوریم.

گزاره ۹.۵.۴. اگر X DFRA باشد، آن‌گاه برای $n=1,2,\ldots$ داریم

$$\bar{I}^w(\bar{F}_{R_{n+1}},\bar{F})-\bar{I}^w(\bar{F}_{R_n},\bar{F}) \geq \sum_{i=1}^{n+1}\mathbb{E}_{R_i}\left[\frac{X}{\lambda(X)}\right]. \qquad (۳۵.۴)$$

برهان. فرض کنید که f_{R_n} تابع چگالی nـامین رکورد R_n باشد. آنگاه، نسبت $\frac{f_{R_{n+1}}(t)}{f_{R_n}(t)}=\frac{-\log\bar{F}(t)}{n}$ در t صعودی است. بنابراین $R_n\leq^{lr}R_{n+1}$ و این نتیجه می‌دهد که $R_n\leq^{st}R_{n+1}$، یعنی $\bar{F}_{R_n}\leq\bar{F}_{R_{n+1}}$ (برای جزییات بیشتر شیکد و شانتی‌کومار ۲۰۰۷، فصل ۱ را ببینید). در

نتیجه، اگر X DFRA و $\lambda(x)$ نرخ خطر باشد، آنگاه $\frac{x}{\lambda(x)}$ تابعی صعودی از x است. بنابراین از (۳۱.۴) داریم

$$\begin{aligned}
\bar{I}^w(\bar{F}_{R_{n+1}}, \bar{F}) &= \sum_{j=0}^{n}(j+1)\mathbb{E}_{R_{j+2}}\left[\frac{X}{\lambda(X)}\right] \\
&\geq \sum_{j=0}^{n}(j+1)\mathbb{E}_{R_{j+1}}\left[\frac{X}{\lambda(X)}\right] \\
&= \sum_{i=-1}^{n-1}(i+2)\mathbb{E}_{R_{i+2}}\left[\frac{X}{\lambda(X)}\right] \\
&= \sum_{i=0}^{n-1}(i+2)\mathbb{E}_{R_{i+2}}\left[\frac{X}{\lambda(X)}\right] + \mathbb{E}_{R_1}\left[\frac{X}{\lambda(X)}\right] \\
&= \sum_{i=0}^{n-1}(i+1)\mathbb{E}_{R_{i+2}}\left[\frac{X}{\lambda(X)}\right] + \sum_{i=0}^{n-1}\mathbb{E}_{R_{i+2}}\left[\frac{X}{\lambda(X)}\right] + \mathbb{E}_{R_1}\left[\frac{X}{\lambda(X)}\right] \\
&= \bar{I}^w(\bar{F}_{R_n}, \bar{F}) + \sum_{i=1}^{n+1}\mathbb{E}_{R_i}\left[\frac{X}{\lambda(X)}\right]. \qquad (۳۶.۴)
\end{aligned}$$

اثبات کامل است. □

گزاره ۱۰.۵.۴. اگر X دارای توزیع نمایی با میانگین $\mu = \frac{1}{\theta}$ باشد، از آنجا که نرخ خطر مقداری ثابت است، به‌دست می‌آوریم $\bar{I}^w(\bar{F}_{R_n}, \bar{F}) = \frac{n(n+1)(n+2)}{3}\mu^2$ که تابعی صعودی از n است.

گزاره ۱۱.۵.۴. فرض کنید X و Y دو متغیر تصادفی نامنفی با تابع بقای $\bar{F}(x)$ و $\bar{G}(x)$ باشند. اگر $X \leq^{hr} Y$ و X DFRA باشد، آن‌گاه

$$\bar{I}^w(\bar{F}_{R_n}, \bar{F}) \leq \bar{I}^w(\bar{G}_{\tilde{R}_n}, \bar{G}), \qquad (۳۷.۴)$$

برای $n = 1, 2, \ldots$.

برهان. می‌دانیم که $X \leq^{hr} Y$ نتیجه می‌دهد $X \leq^{st} Y$. بنابراین

$$\bar{F}_{R_{j+2}} \leq \bar{G}_{\tilde{R}_{j+2}},$$

که $\bar{G}_{\tilde{R}_{j+2}}$ تابع بقای $\tilde{R}_{j+2}$ است. در نتیجه برقرار می‌شود $R_{j+2} \leq^{st} \tilde{R}_{j+2}$. این معادل است با اینکه داشته باشیم

$$\mathbb{E}(\phi(R_{j+2})) \leq \mathbb{E}(\phi(\tilde{R}_{j+2})),$$

برای تمام توابع صعودی ϕ به‌شرطی که امید موجود باشد. بنابراین اگر فرض کنیم که X DFRA و $\lambda(x)$ نرخ شکست آن است، آنگاه $\frac{x}{\lambda(x)}$ صعودی است و داریم

$$\bar{I}^w(\bar{F}_{R_n}, \bar{F}) = \sum_{j=0}^{n-1}(j+1)\mathbb{E}_{R_{j+2}}\left(\frac{X}{\lambda_F(X)}\right) \leq \sum_{j=0}^{n-1}(j+1)\mathbb{E}_{\tilde{R}_{j+2}}\left(\frac{X}{\lambda_F(X)}\right).$$

از طرف دیگر، $X \leq^{hr} Y$ بر این دلالت دارد که توابع نرخ شکست مربوطه در رابطه‌ی $\lambda_F(x) \geq \lambda_G(y)$ صدق می‌کنند. در نتیجه داریم

$$\sum_{j=0}^{n-1}(j+1)\mathbb{E}_{\tilde{R}_{j+2}}\left(\frac{X}{\lambda_F(X)}\right) \leq \sum_{j=0}^{n-1}(j+1)\mathbb{E}_{\tilde{R}_{j+2}}\left(\frac{X}{\lambda_G(Y)}\right) = I^w(\bar{G}_{\tilde{R}_n}, \bar{G}).$$

بنابراین، با استفاده از هر دو عبارت به‌دست می‌آوریم $\bar{I}^w(\bar{F}_{R_n}, \bar{F}) \leq \bar{I}^w(\bar{G}_{\tilde{R}_n}, \bar{G})$. □

گزاره ۱۲.۵.۴. فرض کنید X و Y دو متغیر تصادفی نامنفی با تابع بقای $\bar{F}(x)$ و $\bar{G}(x)$ باشند. اگر $X \leq^{icx} Y$ آنگاه

$$\bar{I}^w(\bar{F}_{R_n}, \bar{F}) \leq \bar{I}^w(\bar{G}_{\tilde{R}_n}, \bar{G}).$$

برهان. از آنجا که $h^w_{j+1}(.)$ یک تابع محدب صعودی برای $j \geq 0$ است، از شیکد و شانتی‌کومار (۲۰۰۷) نتیجه می‌شود که $X \leq^{icx} Y$ به این معنی است که $h^w_{j+1}(X) \leq^{icx} h^w_{j+1}(Y)$. با یادآوری تعریف ترتیب محدب صعودی و گزاره‌ی ۵.۵.۴ اثبات کامل می‌شود. □

گزاره ۱۳.۵.۴. اگر X IFRA (DFRA) باشد، آنگاه برای $n = 1, 2, \dots$ داریم

$$\bar{I}^w(\bar{F}_{R_n}, \bar{F}) \leq (\geq) \sum_{j=0}^{n-1} \frac{1}{j!}\mathbb{E}\left[X^2\left(-\log \bar{F}(X)\right)^j\right]. \tag{۳۸.۴}$$

برهان. از (۳۱.۴) داریم

$$\bar{I}^w(\bar{F}_{R_n}, \bar{F}) = \sum_{j=0}^{n-1}\int_0^{+\infty} x\frac{[-\log \bar{F}(x)]^j}{j!}[-\log \bar{F}(x)]\bar{F}(x)dx. \tag{۳۹.۴}$$

حال از آنجا که X IFRA (DFRA) است، $\frac{-\log \bar{F}(x)}{x}$ نسبت به $x > 0$ صعودی (نزولی) است، که به این معنی است که

$$-\bar{F}(x)\log \bar{F}(x) \leq (\geq) x f(x), \quad x > 0. \tag{۴۰.۴}$$

بنابراین با توجه به (۳۹.۴) و (۴۰.۴) اثبات کامل می‌شود. □

گزاره ۱۴.۵.۴. فرض کنید X و Y دو متغیر تصادفی نامنفی با تابع بقای $\bar{F}(x)$ و $\bar{G}(x)$ باشند. اگر $X \leq^{hr} Y$، آنگاه برای $n = 1, 2, \dots$ برقرار می شود

$$\frac{\bar{I}^w(\bar{F}_{R_n}, \bar{F})}{\mathbb{E}(X)} \leq \frac{\bar{I}^w(\bar{G}_{R_n}, \bar{G})}{\mathbb{E}(Y)}.$$

برهان. با توجه به اینکه تابع $h^w_{j+1}(x) = \int_0^x z[-\log \bar{F}(z)]^{j+1}dz$ یک تابع محدب صعودی است، تحت فرض $X \leq^{hr} Y$، نتیجه می‌شود

$$\sum_{j=0}^{n-1}\frac{1}{j!}\left[\frac{\mathbb{E}\left[h^w_{j+1}(X)\right]}{\mathbb{E}(X)}\right] \leq \sum_{j=0}^{n-1}\frac{1}{j!}\left[\frac{\mathbb{E}\left[h^w_{j+1}(Y)\right]}{\mathbb{E}(Y)}\right].$$

بنابراین، با یادآوری (۳۱.۴) اثبات کامل می‌شود. □

گزاره ۱۵.۵.۴. (۱) فرض کنید X یک متغیر تصادفی پیوسته با تابع بقای $\bar{F}(.)$ باشد که در $[0, b]$ مقدار می‌گیرد و b متناهی است. آن‌گاه

$$\bar{I}^w(\bar{F}_{R_n}, \bar{F}) \leq b\bar{I}(\bar{F}_{R_n}, \bar{F}).$$

(۲) فرض کنید X یک متغیر تصادفی پیوسته و نامنفی با تابع بقای $\bar{F}(.)$ باشد که در $[a, \infty)$ مقدار می‌گیرد و $a > 0$ متناهی است. آن‌گاه

$$\bar{I}^w(\bar{F}_{R_n}, \bar{F}) \geq a\bar{I}(\bar{F}_{R_n}, \bar{F}).$$

فرض کنید X_θ^* یک متغیر تصادفی کاملاً پیوسته و نامنفی با تابع بقای $\bar{H}_\theta(x) = [\bar{F}(x)]^\theta$ و $x \geq 0$ را نشان دهد. این مدل به‌عنوان مدل نرخ خطر نسبی شناخته می‌شود. اکنون ما اندازه عدم دقت باقیمانده‌ی تجمعی وزنی را بین $\bar{H}_{R_n}$ و $\bar{H}$ به‌صورت زیر به‌دست می‌آوریم :

$$\begin{aligned} \bar{I}^w(\bar{H}_{R_n}, \bar{H}) &= -\int_0^{+\infty} x\bar{H}_{R_n}(x)\log\left(\bar{H}(x)\right)dx \\ &= \sum_{j=0}^{n-1}\theta^{j+1}\int_0^{+\infty} x\frac{[-\log \bar{F}(x)]^{j+1}}{j!}[\bar{F}(x)]^\theta dx. \end{aligned} \tag{۴۱.۴}$$

گزاره ۱۶.۵.۴. اگر $\theta \geq (\leq) 1$، آنگاه برای هر $n \geq 1$ داریم

$$\bar{I}^w(\bar{H}_{R_n}, \bar{H}) \leq (\geq) \sum_{j=0}^{n-1}(j+1)\theta^{j+1}\mathcal{E}^w_{j+1}(X),$$

که $\mathcal{E}^w_{j+1}(X)$ آنتروپی باقیمانده‌ی تجمعی تعمیم‌یافته‌ی وزنی از X است که توسط کایال (۲۰۱۸) به‌صورت زیر معرفی شده‌است

$$\mathcal{E}^w_{j+1}(X) = \int_0^{+\infty} x\frac{\bar{F}(x)[-\log \bar{F}(x)]^{j+1}}{(j+1)!}dx.$$

برهان. فرض کنید که $\theta \geq (\leq) 1$، آنگاه واضح است که $[\bar{F}(x)]^\theta \leq (\geq) \bar{F}(x)$ و بنابراین از (۴۱.۴) نتیجه می‌شود

$$\bar{I}^w(\bar{H}_{R_n}, \bar{H}) \leq (\geq) \sum_{j=0}^{n-1}(j+1)\theta^{j+1}\mathcal{E}^w_{j+1}(X).$$

□

گزاره ۱۷.۵.۴. فرض کنید X یک متغیر تصادفی نامنفی با تابع بقای $\bar{F}(.)$ باشد. آنگاه یک عبارت تحلیلی برای $\bar{I}^w(\bar{F}_{R_n}, \bar{F})$ به‌صورت زیر است

$$\bar{I}^w(\bar{F}_{R_n}, \bar{F}) = \sum_{j=0}^{n-1}\int_0^{\infty} x\frac{\left[-\log \bar{F}(x)\right]^{j+1}}{j!}\bar{F}(x)dx = \sum_{j=0}^{n-1}(j+1)\mathcal{E}^w_{j+1}(X). \tag{۴۲.۴}$$

قضیه ۱.۵.۴. $\bar{I}^w(\bar{F}_{R_n}, \bar{F}) = 0$، اگر و تنها اگر، X تباهیده باشد.

برهان. فرض کنید X در نقطه‌ی a تباهیده باشد. با توجه به تعریف تابع تباهیده و تعریف $\bar{I}^w(\bar{F}_{R_n}, \bar{F})$ واضح است که داریم $\bar{I}^w(\bar{F}_{R_n}, \bar{F}) = ۰$.
حال، فرض کنید که $I^w(\bar{F}_{R_n}, \bar{F}) = ۰$، یعنی

$$\mathcal{E}^w_{j+۱}(X) = \int_۰^\infty x\bar{F}(x)\left(-\log \bar{F}(x)\right)^{j+۱} dx = ۰. \qquad (۴۳.۴)$$

آنگاه با توجه به نامنفی بودن (۴۳.۴)، نتیجه می‌گیریم که $x\bar{F}(x)\left(-\log \bar{F}(x)\right)^{j+۱} = ۰$ برای تقریباً تمام $x \in \mathbb{R}^+$. بنابراین، $\bar{F}(x) = ۰ \; or \; ۱$ برای تقریباً هر $x \in \mathbb{R}^+$. □

ملاحظه ۱.۵.۴. فرض کنید X یک متغیر تصادفی کاملاً پیوسته و نامنفی با تابع بقای $\bar{F}(.)$ باشد. آنگاه در مقایسه با اندازه‌ی تعریف شده در رابطه‌ی (۱۶.۴)، WGCRI مرتبه m بین $\bar{F}_{R_n}$ و $\bar{F}$ به‌دست می‌آید

$$\begin{aligned} \bar{I}^w_m(\bar{F}_{R_n}, \bar{F}) &= \frac{۱}{m!}\int_۰^\infty x\bar{F}_{R_n}(x)[-\log \bar{F}(x)]^m dx \\ &= \sum_{j=۰}^{n-۱}\binom{m+j}{m}\mathcal{E}^w_{m+j}(X). \end{aligned}$$

در ادامه‌ی این بخش، نسخه‌ی پویای $\bar{I}^w(\bar{F}_{R_n}, \bar{F})$ را بررسی می‌کنیم. فرض کنید X طول عمر یک سیستم باشد تحت این شرط که سیستم تا سن t سالم بوده است. به‌طور مشابه، می‌توانیم نسخه‌ی پویای $\bar{I}^w(\bar{F}_{R_n}, \bar{F})$ را به‌صورت زیر در نظر بگیریم

$$\begin{aligned} \bar{I}^w(\bar{F}_{R_n}, \bar{F}; t) &= -\int_t^{+\infty} x\frac{\bar{F}_{R_n}(x)}{\bar{F}_{R_n}(t)}\log\left(\frac{\bar{F}(x)}{\bar{F}(t)}\right) dx \\ &= \log \bar{F}(t)m^c_n(t) - \int_t^{+\infty} x\frac{\bar{F}_{R_n}(x)}{\bar{F}_{R_n}(t)}\log\left(\bar{F}(x)\right) dx \\ &= \log \bar{F}(t)m^c_n(t) + \frac{۱}{\bar{F}_{R_n}(t)}\sum_{j=۰}^{n-۱}\int_t^{+\infty} x\frac{[-\log \bar{F}(x)]^{j+۱}}{j!}\bar{F}(x)dx. \end{aligned} \qquad (۴۴.۴)$$

توجه کنید که $\lim_{t\to۰}\bar{I}^w(\bar{F}_{R_n}, \bar{F}; t) = \bar{I}^w(\bar{F}_{R_n}, \bar{F})$. از آنجا که $\log \bar{F}(t) \leq ۰$ برای $t \geq ۰$، داریم

$$\begin{aligned} \bar{I}^w(\bar{F}_{R_n}, \bar{F}; t) &\leq \frac{۱}{\bar{F}_{R_n}(t)}\sum_{j=۰}^{n-۱}\int_t^{+\infty} x\frac{[-\log \bar{F}(x)]^{j+۱}}{j!}\bar{F}(x)dx \\ &\leq \frac{۱}{\bar{F}_{R_n}(t)}\sum_{j=۰}^{n-۱}\int_۰^{+\infty} x\frac{[-\log \bar{F}(x)]^{j+۱}}{j!}\bar{F}(x)dx = \frac{\bar{I}^w(\bar{F}_{R_n}, \bar{F})}{\bar{F}_{R_n}(t)}. \end{aligned}$$

قضیه ۲.۵.۴. فرض کنید X یک متغیر تصادفی پیوسته و نامنفی با تابع توزیع $F(.)$ باشد. فرض کنید عدم دقت تجمعی پویای وزنی مربوط به رکورد nـام را با $\bar{I}^w(\bar{F}_{R_n}, \bar{F}; t) < \infty$ و $t \geq ۰$ نشان دهیم. آنگاه $\bar{I}^w(\bar{F}_{R_n}, \bar{F}; t)$ تابع توزیع را به‌طور یکتا مشخصه‌سازی می‌کند.

برهان. از (۴۴.۴) داریم

$$\bar{I}^w(\bar{F}_{R_n}, \bar{F}; t) = \log \bar{F}(t) m_n^c(t) + \frac{1}{\bar{F}_{R_n}(t)} \sum_{j=0}^{n-1} \int_t^{+\infty} x \frac{[-\log \bar{F}(x)]^{j+1}}{j!} \bar{F}(x) dx. \quad (۴۵.۴)$$

با مشتق‌گیری از دو طرف (۴۵.۴) نسبت به t به‌دست می‌آوریم

$$\begin{aligned} \frac{\partial}{\partial t}[\bar{I}^w(\bar{F}_{R_n}, \bar{F}; t)] &= -\lambda_F(t) m_n^c(t) + \lambda_{F_{R_n}}(t) \bar{I}^w(\bar{F}_{R_n}, \bar{F}; t) \\ &= -\lambda_F(t) m_n^c(t) + c(t) \lambda_F(t) \bar{I}^w(\bar{F}_{R_n}, \bar{F}; t) \\ &= \lambda_F(t) \left[c(t) \bar{I}^w(\bar{F}_{R_n}, \bar{F}; t) - m_n^c(t) \right]. \end{aligned}$$

با مشتق‌گیری مجدد نسبت به t نتیجه می‌شود

$$\acute{\lambda}_F(t) = \frac{(\lambda_F(t))^2 \left[t - c(t)\lambda_F(t) m_n^c(t) + \acute{c}(t) \bar{I}^w(\bar{F}_{R_n}, \bar{F}; t) + c(t) \frac{\partial}{\partial t} \bar{I}^w(\bar{F}_{R_n}, \bar{F}; t) \right]}{\frac{\partial}{\partial t} \bar{I}^w(\bar{F}_{R_n}, \bar{F}; t)} \quad (۴۶.۴)$$

فرض کنید دو تابع F و F^* وجود دارند به‌طوری‌که

$$\bar{I}^w(\bar{F}_{R_n}, \bar{F}; t) = \bar{I}^w(\bar{F}^*_{R_n}, \bar{F}^*; t) = \tilde{z}(t).$$

آنگاه برای تمام مقادیر t، از (۴۶.۴) به‌دست می‌آوریم

$$\acute{\lambda}_F(t) = \varphi(t, \lambda_F(t)), \quad \acute{\lambda}_{F^*}(t) = \varphi(t, \lambda_{F^*}(t)),$$

که

$$\varphi(t, y) = \frac{y^2 \left[\acute{c}(t) \tilde{z}(t) + c(t) \left(\acute{\tilde{z}}(t) - y \tilde{s}(t) \right) + t \right]}{\acute{\tilde{z}}(t)},$$

و $\tilde{s}(t) = m_n^c(t)$. با استفاده از قضیه ۲.۳ و لم ۳.۳ از گوبتا و کرمانی (۲۰۰۸) داریم $\lambda_F(t) = \lambda_{F^*}(t)$، برای تمام مقادیر t. از آنجا که تابع نرخ خطر تابع توزیع را به‌طور یکتا مشخصه‌سازی می‌کند، اثبات کامل می‌شود. □

فصل ۵

معیارهای عدم دقت در متغیرهای همراه از آماره‌های ترتیبی تعمیم‌یافته

۱.۵ مقدمه

فرض کنید (X_i, Y_i)، $i = ۱,۲,\cdots,n$ متغیرهای تصادفی مستقل و همتوزیع از تابع توزیع پیوسته و دومتغیره‌ی $F(x,y)$ باشند. اگر $X_{(r:n)}$ نشاندهنده‌ی rـامین آماره‌ی ترتیبی باشد، آنگاه متغیر Y که مرتبط با $X_{(r:n)}$ است را با $Y_{[r:n]}$ نشان می‌دهیم و به آن همراه rـامین آماره‌ی ترتیبی می‌گوییم. متغیرهای همراه در انتخاب و پیش‌بینی مسایل مورد توجه هستند.

مفهوم آماره‌های ترتیبی تعمیم‌یافته (GOS) به‌عنوان یک روش یکپارچه‌سازی از مدل‌های مختلف متغیرهای تصادفی ترتیبی مانند آماره‌های ترتیبی معمولی، آماره‌های ترتیبی متوالی، سانسور فزاینده نوع ۲ و مقادیر رکورد توسط کمپس (۱۹۹۵) معرفی شده است. متغیرهای تصادفی $X(۱,n,m,k), X(۲,n,m,k), \cdots, X(n,n,m,k)$ را آماره‌های ترتیبی تعمیم‌یافته بر اساس تابع توزیع کاملاً پیوسته‌ی F با تابع چگالی f می‌گوییم اگر تابع چگالی توأم آن‌ها به‌صورت زیر باشد

$$\begin{aligned} f_{X(۱,n,m,k),\dots,X(n,n,m,k)}(x_۱,\dots,x_n) &= k\left(\prod_{j=۱}^{n-۱}\gamma_j\right)\left(\prod_{i=۱}^{n-۱}(۱-F(x_i))^m f(x_i)\right)(۱-F(x_n))^{k-۱}f(x_n), \\ & F^{-۱}(۰) \le x_۱ \le x_۲ \le \dots \le x_n \le F^{-۱}(۱), \end{aligned}$$

با پارامترهای $n \in \mathbb{N}, k > 0, m \in \mathbb{R}$، به‌طوری‌که $\gamma_r = k + (n-r)(m+1) > 0$ برای تمام $1 \leq r \leq n$. به‌طور مشابه، متغیرهای همراه می‌توانند در مورد GOS تعریف شوند.

مورگنسترن[۱] (۱۹۵۶) یک کلاس از توزیع‌های دومتغیره را تعریف کرد که تابع چگالی آن به‌صورت زیر است

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) \left[1 + \alpha(2F_X(x)-1)(2F_Y(y)-1)\right], \quad |\alpha| \leq 1, \tag{۱.۵}$$

که α پارامتر وابستگی است (برای جزییات بیشتر به دورانتی و سمپای[۲]، ۲۰۱۵ و منابع آن مراجعه شود). برای خانواده توزیع‌های مورگنسترن با تابع چگالی داده شده در (۱.۵)، تابع چگالی و تابع توزیع همراه rـامین GOS (با $Y_{[r,n,m,k]}, 1 \leq r \leq n$ نشان داده می‌شود)، توسط بگ و احسان‌الله[۳] (۲۰۰۸) به‌صورت زیر داده شده است:

$$g_{[r,n,m,k]}(y) = f_Y(y) \left[1 + \alpha C^*(r,n,m,k)(1 - 2F_Y(y))\right], \tag{۲.۵}$$

و

$$G_{[r,n,m,k]}(y) = F_Y(y) \left[1 + \alpha C^*(r,n,m,k)(1 - F_Y(y))\right], \tag{۳.۵}$$

که $C^*(r,n,m,k) = \frac{2\prod_{j=1}^{r}\gamma_j}{\prod_{i=1}^{r}(\gamma_i+1)} - 1$. به‌عنوان حالت خاصی از GOS در خانواده‌ی مورگنسترن، اگر $Y_{[r:n]}$ همراه rـامین آماره‌ی ترتیبی $X_{(r:n)}$ را نشان دهد، آنگاه تابع چگالی و توزیع $Y_{[r:n]}$ در خانواده مورگنسترن به‌صورت زیر است

$$f_{Y_{[r:n]}}(y) = f_Y(y) \left[1 + \alpha \left(\frac{n - 2r + 1}{n+1}\right)(1 - 2F_Y(y))\right],$$

و

$$F_{Y_{[r:n]}}(y) = F_Y(y) \left[1 + \alpha \left(\frac{n - 2r + 1}{n+1}\right)(1 - F_Y(y))\right],$$

برای جزییات بیشتر به آرنولد و همکاران (۱۹۹۲) مراجعه کنید.

فرض کنید $(X_1, Y_1), (X_2, Y_2), \cdots$ یک دنباله از متغیرهای تصادفی دومتغیره از یک توزیع پیوسته باشد. اگر $\{R_n, n \geq 1\}$ دنباله‌ی رکوردهای بالا در دنباله‌ی Xها باشد، آنگاه Y مرتبط با رکورد nـام را همراه رکورد nـام می‌خوانیم و با $R_{[n]}$ نشان داده می‌شود. همراه مقادیر رکورد در طیف گسترده‌ای از آزمایش‌های تجربی مانند آزمون تنش صنعتی، آزمایش‌های طول عمر، تحلیل هواشناسی، مسابقات ورزشی و برخی زمینه‌های تجربی دیگر کاربرد دارند. برای کابردهای دیگر مقادیر رکورد و همراه آن‌ها آرنولد و همکاران (۱۹۹۲) را ببینید. تابع چگالی و تابع توزیع $R_{[n]}$ به‌صورت زیر به‌دست می‌آید:

$$f_{R_{[n]}}(y) = f_Y(y)[1 + \alpha_n(1 - 2F_Y(y))], \quad n \geq 1 \tag{۴.۵}$$

[۱] Morgenstern

[۲] Durante and Sempi

[۳] Beg and Ahsanullah

و

$$F_{R_{[n]}}(y) = F_Y(y)[۱ + \alpha_n(۱ - F_Y(y))], \qquad (۵.۵)$$

که $\alpha_n = \alpha(۲^{۱-n} - ۱)$.

۲.۵ اندازه عدم دقت برای همراه GOS

اگر $Y_{[r,n,m,k]}$ همراه rـامین GOS از (۱.۵) باشد، اندازه عدم دقت بین $g_{[r,n,m,k]}(y)$ و $f_Y(y)$ برای $۱ \leq r \leq n$ و $\alpha \neq ۰$ به‌صورت زیر به‌دست می‌آید

$$\begin{aligned} I(g_{[r,n,m,k]}, f_Y) &= -\int_۰^\infty g_{[r,n,m,k]}(y) \log f_Y(y) dy \\ &= [۱ + \alpha C^*(r,n,m,k)]\, H(Y) + ۲\alpha C^*(r,n,m,k) \int_۰^\infty f_Y(y) F_Y(y) \log f_Y(y) dy \\ &= [۱ + \alpha C^*(r,n,m,k)]\, H(Y) + ۲\alpha C^*(r,n,m,k) \int_۰^۱ u \log f_Y(F_Y^{-۱}(u)) du \\ &= [۱ + \alpha C^*(r,n,m,k)]\, H(Y) + ۲\alpha C^*(r,n,m,k) \phi_f(u), \qquad (۶.۵) \end{aligned}$$

که $\phi_f(u) = \int_۰^۱ u \log f_Y(F_Y^{-۱}(u)) du$ و

$$H(Y) = -\int_۰^\infty f_Y(y) \log f_Y(y) dy$$

آنتروپی شانون مربوط به متغیر تصادفی Y است.

اکنون حالت‌های خاص زیر را به‌عنوان کاربردی از رابطه‌ی (۶.۵) در نظر می‌گیریم.

حالت ۱. در (۶.۵) اگر $m = ۰$ و $k = ۱$ قرار دهیم، آنگاه اندازه عدم دقت بین $f_{Y_{[r:n]}}$ (تابع چگالی rـامین همراه آماره ترتیبی) و f_Y در خانواده مورگنسترن به‌صورت زیر به‌دست می‌آید:

$$\begin{aligned} I(f_{Y_{[r:n]}}, f_Y) &= -\int_۰^\infty f_{Y_{[r:n]}}(y) \log f_Y(y) dy \\ &= \left[۱ + \alpha \left(\frac{n - ۲r + ۱}{n + ۱}\right)\right] H(Y) + ۲\alpha \left(\frac{n - ۲r + ۱}{n + ۱}\right) \int_۰^\infty f_Y(y) F_Y(y) \log f_Y(y) dy \\ &= \left[۱ + \alpha \left(\frac{n - ۲r + ۱}{n + ۱}\right)\right] H(Y) + ۲\alpha \left(\frac{n - ۲r + ۱}{n + ۱}\right) \int_۰^۱ u \log f(F^{-۱}(u)) du, \\ &= H(Y) + \frac{۳(n - ۲r + ۱)}{۲(n + ۱)} \left[I(f_{Y_{[۱:۲]}}, f_Y) - I(f_{Y_{[۲:۲]}}, f_Y)\right]. \qquad (۷.۵) \end{aligned}$$

که $H(Y) = -\int_۰^\infty f_Y(y) \log f_Y(y) dy$ آنتروپی شانون متغیر تصادفی Y است. توجه کنید که $I(f_{Y_{[\frac{n+۱}{۲}:n]}}, f_Y) = H(Y)$.

در ادامه برخی مثال‌ها و ویژگی‌های $I(f_{Y_{[r:n]}}, f_Y)$ را بیان می‌کنیم.

مثال ۱.۲.۵. فرض کنید (X_i, Y_i)، $i = 1, 2, ..., n$ یک نمونه تصادفی از توزیع نمایی دومتغیره گامبل (GBED) با تابع توزیع زیر باشد

$$F(x,y) = \left(1 - \exp\left(\frac{-x}{\theta_1}\right)\right)\left(1 - \exp\left(\frac{-y}{\theta_2}\right)\right)\left[1 + \alpha \exp\left(\frac{-x}{\theta_1} - \frac{y}{\theta_2}\right)\right]. \qquad (8.5)$$

از (۷.۵) به‌دست می‌آوریم

$$I(f_{Y_{[r:n]}}, f_Y) = [1 + \log \theta_2] - \frac{\alpha}{2}\left(\frac{n - 2r + 1}{n + 1}\right). \qquad (9.5)$$

با استفاده از (۹.۵) داریم

$$A_\alpha(n) = I(f_{Y_{[n:n]}}, f_Y) - I(f_{Y_{[1:n]}}, f_Y) = \alpha\left(\frac{n-1}{n+1}\right),$$

که مثبت، منفی و صفر است هرگاه $(0 < \alpha \leq 1, n > 1)$، $(-1 \leq \alpha < 0, n > 1)$ یا $(n = 1\ or\ \alpha = 0)$. همچنین اختلاف بین $I(f_{Y_{[r:n]}}, f_Y)$ و $H(Y)$ برابر است با

$$B_{\alpha,n}(r) = I(f_{Y_{[r:n]}}, f_Y) - H(Y) = -\frac{\alpha}{2}\left(\frac{n - 2r + 1}{n + 1}\right).$$

$B_{\alpha,n}(r)$ مثبت است اگر $-1 \leq \alpha < 0$ و $1 \leq r < \frac{n+1}{2}$ (یا $0 < \alpha \leq 1$ و $\frac{n+1}{2} < r \leq n$). همچنین منفی است اگر $-1 \leq \alpha < 0$ و $\frac{n+1}{2} < r \leq n$ (یا $0 < \alpha \leq 1$ و $1 \leq r < \frac{n+1}{2}$).
حال اگر n فرد باشد، محاسبات عددی نشان می‌دهد که $I(f_{Y_{[r:n]}}, f_Y)$ در r برای $1 \leq r < \frac{n+1}{2}$ و $0 < \alpha \leq 1$ ($\frac{n+1}{2} < r \leq n$ و $-1 \leq \alpha < 0$) صعودی (نزولی) است.

مثال ۲.۲.۵. فرض کنید (X_i, Y_i)، $i = 1, 2, \cdots, n$ یک نمونه تصادفی از توزیع لجستیک دومتغیره مورگنشترن با تابع توزیع زیر باشد

$$F(x,y) = (1 + \exp(-x))^{-1}(1 + \exp(-y))^{-1}\left(1 + \frac{\alpha e^{-x-y}}{(1 + e^{-x})(1 + e^{-y})}\right).$$

محاسبات نشان می‌دهد

$$I(f_{Y_{[r:n]}}, f_Y) = 1 - 0/6\alpha\left(\frac{n - 2r + 1}{n + 1}\right). \qquad (10.5)$$

با استفاده از (۱۰.۵) داریم

$$D_\alpha(n) = I(f_{Y_{[n:n]}}, f_Y) - I(f_{Y_{[1:n]}}, f_Y) = 1/2\alpha\left(\frac{n-1}{n+1}\right),$$

که مثبت، منفی یا صفر است هرگاه به‌ترتیب $(0 < \alpha \leq 1, n > 1)$، $(-1 \leq \alpha < 0, n > 1)$ یا $(n = 1\ or\ \alpha = 0)$.

مثال ۳.۲.۵. فرض کنید (X_i, Y_i)، $i = ۱, ۲, ..., n$ یک نمونه تصادفی از توزیع رایلی دومتغیره مورگنسترن با تابع توزیع زیر باشد

$$F(x, y) = \left(۱ - \exp\left(-\frac{x^۲}{۲\sigma_۱^۲}\right)\right)\left(۱ - \exp\left(-\frac{y^۲}{۲\sigma_۲^۲}\right)\right)\left(۱ + \alpha \exp\left(-\frac{x^۲}{۲\sigma_۱^۲} - \frac{y^۲}{۲\sigma_۲^۲}\right)\right)$$

از رابطه (۷.۵) به‌دست می‌آوریم

$$I(f_{Y_{[r:n]}}, f_Y) = \frac{\alpha(n - ۲r + ۱)}{n + ۱}\left(\log\sqrt{۲} - \frac{۱}{۲}\right) + ۱ - \frac{۱}{۲}\psi(۱) + \log\left(\frac{\sigma_۲}{\sqrt{۲}}\right). \qquad (۱۱.۵)$$

با استفاده از (۱۱.۵) داریم

$$W_\alpha(n) = I(f_{Y_{[n:n]}}, f_Y) - I(f_{Y_{[۱:n]}}, f_Y) = ۲\alpha\left(۰/۵ - \log\sqrt{۲}\right)\left(\frac{n - ۱}{n + ۱}\right),$$

که مثبت، منفی یا صفر است هرگاه به‌ترتیب $(۰ < \alpha \leq ۱, n > ۱)$، $(-۱ \leq \alpha < ۰, n > ۱)$ یا $(n = ۱ \; or \alpha = ۰)$.

مثال ۴.۲.۵. فرض کنید (X_i, Y_i)، $i = ۱, ۲, ..., n$ یک نمونه تصادفی از توزیع نمایی تعمیم‌یافته دومتغیره مورگنسترن (MTBGED) با تابع توزیع زیر باشد

$$F_{X,Y}(x, y) = \left\{\left(۱ - e^{-\theta_۱ x}\right)\left(۱ - e^{-\theta_۲ y}\right)\right\}^\lambda \left[۱ + \alpha\left(۱ - \left(۱ - e^{-\theta_۱ x}\right)^\lambda\right)\left(۱ - \left(۱ - e^{-\theta_۲ y}\right)^\lambda\right)\right].$$

با استفاده از (۷.۵) داریم

$$I(f_{Y_{[r:n]}}, f_Y) = -\log(\lambda\theta_۲) + B(\lambda) - \frac{\alpha(n - ۲r + ۱)}{n + ۱}D(\lambda) + \frac{\lambda - ۱}{\lambda}\left[۱ + \frac{\frac{\alpha(n-۲r+۱)}{n+۱}}{۲}\right], \qquad (۱۲.۵)$$

که $B(\lambda) = \psi(\lambda + ۱) - \psi(۱)$ و $D(\lambda) = B(۲\lambda) - B(\lambda)$. با توجه به (۱۲.۵) داریم

$$Q_{\alpha,\lambda}(n) = I(f_{Y_{[n:n]}}, f_Y) - I(f_{Y_{[۱:n]}}, f_Y) = \frac{\alpha(n - ۱)}{n + ۱}\left[۲D(\lambda) - \frac{\lambda - ۱}{\lambda}\right],$$

که مثبت، منفی یا صفر است اگر به‌ترتیب $(۰ < \alpha \leq ۱, n > ۱)$، $(-۱ \leq \alpha < ۰, n > ۱)$ یا $(n = ۱ \; or \alpha = ۰)$.

ملاحظه ۱.۲.۵. فرض کنید $(X_i, Y_i), i = ۱, ۲, \cdots, n$ یک نمونه‌ی تصادفی به حجم n با تابع چگالی (۱.۵) باشد. آنگاه، از (۷.۵) داریم

$$H(Y) = \frac{I(f_{Y_{[n:n]}}, f_Y) + I(f_{Y_{[۱:n]}}, f_Y)}{۲}.$$

ملاحظه ۲.۲.۵. فرض کنید $(X_i, Y_i), i = ۱, ۲, \cdots, n$ یک نمونه‌ی تصادفی به حجم n با تابع چگالی (۱.۵) باشد. اگر $\lambda \geq ۱$ یک عدد صحیح باشد و ما r را به $r\lambda$ و n را به $(n + ۱)\lambda - ۱$ تغییر دهیم، آنگاه از (۷.۵) داریم

$$I(f_{Y_{[r:n]}}, f_Y) = I(f_{Y_{[r\lambda:(n+۱)\lambda-۱]}}, f_Y).$$

ما متغیرهای همراه آماره‌های ترتیبی $(X_۱, Y_۱), (X_۲, Y_۲), \ldots, (X_n, Y_n)$ را مستقل ولی با توزیع‌های متفاوت در نظر می‌گیریم. فرض کنید ما خانواده‌ی مورگنسترن را با تابع توزیع زیر داشته باشیم

$$F_{X_i,Y_i}(x, y) = F_{X_i}(x)F_{Y_i}(y)\left[۱ + \alpha_i(۱ - F_{X_i}(x))(۱ - F_{Y_i}(y))\right]. \qquad (۱۳.۵)$$

حال، فرض کنید که $F_{X_i}(x) = F_X(x)$، $F_{Y_i}(y) = F_Y(y)$ و $|\alpha_i| \leq ۱$. در این حالت خاص تابع چگالی $Y_{[۱:n]}$ و $Y_{[n:n]}$ توسط اریالماز[۴] (۲۰۰۵) به‌صورت زیر داده شده است:

$$f_{[۱:n]}(y) = f_Y(y)\left[۱ + \frac{n-۱}{(n+۱)n}\sum_{j=۱}^{n}\alpha_j(۱ - ۲F_Y(y))\right], \qquad (۱۴.۵)$$

$$f_{[n:n]}(y) = f_Y(y)\left[۱ - \frac{n-۱}{(n+۱)n}\sum_{j=۱}^{n}\alpha_j(۱ - ۲F_Y(y))\right]. \qquad (۱۵.۵)$$

در ادامه، اندازه عدم دقت را برای همراه آماره‌های ترتیبی مینیمم و ماکزیمم نشان می‌دهیم.

مثال ۵.۲.۵. فرض کنید $(X_i, Y_i),\ i = ۱, ۲, \ldots, n$ بردارهای تصادفی مستقل از (۱۳.۵) باشند. اگر $Y_{[۱:n]}$ و $Y_{[n:n]}$ همراه آماره‌های ترتیبی مینیمم و ماکزیمم باشند، آن‌گاه

$$I(f_{[۱:n]}, f_Y) = \left(۱ + \frac{n-۱}{(n+۱)n}\sum_{j=۱}^{n}\alpha_j\right)H(Y) + ۲\frac{n-۱}{(n+۱)n}\sum_{j=۱}^{n}\alpha_j\phi_f(u), \qquad (۱۶.۵)$$

$$I(f_{[n:n]}, f_Y) = \left(۱ - \frac{n-۱}{(n+۱)n}\sum_{j=۱}^{n}\alpha_j\right)H(Y) - ۲\frac{n-۱}{(n+۱)n}\sum_{j=۱}^{n}\alpha_j\phi_f(u). \qquad (۱۷.۵)$$

با استفاده از (۱۶.۵) و (۱۷.۵) داریم

$$A_n = I(f_{[n:n]}, f_Y) - I(f_{[۱:n]}, f_Y) = -\frac{۲(n-۱)}{n(n+۱)}\Delta,$$

که قرار داده‌ایم $\Delta = H(Y)\sum_{j=۱}^{n}\alpha_j + ۲\sum_{j=۱}^{n}\alpha_j\phi_f(u)$. اگر $\Delta > ۰$ ($\Delta < ۰$)، آنگاه $A_n < ۰$ ($A_n > ۰$). در انتها به‌دست می‌آوریم

$$I(f_{[n:n]}, f_Y) + I(f_{[۱:n]}, f_Y) = ۲H(Y).$$

که $\alpha_n = \alpha(۲^{۱-n} - ۱)$. اندازه عدم دقت بین $f_{R_{[n]}}$ (تابع چگالی nـامین همراه رکورد بالا) و f_Y به‌صورت زیر است

$$\begin{aligned} I(f_{R_{[n]}}, f_Y) &= -\int_۰^\infty f_{R_{[n]}}(y)\log f_Y(y)dy \\ &= [۱ + \alpha_n]H(Y) + ۲\alpha_n\int_۰^\infty f_Y(y)F_Y(y)\log f_Y(y)dy. \qquad (۱۸.۵) \end{aligned}$$

[۴] Eryilmaz

حالت ۲. طبق (۶.۵) اگر قرار دهیم $m=-۱$ و $k=۱$، آنگاه اندازه عدم دقت بین $f_{R_{[r]}}$ (تابع چگالی rـامین همراه رکورد بالا) و f_Y در خانواده مورگنسترن به‌صورت زیر به‌دست می‌آید:

$$I(f_{R_{[r]}}, f_Y) = \left(۱+\alpha(۲^{۱-r}-۱)\right) H(Y) + ۲\alpha(۲^{۱-r}-۱)\phi_f(u). \qquad (۱۹.۵)$$

مثال ۶.۲.۵. فرض کنید $(X_i, Y_i),\ i=۱,۲,...,n$ یک نمونه‌ی تصادفی از GBED با تابع توزیع زیر باشد

$$F(x,y) = \left(۱-\exp\left(\frac{-x}{\theta_۱}\right)\right)\left(۱-\exp\left(\frac{-y}{\theta_۲}\right)\right)\left[۱+\alpha\exp\left(\frac{-x}{\theta_۱}-\frac{y}{\theta_۲}\right)\right]. \qquad (۲۰.۵)$$

از (۱۹.۵) به‌دست می‌آوریم

$$I(f_{R_{[r]}}, f_Y) = [۱+\log\theta_۲] + \frac{\alpha}{۲}\left(۲^{۱-r}-۱\right). \qquad (۲۱.۵)$$

با استفاده از (۲۱.۵) داریم

$$A_\alpha(r) = I(f_{R_{[r]}}, f_Y) - I(f_{R_{[r-۱]}}, f_Y) = -\alpha ۲^{-r},$$

که مثبت، منفی یا صفر است هرگاه به‌ترتیب $-۱\leq\alpha<۰, r>۱$؛ $۰<\alpha\leq ۱, r>۱$ یا $\alpha=۰$.
همچنین، اختلاف بین $I(f_{R_{[r]}}, f_Y)$ و $H(Y)$ برابر است با

$$B_{\alpha,n}(r) = I(f_{R_{[r]}}, f_Y) - H(Y) = \frac{\alpha}{۲}\left(۲^{۱-r}-۱\right).$$

$B_{\alpha,n}(r)$ مثبت، منفی یا صفر است هرگاه به‌ترتیب $-۱\leq\alpha<۰, r>۱$؛ $۰<\alpha\leq ۱, r>۱$ یا $\alpha=۰$ $r=۱$.

ملاحظه ۳.۲.۵. در مقایسه با رابطه‌ی (۶.۵)، اندازه عدم دقت مرتبط با $f_Y(y)$ و $g_{[r,n,m,k]}(y)$ به‌صورت زیر به‌دست می‌آید

$$I(f_Y, g_{[r,n,m,k]}) = H(Y) - E\left[\log\left(۱+\alpha C^*(r,n,m,k)\left(۱-۲U\right)\right)\right],$$

که U دارای توزیع یکنواخت در $(۰,۱)$ است.

توابع چندک جایگزین‌های کارآمدی برای تابع توزیع در مدل‌سازی و تحلیل داده‌های آماری هستند. تابع چندک به‌صورت زیر تعریف می‌شود

$$Q(u) = F^{-۱}(u) = \inf\{y : F(y)\geq u\}, \qquad ۰<u<۱.$$

توجه کنید که $F(Q(u))=u$ و مشتق‌گیری نسبت به u نتیجه می‌دهد $Q'(u)f(Q(u))=۱$. فرض کنید که Y یک متغیر تصادفی نامنفی با تابع چگالی $f(\cdot)$ و تابع چندک $Q(\cdot)$ باشد، آنگاه $f(Q(u))$ تابع چگالی چندک خوانده می‌شود و $q(u)=Q'(u)$ به‌عنوان چندک تابع چگالی Y شناخته می‌شود. اکنون با استفاده از (۶.۵)، $I(g_{[r,n,m,k]}, f_Y)$ بر اساس چندک متناظر به این صورت تعریف می‌شود

$$I(g_{[r,n,m,k]}, f_Y) = E(\log q(U)) + \alpha C^*(r,n,m,k) E\left[(۱-۲U)\log q(U)\right]. \qquad (۲۲.۵)$$

۳.۵ CPI بین $Y_{[r,n,m,k]}$ و Y

اگر $Y_{[r,n,m,k]}$ همراه rــامین GOS از (۱.۵) باشد، آنگاه اندازه CPI بین $G_{[r,n,m,k]}(y)$ و $F_Y(y)$ برای $1 \leq r \leq n$ و $\alpha \neq 0$ به‌صورت زیر داده می‌شود

$$\begin{aligned} I(G_{Y_{[r,n,m,k]}}, F_Y) &= -\int_0^{\infty} G_{[r,n,m,k]}(y) \log F_Y(y) dy \\ &= [1+\alpha C^*(r,n,m,k)]\, \mathcal{CE}(Y) + \alpha C^*(r,n,m,k) \int_0^{\infty} F_Y^2(y) \log F_Y(y) dy \\ &= [1+\alpha C^*(r,n,m,k)]\, \mathcal{CE}(Y) - \frac{\alpha}{2} C^*(r,n,m,k) \mathcal{CE}(Y_{(2:2)}), \qquad (۲۳.۵) \end{aligned}$$

که $\mathcal{CE}(Y)$ و $\mathcal{CE}(Y_{(2:2)})$ به ترتیب آنتروپی تجمعی متغیرهای تصادفی Y و $Y_{(2:2)}$ هستند (دی کرشنزو و لونگوبردی، ۲۰۰۹ را ببینید).

ملاحظه ۱.۳.۵. در مقایسه با (۲۳.۵)، اندازه عدم دقت مرتبط با F_Y و $G_{[r,n,m,k]}$ به‌صورت زیر به‌دست می‌آید

$$I(F_Y, G_{[r,n,m,k]}) = \mathcal{CE}(Y) - E\left[\frac{U \log\left(1+\alpha C^*(r,n,m,k)\,(1-U)\right)}{f(F^{-1}(U))}\right].$$

حالت ۱. اگر قرار دهیم $m=0$ و $k=1$، آنگاه اندازه عدم دقت بین $F_{Y_{[r:n]}}$ (تابع توزیع rــامین همراه آماره ترتیبی) و F_Y به‌صورت زیر نشان داده می‌شود

$$\begin{aligned} I(F_{Y_{[r:n]}}, F_Y) &= -\int_0^{\infty} F_{Y_{[r:n]}}(y) \log F_Y(y) dy \\ &= \left[1+\alpha\left(\frac{n-2r+1}{n+1}\right)\right] \mathcal{CE}(Y) + \alpha\left(\frac{n-2r+1}{n+1}\right) \int_0^{\infty} F_Y^2(y) \log F_Y(y) dy \\ &= \left[1+\alpha\left(\frac{n-2r+1}{n+1}\right)\right] \mathcal{CE}(Y) - \alpha\left(\frac{n-2r+1}{2(n+1)}\right) \mathcal{CE}(Y_{(2:2)}). \qquad (۲۴.۵) \end{aligned}$$

مثال ۱.۳.۵. فرض کنید (X_i, Y_i)، $i = 1, 2, \cdots, n$ یک متغیر تصادفی از توزیع یکنواخت دومتغیره مورگنسترن (MTBUD) با تابع توزیع زیر باشد

$$F(x,y) = \frac{xy}{\theta_1\theta_2}\left[1+\alpha\left(1-\frac{x}{\theta_1}\right)\left(1-\frac{y}{\theta_2}\right)\right], \quad 0<x<\theta_1, \quad 0<y<\theta_2.$$

با محاسبه به‌دست می‌آید

$$\begin{aligned} I(F_{Y_{[r:n]}}, F_Y) &= \left[1+\alpha\left(\frac{n-2r+1}{n+1}\right)\right]\frac{\theta_2}{4} - \alpha\left(\frac{n-2r+1}{(n+1)}\right)\frac{\theta_2}{9} \\ &= \frac{\theta_2}{4} + \alpha\left(\frac{n-2r+1}{n+1}\right)\left(\frac{5\theta_2}{36}\right). \qquad (۲۵.۵) \end{aligned}$$

با استفاده از (۲۵.۵) داریم

$$D_{\alpha,\theta_2}(n) = I(F_{Y_{[n:n]}}, F_Y) - I(F_{Y_{[1:n]}}, F_Y) = \frac{5\alpha\theta_2(-n+1)}{18(n+1)}.$$

که مثبت، منفی یا صفر است هرگاه به‌ترتیب $(-1 \leq \alpha < 0)$، $(0 < \alpha \leq 1)$ یا $(\alpha = 0)$.

مثال ۲.۳.۵. فرض کنید (X_i, Y_i)، $i = ۱,۲,\cdots,n$ یک نمونه‌ی تصادفی از GBED باشد. محاسبات نشان می‌دهد که

$$\begin{aligned} I(F_{Y_{[r:n]}}, F_Y) = &= \left[۱+\alpha\left(\frac{n-۲r+۱}{n+۱}\right)\right]\left[\frac{\pi^۲}{۶}-۱\right]\theta_۲ - \alpha\left(\frac{n-۲r+۱}{(n+۱)}\right)\left[\frac{\pi^۲}{۶}-\frac{۵}{۴}\right]\theta_۲ \\ &= \left[\frac{\pi^۲}{۶}-۱\right]\theta_۲ + \frac{\alpha\theta_۲}{۴}\left(\frac{n-۲r+۱}{n+۱}\right). \end{aligned} \tag{۲۶.۵}$$

با استفاده از (۲۶.۵) داریم

$$Q_{\alpha,\theta_۲}(n) = I(F_{Y_{[n:n]}}, F_Y) - I(F_{Y_{[۱:n]}}, F_Y) = \frac{\alpha\theta_۲(-n+۱)}{۲(n+۱)},$$

که به‌ترتیب مثبت، منفی یا صفر است هرگاه $(-۱ \leq \alpha < ۰)$، $(۰ < \alpha \leq ۱)$ یا $(\alpha = ۰)$.

مثال ۳.۳.۵. فرض کنید (X_i, Y_i)، $i = ۱,۲,\cdots,n$ یک نمونه تصادفی از توزیع وایبل معکوس دومتغیره مورگنسترن با تابع توزیع زیر باشد

$$F(x,y) = \exp\left(-\left(\frac{\theta_۱}{x}\right)^{\beta_۱} - \left(\frac{\theta_۲}{y}\right)^{\beta_۲}\right)\left[۱+\alpha\left(۱-\exp\left(-\left(\frac{\theta_۱}{x}\right)^{\beta_۱}\right)\right)\left(۱-\exp\left(-\left(\frac{\theta_۲}{y}\right)^{\beta_۲}\right)\right)\right].$$

محاسبات نشان می‌دهد

$$\begin{aligned} I(F_{Y_{[r:n]}}, F_Y) &= \left[۱+\alpha\left(\frac{n-۲r+۱}{n+۱}\right)\right]\frac{\theta_۲}{\beta_۲}\Gamma\left(\frac{\beta_۲-۱}{\beta_۲}\right) - \alpha\left(\frac{n-۲r+۱}{(n+۱)}\right)\frac{۲^{\frac{۱}{\beta_۲}-۱}\theta_۲}{\beta_۲}\Gamma\left(\frac{\beta_۲-۱}{\beta_۲}\right) \\ &= \frac{\theta_۲}{\beta_۲}\Gamma\left(\frac{\beta_۲-۱}{\beta_۲}\right) + \alpha\left(\frac{n-۲r+۱}{n+۱}\right)\frac{\theta_۲}{\beta_۲}\Gamma\left(\frac{\beta_۲-۱}{\beta_۲}\right)\left(۱-۲^{\frac{۱}{\beta_۲}-۱}\right). \end{aligned} \tag{۲۷.۵}$$

با استفاده از (۲۷.۵) داریم

$$D_{\alpha,\theta_۲}(n) = I(F_{Y_{[n:n]}}, F_Y) - I(F_{Y_{[۱:n]}}, F_Y) = \alpha\frac{\theta_۲(۱-n)}{\beta_۲(n+۱)}\Gamma\left(\frac{\beta_۲-۱}{\beta_۲}\right)\left(۱-۲^{\frac{۱}{\beta_۲}-۱}\right).$$

که به‌ترتیب مثبت، منفی یا صفر است هرگاه $(-۱ \leq \alpha < ۰)$، $(۰ < \alpha \leq ۱)$ یا $(\alpha = ۰)$.

گزاره ۱.۳.۵. فرض کنید (X_i, Y_i)، $i = ۱,۲,\cdots,n$ یک نمونه تصادفی از خانواده مورگنسترن باشد. آنگاه برای $۱ \leq r \leq \frac{n+۱}{۲}$ داریم

$$I(F_{Y_{[r:n]}}, F_Y) \leq (\geq)\mathcal{CE}(Y), \quad -۱ \leq \alpha < ۰(۰ < \alpha \leq ۱). \tag{۲۸.۵}$$

برهان. با یادآوری قضیه ۴.۸ از دی‌کرشنزو و لونگوباردی (۲۰۰۹) اثبات به‌دست می‌آید. □

حالت ۲. اگر قرار دهیم $m = -۱$ و $k = ۱$، آنگاه اندازه عدم دقت بین $F_{R_{[r]}}$ (تابع توزیع rـامین همراه رکورد بالا) و F_Y به‌صورت زیر است

$$\begin{aligned} I(F_{R_{[r]}}, F_Y) &= -\int_۰^\infty F_{R_{[r]}}(y)\log F_Y(y)dy \\ &= \left[۱+\alpha\left(۲^{۱-r}-۱\right)\right]\mathcal{CE}(Y) + \alpha\left(۲^{۱-r}-۱\right)\int_۰^\infty F_Y^۲(y)\log F_Y(y)dy \\ &= \left[۱+\alpha\left(۲^{۱-r}-۱\right)\right]\mathcal{CE}(Y) - \frac{\alpha}{۲}\left(۲^{۱-r}-۱\right)\mathcal{CE}\left(Y_{(۲:۲)}\right). \end{aligned} \tag{۲۹.۵}$$

گزاره ۲.۳.۵. فرض کنید (X_i, Y_i) و $i = ۱,۲,\cdots,n$ یک نمونه تصادفی از خانواده مورگنسترن باشد. آنگاه داریم

$$I(F_{R_{[r]}}, F_Y) \leq (\geq) \mathcal{CE}(Y), \quad ۰ < \alpha \leq ۱ \; (-۱ \leq \alpha < ۰). \tag{۳۰.۵}$$

برهان. با توجه به قضیه ۴.۸ دی‌کرشنزو و لونگوباردی (۲۰۰۹) اثبات می‌شود. □

۴.۵ CPI تجربی برای همراه GOS

در این بخش ما موضوع برآورد CPI برای متغیرهای همراه را با استفاده از از مقدار میانگین CPI تجربی بررسی می‌کنیم. فرض کنید $(X_i, Y_i), i = ۱,۲,\cdots,n$ یک نمونه‌ی تصادفی به حجم n از خانواده‌ی مورگنسترن باشد. آن‌گاه طبق (۲۳.۵)، CPI تجربی بین $G_{Y_{[r,n,m,k]}}$ و F_Y به‌صورت زیر به‌دست می‌آید:

$$\begin{aligned}\hat{I}(G_{Y_{[r,n,m,k]}}, F_Y) &= [۱ + \alpha C^*(r,n,m,k)] \sum_{j=۱}^{n-۱} U_j \left(\frac{j}{n}\right) \left(-\log \frac{j}{n}\right) \\ &- \alpha C^*(r,n,m,k) \sum_{j=۱}^{n-۱} U_j \left(\frac{j}{n}\right)^۲ \left(-\log \frac{j}{n}\right) \\ &= \sum_{j=۱}^{n-۱} U_j \left(\frac{j}{n}\right) \left(-\log \frac{j}{n}\right) \left[۱ + \alpha C^*(r,n,m,k) \left(۱ - \frac{j}{n}\right)\right],\end{aligned} \tag{۳۱.۵}$$

که $U_j = Z_{(j+۱)} - Z_{(j)}, j = ۱,۲,...,n-۱$.

حالت ۱. اگر قرار دهیم $m = ۰$ و $k = ۱$، آنگاه CPI تجربی بین $F_{Y_{[r:n]}}$ و F_Y برابر می‌شود با

$$\hat{I}(F_{Y_{[r:n]}}, F_Y) = \sum_{j=۱}^{n-۱} U_j \frac{j}{n} \left(-\log \frac{j}{n}\right) \left[۱ + \alpha \left(\frac{n - ۲r + ۱}{n+۱}\right) \left(۱ - \frac{j}{n}\right)\right]. \tag{۳۲.۵}$$

حالت ۲. اگر قرار دهیم $m = -۱$ و $k = ۱$، آنگاه CPI تجربی بین $F_{R_{[r]}}$ و F_Y به‌صورت زیر به‌دست می‌آید

$$\hat{I}(F_{R_{[r]}}, F_Y) = \sum_{j=۱}^{n-۱} U_j \frac{j}{n} \left(-\log \frac{j}{n}\right) \left[۱ + \alpha \left(۲^{۱-r} - ۱\right) \left(۱ - \frac{j}{n}\right)\right]. \tag{۳۳.۵}$$

مثال ۱.۴.۵. فرض کنید $(X_i, Y_i), i = ۱,۲,...,n$ یک نمونه‌ی تصادفی از MTBGED با $\lambda = ۱$ باشد. آنگاه U_jها مستقل و دارای توزیع نمایی با میانگین $\frac{۱}{\theta_۲(n-j)}$ است. حال از (۳۳.۵) به‌دست می‌آوریم

$$E[\hat{I}(F_{R_{[r]}}, F_Y)] = \frac{۱}{\theta_۲} \sum_{j=۱}^{n-۱} \frac{j}{n(n-j)} \left(-\log \frac{j}{n}\right) \left[۱ + \alpha \left(۲^{۱-r} - ۱\right) \left(۱ - \frac{j}{n}\right)\right], \tag{۳۴.۵}$$

و

$$Var[\hat{I}(F_{R_{[r]}},F_Y)]=\frac{۱}{\theta_۲^۲}\sum_{j=۱}^{n-۱}\left(\frac{j}{n(n-j)}\left(-\log\frac{j}{n}\right)\left[۱+\alpha\left(۲^{۱-r}-۱\right)\left(۱-\frac{j}{n}\right)\right]\right)^۲. \quad (۳۵.۵)$$

ما مقادیر عددی $E[\hat{I}(F_{R_{[r]}},F_Y)]$ و $Var[\hat{I}(F_{R_{[r]}},F_Y)]$ را برای حجم نمونه‌ی $n=۱۰,۱۵,۲۰$، $\theta_۲=۰/۵,۱,۲$، $\alpha=-۱,-۰/۵,۰/۵,۱$ و $r=۲$ در جدول ۱.۵ حساب کردیم. به آسانی می‌توانیم ببینیم که $E[\hat{I}(F_{R_{[r]}},F_Y)]$ و $Var[\hat{I}(F_{R_{[r]}},F_Y)]$ در α و $\theta_۲$ نزولی هستند. همچنین ملاحظه می‌شود که $\lim_{n\to\infty}Var[\hat{I}(F_{R_{[r]}},F_Y)]=۰$.

مثال ۲.۴.۵ فرض کنید $(X_i,Y_i),\ i=۱,۲,\cdots,n$ یک نمونه‌ی تصادفی از MTBUD باشد که $\theta_۱=\theta_۲=۱$. آنگاه U_jها مستقل و دارای توزیع بتا با پارامترهای ۱ و n است (برای جزییات بیشتر پیک، ۱۹۶۵ را ببینید). حال از (۳۳.۵) به‌دست می‌آوریم

$$E[\hat{I}(F_{R_{[r]}},F_Y)]=\frac{۱}{n+۱}\sum_{j=۱}^{n-۱}\frac{j}{n}\left(-\log\frac{j}{n}\right)\left[۱+\alpha\left(۲^{۱-r}-۱\right)\left(۱-\frac{j}{n}\right)\right], \quad (۳۶.۵)$$

و

$$Var[\hat{I}(F_{R_{[r]}},F_Y)]=\frac{n}{(n+۱)^۲(n+۲)}\sum_{j=۱}^{n-۱}\left(\frac{j}{n}\left(-\log\frac{j}{n}\right)\left[۱+\alpha\left(۲^{۱-r}-۱\right)\left(۱-\frac{j}{n}\right)\right]\right)^۲. \quad (۳۷.۵)$$

ما مقادیر عددی $E[\hat{I}(F_{R_{[r]}},F_Y)]$ و $Var[\hat{I}(F_{R_{[r]}},F_Y)]$ را برای حجم‌های نمونه‌ی $n=۱۰,۱۵,۲۰$، $\alpha=-۱,-۰/۵,۰/۵,۱$ و $r=۲$ در جدول ۲.۵ محاسبه کرده‌ایم. به آسانی می‌توانیم ببینیم که $E[\hat{I}(F_{R_{[r]}},F_Y)]$ و $Var[\hat{I}(F_{R_{[r]}},F_Y)]$ در α نزولی هستند. همچنین ملاحظه می‌شود که $\lim_{n\to\infty}Var[\hat{I}(F_{R_{[r]}},F_Y)]=۰$.

جدول ۱.۵: مقادیر عددی $\mathbb{E}[\hat{I}(F_{R_{[r]}},F_Y)]$ و $Var[\hat{I}(F_{R_{[r]}},F_Y)]$ برای MTBGED با $\lambda=۱$

	$\mathbb{E}[\hat{I}(F_{R_{[r]}},F_Y)]$											
$\theta_۲$	۰.۵	۱	۲	۰.۵	۱	۲	۰.۵	۱	۲	۰.۵	۱	۲
n	$\alpha=-۱$			$\alpha=-۰/۵$			$\alpha=۰/۵$			$\alpha=۱$		
۱۰	۱.۴۲۹	۰.۷۱۴	۰.۳۵۷	۱.۳۰۶	۰.۶۵۳	۰.۳۲۶	۱.۰۶۱	۰.۵۳۰	۰.۲۶۵	۰.۹۳۸	۰.۴۶۹	۰.۲۳۴
۱۵	۱.۴۶۸	۰.۷۳۴	۰.۳۶۷	۱.۳۴۴	۰.۶۷۲	۰.۳۳۶	۱.۰۹۶	۰.۵۴۸	۰.۲۷۴	۰.۹۷۲	۰.۴۸۶	۰.۲۴۳
۲۰	۱.۴۸۷	۰.۷۴۳	۰.۳۷۲	۱.۳۶۲	۰.۶۸۱	۰.۳۴۰	۱.۱۱۴	۰.۵۵۷	۰.۲۷۸	۰.۹۸۹	۰.۴۹۴	۰.۲۴۷

	$Var[\hat{I}(F_{R_{[r]}},F_Y)]$											
$\theta_۲$	۰.۵	۱	۲	۰.۵	۱	۲	۰.۵	۱	۲	۰.۵	۱	۲
n	$\alpha=-۱$			$\alpha=-۰/۵$			$\alpha=۰/۵$			$\alpha=۱$		
۱۰	۰.۲۴۱	۰.۰۶۰	۰.۰۱۵	۰.۲۰۵	۰.۰۵۱	۰.۰۱۳	۰.۱۴۴	۰.۰۳۶	۰.۰۰۹	۰.۱۱۹	۰.۰۳۰	۰.۰۰۷
۱۵	۰.۱۶۵	۰.۰۴۱	۰.۰۱۰	۰.۱۴۱	۰.۰۳۵	۰.۰۰۹	۰.۱۰۰	۰.۰۲۵	۰.۰۰۶	۰.۰۸۳	۰.۰۲۱	۰.۰۰۵
۲۰	۰.۱۲۶	۰.۰۳۱	۰.۰۰۸	۰.۱۰۸	۰.۰۲۷	۰.۰۰۷	۰.۰۷۷	۰.۰۱۹	۰.۰۰۵	۰.۰۶۴	۰.۰۱۶	۰.۰۰۴

جدول ۲.۵: مقادیر عددی $\mathbb{E}[\hat{I}(F_{R_{[r]}}, F_Y)]$ و $Var[\hat{I}(F_{R_{[r]}}, F_Y)]$ برای MTBUD با $\theta_1 = \theta_2 = 1$

	$\mathbb{E}[\hat{I}(F_{R_{[r]}}, F_Y)]$				$Var[\hat{I}(F_{R_{[r]}}, F_Y)]$			
n	$\alpha = -1$	$\alpha = -0/5$	$\alpha = 0/5$	$\alpha = 1$	$\alpha = -1$	$\alpha = -0/5$	$\alpha = 0/5$	$\alpha = 1$
۱۰	۰.۲۸۵	۰.۲۵۴	۰.۱۹۲	۰.۱۶۲	۰.۰۰۸	۰.۰۰۷	۰.۰۰۴	۰.۰۰۳
۱۵	۰.۲۹۷	۰.۲۶۴	۰.۲۰۰	۰.۱۶۸	۰.۰۰۶	۰.۰۰۵	۰.۰۰۳	۰.۰۰۲
۲۰	۰.۳۰۲	۰.۲۷۰	۰.۲۰۴	۰.۱۷۱	۰.۰۰۵	۰.۰۰۴	۰.۰۰۲	۰.۰۰۱

قضیه ۱.۴.۵. فرض کنید $(X_i, Y_i),\ i = 1, 2, \cdots, n$ یک نمونه‌ی تصادفی با حجم n از خانواده‌ی مورگنسترن باشد. آن‌گاه

$$\hat{I}(F_{R_{[r]}}, F_Y) \longrightarrow I(F_{R_{[r]}}, F_Y) \quad a.s. \;\; as\; n \to \infty.$$

برهان. از رابطه‌ی (۳۳.۵) به‌دست می‌آوریم

$$\hat{I}(F_{R_{[r]}}, F_Y) = \left[1 + \alpha\left(2^{1-r} - 1\right)\right] \hat{\mathcal{CE}}(Y) - \frac{\alpha}{2}\left(2^{1-r} - 1\right) \hat{\mathcal{CE}}(Y_{(2:2)}).$$

از آنجا که $\hat{\mathcal{CE}}(Y) \longrightarrow \mathcal{CE}(Y)$ و $\hat{\mathcal{CE}}(Y_{(2:2)}) \longrightarrow \mathcal{CE}(Y_{(2:2)})$، از دی‌کرشنزو و لونگوبردی (۲۰۰۹) اثبات به‌دست می‌آید.

□

قضیه ۲.۴.۵. فرض کنید $(X_i, Y_i),\ i = 1, 2, ..., n$ یک نمونه ی تصادفی از MTBGED با $\lambda = 1$ باشد. آن‌گاه

$$Z_n := \frac{\hat{I}(F_{R_{[r]}}, F_Y) - E\left[\hat{I}(F_{R_{[r]}}, F_Y)\right]}{\sqrt{Var\left[\hat{I}(F_{R_{[r]}}, F_Y)\right]}}$$

در توزیع به یک متغیر نرمال استاندارد میل می‌کند هرگاه $n \to \infty$.

برهان. در ابتدا اندازه‌ی تجربی $\hat{I}(F_{R_{[r]}}, F_Y)$ می‌تواند به‌صورت جمع متغیرهای تصادفی مستقل، به شکل زیر نوشته شود

$$\hat{I}(F_{R_{[r]}}, F_Y) = \sum_{j=1}^{n-1} W_j,$$

که $W_j = U_j \frac{j}{n}\left(-\log\frac{j}{n}\right)\left[1 + \alpha\left(2^{1-r} - 1\right)\left(1 - \frac{j}{n}\right)\right]$ متغیرهای تصادفی مستقل با میانگین و واریانس زیر هستند

$$E[W_j] = \frac{1}{n\theta_2\left(1 - \frac{1}{j/n}\right)}\left(\log\frac{j}{n}\right)\left[1 + \alpha\left(2^{1-r} - 1\right)\left(1 - \frac{j}{n}\right)\right],$$

$$Var[W_j] = \frac{1}{n^2\theta_2^2\left(1 - \frac{1}{j/n}\right)^2}\left(\log\frac{j}{n}\right)^2\left[1 + \alpha\left(2^{1-r} - 1\right)\left(1 - \frac{j}{n}\right)\right]^2.$$

ازآنجاکه برای هر متغیر تصادفی از توزیع نمایی داریم $E[|W_j-E(W_j)|^{۳}] = ۲e^{-۱}(۶-e)[E(W_j)]^{۳}$، با قرار دادن $\alpha_{j,k} = E[|W_j - E(W_j)|^k]$ تقریب زیر برای n بزرگ برقرار می‌شود:

$$\sum_{j=۱}^{n} \alpha_{j,۲} = \frac{۱}{n^{۲}\theta_{۲}^{۲}} \sum_{j=۱}^{n} \frac{۱}{\left(۱-\frac{۱}{j/n}\right)^{۲}} \left(\log \frac{j}{n}\right)^{۲} \left[۱+\alpha\left(۲^{۱-r}-۱\right)\left(۱-\frac{j}{n}\right)\right]^{۲} \approx \frac{c_{۲}}{n\theta_{۲}^{۲}},$$

$$\sum_{j=۱}^{n} \alpha_{j,۳} = \frac{۲(۶-e)}{en^{۳}\theta_{۲}^{۳}} \sum_{j=۱}^{n} \frac{۱}{\left(۱-\frac{۱}{j/n}\right)^{۳}} \left(\log \frac{j}{n}\right)^{۳} \left[۱+\alpha\left(۲^{۱-r}-۱\right)\left(۱-\frac{j}{n}\right)\right]^{۳} \approx \frac{۲(۶-e)c_{۳}}{en^{۲}\theta_{۲}^{۳}},$$

که

$$c_k := \int_{۰}^{۱} \left(\frac{\log x}{۱-۱/x}\right)^{k} \left[۱+\alpha\left(۲^{۱-r}-۱\right)\left(۱-\frac{j}{n}\right)\right]^{k}.$$

بنابراین، شرط لیاپانوف از قضیه‌ی حد مرکزی برقرار می‌شود

$$\frac{(\alpha_{۱,۳}+\cdots+\alpha_{n,۳})^{۱/۳}}{(\alpha_{۱,۲}+\cdots+\alpha_{n,۲})^{۱/۲}} \approx \frac{[۲(۶-e)c_{۳}]^{۱/۳}}{e^{۱/۳}c_{۲}^{۱/۲}} n^{-۱/۶} \to ۰ \quad as\ \ n\to\infty,$$

که اثبات را کامل می‌کند. □

کاربردی از مفهوم عدم دقت تجمعی در پردازش تصاویر

در این بخش سعی داریم که بر روی کاربرد مناسب اطلاع کولبک–لایبلر تجمعی در تحلیل تصویر تمرکز کنیم؛ برای این منظور ابتدا به مسئله تخمین اطلاع کولبک–لایبلر تجمعی با استفاده از یک اندازه‌گیری تجربی از تشخیص می‌پردازیم.

تعریف: فرض کنید X و Y متغیرهای تصادفی با توابع توزیع $F(\cdot)$ و $G(\cdot)$ باشند، اطلاع کولبک–لایبلر تجمعی تجربی بین X و Y به‌صورت زیر محاسبه می‌شود

$$K(\hat{F}_n|\hat{G}_m) = \int_{۰}^{+\infty} \hat{F}_n(x)\log\left(\frac{\hat{F}_n(x)}{\hat{G}_m(x)}\right) dx + \bar{X}_n - \bar{Y}_m,$$

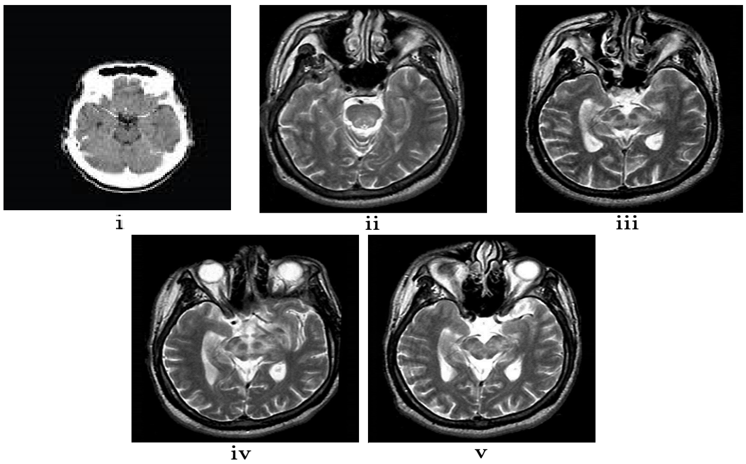
که در آن $\hat{F}_n(x) = \frac{۱}{n}\sum_{i=۱}^{n} I_{X_i\leqslant x}$ و $\hat{G}_m(x) = \frac{۱}{m}\sum_{i=۱}^{m} I_{Y_i\leqslant x}$ توزیع تجربی متغیرهاست و $\bar{X}_n$، $\bar{Y}_m$ میانگین‌های نمونه هستند. برای بدست آوردن یک فرمول موثر برای $K(\hat{F}_n|\hat{G}_m)$ آماره‌های ترتیبی را به‌صورت $X_{(۱)} \leqslant X_{(۲)} \leqslant \ldots \leqslant X_{(n)}$ نشان می‌دهیم. به‌طور مشابه آماره‌های ترتیبی Y نیز به‌صورت $Y_{(۱)} \leqslant Y_{(۲)} \leqslant \ldots \leqslant Y_{(m)}$ تعریف می‌شود.

نکته: طبق رابطه (۱۲.۱) می‌دانیم که اطلاع کولبک–لیبلر تجمعی را می‌توان به‌صورت زیر بازنویسی کرد

$$K(\hat{F}_n|\hat{G}_m) = I[\hat{F}_n, \hat{G}_m)] - \mathcal{CE}(\hat{F}_n) + \bar{X}_n - \bar{Y}_m. \qquad (۳۸.۵)$$

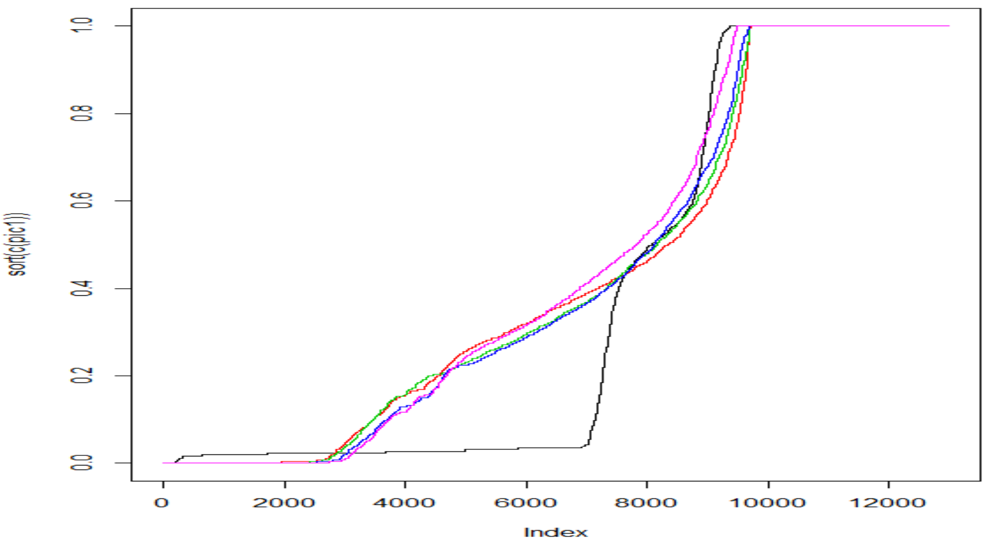
مثال کاربردی ۱: حال می‌خواهیم اطلاع کولبک–لایبلر تجمعی تجربی را برای اندازه‌گیری اختلافات در سطح خاکستری تصاویر دیجیتالی شده در شکل ۱.۵ بکار بگیریم.

تصاویر توسط ۱۶۰۰ سلول تشکیل شده‌است. سطح خاکستری از هر سلول توسط یک عدد واقعی از صفر (سیاه) تا یک (سفید) اندازه‌گیری شده‌اند. شکل ۲.۵ سطح خاکستری مرتب شده در جهت افزایش را نشان می‌دهد. بنابراین میانگین سطح خاکستری به سمت مقدار یک، وضوح تصویر بیشتر و به سمت صفر عدم وضوح را نشان می‌دهد. علاوه بر این متوسط سطح خاکستری از پنج تصویر در جدول ۳.۵ ارائه شده‌است.
اطلاع کولبک ـ لایبلر تجمعی تجربی از سطوح خاکستری برای جفت تصاویر با استفاده از معادله (۳۸.۵) برای $n = m = ۱۶۰۰$ در جدول ۴.۵ ارزیابی می‌شود. حداکثر اختلاف بین شکل‌های $(i), (v)$ مشاهده می‌شود، در حالیکه حداقل مقدار $K(\hat{F}_n|\hat{G}_m)$ مربوط به شکل‌های $(vi), (v)$ هستند.
یک کاربرد مناسب از اطلاع کولبک_لایبلر تجمعی تجربی در تحلیل تصویر را نشان دادیم. در این زمینه، با شروع از سطوح خاکستری تصاویر در نمونه‌های تصادفی مقادیر بالاتر از $K(\hat{F}_n|\hat{G}_m)$ تفاوت قابل توجهی در شکل‌ها داشتند. بنابراین یک روش خودکار که اطلاع کولبک_لایبلر تجمعی تجربی را ارزیابی می‌کند قادر است تصاویر را براساس سطح خاکستری یا (رنگی) طبقه‌بندی کند.



شکل ۱.۵: تصاویر اسکن مغز

برای درک بهتر کاربرد اطلاع کولبک_لایبلر تجمعی تجربی در اندازه‌گیری اختلافات سطح خاکستری تصاویر به مثال دیگری می‌پردازیم.
مثال کاربردی ۲: همانند مثال قبل سطح خاکستری از هر سلول توسط یک عدد واقعی از صفر (سیاه) تا یک (سفید) اندازه‌گیری شده‌است. هرچه میانگین سطح خاکستری بیشتر باشد اختلاف بین دو تصویر کمتر است به این معنا که هر دو تصویر از نظر وضوح تصویر و سطح خاکستری نزدیک به هم هستند. در جدول ۵.۵ همانطور که مشاهده می‌کنیم شکل (i) دارای بیشترین سطح خاکستری می‌باشد. این به آن معنی است که وضوح شکل (i) به سلول‌های



شکل ۲.۵: میزان سطح خاکستری مرتب‌شده در شکل ۱.۵

شکل	میانگین سطح خاکستری
i	۰/۲۸۰۷۱
ii	۰/۵۶۲۹۷
iii	۰/۵۹۲۳۴
iv	۰/۶۳۲۱۹
v	۰/۷۸۵۰۲

جدول ۳.۵: میانگین سطح خاکستری برای تصاویر شکل ۱.۵

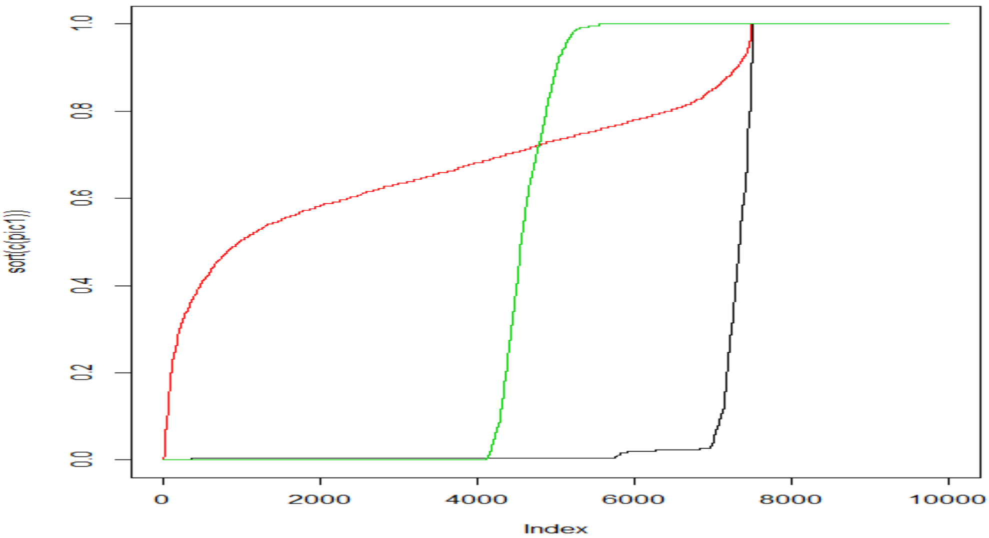
X,Y	i	ii	iii	iv	v
i	۰	۰/۰۹۳۷۲	۰/۱۲۸۵۹	۰/۲۱۸۴۲	۰/۴۰۲۳۵
ii	۰/۰۷۵۵۰	۰	۰/۰۷۲۹۵	۰/۰۸۹۹۰	۰/۰۵۴۳۵
iii	۰/۱۸۲۹۱	۰/۳۴۴۱۵	۰	۰/۰۱۲۸۷	۰/۰۷۲۹۵
iv	۰/۱۲۷۵۱	۰/۰۲۱۶۸	۰/۱۴۹۴۲	۰	۰/۰۰۱۴۸
v	۰/۳۳۲۱۷	۰/۰۶۶۲۴	۰/۳۳۲۱۷	/۰۱۵۱۲	۰

جدول ۴.۵: اندازه کولبک_لایبلر تجمعی تجربی برای تصاویر شکل ۱.۵

سفید مربوط می‌شود، همچنین اطلاع کولبک_لایبلر تجمعی تجربی از سطح خاکستری برای جفت تصاویر با استفاده از معادله (۳۸.۵) ارزیابی شده‌است. از جدول ۶.۵ می‌بینیم که حداکثر اختلاف بین شکل‌های (ii) و(i) می‌باشد، در حالیکه حداقل مقدار $K(\hat{F}_n|\hat{G}_m)$ مربوط

به شکل‌های (ii) و (iii) هستند. به این منظور که اگر اختلاف کولبک_لایبلر تجمعی تجربی بین دو شکل زیاد باشد یعنی تفاوت زیادی در سطح خاکستری تصاویر و وضوح تصویر دارند و اگر تفاوت بین دو تصویر کم باشد بدین معنی است که سطح خاکستری در دو تصویر نزدیک به هم است و وضوح هر دو تصویر یکسان است.

شکل ۳.۵: تصاویر منتخب



شکل ۴.۵: سطح خاکستری مرتب‌شده برای تصاویر شکل ۳.۵

شکل	میانگین سطح خاکستری
i	۰/۷۴۰۶۱۸۸
ii	۰/۲۷۳۷۰۳۵
iii	۰/۵۳۹۰۸۲۴

جدول ۵.۵: میانگین سطح خاکستری برای تصاویر شکل ۳.۵

X, Y	i	ii	iii
i	۰	۰/۳۵۷۷۷۶۵	۰/۲۰۱۴۲۶۲
ii	۱/۱۱۴۲۵۶	۰	۰/۶۷۵۵۰۱۷
iii	۰/۵۷۳۷۰۷۰	۰/۰۵۹۳۲۵۸۷	۰

جدول ۶.۵: اندازه‌های کولبک‌ـ‌لایبلر تجمعی تجربی برای تصاویر شکل ۳.۵

آینده‌ی تحقیق

در این بخش به طرح مسائل و پیشنهاداتی روی معیار عدم دقت می‌پردازیم:

۱. خواص و ویژگی‌های دیگری از تابع عدم دقت $I(f_{i:n}, f)$ در سیستم‌های سری، موازی و منسجم چیست؟

۲. تعمیم‌های جدید از معیار عدم دقت در متغیرهای ترتیبی چگونه است؟

۳. تابع عدم دقت برای سانسورها چیست و چه ویژگی‌هایی دارد؟

۴. کاربرد معیارهای جدید عدم دقت در پردازش تصاویر چیست؟

۵. کاربرد معیارهای عدم دقت به عنوان یک ریسک در داده‌های مالی و اقتصادی چیست؟

۶. معیارهای عدم دقت تجمعی در سیستم‌های پیچیده (منسجم) چیست و چه ویژگی‌هایی دارد؟

۷. معیارهای عدم دقت وزنی در توابع دو متغیره چگونه است؟

مراجع

[1] Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (1992), ”**A First Course in Order Statistics**”, John Wiley and Sons , New York.

[2] Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (1998), ”**Records**”, John Wiley and Sons, New York.

[3] Asadi, M. and Zohrevand, Y. (2007), ”On the dynamic cumulative residual entropy”, ”**Journal of Statistical Planning and Inference**”, 137(6), pp. 1931-41.

[4] Baratpour, S. (2010), ”Characterizations based on cumulative residual entropy of first-order statistics”, ”**Communications in Statistics-Theory and Methods**”, 39(20), pp. 3645-3651.

[5] Baratpour, S., Ahmadi, J. and Arghami, N.R. (2007a), ”Entropy properties of record statistics”, ”**Statistical Papers**”, 48(1), pp. 197-213.

[6] Baratpour, S., Ahmadi, J. and Arghami, N.R. (2007b), ”Some characterizations based on entropy of order statistics and record values”, ”**Communications in Statistics - Theory and Methods**”, 36(1), pp. 47-57.

[7] Baratpour, S., Ahmadi, J., Arghami, N.R. (2008), ”Characterizations based on Renyi entropy of order statistics and record values”, **Journal of statistical planning and inference** 138, 2544–2551.

[8] Barnett, V. and Lewis, T. (1984), ”**Outliers in Statistical Data**”, (2nd ed.), Wiley, New York.

[9] Basak, P. and Balakrishnan, N. (2003), ”Maximum likelihood prediction of future record statistic”, ”**IMathematical Statistical Methods in Reliability. In: Series on Quality, Reliability and Engineering Statistics, World Scientific Publishing, Singapore**”, 7, 159-175.

[10] Beg, M.I. and Ahsanullah, M. (2008), ”Concomitants of generalized order statistics from Farlie-Gumbel-Morgenstern distributions”, ”**Statistical Methodology**”, 5(1), pp. 1–20.

[11] Bdair, O.M. and Raqab, M.Z. (2012), ”Sharp upper bounds for the mean residual waiting time of records”, ”**Statistics**”, 46(1), pp. 69-84.

[12] Billingsley, P. (1986), ”**Probability and Measure (2nd ed.)**”, Wiley, New York.

[13] Cali, C., Longobardi, M. and Navarro, J. (2018), ”Properties for generalized cumulative past measures of information”, ”**Probability in the Engineering and Informational Sciences**”, doi.org/10.1017/S0269964818000360

[14] Castillo, E. (2012), ”**Extreme Value Theory in Engineering**”, Elsevier.

[15] Chandler, K.N. (1952), ”The distribution and frequency of record values”, ”**Journal of the Royal Statistical Society. Series B (Methodological)**”, 14(2), pp. 220-228.

[16] Cover, T.M. and Thomas, J.A. (1991), ”**Elements of information theory**”, John Wiley and Sons, New York.

[17] Daneshi, S., Nezakati, A. and Tahmasebi, S. (2019), ”On weighted cumulative past (residual) inaccuracy for record values”, ”**Journal of Inequalities and Applications**”, 1, pp. 134.

[18] David, H.A. (1970), ”**Order Statistics**”, (1st ed.), Wiley, New York.

[19] David, H.A. (1981), ”**Order Statistics**”, (2nd ed.), Wiley, New York.

[20] David, H.A. and Nagaraja, H.N. (2003), ”**Order Statistics**”, Wiley, New York.

[21] Di Crescenzo, A. and Longobardi, M. (2006), ”On weighted residual and past entropies”, ”**Scientiae Mathematicae Japonicae**”, 64(2), pp. 255-266.

[22] Di Crescenzo, A. and Longobardi, M. (2009), ”On cumulative entropies”, ”**Journal of Statistical Planning and Inference**”, 139, pp. 4072-4087.

[23] Durante, F., Sempi, C. (2015), ”**Principles of Copula Theory**”, CRC Press, Boca Raton, FL.

[24] Ebrahimi, N. (1996), ”How to measure uncertainty in the residual lifetime distributions”, ”**Sankhya Ser**”, A 58, 48–57.

[25] Ebrahimi, N., Soofi, E.S. and Zahedi, H. (2004), ”Information properties of order statistics and spacings”, ”**IEEE Transactions on Information Theory**”, 50, 177–183.

[26] Eryilmaz, S. (2005), ”Concomitants in a sequence of independent nonidentically distributed random vectors”, ”**Communications in Statistics—Theory and Methods**”, 34(9-10), pp. 1925-1933.

[27] Eskandarzadeh M., Di Crescenzo, A. and Tahmasebi, S. (2019), ”Cumulative measure of inaccuracy and mutual information in k-th lower record values”, ”**Mathematics**”, 7(2), pp.175, https://doi.org/10.3390/math7020175.

[28] Galambos, J. (1978,1987), ”**The Asymptotic Theory of Extreme Order Statistics**”, Wiley, New York (1st ed.). Kreiger, FL (2nd ed.).

[29] Galambos, J. and Kotz, S. (1978), ”**Characterizations of Probability Distributions**”, Lecture Notes in Mathematics No. 675, Springer-Verlag, New York.

[30] Ghosh, A. and Kundu, C. (2017), ”Bivariate extension of (dynamic) cumulative residual and past inaccuracy measures”, ”**Stat Papers**”, DOI 10.1007/s00362-017-0917-5.

[31] Gupta, R.C. and Kirmani, S.N.U.A. (2008), ”Characterizations based on conditional mean function”, ”**Journal of Statistical Planning and Inference**”, 138, pp. 964-970.

[32] Gupta, R.C., Taneja, H.C. and Thapliyal, R. (2014), ” Stochastic comparisons based on residual entropy of order statistics and some characterization results”, ”**Journal of Statistical Theory and Applications**”, 13 (1), 27–37.

[33] Harter, H.L. (1970), ”**Order Statistics and Their Use in Testing and Estimation**”, (Vol. 1 and 2), U.S. Government Printing Office, Washington, D.C.

[34] Hartley, R.V. (1928), ”Transmission of information1”, ”**Bell System technical journal**”, 7(3), 535-563.

[35] Kamps, U. (1995), ”**A Concept of Generalized Order Statistics**”, Teubner Skripten zur Mathematischen Stochastik. Teubner, Stuttgart.

[36] Kayal, S. (2016), ”On generalized cumulative entropies”, ”**Probability in the Engineering and Informational Sciences**”, 30(4), pp. 640-662.

[37] Kayal, S. (2018), ”On weighted generalized cumulative residual entropy of order n”, ”**Methodology and Computing in Applied Probability**”, 20(2), pp. 487-503.

[38] Kayal, S. and Moharana, R. (2017), ”On weighted measures of cumulative entropy”, ”**International Journal of Mathematics and Statistics**”, 18(3), pp. 26-46.

[39] Kayal, S. and Moharana, R. (2018), ”A shift-dependent generalized cumulative entropy of order n”, ”**Communications in Statistics - Simulation and Computation**”, doi.org/10.1080/03610918.2018.1423692.

[40] Kayid, M., Izadkhah, S. and Alhalees, H. (2016), ”Combination of mean residual life order with reliability applications”, ”**Statistical Methodology**”, 29, pp. 51-69.

[41] Kerridge, D.F. (1961), ”Inaccuracy and inference”, ”**Journal of the Royal Statistical Society, Series B (Methodological)**”, 23, pp. 184-194.

[42] Kullback, S., (1959), ”**Information Theory and Statistics**”, Wiley, New York.

[43] Kullback, S. and Leibler, R.A. (1951), ”On information and sufficiency”, ”**The Annals of Mathematical Statistics**”, 22(1), 79-86.

[44] Kumar, V. and Taneja, H.C. (2015), ”Dynamic cumulative residual and past inaccuracy measures”, ”**Journal of Statistical Theory and Applications**, 14(4), pp. 399-412.

[45] Kundu, C. (2017), ”On weighted measure of inaccuracy for doubly truncated random variables”, ”**Communications in Statistics - Theory and Methods**”, 46(7), pp. 3135-3147.

[46] Kundu, C., Di Crescenzo, A. and Longobardi, M. (2016), ”On cumulative residual (past) inaccuracy for truncated random variables”, ”**Metrika**”, 79 , pp. 335-356.

[47] Lindley, D.V. (1957), ”A statistical paradox”, ”**Biometrika**”, 44(1/2), 187-192.

[48] Morgenstern, D. (1956), ”Einfache Beispiele zweidimensionaler verteilungen”, ”**Mitteilungsblatt für Mathematische Statistik**”, 8, pp. 234-235.

[49] Misagh, F. and G.H. Yari, G.H. (2011), ”On weighted interval entropy”, ”**Statistics and Probability Letters**”, 81(2), pp. 188-194, 2011.

[50] Navarro, J., del Aguila, Y. and Asadi, M. (2010), ”Some new results on the cumulative residual entropy”, ”**Journal of Statistical Planning and Inference**, 140, pp. 310-322.

[51] Nyquist, H. (1932), ”Regeneration theory”, ”**Bell system technical journal**”, 11(1), 126-147.

[52] Park, S. (1995), ”The entropy of consecutive order statistics”, ”**IEEE Transactions on Information Theory**”, 41(6), 2003-2007.

[53] Park, S. (1996), ”Fisher information in order statistics”, ”**Journal of the American Statistical Association**”, 91(433), 385-390.

[54] Park, S. and Park, D. (2003), ”Correcting moments for goodness of fit tests based on two entropy estimates”, ”**Journal of Statistical Computation and Simulation**”, 73(9), 685-694.

[55] Psarrakos, G. and Navarro, J. (2013), ”Generalized cumulative residual entropy and record values”, ”**Metrika**”, 76, pp. 623-640.

[56] Pyke, R. (1965), ”Spacings”, ”**Journal of Royal Statistical Society, Series B (Methodological)**”, pp. 395-449.

[57] Rao, M. (2005), ”More on a new concept of entropy and information”, ”**Journal of Theoretical Probability**”, 18, pp. 967-981.

[58] Rao, M., Chen, Y., Vemuri, B.C. and Wang, F. (2004), ”Cumulative Residual Entropy: A New Measure of Information”,”**IEEE transactions on Information Theory**”, 50(6), pp. 1220-1228.

[59] Renyi, A. (1961), ”On measures of entropy and information”, ”**In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics**”, The Regents of the University of California, pp. 547–561.

[60] Sarhan, A.E. and Greenberg, B.G. (Eds.) (1962a), ”**Contributions to Order Statistics**”, Wiley, New York.

[61] Shaked, M., Shanthikumar, J.G., 2007. ”**Stochastic Orders**”, Springer, New York.

[62] Shannon, C.E. (1948), ”A mathematical theory of communication”, ”**Bell System Technical Journal**, 37, pp. 379-432.

[63] Tahmasebi, S. and Daneshi, S. (2018), ”Measures of inaccuracy in record values”, ”**Communications in Statistics - Theory and Methods**”, 47 (24), pp. 6002-6018.

[64] Tahmasebi, S., Nezakati, A. and Daneshi, S. (2018), ”Results on cumulative measure of inaccuracy in record values”, ”**Journal of Statistical Theory and Applications**”, 17 (1), pp. 15-28.

[65] Taneja, H.C., Kumar, V. and Srivastava, R. (2009), ”A dynamic measure of inaccuracy between two residual lifetime distributions”, ”**International Mathematical Fourm**”, 4(25), pp. 1213-1220.

[66] Thapliyal, R. and Taneja, H.C. (2013), ”A measure of inaccuracy in order statistics”, ”**Journal of Statistical Theory and Applications**”, 12(2), pp. 200-207.

[67] Thapliyal, R. and Taneja, H.C. (2015a), ”Dynamic cumulative residual and past inaccuracy measures”, ”**Journal of Statistical Theory and Applications**”, 14(4), pp. 399-412.

[68] Thapliyal, R. and Taneja, H.C. (2015b), ”On residual inaccuracy of order statistics”, ”**Statistics Probability Letters**”, 97, pp. 125-131.

[69] Wang, S. (1998), ”An actuarial index of the right-tail risk”, ”**North American Actuarial Journal**”, 2, pp. 88-101.

[70] Wilks, S.S. (1959), ”Recurrence of extreme observations”, ”**Journal of the Australian Mathematical Society**”, 1(01), 106-112.

[71] Wong, K.M. and Chen, S. (1990), ”The entropy of ordered sequences and order statistics”, ”**IEEE Transactions on Information Theory**”, 36(2), pp. 276–284.

[72] Zardasht, V. (2019), ”Results on relative mean residual life and relative cumulative residual entropy”,”**Statistics, Optimization and Information Computing**”, 7(1), pp. 150-159.

پیوست آ

تعاریف

آ.۱ قضیه همگرایی مغلوب

اگر $\{f_n\}$ دنباله‌ای از توابع اندازه‌پذیر بر یک فضای اندازه (S, Σ, μ) باشد و

$$\lim_{n\to\infty} f_n(x) = f(x) \quad a.e.\, \mu$$

و یک تابع انتگرال‌پذیر g موجود باشد که

$$|f_n(x)| \leq g(x) \quad \forall x$$

آنگاه $\{f_n\}$ و f انتگرال‌پذیر هستند و

$$\int f d\mu = \lim_{n\to\infty} \int f_n d\mu.$$

آ.۲ قضیه گلیونکو_ کانتلی

فرض کنید $X_۱, X_۲, ..., X_n$ متغیرهای تصادفی مستقل و هم‌توزیع با تابع توزیع $F(x)$ باشند. اگر تابع توزیع تجربی را با $F_n(X)$ نشان دهیم، آن‌گاه

$$\sup_x |F_n(x) - F(x)| \to^{a.s.} ۰ \quad as \quad n \to \infty.$$

آ.۳ شاخص جینی

شاخص جینی معیاری مشهور از نابرابری درآمد است و به صورت زیر تعریف می‌شود (وانگ (۱۹۹۸) را ببینید)

$$gini[X] = ۱ - \frac{\int_{۰}^{\infty}[\bar{F}(x)]^{۲}dx}{E(X)}.$$

آ.۴ قضیه فوبینی

اگر تابع دو متغیره $f(x,y)$ بر روی ناحیه مستطیلی $R = \{(x,y)|a \leq x \leq b, c \leq y \leq d\}$ پیوسته باشد، آن‌گاه داریم

$$\int_c^d \int_a^b f(x,y)dxdy = \int_a^b \int_c^d f(x,y)dydx.$$

Aabstract

One of the main subjects of statistics, is the information theory. Nowadays the information and inaccuracy measures have a special place in reliability. In this thesis, after expressing preliminaries of inaccuracy, introduce the information measures and ordered random variables, we consider a measure of inaccuracy between distributions of the nth upper (lower) record value and parent random variable. We also express a measure of inaccuracy between distributions of the ith order statistics and parent random variable. We propose weighted cumulative past (residual) inaccuracy in record values and order statistics and we obtain a measure of inaccuracy between rth concomitant of generalized order statistic and the parent random variable in Morgenstern family. For these concepts, we obtain some properties and characterization results such as relationships with other reliability functions, bounds, stochastic ordering and effect of linear transformation. Dynamic versions of these measures are considered. We study on a problem of estimating the measure of inaccuracy by means of the empirical cumulative inaccuracy and weighted cumulative past (residual) inaccuracy in lower record values.

Keywords: Measure of inaccuracy, Cumulative inaccuracy, Cumulative entropy, Record values, Order statistics, Weighted inaccuracy, Empirical approach.

Shahrood University of Technology

Faculty of Mathematical Sciences

PhD Thesis in: Probability

Measures of inaccuracy in some ordered random variables

By: Safieh Daneshi

Supervisors

Dr. Ahmad Nezakati
Dr. Saeid Tahmasebi

September 2019