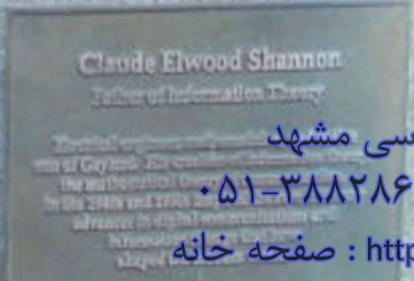
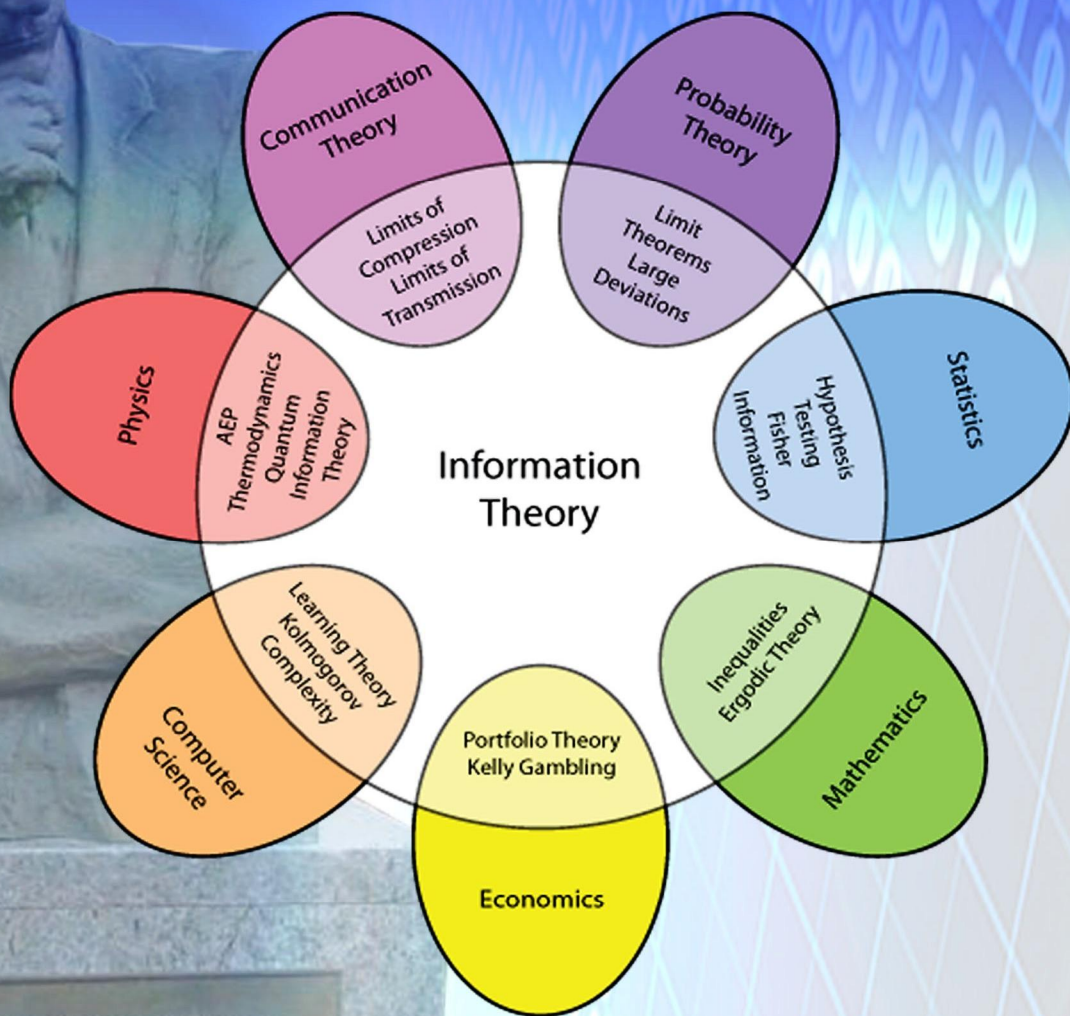


# دومین کارگاه اندازه های اطلاعات و کاربرد های آن

قطب علمی داده های ترتیبی و فضایی  
گروه آمار، دانشگاه فردوسی مشهد  
هشتم و نهم بهمن ماه ۱۳۹۳



آدرس : گروه آمار، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد  
صندوق پستی ۹۱۷۷۵-۱۱۵۹  
تلفن و فکس ۰۵۱-۳۸۸۲۸۶۰۵  
پست الکترونیک : [osdce@um.ac.ir](mailto:osdce@um.ac.ir)  
<http://osdce.um.ac.ir> : صفحه خانه

باسمه تعالی



مجموعه مقالات

دومین کارگاه آموزشی

# اندازه های اطلاعات و کاربردهای آن

قطب علمی داده های ترتیبی و فضایی

گروه آمار، دانشکده علوم ریاضی

دانشگاه فردوسی مشهد

۸ و ۹ بهمن ماه ۱۳۹۳

## پیشگفتار

به منظور تبادل نظر اعضای هیأت علمی و ارتقای سطح علمی دانشجویان تحصیلات تکمیلی دانشگاه‌ها، مراکز آموزش عالی کشور و کاربران نظریه اطلاعات و کاربردهای آن، دومین کارگاه آموزشی روزهای ۸ و ۹ بهمن ماه سال ۱۳۹۳ توسط قطب علمی داده‌های ترتیبی و فضایی دانشگاه فردوسی مشهد و با همکاری گروه‌های آمار دانشگاه‌های فردوسی مشهد، اصفهان، تربیت مدرس، تهران، رازی و انجمن آمار ایران در محل دانشکده علوم ریاضی دانشگاه فردوسی مشهد، برگزار گردید. در این کارگاه دو روزه بیش از ۶۰ نفر از اعضای هیأت علمی، دانشجویان تحصیلات تکمیلی و کاربران مباحث مرتبط با نظریه اطلاعات از دانشگاه‌ها و مراکز علمی شرکت نمودند. پس از فراخوان کارگاه، کمیته علمی اقدام به بررسی مقالات دریافتی نموده و ارائه بیست و سه مقاله را مورد موافقت قرار داد.

خلاصه سخنرانی‌های عمومی این کارگاه در پایگاه اطلاعاتی قطب علمی داده‌های ترتیبی و فضایی دانشگاه فردوسی مشهد به نشانی <http://osdce.um.ac.ir> قابل دسترسی است. متن کامل تعداد هفت مقاله‌ی فارسی و پانزده مقاله‌ی انگلیسی پذیرفته شده پس از تأیید کمیته علمی در مجموعه مقالات تدوین شده است که شامل دو بخش مقالات انگلیسی و مقالات فارسی می‌باشد.

این مجموعه شامل مقالات سخنرانی و پوستره‌های فارسی و انگلیسی ارائه شده در دومین کارگاه آموزشی اندازه‌های اطلاعات و کاربردهای آن است. بدیهی است که ارائه دهنده سخنرانی یا پوستر مسئول صحت مطالب، اصالت آنها و ویرایش علمی-ادبی مقاله خود بوده است و قطب علمی داده‌های ترتیبی و فضایی مسئولیتی در این خصوص ندارد.

برگزاری این کارگاه علاوه بر تلاش خستگی ناپذیر کمیته‌های علمی و اجرایی، مرهون پشتیبانی گروه آمار دانشگاه فردوسی مشهد و مسئولین محترم دانشکده علوم ریاضی بوده است. بدین وسیله مراتب سپاس و قدردانی قطب علمی داده‌های ترتیبی و فضایی و گروه آمار دانشگاه فردوسی مشهد را ابراز می‌نماییم.

با آرزوی توفیق الهی

قطب علمی داده‌های ترتیبی و فضایی

### دبیر کارگاه:

دکتر غلامرضا محتشمی برزادران، دانشگاه فردوسی مشهد

### اعضای کمیته علمی (به ترتیب حروف الفبا):

- دکتر مجید اسدی، دانشگاه اصفهان
- دکتر بهاءالدین خالدی، دانشگاه رازی
- دکتر قوشه عابد هدتنی، دانشگاه فردوسی مشهد
- دکتر ماریا لنگ باردی، دانشگاه ناپلی، ایتالیا
- دکتر غلامرضا محتشمی برزادران، دانشگاه فردوسی مشهد

### اعضای کمیته اجرایی (به ترتیب حروف الفبا):

- دکتر جعفر احمدی، دانشگاه فردوسی مشهد
- دکتر هادی جباری نوقایی، دانشگاه فردوسی مشهد
- دکتر معصومه فشندی، دانشگاه فردوسی مشهد
- دکتر مهدی عمادی، دانشگاه فردوسی مشهد

### اعضای کمیته دانشجویی (به ترتیب حروف الفبا):

- زهره پاکدامن، دانشگاه فردوسی مشهد
- فاطمه صفائی، دانشگاه فردوسی مشهد
- حسینعلی محتشمی برزادران، دانشگاه فردوسی مشهد

### محورها:

- معیارهای اساسی در نظریه اطلاعات
- استنباط آماری بر مبنای اندازه‌های اطلاعات
- اندازه‌های اطلاعات و نظریه احتمال
- اندازه‌های اطلاعات و قابلیت اعتماد
- کاربرد اندازه‌های اطلاعات در مهندسی و علوم

# فهرست

## طراحی نمونه گیری فضایی بهینه با ملاک آنتروپی

بیدار بخت، ح.، محمدزاده، م.، محمدی، الف. و مرید، س. ..... ۱

## مشخص سازی بر اساس اندازه آنتروپی

حوتی، ف. ..... ۱۴

## آنتروپی بیشینه در مدل های صف بندی یک باجه ای با ورودی گروهی

شمس، م. و حسامیان، غ. .... ۲۳

## نظریه اطلاعات و جایگاه آن در علوم و مهندسی

عابد هدتنی، ق. .... ۳۱

## ماکسیمم اندازه آنتروپی تعمیم یافته تحت قید پارامتری کردن اندازه واگرایی تعمیم یافته

صانعی طبس، م. و محتشمی برزادران، غ. ر. .... ۳۲

## شاخص تایل بر اساس توزیع لمبدای تعمیم یافته و کاربردی از آن

کدخدا، ا.، محتشمی برزادران، غ. ر. و امینی، م. .... ۴۰

## برآورد اطلاع متقابل بر اساس تابع مفصل

محمدی، م.، عمادی، م.، امینی، م. و بزرگ، م. .... ۴۸

## کنکاشی در آنتروپی مفصل

ولی پور مقدم، ح. و محتشمی برزادران، غ. ر. .... ۵۷



## طراحی نمونه گیری فضایی بهینه با ملاک آنتروپی

بیداربخت، ح.ا<sup>۱</sup>، محمدزاده، م.ا<sup>۱</sup>، محمدی، الف.ا<sup>۱</sup>، مرید، س.ا<sup>۲</sup>

<sup>۱</sup>گروه آمار، دانشکده علوم ریاضی، دانشگاه تربیت مدرس

<sup>۲</sup>گروه سازه های آبی، دانشکده کشاورزی، دانشگاه تربیت مدرس

### چکیده

داده های بارندگی عناصر اصلی برای برنامه ریزی، ساخت و ساز و بهره برداری از سازه های هیدرولیکی هستند. یک شبکه بارندگی با طراحی خوب می‌تواند اطلاعات بارندگی در یک حوضه آبریز را به درستی فراهم کند. از آنجا که ایستگاه های باران سنجی موجود در حوزه آبریز استان خوزستان به مرور زمان ایجاد شده و به لحاظ تعداد و موقعیت از الگوی مشخصی پیروی نمی کنند، برای فراهم آوردن نتایج تحلیلی بهینه لازم است شبکه موجود مورد ارزیابی قرار گرفته و یک طرح نمونه گیری فضایی بهینه برای آن ارائه گردد. در این مقاله برای اصلاح شبکه ایستگاه های باران سنجی و به دست آوردن طرحی بهینه با حداقل تعداد و موقعیت های مناسب، روشی مبتنی بر میزان حتمیت پیشگویی های فضایی بارش باران بر اساس آنتروپی ارائه می شود، که ضمن ارزیابی شبکه موجود می توان راه حلی برای شناسایی و حذف ایستگاه های غیرضرور و ایجاد ایستگاه های مناسب ارائه و شبکه باران سنجی را بهینه نمود.

**کلمات کلیدی:** آنتروپی، بهینه سازی، پیشگویی فضایی، طرح نمونه گیری، استان خوزستان.

**مقدمه:**

داده‌های بارندگی اطلاعات بسیار مهمی برای برنامه‌ریزی، ساخت و راه‌اندازی سازه‌های آبی، تحلیل چرخه هیدرولوژیکی، توسعه و مدیریت منابع آب، مطالعه تغییر آب و هوا، مطالعه انحراف بارش و ارزیابی عدم قطعیت زیست محیطی فراهم می‌سازند. اگر مجموعه‌ای جامع از داده‌های بارندگی در یک حوضه آبریز طی یک دوره طولانی مدت جمع‌آوری شود امکان ایجاد و اجرای یک طرح کامل برای مدیریت منابع آب در منطقه با حتمیت و موفقیت بالا افزایش می‌یابد. از این‌رو، هدف از یک شبکه بارش، طراحی ساختارهای هیدرولیکی کارآمد و اقتصادی، بر اساس داده‌های بارش است (هاکت، ۱۹۶۶). چون، توزیع مکانی و زمانی بارش با توجه به توپوگرافی، الگوهای باران و اثراترمان، ناهموار است. طراحی شبکه ایستگاه‌های باران‌سنجی بهینه معمولاً چالش برانگیز است. حتی اگر دو ایستگاه باران‌سنجی نزدیک به هم باشند ضریب همبستگی دنباله‌های زمانی بارش کمتر از ۰/۵ می‌شود (کمپبل، ۱۹۸۳). این موضوع طراحی و تعدیل شبکه‌های بارندگی را پیچیده‌تر می‌سازد. بنابراین برای بهینه‌سازی شبکه بارندگی نیاز به برآوردند و معیار فضایی و زمانی است. اغلب مطالعات قبلی در زمینه تعیین موقعیت ایستگاه‌های باران‌سنجی بر اساس نحوه توزیع و حتی امکان پوشش کامل آنها در حوزه آبریز متمرکز بوده‌اند. با این حال، طراحی برحسب حداقل تعداد و تعیین موقعیت بهینه ایستگاه‌های باران‌سنجی نیازمند مطالعه‌های دقیق براساس ویژگی‌های منطقه و ساختار همبستگی فضایی میزان بارش باران در ایستگاه‌های مختلف است. ایستگاه‌های باران‌سنجی، با طرح‌های مبتنی بر نیازهای فناورانه و برنامه‌های کاربردی عملی، داده‌های بارش را فراهم می‌کنند، کیفیت داده‌های بارش به طور قابل توجهی تحت تاثیر جای گذاری و تنظیم ایستگاه‌های باران‌سنجی قرار دارند. یک شبکه بارندگی با طراحی خوب داده‌های بارش قابل اطمینانی را فراهم می‌کند و دقت آن نیز منعکس کننده ویژگی‌های منطقه زهکشی شده است. طراحی یک شبکه بارندگی، به عوامل مختلف از جمله توپوگرافی، الگوی بارش، مشخصات حوضه آبریز، ترافیک، کار و بودجه بستگی دارد به همین دلیل باید هر حوضه به صورت جداگانه و نه به صورت جمعی در نظر گرفته شود. یک شبکه بارندگی معتبر می‌تواند داده‌های بارش را به صورت فوری و دقیق برای طراحی و قرار دادن سازه‌های هیدرولیکی مانند زهکشی پیشگیری از سیل فراهم نماید. تعداد ایستگاه‌های باران‌سنجی مورد نیاز در یک حوضه آبریز ممکن است به طور قابل توجهی متفاوت از دیگر حوضه‌ها باشد. با پیشنهاد لانگین (۱۹۶۰) تراکم ایستگاه‌های باران‌سنجی به طور مستقیم با توجه به جمعیت منطقه تعیین می‌شود، که در آن ایستگاه‌ها بدون هیچ‌گونه استانداردی، تنها با توجه به تراکم جمعیت و بودجه اختصاص داده شده به هر منطقه ساخته می‌شدند. سازمان جهانی هواشناسی در سال ۱۹۷۰ حداقل تراکم شبکه‌های بارندگی را تحت شرایط مختلف توپوگرافی و هواشناسی به شرح زیر تعیین کرده است:

۱. در مناطق مسطح با شدت متوسط و آب و هوای گرمسیری و مدیترانه‌ای به ازای هر ۶۰۰ تا ۹۰۰ کیلومتر مربع یک ایستگاه.
۲. در مناطق کوهستانی با شدت متوسط و آب و هوای گرمسیری و مدیترانه‌ای به ازای هر ۲۵۰ تا ۱۰۰۰ کیلومتر

مربع یک ایستگاه.

۳. در جزایر کوهستانی با بارش نامنظم به ازای هر ۲۵ کیلومتر مربع یک ایستگاه.

۴. در مناطق قطبی به ازای هر ۱۰ تا ۱۵۰۰ کیلومتر مربع یک ایستگاه.

فنون مورد استفاده در بررسی شبکه بارندگی شامل روش‌های اقتصادی کارآمد (داک استاین، ۱۹۷۴)، روش‌های دسته‌بندی (بورن و گلتر، ۱۹۹۱)، روش‌های کمترین توان‌های دوم (مارکوس و همکاران، ۲۰۰۳)، تحلیل واریانس (شی، ۱۹۸۲؛ رودریگز-روب و میجیا، ۱۹۷۴) کریگیدن (هوگزو همکاران، ۱۹۸۱؛ چنگ و همکاران، ۲۰۰۸)، اطلاع آنتروپی (هارمونچیکل، ۱۹۸۱؛ کرستاناویک و سینک، ۱۹۹۲؛ الزهرانی و حسین، ۱۹۹۸؛ آزکل و همکاران، ۲۰۰۰) و ترکیب کریگیدن و اطلاع آنتروپی (چن و همکاران، ۲۰۰۸) بوده است. به طور خاص، رویکرد اطلاع آنتروپی به طور گسترده‌ای از سال ۱۹۷۰ برای طراحی شبکه مجموعه داده‌های بارندگی و ارزیابی عدم قطعیت مورد استفاده قرار گرفته است. بر اساس گزارشات سازمان جهانی هواشناسی ملاک‌های طراحی شبکه‌های باران‌سنجی، بیشتر بر سطح پوشش منطقه توسط ایستگاه‌ها تاکید دارند. شبکه ایستگاه‌های باران‌سنجی موجود که توسط وزارت نیرو احداث شده و مورد بهره‌برداری قرار دارند از این ویژگی برخوردار نیستند. بنابراین لازم است شبکه موجود بر اساس ملاک‌هایی مورد ارزیابی قرار گرفته و به لحاظ تعداد و موقعیت بهینه شود. هدف این مقاله بهینه‌سازی شبکه ایستگاه‌های باران‌سنجی استان خوزستان است به طوری که پیش‌گویی فضایی میزان بارندگی در هر یک از نقاط حوزه براساس کریگیدن از حتمیت قابل قبولی مبتنی بر آنتروپی پیش‌گویی برخوردار باشند. برای این منظور ابتدا فهرستی از موقعیت‌های منطقه که واجد شرایط ایجاد ایستگاه باران‌سنجی هستند، تعیین می‌شوند. سپس با استفاده از کریگیدن در موقعیت‌هایی که پتانسیل احداث ایستگاه باران‌سنجی دارند پیش‌گویی می‌شود. آنگاه بر اساس آنتروپی، ایستگاه‌های غیر ضرور حذف و موقعیت‌های کاندید اضافه کردن ایستگاه به شبکه تعیین و اولویت‌بندی می‌شوند.

## پیش‌گویی فضایی داده‌ها

داده‌های فضایی، مشاهداتی هستند که وابستگی آنها ناشی از موقعیت مکانی آن است و این وابستگی نوعاً تابعی از فاصله قرار گرفتن مشاهدات از یکدیگر در ناحیه یا فضای مورد مطالعه است. به دلیل وابستگی داده‌های میزان بارندگی در موقعیت‌های مختلف، قطعاً این‌گونه داده‌ها مصداقی از داده‌های فضایی هستند و برای تحلیل آنها باید از روش‌های آمار فضایی استفاده شود. کریگیدن<sup>۲</sup> روشی برای پیش‌گویی فضایی در موقعیت‌های فاقد مشاهده ایستگاه بر اساس مشاهدات موجود است. به طور معمول برای مدل‌بندی داده‌ها یک میدان تصادفی مانند  $s \in Z(s)$

<sup>۲</sup> Kriging



$D \subseteq R^d$  در نظر گرفته می‌شود، که در آن  $d \geq 1$  و مجموعه اندیس‌گذار  $D$  زی مجموعه‌ای از فضای اقلیدسی  $R^d$  است. با فرض مانایی میدان تصادفی، معلوم بودن ساختار همبستگی فضایی از طریق تغییرنگار  $\gamma(h) = \text{Var}[Z(s+h) - Z(s)]$  و در نظر گرفتن تابع زیان درجه دوم، پیش‌گویی فضایی بهینه از مینیم کردن میانگین توان دوم خطای پیش‌گو به دست آورده می‌شود. برای پیش‌گویی میدان تصادفی در موقعیت  $s$  با استفاده از داده‌های مشاهده شده  $Z = (Z(s_1), \dots, Z(s_n))$  در موقعیت‌های نمونه‌گیری  $s_1, \dots, s_n$  در ناحیه تحت مطالعه  $D$ ، بهترین پیش‌گویی خطی نااریب  $Z(s)$  تحت عنوان کریگیدن عادی به صورت

$$\hat{Z}(s) = \sum_{i=1}^n \lambda_i Z(s_i) = \lambda' Z$$

تعیین می‌شود، که در آن بردار ضرایب  $\lambda$  با مینیم کردن میانگین توان‌های دوم خطای پیش‌گو، یعنی  $E(Z(s) - \hat{Z}(s))^2$  برآورد می‌شود. برآورد ضرایب این پیش‌گو به صورت

$$\hat{\lambda}' = (\gamma + \frac{1 - \frac{1}{\Gamma^{-1}} \gamma}{\frac{1}{\Gamma^{-1}} \frac{1}{\Gamma^{-1}}}) \Gamma^{-1}$$

است، که در آن  $\gamma = (\gamma(s - s_1), \dots, \gamma(s - s_n))'$  و  $\Gamma_{n \times n} = (\gamma(s_i - s_j))_{n \times n}$ . واریانس این پیش‌گو، که بیان‌گر میزان دقت آن است، به صورت

$$\hat{\sigma}_K^2(s) = \gamma' \Gamma^{-1} \gamma - \frac{(1 - \frac{1}{\Gamma^{-1}} \gamma)^2}{\frac{1}{\Gamma^{-1}} \frac{1}{\Gamma^{-1}}}$$

محاسبه می‌شود. در روش کریگیدن برای پیش‌گویی در یک موقعیت مشخص، به مشاهدات واقع در موقعیت‌های نزدیک وزن بیشتر و به مشاهدات دورتر وزن کمتری اختصاص داده می‌شود. وزن‌ها به گونه‌ای انتخاب می‌شوند که میانگین توان دوم خطای پیش‌گو مینیمم شود. اگر متغیرهای دیگری که با میدان تصادفی ارتباط دارند در اختیار باشند، به جای کریگیدن از هم کریگیدن<sup>۳</sup> برای پیش‌گویی دقیق‌تر استفاده می‌شود. فرض کنید در هر یک از  $n$  موقعیت  $s_1, \dots, s_n$  به جای یک متغیر، بردار  $k$  بعدی  $Z(s_i) = (Z_1(s_i), \dots, Z_k(s_i))$  مشاهده شده باشد. پیش‌گویی هم کریگیدن ترکیبی خطی از تمام مشاهدات  $k$  متغیر به صورت

$$Z(s) = \sum_{i=1}^n \sum_{j=1}^k \lambda_{ji} Z_j(s_i)$$

است، که ضرایب آن از حل  $k(n+1)$  معادله زیر محاسبه می‌شوند:

$$\sum_{i=1}^n \sum_{j=1}^k \lambda_{ij} c_{jj'} - m_{j'} = c_{1j'}(s, s_i), \quad i = 1, \dots, n, \quad j' = 1, \dots, k$$

$$\sum_{i=1}^n \lambda_{1j} = 1, \quad \sum_{i=1}^n \lambda_{ji} = 0, \quad j = 2, \dots, k$$

<sup>۳</sup> Cokriging

واریانس این پیش‌گو نیز به صورت

$$\sigma_k^Y(s.) = C_{11}(s., s.) - \sum_{i=1}^n \sum_{j=1}^k \lambda_{ji} C_{1j}(s., s_i) + m_1$$

$$\sum_{i=1}^n \lambda_{1j} = 1, \quad \sum_{i=1}^n \lambda_{ji} = 0, \quad j = 2, \dots, k$$

محاسبه می‌شوند. (کرسی، ۱۹۹۳).

## طرح نمونه‌گیری فضایی بهینه

یکی از ملاک‌های اندازه‌گیری عدم حتمیت یک متغیر تصادفی آنتروپی است. که میزان آشفتگی، بی‌نظمی، عدم حتمیت و میزان تصادفی بودن متغیر تصادفی را تشریح می‌کند (شانون، ۱۹۴۸). فرض کنید متغیر تصادفی  $X$  با تکیه‌گاه  $\chi$  دارای تابع جرم احتمال  $f(x) = P(X = x)$  باشد. متوسط اندازه تعجب از اطلاع در مورد وقوع مقادیر متغیر تصادفی  $X$  به صورت

$$H(X) = - \sum_{x \in \chi} \log(f(x)) f(x) = E \{-\log(f(x))\}$$

تعریف می‌شود که آنتروپی  $X$  نام دارد.

طرح‌های نمونه‌گیری که نسبت به تغییرات واحدهای نمونه‌گیری حساسیت کمتری دارند طرح‌های نمونه‌گیری استوار نامیده می‌شوند. بنابراین طرح نمونه‌گیری با میزان تصادفی بودن بالا در انتخاب واحدهای نمونه‌گیری، طرح نمونه‌گیری استوار است. آنتروپی طرح نمونه‌گیری اندازه‌های از میزان تصادفی بودن و گستردگی آن را نشان می‌دهد. هنگامی که طرح نمونه‌گیری آنتروپی بالایی دارد، عدم حتمیت زیادی در نمونه‌انتخابی وجود دارد. فرض کنید  $Q_n$ ، تکیه‌گاه یک طرح نمونه‌گیری، یعنی مجموعه تمام نمونه‌های ممکن با حجم ثابت  $n$  و احتمال انتخاب مثبت باشد. چنانچه احتمال به دست آمدن نمونه  $Z = (Z(s_1), \dots, Z(s_n)) \in Q_n$  با  $p(Z)$  نشان داده شود، به طوری که  $\sum_{Z \in Q_n} p(Z) = 1$ ، آنگاه آنتروپی طرح نمونه‌گیری به صورت

$$H(Z) = - \sum_{Z \in Q} p(Z) \log p(Z) = -E_P[\log p(Z)]$$

تعریف می‌شود (هاجک، ۱۹۸۱) و هر چه مقدار آن بیشتر باشد طرح نمونه‌گیری از میزان تصادفی بودن بیشتری برخوردار است.

فرض کنید برای میدان تصادفی  $\{Z(s); s \in D\}$  در  $n$  موقعیت  $S_n = \{s_1, \dots, s_n\}$  تحقق‌هایی به صورت  $Z_1 = (Z(s_1), \dots, Z(s_n))$  در اختیار باشد و بخواهیم در  $m$  موقعیت جدید  $S_m = \{s'_1, \dots, s'_m\}$  نمونه  $Z_2 = (Z(s'_1), \dots, Z(s'_m))$  را به گونه‌ای اختیار کنیم که پیش‌گویی میدان تصادفی بر اساس مشاهدات در دو مجموعه موقعیت‌های موجود و جدید بهینه باشد. موقعیت‌های جدید باید به گونه‌ای انتخاب شوند که عدم حتمیت یا آنتروپی پیش‌گویی در آنها

ماکسیمم شود، زیرا هر چه میزان عدم حتمیت در این موقعیت‌ها بیشتر باشد، پیش‌گویی در موقعیت‌های باقیمانده از حتمیت بیشتری برخوردار خواهند شد. با قرار دادن  $Z = (Z_1, Z_2)$ ، عدم حتمیت یا آنترپی  $Z$  عبارتست از

$$\begin{aligned} H(Z) &= H(Z_1, Z_2) \\ &= E\{-\log[f(Z_1, Z_2)]\} \\ &= E\{-\log[f(Z_2|Z_1)f(Z_1)]\} \\ &= E\{-\log[f(Z_2|Z_1) - \log f(Z_1)]\} \\ &= H(Z_2|Z_1) + H(Z_1) \end{aligned}$$

به دلیل ثابت بودن موقعیت‌های موجود  $S_n$ ،  $H(Z_1)$  ثابت است و کافی است آنترپی یعنی  $H(Z_2|Z_1)$  ماکسیمم شود. در واقع با در اختیار داشتن مشاهدات در موقعیت‌های جاری  $S_n$ ، باید موقعیت‌های جدید  $S_m$  طوری انتخاب شوند که آنترپی  $Z_2$  در موقعیت‌های جدید به شرط موقعیت‌های جاری ماکسیمم شود. در این صورت عدم حتمیت پیش‌گویی در موقعیت‌های جدید بیشترین است، در نتیجه پیش‌گویی در این موقعیت‌ها از دقت کمتری نسبت به موقعیت‌های دیگر برخوردار است. لذا سعی می‌شود در این موقعیت‌ها نمونه‌گیری شود تا نیازی به پیش‌گویی در آنها نباشد. بنابراین طرح بهینه شامل موقعیت‌هایی می‌شود که پیش‌گویی در آنها از دقت بالایی برخوردار نیست و نمونه‌گیری در آنها موجب افزایش دقت پیش‌گویی در سایر موقعیت‌ها می‌شوند.

## منطقه مورد مطالعه و داده‌های بارندگی

حوضه‌های آبریز مشرف به دشت خوزستان، به لحاظ حجم آب‌دهی و طرح‌های عظیم توسعه و بهره‌برداری از منابع جایگاه ویژه‌ای در سطح کشور دارند. نظر به این که احداث شبکه‌های عظیم آبیاری در سال‌های اخیر منجر به بروز مشکلاتی در تامین آب شرب و خشک شدن نخل‌های منطقه شده است، توجه خاص به مدیریت منابع آب استان خوزستان را ضروری ساخته است. در مطالعات بررسی شبکه باران‌سنجی استان خوزستان دو سری موقعیت ایستگاه انتخاب شده‌اند که عبارتند از:

- شبکه ایستگاه‌های موجود وزارت نیرو در محدوده طرح شامل ۱۳۷ ایستگاه که استان خوزستان و یک نوار ۵۰ کیلومتری خارج از آن (در مرز زمینی کشور) را پوشش می‌دهد. مشخصات و موقعیت نمونه‌ای از این ایستگاه‌ها در جدول ۱ قابل ملاحظه است.
- علاوه بر ایستگاه‌های موجود با هدف ایجاد پوشش کامل تعداد ۸۸ موقعیت براساس استانداردهای ارائه شده توسط سازمان هواشناسی جهانی و وزارت نیرو در گستره محدوده مطالعاتی انتخاب شده است. جدول ۲ نمونه‌ای از این موقعیت‌ها را نشان می‌دهد.

جدول ۱: داده‌های ایستگاه‌های وزارت نیرو

ردیف	طول	عرض	ارتفاع	باران	موقعیت
۱	۴۲۵۷۹۵/۶	۳۵۱۱۱۴۵	۱۳۵۰	۷۳۲/۶	$s_1$
۲	۴۸۰۰۰۹/۵	۳۵۰۱۲۰۰	۱۲۹۵	۶۰۰/۷	$s_2$
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
۸۸	۳۷۶۵۱۰/۹	۳۳۴۵۳۳۸	۱۴	۲۲۰/۸	$s_{88}$

جدول ۲: داده‌های ۸۸ موقعیت انتخابی

ردیف	طول	عرض	ارتفاع	باران	موقعیت
۱	۲۵۴۸۹۴/۰	۳۳۲۳۲۶۲	۳	.	$P_1$
۲	۱۸۷۸۵۷/۰	۳۴۶۲۷۲۳	۲۹	.	$P_2$
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
۸۸	۳۹۳۶۲۴/۰	۳۵۵۹۴۳۶	۱۶۶۸	.	$P_{88}$

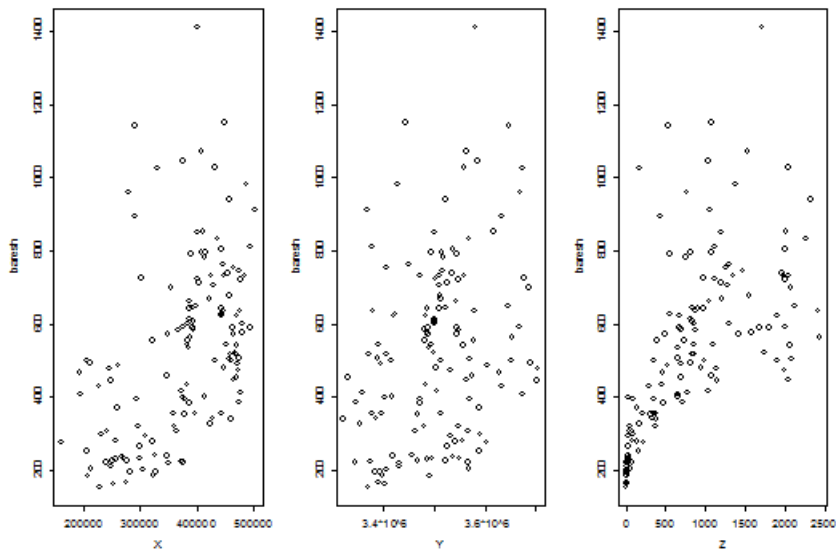
## تحلیل اکتشافی داده‌ها

در تحلیل داده‌های فضایی، مفروضات بنیادی مانند گاوسی بودن، مانابی<sup>۴</sup> در میانگین، هم‌سانگردی، یعنی یکسان بودن ساختار همبستگی در جهات مختلف و همچنین وجود داده‌های پرت، مورد بررسی قرار می‌گیرند. آزمون کولموگروف اسمیرنوف با  $p$ -مقدار ۰/۲۸۹ بیان‌گر نرمال بودن توزیع مشاهدات است. نمودار پراکنش مشاهدات در مقابل هر یک از مختصات طول و عرض و ارتفاع جغرافیایی در شکل ۱ بیان‌گر وجود روندی کاهشی از شمال به جنوب، روندی افزایشی از غرب به شرق و روندی افزایشی از سطح به ارتفاع است.

برای تعیین ساختار همبستگی فضایی داده‌ها، لازم است داده‌ها روند زدوده شوند. مدل رگرسیونی

$$\mu(s) = 1953/104 - 0/0116x - 0/0005y + 3/467 \times 10^{-9}xy + 11/236h^2 \quad (1)$$

<sup>۴</sup> stationary



شکل ۱: نمودار پراکنش متوسط سالانه بارندگی در برابر (الف) طول، (ب) عرض و (ج) ارتفاع جغرافیایی موقعیت ایستگاه‌های باران‌سنجی

با ضریب تعیین  $0.8667$  به روند داده‌ها برازش داده شد، که در آن  $s = (x, y, h)$  موقعیت جغرافیایی هر ایستگاه است. سپس داده‌های روند زدوده از رابطه  $\delta(s) = Z(s) - \mu(s)$  محاسبه شدند.

برای بررسی هم‌سانگردی ساختار همبستگی داده‌ها تغییرنگار تجربی داده‌های روند زدوده شده در چهار جهت صفر،  $45^\circ$ ،  $90^\circ$ ،  $135^\circ$  درجه (شکل ۲) شباهت تقریبی رفتار تغییرنگار در جهت‌های مختلف بیان‌گر هم‌سانگردی برآورد تغییرنگار داده‌ها است.

برآورد استوار تغییرنگار تجربی از رابطه

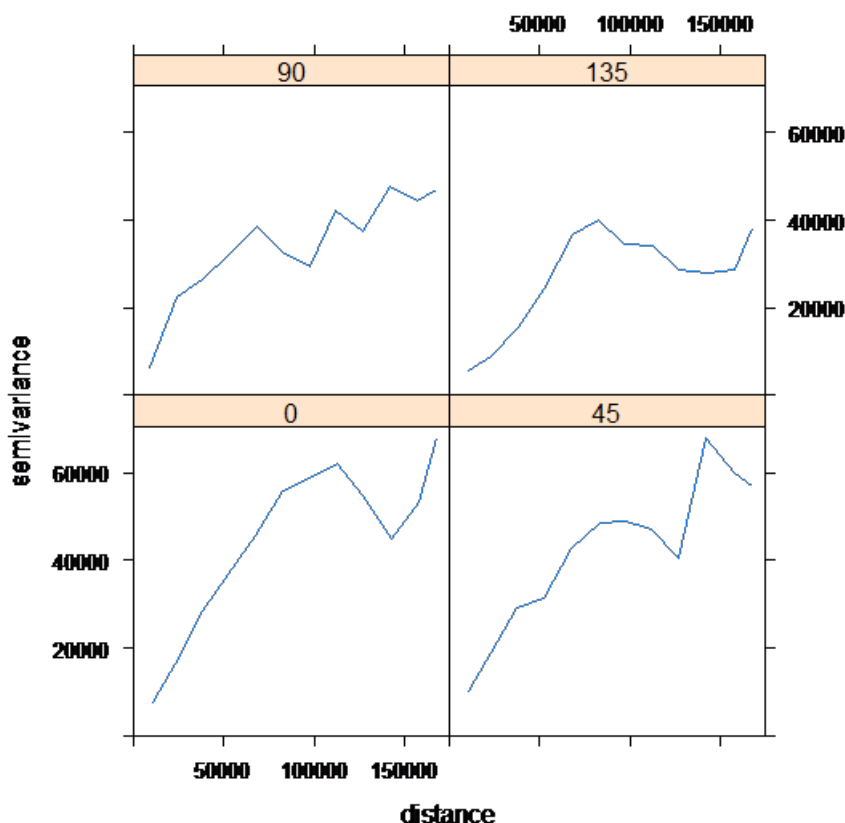
$$\hat{\gamma}(h) = \frac{\left\{ \frac{1}{N_h} \sum_{N_h} |Z(s_i) - Z(s_j)|^{\frac{1}{d}} \right\}^d}{0.457 + \frac{0.494}{N_h}}, \quad h \in R$$

محاسبه شده است، که در آن  $N_h = \{(s_i, s_j); s_i - s_j = h; i, j = 1, \dots, n\}$  مجموعه زوج موقعیت‌هایی است که در فاصله  $h$  از یکدیگر قرار دارند و  $N_h$  تعداد اعضای مجموعه  $N_h$  است. سپس مدل‌های مختلف تغییرنگار معتبر

به داده‌ها برازش داده شد و از بین آنها مدل نیم تغییرنگار کروی

$$\gamma(h) = \begin{cases} c + c \left( \frac{\|h\|}{a} - \frac{1}{\frac{\|h\|}{a}} \right) & 0 < \|h\| < a, h \in R^d, \quad d = 1, 2, 3 \\ c + c & \|h\| > a \end{cases}$$

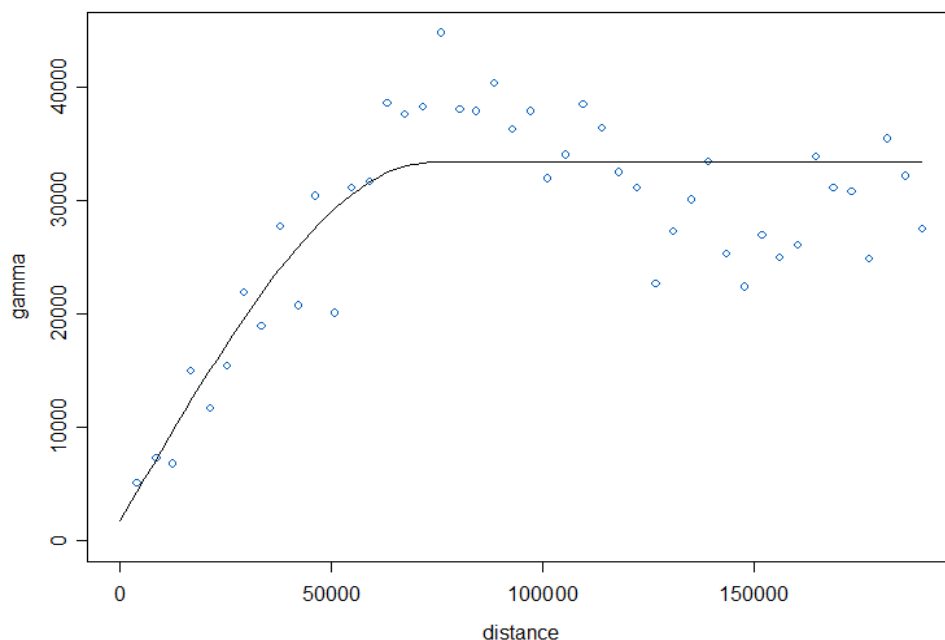
که شامل سه پارامتر با برآوردهای  $177.0/83$ ،  $c = 3165.0/20$ ،  $c = 73715/74$ ،  $a = 73715/74$  است، تعیین شد. این مدل که نمودار آن در شکل ۳ رسم شده است، به دلیل دارا بودن کمترین مجموع توان‌های دوم خطا (RSS) از میان انواع مدل‌های برازنده شده، به عنوان تغییرنگار داده‌ها برگزیده شد.



شکل ۲: نمودار برآورد استوار تغییرنگار داده‌های روند زدوده در مقابل موقعیت‌ها

## تعیین شبکه ایستگاه‌های بهینه

ایستگاه‌هایی که پیش‌گویی در آنها بر اساس سایر مشاهدات از بیشترین دقت یا به عبارت دیگر کمترین آنتروپی برخوردار باشند می‌توانند در الویت حذف قرار گیرند. مجموعه ۲۲۵ ایستگاه متشکل از ۱۳۷ ایستگاه باران‌سنجی و ۸۸ موقعیت پیشنهادی در نظر گرفته شده و آنتروپی پیش‌گویی تمام ایستگاه‌ها بر اساس مشاهدات در ۲۲۴ ایستگاه باقیمانده محاسبه شده و اولین ایستگاه دارای کمترین آنتروپی از مجموعه ایستگاه‌ها حذف می‌شود در مرحله بعد پس از حذف ایستگاه دارای کمترین آنتروپی، مجدداً آنتروپی پیش‌گویی در هر یک از ۲۲۴ موقعیت بر اساس ۲۲۳ موقعیت باقیمانده محاسبه و مقایسه می‌شوند. سپس ایستگاه دوم با کمترین آنتروپی از مجموعه حاوی ۲۲۴ ایستگاه حذف می‌شود. این فرایند برای شناسایی اولویت ایستگاه‌های بعدی برای حذف تکرار می‌شود. شکل ۴ نمودار مجموع آنتروپی پیش‌گویی در ایستگاه‌های باقیمانده پس از حذف متوالی ایستگاه‌ها به روش گام به گام در مقابل تعداد ایستگاه‌های حذف شده را نشان می‌دهد. با انتخاب یک سرحد برای مجموع آنتروپی پیش‌گویی‌ها

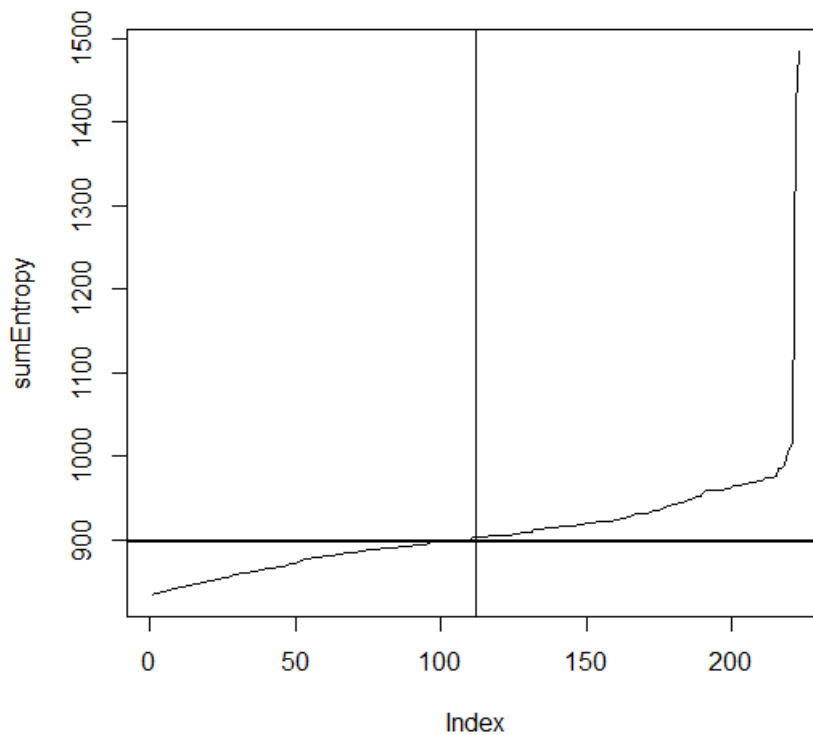


شکل ۳: برآورد تجربی تغییرنگار و مدل کروی برازانده شده به داده‌های روند زدوده

می‌توان با استفاده از این نمودار در خصوص تعداد ایستگاه‌های قابل حذف تصمیم گرفت. به عنوان مثال اگر سرحد ۹۰۰ اختیار شود، باید ۱۱۰ ایستگاه از ۲۲۵ ایستگاه حذف شوند. در این صورت ۱۱۵ ایستگاه خواهیم داشت که برای آن لازم است ۶۵ ایستگاه وزارت نیرو حذف شوند و ۴۵ ایستگاه از بین موقعیت‌های پیشنهادی ایجاد شوند، که مشخصات آنها در جدول ۲ ارائه شده‌اند.

## بحث و نتیجه‌گیری

داده‌های بارندگی منبع بنیادی از اطلاعات برای برنامه‌ریزی سازه‌های هیدرولیکی و سیستم‌های پیشگیری از سیل هستند. چگالی باران‌سنجی مهم‌ترین عامل در برآورد عدم قطعیت است. با این حال، ساخت و استقرار شبکه ایستگاه‌های باران‌سنجی معمولاً مستلزم صرف مقدار زیادی نیروی انسانی و منابع است. کریگیدن روشی برای پیش‌گویی داده‌های بارندگی در یک ایستگاه پیشنهادی است. آنتروپی برای ارزیابی اطلاعات فضایی هر یک از ایستگاه‌های پیشنهادی اعمال می‌شود. این داده‌ها، به نوبه خود، می‌توانند برای بررسی اهمیت هر یک از ایستگاه‌ها مورد استفاده قرار گیرند. برای شبکه ایستگاه‌های بارش استان خوزستان یک طرح نمونه‌گیری فضایی بهینه با استفاده از رهیافت آنتروپی به دست آورده شد که طرح جدید نسبت به طرح نمونه موجود پیش‌گویی‌های دقیق‌تری را مسلماً با هزینه و امکانات کمتر فراهم می‌سازد.



شکل ۴: نمودار مقادیر مرتب شده آنتروپی پیش‌گویی در ۲۲۵ ایستگاه

## مراجع

- [1] Al-Zahrani, M. and Husain, T. (1998), An Algorithm for Designing a Precipitation Network in the South-Western Region of Saudi Arabia, *Journal of Hydrology*, **205**, 205–216.
- [2] Burn, D. and Goulter I. (1991), An Approach to the Rationalization of Streamflow Data Collection Networks, *Journal of Hydrology*, **122**, 71–91.
- [3] Campbell, S. (1983), Sampling and Analysis of Rain, *American Society of Testing and Materials, Philadelphia*.
- [4] Chen, Y., Chiang, W. and Yeh, H. (2008), Rainfall Network Design Using Kriging and Entropy, *Hydrol Process*, **22**, 340–346.



- [5] Cheng, K., Lin, Y., Kiou, J. (2008), Rain-Gauge Network Evaluation and Augmentation Using Geostatistics, *Hydrological Processes*, **22**, 2554–2564.
- [6] Cressie, N. (1993), *Statistics for Spatial Data*, John Wiley, New York.
- [7] Duckstein, L., Davis, D. and Bogardi, I. (1974), Applications of Decision Theory to Hydraulic Engineering, *Symposium of American Society of Civil Engineers hydraulic division specialty conference*.
- [8] Hackett, O. (1966), National Water Data Program, *Journal of the American Water Works Association*, **58**, 786–792.
- [9] Hajek, j. (1964), Asymptotic Theory of Rejective Sampling with Varying Probabilities from a Finite Population, *Annals of Mathematical Statistics*, **35**, 1491-1523.
- [10] Harmancioglu, N. (1981), Measuring the Information Content of Hydrological Processes by the Entropy Concept, *Issue for the Centennial of Ataturk's Birth*, Turkey, 13–40.
- [11] Hughes, J. and Lettenmaier, D. (1981), *Data Requirements for Kriging Estimation and Network Design*, *Water Resources Research*, **17**, 1641–1650.
- [12] Krstanovic, P. and Singh, V. (1992), Evaluation of Rainfall Network Using Entropy II: application, *Water Resources Management*, **6**, 295–314.
- [13] Langbein, W. (1960), Hydrologic Data Networks and Methods of Extrapolating or Extending Available Hydrologic Data, *Hydrologic Networks and Methods*.
- [14] Markus M., Knapp, V. and Tasker, G. (2003), Entropy and Generalized Least Square Methods in Assessment of the Regional Value of Streamgauges, *Journal of Hydrology*, **283**, 107–121.
- [15] Ozkul, D., Harmancioglu, N. and Singh, V. (2000), Entropy-Based Assessment of Water Quality Monitoring Networks, *Journal of Hydrologic Engineering*, **5**, 90–100.
- [16] Rodriguez, I., Mejia, J. (1974), The Design of Rainfall Networks in Time and Space, *Water Resources Research*, **10**, 713–728.

- [17] Shannon, C. (1948), A Mathematical Theory of Communication, *Bell System Technical Journal*, **27**, 623–656.
- [18] Shih, S. (1982), Rainfall Variation Analysis and Optimization of Gauging Systems, *Water Resources Research*, **18**, 1269–1277.



## مشخص سازی بر اساس اندازه آنروپی

### حوتی، ف. ۱

گروه آمار، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد

### چکیده

یکی از بخش های مهم در آمار و احتمال، شناسایی مدل آماری جامعه توسط داده هایی است که از آن جامعه در اختیار داریم. اغلب توزیع های آماری خواص و ویژگی های منحصر به فرد دارند. هدف از مشخص سازی توزیع ها نیز شناسایی توزیع جامعه به وسیله ویژگی های خاص آن است. این ویژگی ها را می توان اندازه های معرفی شده در مباحث قابلیت اعتماد و نیز اندازه های اطلاع در نظر گرفت. در واقع یکی از مهم ترین جنبه های مسائل طول عمر، پیدا کردن توزیع طول عمری است که بتواند به اندازه کافی رفتار تصادفی طول عمر قطعه مورد نظر را توصیف کند. از طرفی استفاده از نظریه اطلاع در مباحث قابلیت اعتماد، استنباط ناپارامتری و مشخص سازی توزیع های آماری جایگاه ویژه ای به خود اختصاص داده است. بر این اساس، در این مقاله برخی از نتایج مشخص سازی با استفاده از اندازه های اطلاع مانند، آنروپی گذشته و آنروپی باقی مانده تجمعی پویا و ارتباط آن ها با اندازه های قابلیت اعتماد، ارائه خواهد شد.

**کلمات کلیدی:** مشخص سازی، آنروپی گذشته، آنروپی باقی مانده تجمعی، اندازه های قابلیت اعتماد.

## ۱ پیش‌گفتار

در سال‌های اخیر توجه زیادی به اندازه‌گیری میزان عدم حتمیت و آنتروپی متغیر تصادفی و مشخص‌سازی توزیع‌های آماری بر اساس آن شده است. به عنوان مثال در داخل کشور اسدی و ابراهیمی (۲۰۰۰) نتایجی از خواص یکنوایی و مشخص‌سازی توزیع‌های آماری بر حسب آنتروپی باقی‌مانده و تابع نرخ خطر و میانگین طول عمر باقی‌مانده بدست آوردند. اسدی و زهره‌وند (۲۰۰۷) به مشخص‌سازی توزیع‌های آماری بر اساس آنتروپی باقی‌مانده تجمعی پرداختند. فشندی و احمدی (۲۰۱۲) نتایجی از مشخص‌سازی توزیع‌های متقارن را بر پایه آنتروپی رنی بدست آوردند. همچنین براتپور و همکاران (۲۰۰۷) و براتپور و همکاران (۲۰۰۸) به مشخص‌سازی بر اساس آنتروپی آماره‌های مرتب و رکوردها پرداختند. در خارج از کشور در خصوص مشخص‌سازی توزیع‌ها بر اساس اندازه‌های آنتروپی تحقیقات زیادی انجام شده است از جمله بلزنس و همکاران (۲۰۰۴)، ناندا و پااول (۲۰۰۶a, ۲۰۰۶b) و کریشنزو و لنگ باردی (۲۰۰۴).

در بخش دوم آنتروپی گذشته و مشخص‌سازی بر پایه ارتباط بین این آنتروپی، میانگین طول عمر گذشته و تابع نرخ خطر معکوس بیان می‌شود. بخش سوم شامل معرفی آنتروپی گذشته تعمیم یافته و مشخص‌سازی توزیع جامعه بر اساس آن می‌باشد. در بخش چهارم آنتروپی باقی‌مانده تجمعی تعمیم یافته را بیان نموده و با استفاده از ارتباط بین این آنتروپی، میانگین طول عمر باقی‌مانده و تابع نرخ خطر، توزیع نمایی را مشخص خواهیم کرد.

## ۲ مشخص‌سازی بر اساس آنتروپی گذشته

فرض کنید  $X$  یک متغیر تصادفی پیوسته نامنفی با تابع توزیع  $F$  و تابع چگالی  $f$  باشد. کریشنزو و لنگ باردی (۲۰۰۲)، آنتروپی گذشته را برای اندازه‌گیری میزان عدم حتمیت متغیر طول عمر گذشته یعنی  $X^t = [t - X | X < t]$ ، به صورت

$$H^*(X; t) = - \int_0^t \frac{f(x)}{F(t)} \ln \frac{f(x)}{F(t)} dx, \quad (1)$$

تعریف کردند. در این بخش می‌خواهیم به مشخص‌سازی بر اساس ارتباط بین این آنتروپی و میانگین طول عمر گذشته و تابع نرخ خطر معکوس بپردازیم. می‌دانیم تابع نرخ خطر معکوس به صورت  $r^*(t) = \frac{f(t)}{F(t)}$  تعریف می‌شود. در اینصورت رابطه (۱) را می‌توان بر اساس تابع نرخ خطر معکوس به صورت زیر بازنویسی کرد

$$\begin{aligned} H^*(X; t) &= - \int_0^t \frac{f(x)}{F(t)} \left[ \ln \frac{f(x)}{F(x)} + \ln \frac{F(x)}{F(t)} \right] dx \\ &= 1 - \frac{1}{F(t)} \int_0^t f(x) \ln(r^*(x)) dx. \end{aligned} \quad (2)$$

**نکته ۱.۰۲.** لازم به ذکر است که آنتروپی گذشته توزیع متغیر تصادفی را به طور یکتا مشخص نمی‌کند. یعنی ممکن

است توزیع‌های مختلف دارای آنتروپی گذشته یکسان باشند. در این خصوص به مثال بعد که با جزئیات بیشتر در [۱۱] آمده است، توجه نمایید.

**مثال ۱.** فرض کنید  $X$  و  $Y$  متغیرهای تصادفی طول عمر به ترتیب با توابع چگالی زیر باشند

$$f_X(t) = 2t, \quad 0 < t < 1,$$

$$f_Y(t) = \frac{2t}{c^2}, \quad 0 < t < c, \quad c \neq 1.$$

در اینصورت از رابطه (۱) به آسانی مشاهده می‌شود که آنتروپی گذشته برای این متغیرها برابر است با

$$H^*(X; t) = \frac{1}{2} + \ln\left(\frac{t}{2}\right) = H^*(Y; t), \quad \forall t \in (0, \min(c, 1)).$$

بنابراین مشاهده می‌شود که هر دو متغیر  $X$  و  $Y$  برای هر  $t$  در دامنه مشترک دارای آنتروپی گذشته یکسان می‌باشند.

لازم به یادآوری است که میانگین طول عمر گذشته به صورت

$$m^*(t) = E(X^t) = \int_0^t \frac{F(u)}{F(t)} du,$$

تعریف می‌شود.

**قضیه ۱.۲.** فرض کنید  $X$  یک متغیر تصادفی پیوسته با تکیه‌گاه  $-\infty < t < b < \infty$  و میانگین متناهی  $\mu$  باشد.

رابطه  $m^*(t) r^*(t) = 1$  برقرار است، اگر و تنها اگر

$$F(t) = \exp\left(\frac{t-b}{b-\mu}\right), \quad t \in (-\infty, b].$$

□

**برهان.** جزئیات در [۶] آمده است.

**قضیه ۲.۲.** فرض کنید  $X$  یک متغیر تصادفی پیوسته با میانگین متناهی  $\mu$ ، میانگین طول عمر گذشته  $m^*(t)$ ، نرخ

خطر معکوس  $r^*(t)$  و آنتروپی گذشته  $H^*(X; t)$  باشد. در اینصورت هر یک از گزاره‌های زیر در مشخص‌سازی

توزیع

$$F(t) = \exp\left(\frac{t-b}{b-\mu}\right), \quad t \in (-\infty, b], \quad (3)$$

معادل‌اند.

الف)  $r^*(t)m^*(t) = 1$ .

ب)  $H^*(X; t) = 1 + \ln m^*(t)$ .

**برهان.** ابتدا فرض کنید الف) برقرار باشد. در اینصورت با توجه به قضیه ۱.۲، توزیع (۳) مشخص می‌شود.

بنابراین میانگین طول عمر گذشته و آنتروپی گذشته برای این توزیع به صورت

$$m^*(t) = b - \mu,$$

و

$$H^*(X; t) = 1 + \ln(b - \mu)$$

خواهد بود که در اینصورت قسمت ب) برقرار است. برای اثبات عکس رابطه، فرض کنید ب) برقرار باشد. در اینصورت داریم

$$m^{*'}(t) = m^*(t) H^{*'}(X; t). \quad (۴)$$

با جایگذاری مشتق آنتروپی گذشته نسبت به  $t$ ، در (۴) نتیجه می‌شود

$$m^{*'}(t) = m^*(t) r^*(t) (1 - 1 - \ln(m^*(t) r^*(t))), \quad (۵)$$

که در این تساوی از فرض ب) استفاده شده است. با قرار دادن  $c(t) = m^*(t) r^*(t)$  و استفاده از این نکته که

$$r^*(t) = \frac{1 - m^{*'}(t)}{m^*(t)}, \text{ داریم}$$

$$1 - c(t) = c(t) (-\ln c(t)).$$

اکنون از طرفین تساوی فوق نسبت به  $t$  مشتق می‌گیریم

$$c'(t) (-\ln c(t)) = 0,$$

که از این معادله نتیجه می‌شود

$$c'(t) = 0,$$

و یا

$$c(t) = 1,$$

□ که این یعنی  $r^*(t)m^*(t) = 1$  و با توجه به قضیه ۱.۲، توزیع (۳) برقرار است و اثبات تمام می‌شود.

### ۳ مشخص سازی بر اساس آنتروپی گذشته تعمیم یافته

اگر  $X$  یک متغیر تصادفی پیوسته با تکیه‌گاه  $-\infty < a < b < \infty$ ، تابع توزیع  $F$  و تابع چگالی  $f$  باشد، آن‌گاه آنتروپی گذشته تعمیم یافته به صورت

$$H^{*\alpha}(X; t) = \frac{1}{1 - \alpha} \ln \int_a^t \left( \frac{f(x)}{F(t)} \right)^\alpha dx, \quad \alpha > 0, \quad (۶)$$

تعریف می‌شود. عبارت (۶) به آنتروپی گذشته رنی نیز معروف است. می‌توان نشان داد اگر  $\alpha \rightarrow 1$ ، آن‌گاه آنتروپی گذشته تعمیم یافته به آنتروپی گذشته تبدیل می‌شود.

با استفاده از قاعده هوییتال داریم

$$\begin{aligned} \lim_{\alpha \rightarrow 1} H^{*\alpha}(X; t) &= \lim_{\alpha \rightarrow 1} \frac{1}{1-\alpha} \ln \int_a^t \left( \frac{f(x)}{F(t)} \right)^\alpha dx \\ &= - \int_a^t \left( \frac{f(x)}{F(t)} \right) \ln \left( \frac{f(x)}{F(t)} \right) dx \\ &= H^*(X; t). \end{aligned}$$

قضیه بعد به مشخص‌سازی توزیع نمایی با استفاده از آنتروپی گذشته تعمیم یافته می‌پردازد. که برهان آن در [۱۱] آمده است.

قضیه ۱.۳. متغیر تصادفی نامنفی  $X$  دارای توزیع نمایی با میانگین  $\mu$  است اگر و تنها اگر

$$H^{*\alpha}(X; t) = \frac{1}{1-\alpha} \ln \left( \mu (r^*(t))^\alpha \left( \frac{e^{\alpha t/\mu} - 1}{\alpha} \right) \right), \quad t > 0. \quad (7)$$

#### ۴ مشخص‌سازی بر اساس آنتروپی باقی‌مانده تجمعی پویا تعمیم یافته

رائو و همکاران (۲۰۰۴) یک اندازه عدم حتمیت، جایگزین آنتروپی شانون با نام آنتروپی باقی‌مانده تجمعی ( $CRE$ ) معرفی نمودند و نیز نشان دادند که این کمیت مزایای بهتری نسبت به آنتروپی شانون دارد. سپس اسدی و زهره‌وند (۲۰۰۷) ضمن بررسی برخی خواص این آنتروپی، آنتروپی باقی‌مانده تجمعی پویا را معرفی نموده و به مشخص‌سازی بر پایه آن پرداختند. از دیگر منابع موجود در این زمینه می‌توان به کریشنزو و لنگ باردی (۲۰۰۹) اشاره کرد. در این بخش آنتروپی باقی‌مانده تجمعی پویا تعمیم یافته را معرفی خواهیم کرد و سپس به مشخص‌سازی بر اساس ارتباط آن با میانگین طول عمر باقی‌مانده و تابع نرخ خطر می‌پردازیم.

فرض کنید  $X$  یک متغیر تصادفی نامنفی با تابع چگالی  $f$  و تابع توزیع  $F$  باشد. آنتروپی باقی‌مانده تجمعی ( $CRE$ ) متغیر  $X$  به صورت

$$\xi(X) = - \int_0^\infty \bar{F}(x) \ln \bar{F}(x) dx, \quad (8)$$

تعریف می‌شود. اسدی و زهره‌وند (۲۰۰۷) نشان دادند

$$\xi(X) = E(m(X)),$$

که در آن  $m(t)$  میانگین طول عمر باقی‌مانده یعنی  $X_t = [X - t | X > t]$  است که به صورت

$$m(t) = E(X_t) = \int_t^\infty \frac{\bar{F}(u)}{\bar{F}(t)} du,$$

تعریف می‌شود.

آنتروپی باقی‌مانده تجمعی پویا مرتبه  $\alpha$  برای متغیر تصادفی طول عمر  $X$  عبارتست از

$$\xi_\alpha(X; t) = \frac{1}{1-\alpha} \ln \left( \frac{\int_t^\infty \bar{F}^\alpha(x) dx}{\bar{F}^\alpha(t)} \right), \quad \alpha > 0. \quad (9)$$

در قضایای بعد می‌خواهیم به مشخص‌سازی بر اساس رابطه بین آنتروپی باقی‌مانده تجمعی پویا تعمیم یافته و میانگین طول عمر باقی‌مانده و تابع نرخ خطر پردازیم. قبل از بیان این قضایا، یادآوری می‌شود که تابع نرخ خطر به صورت  $r(t) = \frac{f(t)}{F(t)}$  تعریف می‌شود.

**قضیه ۱۰۴.** فرض کنید  $X$  یک متغیر تصادفی نامنفی با تابع قابلیت اعتماد  $\bar{F}$  باشد. اگر رابطه

$$(1 - \alpha) \xi_\alpha(X; t) = \ln k + \ln m(t), \quad (10)$$

برقرار باشد آن‌گاه  $k = \frac{1}{\alpha}$ ، اگر و تنها اگر  $X$  دارای توزیع نمایی باشد.

**برهان.** فرض کنید (۱۰) برقرار باشد. با مشتق گرفتن از طرفین آن نسبت به  $t$  داریم

$$(1 - \alpha) \xi'_\alpha(X; t) = \frac{m'(t)}{m(t)}.$$

با قرار دادن مشتق  $\xi_\alpha(X; t)$  در عبارت فوق نتیجه می‌شود

$$\alpha r(t) - e^{-(1-\alpha)\xi_\alpha(X;t)} = \frac{m'(t)}{m(t)}. \quad (11)$$

از طرفی با توجه به (۱۰) واضح است که

$$e^{-(1-\alpha)\xi_\alpha(X;t)} = \frac{1}{km(t)}. \quad (12)$$

بنابراین با جایگذاری (۱۲) در رابطه (۱۱) عبارت

$$\alpha kr(t)m(t) - 1 = km'(t). \quad (13)$$

حاصل می‌گردد. با استفاده از ارتباط بین تابع نرخ خطر و میانگین طول عمر باقی‌مانده یعنی  $r(t) = \frac{m'(t)+1}{m(t)}$  و جایگذاری آن در (۱۳) خواهیم داشت

$$m'(t) = \frac{1 - k\alpha}{k(\alpha - 1)}.$$

قرار می‌دهیم  $k_1 = \frac{1-k\alpha}{k(\alpha-1)}$ ، در اینصورت داریم

$$m(t) = k_1 t + k_2,$$

که در آن  $k_2$  ثابت انتگرال‌گیری است. می‌دانیم میانگین طول عمر باقی‌مانده از رابطه

$$\bar{F}(t) = \frac{m(\cdot)}{m(t)} \exp\left(-\int_0^t \frac{dx}{m(x)}\right),$$

تابع توزیع را به طور یکتا مشخص می‌کند. بنابراین با در نظر گرفتن  $k_1 = 0$ ، توزیع نمایی نتیجه می‌شود. برای اثبات عکس قضیه اگر فرض کنیم  $X$  دارای توزیع نمایی با میانگین  $\frac{1}{\theta}$  باشد، آن‌گاه میانگین طول عمر باقی‌مانده برابر است با  $\frac{1}{\theta}$  و نیز با محاسبه آنتروپی باقی‌مانده تجمعی پویا تعمیم یافته، رابطه (۹)، نتیجه می‌شود

$$(1 - \alpha) \xi_\alpha(X; t) = -\ln(\alpha) + \ln \theta.$$

□

بنابراین واضح است که  $k = \frac{1}{\alpha}$ .



قضیه بعد به مشخص‌سازی بر اساس ارتباط بین آنتروپی باقی‌مانده تجمعی پویا تعمیم یافته و تابع نرخ خطر می‌پردازد.

**قضیه ۲.۴.** فرض کنید  $X$  یک متغیر تصادفی پیوسته و نامنفی با تابع قابلیت اعتماد  $\bar{F}$  باشد. در اینصورت

$$\xi_{\alpha}(X; t) = k, \quad (14)$$

که در آن  $k$  مقداری ثابت است، برقرار است اگر و تنها اگر  $X$  دارای توزیع نمایی با تابع قابلیت اعتماد زیر باشد.

$$\bar{F}(t) = e^{-\theta t}, \quad \theta > 0, \quad t > 0.$$

**برهان.** ابتدا فرض کنید تساوی (۱۴) برقرار باشد. با استفاده از مشتق  $\xi_{\alpha}(X; t)$  داریم

$$\alpha r(t) - e^{-(1-\alpha)\xi_{\alpha}(X;t)} = 0. \quad (15)$$

با جایگذاری (۹) در رابطه (۱۵) نتیجه می‌شود

$$\alpha f(t) \int_t^{\infty} \bar{F}^{\alpha}(x) dx = \bar{F}^{\alpha+1}(t). \quad (16)$$

اکنون با مشتق گرفتن از طرفین تساوی فوق نسبت به  $t$  خواهیم داشت

$$\alpha \left[ f'(t) \int_t^{\infty} \bar{F}^{\alpha}(x) dx - f(t) \bar{F}^{\alpha}(t) \right] = -(\alpha + 1) f(t) \bar{F}^{\alpha}(t). \quad (17)$$

با جایگذاری (۱۶) در عبارت (۱۷)، تساوی

$$\frac{f'(t) \bar{F}^{\alpha+1}(t)}{f(t)} - \alpha f(t) \bar{F}^{\alpha}(t) = -(\alpha + 1) \bar{F}^{\alpha}(t) f(t),$$

حاصل می‌گردد. بنابراین از رابطه اخیر نتیجه می‌شود

$$\frac{f'(t)}{f(t)} = -\frac{f(t)}{\bar{F}(t)}.$$

پس از ضرب کردن طرفین تساوی فوق در  $\frac{f(t)}{\bar{F}^{\alpha}(t)}$  و نیز از آنجا که  $r'(t) = r(t) + r^2(t)$ ، معادله

$$r'(t) = 0,$$

بدست می‌آید. پس از حل معادله دیفرانسیل فوق داریم

$$r(t) = d,$$

که در آن  $d$  ثابت انتگرال‌گیری است. بنابراین از حل این معادله دیفرانسیل، تابع قابلیت اعتماد به صورت  $\bar{F}(t) = e^{-(\theta t + k_1)}$  حاصل می‌شود طوری که  $k_1$  ثابت انتگرال‌گیری و  $\theta > 0$  مقداری ثابت است. چون  $\bar{F}(t)$  تابع قابلیت اعتماد است، مقدار  $k_1 = 0$  و تکیه‌گاه متغیر برابر  $t > 0$  بدست می‌آید، که این یعنی  $X$  دارای توزیع نمایی است. برای اثبات عکس قضیه کافی است فرض کنید  $X$  دارای توزیع نمایی با میانگین  $\frac{1}{\theta}$  است. بنابراین از رابطه (۹) داریم

$$(1 - \alpha) \xi_{\alpha}(X; t) = -\ln(\alpha \theta),$$

□ که از این رابطه نتیجه می‌شود  $\xi_{\alpha}(X; t)$  مقداری ثابت است و بنابراین اثبات تمام می‌شود.

## مراجع

- [1] Asadi, M. and Ebrahimi, N., Residual entropy and its characterizations in terms of hazard function and mean residual life function, *Statistics and Probability Letters*, (2000), **49**, 263-269.
- [2] Asadi, M. and Zohrevand, Y., On the dynamic cumulative residual entropy, *Journal of Statistical Planning and Inference*, (2007), **137**, 1931-1941.
- [3] Baratpour, S., Ahmadi, J. and Arghami, N. R., Some characterizations based on entropy of order statistics and record values, *Communications in Statistics-Theory and Methods*, (2007), **36**, 47-57.
- [4] Baratpour, S., Ahmadi, J. and Arghami, N. R., Characterizations based on Renyi entropy of order statistics and record values, *Journal of Statistical Planning and Inference*, (2008), **138**, 2544-2551.
- [5] Belzunce, F., Navarro, J., Ruiz, J. M. and Del Aguila, Y., Some results on residual entropy function, *Metrika*, (2004), **59**, 147-161.
- [6] Chandra, N. K. and Roy, D., Some results on reversed hazard rate, *Probability in the Engineering and Informational Sciences*, (2001), **15**, 95-102.
- [7] Di Crescenzo, A. and Longobardi, M., Entropy-based measure of uncertainty in past lifetime distribution, *Journal of Applied Probability*, (2002), **39**, 434-440.
- [8] Di Crescenzo, A. and Longobardi, M., A measure of discrimination between past lifetime distributions, *Statistics and Probability Letters*, (2004), **67**, 173-182.
- [9] Di Crescenzo, A. and Longobardi, M., On cumulative entropies, *Journal of Statistical Planning and Inference*, (2009), **139**, 4072-4087.
- [10] Fashandi, M. and Ahmadi, J., Characterizations of symmetric distributions based on Renyi entropy, *Statistics and Probability Letters*, (2012), **82**, 798-804.
- [11] Kundu, C., Nanda, A. K. and Maiti, S., Some distributional results through past entropy, *Journal of Statistical Planning and Inference*, (2010), **140**, 1280-1291.
- [12] Nanda, A. K. and Paul, P., Some properties of past entropy and their applications, *Metrika*, (2006a), **64**, 47-61.

- [13] Nanda, A. K. and Paul, P., Some results on generalized past entropy, *Journal of Statistical Planning and Inference*, (2006b), **136**, 3659-3674.
- [14] Rao, M., Chen, Y., Vemuri, B. C. and Wang, F., Cumulative residual entropy: a new measure of information, *IEEE Transactions on Information Theory*, (2004), **50**, 1220-1228.



## آنتروپی بیشینه در مدل های صف بندی یک باجه ای با ورودی گروهی

شمس، م. ا و حسامیان، غ.

گروه آمار، دانشکده علوم ریاضی، دانشگاه کاشان

### چکیده

در این مقاله با استفاده از اصل آنتروپی بیشینه، تقریبی برای زمان انتظار در یک مدل صف بندی یک باجه ای با ورودی گروهی که یک سرویس اصلی و  $k-1$  سرویس جانبی دارد پیدا می کنیم و سپس مقدار تقریبی با مقدار دقیق مقایسه می شود.

**کلمات کلیدی:** اندازه آنتروپی تعمیم یافته، اندازه واگرایی تعمیم یافته، ماکسیمم اندازه آنتروپی تعمیم یافته، مینیمم اندازه واگرایی، مدل صف بندی.

### ۱ مقدمه

اصل آنتروپی بیشینه توسط افراد متعددی نظیر بارد [۲] و گیاشو [۵] و به طور گسترده در سیستم های صف با زمان های ورودی و سرویس دلخواه مورد استفاده قرار گرفته است. شور [۱۳] به بررسی مدل های صف بندی  $M/M/\infty$ ،  $M/M/\infty/N$  و  $M/G/1$  پرداخت. افندی و کوواتسس [۴] از اصل آنتروپی بیشینه به منظور تحلیل مدل های صف بندی  $M/G/1$  و  $G/M/1$  بهره بردند. کوواتسس [۸] توزیع احتمال طول مدل های صف بندی  $G/G/C/N$  و  $G/G/1/N$  را بررسی کرد. آریزونا و کوواتسس [۱] یک جواب تقریبی برای مدل های صف بندی  $M/M/R$  پیدا کردند. وانگ و همکاران [۱۴، ۱۵] در مورد مدل های  $M/G/1$  با سرویس دهنده قابل انتقال، جین و جین [۶] در مورد مدل های صف بندی  $G/G/1$  با تعطیلی، ک و لین [۷] در مورد مدل های صف بندی  $M^X/G/1$

<sup>۱</sup>Email: mehdishams@kashanu.ac.ir

با یک تعطیلی و بالاخره مایریا [۱۱] درباره مدل‌های صف‌بندی  $M^X/G_1, G_2/1$  با دو مرحله و با تعطیلی برنولی تحقیقات ارزنده‌ای انجام داده‌اند. اکثر کارهای انجام شده برای سرویس واحد است در صورتی که در زندگی روزمره با صف‌های متعددی رویارو می‌شویم که مشتریان علاوه بر سرویس اصلی نیاز به سرویس‌های جانبی که توسط سرویس‌دهنده فراهم می‌شود دارند. البته در این زمینه تحقیقاتی نیز انجام شده است. به عنوان نمونه کریشنا و لی [۹] سرویس‌های دو مرحله‌ای را بررسی کرده‌اند. سرویس‌دهنده واحد با وردوی پواسن و سرویس اختیاری دوم توسط مدھی [۱۲] معرفی شد. سرویس دو مرحله‌ای برای مدل صف‌بندی  $M/G/1$  که سرعت سرویس‌دهنده وابسته به تعداد مراجعه‌کنندگان است توسط لی و کیم [۱۰] مطرح شد. کودهاری و تادج [۳] مدل  $M^X/G/1$  با تعطیلی برنولی و سرویس دو مرحله‌ای را بررسی کردند.

در این مقاله به بررسی مدل‌های صف‌بندی گروهی یک باجه‌ای  $M^X/G/1$  با سرویس اختیاری چندگانه و سرویس‌دهنده‌های از کار افتاده می‌پردازیم و زمان‌های انتظار اصلی و تقریبی را با هم مقایسه می‌کنیم. همچنین از اصل آنتروپی بیشینه برای یافتن احتمال‌های حالت پایا و متوسط زمان انتظار در یک مدل  $M^X/G/1$  با  $k$  سرویس که یکی اصلی و بقیه فرعی هستند استفاده می‌کنیم. سرویس‌دهنده با احتمال‌های  $r_i$ ،  $i = 1, \dots, k-1$ ، سرویس‌های فرعی را انجام می‌دهد و با احتمال  $1 - r_i$  بی‌کار است. توزیع اولین سرویس دلخواه و توزیع سرویس‌های فرعی نمایی است. طول عمر و مدت تعمیر سرویس‌دهنده سرویس اصلی و فرعی عمومی است. مشتریان با نرخ  $\lambda$  وارد می‌شوند و  $\mu_i$  نرخ سرویس  $i$  امین ( $i = 1, \dots, k$ ) سرویس اختیاری چندگانه است. اگر متغیر تصادفی  $X$  نشان‌دهنده اندازه دسته‌ای مشتریان،  $B(\cdot)$  و  $b(\cdot)$  به ترتیب تابع توزیع تجمعی و تابع چگالی احتمال زمان سرویس اصلی و  $R_l(\cdot)$  و  $g_l(\cdot)$  به ترتیب تابع توزیع تجمعی و تابع چگالی احتمال زمان  $l$  امین ( $l = 1, \dots, k$ ) سرویس فرعی باشد و همچنین  $\alpha_l$  نرخ خرابی سرویس‌دهنده در طول  $l$  امین ( $l = 1, \dots, k$ ) خرابی و  $\gamma_l$  اولین گشتاور  $l$  امین ( $l = 1, \dots, k$ ) زمان تعمیر باشد، قرار می‌دهیم،  $\rho_l = \frac{\lambda EX}{\mu_l}$ ،  $\delta_l = 1 + \alpha_l \gamma_l$ ،  $\beta_l(x) = \frac{b(x)}{1 - B(x)}$  و  $\gamma_l(x) = \frac{g_l(x)}{1 - R_l(x)}$  که در آن  $l = 1, \dots, k$ . گشاوهرهای دوم زمان سرویس اصلی و زمان تعمیر به ترتیب به صورت  $\beta_l^{(2)} = (-1)^2 \tilde{b}^{(2)}(\cdot)$  و  $\gamma_l^{(2)} = (-1)^2 \tilde{g}^{(2)}(\cdot)$  می‌باشند. برای حالت پایا احتمال‌های زیر را تعریف می‌کنیم:

$(P_l(n))P_1(n)$ ،  $n > 0$  مشتری در صف از سرویس اصلی (فرعی  $l$  ام،  $l = 2, \dots, k$ ) محروم می‌شوند و سرویس‌دهنده مشغول است.

$(Q_l(n))Q_1(n)$ ،  $n > 0$  مشتری در صف از سرویس اصلی (فرعی  $l$  ام،  $l = 2, \dots, k$ ) محروم می‌شوند و سرویس‌دهنده در حال تعمیر است.

**قضیه ۱.۱.** اگر  $X(t)$  و  $Y(t)$  به ترتیب متغیرهای تصادفی مربوط به زمان سرویس سپری شده برای سرویس اصلی و زمان تعمیر در زمان  $t$  باشند:

الف) احتمال مشغول بودن سرویس‌دهنده برای سرویس اصلی (فرعی) در دراز مدت برابر است با  $P_{B1} = \sum_{n=1}^{\infty} P_1(n) =$

$$P_{B1} = \sum_{n=1}^{\infty} P_1(n) = \left( \prod_{n=1}^{l-1} r_n \right) \rho_l \quad (l = 2, \dots, k)$$

ب) احتمال تحت تعمیر بودن سرویس دهنده برای سرویس اصلی (فرعی) در دراز مدت برابر  $P_{Rl} = \sum_{n=1}^{\infty} Q_l(n)$  است.  $(l = 2, \dots, k, P_{Rl} = \sum_{n=1}^{\infty} Q_l(n) = (\prod_{n=1}^{l-1} r_n) \rho_l \alpha_l \gamma_l)$

**قضیه ۲.۰۱.** متوسط تعداد مشتریان در سیستم عبارت است از

$$L_s = \rho_1 \delta_1 + r_1 \rho_2 \delta_2 + \sum_{l=2}^k \left( \prod_{n=1}^{l-1} r_n \right) \rho_l \delta_l + E(Q)$$

که در آن

$$E(Q) = \frac{1}{\gamma} \left[ (\lambda E(X))^2 \delta_1^2 \beta_1^{(\gamma)} + \Delta_1 \beta_1 + r_1 \left( 2 \rho_1 \delta_1 \rho_2 \delta_2 + \frac{\Delta_2}{\mu_2} + 2(\rho_2 \delta_2)^2 \right) \right. \\ \left. + \sum_{l=2}^k \left( \prod_{n=1}^{l-1} r_n \right) \left( 2 \rho_l \delta_l \sum_{n=1}^{l-1} \rho_n \delta_n + \frac{\Delta_l}{\mu_l} + 2(\rho_l \delta_l)^2 \right) \right] \\ \times \left[ 1 - \rho_1 \delta_1 - \sum_{l=2}^k \left( \prod_{n=1}^{l-1} r_n \right) \rho_l \delta_l \right]^{-1}$$

$$.i = 1, \dots, k, \Delta_i = \lambda E(X(X-1))(1 + \alpha_i \gamma_i) + (\lambda E(X))^2 \alpha_i \gamma_i \text{ و}$$

## ۲ اصل آنتروپی بیشینه

این اصل برای اطمینان دادن به عدم وجود فرضیات غیر لازم در مورد سیستم است. بر اساس این اصل، تابع آنتروپی بیشینه برای مدل صف بندی  $M^X/G/1$  با سرویس اصلی و فرعی را به صورت زیر تشکیل می دهیم:

$$y = - \sum_{n=1}^{\infty} P_1(n) \ln P_1(n) - \sum_{l=2}^k \sum_{n=1}^{\infty} P_l(n) \ln P_l(n) \\ - \sum_{n=1}^{\infty} Q_1(n) \ln Q_1(n) - \sum_{l=2}^k \sum_{n=1}^{\infty} Q_l(n) \ln Q_l(n).$$

آنتروپی بیشینه توسط بیشینه کردن تساوی بالا تحت محدودیت های زیر به دست می آید:

۱. احتمال این که سرویس دهنده سرویس اصلی (فرعی) را انجام دهد:

$$\sum_{n=1}^{\infty} P_1(n) = \rho_1, \quad \left( \sum_{n=1}^{\infty} P_l(n) = \left( \prod_{n=1}^{l-1} r_n \right) \rho_l, \quad l = 2, \dots, k \right).$$

۲. احتمال این است که سرویس دهنده در طول سرویس اصلی (فرعی) از کار بیفتد:

$$\sum_{n=1}^{\infty} Q_1(n) = \rho_1 \alpha_1 \gamma_1, \quad \left( \sum_{n=1}^{\infty} Q_l(n) = \left( \prod_{n=1}^{l-1} r_n \right) \rho_l \alpha_l \gamma_l, \quad l = 2, \dots, k \right).$$

۳. متوسط تعداد مشتریان در سیستم:

$$\sum_{n=1}^{\infty} n \left[ P_1(n) + \sum_{l=2}^k P_l(n) + Q_1(n) + \sum_{l=2}^k Q_l(n) \right] = L_s$$

که  $L_s$  از قضیه ۲.۱ محاسبه می‌شود.

با ضرب هر محدودیت در  $\theta_i$  ( $i = 1, \dots, 2k+1$ )، تابع لاگرانژ به صورت زیر به دست می‌آید:

$$\begin{aligned} y = & - \sum_{n=1}^{\infty} P_1(n) \ln P_1(n) - \sum_{l=2}^k \sum_{n=1}^{\infty} P_l(n) \ln P_l(n) \\ & - \sum_{n=1}^{\infty} Q_1(n) \ln Q_1(n) - \sum_{l=2}^k \sum_{n=1}^{\infty} Q_l(n) \ln Q_l(n) \\ & - \theta_1 \left[ \sum_{n=1}^{\infty} P_1(n) - \rho_1 \right] - \sum_{l=2}^k \theta_l \left[ \sum_{n=1}^{\infty} P_l(n) - \left( \prod_{n=1}^{l-1} r_n \right) \rho_l \right] \\ & - \theta_{k+1} \left[ \sum_{n=1}^{\infty} Q_1(n) - \rho_1 \alpha_1 \gamma_1 \right] \\ & - \sum_{l=k+2}^{2k} \theta_l \left[ \sum_{n=1}^{\infty} Q_l(n) - \left( \prod_{n=1}^{l-1} r_n \right) \rho_l \alpha_l \gamma_l \right] \\ & - \theta_{2k+1} \left[ \sum_{n=1}^{\infty} n \left\{ P_1(n) + \sum_{l=2}^k P_l(n) + Q_1(n) + \sum_{l=2}^k Q_l(n) \right\} - L_s \right]. \end{aligned}$$

برای حل آن‌تروپی بیشینه برای  $P_j(n)$  و  $Q_j(n)$  ( $j = 1, \dots, k$ ) نیاز به بیشینه کردن تابع بالا داریم. پس از گرفتن مشتق جزئی و قرار دادن آن برابر با صفر داریم:

$$\begin{aligned} \frac{\partial y}{\partial P_j(n)} &= -\ln P_j(n) - 1 - \theta_j - \theta_{2k+1} n = 0, \quad j = 1, \dots, k \\ \frac{\partial y}{\partial Q_j(n)} &= -\ln Q_j(n) - 1 - \theta_{k+j} - \theta_{2k+1} n = 0, \quad j = 1, \dots, k \end{aligned}$$

که حل معادلات بالا منجر به جواب‌های

$$P_j(n) = e^{-(1+\theta_j+\theta_{2k+1}n)}, \quad Q_j(n) = e^{-(1+\theta_{k+j}+\theta_{2k+1}n)} \quad j = 1, \dots, k$$

می‌شود که با جایگذاری آن‌ها در محدودیت‌های (۱) تا (۳) داریم:

$$\begin{aligned} P_1(n) &= \frac{\rho_1 \rho_H (L_s - \rho_H)^{n-1}}{(L_s)^n}, \\ P_l(n) &= \frac{\left( \prod_{n=1}^{l-1} r_n \right) \rho_l \rho_H (L_s - \rho_H)^{n-1}}{(L_s)^n}, \quad l = 2, \dots, k \\ Q_1(n) &= \frac{\rho_1 \alpha_1 \gamma_1 \rho_H (L_s - \rho_H)^{n-1}}{(L_s)^n}, \\ Q_l(n) &= \frac{\left( \prod_{n=1}^{l-1} r_n \right) \rho_l \alpha_l \gamma_l \rho_H (L_s - \rho_H)^{n-1}}{(L_s)^n}, \quad l = 2, \dots, k \end{aligned}$$

که در آن

$$\rho_H = \rho_1 \delta_1 + \sum_{l=2}^k \left( \prod_{n=1}^{l-1} r_n \right) \rho_l \delta_n, \quad e^{-\theta_{\tau_{k+1}}} = \frac{L_s - \rho_H}{L_s}.$$

### ۳ متوسط زمان انتظار در صف

در این بخش فرمول دقیق و تقریبی متوسط زمان انتظار در صف محاسبه می‌شود. مقدار دقیق متوسط زمان انتظار در صف توسط فرمول لیتل به صورت  $W = \frac{E(Q)}{\lambda EX}$  است. برای محاسبه متوسط زمان انتظار یک مشتری فرضی در حالت مشغول بودن یا در حال تعمیر بودن سیستم به این صورت عمل می‌کنم که در هنگام ورود مشتری فرضی،  $n$  متقاضی را در صف و در حال سرویس اصلی و فرعی می‌بینید که این سرویس‌دهنده‌ها در حال سرویس‌دهی یا تعمیر هستند که در این صورت:

۱- در حالت مشغول بودن از این که سرویس‌دهنده‌ها در حال کار هستند، مشتری فرضی تا زمانی که  $n$  مشتری سرویس اصلی و فرعی خود را انجام می‌دهند منتظر می‌ماند و در این حالت متوسط زمان انتظار برای او برای سرویس اصلی  $n\beta_1$  و برای سرویس فرعی  $l$  ام  $(l = 2, \dots, k)$ ، برابر با  $\frac{n}{\mu_l}$  است.

۲- در حالت تعمیر اگر  $E(R_i)$  ( $i = 1, \dots, k$ ) متوسط زمان باقیمانده باشد، داریم:  $E(R_i) = \frac{\gamma_i^2}{2\gamma_i}$ . بنابراین متوسط زمان انتظار برای مشتری خاص هنگامی که سرویس در حال تعمیر است  $\frac{n}{\mu_l} + \frac{\gamma_i^2}{2\gamma_i}$ ،  $i = 1, \dots, k$  می‌باشد. اکنون با استفاده از نتیجه بالا مقدار تقریبی متوسط زمان انتظار در صف به صورت زیر خواهد بود:

$$\begin{aligned} W_q^* &= \sum_{n=1}^{\infty} n\beta_1 P_1(n) + \sum_{l=2}^k \sum_{n=1}^{\infty} \frac{n}{\mu_l} P_l(n) + \sum_{n=1}^{\infty} \left( n\beta_1 + \frac{\gamma_1^2}{\gamma_1} \right) Q_1(n) \\ &+ \sum_{l=2}^k \sum_{n=1}^{\infty} \left( \frac{n}{\mu_l} + \frac{\gamma_l^2}{\gamma_l} \right) Q_l(n) = \frac{L_s}{\rho_H} \left[ \beta_1 \rho_1 \delta_1 + \sum_{l=2}^k \left( \prod_{n=1}^{l-1} r_n \right) \frac{\rho_l \delta_l}{\mu_l} \right] \\ &+ \frac{1}{2} \left[ \alpha_1 \rho_1 \gamma_1^2 + \sum_{l=2}^k \left( \prod_{n=1}^{l-1} r_n \right) \alpha_l \rho_l \gamma_l^2 \right]. \end{aligned}$$

### ۴ تحلیل حساسیت

در این بخش میزان دقت نتایج تقریبی با مقادیر  $\lambda = 0.5$ ،  $r_1 = 0.2$ ،  $r_2 = 0.8$  و  $(\alpha_1, \alpha_2, \alpha_3) = (0.1, 0.5, 0.3)$  با سرویس اصلی و دو سرویس فرعی و اندازه متوسط دسته  $EX = 2$  و  $EX = 3$  برای دو مدل صف‌بندی زیر بررسی می‌شود:

$$\text{مدل } M/M/1: (\mu_1, \mu_2, \mu_3) = (2, 10, 10) \text{ و } (\beta_1, \beta_2, \beta_3) = (3, 2, 2)$$

$$\text{مدل } M/E_2/1: (\mu_1, \mu_2, \mu_3) = (3, 4, 9) \text{ و } (\beta_1, \beta_2, \beta_3) = (1, 2, 1)$$



جدول ۳: مقایسه  $W_q$  و  $W_q^*$  بر اساس نرخ ورود  $\lambda$ 

Models	$\lambda$	$E(X) = 2$			$E(X) = 3$		
		$W_q$	$W_q^*$	% Error	$W_q$	$W_q^*$	% Error
$M^X/M/1$	۰/۲	۰/۱۵	۰/۱۳	۱۷/۸۲	۰/۲۶	۰/۲۳	۱۱/۸۹
	۰/۳	۰/۲۷	۰/۲۳	۱۴/۶۷	۰/۴۸	۰/۴۴	۹/۳۵
	۰/۴	۰/۴۱	۰/۳۶	۱۱/۸۴	۰/۹	۰/۸۳	۸/۲۵
	۰/۵	۰/۶۱	۰/۵۵	۱۰/۱۶	۱/۹۳	۱/۷۸	۷/۷۶
	۰/۶	۰/۹۳	۰/۸۴	۹/۱۱	۸/۵	۷/۸۶	۷/۵۹
$M^X/E_T/1$	۰/۲	۰/۱۵	۰/۱۳	۱۰/۶۷	۰/۲۳	۰/۲۲	۲/۰۵
	۰/۳	۰/۲۴	۰/۲۳	۵/۹۳	۰/۴۴	۰/۴۳	۲/۵۹
	۰/۴	۰/۳۷	۰/۳۵	۶/۷	۰/۸۵	۰/۸	۵/۲۱
	۰/۵	۰/۵۶	۰/۵۳	۵/۰۸	۲	۱/۸۳	۸/۷۳
	۰/۶	۰/۸۷	۰/۸۲	۶/۳۶	۳/۵۸	۲۶/۶۹	۱۲/۷۲

جدول ۴: مقایسه  $W_q$  و  $W_q^*$  بر اساس نرخ سرویس  $\mu$ 

Models	$\mu$	$E(X) = 2$			$E(X) = 3$		
		$W_q$	$W_q^*$	% Error	$W_q$	$W_q^*$	% Error
$M^X/M/1$	۲	۱/۶۱	۱/۵	۷/۲۷	۱/۹۳	۱/۷۸	۷/۷۶
	۲/۵	۰/۶۱	۰/۵۵	۱۰/۱۶	۰/۷۵	۰/۶۸	۹/۴۸
	۰/۳	۰/۳۴	۰/۲۹	۱۲/۸۸	۰/۴۲	۰/۳۷	۱۱/۱۲
	۳/۵	۰/۲۲	۰/۱۸	۱۵/۴۳	۰/۲۷	۰/۲۴	۱۲/۶۹
	۴	۰/۱۵	۰/۱۳	۱۷/۸۲	۰/۲	۰/۱۷	۱۴/۱۸
$M^X/E_T/1$	۲	۰/۵۶	۰/۵۳	۵/۰۸	۲	۱/۸۳	۸/۷۳
	۲/۵	۰/۳۲	۰/۲۹	۶/۹۷	۰/۷۴	۰/۶۸	۷/۶۴
	۳	۰/۲۱	۰/۱۹	۹/۳۹	۰/۴۲	۰/۳۸	۷/۸۹
	۳/۵	۰/۱۵	۰/۱۴	۱۱/۹۱	۰/۲۸	۰/۲۶	۸/۷۶
	۴	۰/۱۲	۰/۱	۱۴/۳۵	۰/۲۱	۰/۱۹	۹/۹

در جدول ۳ مشاهده می‌شود که با افزایش نرخ ورودی (و افزایش اندازه متوسط دسته‌ها)، زمان انتظار دقیق ( $W_q$ ) و زمان انتظار تقریبی ( $W_q^*$ ) هر دو افزایش می‌یابند. جدول ۴ تغییرات  $W_q$  و  $W_q^*$  توسط افزایش نرخ سرویس را نشان می‌دهد. زمان انتظار دقیق و تقریبی هر دو با افزایش نرخ سرویس در همه حالت‌ها کاهش می‌یابند. جدول ۵ نشان‌دهنده تحلیل زمان انتظار توسط تغییر نرخ خرابی سرویس دهنده است که در این جا نیز مشاهده می‌شود با افزایش نرخ خرابی  $W_q$  و  $W_q^*$  در همه حالت‌ها افزایش می‌یابند. در جدول ۶ مشاهده می‌شود که با افزایش نرخ تعمیر سرویس دهنده،  $W_q$  و  $W_q^*$  در همه حالت‌ها کاهش می‌یابند. بنابراین نتیجه گرفته می‌شود که با نرخ ورود و نرخ خرابی کم و افزایش نرخ سرویس و نرخ تعمیر، می‌توان زمان انتظار را کمینه کرد.

جدول ۵: مقایسه  $W_q$  و  $W_q^*$  بر اساس نرخ خرابی  $\alpha$

Models	$\alpha$	$E(X) = 2$			$E(X) = 3$		
		$W_q$	$W_q^*$	% Error	$W_q$	$W_q^*$	% Error
$M^X/M/1$	۰.۰۱	۰.۶۱	۰.۵۵	۱۰.۱۶	۱.۹۳	۱.۷۸	۷.۷۶
	۰.۰۲	۰.۶۲	۰.۵۵	۱۰.۴۱	۱.۹۷	۱.۸۱	۸.۰۴
	۰.۰۳	۰.۶۳	۰.۵۶	۱۰.۶۵	۲.۰۱	۱.۸۴	۸.۳۱
	۰.۰۴	۰.۶۳	۰.۵۶	۱۰.۹	۲.۰۵	۱.۸۸	۸.۵۸
	۰.۰۵	۰.۶۴	۰.۵۷	۱۱.۱۴	۲.۰۹	۱.۹۱	۸.۸۶
$M^X/E_2/1$	۰.۰۱	۰.۵۶	۰.۵۳	۵.۰۸	۲	۱.۸۳	۸.۷۳
	۰.۰۲	۰.۵۸	۰.۵۵	۵.۹۵	۲.۱۶	۱.۹۵	۹.۷۱
	۰.۰۳	۰.۶۱	۰.۵۷	۶.۸	۲.۳۳	۲.۰۹	۱۰.۶۶
	۰.۰۴	۰.۶۴	۰.۵۹	۷.۶۲	۲.۵۲	۲.۲۳	۱۱.۵۹
	۰.۰۵	۰.۶۶	۰.۶۱	۸.۴۲	۲.۷۳	۲.۳۹	۱۲.۵۱

جدول ۶: مقایسه  $W_q$  و  $W_q^*$  بر اساس نرخ تعمیر  $\beta$

Models	$1/\gamma$	$E(X) = 2$			$E(X) = 3$		
		$W_q$	$W_q^*$	% Error	$W_q$	$W_q^*$	% Error
$M^X/M/1$	۱	۰.۶۳	۰.۵۷	۱۰.۵۸	۲.۰۳	۱.۸۷	۸.۲۹
	۲	۰.۶۲	۰.۵۵	۱۰.۲۸	۱.۹۵	۱.۸	۷.۹
	۳	۰.۶۱	۰.۵۵	۱۰.۱۶	۱.۹۳	۱.۷۸	۷.۷۶
	۴	۰.۶۱	۰.۵۵	۱۰.۱۱	۱.۹۲	۱.۷۷	۷.۶۹
	۵	۰.۶۱	۰.۵۵	۱۰.۰۷	۱.۹۲	۱.۷۷	۷.۶۵
$M^X/E_2/1$	۱	۰.۲۳	۰.۲	۱۱.۶۴	۰.۴۶	۰.۴۲	۱۰.۳
	۲	۰.۲۲	۰.۲	۱۱.۳۹	۰.۴۶	۰.۴۱	۹.۹۲
	۳	۰.۲۲	۰.۲	۱۱.۲۹	۰.۴۵	۰.۴	۹.۷۹
	۴	۰.۲۲	۰.۲	۱۱.۲۴	۰.۴۵	۰.۴	۹.۷۳
	۵	۰.۲۲	۰.۲	۱۱.۲۱	۰.۴۵	۰.۴	۹.۶۹

## مراجع

- [1] Arizono, L., Cui, Y. and Ohta, H. (1991): An analysis of  $M/M/S$  queueing systems based on the maximum entropy principle, J. Oper. Res. Soc., Vol. 42, 69-73.
- [2] Bard, Y. (1980): Estimation of state Probabilities using the maximum entropy principle, IBM J. Res. Develop., Vol. 24, 563-569.
- [3] Choudhury, G. and Tadj, L. (2011): The optimal control of an unreliable server queue with two phases of service and Bernoulli vacation schedule, Math. & Comp. Mod., Vol. 54, No. 1-2, 673-688.
- [4] El-Attendi, M. A. and Kouvatso, D.D. (1983): A maximum entropy analysis of the  $M/G/1$  and  $G/M/1$  queueing systems at equilibrium, Acta Inform., Vol. 19, 339-355.
- [5] Guiasu, S. (1986): Maximum entropy condition in queueing theory, J. Oper. Res. Soc., Vol. 37, 293-301.

- [6] Jain, M. and Jain, A. (2006): Principle of maximum entropy for  $G/G/1$  queue with vacation under N-policy, IJICS, Vol. 9, No. 1, 28-37.
- [7] Ke, J. C. (2007): Batch arrival queues under vacation policies with server breakdowns and startup/closedown times, Appl. Math. Model., Vol. 31, No.7, 1282-1292.
- [8] Kouvatsos, D.D. (1989): A maximum entropy analysis of the  $G/G/1$  queue at equilibrium, J. Oper. Res. Soc., Vol. 39, 183-200.
- [9] Krishna, C.M. and Lee, Y.H. (1990): A study of two-phase service, Oper. Res. Lett., Vol. 9, 91-97.
- [10] Lee, J. and Kim, J. (2006): A work load dependent  $M/G/1$  queue under a two stage service policy, Oper. Res. Lett., Vol. 34, No. 5, 531-538.
- [11] Maurya, V. N, (2013): Maximum entropy Analysis of  $M^X/(G_1, G_2)/1$  retrial queueing model with second phase optional service and Bernoulli vacation schedule. AJOR, Vol. 3, No. 1, 1-12.
- [12] Medhi, J. (2002): A single server Poisson input queue with a second optional channel, Queueing System, Vol. 42, No. 3, 239-242.
- [13] Shore, J. E. (1978): Derivation of equilibrium and time dependent solutions to  $M/M/N$  and  $M/M$  queueing systems using entropy maximization, In proceedings, Nat. Comp. Con., AFIPS, 183-187.
- [14] Wang, K.-H. , Dang, D.-Y. , and Pearn, W.L. (2011) : comparative analysis of a randomized  $N$  policy queue: An improved maximum entropy method. Expert systems with application Vol. 38, No. 8, 9461-9471.
- [15] Wang, K.-H. , Wang, T.-Y. , and Pearn, W.L. (2005) :Maximum entropy analysis to the  $N$  policy  $M/G/1$  queueing system with server breakdowns and general startup times, Appl. Math.& Compt. Vol.165, 45-61.



## نظریه اطلاعات و جایگاه آن در مهندسی و علوم

عابد هدتنی، ق. ۱

گروه برق، دانشکده مهندسی، دانشگاه فردوسی مشهد

### چکیده

ضمن مرور ماهیت عمومی و مبانی نظریه های علمی و مبحث عمومی آنها (ماده یا شی فیزیکی - ماده با ماده های دیگر) و خواص عمومی ماده - عصاره نظریه اطلاعات را با عنوان خبر و اطلاعات یا جان ماده مطرح می کنیم و کلمات یا مفاهیم اصولی نظریه (آنتروپی - اطلاعات متقابل و دیورژانس) را توضیح می دهیم. سپس با نگاه آنالیز تابعی و هم نگاه آماری هندسی - نحوه تغییرات آنها و ماکزیمم - مینیمم شدن آنها را تشریح کرده به وظیفه اصلی نظریه (تعیین قله ها و ممکنات) می رسمیم. آنگاه با مرور اطلاعات دنباله ها و اطلاعات جهت دار - توانایی بالقوه نظریه در مطالعه علیت - حافظه و فیدبک - تاثیر و تاثر متقابل اجزای سیستم ها و انواع متغیرها بررسی می شود. در قسمت بعدی جایگاه امروز نظریه اطلاعات را در رشته های مختلف مهندسی و علوم با توجه به آخرین تحقیقات بین المللی ترسیم می کنیم.

**الف - مهندسی برق:**

مخابرات - سیستم: کانالهای نقطه به نقطه و چند کاربره و شبکه مخابراتی - الفبا گسسته و الفبا پیوسته.  
مخابرات میدان

**ب - گرایش های دیگر برق:**

کامپیوتر - مهندسی پزشکی - بیولوژی - اقتصاد (یادگیری نظریه اطلاعاتی - علیت - تولید ماکزیمم آنتروپی)  
مواد (کریستالو گرافی مدرن).



## ماکسیمم اندازه آنتروپی تعمیم یافته تحت قید پارامتری کردن اندازه واگرایی تعمیم یافته

صانعی طبس، م. <sup>۱</sup> و محتشمی برزادران، غ. ر.  
آمار، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد

### چکیده

در این مقاله ضمن یادآوری اندازه آنتروپی تعمیم یافته و اندازه واگرایی تعمیم یافته، به مسئله ماکسیمم آنتروپی تعمیم یافته تحت قید پارامتری کردن اندازه واگرایی تعمیم یافته و نیز به طور معادل، مسئله مینیمم اندازه واگرایی تعمیم یافته تحت قید پارامتری کردن اندازه آنتروپی تعمیم یافته می پردازیم و نشان می دهیم که این دو مسئله جواب های یکسان خواهند داشت. این نکته را برای برخی از آنتروپی های خاص را مورد بررسی قرار داده مشاهده نمودیم که ضرایب لاگرانژ دو مسئله معکوس همدیگر هستند و بنابراین توابع با ماکسیمم آنتروپی و مینیمم اندازه واگرایی معادل خواهند بود.

**کلمات کلیدی:** آنتروپی شانون، اطلاع متقابل، تابع مفصل، نرمال دو متغیره.

---

<sup>۱</sup>Email: Sanei@um.ac.ir

## ۱ مقدمه

در متون آماری نظریه اطلاع، اندازه‌های آنتروپی زیادی غیر از آنتروپی شانون<sup>۲</sup> تعریف شده‌اند از جمله آنها تلاش‌هایی بود که توسط رنی<sup>۳</sup> انجام شد وی اندازه آنتروپی رنی مرتبه  $\alpha$  را در سال ۱۹۶۱ معرفی نمود. به منظور اهداف کاربردی تر، هارودا و چاروات<sup>۴</sup> اندازه آنتروپی مرتبه  $s$  را در سال ۱۹۶۷ تعریف کردند که خود یک تابع صعودی از آنتروپی رنی می باشد. همچنین شارما و میتال<sup>۵</sup> آنتروپی مرتبه ۱ و درجه  $s$  و نیز آریموتو<sup>۶</sup> آنتروپی نوع  $t$  را معرفی کردند. میندز و همکاران<sup>۷</sup> شکل کلی از اندازه آنتروپی را به صورت تابعی از  $h, \phi$  تعریف کردند که برای توابع مختلف  $h, \phi$  هر یک از اندازه‌های آنتروپی ذکر شده در بالا را نتیجه می دهد.

$$H_{\phi}^h(p(x)) = h\left(\sum_{i=1}^n \phi(p_i(x))\right) \quad (1)$$

که در این رابطه  $\phi: [0, \infty) \rightarrow R$  یک تابع مقعر و  $h: R \rightarrow R$  یک تابع صعودی است. و یا اینکه  $\phi: [0, \infty) \rightarrow R$  یک تابع محدب و  $h: R \rightarrow R$  یک تابع نزولی می باشد. در ادامه فرض خواهیم کرد  $\phi$  یک تابع مقعر و  $h$  یک تابع صعودی است. در جدول زیر برخی از اندازه‌های آنتروپی آورده شده اند.

جدول ۷: مثال هایی از اندازه های مختلف آنتروپی

$\phi(x)$	$h(x)$	اندازه آنتروپی
$-x \log x$	$x$	Shannon
$x^{\alpha}$	$\frac{1}{1-\alpha} \log x$	Renyi
$\frac{1}{1-s}(x^s - x)$	$x$	Harvda and Charvat
$x \log x$	$\frac{e^{(s-1)x} - 1}{1-s}$	Sharma and Mittal
$x^{\frac{1}{t}}$	$\frac{x^t - 1}{t - 1}$	Arimoto
$x^{r \log x}$	$-2^{r-1} x$	Taneja
$(1 + \lambda x) \log(1 + \lambda x)$	$(1 + \frac{1}{\lambda}) \log(1 + \lambda) - \frac{x}{\lambda}$	Ferreri
$x^{r-m+1}$	$\frac{1}{m-r} \log x$	Varmma

اصل ماکسیم آنتروپی جینز در جواب یک طیف گسترده از سیستم‌های احتمالی نهفته بوده است. این اصل مفاهیم

<sup>۲</sup> Shannon

<sup>۳</sup> Renyi

<sup>۴</sup> Harvda and Charvat

<sup>۵</sup> Sharma and Mittal

<sup>۶</sup> Arimoto

<sup>۷</sup> Menendez et. al.

نظریه اطلاع، استنباط آماری، بهینه‌سازی و ... را در هم می‌آمیزد. توزیع دارای ماکسیمم آنتروپی معمولاً به فرم تابع نمایی از محدودیت‌ها بیان می‌شود. و هنگامی که گشتاورهای مراتب مختلف معلوم باشد به ازای دامنه تغییرات متغیر تصادفی، توزیع‌های مشهوری را به عنوان توزیع آنتروپی ماکسیمم نتیجه می‌دهد. آنتروپی رنی ماکسیمم و آنتروپی تی سالیس ماکسیمم نیز توسیع ایده آنتروپی ماکسیمم به کلاس بزرگتری از آنتروپی شانون است. کستا و همکاران،<sup>۸</sup> بشیرو<sup>۹</sup> و یگنات و جانسون<sup>۱۰</sup>، برچر<sup>۱۱</sup> تعبیرها و مشخصه‌هایی بر اساس آنتروپی رنی ماکسیمم و یا آنتروپی تی سالیس ماکسیمم در حالت‌های یک‌متغیره و چندمتغیره ارائه داده‌اند.

فرض کنید دو تابع چگالی احتمال  $f, g$  داشته باشیم هدف ما محاسبه میزان رابطه موجود بین دو توزیع می‌باشد. اگر  $\phi$  یک تابع محدب باشد برای دو تابع چگالی احتمال  $f, g$  اندازه واگرایی تعمیم‌یافته سیزار<sup>۱۲</sup> عبارت است از:

$$D_{\phi}(f||g) = \sum_{i=1}^n g_i \phi\left(\frac{f_i}{g_i}\right) \quad (2)$$

بر حسب اینکه تابع  $\phi$  چگونه تعریف شده باشد اندازه اطلاع سیزار اندازه‌های واگرایی مختلف را نتیجه می‌دهد. برخی از این اندازه‌ها در جدول ۲ خلاصه شده‌اند.

جدول ۸: مثال‌هایی از اندازه‌های مختلف واگرایی

$\phi(x)$	نام اندازه واگرایی
$x \log x$	Kullback-Libeler
$x^{\alpha}$	Renyi divergence
$x^{\alpha}$	Tsallis divergence
$\frac{(1-\sqrt{t})^2}{t}$	Hlinger divergence
$(t-1)^2$	$\chi^2$ divergence

برچر<sup>۲۰۰۸</sup> با ماکسیمم کردن آنتروپی شانون تحت پارامتری کردن اطلاع کولبک لایبر توزیع‌های تی سالیس را استخراج کرد و سپس با اضافه کردن محدودیت میانگین به صورت مسئله، در نهایت به موضوع ماکسیمم آنتروپی رنی تحت محدودیت میانگین رسید و در ادامه نشان داد که توزیع پاراتوی تعمیم‌یافته، ماکسیمم‌کننده آنتروپی رنی تحت محدودیت یادشده می‌باشد. ونیز بیان می‌کند که مسئله مینیمم کردن اطلاع کولبک لایبر تحت پارامتری کردن اندازه آنتروپی شانون نیز به جواب یکسانی منجر می‌شود.

<sup>۸</sup>Costa, et. al

<sup>۹</sup>Bashkirov

<sup>۱۰</sup> Johnson and Vignat

<sup>۱۱</sup>Berecher

<sup>۱۲</sup>Csiszar

## ۲ دست‌آوردهای پژوهش

ادعای اساسی اصل ماکسیم آنتروپی جینز، یافتن نااریب ترین توزیع احتمال در بین همه توزیع‌هایی است که در یک سری از محدودیت‌ها صدق می‌کند. هدف ما کاهش عدم قطعیت با بدست آوردن اطلاعات است. در واقع باید گفت که باید از همه اطلاعاتی که به ما داده شده استفاده کنیم و محتاطانه از آنچه به ما داده نشده است بگذریم. این یک اصل از دانش و معرفت باستان است. به عبارت دیگر عدم قطعیت در مورد آنچه را نمیدانیم ماکسیم می‌کنیم که این درست‌ترین کار برای انجام است.

فرض کنید دو تابع چگالی احتمال  $f, g$  داشته باشیم اینکه توزیع  $f$  را به عنوان توزیع دمی  $g$  تصور کنیم خیلی مفید به نظر می‌رسد. در یک چنین مواردی این توزیع‌ها باهم رابطه نزدیکی پیدا می‌کنند و هدف ما کشف رابطه موجود بین این دو توزیع است. همانگونه که در مقدمه ذکر شد برچر ۲۰۰۸ مسئله ماکسیم آنتروپی شانون تحت قید پارامتری کردن اطلاع کولبک لایبر و بطور معادل مینیم اطلاع کولبک لایبر تحت قید پارامتری کردن آنتروپی شانون را انجام داده و نشان داد که این دو مسئله منجر به جواب‌های یکسان و معادل می‌شوند. در این مقاله سعی بر این داریم که به جای آنتروپی شانون از اندازه آنتروپی تعمیم یافته و نیز به جای اطلاع کولبک لایبر، اندازه واگرایی تعمیم یافته سیزار را به کار گرفته و این موضوع را برای این اندازه‌های تعمیم یافته بررسی کنیم. همچنین کار برچر را به عنوان یک حالت خاص از آن نتیجه می‌گیریم.

بافرض اینکه  $h$  یک تابع صعودی باشد و برای راحتی کار اندازه آنتروپی تعمیم یافته را به صورت زیر در نظر می‌گیریم

$$H_{\phi}(f) = \int \phi(f(x))dx \quad (۳)$$

اولین مسئله عبارت است از:

$$\begin{aligned} \max H_{\phi}(f(x)) &= - \int \phi(f(x))dx \\ \text{s.t.} \end{aligned} \quad (۴)$$

$$D_{\phi}(f(x)||g(x)) = \int g(x)\phi\left(\frac{f(x)}{g(x)}\right)dx = \theta$$

تعریف  $D_{\phi}(f||g)$  مستلزم این است که  $f$  نسبت به  $g$  مطلقا پیوسته باشد. متناظر با این مسئله دیگری وجود دارد مبنی بر اینکه در جستجوی توزیعی با مینیم اندازه واگرایی تعمیم یافته در بین همه توزیع‌هایی است که برای آن‌ها داریم

$$H_{\phi}(f(x)) = \theta'$$

به عبارت دیگر:

$$\begin{aligned} \min D_{\phi}(f(x)||g(x)) &= \int g(x)\phi\left(\frac{f(x)}{g(x)}\right)dx \\ \text{s.t.} \end{aligned} \quad (۵)$$

$$H_{\phi}(f(x)) = - \int \phi(f(x))dx = \theta'$$



در ادامه نشان می‌دهیم که حل هر یک از دو مسئله بالا ما را به جواب‌های یکسان می‌رساند.

## ۱.۲ حل اولین مسئله

برای حل ابتدا معادله لاگرانژ را تشکیل داده و نسبت به تابع  $f$  از طرفین این معادله مشتق گرفته و بعد از یک سری عملیات جبری خواهیم داشت:

$$\begin{aligned}
 & - \int \phi(f(x))dx - \beta \left( \int g(x) \phi\left(\frac{f(x)}{g(x)}\right) dx - \theta \right) \\
 & - \lambda \left( \int f(x) dx - 1 \right) = 0 \\
 \Rightarrow & -\phi'(f(x)) - \beta \phi'\left(\frac{f(x)}{g(x)}\right) - \lambda = 0 \\
 \Rightarrow & -\phi'(f(x)) + \int f(x) \phi'(f(x)) \\
 & = \beta \left[ - \int f(x) \phi'\left(\frac{f(x)}{g(x)}\right) + \phi'\left(\frac{f(x)}{g(x)}\right) \right] \\
 \Rightarrow & \beta = \frac{-\phi'(f(x)) + \int f(x) \phi'(f(x))}{- \int f(x) \phi'\left(\frac{f(x)}{g(x)}\right) + \phi'\left(\frac{f(x)}{g(x)}\right)} \quad (۶)
 \end{aligned}$$

بدین ترتیب ضرایب لاگرانژ  $\lambda, \beta$  به دست آمده اند با جایگذاری این ضرایب، تابع  $\phi$  و نیز به کمک روش‌های عددی تابع ماکسیم‌کننده معادله بالا را به دست می‌آوریم. حال به سراغ حل دومین مسئله قیدشده در بالا می‌رویم.

## ۲.۲ حل دومین مسئله

در اینجا نیز اولین مرحله تشکیل معادله لاگرانژ و سپس مشتق‌گیری از طرفین و بدست آوردن ضرایب لاگرانژ می‌باشد.

$$\begin{aligned}
 & \int g(x) \phi\left(\frac{f(x)}{g(x)}\right) dx + \beta' \left( - \int \phi(f(x)) dx - \theta' \right) \\
 & - \lambda \left( \int f(x) dx - 1 \right) = 0 \\
 \Rightarrow & \phi'\left(\frac{f(x)}{g(x)}\right) - \beta' \phi'(f(x)) - \lambda = 0 \\
 \Rightarrow & \phi'\left(\frac{f(x)}{g(x)}\right) - \int f(x) \phi'\left(\frac{f(x)}{g(x)}\right) \\
 & = \beta' \left( \phi'(f(x)) - \int f(x) \phi'(f(x)) \right) \\
 \Rightarrow & \beta' = \frac{\phi'\left(\frac{f(x)}{g(x)}\right) - \int f(x) \phi'\left(\frac{f(x)}{g(x)}\right)}{\phi'(f(x)) - \int f(x) \phi'(f(x))} \quad (۷)
 \end{aligned}$$

با کمی دقت و مقایسه می یابیم که  $\beta' = -\frac{1}{\beta}$  و بر این اساس نتیجه می گیریم که تابع دارای ماکسیمم اندازه آنتروپی تعمیم یافته با تابع دارای مینیمم اندازه واگرایی تعمیم یافته معادل خواهند بود. در ادامه با قراردادن برخی از توابع  $\phi(x)$  از جداول ۱ و ۲ در دو مسئله بالا بیشتر به این موضوع می پردازیم.

**مثال ۱.** اگر فرادهم  $\phi(x) = x \log x$  یعنی اندازه آنتروپی شانون و اندازه اطلاع کولبک لایبر را داریم. در این صورت  $\phi'(x) = 1 + \log x$  و مقدار ضریب لاگرانژ عبارت است از:

$$\beta = \frac{\int f(x) \log f(x) - \log f(x)}{\log \frac{f(x)}{g(x)} - \int f(x) \log \frac{f(x)}{g(x)}}$$

$$\beta' = \frac{\log \frac{f(x)}{g(x)} - \int f(x) \log \frac{f(x)}{g(x)}}{-\int f(x) \log f(x) + \log f(x)}$$

همانگونه که قبلا ذکر شد برچر هم به این نتیجه مشابه رسیده بود که حل این دو مسئله برای آنتروپی شانون و اطلاع کولبک لایبر منجر به نتایج یکسان می شود.

**مثال ۲.** فرض کنید  $\phi(x) = x^\alpha$  یعنی اندازه آنتروپی رنی و اندازه واگرایی رنی را داریم در این حالت  $\phi'(x) = \alpha x^{\alpha-1}$  همچنین مقادیر ضرایب لاگرانژ برای این مسئله عبارت است از:

$$\beta = \frac{\int f(x)^\alpha - f(x)^{\alpha-1}}{\left(\frac{f(x)}{g(x)}\right)^{\alpha-1} - \int \left(\frac{f(x)}{g(x)}\right)^{\alpha-1} f(x)}$$

$$\beta' = \frac{\left(\frac{f(x)}{g(x)}\right)^{\alpha-1} - \int \left(\frac{f(x)}{g(x)}\right)^{\alpha-1} f(x)}{-\int f(x)^\alpha + f(x)^{\alpha-1}}$$

ملاحظه می کنید که برای آنتروپی رنی و اندازه واگرایی رنی نیز موضوع بالا صادق است.

**مثال ۳.** فرض کنید  $\phi(x) = \frac{1}{1-s}(x^s - s)$  یعنی آنتروپی تی سالیس (آنتروپی هارودا چاروات) و اندازه واگرایی تی سالیس را داریم در این حالت  $\phi'(x) = \frac{s}{1-s}(x^{s-1} - 1)$  محاسبه مقادیر ضرایب لاگرانژ برای این مسئله ما را به عبارات زیر می رساند:

$$\beta = \frac{f(x)^{s-1} - \int f(x)^s}{\int f(x) \left(\frac{f(x)}{g(x)}\right)^{s-1} - \left(\frac{f(x)}{g(x)}\right)^{s-1}}$$

$$\beta' = \frac{\left(\frac{f(x)}{g(x)}\right)^{s-1} - \int \left(\frac{f(x)}{g(x)}\right)^{s-1} f(x)}{-\int f(x)^s + f(x)^{s-1}}$$

ملاحظه می کنید که برای آنتروپی تی سالیس و اندازه واگرایی تی سالیس نیز حل دو مسئله جواب یکسان به دست می دهد.

نکته قابل توجه در دو مثال (۲.۲) و (۳.۲) این هست که اگر  $s = \alpha$  مقادیر ضرایب لاگرانژ متناظر برای این دو مثال برابر می باشند. البته این نکته را هم باید اضافه کرد که آنتروپی رنی تابعی صعودی از آنتروپی تی سالیس می باشد.

## ۳ نتیجه‌گیری

در ابتدا اندازه آنتروپی تعمیم یافته و اندازه واگرایی تعمیم یافته را یادآوری کردیم. در ادامه به مسئله ماکسیمم آنتروپی تعمیم یافته تحت قید پارامتری کردن اندازه واگرایی تعمیم یافته و نیز به طور معادل مسئله مینیمم اندازه واگرایی تعمیم یافته تحت قید پارامتری کردن اندازه آنتروپی تعمیم یافته پرداختیم و نشان دادیم که این دو مسئله جواب‌های یکسان خواهند داشت. این مسئله را برای بعضی از اندازه‌های آنتروپی خاص مورد بررسی قرار داده‌ایم. مشاهده نمودیم که ضرایب لاگرانژ دو مسئله قرینه معکوس همدیگر هستند و بنابراین توابع با ماکسیمم آنتروپی و مینیمم اندازه واگرایی معادل خواهند بود.

## مراجع

- [1] Menendez, M. L., Pardo, J. A. and Pardo, M. C. (1998), *Estimators based on sample Quantiles using  $(h, \phi)$ -Entropy measures*, Appl, Math, Lett, 11(4), 99-104.
- [2] Bashkirov, A. G. (2004), *Maximum Renyi Entropy Principle for systems with Power-Low Hamiltonians*, Physics A, 93, arXiv:cond-mat/0402404v3 [cond-mat.stat-mech] 5 Mar 2004.
- [3] Bashkirov, A. G. (2006), *Renyi Entropy as a Statistical Entropy for Complex Systems*, Theoretical and Mathematical Physics, 149(2), 1559-1573.
- [4] Bercher, J. F. (2008a), *On some entropy functionals from Renyi information divergence*, Information Sciences, 178, 2489-2506.
- [5] Csiszar, I. (1967). *On Topological properties of  $f$ -divergences*, studia Math. Hungaria, 2, 329-339.
- [6] Costa, J. A., Hero, A. O. and Vignat, C. (2006), *A Geometric Characterization of Maximum Renyi Entropy Distributions*, IEEE International Symposium on Information Theory, ISIT, Lausanne, pp. 1822-1826.
- [7] Jaynes, E. T. (1957). *Information theory and statistical mechanics*, Physical Reviews, 106, 620-630.
- [8] Johnson, O. and Vignat, C. (2007), *Annales de l'institut Henri Poincare (B) Probabilites et Statistiques*, 43, No.3, 339-351.

- [9] Renyi, A. (1961), *On measures of entropy and information*, Proc, Berekeley Symposium, Statist, Probability, 1, 547-561.
- [10] Shannon, C. E. (1948). *A Mathimatical Theory Of Communication*, Bell System Tech. 27, 379-423.
- [11] Tsallis, C. (1988). *Possible generalizations of Boltzmann-Gibbs statistics*, Journal of Statistical Physics, 52, 479-487.



## شاخص تایل بر اساس توزیع لمبدای تعمیم یافته و کاربردی از آن

کدخدا، ا. <sup>۱</sup>، محتشمی برزادران، غ. ر. <sup>۱</sup>، امینی، م. <sup>۱</sup> و محمدی، م. <sup>۲</sup>

<sup>۱</sup> گروه آمار، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد

<sup>۲</sup> گروه آمار، دانشکده علوم پایه، دانشگاه زابل

### چکیده

هدف اصلی این مقاله محاسبه‌ی شاخص تایل برای توزیع لمبدای تعمیم یافته است. در این راستا ابتدا مروری خواهیم داشت بر توزیع لمبدای تعمیم یافته و پس از آن آنتروپی و شاخص تایل را معرفی خواهیم کرد. در انتها شاخص تایل متناظر با توزیع لمبدای تعمیم یافته را برای داده‌های هزینه، درآمد و سود شرکت‌های پذیرفته شده در بازار بورس اوراق بهادار تهران محاسبه می‌نماییم.

**کلمات کلیدی:** توزیع لمبدای تعمیم یافته، نظریه اطلاع، شاخص تایل.

### ۱ مقدمه

یکی از موضوعاتی که جوامع امروزی را با مشکل مواجه کرده، توزیع ناعادلانه درآمد و ثروت در بین اقشار مختلف جامعه و بروز نابرابری اقتصادی در جامعه است. از این رو محاسبه‌ی میزان این نابرابری توجه بسیاری از اقتصاددانان و سیاست‌گذاران را به خود جلب کرده است. با توجه به اهمیت این موضوع، معیارها و شاخص‌های متعددی برای اندازه‌گیری میزان نابرابری پیشنهاد شده‌اند. در سال ۱۹۰۵ ماکس اوتو لورنتز منحنی‌ای را برای نمایش توزیع درآمد پیشنهاد کرد که به منحنی لورنتس مشهور است. معیارهای سنجش نابرابری زیادی بر اساس منحنی لورنتس تعریف شده است که از میان آنها می‌توان به ضریب جینی، ضریب بن فرونی و ضریب زنگا اشاره کرد. با افزایش حساسیت‌ها نسبت به توزیع درآمد و اندازه‌گیری نابرابری شاخص‌ها و معیارهای دیگری نظیر شاخص تایل، شاخص اتکینسون

<sup>۱</sup>Email: ekadkhoda@gmail.com

و شاخص هرفیندال، معرفی گردیدند که از این میان ما بر روی شاخص تایل تمرکز خواهیم کرد. در واقع در این شاخص از مفهوم آنتروپی و نظریه اطلاع برای اندازه گیری میزان نابرابری جامعه استفاده شده است. یکی از روش های محاسبه ی این معیارهای نابرابری بدین ترتیب است که ابتدا یک توزیع درآمدی مناسب به داده ها برازش داده و پارامترهای توزیع را برآورد نماییم سپس معیارهای نابرابری را محاسبه کنیم برای این منظور از توزیع لمبدای تعمیم یافته که توزیعی پرکاربرد در زمینه ی داده های درآمد و دارایی می باشد، استفاده می نماییم. روش دیگر این است که بدون در نظر گرفتن توزیع خاصی برای داده ها به صورت مستقیم آنها را محاسبه نماییم. در اینجا برای داده های سود حاصل از فروش اوراق بهادار کلیه شرکت های پذیرفته شده در بازار بورس تهران با استفاده از هر دو روش ذکر شده مقدار شاخص تایل را محاسبه خواهیم کرد.

## ۲ توزیع لمبدای تعمیم یافته و روش های برآورد پارامتر

توزیع لمبدای تعمیم یافته GLD تعمیمی از توزیع تک پارامتری توکی است که در سال ۱۹۶۰ توسط جان توکی معرفی شده است. اولین روش پارامتری کردن این تعمیم چهار پارامتری که به  $RS - GLD$  مشهور است را در سال ۱۹۷۴ رمبرگ و اشمیزر پیشنهاد کردند و با تابع چندکی اش مشخص می شود که برابر با

$$Q_{RS}(u) = \lambda_1 + \frac{u^{\lambda_3} - (1-u)^{\lambda_4}}{\lambda_2} \quad (1)$$

است و تابع چگالی آن در نقطه ی  $x = F^{-1}(u)$  از رابطه ی  $f_{RS}(x) = \frac{\lambda_2}{\lambda_3 u^{\lambda_3-1} + \lambda_4 (1-u)^{\lambda_4-1}}$  بدست می آید که  $u \in [0, 1]$  (کاریان و دودیکس ۲۰۰۰). رمبرگ و همکاران (۱۹۷۹) ذکر کردند که ترکیبات معلومی از  $\lambda_3$  و  $\lambda_4$  هستند که تابع توزیع فوق در آنها یک توزیع احتمالی معتبر نمی باشد. برای مطالعه دقیقتر خواننده را به کاریان و دودویکنز [۳] (۲۰۰۰) ارجاع می دهیم. برای توسیع تکیه گاه توزیع روی همه ی صفحه ی  $(\lambda_3, \lambda_4)$ ، فریمر و همکاران روش پارامتری کردن FMKL را معرفی کردند که برابر است با

$$Q_{FMKL}(u) = \lambda_1 + \frac{1}{\lambda_2} \left( \frac{u^{\lambda_3} - 1}{\lambda_3} - \frac{(1-u)^{\lambda_4} - 1}{\lambda_4} \right) \quad (2)$$

و روی همه ی صفحه  $(\lambda_3, \lambda_4)$  تعریف می شود و تابع چگالی آن برابر است با:

$$f_{FMKL}(x) = \frac{\lambda_2}{u^{\lambda_3-1} + (1-u)^{\lambda_4-1}}, \quad 0 \leq u \leq 1.$$

**قضیه ۱.۲.** گشتاور مرتبه  $k$ ،  $GLD$  وجود دارد اگر و تنها اگر:

$$\min(\lambda_3, \lambda_4) > -\frac{1}{k}.$$

از طرفی امید ریاضی توزیع لمبدای تعمیم یافته برای هر دو روش پارامتری کردن آن به ترتیب از روابط

$$\mu_{RS} = \lambda_1 + \frac{1}{\lambda_2} \left( \frac{1}{1+\lambda_3} - \frac{1}{1+\lambda_4} \right)$$

و

$$\mu_{FMKL} = \lambda_1 - \frac{1}{\lambda_2 \lambda_3} + \frac{1}{\lambda_2 \lambda_4} + \frac{1}{\lambda_2} \left( \frac{1}{\lambda_3(\lambda_3 + 1)} - \frac{1}{\lambda_4(\lambda_4 + 1)} \right)$$

بدست می‌آیند که با توجه به قضیه ۱.۲ این مقادیر تنها در صورتی وجود دارند که

$$\min(\lambda_3, \lambda_4) > -1.$$

برای مطالعه دقیقتر به لاخانی و موزر [۱] (۲۰۰۰) مراجعه کنید.

همانطور که اشاره شد هدف برازش  $GLD$  به یک مجموعه داده و سپس محاسبه شاخص تایل است. لذا باید در ابتدا پارامترهای توزیع را برآورد نماییم. برای برآورد پارامترهای این توزیع روشهای متفاوتی ارائه شده است که از میان آنها می‌توان به روش گشتاورها، صدکها،  $starship$  و درستنمایی ماکسیمم اشاره کرد. ما در این تحقیق از روش درستنمایی ماکسیمم استفاده می‌کنیم. در این روش که توسط سو [۴] (۲۰۰۷) ارائه شده است، از اعداد شبه تصادفی و روش صدکها برای  $RS - GLD$  و روش گشتاورها برای  $FMKL - GLD$  جهت یافتن مقادیر ورودی استفاده می‌کنیم. سپس این مقادیر ورودی را برای ماکسیمم کردن لگاریتم عددی درستنمایی بکار می‌بندیم تا پارامترهای  $GLD$  مناسب را بیابیم. الگوریتم روش درستنمایی ماکسیمم در ۴ گام خلاصه می‌شود:

گام ۱: محدوده‌ای از مقادیر ورودی برای  $\lambda_3$  و  $\lambda_4$  و همچنین تعداد مقادیر ورودی انتخاب شده را تعیین کنید. در نرم افزار  $R$  از دنباله سوبول برای تولید اعداد شبه تصادفی استفاده می‌شود. به طور پیش فرض تعداد مقادیر پیش فرض ۱۰۰۰۰ در نظر گرفته می‌شود.

گام ۲:  $\lambda_1$  و  $\lambda_2$  را برای هرکدام از  $\lambda_3$  و  $\lambda_4$  محاسبه کنید. مجموعه‌ی مقادیر ورودی که در پارامتری کردن  $GLD$  صدق نکنند یا همه‌ی فضای مجموعه داده‌ها را پوشش ندهند، غیرقابل قبول تلقی کرده و آنها را حذف کنید. در میان مجموعه نقاط شروع باقیمانده مجموعه‌ای را بیابید که کمترین مقدار  $\sqrt{(\hat{\rho}_3 - \rho_3)^2 - (\hat{\rho}_4 - \rho_4)^2}$  و  $\sqrt{(\hat{\alpha}_3 - \alpha_3)^2 - (\hat{\alpha}_4 - \alpha_4)^2}$  را تولید کند. حال این مجموعه مقادیر برای فرآیند بهینه‌سازی مورد استفاده قرار می‌گیرد.

گام ۳: با جایگذاری مقادیر ورودی در فرمول‌های (۱) و (۳) صدکهای  $u_i$  را برای  $RS - GLD$  و  $FMKL - GLD$  بیابید.

گام ۴:  $u_i$  بدست آمده در گام ۳ را در فرمولهای لگاریتم عددی درستنمایی (عبارات (۴) و (۳)) جایگذاری کنید.

$$ML_{RS} = \sum_{i=1}^n \log \left[ \frac{\lambda_2}{\lambda_3 u_i^{\lambda_2 - 1} + \lambda_4 (1 - u_i)^{\lambda_2 - 1}} \right] \quad (۳)$$

$$ML_{FMKL} = \sum_{i=1}^n \log \left[ \frac{\lambda_2}{u_i^{\lambda_2 - 1} + (1 - u_i)^{\lambda_2 - 1}} \right] \quad (۴)$$

گام ۵: مقدار بهینه با استفاده از الگوریتم سیمپلکس نلدر-مید یا هر الگوریتم مینیم سازی عددی دیگر بدست می‌آید.

برای برازش  $GLD$  و یافتن پارامترهای مناسب آن با استفاده از روش درستنمایی ماکسیمم در بسته  $GLDEX$  در نرم افزار  $R$  از تابع  $fun.data.fit.mm$  استفاده می‌کنیم.

### ۳ آنتروپی و نظریه اطلاع

فرض کنید  $p_i, i = 1, 2, \dots, n$  احتمال وقوع یک پیشامد گسسته باشد، در نظریه اطلاع تابع  $H$  مدنظر است که با احتمال وقوع پیشامد رابطه‌ی عکس دارد و میزان اطلاع موجود در یک توزیع احتمال را نشان می‌دهد. این تابع که آنتروپی نام دارد، نخستین بار توسط شانن [۲] ۱۹۴۸ معرفی شد و به صورت زیر تعریف می‌شود:

$$H = - \sum_{i=1}^n p_i \log p_i \quad (5)$$

و در آن  $\sum_{i=1}^n p_i = 1$  است. آنتروپی برای متغیر گسسته همواره نامنفی است و بیشترین مقدار آن  $\log n$  می‌باشد که زمانی رخ میدهد که احتمال وقوع همه‌ی پیشامدها برابر باشد در واقع در این حالت بیشترین عدم حتمیت را داشته خواهیم داشت و عدم حتمیت از این جهت است که نمی‌دانیم کدام پیشامد رخ خواهد داد. برای متغیر تصادفی مطلقاً پیوسته  $X$  با تابع چگالی احتمال  $f(x)$ ، آنتروپی برابر است با

$$H = \int_{-\infty}^{\infty} f(x) \log f(x) dx \quad (6)$$

که میزان بی‌نظمی و اغتشاش تابع چگالی  $f(x)$  را اندازه می‌گیرد، به عبارتی هر چه چگالی به تابع چگالی یکنواخت نزدیک‌تر باشد مقدار آنتروپی نیز کمتر می‌شود و تصمیم‌گیری براساس آن مشکل‌تر خواهد شد.

### ۴ شاخص تایل

در سال ۱۹۶۷ تایل [۵] یک شاخص جدید برای ارزیابی توزیع درآمد معرفی کرد که بر پایه مفهوم آنتروپی در نظریه اطلاع استوار است و به صورت تفاضل میان آنتروپی توزیع درآمد جامعه از آنتروپی درآمد عادلانه تعریف می‌شود. تایل در بحث کاربرد نظریه اطلاع در توزیع درآمد به جای استفاده از احتمال رخداد یک حادثه از سهم درآمد در جامعه استفاده کرد.

فرض کنید  $x_1, x_2, \dots$  و  $x_n$  درآمدهای افراد جامعه را نشان دهد که  $x_i \geq 0$ . در این صورت سهم درآمد فرد  $i$  ام برابر با  $\frac{x_i}{n\bar{x}}$  می‌باشد که  $\bar{x}$  میانگین درآمد کل جامعه را نشان می‌دهد. طبق تعریف شاخص تایل تفاضل میان آنتروپی



سهم درآمد جامعه از آنتروپی درآمد عادلانه یعنی زمانی که سهم درآمد هر فرد  $\frac{1}{n}$  باشد را اندازه می‌گیرد و برابر است با:

$$\begin{aligned} T &= \frac{1}{n} \sum_{i=1}^n \log \frac{1}{n} - \sum_{i=1}^n \frac{x_i}{n\bar{x}} \log \frac{n\bar{x}}{x_i} \\ &= \frac{1}{n} \sum_{i=1}^n \frac{x_i}{\bar{x}} \log \frac{x_i}{\bar{x}} \end{aligned} \quad (7)$$

شاخص تایل بین صفر (حالت برابری کامل توزیع درآمد) و  $\log n$  (حالت نابرابری کامل توزیع درآمد) تغییر می‌کند. اگر برای افراد جامعه توزیع درآمد پیوسته در نظر گرفته شود در این صورت شاخص تایل برابر خواهد بود با:

$$T = \int_0^1 \frac{x}{\mu} \log \frac{x}{\mu} dF(x) = \int_0^1 \frac{Q(u)}{\mu} \log \frac{Q(u)}{\mu} du. \quad (8)$$

که در آن  $F(x)$  تابع توزیع تجمعی با دامنه‌ی زیر مجموعه‌ی اعداد حقیقی نامنفی و میانگین متناهی  $\mu$  و  $Q(u)$  تابع چندکی متغیر تصادفی  $X$  می‌باشد.

## ۵ شاخص تایل بر اساس توزیع لمبدای تعمیم یافته

اگر از توزیع لمبدای تعمیم یافته برای توزیع داده‌های درآمد استفاده کنیم آنگاه با استفاده از فرمول (۸) شاخص تایل برای هر دو نوع پارامتری کردن این توزیع بدست می‌آید:

$$T_{RS} = \int_0^1 \left( \frac{\lambda_1 \lambda_2 + u^{\lambda_2} - (1-u)^{\lambda_2}}{\mu_{RS} \lambda_2} \right) \log \left( \frac{\lambda_1 \lambda_2 + u^{\lambda_2} - (1-u)^{\lambda_2}}{\mu_{RS} \lambda_2} \right) du, \quad (9)$$

$$\begin{aligned} T_{FMKL} &= \int_0^1 \left( \frac{\lambda_1}{\mu_{FMKL}} + \frac{1}{\lambda_2 \mu_{FMKL}} \left( \frac{u^{\lambda_2} - 1}{\lambda_2} - \frac{(1-u)^{\lambda_2} - 1}{\lambda_2} \right) \right) \\ &\quad \log \left( \frac{\lambda_1}{\mu_{FMKL}} + \frac{1}{\lambda_2 \mu_{FMKL}} \left( \frac{u^{\lambda_2} - 1}{\lambda_2} - \frac{(1-u)^{\lambda_2} - 1}{\lambda_2} \right) \right) du, \end{aligned} \quad (10)$$

که  $\mu_{RS}$  و  $\mu_{FMKL}$  به ترتیب میانگین روش پارامتری  $RS$  و  $FMKL$  می‌باشند. بدست آوردن انتگرال‌های (۹) و (۱۰) به روش‌های مستقیم تا حدی مشکل است لذا با شبیه‌سازی مونت کارلو آنها را بدست می‌آوریم.

## ۶ یافته‌های عددی

در این تحقیق هزینه، درآمد و سود حاصل از فروش کلیه شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران که از ابتدای سال ۱۳۹۰ تا پایان سال ۱۳۹۱ فعالیت مستمر داشته‌اند، را در نظر گرفته‌ایم. ابتدا جهت ارائه‌ی اطلاعات

کلی از داده‌های مورد استفاده، به آمار توصیفی متغیرهای هزینه، درآمد و سود پرداخته‌ایم که نتایج به شرح جدول ۹ است.

**جدول ۹: آمارتوصیفی داده‌ها**

سال	میانگین	انحراف معیار	چولگی	کشیدگی
هزینه				
۱۳۹۰	۸۲۷۶۰	۵۱۱۴۸۳/۲	۱۴/۰۴۰۵	۲۱۹/۳۰۶۹
۱۳۹۱	۴۶۵۱۰	۱۱۳۰۴۴/۸	۴/۳۹۵۷	۲۴/۲۰۹۵
درآمد				
۱۳۹۰	۱۸۸۹۰۰۰	۱۰۶۷۱۲۳۱	۱۳/۰۰۵۷	۱۹۹/۰۲۴
۱۳۹۱	۶۹۹۶۰۰	۲۴۴۶۲۹۶	۵/۳۵۰۱	۳۱/۴۵۴۴
سود				
۱۳۹۰	۳۹۳۴۰۰	۱۳۸۸۵۱۸	۷/۸۶۶۶	۸۱/۳۱۸۶
۱۳۹۱	۳۹۷۷۰۰	۱۵۸۳۰۳۹	۶/۵۹۰۶	۴۸/۱۰۵۹

همانطور که در جدول ۹ مشاهده می‌کنیم مقادیر چولگی و کشیدگی مربوط به داده‌ها زیاد می‌باشد که بیانگر این حقیقت است که داده‌ها انحراف زیادی از توزیع نرمال دارند.

سپس با استفاده از روش درست‌نمایی ماکسیمم پارامترهای  $GLD$  را برای هر دو روش پارامتری  $RS$  و  $FMKL$  برای هر سه متغیر هزینه، درآمد و سود محاسبه نموده و توزیع  $GLD$  را به این داده‌ها برازش می‌دهیم و مقدار شاخص تایل را با استفاده از فرمول‌های (۷)، (۹) و (۱۰) برای آنها محاسبه می‌کنیم که نتایج به شرح جدول ۱۲ می‌باشد.

مقادیر مبهم گزارش شده در جدول ۱۲ به دلیل آن است که مقادیر  $\lambda_3$  و  $\lambda_4$  در شرط ذکر شده در قضیه ۱.۲ صدق نمی‌کنند و لذا در این موارد امیدریاضی توزیع تعریف نمی‌شود و نمی‌توان با استفاده از  $GLD$  شاخص تایل را برای آنها حساب کرد. همچنین مقادیر \* جدول نیز که در ستون  $T_{RS}$  دیده می‌شود به این دلیل است که پارامترهای حاصل از برازش داده‌ها در روش پارامتری  $RS$  یک توزیع معتبر را تولید نمی‌کند و نشان‌دهنده‌ی محدودیت برازش  $RS - GLD$  می‌باشد. در حالت کلی  $FMKL - GLD$  برازش بهتری را نتیجه می‌دهد و در اکثر موارد می‌توان شاخص تایل را بر اساس آن محاسبه کرد.

جدول ۱۰: مقادیر شاخص تایل برای میزان فروش و درآمد و سود شرکت‌های بورس

$T_{FMKL}$	$T_{RS}$	$T$	سال
۴,۶۷۸۶	*	۲,۳۷۹۳	۱۳۹۰
مبهم	*	۱,۳۱۲۴	۱۳۹۱
هزینه			
۳,۲۵۸۱	مبهم	۲,۳۴۶۳	۱۳۹۰
۲,۳۸۸۰	*	۱,۸۳۳۷	۱۳۹۱
درآمد			
۲,۰۷۲۴	*	۱,۷۴۱۵	۱۳۹۰
۲,۲۳۸۴	مبهم	۲,۰۵۷۵	۱۳۹۱
سود			

## ۷ دست‌آوردهای پژوهش

در این پژوهش ضمن معرفی توزیع لمبدای تعمیم‌یافته و یکی از روش‌های برآورد پارامترهای آن، شاخص نابرابری تایل را بر اساس هر دو روش پارامتری کردن  $RS - GLD$  و  $FMKL - GLD$  برای داده‌های هزینه، درآمد و سود حاصل از فروش کلیه‌ی شرکت‌های پذیرفته شده در بورس اوراق بهادار تهران، محاسبه کردیم.

## مراجع

- [1] Lakhany, A. and Massuer, H. (2000). Estimating the parameters of the generalised lambda distribution. *Algo Research Quarterly*, 3(3), 47-58.
- [2] Shannon, C. E. (1948), A mathematical theory of communication. *Bell System Technical Journal*, 27, 370-432.
- [3] Karian, Z.A. and Dudewicz. E.J. (2000). Fitting Statistical Distributions to Data: Generalized Lambda Distribution and Generalized Bootstrap Methods. *CRC Press, New York*.

[4] Su, S. (2007). Numerical maximum log likelihood estimation for generalized lambda distributions. *Journal of Computational statistics and data analysis* 51(8), 3983-3998.

[5] Theil, H. (1967). Economics and Information Theory. *Amsterdam: North Holland.*



## برآورد اطلاع متقابل براساس تابع مفصل

محمدی، م. ۱، عمادی، م. ۲، امینی، م. ۲ و بزرگ، م. ۱

<sup>۱</sup>گروه آمار، دانشکده علوم پایه، دانشگاه زابل

<sup>۲</sup>گروه آمار، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد

### چکیده

اطلاع متقابل و فاصله کولبک-لیبلر معیارهایی برای اندازه گیری وابستگی چندگانه بین برداری از متغیرهای تصادفی هستند. این معیارها برحسب تابع چگالی توأم و توابع چگالی حاشیه ای به دست می آیند. در این مقاله دو معیار اطلاع متقابل و فاصله کولبک-لیبلر را برحسب تابع مفصل محاسبه می نمائیم و نشان می دهیم که فقط به تابع چگالی توأم مفصل بستگی دارند. در انتها یک روش جدید برای برآورد میزان اطلاع متقابل برحسب تابع مفصل ارائه و مشاهده نمودیم که میزان خطای برآورد بسیار اندک است. همچنین میزان دقت روش برآورد را مورد بررسی قرار داده ایم.

**کلمات کلیدی:** آنتروپی شانون، اطلاع متقابل، تابع مفصل، نرمال دو متغیره.

## ۱ پیش گفتار

مطالعه تابع مفصل<sup>۲</sup> و کاربردهای آن یکی از پدیده های اخیر علم آمار است. تاریخچه تابع مفصل با کارهای فرشه (۱۹۵۱) آغاز شده است. نظریه مفصل در سال ۱۹۵۹ توسط اسکالر بدین گونه مطرح گردید که توزیع توأم را می توان به دو بخش توزیع های حاشیه ای و تابع مفصل، که ساختار وابستگی متغیرها را نشان می دهد، تقسیم نمود.

<sup>۱</sup> Email: Mo.Mohammadi@uoz.ac.ir

<sup>۲</sup> Copula

مفهوم آنتروپی<sup>۳</sup> اولین بار در علم فیزیک معرفی و برای کمی سازی مفاهیمی چون عدم حتمیت و بی نظمی به کار برده شد. بعدها از این مفهوم برای اندازه گیری اطلاعات حاصل از مشاهدات متغیرهای تصادفی استفاده شد. آنتروپی یک پیشامد میزان اطلاعی را که وقوع آن پیشامد در بردارد، اندازه گیری می کند. این میزان تابعی از احتمال رخ دادن پیشامد است. فرض کنید  $X$  یک متغیر تصادفی پیوسته با تابع توزیع  $F_X(x)$  و تابع چگالی  $f_X(x)$  باشد. شانون<sup>۴</sup> آنتروپی (مقدار اطلاع) متغیر تصادفی  $X$  را به صورت زیر تعریف کرد:

$$H(X) = - \int_X f(x) \log f(x) dx. \quad (1)$$

برآورد آنتروپی بر اساس مشاهدات  $x_1, x_2, \dots, x_n$  را آنتروپی نمونه ای گویند. مبنای لگاریتم در عبارت فوق دلخواه است. ما در این مقاله از لگاریتم طبیعی استفاده می کنیم. هدف ما در این مقاله این است که ارتباطی بین نظریه اطلاع و تابع مفصل بیان می کنیم. در واقع معیارهای اطلاع متقابل و فاصله کولبک-لیبلر را بر حسب تابع مفصل به دست می آوریم و خواص آن را برای مفصل نرمال مورد بررسی قرار می دهیم. بنابراین جهت سهولت کار فقط حالت دو متغیره پیوسته را بررسی خواهیم نمود.

## ۲ تابع مفصل و خواص آن

**قضیه ۱.۲.** (اسکلار، ۱۹۵۹) فرض کنید  $X$  و  $Y$  متغیرهای تصادفی به ترتیب با توابع توزیع  $F_X$  و  $F_Y$  و تابع توزیع توأم  $F$  باشند. در این صورت مفصلی مانند  $C$  وجود دارد به طوری که،

$$F(x, y) = C(F_X(x), F_Y(y)). \quad (2)$$

اگر  $F_X$  و  $F_Y$  پیوسته باشند، آن گاه  $C$  منحصر به فرد می باشد، در غیر این صورت  $C$  روی  $RanF \times RanG$  یکتا خواهد بود، که  $RanF_X$  و  $RanF_Y$  به ترتیب برد توابع توزیع حاشیه ای  $F_X$  و  $F_Y$  می باشند. بالعکس اگر  $C$  یک مفصل و  $F_X$  و  $F_Y$  توابع حاشیه ای باشند، آن گاه تابع  $F$  تعریف شده در رابطه (۲) یک تابع توزیع توأم با حاشیه ای های  $F_X$  و  $F_Y$  می باشد.

فرض کنید  $X$  و  $Y$  متغیرهای تصادفی با توابع توزیع  $F_X$  و  $F_Y$ ، توزیع توأم  $F$  و مفصل متناظر  $C$  باشند. آن گاه  $X$  و  $Y$  مستقل اند اگر و تنها اگر،

$$C(u, v) = u.v = \Pi(u, v).$$

$\Pi(u, v)$  مفصل حاصلضرب (استقلال) نام دارد. فرض کنید  $X$  و  $Y$  دو متغیر تصادفی به ترتیب با توابع چگالی حاشیه ای  $f_X(x)$  و  $f_Y(y)$  و توابع توزیع حاشیه ای  $u = F_X(x)$  و  $v = F_Y(y)$  و تابع مفصل  $C$  باشند. آن گاه تابع

<sup>۳</sup>Entropy

<sup>۴</sup>Shannon

چگالی مفصل به صورت زیر تعریف می‌شود،

$$c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v},$$

یا به عبارت دیگر، اگر  $f(x, y)$  تابع چگالی توأم دو متغیر تصادفی  $X$  و  $Y$  باشد، آن‌گاه،

$$c(u, v) = \frac{f(x, y)}{f_X(x) \cdot f_Y(y)}.$$

(تابع مفصل نرمال) این تابع مفصل از خانواده مفصل‌های بیضوی است که شکل پارامتری آن به صورت زیر است:

$$\begin{aligned} C^{Gu}(u, v, \rho) &= \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v)) \\ &= \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} \exp\left\{-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right\} dx dy \end{aligned}$$

به طوری که  $\Phi_\rho$  تابع توزیع نرمال استاندارد دو متغیره با ضریب همبستگی  $\rho$  می‌باشد. برای مطالعه بیشتر توابع مفصل و ویژگی‌های آن می‌توان به [۵] مراجعه نمود.

### ۳ اطلاع توأم، شرطی، متقابل و فاصله کولبک-لیبلر

در حالتی که دو متغیر تصادفی پیوسته  $X$  و  $Y$  با چگالی احتمال توأم  $f(x, y)$  وجود دارند، مقدار اطلاع توأم به صورت زیر تعریف می‌شود:

$$H(X, Y) = - \int_Y \int_X f(x, y) \log f(x, y) dx dy.$$

مقدار اطلاع شرطی  $x$  به شرط  $y$  به صورت زیر تعریف می‌شود:

$$H(X|Y) = - \int_Y \int_X f(x, y) \log f(x|y) dx dy,$$

به طور مشابه،

$$H(Y|X) = - \int_Y \int_X f(x, y) \log f(y|x) dx dy.$$

به سادگی می‌توان نتیجه گرفت که،

$$\begin{aligned} H(X, Y) &= H(X) + H(Y|X) \\ &= H(Y) + H(X|Y) \\ &\leq H(X) + H(Y). \end{aligned}$$

اگر  $X$  و  $Y$  به طور آماری مستقل باشند، داریم،

$$H(X, Y) = H(X) + H(Y).$$

مقدار اطلاع متقابل<sup>۵</sup> برای دو متغیر تصادفی پیوسته  $X$  و  $Y$  با تابع چگالی توام  $f(x, y)$  و به ترتیب توزیع‌های حاشیه‌ای  $f_X(x)$  و  $f_Y(y)$  به صورت زیر تعریف می‌شود:

$$I(X; Y) = \int_Y \int_X f(x, y) \log \left( \frac{f(x, y)}{f_X(x) \cdot f_Y(y)} \right) dx dy.$$

روابط زیر برای دو متغیر تصادفی  $X$  و  $Y$  بین مقدار اطلاع توام، شرطی و متقابل برقرار هستند.

$$I(X; Y) = H(X) + H(Y) - H(X, Y) \quad (۱)$$

$$I(X; Y) = I(Y; X) \quad (۲)$$

$$I(X; Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \quad (۳)$$

$$0 \leq I(X; Y) \leq \min(H(X), H(Y)) \quad (۴)$$

**نکته ۱.۳.** اطلاع متقابل را می‌توان به عنوان ابزاری برای اندازه‌گیری وابستگی بین متغیرهای تصادفی تفسیر نمود. اگر  $X$  و  $Y$  مستقل باشند آن‌گاه  $I(X; Y) = 0$  و کم‌ترین میزان اطلاع متقابل را خواهیم داشت. در صورتی که  $X$  و  $Y$  کاملاً وابسته باشند آن‌گاه  $H(Y|X) = 0$ ،  $H(X; Y) = H(Y)$  و بیشترین مقدار ممکن خود را می‌گیرد. برای اثبات روابط بیان شده می‌توان به [۱] مراجعه نمود.

فاصله کولبک-لیبلر<sup>۶</sup> جهت اندازه‌گیری فاصله میان دو تابع چگالی  $f_1(x)$  و  $f_2(x)$  به صورت زیر به صورت زیر تعریف می‌شود:

$$KL(f_1|f_2) = \int_X f_1(x) \log \frac{f_1(x)}{f_2(x)} dx.$$

این فاصله برابر صفر است اگر و فقط اگر  $f_1(x) = f_2(x)$ . در صورتی که فاصله کولبک-لیبلر بین تابع چگالی توام دو متغیر تصادفی پیوسته  $X$  و  $Y$  و توابع چگالی حاشیه‌ای آن‌ها را بخواهیم، می‌توان نوشت:

$$KL(f(x, y) \| f_X(x) \cdot f_Y(y)) = \int_Y \int_X f(x, y) \log \left( \frac{f(x, y)}{f_X(x) f_Y(y)} \right) dx dy.$$

در واقع این فاصله معیاری برای تعیین میزان وابستگی دو متغیر تصادفی پیوسته است. اگر  $X$  و  $Y$  مستقل باشند آن‌گاه  $f(x, y) = f_X(x) \cdot f_Y(y)$  و در نتیجه،

$$KL(f(x, y) \| f_X(x) \cdot f_Y(y)) = 0.$$

در حالت کلی این فاصله عددی در بازه  $[0, +\infty)$  است.

**مثال ۴.** فرض کنید  $(X, Y) \sim N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho^2)$ . به راحتی می‌توان نتایج زیر را به دست آورد.

$$H(X) = \log(\sigma_1 \sqrt{2\pi e}) \quad (۱)$$

<sup>۵</sup>Mutual Information

<sup>۶</sup>Kulback-Leibler



$$H(X|Y) = \log \left( \sigma_1 \sqrt{1 - \rho^2} \sqrt{2\pi e} \right) \quad (۲)$$

$$H(X, Y) = \log \left( 2\pi e \sigma_1 \sigma_2 \sqrt{1 - \rho^2} \right) \quad (۳)$$

در نهایت،

$$\begin{aligned} I(X; Y) &= H(X) - H(X|Y) \\ &= -\frac{1}{2} \log(1 - \rho^2). \end{aligned} \quad (۳)$$

همانطور که در رابطه (۳) مشاهده می‌کنیم بین پارامتر  $\rho$  در توزیع نرمال دومتغیره و میزان اطلاع متقابل رابطه مستقیم وجود دارد.

در صورتی که  $\rho = 0$ ، مقدار اطلاع متقابل برابر صفر خواهد شد و از این مطلب نیز می‌توان نتیجه گرفت که اطلاع متقابل ساختار وابستگی بین دو متغیر تصادفی را بیان می‌کند.

## ۴ اطلاع متقابل و فاصله کولبک-لیبلر بر اساس تابع چگالی مفصل

در این بخش معیارهای سنجش وابستگی اطلاع متقابل و فاصله کولبک-لیبلر را بر اساس تابع چگالی مفصل به دست می‌آوریم. لازم به ذکر است، مطالب این بخش بر اساس [۲]، [۳] و [۴] ارائه می‌گردد. (آنتروپی مفصل) فرض کنید  $X$  و  $Y$  دو متغیر تصادفی به ترتیب با تابع توزیع‌های حاشیه‌ای  $U = F_X(x)$  و  $V = F_Y(y)$  و تابع چگالی مفصل  $c(u, v)$  باشند. در این صورت آنتروپی مفصل به صورت زیر تعریف می‌شود:

$$H_c(X, Y) = - \int \int c(u, v) \log c(u, v) du dv = -E(\log c(u, v)).$$

آنتروپی مفصل فقط تابعی بر حسب تابع چگالی مفصل است و مستقل از توزیع‌های حاشیه‌ای می‌باشد.

قضیه ۱.۴. اطلاع متقابل دو متغیر تصادفی برابر با منفی آنتروپی مفصل آنهاست. به عبارت دیگر،

$$I(X; Y) = H_c(X, Y).$$

برهان.

$$I(X; Y) = \int_Y \int_X f(x, y) \log \frac{f(x, y)}{f_X(x) \cdot f_Y(y)} dx dy.$$

با توجه به تعریف تابع چگالی مفصل داریم،

$$I(X; Y) = \int_Y \int_X c(F_X(x), F_Y(y)) f_X(x) \cdot f_Y(y) \log \left( c(F_X(x), F_Y(y)) \right) dx dy.$$

با اعمال تبدیل  $F_X(x) = U$  و  $F_Y(y) = V$  می‌توان نوشت:

$$\begin{aligned} I(X; Y) &= \int \int c(u, v) \log c(u, v) du dv \\ &= -H_c(X, Y). \end{aligned}$$

□

همان‌طور که مشاهده می‌کنیم برای یافتن آنتروپی مفصل دو متغیر تصادفی کافی است که میزان اطلاع متقابل آن‌ها را محاسبه کنیم و یا برعکس. اطلاع متقابل بر اساس تابع مفصل یک معیار بسیار خوب برای یافتن ساختار وابستگی بین متغیرها است. مقدار این معیار نامنفی است و کم‌ترین مقدار خود را در حالت استقلال بین متغیرهای تصادفی خواهد پذیرفت.

**مثال ۵.** آنتروپی مفصل نرمال را با توجه به قضیه قبل و مثال ۳ به راحتی می‌توان به دست آورد:

$$H_c(X, Y) = \frac{1}{\rho} \log(1 - \rho^2),$$

که در آن  $\rho$  پارامتر مفصل نرمال است.

رابطه بین آنتروپی شانون و آنتروپی مفصل در حالت دومتغیره به صورت زیر است:

$$H(X, Y) = H(X) + H(Y) + H_c(X, Y).$$

فاصله کولبک-لیبلر بین دو تابع مفصل به صورت زیر تعریف می‌شود:

$$KL(C_1 \parallel C_2) = \int \int c_1(u, v) \log \frac{c_1(u, v)}{c_2(u, v)} du dv.$$

فاصله کولبک-لیبلر بین هر تابع مفصل با مفصل حاصلضرب نشان‌دهنده ساختار وابستگی بین دو متغیر تصادفی

است زیرا،

$$KL(C \parallel \Pi) = \int \int c(u, v) \log \frac{c(u, v)}{\pi(u, v)} du dv.$$

از طرفی با توجه به تعریف مفصل حاصلضرب می‌دانیم،

$$\pi(u, v) = C_\pi(u, v) = \frac{\partial^2 \Pi(u, v)}{\partial u \partial v} = 1.$$

در نتیجه،

$$\begin{aligned} KL(C \parallel \Pi) &= \int \int c(u, v) \log c(u, v) du dv \\ &= -H_c(X, Y) = I(X; Y). \end{aligned}$$

## ۵ برآورد اطلاع متقابل و فاصله کولبک بر اساس تابع مفصل

در این بخش به کمک شبیه‌سازی مونت-کارلو، بسته نرم افزاری تابع مفصل در نرم‌افزار R و مراحل که توسط نویسندگان مقاله ارائه گردیده است می‌توان اطلاع متقابل را بر اساس تابع مفصل برای هر تعداد متغیر تصادفی برآورد

نمود.

فرض کنید نمونه‌ای تصادفی به حجم  $n$  از دو متغیر تصادفی  $X$  و  $Y$  در اختیار داریم. برای برآورد میزان اطلاع متقابل به ترتیب مراحل زیر را اجرا می‌کنیم:

(۱) تابع توزیع تجربی  $X$  و  $Y$  یعنی به ترتیب  $\hat{F}_{X_n}(x)$  و  $\hat{F}_{Y_n}(y)$  را محاسبه می‌کنیم.

(۲) قرار می‌دهیم  $\hat{u}_n = \hat{F}_{X_n}(x)$  و  $\hat{v}_n = \hat{F}_{Y_n}(y)$ .

(۳) تابع مفصل مناسب داده‌ها را به کمک آزمون نیکویی برازش تشخیص می‌دهیم.

(۴) تابع چگالی مفصل  $(\hat{c}_n(\hat{u}_n, \hat{v}_n))$  را به ازای مقادیر  $X$  و  $Y$  به دست می‌آوریم.

به کمک شبیه‌سازی مونت-کارلو اطلاع متقابل را توسط رابطه زیر برآورد می‌کنیم:

$$\hat{K}L_n(C \parallel \Pi) = \hat{I}_n(X; Y) = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \log \hat{c}_{ij}(\hat{u}_j, \hat{v}_j),$$

که در آن  $m$  تعداد تکرارها در شبیه‌سازی و  $n$  حجم نمونه است.

برای مثال فرض کنید  $X$  و  $Y$  دو متغیر تصادفی از توزیع نرمال دو متغیره استاندارد با ضریب همبستگی  $\rho$  باشند. می‌دانیم که میزان اطلاع متقابل برای این دو متغیر تصادفی به صورت تحلیلی (جبری) برابر با  $-\frac{1}{2} \log(1 - \rho^2)$  است. حال این مقدار را به کمک مراحل فوق برای تابع مفصل نرمال و مقادیر مختلف پارامتر  $\rho$  و حجم نمونه برآورد می‌کنیم. نتایج برآورد اطلاع متقابل بر اساس تابع مفصل نرمال در جدول زیر ارائه گردیده است.

در جدول ۱۱  $\hat{I}_n(X; Y)$  در واقع فاصله بین مفصل نرمال و مفصل حاصلضرب است. همان‌طور که مشاهده می‌کنیم در  $\rho = 0$  مقدار  $\hat{I}_n(X; Y)$  بسیار کوچک و نزدیک به صفر است. با افزایش مقدار  $\rho$ ، فاصله بین مفصل نرمال و مفصل حاصلضرب افزایش می‌یابد، زیرا وابستگی بین متغیرها افزایش یافته است. همان‌طور که ملاحظه می‌کنیم مقدار خطای برآورد که در واقع تفاضل مقدار اطلاع متقابل تحلیلی و برآورد اطلاع متقابل بر اساس تابع مفصل می‌باشد بسیار اندک است و این مقدار با افزایش حجم نمونه نیز کاهش می‌یابد.

در انتها میزان اریبی (*Bias*)، میانگین مربعات خطا (*MSE*) و میانگین قدرمطلق خطا (*MAD*) مربوط به مراحل ارائه شده در ابتدای این بخش را به ازای مقادیر مختلف پارامتر مفصل نرمال و حجم نمونه‌های مختلف در جدول ۱۲ محاسبه نموده‌ایم که این قسمت نیز توسط نویسنده این مقاله ارائه طراحی شده است. مشاهده می‌کنیم که میزان اریبی، میانگین مربعات خطا و میانگین قدرمطلق خطا به طور کلی بسیار کم و ناچیز است. این مقادیر با افزایش حجم نمونه کاهش می‌یابند.

جدول ۱۱: نتایج ۱۰۰۰ بار شبیه‌سازی برای برآورد اطلاع متقابل تابع توزیع نرمال دومتغیره و مفصل نرمال

قدرمطلق خطای برآورد	$I_n(X; Y)$	$\hat{I}_n(X; Y)$	$n$	$\rho$
$2/80858 \times 10^{-1}$	۰/۱۴۳۸۴	۰/۱۳۷۰۱	۱۰	-۰/۵
$2/90126 \times 10^{-1}$	۰/۱۴۳۸۴	۰/۱۴۶۲۸	۲۰	
$2/99049 \times 10^{-3}$	۰/۱۴۳۸۴	۰/۱۴۶۸۳	۱۰۰	
$8/50335 \times 10^{-6}$	۰/۱۴۳۸۴	۰/۱۴۳۸۳	۵۰۰	
$2/19979 \times 10^{-16}$	۰	$2/19979 \times 10^{-16}$	۱۰	۰
$2/18581 \times 10^{-16}$	۰	$2/18581 \times 10^{-16}$	۲۰	
$2/19362 \times 10^{-16}$	۰	$2/19362 \times 10^{-16}$	۱۰۰	
$2/19817 \times 10^{-16}$	۰	$2/19817 \times 10^{-16}$	۵۰۰	
$2/87186 \times 10^{-1}$	۰/۱۴۳۸۴	۰/۱۴۳۳۴	۱۰	۰/۵
$2/82774 \times 10^{-1}$	۰/۱۴۳۸۴	۰/۱۳۸۹۳	۲۰	
$2/62068 \times 10^{-3}$	۰/۱۴۳۸۴	۰/۱۴۶۴۶	۱۰۰	
$1/17041 \times 10^{-3}$	۰/۱۴۳۸۴	۰/۱۴۵۰۱	۵۰۰	

## ۶ نتیجه‌گیری

در این مقاله ارتباطی بین تابع مفصل و نظریه اطلاع بیان کردیم. همچنین یک روش جدید برای برآورد اطلاع متقابل و فاصله کولبک-لیبلر بر اساس تابع مفصل ارائه کردیم. به راحتی می‌توان میزان اطلاع متقابل و فاصله کولبک-لیبلر را برای سایر تابع مفصل‌ها برآورد کرد. روش ارائه شده در این مقاله برای یافتن ساختار وابستگی بسیار دقیق‌تر از روش‌های معمول همبستگی از قبیل همبستگی پیرسون است و محدود به بررسی روابط خطی نمی‌باشد. اطلاع متقابل بر اساس تابع مفصل برای بسیاری از تابع مفصل‌ها شکل بسته‌ای ندارد اما به کمک الگوریتم بیان‌شده می‌توان این معیار را برای هر تابع مفصلی محاسبه کرد. با توجه به فاصله کولبک-لیبلر بیان‌شده بر اساس تابع مفصل، می‌توان فاصله بین هر دو تابع مفصل دیگر را نیز محاسبه نمود.

جدول ۱۲: بررسی دقت برآورد اطلاع متقابل بر اساس تابع مفصل نرمال

$MAD$	$MSE$	$Bias$	$n$
۰/۰۰۰۳۹۲۴	۰/۰۰۰۰۰۳۲	۰/۰۰۰۰۲۴۲	۱۰
۰/۰۰۰۲۷۳۸	۰/۰۰۰۰۰۱۹	۰/۰۰۰۰۰۹۱	۲۰
۰/۰۰۰۱۲۸۵	۰/۰۰۰۰۰۰۳	۰/۰۰۰۰۰۴۷	۱۰۰
۰/۰۰۰۰۵۵۷	۰/۰۰۰۰۰۰۱	۰/۰۰۰۰۰۳۷	۵۰۰

## مراجع

- [۱] جان سی. الف. وان در لوب؛ ترجمه حسنعلی آذرنوش، نظریه اطلاع، انتشارات دانشگاه فردوسی مشهد، ۱۳۸۰.
- [2] Blumentritt, T. and Schmid, F. (2012), *Mutual information as a measure of multivariate association: analytical properties and statistical estimation*, Journal of Statistical Computation and Simulation. 82, 1257–1274.
- [3] Jian, M. and Zengqi, S. (2011), *Mutual Information Is Copula Entropy*, Tsinghua Science and Technology. 16, 51-54.
- [4] Kumar, P. (2012), *Statistical Dependence: Copula Functions and Mutual Information Based Measures*, Journal of Statistics Applications Probability. 1, 1-14.
- [5] Nelsen, R.B. (2006). *An Introduction to Copulas*. Springer.



## کنکاشی در آنتروپی مفصل

ولی پور مقدم، ح. <sup>۱</sup> و محتشمی برزادران، غ. ر.  
گروه آمار، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد

### چکیده

در این مقاله پس از بیان مقدماتی از نظریه اطلاع و تعریف آنتروپی، از مفهوم تابع مفصل برای توصیف عدم قطعیت مرتبط با سیستم‌های احتمالی استفاده می‌شود. سپس ارتباط بین آنتروپی توأم با آنتروپی مفصل، بیان و معیار بهینه‌سازی آنتروپی مفصل بررسی می‌شود. در ادامه با مثالی نشان داده می‌شود ماکسیم آنتروپی توأم با آنتروپی‌های حاشیه‌ای معلوم، با ماکسیم نمودن آنتروپی مفصل معادل است. همچنین ارتباط بین آنتروپی مفصل و اندازه‌ی اطلاع متقابل بررسی می‌شود.

**کلمات کلیدی:** تابع توزیع، مفصل، اندازه‌های اطلاع، آنتروپی، آنتروپی مفصل.

## ۱ مفاهیمی از نظریه اطلاع

عدم قطعیت نقش حیاتی برای داشتن اطلاع از پدیده‌های اطراف و آنچه که مشاهده می‌شود، ایفا می‌کند. کاهش عدم قطعیت نیازمند افزایش اطلاع است. به بیان وان درلوپ (۱۹۹۷)، شانون (۱۹۴۸) اولین کسی است که اطلاع را با مفهوم احتمال، پیوند داد و اندازه‌ی احتمالی عدم قطعیت، را که اشاره به آنتروپی دارد، تعریف نمود. رابطه‌ی زیر اندازه‌ی اطلاع (آنتروپی) شانون برای متغیر پیوسته  $X$  را نشان می‌دهد.

<sup>۱</sup>Email: hosein.valipourmoqhaddam@stu.um.ac.ir

$$H(X) = - \int f(x) \ln f(x) dx = E(\ln f(X)) \quad (۱)$$

توجه شود  $H(X)$  تعریف شده برای یک متغیر تصادفی پیوسته، می‌تواند منفی شود. در حالتی که دو متغیر تصادفی  $X$  و  $Y$  با چگالی احتمال توأم  $f(x, y)$  وجود داشته باشند مقدار اطلاع توأم (آنتروپی توأم) به صورت زیر تعریف می‌شود:

$$H(X, Y) = - \int \int f(x, y) \ln f(x, y) dx dy \quad (۲)$$

رابطه‌ی زیر به آسانی تحقیق می‌شود.

$$H(X, Y) = H(Y) + H(X|Y) = H(X) + H(Y|X) \quad (۳)$$

زمانی که  $X$  و  $Y$  مستقل باشند، رابطه‌ی زیر برقرار است.

$$H(X, Y) = H(X) + H(Y) \quad (۴)$$

برای متغیرهای تصادفی پیوسته که به یکدیگر وابسته‌اند به جای استفاده از آنتروپی شرطی می‌توان از آنتروپی مفصل استفاده نمود. در حالت کلی، وابستگی تصادفی متغیرهای تصادفی  $X$  و  $Y$  کاهش آنتروپی توأم را نتیجه می‌دهد.

## ۲ آنتروپی مفصل

بر اساس پژوهش جینز (۱۹۵۷) اگر متغیرهای آنتروپی در مدل آنتروپی به طور متقابل مستقل باشند، فرضیه‌ی اساسی اصل آنتروپی ماکسیمم، وجود دارد. برای توصیف متغیرهای آنتروپی دارای وابستگی، می‌توان از نظریه‌ی مفصل کمک گرفت. (ژائو ولین (۲۰۱۱))

فرض کنید  $u_i = F_i(x_i)$  که  $i = 1, \dots, n$  باشد، می‌توان آنتروپی مفصل را به صورت زیر

$$H_c(U_1, \dots, U_n) = - \int_0^1 \dots \int_0^1 c(u_1, \dots, u_n) \ln c(u_1, \dots, u_n) du_1, \dots, du_n$$

تعریف نمود. برای سادگی از این پس  $(U_1, \dots, U_n) = \mathbf{U}$  و  $I = [0, 1]$  را در نظر می‌گیریم.

$$H_c(\mathbf{U}) = - \int_{I^n} c(\mathbf{u}) \ln c(\mathbf{u}) d\mathbf{u}$$

بر اساس نظریه‌ی مفصل و پژوهش‌های اسکالر (۱۹۵۹) می‌توان آنتروپی توأم  $H(\mathbf{X})$  را به صورت مجموع  $n$  آنتروپی حاشیه‌ای  $H(X_i)$  و آنتروپی توأم مفصل  $H_c(\mathbf{U})$  نوشت.

$$H(\mathbf{X}) = \sum_{i=1}^n H(X_i) + H_c(\mathbf{U}) \quad (۵)$$

ماکسیم نمودن آنتروپی توأم با آنتروپی‌های حاشیه‌ای معلوم، معادل با ماکسیم نمودن آنتروپی مفصل  $H_c(\mathbf{U})$  است. قبل از بیان برهان رابطه‌ی فوق، بهتر است مطالب زیر یادآوری شوند (برای اطلاعات بیشتر به نلسن (۲۰۰۶) رجوع کنید).

با استفاده قضیه‌ی اسکالر در حالت  $n$  بعدی، می‌توان نتیجه گرفت:

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$$

با گرفتن مشتق جزئی از رابطه‌ی فوق، چون  $u_i = F_i(x_i)$ ، تابع چگالی احتمال توأم به صورت زیر حاصل می‌شود:

$$f(x_1, \dots, x_n) = c(u_1, \dots, u_n) \prod_{i=1}^n f_i(x_i) \quad (۶)$$

که  $c(u_1, \dots, u_n)$  تابع چگالی مفصل است.

دیفرانسیل‌ها به صورت  $du_i = dF_i(x_i) = f_i(x_i)dx_i$  و  $d\mathbf{x} = \prod_{i=1}^n dx_i$  نوشته می‌شوند همچنین  $d\mathbf{u} = \prod_{i=1}^n f_i(x_i)dx_i$  است و می‌توان نشان داد:

$$\int_{I^{n-1}} c(\mathbf{u}) \prod_{j=1, j \neq i}^n du_j = \int_{\mathbb{R}^{n-1}} \frac{f(x_1, \dots, x_n)}{f_i(x_i)} \prod_{j=1, j \neq i}^n dx_j = \frac{f_i(x_i)}{f_i(x_i)} = 1$$

برهان:



با استفاده از تعریف آنتروپی شانون و رابطه‌ی (۲.۲) می‌توان نوشت.

$$\begin{aligned}
 H(\mathbf{X}) &= - \int_{\mathbb{R}^n} \left( c(\mathbf{u}) \prod_{i=1}^n f_i(x_i) \right) \ln \left( c(\mathbf{u}) \prod_{i=1}^n f_i(x_i) \right) d\mathbf{x} \\
 &= - \int_{\mathbb{R}^n} \left( c(\mathbf{u}) \prod_{i=1}^n f_i(x_i) \right) \left( \sum_{i=1}^n \ln f_i(x_i) \right) \prod_{i=1}^n dx_i \\
 &\quad - \int_{\mathbb{R}^n} c(\mathbf{u}) \ln c(\mathbf{u}) \prod_{i=1}^n f_i(x_i) dx_i \\
 &= - \sum_{i=1}^n \int_{\mathbb{R}^n} \left( c(\mathbf{u}) \prod_{j=1, j \neq i}^n f_j(x_j) dx_j \right) f_i(x_i) \ln f_i(x_i) dx_i \\
 &\quad - \int_{I^n} c(\mathbf{u}) \ln c(\mathbf{u}) d\mathbf{u} \\
 &= - \sum_{i=1}^n \left( \int_{I^{n-1}} c(\mathbf{u}) \prod_{j=1, j \neq i}^n du_j \right) \left( \int_{\mathbb{R}} f_i(x_i) \ln f_i(x_i) dx_i \right) + H_c(\mathbf{U}) \\
 &= - \sum_{i=1}^n \int_{\mathbb{R}} f_i(x_i) \ln f_i(x_i) dx_i + H_c(\mathbf{U}) \\
 &= \sum_{i=1}^n H(X_i) + H_c(\mathbf{U})
 \end{aligned}$$

اثبات‌های دیگری برای رابطه‌ی (۱.۲) توسط لوو و همکاران (۲۰۱۳) برای حالت دو متغیره و ژائو و لین (۲۰۱۱) برای حالت  $n$  متغیره بیان شده است. با توجه به رابطه‌ی (۱.۲) می‌توان با استفاده از آنتروپی مفصل  $H_c(\mathbf{U})$  ساختار وابستگی را تحلیل نمود.

در این قسمت معیار بهینه‌سازی (ماکسیم‌سازی) آنتروپی مفصل، با مثالی شرح داده می‌شود. با توجه به پژوهش‌های پوگازو و همکاران (۲۰۱۰) یکی از کاربردهای آنتروپی مفصل، بازسازی تصاویر برش نگاری رایانه‌ای است. برای بازسازی تصاویر برش نگاری رایانه‌ای به روش تابع مفصل، باید به دنبال تعیین چگالی احتمال متغیر تصادفی پیوسته‌ای که به ماکسیم مقدار اطلاع منجر می‌شود، باشیم. تابع چگالی احتمال  $f(x, y)$  را با استفاده از ماکسیم نمودن آنتروپی شانون می‌یابیم.

وان درلوپ (۱۹۹۷) بیان می‌کند، محدودیت‌ها توأما، طبیعت چگالی احتمال را که به ماکسیم مقدار اطلاع منجر می‌شود تعیین می‌کنند. در واقع ماکسیم مقدار اطلاعی که یک منبع می‌تواند بدهد بستگی به محدودیت‌ها دارد. برای همین توزیع‌های حاشیه‌ای را به عنوان محدودیت مطرح می‌کنیم تا تابع چگالی مورد نظر برای بازسازی تصاویر برش نگاری رایانه‌ای مناسب باشد.

## ۱.۲ مثال

مقدار آنتروپی ماکسیم تحت قیود زیر با  $\hat{f}$  نشان داده می‌شود. محدودیت‌های مسئله به شرح زیر است:

$$\begin{cases} \int f(x, y) dy = f_1(x), & \text{به ازای هر } y \\ \int f(x, y) dx = f_2(y), & \text{به ازای هر } x \\ \iint f(x, y) dx dy = 1, \\ f(x, y) \geq 0 \end{cases}$$

از آنجایی که محدودیت‌ها خطی اند ملاک تابع مقعر، انتخاب می‌شود تا راه حل منحصر به فردی برای حل مسئله ارائه کند. تابعک آنتروپی شانون به عنوان تابع هدف در نظر گرفته می‌شود.

$$J(f) = H(X, Y) = - \iint f(x, y) \ln f(x, y) dx dy \quad (7)$$

ابزار اصلی، تعریف تکنیک ضرایب لاگرانژ است. برای حل معادله‌ی تابعک لاگرانژ باید فرض کنید تنها یک  $f \geq 0$  شدنی با آنتروپی محدود قانع کننده، وجود دارد.

$$\begin{aligned} \mathcal{L}_g(f, \lambda_0, \lambda_1, \lambda_2) = & J(f) + \lambda_0 \left( 1 - \iint f(x, y) dx dy \right) \\ & + \int \lambda_1(x) \left( f_1(x) - \int f(x, y) dy \right) dx \\ & + \int \lambda_2(y) \left( f_2(y) - \int f(x, y) dx \right) dy \end{aligned}$$

با توجه به دستگاه معادلات زیر نقاط بحرانی پیدا می‌شود.

$$\begin{cases} \frac{\partial \mathcal{L}_g(f, \lambda_0, \lambda_1, \lambda_2)}{\partial f} = 0, \\ \frac{\partial \mathcal{L}_g(f, \lambda_0, \lambda_1, \lambda_2)}{\partial \lambda_i} = 0. \end{cases}$$

فرض کنید انتگرال‌ها بر بازه  $[0, 1]$  همگرا باشند، از حل دستگاه معادلات فوق، نتیجه‌ی زیر حاصل می‌شود.

$$\hat{f}(x, y) = \exp(-\lambda_1(x) - \lambda_2(y) - \lambda_0)$$

در اینجا  $\lambda_1, \lambda_2, \lambda$  به کمک محدودیت‌های ذکر شده محاسبه می‌شوند.

$$\begin{aligned}\lambda_1(x) &= -\ln \left( f_1(x) \int \lambda_1(x) dx \right) \\ \lambda_2(y) &= -\ln \left( f_2(y) \int \lambda_2(y) dy \right) \\ \lambda &= \ln \left( \int \lambda_1(x) dx \int \lambda_2(y) dy \right)\end{aligned}$$

که به رابطه‌ی زیر منجر می‌شود:

$$\hat{f}(x, y) = f_1(x)f_2(y) \quad (۸)$$

رابطه‌ی فوق تابع چگالی‌ای که آنتروپی شانون را ماکسیم می‌کند، نشان می‌دهد. مفصل ماکسیم آنتروپی شانون دارای توزیع یکنواخت ( $c(u, v) = 1$ ) است. با چگالی بدست آمده از اصل ماکسیم آنتروپی شانون، می‌توان مفصل دو متغیره متناظر را پیدا کرد.

$$\begin{aligned}F(x, y) &= \int_0^x \int_0^y f(s, t) ds dt \\ &= \int_0^x \int_0^y f_1(s)f_2(t) ds dt \\ &= \int_0^x f_1(s) dt \int_0^y f_2(t) ds\end{aligned}$$

تابع توزیع تجمعی و تابع مفصل را به ترتیب می‌توان به صورت زیر نوشت.

$$\begin{aligned}F(x, y) &= F_1(x)F_2(y) \\ C(u, v) &= F(F_1^{-1}(u), F_2^{-1}(v)) = uv\end{aligned}$$

تابع مفصلی که از آنتروپی ماکسیم شانون بدست آمد معروف به مفصل استقلال است که استقلال بین دو متغیر تصادفی را توصیف می‌کند. آنتروپی توأم شانون زمانی به ماکسیم مقدار خود می‌رسد که تابع چگالی توأم برابر، حاصلضرب چگالی‌های حاشیه‌ای باشد پس به عبارتی باید متغیرهای تصادفی از یکدیگر مستقل باشند تا آنتروپی توأم ماکسیم شود.

### ۳ آنتروپی مفصل و اندازه‌ی اطلاع متقابل

در نظریه‌ی احتمال و نظریه‌ی اطلاع، اطلاع متقابل دو متغیر تصادفی، کمیتی است که میزان وابستگی متقابل دو متغیر تصادفی را بیان می‌کند. اندازه‌ی اطلاع متقابل (فاصله از استقلال)  $X$  و  $Y$  نامیده می‌شود زیرا اگر  $X$  و  $Y$

مستقل باشند، اندازه‌ی اطلاع متقابل صفر می‌شود و هر چقدر اندازه‌ی اطلاع متقابل از صفر فاصله بگیرد، از استقلال فاصله می‌گیرد. برای حالت پیوسته اطلاع متقابل به صورت زیر

$$I(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \ln \left( \frac{f(x, y)}{f_1(x) \cdot f_2(x)} \right) dx dy \quad (9)$$

تعریف می‌شود.

**برهان.** اگر تابع آنتروپی مفصل دو متغیره به صورت زیر

$$H_c(U, V) = - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c(u, v) \ln c(u, v) dudv \quad (10)$$

باشد که  $u = F_1(x)$  و  $v = F_2(y)$  آنگاه تابع چگالی توأم به صورت زیر

$$f(x, y) = f_1(x) \cdot f_2(y) \cdot c(f_1(x), f_2(y))$$

تعریف می‌شود.

$$\begin{aligned} I(X, Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \ln \left( \frac{f(x, y)}{f_1(x) \cdot f_2(x)} \right) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_1(x) \cdot f_2(y) \cdot c(u, v) \ln \left( \frac{f_1(x) \cdot f_2(y) \cdot c(u, v)}{f_1(x) \cdot f_2(x)} \right) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_1(x) \cdot f_2(y) \cdot c(u, v) \ln c(u, v) dx dy \end{aligned}$$

که به راحتی رابطه‌ی زیر نتیجه می‌شود.

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c(u, v) \ln c(u, v) dudv = -H_c(U, V) \quad (11)$$

□

می‌توان اطلاع متقابل را به صورت زیر نیز تعریف نمود

$$I(X, Y) = H(X) + H(Y) - H(X, Y) = -H_c(U, V) \quad (12)$$

با توجه به این رابطه نیز می‌توان نتیجه گرفت تابع دو متغیره‌ی اطلاع متقابل نوع خاصی از آنتروپی مفصل است. پس به عبارتی آنتروپی مفصل، معیاری برای نشان دادن فاصله‌ی استقلال دو متغیر تصادفی است. رابطه‌ی

(۳.۳) نشان می‌دهد زمانی اندازه‌ی اطلاع متقابل صفر می‌شود که دو متغیر تصادفی از یکدیگر مستقل یا تابع چگالی مفصل‌شان برابر یک شود. همچنین از رابطه‌ی (۴.۳) می‌توان نتیجه گرفت که آنتروپی مفصل در حالت استقلال متغیرها باید صفر شود.

### ۱.۳ توجه

در مثال فوق تابع چگالی توأمی را که دارای حداکثر مقدار اطلاع شانون بود، یافتیم. تابع چگالی توأم به صورت حاصلضرب حاشیه‌ها بدست آمد که نشان می‌داد زمانی تابع چگالی توأم دارای حداکثر اطلاع شانون است که متغیرهای تصادفی مستقل باشند. سپس مفصل متناظر آن را یافتیم. مفصل مورد نظر، مفصل استقلال بود. با توجه به بخش ۳ چون متغیرهای تصادفی مستقل بودند، اندازه اطلاع متقابل برای مثال فوق برابر با صفر حاصل شد. از آنجایی که در بخش ۳ بیان شد که آنتروپی مفصل به نوعی همان اندازه متقابل است، می‌توانیم نتیجه بگیریم که آنتروپی مفصل مثال فوق برابر صفر است. حال می‌توانیم بیان کنیم در مثال فوق، آنتروپی توأم برابر، مجموع آنتروپی‌های حاشیه‌ای است یعنی زمانی آنتروپی توأم ماکسیمم می‌شود که آنتروپی‌های حاشیه‌ای ماکسیمم شوند.

### ۴ نتیجه‌گیری

در این مقاله ضمن معرفی آنتروپی مفصل و ویژگی‌های آن، به دنبال یافتن توزیعی بودیم که دارای آنتروپی ماکسیمم شانون تحت شرایط و محدودیت‌های خاص است و با حل معادلاتی که از محدودیت‌ها حاصل می‌شود، مقادیر آنتروپی ماکسیمم و فرم تابع چگالی دارای آنتروپی ماکسیمم را بدست آوردیم. سپس با نوشتن تابع توزیع تجمعی و با استفاده از قضیه‌ی اسکالر تابع مفصل متناظر را محاسبه نمودیم. با توجه به مفصل بدست آمده، نتیجه می‌گیریم از این پس به جای ماکسیمم نمودن آنتروپی تابع چگالی احتمال توأم، می‌توان ماکسیمم‌سازی آنتروپی را با توجه به دیدگاه نظریه‌ی مفصل بررسی نمود.

### مراجع

- [1] Lu, Ch. Vijay, P. Shenglian, Z. Jianzhong, Lei Y., (2013). Copula entropy coupled with artificial neural network for rainfall-runoff simulation, Stoch Environ Res Risk Assess.
- [2] Jaynes, E.T., (1957). Information theory and statistical mechanics, Phys. Rev, Vol:106, 620-630.
- [3] Nelsen, R.B., (2006). An Introduction to Copulas, Springer Series in Statistics, second ed. Springer, NewYork.

- [4] Pougaza , D.B. Djafari , A.M. Bercher, J.F., (2009). Using the notion of copula in tomography, in XXIIe colloque GRETSI, Dijon, France.
- [5] Shannon, C.E. (1948). A Mathematical Theory of Communication, Bell. Syst. Tech. J, Vol:27, 379–423.
- [6] Van der Lubbe, J.C.A., (1997). Information Theory, Cambridge University Press.
- [7] Zhao, N.L. Winston, T., (2011). A copula entropy approach to correlation measurement at the country level, Applied Mathematics and Computation, Vol: 218, 628–642.