
بسم الله الرحمن الرحیم

دانشگاه یزد

دانشکده ریاضی

گروه آمار

پایان‌نامه

برای دریافت درجه کارشناسی ارشد

آمار ریاضی

ماکسیمم آنتروپی چند متغیره: شناسایی، تبدیل و وابستگی

استاد راهنما:

دکتر حجت‌اله ذاکرزاده

استاد مشاور:

دکتر علی دولتی

پژوهش و نگارش:

سیدمجتبی نجفی‌فراشاه

اسفند ۱۳۸۹

تقدیم به:

پدر دلسوزم

که مسیر سربلندی و مردانگی را به شیواترین روش به من آموخت

و

مادر مهربانم

که بهشت خدا با حقارت زیر پایش جای گرفت.

شکرگزارم او را که مسیر رفتنم را به نورهای مکرر روشن ساخت، چرا که اگر خواست او نبود این روشنایی‌ها را به‌همراه نداشتم. هر چند قلم زدن در حیطه اساتید دلسوزم که این اندک نوشته‌های ناچیز حاصل بهره‌وری این حقیر از روشنائی تابنده این بزرگواران است درخور کلام و بضاعت این قلم نیست ولی

آب دریا را اگر نتوان کشید   هم به قدر تشنگی باید چشید

سپاسگزارم از استاد راهنمای ارجمندم آقای دکتر حجت اله ذاکر زاده که بادقت و حوصله زائدالوصفی، راهنمائی‌های ارزنده‌ای را در جهت تدوین ونگارش این تحقیق ارائه نمودند و از استاد مشاور محترم آقای دکتر علی دولتی که تصحیح آن را متقبل شدند و نظرات ارزشمندی در تکوین این پایان‌نامه ابراز نمودند.

سپاسگزارم از آقای دکتر حمزه ترابی که داوری این پایان‌نامه را پذیرفته‌اند همچنین از آقای دکتر غلام رضا محتشمی که قبول زحمت فرموده و برای داوری این پایان‌نامه تشریف آورده‌اند که حضور ایشان برای اینجانب موجب افتخار است.

سپاسگزارم از استادان بزرگواری که پیش از این مفتخر به شاگردی در خدمتشان بوده‌ام: دکتر حمزه ترابی، دکتر عیسی محمودی و دکتر سید محسن میرحسینی.

سپاسگزارم از پدر و مادر صبورم و برادر و خواهر مهربانم که حضورشان مایه دلگرمی و یادشان مایه شادمانی من است.

چکیده

در این پایان‌نامه به بررسی توزیع‌های ماکسیمم آنتروپی، تحت قیود گشتاور می‌پردازیم. بعد از تعریف اصل ماکسیمم آنتروپی، مثال‌هایی از توزیع‌های ماکسیمم آنتروپی نسبت به این قیود ارائه می‌شود. نشان داده می‌شود که توزیع‌های چند متغیره ماکسیمم آنتروپی را می‌توان با استفاده از یک فرم نمایش کلی، مشخصه‌سازی کرد. در این رویکرد، مسئله مشخصه‌سازی ماکسیمم آنتروپی که قبلاً بر اساس اثبات‌های پیچیده‌ی حساب تغییرات صورت می‌گرفت، به مسئله ساده نمایش تابع چگالی به یک فرم مشخص تبدیل می‌شود. ابزار اصلی مشخصه سازی توزیع‌های ماکسیمم آنتروپی در این راهبرد، روابط تشخیص پذیری اطلاع است که به سادگی به حالت چند متغیره قابل تعمیم هستند. ساختار وابستگی توزیع ماکسیمم آنتروپی برحسب گشتاورها و تکیه‌گاه توزیع نیز مطالعه می‌شود. ارتباط خانواده توزیع‌های ماکسیمم آنتروپی با خانواده توزیع‌های نمایی و نمایی شرطی بررسی شده است. به عنوان کاربرد نتایج به دست آمده، مشخصه‌سازی ماکسیمم آنتروپی برای تعدادی از توزیع‌های دومتغیره ارائه شده است.

فهرست مطالب

۱ **مفاهیم و تعاریف اولیه** ۳

۱.۱ مقدمه ۴

۲.۱ آنتروپی شانون ۴

۳.۱ ویژگی‌های آنتروپی ۵

۴.۱ اطلاع کولبک ـ لیبلر ۸

۱.۴.۱ ویژگی‌های اطلاع کولبک لیبلر ۹

۲.۴.۱ معیارهایی برای تمایز بین توزیع‌ها ۱۰

۵.۱ اطلاع متقابل ۱۱

۶.۱ قواعد زنجیره‌ای ۱۳

۱.۶.۱ قاعده زنجیره‌ای برای آنتروپی ۱۳

۷.۱ تابع محدب ۱۶

۸.۱ ضرایب لاگرانژ ۱۶

۲ **اصل ماکسیمم آنتروپی** ۱۹

۱.۲ مقدمه ۲۰

۱.۱.۲ بهینه‌سازی چیست؟ ۲۰
۲.۱.۲ چرا آنتروپی را بهینه می‌کنیم؟ ۲۱
۲.۲ اصل ماکسیمم آنتروپی ۲۲
۱.۲.۲ چرا از توزیع ماکسیمم آنتروپی استفاده می‌کنیم؟ ۲۲
۳.۲ نمونه‌هایی از توزیع‌های ماکسیمم آنتروپی ۲۴
۱.۳.۲ نا مساوی حداکثر آنتروپی ۲۶
۴.۲ نمونه‌های بیشتری از توزیع‌های ماکسیمم آنتروپی ۲۹
۵.۲ مثبت بودن و یکتایی توزیع ماکسیمم آنتروپی ۳۴

۳ شناسایی ماکسیمم آنتروپی **۳۹**

۱.۳ مقدمه ۴۰
۲.۳ مینیمم اطلاع تشخیص ۴۱
۱.۲.۳ شناسایی ماکسیمم آنتروپی ۴۶
۲.۲.۳ قانون زنجیره‌ای ماکسیمم آنتروپی ۴۷
۳.۳ ماکسیمم آنتروپی در خانواده نمایی ۵۱
۱.۳.۳ خانواده شرطی نمایی و ماکسیمم آنتروپی ۵۴
۴.۳ کاربردها: مشخصه ماکسیمم آنتروپی خانواده‌های دومتغیره ۵۹
۵.۳ توزیع‌هایی با تکیه‌گاه نامحدود ۵۹
۱.۵.۳ توزیع نمایی دو متغیره مطلقاً پیوسته ۵۹
۲.۵.۳ توزیع دو متغیره شرطی پارتو ۶۰
۳.۵.۳ توزیع دو متغیره شرطی نمایی ۶۱

۴.۵.۳ توزیع دو متغیره نوع ۲ با شرطی گاما ۶۲
۶.۳ توزیع ماکسیمم آنتروپی با تکیه‌گاه محدود ۶۲
۱.۶.۳ تابع چگالی توزیع دو متغیره نمایی با تکیه گاه مثلثی ۶۳
۲.۶.۳ توزیع دو متغیره با شرطی پارتو ۶۳
۳.۶.۳ توزیع دو متغیره با شرطی بتا ۶۳
۷.۳ توزیع‌ها با تکیه گاه در $\mathbb{R}^2 \cup \mathbb{R}$ ۶۴
۱.۷.۳ توزیع نمایی دو متغیره مارشال ـ الکین ۶۵
۲.۷.۳ توزیع دو متغیره کادراس ـ آگه ۶۷
۳.۷.۳ توزیع نمایی دو متغیره طبیعی ۶۸
۴.۷.۳ توزیع دو متغیره اتورگرسیو نمایی ۶۹

۴ تبدیلات و وابستگی در توزیع‌های ماکسیمم آنتروپی ۷۱
۱.۴ مقدمه ۷۲
۲.۴ تبدیل ماکسیمم آنتروپی ۷۲
۳.۴ اطلاع وابسته ۷۷

الف پیوست ۸۷

ب واژه‌نامه‌ی فارسی به انگلیسی ۸۹

پ واژه‌نامه‌ی انگلیسی به فارسی ۹۳

منابع و مراجع ۹۶

نمادها

مجموعه‌ی اعداد حقیقی	$\mathbb{R}$
اطلاع متقابل	I
آنتروپی شانون	H
اطلاع کولبک لیبلر	K
تابع نشانگر	I_A
تابع گاما	Γ
مجموعه گشتاور اطلاع	$\mathcal{T}$
ضرایب لاگرانژ	λ

پیشگفتار

بعد از معرفی مفهوم آنتروپی توسط شانون [۱] [۵۰]، طی شصت سال گذشته تاکنون، مقالات زیادی پیرامون کار شانون و تعمیم‌های آن نوشته شده است. به عنوان نمونه می‌توان به کاپور [۲][۳۷]، کولبک[۳][۳۸] اشاره کرد.

اصل ماکسیمم آنتروپی اولین بار توسط جینز[۴][۳۱] بیان گردید. در کلاسی از توزیع‌ها، ماکسیمم کردن آنتروپی تحت قیودی خاص، نقشی اساسی در آمار و نظریه اطلاع دارد. ایده او این مطلب بوده است که، ماکسیمم آنتروپی، به انتخاب یک تابع چگالی احتمال، با آگاهی‌های موجود سازگار است کمک می‌کند و اطلاعات توجیه شده و تضمین شده‌ای را معرفی می‌کند.

تعیین توزیع‌های یک متغیره ماکسیمم آنتروپی توسط تعدادی از محققین از جمله رضا[۵][۴۹]، گیاسو[۶][۳۰] مورد مطالعه قرار گرفت. کگان[۷] و همکاران [۳۴]، برای تعداد زیادی از توزیع‌های گسسته و پیوسته ویژگی‌های ماکسیمم آنتروپی را مطالعه نموده‌اند. کاپور [۳۵] و محتشمی برزادران [۸][۴۳]، مسئله‌ی مشخص سازی توزیع‌های چند متغیره با استفاده از آنتروپی شانون را مورد بررسی قرار داده‌اند. تابع توزیع ماکسیمم آنتروپی نیز تحت قیودی خاص، برای توزیع‌های چند متغیره پیوسته و گسسته، توسط اورزو [۹][۵۴]، کاپور [۳۵]، زوگرافوس[۱۰][۵۷] و ابراهیمی[۱۱] و همکاران [۲۶] مورد بررسی قرار گرفته است. ابراهیمی [۲۵]، کاربردهایی از ماکسیمم آنتروپی، همچنین اسدی[۱۲] و

[۱] Shannon

[۲] Kapur

[۳] Kullback

[۴] Jaynes

[۵] Reza

[۶] Guiasu

[۷] Kagan

[۸] Mohtashami Borzadaran

[۹] Urzua

[۱۰] Zografos

[۱۱] Ebrahimi

[۱۲] Asadi

همکاران [۸] توزیع‌های ماکسیمم آنتروپی را بر اساس قیودی از تابع مخاطره و بقاء بررسی کردند. روش‌های ماکسیمم آنتروپی مدل‌های احتمالی را برای زمینه‌های تصادفی که از اطلاعات داده شده متشکل شده‌اند، تهیه می‌کند [۳۱، ۳۲، ۵۱، ۵۲]. در آمار، شمار قابل توجهی از روش های تشخیص و آزمون فرض‌ها برای توزیع‌ها، بر مبنای مشخصه ماکسیمم آنتروپی، پایه‌گذاری و گسترده شده‌اند [۱۳، ۱۶، ۲۰، ۲۱، ۲۲، ۲۴، ۲۸، ۴۰، ۴۴، ۴۶، ۴۷، ۴۸، ۵۳، ۵۵].

این پایان‌نامه به شرح زیر مرتب شده است در فصل اول به بیان مفاهیم اولیه و تعاریف مورد نیاز و استفاده شده، در این پایان‌نامه پرداخته شده است. در فصل دوم آنتروپی و اصل ماکسیمم آنتروپی معرفی و برای روشن‌تر شدن این تعاریف چندین مثال ارائه می‌شود. فصل سوم نتیجه‌هایی برای شناسایی توزیع‌های ماکسیمم آنتروپی با توجه به قیود گشتاور بیان شده است و جستجوی رابطه خانواده توزیع‌های ماکسیمم آنتروپی با خانواده توزیع‌های نمایی و نمایی شرطی در این فصل آورده شده است. در قسمت دیگر از فصل سوم مشخصه‌های ماکسیمم آنتروپی را برای تعدادی از خانواده‌های دو متغیره ارائه شده است. در فصل چهارم نتایجی از تبدیل ماکسیمم آنتروپی و کاربرد آن در مشخصه ماکسیمم آنتروپی در چندین خانواده دو متغیره آورده شده است همچنین به بررسی و بحث در مورد ساختار وابستگی توزیع‌های ماکسیمم آنتروپی می‌پردازیم.

فصل ۱

مفاهیم و تعاریف اولیه

۱.۱ مقدمه

در این فصل به بیان مفاهیم اولیه مانند تعریف آنتروپی شانون، اطلاع کولبک لیبلر، اطلاع متقابل و دیگر تعاریف مورد نیاز و استفاده شده، در این پایان‌نامه می‌پردازیم. در اکثر بخش‌های این فصل روابط و اثبات‌ها برای حالت گسسته بیان شده است که با استفاده از انتگرال به جای مجموع به حالت پیوسته، قابل تبدیل است.

۲.۱ آنتروپی شانون

برای تابع چگالی احتمال پیوسته $f(x)$ روی مجموعه S، آنتروپی به صورت زیر است

$$H(f) = -\int_S f(x)\log f(x)dx. \tag{۱.۱}$$

برای تابع احتمال گسسته $p(x) = p = (p_۱, ..., p_n)$ روی مجموعه $\{x_۱, x_۲, \dots\}$، $p_i = p(x_i)$، آنتروپی به صورت زیر تعریف می‌شود:

$$H(p) = -\sum_{i\geq ۱} p_i \log\, p_i. \tag{۲.۱}$$

قرار می‌دهیم $۰ \log ۰ = ۰$ که با توجه به نمودار $x \log x$ طبیعی است. نمادهای دیگری از جمله $H(X)$ و $H(F)$ را نیز برای آنتروپی استفاده می‌شود. این تعریف از آنتروپی که توسط شانون معرفی شده است [۵۰] با فرمولی که از تصویر ترمودینامیکی آنتروپی وجود دارد نیز شباهت دارد. از نظر فیزیکی، سیستمها باید به حالت‌هایی با بیشترین آنتروپی ظاهر شوند تا به تعادل برسند. در احتمال، به $H(p)$، به عنوان یک اندازه از اطلاعی که p دارد، نگاه می‌شود. آنتروپی بیشتر به معنی عدم قطعیت بیشتر یا بیشتر بودن کمبود اطلاع در مورد یک پدیده است.

مثال ۱.۲.۱ . *مجموعه متناهی $\{x_۱, x_۲, \dots, x_n\}$ را در نظر بگیرید. اگر $p(x_۱) = ۱$ و برای $j > ۱$، داریم $p(x_j) = ۰$ آنگاه $H(p) = -۱ \log ۱ = ۰$ است. تابع آماری، به‌وسیله p تنها یک برآمد ممکن $x_۱$ را ارائه می‌دهد و آگاهی کاملی از آنچه اتفاق خواهد افتاد داریم. از طرف دیگر، اگر p تابع چگالی احتمال یکنواخت گسسته باشد که در آن $p(x_j) = \frac{۱}{n}$ برای همه j ها، آنگاه*