بسم الله الرحمن الرحیم

دانشگاه رازی
دانشکده علوم
گروه آمار

پایان‌نامه جهت اخذ درجه دکتری رشته آمار

عنوان پایان نامه

مقایسه تصادفی چند متغیره متغیرهای همراه و کاربردهای آنها

استاد راهنما:
دکتر بهاءالدین خالدی

نگارش:
ابراهیم امینی سرشت

اسفند ۱۳۹۳

دانشکده علوم پایه
گروه آمار

پایان‌نامه جهت اخذ درجه دکتری
رشته آمار

نام دانشجو:
ابراهیم امینی سرشت

تحت عنوان :

مقایسه های تصادفی چند متغیره متغیرهای همراه و کاربردهای آنها

در تاریخ توسط هیأت داوران زیر بررسی و با درجه به تصویب نهایی رسید.

استاد راهنمای پایان‌نامه دکتر بهاءالدین خالدی با مرتبه‌ی علمی استاد امضاء:

استاد داور دکتر جعفر احمدی با مرتبه‌ی علمی استاد امضاء:

استاد داور دکتر عبدالرضا سیاره با مرتبه‌ی علمی دانشیار امضاء:

استاد داور دکتر مهرداد نیاپرست با مرتبه‌ی علمی استادیار امضاء:

کلیه حقوق مادی مترتب بر نتایج مطالعات، ابتکارات و
نوآوری‌های ناشی از تحقیق موضوع این پایان‌نامه
متعلق به دانشگاه رازی است.

خدایا...

به من زیستنی عطا کن که در لحظه مرگ، بر بی‌ثمری لحظه‌ای که برای زیستن گذشته است، حسرت نخورم و مُردنی عطا کن که بر بیهودگیش، سوگوار نباشم. بگذار تا آن را، خود انتخاب کنم، اما آنچنان که تو دوست می‌داری.
تو می‌دانی و همه می‌دانند که شکنجه دیدن بخاطر تو، زندانی کشیدن بخاطر تو و رنج بردن به پای تو تنها لذت بزرگ زندگی من است، از شادی توست که من در دل می‌خندم، از امید رهایی توست که برق امید در چشمان خسته‌ام می‌درخشد و از خوشبختی توست که هوای پاک سعادت را در ریه‌هایم احساس می‌کنم. نمی‌توانم خوب حرف بزنم. نیروی شگفتی را که در زیر کلمات ساده و جمله‌های ضعیف و افتاده، پنهان کرده‌ام دریاب، دریاب.
تو می‌دانی و همه می‌دانند که زندگی از تحمیل لبخندی بر لبان من، از آوردن برق امیدی در نگاه من، از برانگیختن موج شعفی در دل من، عاجز است.
تو، چگونه زیستن را به من بیاموز، چگونه مردن را خود خواهم آموخت.
به من توفیق تلاش در شکست، صبر در نومیدی، رفتن بی‌همراه، جهاد بی‌سلاح، کار بی‌پاداش، فداکاری در سکوت، دین بی‌دنیا، مذهب بی‌عوام، عظمت بی‌نام، خدمت بی‌نان، ایمان بی‌ریا، خوبی بی‌نمود، گستاخی بی‌خامی، قناعت بی‌غرور، عشق بی‌هوس، تنهایی در انبوه جمعیت، و دوست داشتن بی‌آنکه دوست بداند، روزی کن.

اگر تنهاترین تنهاشوم، باز خدا هست

او جانشین همه نداشتن‌هاست...

سپاس گزاری...

سپاس خداوندگار حکیم را که با لطف بی‌کران خود، آدمی را زیور عقل آراست.

در آغاز وظیفه خود می‌دانم از زحمات بی‌دریغ استاد فرزانه و بزرگوارم، جناب آقای دکتر خالدی، که از همه ی جنبه های مختلف علمی و اخلاقی در کنار ایشان همیشه دریچه های تازه ای از دانایی برایم گشوده میشد و کلام در وصف خوبی های ایشان قاصر است، صمیمانه و از صمیم قلب تشکر و قدردانی کنم که قطعاً بدون راهنمایی‌های ارزنده و ارزشمند ایشان، این مجموعه به انجام نمی‌رسید. همچنین از همه ی اساتید محترم گروه آمار که در این چند سال در کنار این عزیزان کسب علم و تجربه نموده ام کمال تشکر و قدردانی را دارم. در پایان، بوسه می‌زنم بر دستان خداوندگاران مهر و مهربانی، پدر و مادر عزیزم و بعد از خدا، ستایش می‌کنم وجود مقدس‌شان را و تشکر می‌کنم از همسر عزیزم به پاس عاطفه سرشار و گرمای امیدبخش وجودش، که در این سردترین روزگاران، بهترین پشتیبان من بودند.

تقدیم به

مکمل وجودم، همسرم سمانه و پدر

و مادر عزیزم

چکیده

درمتون قابلیت اعتماد ممکن است یک سیستم دارای مؤلفه‌های مستقلِ غیر هم‌توزیع باشد. همچنین ممکن است این سیستم تحت تاثیر یک یا چند متغیر محیطی قرار بگیرد. حالتی را در نظر می‌گیریم که این مدل سیستم‌ها تحت تاثیر یک متغیر محیطی یکسان قرار بگیرند. توزیع طول عمر چنین سیستم‌هایی با مدل‌های آمیخته معرفی می‌شوند. مقایسه‌های تصادفی چند متغیره طول عمر این مدل‌ها بر پایه ترتیب‌های چند متغیره تصادفی مانند ترتیب تصادفی نسبت درستنمایی چند متغیره، ترتیب تصادفی نرخ خطر معکوس چند متغیره و ترتیب تصادفی نرخ خطر چند متغیره مورد بررسی قرار می‌دهیم. همچنین مدل تک متغیره مدل‌های آمیخته با متغیر آمیخته کننده یکسان در نظر می‌گیریم و با استفاده از ترتیب تصادفی مانده طول عمر به مقایسه آن‌ها می‌پردازیم. در ادامه یک سیستم $n-k+1$ از n را در نظر می‌گیریم. پس از اینکه سیستم از کار می‌افتد توزیع طول عمر مؤلفه‌های باقیمانده به عنوان یک مدل آمیخته معرفی می‌شوند. از آنجا که اغلب منطقی به نظر می‌رسد که از مؤلفه‌های باقیمانده در سیستم‌های دیگر استفاده شود، لذا از این حیث به مقایسه تصادفی چند متغیره دو بردار از متغیرهای باقیمانده در مسائل دو نمونه‌ای مورد مطالعه قرار می‌دهیم. نهایتاً با توجه به اینکه توزیع متغیرهای همراه آماره‌های مرتب نیز حالت خاصی از مدل‌های آمیخته است، لذا در مسائل تک نمونه‌ای و دو نمونه‌ای به مقایسه تصادفی چند متغیره بردار تصادفی آماره‌های مرتب و متغیرهای همراه آن‌ها می‌پردازیم.

کلمات کلیدی:

آماره‌های مرتب، ترتیب تصادفی نسبت درستنمایی چند متغیره، ترتیب تصادفی نرخ خطر معکوس چند متغیره، ترتیب تصادفی نرخ خطر چند متغیره، ترتیب تصادفی میانگین مانده عمر، متغیرهای همراه، مدل آمیخته.

فهرست مطالب

فهرست نشانه‌ها و نمادها

hr نرخ خطر

rh نرخ خطر معکوس

lr نسبت درستنمایی

MRL میانگین مانده عمر

ILR صعودی در نسبت درستنمایی

DLR نزولی در نسبت درستنمایی

IFR صعودی در نرخ خطر

DFR نزولی در نرخ خطر

$IRFR$ صعودی در نرخ خطر معکوس

$DRFR$ نزولی در نرخ خطر معکوس

$IMRL$ صعودی در میانگین مانده عمر

$DMRL$ نزولی در میانگین مانده عمر

TP_2 مثبت کل از مرتبه ۲

MTP_2 وابستگی چند متغیره مثبت کل از مرتبه ۲

RR_2 منظم معکوس از مرتبه ۲

PQD وابستگی درجه دوم مثبت

$h_X(t)$ تابع نرخ خطر متغیر تصادفی X

$r_X(t)$ تابع نرخ خطر معکوس متغیر تصادفی X

$F_X(t)$ تابع توزیع متغیر تصادفی X

$F_{Y|X}(y|x)$ تابع توزیع متغیر تصادفی $[Y|X=x]$

SI صعودی بطور تصادفی

$\leq_{st}$ ترتیب تصادفی معمولی

$\leq_{rh}$ ترتیب تصادفی نرخ خطر معکوس

$\leq_{hr}$ ترتیب تصادفی نرخ خطر

$\leq_{lr}$ ترتیب تصادفی نسبت درستنمایی

$\leq_{mrl}$ ترتیب تصادفی میانگین مانده عمر

$\leq_{whr}$ ترتیب تصادفی نرخ خطر ضعیف

$X \overset{\mathrm{d}}{=} Y$ متغیرهای X و Y هم توزیع‌اند

$\mathbf{X}_{()}$ بردار تصادفی آماره‌های مرتب از نمونه X

پیشگفتار

امروزه در بخش صنعت و دیگر بخش ها با سیستم‌هایی سرو کار داریم که متشکل از تعدادی واحد است که برای انجام یک هدف به یکدیگر متصل شده‌اند. مشخصه مورد مطالعه در سیستم‌ها زمان شکست یا طول عمر سیستم و یا مؤلفه‌های آن است. منظور از شکست وضعیتی است که مؤلفه‌ی سیستم نتواند وظیفه محوله را در زمان مورد نظر به طور رضایت بخشی انجام دهد. بنابراین موضوعی که می‌تواند حائز اهمیت باشد مطالعه طول عمر مؤلفه‌های یک سیستم است. تاکنون مطالعه‌های وسیعی در مورد مقایسه قابلیت اعتماد طول عمر سیستم‌های منسجم[۱] صورت گرفته است. منظور از سیستم منسجم سیستمی است که دارای ساختاری است که نسبت به هر مؤلفه خود صعودی و شامل مؤلفه نامرتبط نباشد. برای حالتی که مؤلفه‌های سیستم دارای توزیع‌های احتمال مستقل و مشابه باشند مطالب بسیاری تحت بررسی قرار گرفته است. مقایسه طول عمر دو سیستم A و B اغلب برای طراحی و برنامه ریزی قابلیت اعتماد سیستم‌ها مورد نیاز است. به هر حال وقتی می‌گوییم یک سیستم بهتر از یک سیستم دیگر عمل می‌کند منظور چیست؟ و نیاز به دانستن چه چیزی است؟ برای مثال یک طراح شاید نیاز به چگونگی تخصیص اجزاء سیستم در سطح مؤلفه‌ها در مقابل به سطح سیستم باشد. برای مقایسه معناداری بین سیستم‌ها، به مفاهیم مختلفی از ترتیب‌های تصادفی [۲] از قبیل ترتیب تصادفی معمولی [۳]، ترتیب تصادفی نرخ خطر [۴] و غیره متوسل می شویم. ترتیب‌های تصادفی نقش ویژه ای در شاخه های مختلف آمار دارند. بولند[۵] (۱۹۹۸) استدلال کرد وقتی مردم می‌گویند محصول A بهتر از محصول B است، احتمالاً منظور آن‌ها این است که نرخ خطر محصول B نسبت به محصول A بیشتر است. بنابراین مقایسه سیستم‌ها با استفاده از میزان شکست موجه به نظر می‌رسد.

مطالعه بر روی خواص سیستم‌ها و مقایسه آن‌ها با سیستم‌های مشابه دیگر می‌تواند حائز اهمیت باشد. برای حالتی که مؤلفه‌های سیستم دارای توزیع‌های احتمال مستقل و مشابه باشند مطالب بسیاری در ارتباط با مقایسه‌های تصادفی بین دو سیستم متفاوت، تحت بررسی قرار گرفته است. در مواردی نیز ممکن است سیستم‌ها از مؤلفه‌هایی تشکیل شوند که طول عمرشان بطور مشابه توزیع نشده باشد و بیشتر آن‌ها وابسته باشند، مانند مؤلفه‌هایی که در یک محیط مشترک کار می‌کنند. مقایسه تصادفی [۶]چنین سیستم‌هایی نیز مورد مطالعه قرار گرفته است. در بسیاری از موارد ممکن است با سیستم‌هایی مواجه شویم که شامل n مولفه باشند که هر مؤلفه خود نیز شامل k مولفه وابسته

[۱] Coherent systems
[۲] Stochastic orderings
[۳] Usual stochastic ordering
[۴] Stochastic hazard rate ordering
[۵] Boland
[۶] Stochastic comparison

باشد. قابلیت اعتماد یک جزء از چنین سیستم‌هایی می‌تواند نقش بسیار مهمی در طول عمر سیستم داشته باشد. در مسائل یک نمونه‌ای مقایسه اجزاء مختلف چنین سیستم‌هایی در تجزیه تحلیل قابلیت اعتماد سیستم می‌تواند مورد علاقه باشد. همچنین در مسائل دو نمونه ای مقایسه تصادفی مولفه های چنین سیستم‌هایی با سیستم مشابه دیگر می‌تواند حائز اهمیت باشد. در بعضی از حالات ممکن است سیستم‌های مورد بررسی تحت یک متغیر محیطی [۷] (اثر تصادفی) یکسان مورد مطالعه قرار بگیرند. در این حالت طول عمر سیستم‌ها وابسته به عملکرد متغیر وابسته محیطی است. همچنین مقایسه دو سیستم، وابسته به نوع وابستگی بین متغیر محیطی و مؤلفه‌های سیستم خواهد بود. در این رساله به مطالعه و مقایسه تصادفی سیستم‌هایی می‌پردازیم که در یک محیط مشترک کار می‌کنند و به عبارتی تحت یک متغیر محیطی یکسان قرار گرفته‌اند.

داده‌های ترتیبی نقش ویژه‌ای در شاخه‌های مختلف آمار دارند. کاربرد گسترده متغیرهای تصادفی مرتب شده در مباحث استنباطی آمار به دلیل آن است که این متغیرها شامل حجم بالایی از اطلاعات نمونه می‌باشند. تاکنون مدل‌های مختلفی از متغیرهای تصادفی مرتب شده معرفی و مورد مطالعه قرار گرفته است که از جمله می‌توان آماره‌های مرتب، مقادیر رکورد، آماره‌های مرتب[۸] دنباله‌ای اشاره کرد. هر یک از این مدل ها با توجه به ساختاری که دارند برای تحلیل نوع خاصی از داده‌ها پیشنهاد شده اند. آماره‌های مرتب از مشهورترین متغیرهای تصادفی مرتب شده است که کاربرد فراوانی در مدل سازی و تجزیه و تحلیل قابلیت اعتماد سیستم‌ها دارد، به عنوان مثال طول عمر سیستم‌های منسجم، متشکل از n مؤلفه مستقل و هم توزیع در یکی از آماره‌های مرتب آن رخ می‌دهد. لذا آماره‌های مرتب زیر بنای مطالعه انواع شاخص های قابلیت اعتماد سیستم‌ها می باشند. برای گسترش مفهوم آماره‌های مرتب از حالت تک متغیره به چند متغیره روش سر راستی وجود ندارد و در این حالت انتظار نمی رود که مشابه آماره‌های مرتب در نمونه های تک متغیره بتوان نمونه هایی که از توزیع های چند متغیره استخراج می شوند را مرتب کرد. بنابراین در این حالت نمونه تصادفی را بر حسب یکی از مؤلفه‌های نمونه مرتب می‌کنیم و متغیرهای باقیمانده را به عنوان متغیر همراه [۹] آماره‌های مرتب شده در نظر می‌گیریم. در آزمایش‌های طول عمر، وقتی واحدهای مورد آزمایش، خود شامل مؤلفه‌های وابسته باشند، برای تعیین طول عمر مؤلفه‌ها نیاز به استفاده از متغیرهای همراه آماره‌های مرتب است. بنابراین در توزیع های طول عمر چند متغیره مقایسه تصادفی متغیرها بر پایه آماره‌های همراه می‌تواند مورد توجه باشد. متغیرهای همراه آماره‌های مرتب در زمینه‌های مختلفی کاربرد دارند. در روش انتخاب ها، آیتم ها و یا افراد ممکن است بر اساس مشخصه X، انتخاب شوند و مشخصه وابسته Y که اندازه‌گیری آن مشکل است ممکن است برای مشاهده، بعداً مورد علاقه باشد. به عنوان مثال مشخصه X ممکن است نمره یک نامزد انتخاباتی در یک آزمون به

[۷] Environment variable
[۸] Stochastic order
[۹] Concomitant variable

منظور تعیین عملکرد او باشد و مشخصه Y اندازه‌گیری نهایی عملکرد نامزد پس از انتخاب او باشد و یا ممکن است مقدار مشخصه X نمره یک فرد قبل از استخدام در یک شرکت باشد و مقدار مشخصه Y نمره همان فرد بر اساس عملکرد شغلی او در یک نمونه تصادفی از کارمندان باشد. برای مطالعه بیشتر در این زمینه می‌توان به یو و دیوید[۱۰] (۱۹۸۴) و دیوید (۱۹۹۳) مراجعه کرد.

در این رساله مدل‌های مختلفی از توزیع‌های آمیخته چند متغیره[۱۱] در نظر گرفته می شود و به مقایسه تصادفی آن‌ها به معانی مختلف ترتیب‌های تصادفی چند متغیره پرداخته می‌شود. در فصل ۱ به تعاریف و مفاهیم اساسی، که مورد نیاز در فصل های بعدی است می‌پردازیم و مدل‌هایی از توزیع های آمیخته چند متغیره که در فصل های بعدی مورد مطالعه قرار می‌گیرند را معرفی می‌کنیم. در فصل ۲ ، در مسائل دو نمونه‌ای حالتی را در نظر می گیریم که متغیرها تحت یک متغیر محیطی یکسان قرار بگیرند، به عبارتی در یک محیط مشترک کار کنند و سپس به مقایسه تصادفی آن‌ها به معانی مختلف ترتیب‌های تصادفی چند متغیره می‌پردازیم. در فصل ۳ متغیرهای بر جای مانده از یک سیستم منسجم، بعد از اینکه سیستم شکست می‌خورد را به عنوان حالتی خاص از مدل‌های آمیخته معرفی می‌کنیم و سپس در مسائل دو نمونه‌ای به مقایسه تصادفی آن‌ها به معانی مختلف ترتیب‌های چند متغیره پرداخته می شود و در فصل ۴ متغیرهای همراه آماره‌های مرتب در حالت چند متغیره معرفی می‌شود و سپس در مسائل یک نمونه‌ای و دو نمونه‌ای به مقایسه تصادفی آن‌ها می‌پردازیم. نهایتاً در فصل ۵ به جمع‌بندی مطالبی که در رساله بدست آورده شد می‌پردازیم.

[۱۰] Yuo and Daivid
[۱۱] Multivariate mixture

فصل ۱
تعاریف و مفاهیم اولیه

۱.۱ مقدمه

در زندگی روزمره با مدل‌هایی مواجه می‌شویم که تحت یک یا چند اثر تصادفی محیطی قرار دارند. چنین مدل‌هایی را مدل‌های آمیخته [۱] می‌نامند. این نمونه مدل‌ها در بسیاری از حالت‌ها در قابلیت اعتماد کاربرد فراوان دارد. به عنوان مثال در تئوری بیمه یا در مدیریت ریسک‌ها ممکن است فرض شود شکست یک ریسک وابسته به یک یا چند اثر تصادفی باشد. در قابلیت اعتماد نیز ممکن است متغیرهای طول عمر تحت تاثیر متغیرهای محیطی قرار بگیرند. لذا مدل‌های آمیخته از مهمترین مباحث آمار و احتمال کاربردی است که بسیاری از پدیده‌های تصادفی واقعی را توصیف می‌نمایند. بردار تصادفی $\mathbf{X}$ با تابع توزیع $\mathbf{F}$ را در نظر بگیرید. فرض کنید $(\mathbf{X}, \mathbf{F})$ مدل تصادفی باشد که پدیده تصادفی مورد علاقه را توصیف می‌نماید. اگر چه فرض استقلال و هم توزیعی بین مؤلفه‌های بردار $\mathbf{X}$ مراحل انجام توصیف پدیده مانند محاسبه شاخص‌های مهم مرکزی، پراکندگی و چولگی که بصورت تابعک‌هایی از $\mathbf{F}$ ظاهر می‌شوند را ساده می‌نماید، اما در بسیاری از پدیده‌های واقعی این فرض‌ها مناسب نیستند و می‌توان فرض نمود مؤلفه‌های $\mathbf{X}$ بصورت مستقیم یا غیر مستقیم از طریق مؤلفه‌های تصادفی خارجی(محیطی) وابسته هستند. به عبارتی دیگر فرض می‌شود $\mathbf{X} \stackrel{\mathrm{d}}{=} (Y_1(\Theta), \ldots, Y_n(\Theta))$ که بردار تصادفی Θ یکی از عوامل محیطی است که آن را متغیر محیطی نیز می‌نامند، علت وابستگی میان مؤلفه‌های $\mathbf{X}$ را ایجاد می‌نماید. در بعضی از حالات ممکن است وضعیتی پیش بیاید که علت وابستگی میان مؤلفه‌های $\mathbf{X}$ به چند عامل محیطی بستگی داشته باشد و یا به عبارتی عامل محیطی، $\boldsymbol{\Theta}$، بصورت یک بردار تعریف شود. به شرط اینکه $\theta = \Theta$ باشد، بردار تصادفی $(Y_1(\theta), \ldots, Y_n(\theta))$ دارای تابع توزیع $\mathbf{F}(.|\theta)$، از دیگر عوامل وابستگی به مؤلفه‌های $\mathbf{X}$ است. بنابراین مدل آماری مورد استفاده برای مطالعه پدیده مورد علاقه بصورت زیر در نظر گرفته می‌شود

$$\{(Y_1(\Theta), \ldots, Y_n(\Theta)),\ \mathbf{F}(.|\theta),\ \theta \in \mathcal{X}\}.$$

در این رساله دو مدل بصورت فوق با مؤلفه‌های متفاوت و عامل وابستگی محیطی یکسان Θ در نظر می گیریم و سپس به معانی مختلف ترتیب‌های تصادفی به مقایسه آن‌ها پرداخته می‌شود. حالت‌های دیگری نیز از مدل‌های آمیخته وجود دارند که در همین بخش به معرفی آن‌ها پرداخته می‌شود. در

[۱] Mixture models

حالتی که سیستم‌ها شامل مؤلفه‌هایی با طول عمرهای مستقل و هم توزیع باشند، با توجه به اینکه محاسبه برخی از شاخص‌های مهم در قابلیت اعتماد، نظیر تابع توزیع، تابع بقا، تابع نرخ خطر و غیره برای این سیستم‌ها زیاد پیچیده نیست، نتایج زیادی در دسترس می باشد، اما در حالتی که شامل مؤلفه‌هایی با طول عمر مستقل و غیر هم توزیع باشند با توجه به ساختار پیچیده قابلیت اعتماد آن‌ها یافته های زیادی در دسترس نیست. مقایسه تصادفی چنین سیستم‌هایی تحت متغیر محیطی از جمله مسائلی است که در این رساله به آن‌ها پرداخته می‌شود. در حالتی که سیستم‌ها تحت اثرهای محیطی متفاوت قرار بگیرند نتایج زیادی در ارتباط با مقایسه‌های تصادفی آن‌ها در حالت‌های تک متغیره و چند متغیره در دست رس است که برای مطالعه بیشتر در این زمینه می‌توان به خالدی و شیکد[۲] (۲۰۱۰)، بلزونسه [۳] و همکاران (۲۰۰۹) و میشرا و میشرا[۴] (۲۰۱۲) مراجعه کرد. در حالتی که سیستم‌ها تحت اثر متغیر محیطی یکسان قرار بگیرند تا به حال نتایجی در دسترس نیست که در فصل‌های بعدی مورد مطالعه قرار می‌گیرند و نتایجی در ارتباط با مقایسه‌های تصادفی آن‌ها به معانی مختلف در حالت چند متغیره بیان خواهد شد.

تولید کنندگان به طور معمول علاقمند به محاسبه و کنترل طول عمر محصولات تولیدی خود هستند، قابلیت اعتماد ممکن است یک ارزیابی خوب در این زمینه باشد. برای این منظور n واحد را بطور تصادفی از خط تولید انتخاب می کنیم و در آزمایش طول عمر قرار می‌دهیم. برای این داده های طول عمر، علاقمند به استنباط در مورد ویژگی های قابلیت اعتماد محصولات هستیم. اکنون حالتی را در نظر بگیرید که آزمایش طول عمر در وقتی که k-امین مولفه خراب مشاهده شود، پایان دهیم. این موضوع منطقی به نظر می رسد که ممکن است از نظر اقتصادی مقرون به صرفه باشد که از مؤلفه‌های بر جای مانده برای ساخت سیستم‌هایی دیگر استفاده کرد. در ادامه همین فصل نشان خواهیم داد که توزیع مؤلفه‌های باقیمانده به عنوان یک مدل آمیخته قابل بررسی و مطالعه است. لذا مقایسه تصادفی مؤلفه‌های بر جای مانده از دو سیستم متفاوت در حالت چند متغیره نیز از جمله مسائلی است که در فصل‌های بعدی به آن پرداخته می‌شود. برای مطالعه بیشتر می‌توان به بایراموو و آرنولد[۵] (۲۰۰۸) مراجعه کرد. همچنین در ادامه همین فصل خواهیم دید که توزیع متغیرهای همراه نیز می‌تواند به عنوان یک مدل آمیخته معرفی شود و لذا از این حیث به مطالعه برخی خواص آن‌ها و مقایسه تصادفی آن‌ها در مسائل تک نمونه‌ای و دو نمونه‌ای بر حسب ترتیب‌های تصادفی مختلف، پرداخته می‌شود و در ادامه فصل به تعاریف و مفاهیم مورد نیاز در این رساله می‌پردازیم. در ابتدا در بخش ۱-۲ مدل‌های آمیخته را معرفی می‌کنیم و توزیع مؤلفه‌های بر جای مانده از یک آزمایش طول عمر در بخش ۱-۳ مورد مطالعه و بررسی قرار می‌دهیم. در بخش ۱-۴ متغیرهای همراه آماره‌های مرتب معرفی و به برخی از کاربردهای آن‌ها می پردازیم، سپس برخی از توابع مهم

[۲] Khaledi and Shaked
[۳] Belzunce
[۴] Misra and Misra
[۵] Bairamov and Arnold

در تئوری قابلیت اعتماد و طول عمر را در بخش ۱-۵ بیان می‌کنیم. بررسی و مطالعه برخی از ترتیب‌های تصادفی و روابط میان آن‌ها در حالت تک متغیره و چند متغیره و انواع مفاهیم وابستگی مورد نیاز در این رساله بترتیب در بخش های ۱-۶ و ۱-۷ از جمله مفاهیمی هستند که در ادامه این فصل به آن‌ها پرداخته خواهد شد.

۲.۱ مدل‌های آمیخته

در چند دهه ی اخیر مقایسه تصادفی و وابستگی بردارهایی از متغیرهای تصادفی وابسته با نمونه هایی از تک جامعه و دو جامعه مورد توجه بسیاری از آمار دانان قرار گرفت. خالدی و شیکد (۲۰۱۰) به مقایسه تصادفی توزیع‌های آمیخته چندگانه پرداختند و نتایجی در ترتیب‌های تصادفی چند متغیره برای نمونه هایی تصادفی از مدل آمیخته اثبات کردند. میشرا و میشرا (۲۰۱۲) نیز خواص تصادفی استقلال شرطی مدل‌های آمیخته را مورد مطالعه قرار دادند و برخی مقایسه‌های تصادفی چند متغیره و برخی از خواص سالخوردگی را برای این مدل از توزیع‌ها اثبات کردند. همانطور که قبلاً گفتیم مدل‌های زیادی در آمار وجود دارند که منطبق بر مدل‌های آمیخته هستند. در ادامه همین فصل در بخش های بعدی دو مدل از توزیع‌ها را معرفی می‌کنیم که هر کدام نوع خاصی از مدل‌های آمیخته هستند. لذا مطالعه و بررسی این مدل‌ها می‌تواند حائز اهمیت باشد. در ادامه به معرفی مدل‌های آمیخته چند گانه می پردازیم.

فرض کنید برای $i = 1, 2$، $\{F_i(\mathbf{x}|\theta)\ \theta \in \mathcal{X}\}$ خانواده ای از تابع توزیع‌های n-بعدی باشد که در آن فضای پارامتر تعریف شده، $\mathcal{X}$، زیر مجموعه‌ای از $\mathcal{R}$ است. برای هر $\theta \in \mathcal{X}$ فرض کنید برای $i = 1, 2$، $\mathbf{X}_i(\theta) = (X_{i,1}(\theta), ..., X_{i,n}(\theta))$ برداری تصادفی از متغیرها با تابع توزیع $F_i(\mathbf{x}|\theta)$ باشد. همچنین فرض کنید Θ_i متغیری تصادفی با تابع توزیع تجمعی H_i که متعلق به $\mathcal{X}$ زیر مجموعه ای از $\mathcal{R}$ باشد. آنگاه بردار تصادفی $\mathbf{X}_i(\Theta) = (X_{i,1}(\Theta), ..., X_{i,n}(\Theta))$ دارای تابع توزیعی خواهد بود که آمیخته‌ای از تابع توزیع $F_i(\mathbf{x}|\theta)$ است. حال فرض کنید که تابع توزیع بردار تصادفی $\mathbf{X}_i(\Theta)$ به شرط $\Theta = \theta$ بصورت زیر باشد

$$\mathbf{F}_i(\mathbf{x}|\theta) \equiv \mathbf{F}_i(x_1, ..., x_n|\theta) = \prod_{j=1}^{n} F_{i,j}(x_j|\theta),$$

که در آن $F_{i,j}(x_j|\theta)$ برای هر $j \in \{1, 2, ..., n\}$ و هر $\theta \in \mathcal{X}$، تابع توزیع تجمعی متغیر تصادفی $X_{i,j}(\theta)$ است. به عبارتی فرض می‌کنیم متغیرهای تصادفی $X_n, ..., X_1$ به شرط متغیر تصادفی Θ از هم مستقل باشند. بنابراین تابع توزیع تجمعی بردار تصادفی $\mathbf{X}_i(\Theta_i)$ بصورت زیر ارائه می‌شود

$$G_i(\mathbf{x}) = \int_{\mathcal{X}} \prod_{j=1}^{n} F_{i,j}(x_j|\theta) dH_i(\theta), \qquad (۱.۱)$$

مقایسه تصادفی دو بردار تصادفی $\mathbf{X}_1(\Theta_1)$ و $\mathbf{X}_1(\Theta_2)$ که در آن Θ_1 و Θ_2 دو متغیر تصادفی با توابع توزیع متفاوت هستند، مورد مطالعه مؤلفین قرار گرفته است و نتایج بسیاری دربارهٔ مقایسه‌های تصادفی آن‌ها بر طبق نظریه‌های ترتیب‌های تصادفی چند متغیره در میشرا و میشرا (۲۰۱۲)، خالدی و شیکد (۲۰۱۰)، بلزونسه و همکاران (۲۰۰۹)، لی و دا[۶] (۲۰۱۰)، گوپتا[۷] و همکاران (۲۰۱۱)، ناوارو[۸] (۲۰۰۸)، گاپتا و گاپتا (۲۰۰۹) و لی و زاو[۹] (۲۰۱۱) وجود دارد. مقایسه تصادفی بردارهای تصادفی $\mathbf{X}_1(\Theta)$ و $\mathbf{X}_2(\Theta)$ که تاکنون مورد مطالعه قرار نگرفته است در فصل دوم با استفاده از ترتیب‌های تصادفی چند متغیره نسبت درستنمایی، نرخ خطر و نرخ خطر معکوس که تعاریف آن‌ها در بخش ۱.۶ آورده شده است، مطالعه می‌شود.

۳.۱ مؤلفه‌های بر جای مانده از یک آزمایش طول عمر

فرض کنید متغیرهای تصادفی $X_1, ..., X_n$، معرف طول عمر مؤلفه‌های یک سیستم منسجم باشند. یک سیستم $n-k+1$ از n تا زمانی که k مولفه از آن فعال باشد به فعالیت خود ادامه می دهد، لذا طول عمر این سیستم برابر با $X_{k:n}$، kامین آماره مرتب است، به عبارت دیگر سیستم با مشاهده kامین مؤلفه از کار افتاده، از کار می‌افتد. در تئوری قابلیت اعتماد معمولاً فرض بر این است که مؤلفه‌های سیستم از هم مستقل و هم توزیع باشند. در چنین سیستم‌هایی خراب شدن یک مؤلفه باعث افزایش فشار و کار اضافی بر مؤلفه‌های بر جای مانده نمی‌شود. مسئله ای که می‌تواند حائز اهمیت باشد این است که ممکن است مؤلفه‌های سیستم گران قیمت باشند و مقرون به صرفه باشد که وقتی سیستم شکست بخورد بتوان از مؤلفه‌های سالم بر جای مانده برای ساختن سیستم‌های دیگر استفاده کرد. برای مثال اگر بدانیم که طول عمر مؤلفه‌های سیستم از مدل نمایی تبعیت می کنند، می‌توانیم نتیجه بگیریم که مؤلفه‌های بر جای مانده همانند مؤلفه‌های اولیه کار خواهند کرد. مسلماً استفاده مجدد از آن‌ها مناسب خواهد بود. لذا مشخص کردن توزیع مؤلفه‌های بر جای مانده حائز اهمیت است. در زندگی روزمره ممکن است این امکان بوجود آید که مؤلفه‌های کهنه بهتر از مؤلفه‌های نو عمل کنند. لذا می‌توان از مطلوبیت بکار گرفتن مؤلفه ای کهنه در سیستم‌های دیگر صحبت کرد و نتیجه گرفت آیا می‌توان از این مؤلفه ها بطور بهینه استفاده کرد یا خیر؟ بنابراین با توجه به اهمیت موضوع در ادامه به مطالعه نظریه توزیعی مؤلفه‌های بر جای مانده از یک سیستم می‌پردازیم.

یک سیستم $n-k+1$ از n را با مؤلفه‌های مستقل از هم و هم توزیع در نظر بگیرید، سیستم با مشاهده ی k-امین شکست و به عبارتی با مشاهده ی آمارهٔ مرتب k-ام از کار خواهد افتاد. برای هر $k \in \{1, 2, ..., n\}$ مؤلفه‌های بر جای مانده از این سیستم را با $X_1^{(k)}, ..., X_{n-k}^{(k)}$ نمایش می‌دهیم

[۶] Li and Da
[۷] Gupta
[۸] Navaro
[۹] Li and Zhao

که در زمان k-امین شکست هنوز فعال بوده اند. در ادامه، توزیع توام و حاشیه ای باقیمانده‌ی عمر مؤلفه‌های بر جای مانده و برخی خواص آن‌ها را مطالعه می‌کنیم.

فرض کنید $X_1, ..., X_n$ متغیرهای تصادفی مستقل و هم توزیع با تابع توزیع تجمعی F و تابع چگالی احتمال f باشند. اگر قرار دهیم $X_{k:n} = t$ آنگاه می‌توان گفت که توزیع شرطی آماره‌های مرتب $X_{k+1:n}, ..., X_{n:n}$ هم توزیع با آماره‌های مرتب $X^{(k)}_{1:n-k}, \ldots, X^{(k)}_{n-k:n-k}$ متناظر با نمونه تصادفی با اندازه $n-k$ از متغیر تصادفی مانده عمر $X_t = [X - t|X > t]$ است. جمالاما داکا[10] و تاوفر[11] (۲۰۰۲)، با استفاده از این نمونه تصادفی توانستند آزمونی برای تست توزیع نمایی بدست آورند. حال اگر فرض کنیم $Y^{(k)}_1, ..., Y^{(k)}_{n-k}$ متغیر های مستقل و هم توزیع با تابع بقای مشترک $\frac{\overline{F}(y)}{\overline{F}(t)}$ باشند، آنگاه مؤلفه‌های مانده عمر بعد از k-امین شکست را می‌توان بصورت زیر نمایش داد

$$X^{(k)}_i \stackrel{\mathrm{d}}{=} Y^{(k)}_i - X_{k:n},\ i = 1, ..., n-k,$$

که درآن $\stackrel{\mathrm{d}}{=}$ به معنای هم توزیعی دو متغیر تصادفی است. تابع توزیع تجمعی $X_{k:n}$، را با $F_{k:n}$ نمایش می‌دهیم. تابع بقای توام متغیرهای تصادفی بر جای مانده بصورت

$$\begin{aligned}
\overline{F}^{(k)}(x_1, ..., x_{n-k}) &= P(X^{(k)}_1 > x_1, ..., X^{(k)}_{n-k} > x_{n-k}) \\
&= \int_0^\infty P(X^{(k)}_1 > x_1, ..., X^{(k)}_{n-k} > x_{n-k}|X_{k:n} = t)dF_{k:n}(t) \\
&= \int_0^\infty P(Y^{(k)}_1 > x_1 + t, ..., Y^{(k)}_{n-k} > x_{n-k} + t|X_{k:n} = t)dF_{k:n}(t) \\
&= \int_0^\infty \prod_{j=1}^{n-k} \frac{\overline{F}(x_j + t)}{\overline{F}(t)} dF_{k:n}(t) \\
&= c_{k,n} \int_0^\infty \prod_{j=1}^{n-k} \overline{F}(x_j + t) dF_{k:k}(t) \qquad (۲.۱)
\end{aligned}$$

است که در آن $c_{k,n} = k\binom{n}{k}$، و $F_{k:k}$ تابع توزیع آماره‌ی مرتب k-ام از یک نمونه تصادفی به حجم k است.

با توجه به معرفی مدل‌های آمیخته در بخش ۲.۱، کاملاً مشهود است که مدل فوق را می‌توان مدلی آمیخته در نظر گرفت.

در صورتی که F مطلقاً پیوسته باشد با مشتق گیری از رابطه‌ی فوق نسبت $x_1, ..., x_{n-k}$ تابع چگالی توام متغیرهای تصادفی مانده عمر بصورت

$$f^{(k)}(x_1, ..., x_{n-k}) = \int_0^\infty \prod_{j=1}^{n-k} \frac{f(x_j + t)}{\overline{F}(t)} dF_{k:n}(t) \qquad (۳.۱)$$

[10] Jammalamadaka
[11] Taufer