بسم الله الرحمن الرحیم

دانشگاه اصفهان

دانشکده علوم

گروه آمار

پایان نامه ی کارشناسی ارشد رشته‌ی آمار گرایش ریاضی

توزیع بیرنبام–ساندرز و تعمیم‌های آن

استاد راهنما:

دکتر محمد حسین علامت ساز

پژوهشگر:

زینب آقابزاز

آذر ماه ۱۳۹۱

کلیه حقوق مادی مترتب بر نتایج مطالعات، ابتکارات و نوآوری های ناشی از تحقیق موضوع این پایان نامه متعلق به دانشگاه اصفهان است.

دانشگاه اصفهان

دانشکده علوم

گروه آمار

پایان‌نامه‌ی کارشناسی ارشد رشته‌ی آمارگرایش آمار ریاضی

خانم زینب آقابزاز

تحت عنوان

توزیع بیر نبام – ساندرز و تعمیم های آن

در تاریخ ۹۱/۹/۲۷ توسط هیأت داوران زیر بررسی با درجه **بسیار خوب** به تصویب نهایی رسید.

۱- استاد راهنمای پایان‌نامه دکتر محمد حسین علامت ساز با مرتبه‌ی علمی استاد امضاء

۲- استاد داور داخل گروه پایان‌نامه دکتر محمد بهرامی با مرتبه‌ی علمی استادیار امضاء

۳- استاد داور داخل گروه پایان نامه دکتر حمید بیدرام با مرتبه‌ی علمی استادیار امضاء

مدیر گروه

تقدیم به

پدر و مادر مهربانم

تقدیر و تشکر

حمد و سپاس مخصوص ذات حق تعالی است که هرچه هست از اوست.

خداوند را شاکرم که به بنده ی کمترین خود نعمت سلامتی و کسب علم را عطا فرمود و در مسیر بندگان خاصش که همان تعلیم و تعلم است هدایت نمود. لازم می‌دانم از تمامی کسانی که در رسیدن به این هدف عالی کمک ها و راهنمایی های لازم را نموده اند و همدل و همراهم بودند، تشکر نمایم. خصوصاً از زحمات بی دریغ استاد گرانقدر جناب آقای دکتر محمد حسین علامت ساز که در تدوین این پایان‌نامه راهنمایم بوده‌اند، به طور ویژه تقدیر و تشکر می نمایم. همچنین از خانواده عزیزم که همواره در تمامی مراحل پشتیبانم بوده‌اند و مرا در رسیدن به این هدف یاری نموده‌اند نهایت تشکر و قدردانی را دارم.

زینب آقابزاز

پاییز ۹

چکیده

یکی از جنبه‌های بسیار مهم تحلیلهای پارامتری طول عمر، تعیین توزیع طول عمر مناسب است. بیرنبام و ساندرز (1958) علاقهمند به یافتن مدلی برای توصیف زمان فرسودگی مواد بودند که ارتباط بین نمونه مواد و زمان فرسودگی آن را نشان دهد. آن‌ها مقاله‌ای تحت عنوان "خانواده جدیدی از توزیعهای طول عمر" ارائه کردند و به همراهی گروهی از آماردانان مانند اساری و همکاران (1973) تحقیقات خود را در مورد این مدل ادامه دادند که منجر به معرفی و بررسی ویژگی‌های توزیع بیرنبام-ساندرز گردید. این توزیع طول عمر، توزیعی دو پارامتری بر اساس تحلیلهای فیزیکی بود که از قضیه تجدید نشأت میگرفت. ایده این توزیع از تعداد سیکل‌های لازم برای فرسودگی ناشی از رشد ترکهای مواد به وجود آمد و مبنای آن استدلال‌های فیزیکی برای خسارات تجمعی بود که باعث ایجاد فرسودگی مواد می شد . درحقیقت آن‌ها تعبیراحتمالی آن را بهدست آوردند.

دزموند (1985) این توجیه فیزیکی را با ساده‌تر کردن فوضهای اولیه که توسط بیرنبام و ساندرز مطرح شده بود، تحکیم بخشید و به‌دنبال آن تحقیقات توسط جانسون و همکاران (1995) و دیگر محققین ادامه یافت. توزیع فرسودگی عمری که توسط آن‌ها مطرح شد مبتنی بر کل زمان و خسارات تجمعی وارد شده بر سیستم بود. این خسارات باعث افزایش دامنه ترکها و کاهش آستانه تحمل مواد و در نهایت تخریب و فروپاشی آن‌ها می‌شد. درحقیقت، عمر فرسودگی یکی از دلایل اصلی تخریب فلزات و بتنهای ساختمانی و یا در مواد مدرن فیبرهای کربنی به حساب می‌آید. چندین توزیع آماری برای توصیف دادههای طول عمر فرسودگی بهکار می‌رود اگرچه توافق نظری در رابطه با تأثیرگذاری بیشتر آنها در تحلیل دادههای فرسودگی وجود ندارد. از جمله آن‌ها توزیعهای گاما، لگ نرمال، گوسین معکوس و توزیع بیرنبام- ساندرز است.

در این پایان‌نامه سعی داریم به نحوه پیدایش توزیع بیرنبام- ساندرز و معرفی این توزیع بپردازیم. برخی از خواص مهم توزیع مانند تحلیلهای طول عمر و رفتار تابع نرخ خطر را بررسی میکنیم. گشتاورهای توزیع بیرنبام- ساندرز را مطالعه خواهیم کرد و برآورد پارامترهای آن را مطرح میکنیم. همچنین این توزیع را با دیگر توزیعهایی که برای توصیف خسارات تجمعی بهکار میرود، مقایسه می‌کنیم. توزیع بیرنبام-ساندرز بریده شده را معرفیکرده وکاربردهای آن را در ریسک‌های مالی بیان میکنیم سپس با استفاده از توزیع‌های بیضوی تراز و چوله-بیضوی تراز به معرفی تعمیمهای مختلف توزیع بیرنبام-ساندرز میپردازیم و رفتارهای توزیع را در حالات خاص بررسی می کنیم. همچنین حالت دو متغیره و چند متغیره این توزیع را معرفی کرده و به استنباطهایی درخصوص پارامترهای آنها میپردازیم.

واژه‌های کلیدی: تابع خطر، توزیع طول عمر، توزیع‌های بیضوی تراز، توزیع‌های چوله، عمر فرسودگی

فهرست مطالب

عنوان **صفحه**

عنوان **صفحه**

فصل سوم: توزیع BS بریده شده

عنوان **صفحه**

عنوان **صفحه**

فهرست شکل‌ها

عنوان **صفحه**

فهرست جدول‌ها

عنوان **صفحه**

کوتاه نوشت‌ها

ACI	Approximate confidence interval
AIC	Akaike information criterian
a.s	Almost surely
BS	Birnbaum-Saunders
BFR	Bathtub failure rate
BVBS	Bivariate Birnbaum-Saunders
Cdf	Cumulative distribution function
CI	Confidence interval
CK	Coefficient of kurtosis
CS	Coefficient of skewness
CV	Coefficient of variation
DF	Decreasing failure rate
DFRA	Decreasing failure rate average
DMRL	Decreasing mean residual life
DRL	Decreasing residual life
EC	Elliptical contoured
GBS	Generalized Birnbaum-Saunders
GBS''	Doubly generalized Birnbaum-Saunders
IBFR	Inverse bathtub failure rate
IFR	Increasing failure rate
IFRA	Increasing failure rate average
IG	Inverse Gaussian
iid	Independent and identically distribution
IMRL	Increasing mean residual life
IRL	Increasing residual life
JMLE	Jackknife maximum likelihood estimate
JMME	Jackknife modified moment estimate
KT	Kotz type
KS	Kolmogorov-Smirnov
LIII	Type-III generalized logistic
LN	Lognormal
ML	Maximum likelihood
MM	Modified moment
NBU	New better than used
NBUE	New better than used in expectation

NWU	New worse than used
NWUE	New worse than used in expectation
Pdf	Probability density function
PE	Power exponential
Psi	Pounds per square inch
PVII	Pearson type VII
r.v	Random variable
SBVBS	Skewed bivariate Birnbaum-Saunders
SD	Standard deviation
SEC	Skew-elliptical contoured
SHN	Sinh-Normal
SIC	Schwarz information criterian
SMVBS	Skewed multivariate Birnbaum-Saunders
SN	Skew-Normal
SN-BS	Skew-normal- Birnbaum-Saunders
TBS	Truncated Birnbaum-Saunders
TVaR	Tail value at risk
VaR	Value at risk

فصل اول

مفاهیم و مقدمات اولیه

برای درک بهتر مطالب عنوان شده در این پایان‌نامه، لازم است ابتدا مفاهیم اولیه ضروری را شرح دهیم. به‌این منظور مختصری در مورد توزیع‌های طول عمر، مفاهیم قابلیت اعتماد و نحوه انتخاب توزیع‌های طول عمر مطالبی را شرح می‌دهیم. از آنجایی‌که توزیع بیرنبام-ساندرز از جمله توزیع‌های عمر فرسودگی محسوب می‌شود، مفاهیم فرسودگی، توزیع‌های عمر فرسودگی و سیر تکاملی توزیع بیرنبام-ساندرز نیز در این فصل مطرح می‌شود.

1-1 توزیع‌های طول عمر

به‌طورکلی، تحقیقات مرتبط با روش‌های آماری و مدل سازی‌های داده‌های طول عمر این تصور را به‌وجود می‌آورد که متغیر طول عمر[1]، متغیر تصادفی پیوسته و مثبتی است که زمان سپری شده تا وقوع پیشامد مورد نظر را نشان می‌دهد. از آنجایی‌که سالخوردگی اقلام [2] (به‌عنوان مثال، مؤلفه‌ها، سیستم‌ها مواد، ساختارها، ارگان‌ها، واحدها و...) همیشه در قالب زمان قابل اندازه‌گیری نیستند، در برخی موارد طول عمر با تعابیر دیگری از متغیرها اندازه‌گیری می‌شوند. برای مثال، تعداد کیلومترهای طی شده، دوام نمونه مواد تا فروپاشی آن، سطح فرسودگی، انعطاف‌پذیری مواد چسبنده و تعداد دوره‌هایی که مواد در طی آن فرسوده می‌شوند. به‌علاوه اصطلاح متغیر طول عمر به هر متغیر تصادفی مثبت نیز اتلاق می‌شود (مانند میزان بارش باران). مدل‌های آماری مرتبط با متغیرهای طول عمر، توزیع‌های طول عمر نامیده می‌شوند. برای جزئیات بیشتر در مورد توزیع‌های طول عمر به مارشال و الکین[3] (2007) مراجعه کنید.

---

[1] -Lifetime
[2] - Aging of items
[3] -Marshall and Olkin

نظریه قابلیت اعتماد مسائل آماری و احتمالی توزیع‌های طول عمر اقلام شکست [1] را مورد بحث قرار می‌دهد. روش‌های احتمالی این نظریه کارایی و از کارافتادگی مؤلفه‌ها را در قالب متغیرهای طول عمر مدل‌سازی شده توصیف می‌کند و به این ترتیب می‌توان تابع بقا، نرخ خطر و ... مؤلفه‌ها را تعیین کرد.

جنبه آماری این نظریه در حل مسائل برآورد پارامترهای توزیع‌های طول عمر است. قابلیت اعتماد یا بقا احتمالی است که متغیر طول عمر $T > 0$ از زمان ثابت t بیشتر شود. به‌عبارتی قابلیت اعتماد مؤلفه‌ها برحسب زمان t به‌صورت $R_T(t) = \bar{F}_T(t) = P(T>t)$ به‌ازای $t \geq 0$، تعریف می‌شود.

2-1 تحلیل‌های طول عمر

در این بخش مفاهیم قابلیت اعتماد متغیرهای طول عمر بیان می‌شود.

1-2-1 تابع چگالی احتمال

فرض کنید $f_T(.)$ و $F_T(.)$ به ترتیب تابع چگالی احتمال(pdf) و تابع توزیع تجمعی(cdf) متغیر تصادفی T باشند. در این صورت داریم

$$F_T(t) = P(T \leq t) = \int_0^t f_T(x)dx \quad ; \quad t > 0,$$

که در آن $f_T(t)$ را می‌توان به عنوان تغییرات لحظه‌ای احتمال شکست در بازه $[t, t + \Delta t]$ تعبیر کرد. زیرا داریم

$$f_T(t) = \frac{dF_T(t)}{dt} = \lim_{\Delta t \to 0} \frac{F_T(t+\Delta t) - F_T(t)}{\Delta t} = \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t+\Delta t)}{\Delta t} \quad ; \quad t > 0.$$

موضوع بااهمیت دیگر مرتبط با تحلیل‌های طول عمر، چندک q-ام توزیع طول عمر است که با t_q، به‌ازای $0 < q < 1$، نشان داده می‌شود. به عبارتی t_q وارون تابع توزیع تجمعی است و بیانگر زمانی است که $q \times 100\%$ جامعه شکست خورده است. چندک q-ام $F_T(.)$ کوچکترین طول عمری است که $P(T \leq t_q) \geq q$ باشد. $F_T(t)$ برحسب t غیرنزولی است در حالی که با افزایش t، t_q تنها مقدار ثابت $F_T(t_q) = q$ را اختیار می‌کند. این مقدار به‌صورت $t_q = F_T^{-1}(q)$ به‌دست می‌آید. زمانی‌که تابع توزیع در برخی فواصل ثابت باشد ممکن است بیش از یک جواب برای $F_T(t) = q$ وجود داشته باشد. در این حالت، t_q کمترین مقدار t است که $F_T(t) = q$ باشد.

---

[1] -Failure

2-2-1 تابع قابلیت اعتماد

قابلیت اعتماد[۱] متغیرتصادفی T ، احتمالی است که T از مقدار t بیشتر باشد. در این صورت داریم

$$R_T(t) = P(T > t) = 1 - F_T(t) = \bar{F}_T(t) = \int_t^\infty f_T(x)dx \quad ; \quad t > 0 , \qquad (۱ ـ۱)$$

به‌علاوه، $R_T(0) = 1$ و $R_T(t)$ تابعی پیوسته و غیر صعودی است.

تابع قابلیت اعتماد شرطی در زمان $(t + x)$ به شرطی که مؤلفه‌ها تا زمان x کار کنند به‌صورت

$$R_T(t \mid x) = \frac{R_T(t+x)}{R_T(x)} \quad ; \quad t > 0 , x > 0 , \qquad (۱ ـ۲)$$

است، که در آن $0 < R_T(.) < 1$.

3-2-1 تابع نرخ خطر

علاوه بر تابع بقا یا قابلیت اعتماد، شاخص‌های دیگری ازجمله تابع نرخ خطر [۲] در تحلیل‌های قابلیت اعتماد به‌کارمی‌رود. این شاخص تابع شانس[۳] ، نرخ شکست[۴] ، شدت مرگ و میر[۵] یا نرخ ریسک نیز نامیده می‌شود. جزئیات بیشتر در این مورد در کتاب مارشال و الکین(2007) بیان شده‌است.

نرخ خطر اهمیت ویژه‌ای در قابلیت اعتماد دارد چراکه در مهندسی، اطلاعات زیادی در خصوص فرسودگی مؤلفه‌ها در اختیار ما قرار می‌دهد. از طرف دیگر، در علوم پزشکی نیز بسیاری از نتیجه‌گیری‌های بقا براساس نرخ خطر صورت می‌گیرد. نرخ خطر به‌عنوان نرخ خطر آنی یا تغییرات شکست در فاصله زمانی t تا $t + \Delta t$ ، زمانی که مؤلفه‌ها تا لحظه t کارمی‌کند، تعریف می‌شود یعنی

$$h_T(t) = \lim_{\Delta t \to 0} \frac{P(t \le T \le t + \Delta t \mid T > t)}{\Delta t} = \frac{f_T(t)}{R_T(t)} \quad ; \quad t > 0 , \qquad (۱ ـ۳)$$

به‌این ترتیب $h_T(.) \ge 0$ و $\int_0^\infty h_T(t)dt = \infty$.

اگر $\Delta t > 0$ ناچیز باشد، $h_T(t) \times \Delta t$ به‌طور تقریبی احتمال ازکارافتادگی در بازه $[t, t + \Delta t]$ را با فرض اینکه مؤلفه‌ها تا زمان t در حال کارکردن باشد نشان می‌دهد. به‌عبارتی داریم

$$h_T(t) \times \Delta t \approx P(t \le T \le t + \Delta t \mid T > t) \quad ; \quad t > 0.$$

---

[1] -Reliability
[2] -Hazazrd rate function
[3] -Chance function
[4] -Failure rate
[5] -Force of mortality

اگر نرخ خطر معلوم باشد توابع چگالی، توزیع و قابلیت اعتماد با فرض داشتن یکی از آن‌ها به‌دست می‌آید. اهمیت این موضوع به‌این خاطر است که رفتار نرخ خطر به‌طور تجربی در مطالعات معلوم است. به این ترتیب دیگر توابع ذکر شده از نرخ خطر به‌دست می‌آیند.

4-2-1 تابع خطر [1]

می‌دانیم با انتگرال‌گیری از تابع چگالی احتمال، تابع توزیع تجمعی به‌دست می‌آید. به‌طور مشابه می‌توان با انتگرال‌گیری از نرخ خطر، نرخ خطر تجمعی یا تابع خطر را به‌دست‌آورد. نرخ خطر تجمعی متغیر تصادفی T به‌صورت زیر به‌دست می‌آید

$$H_T(t) = \int_0^t h_T(u)du = -\log\big(R_T(t)\big) \quad ; \quad t > 0.$$

با استفاده از نرخ خطر تجمعی می‌توان رفتار نرخ خطر را در بازه زمانی ثابت $[0, t]$ نمایش داد. بنابراین میانگین نرخ خطر [2] (FRA) متغیر تصادفی T به‌صورت زیر به‌دست می‌آید

$$FRA(t) = \frac{H_T(t)}{t} = \frac{-\log\big(R_T(t)\big)}{t} \quad ; \quad t > 0\,, \qquad (۱ ۴)$$

به ازای $0 < R_T(.) < 1$.

5-2-1 میانگین باقیمانده عمر

میانگین زمان تا وقوع شکست، میانگین طول عمر نامیده می‌شود و به‌صورت زیر به‌دست می‌آید

$$\mu = E(T) = \int_0^\infty t\, f_T(t)dt = \int_0^\infty R_T(t)dt\,, \qquad (۱ ۵)$$

میانگین باقیمانده عمر برحسب تابع قابلیت اعتماد تعریف می‌شود. میانگین باقیمانده عمر، متوسط زمانی است که مؤلفه پس از زمان x فعال است مشروط برآنکه تازمان x سالم باشد. یعنی

$$\mu_x = E(T - x \mid T > x) = \int_0^\infty R_T(t \mid x)dt = \frac{\int_x^\infty R_T(u)du}{R_T(x)} \quad ; \quad x > 0\,,$$

به ازای $0 < R_T(.) < 1$.

---
[1] - Failure function
[2] -Failure rate average