بسم الله الرحمن الرحیم

دانشگاه شیراز

دانشکده علوم

پایان‌نامه کارشناسی ارشد در رشته‌ی آمار

استنباطهایی درباره توزیعهای دمی-وایبل

توسط:

اکرم فخاری اسفریزی

استاد راهنما:

دکتر ناهید سنجری فارسی پور

شهریور ماه ۱۳۸۶

به نام خدا

استنباطهایی درباره توزیعهای دمی- وایبل

به وسیله‌ی :

اکرم فخاری اسفریزی

پایان‌نامه

ارائه‌شده به تحصیلات تکمیلی دانشگاه به‌عنوان بخشی

از فعالیت های لازم برای اخذ درجه کارشناسی ارشد

در رشته‌ی:

آمار ریاضی

از دانشگاه شیراز

شیراز

جمهوری اسلامی ایران

ارزیابی شده توسط کمیته پایان‌نامه با درجه : عالی

دکتر ناهید سنجری فارسی پور، استاد بخش آمار (رئیس کمیته)..

دکتر کاووس خورشیدیان، استادیار بخش آمار..

دکتر مینا توحیدی، استادیار بخش آمار..

شهریورماه ۱۳۸۶

تقدیم به

پدر و مادر عزیزم که مرا در راه کسب علم و دانش و معرفت تشویق و ترغیب نموده وهمچنین از همسرم که از هیچ کوششی در این راه دریغ ننموده و همیشه سختیهای این راه را همراه من تحمل نموده و امید بخش من بوده است.

سپاسگزاری

سپاس خداوندی را که منبع همه علوم است. اکنون که به لطف خداوند بزرگ و متعال توانسته‌ام پله دیگری از نردبان علم بالاتر روم، وظیفه خود می‌دانم از تمامی اساتید، بزرگواران و دانشجویان عزیزی که در طول دوره این تحقیق مرا یاری فرموده‌اند، تشکر و قدردانی نمایم. به ویژه از استاد راهنمای عزیز خانم دکتر ناهید سنجری فارسی پورکه با دقت و ظرافت و با ارائه نقطه نظرات خود بر غنای این پایان نامه افزوده اند، سپاسگزارم. از اساتید گرامی خانم دکتر مینا توحیدی و آقای دکتر کاووس خورشیدیان که با کمال صبر و حوصله و دقت، پاسخگوی سؤالات و مشکلات بنده در این تحقیق بوده‌اند، کمال تشکر و قدردانی را دارم. همچنین از دانشجویان گرامی که مرا در طول دوران تحصیل یاری داده‌اند، کمال سپاسگزاری را دارم. برای تمام این عزیزان از ایزد یگانه موفقیت روزافزون را خواستارم.

چکیده

استنباطهایی درباره توزیعهای دمی – وایبل

به وسیله ی:

اکرم فخاری اسفریزی

توزیعهای دمی- وایبل از جمله توزیعهایی هستند که در امور بیمه و ورزش و قابلیت اعتماد مورد استفاده قرار می گیرند. دم این توزیعها به فرم نمایی نوشته می شود. ضریب دمی- وایبل به عنوان ضریب تغییرات با آهنگ منظم مربوط به معکوس تابع نرخ شکست تعریف می شود. برآوردگر این پارامتر براساس فضای لگاریتمی آماره های ترتیبی بالایی می باشد.

برآوردگر با اریبی کاهش یافته براساس مدل رگرسیون نمایی به دست می آید. یک مطالعه شبیه سازی برای مقایسه روشهایی که در بالا ذکر شد، فراهم شده است. برای انتخاب کردن کسر نمونه بهینه ، ما یک روش گرافیکی ارایه کرده ایم.

فهرست مطالب

فصل پنجم : روشهای انتخاب بهینه برای k_n

فهرست جدول ها

عنوان و شماره **صفحه**

فهرست شکل ها

عنوان **صفحه**

فصل اول :مقدمه

فصل اول :مقدمه

۱-۱- مقدمه

در بررسی پدیده های نهایی و رکوردها ، رفتار دم توزیع برای ما مهم است . اینکه احتمال در دمها بیشتر باشد (دنباله سنگین) یا دمها دیر به صفر همگرا شوند (دم بلند) یا اطلاعاتی از این قبیل می تواند سمت و سوی خاصی به تجزیه و تحلیل این پدیده ها بدهد.

در چنین بررسیهایی ، بدون در نظر گرفتن فرم پارامتری خاص ، فرمی کلی را برای دم توزیع در نظر می گیرند. مثلا فرم خطی (توزیعهای دمی- نمایی) یا فرم نمایی (توزیعهای دمی- وایبل) و ...

در این تحقیق ، با توزیعهایی که دم آنها به فرم نمایی نوشته شود و برای نمای آن یعنی تابع انباشتگی نرخ شکست یک فرم خاص در نظر گرفته شود ، کار می کنیم. در فرمی که برای معکوس تابع انباشتگی نرخ شکست در نظر می گیریم ، پارامتری وجود دارد که این پارامتر را ضریب دمی- وایبل می نامند ، همچنین توزیعهایی که در شرایط فوق صدق می کنند را توزیعهای دمی - وایبل می نامند. این گونه از توزیعها در امور بیمه ، امور ورزشی و مطالعات قابلیت اعتماد کاربرد دارند.

پس لازم است که پارامتر ضریب دمی- وایبل برآورد شود و خواص آن مورد بررسی قرار گیرد. برآوردگر این پارامتر، با توجه به اینکه دم توزیع بررسی می شود ، بر اساس آماره های ترتیبی نهایی صورت می گیرد یا به عبارتی از قسمت بالایی آماره های ترتیبی استفاده کرده و برآوردگر را به دست می آورد. پس از برآورد ، تحت فرضیاتی توزیع حدی نرمال و خاصیت سازگاری ضعیف برای این برآوردگر در قالب قضایایی مطرح می شود. در اثبات این قضایا از قضیه حد مرکزی ، قانون قوی اعداد بزرگ ، قانون ضعیف اعداد بزرگ ، نامساوی چبیشف و قضایای معروف دیگری استفاده شده است. با استفاده از این قضایا می توان میانگین مربعات خطای مجانبی برآوردگر را به دست آورد.

از جمله توزیعهایی که به عنوان توزیعهای دمی – وایبل می توان نام برد : توزیع نرمال (با پارامتر ضریب دمی- وایبل برابر ۰/۵) ، توزیع گاما (با پارامترضریب دمی- وایبل برابر ۱) ، توزیع وایبل، توزیع وایبل توسیع یافته[۱]، توزیع وایبل اصلاح یافته[۲]می باشند. وقتی علاقه مندیم در این توزیعها ، پدیده هایی که احتمالشان از یک سطح معین تجاوز می کند را بررسی کنیم با مبحث چندکهای نهایی مواجه می شویم. برای چندکهای نهایی توزیعهای دمی – وایبل برآوردگرهای مختلفی مطرح می شود. باز هم در اینجا با اثبات قضایایی نرمال بودن مجانبی و خاصیت سازگاری ضعیف آنها بررسی می شود.

۱-۲- مروری بر تحقیقات گذشته

یکی از اولین برآوردگرهایی که برای پارامتر ضریب دمی – وایبل پیشنهاد شد ، براساس مقادیر رکورد بود که این کار توسط برد[۳] در سال ۱۹۹۱ انجام گرفت. با استفاده از تابع چندک و قضیه مقدار میانگین و نیز آماره های ترتیبی نهایی برآوردگری برای ضریب دمی – وایبل توسط برونیاتوسکی[۴] در سال ۱۹۹۳ ارائه گردید.

با استفاده از تابع باقیمانده عمر ، برآوردگر دیگری در سال ۱۹۹۵ توسط بیرلنت[۵] و همکارانش پیشنهاد شد که این برآوردگر دارای توزیع مجانبی نرمال است. در این مقاله تابعی از ضریب دمی – وایبل به عنوان اندازه سنگینی دم توزیع مطرح شده است.

در سال ۲۰۰۴ برآوردگری براساس لگاریتم آماره های ترتیبی نهایی توسط گرارد[۶] ارائه شد . این برآوردگر خیلی شبیه به برآوردگری که هیل[۷] در سال ۱۹۷۵ برای برآورد شاخص مقدار نهایی در توزیعهای نوع پاراتو ارائه داد ، می باشد. گرارد مقایسه هایی را بین برآوردگر خود و بعضی برآوردگرهای دیگر از دیدگاه تئوری و از دیدگاه کاربردی با استفاده از شبیه سازی انجام داده است که تحت فرضیات خاصی برآوردگر گرارد به نتایج بهتری می انجامد.

در مقاله دیابلت[۸] و همکارانش (۲۰۰۶) برآوردگر بهتری برای پارامتر ضریب دمی – وایبل ارائه گردید. این برآوردگر ، برآوردگری با اریبی- کاهش یافته می باشد. برای به دست آوردن این برآوردگر از مدل رگرسیونی و برآوردهای حداقل مربعات استفاده می شود.

---

[۱] *Extended weibull*
[۲] Modified Weibull
[۳] Berred
[۴] Broniatowski
[۵] Beirlant
[۶] Girard
[۷] Hill
[۸] Diebolt

اینکه چه قسمتی از آماره های ترتیبی بالایی را انتخاب کنیم و براساس آن برآوردیابی انجام شود، می تواند در یافتن برآوردی بهتر تاثیر بسزایی داشته باشد. برای اینکه این انتخـاب را بـه صورت بهینه انجام دهیم راه حلهایی پیشنهاد مـی شـود کـه از آن جملـه مـی تـوان راه حـل گرافیکی را مطرح کرد. دراین روش از نمودار چندک[۹] استفاده مـی شـود و جـایی کـه نمـودار شروع به خطی شدن می کند را انتخاب بهینه می دانیم. روشهای دیگری نیز وجود دارند که در یکی از این روشها به ازای انتخابهای مختلف ، میانگین مربعات خطا را محاسبه می کنـیم ، آن انتخابی که کمترین میانگین مربعات خطا را می دهد انتخاب بهینه می باشد.

در این پایان نامه، سعی بر آن است که با استفاده از مقاله های موجود و همچنین کتابهـای معتبر مبحث توزیعهای دمی – وایبل به طور وسیعی بسط داده شود. بررسیهای انجام شده هم از دیدگاه استنباطی و هم از دیدگاه احتمالی می با شد که کـارایی مبحـث احتمـال در اثبـات قضایا بیشتر مشهود است .

برای بسط موضوع از دیدگاه کاربردی ، از توزیعهای دمـی – وایبـل شـبیه سـازی نمـوده و برآورد ها را انجام می دهیم .برآوردهای مختلفی از نظـر اریبـی ، واریـانس ، فواصـل اطمینـان تجربی و میانگین مربعات خطا مقایسه می شوند. بدین منظور از شبیه سازی و برنامه نویسی در SPLUS استفاده شده است .

در این پایان نامه در فصل دوم به معرفی توزیعهای دمی – وایبل و برآورد پارامتر آن می پردازیم.در فصل سوم به برآورد چندکهای نهایی توزیع دمی – وایبل می پردازیم. در فصل چهارم برآوردگرهای اریبی - کاهش یافته ضریب دمی – وایبل را بررسی می کنیم و نیز در فصل پنجم روشهای انتخاب بهینه k_n را بررسی می کنیم.

---

[۹] Q-Q Plot

فصل دوم: معرفی توزیعهای دمی – وایبل و برآورد پارامتر آن

فصل دوم: معرفی توزیعهای دمی – وایبل و برآورد پارامتر آن

۲-۱- مقدمه

اگر معکوس تابع انباشتگی نرخ شکست را در نظر بگیریم، در بعضی از توزیعها این تابع به فرم خاصی در خواهد آمد ، که در این فرم خاص، پارامتری به نام ضریب دمی- وایبل وجود دارد. به منظور تعریف ضریب دمی- وایبل و استفاده از آن به مفاهیم اولیه زیر نیازمندیم:

تعریف۲-۱: تابع با آهنگ تغییر تدریجی *(Slowly varing function)*

تابع ℓ را یک تابع با آهنگ تغییر تدریجی گوئیم اگر برای هر $\lambda > 0$ وقتی که $x \to \infty$ داشته باشیم:

$$\frac{\ell(\lambda x)}{\ell(x)} \to 1$$

تعریف۲-۲: تابع با آهنگ تغییر منظم با شاخص α

(Regularly varing function with index α)

تابع f را تابع با آهنگ تغییر منظم با شاخص α گوییم اگر برای هر $\lambda > 0$ وقتی که $x \to \infty$ داشته باشیم:

$$\frac{f(\lambda x)}{f(x)} \to \lambda^{\alpha}$$

که این را با $f \in R_\alpha$ نمایش می دهند.

نکته۲-۱: اگر $f \in R_\alpha$ و ℓ یک تابع با آهنگ تغییر تدریجی باشد ، آنگاه می توان f را بدین صورت نوشت:

$$f(x) = x^{\alpha}\ell(x)$$

نکته۲-۲: هر تابع با آهنگ تغییر تدریجی یک تابع با آهنگ تغییر منظم با شاخص صفر است یعنی $\ell \in R_0$.

فرض کنید $X_1, X_2, ..., X_n$ یک دنباله از متغیرهای تصادفی مستقل و هم توزیع با تابع توزیع F باشد.

تعریف۲-۳: ضریب دمی- وایبل $(Weibull\, Tail-Coefficent)$

وقتی که دم توزیع در روابط زیر صدق کند:

$$1-F(x)=e^{-H(x)} \qquad x \geq x_0 \geq 0 \tag{A.۱}$$

$$H^{\leftarrow}(t)=\inf\{\, x : H(x) \geq t\} = t^{\theta}\ell(t)$$

بطوریکه ℓ یک تابع با آهنگ تغییر تدریجی باشد، آنگاه چنین توزیعی که در روابط بالا صدق می کند، را توزیع دمی – وایبل گویند و پارامتر θ را ضریب دمی- وایبل گویند. (در اینجا علامت← نشان دهنده معکوس تابع است.)طبق نکته۲-۱، چون $H^{\leftarrow}(t)=t^{\theta}\ell(t)$ تابع $H^{\leftarrow}$یک تابع با آهنگ تغییر منظم با شاخص θ دربی نهایت است و می نویسیم $H^{\leftarrow}\in R_{\theta}$.

در (A.۱) تابع ℓ می تواند یک تابع ثابت باشد که در این حالت پارامتر θ ، ضریب دمی- وایبل، متناسب با پارامتر شکل توزیع وایبل می شود . توزیعهایی هم با ℓغیر ثابت وجود دارند مثل نرمال ، گاما و وایبل توسیع یافته که در (A.۱) صدق می کنند. چنین توزیعهایی در بیمه و مطالعات قابلیت اعتماد کاربرد دارند.

به عنوان مثال یک شرکت بیمه باید بتواند خودش را در مقابل پرداخت حق بیمه های بزرگ بیمه کند که یک نوع دوباره بیمه کردن است تا بتواند در مقابل چنین مواردی دچار ورشکستگی نشود که این سیاست را $ECOMOR$ گویند. در این سیاست شرکت بیمه K تا از بزرگترین خسارتهای یک سال را بیمه می کند، که شرکت بیمه روی مقدار K در ابتدای سال توافق می کند که مقدار K معمولاً کمتر از ۱۰ می باشد.

اولین مساله در مواجه با چنین توزیعهایی ، مساله برآورد پارامتر θ است. در مورد برآورد θ، یکی از اولین پژوهشها ، توسط برد در سال ۱۹۹۱ بود که یک برآوردگر بر اساس مقادیر رکورد را پیشنهاد کرد. روش دیگر برآورد، از k_n تای بالایی آماره های ترتیبی استفاده می کند یعنی برآورد بر اساس $X_{n-k_n+1,n} \leq \ldots \leq X_{n,n}$ که k_n یک دنباله از اعداد صحیح است به طوری که $1 \leq k_n < n$باشد.

برآوردگرهایی برای θ بر اساس این روش ، توسط بیرلنت در سال۱۹۹۵ ، برانیاتوسکی در سال ۱۹۹۳ ، کلاپلبرگ[۱] و ویلاسنور[۲] در سال ۱۹۹۳ ارائه شده اند.

در این فصل به ارائه برآوردگری جدید که توسط گرارد در سال (۲۰۰۴) پیشنهاد شده است می پردازیم و خواص مجانبی آن را مطرح کرده و سپس این برآوردگر را با بعضی برآوردگرها که در بالا ذکر شد ، مقایسه می کنیم که این مقایسه هم از دیدگاه تئوری و هم از دیدگاه کاربردی می باشد.

---

۱ Kluppelberg
۲ Villasenor

۲-۲- برآوردگر گرارد برای θ و خواص مجانبی آن

برآوردگری که برای θ ارائه می شود بدین صورت است:

$$\hat{\theta}^{G}_{n} = \frac{\sum_{i=1}^{k_n-1} (\log(X_{n-i+1,n}) - \log(X_{n-k_n+1,n}))}{\sum_{i=1}^{k_n-1} (\log_2(\frac{n}{i}) - \log_2(\frac{n}{k_n}))} \qquad (۱-۲-۲)$$

که

$$\log_2(t) = \log(\log(t)) \qquad ; t > 1$$

این برآوردگر به برآوردگر هیل (۱۹۷۵) برای توزیعهای نوع پاراتو شباهت بسیاری دارد.

$$\hat{\gamma}_n = \frac{1}{k_n - 1} \sum_{i=1}^{k_n-1} (\log(X_{n-i+1,n}) - \log(X_{n-k_n+1,n}))$$

که γ شاخص مقدار نهایی برای توزیعهای نوع پاراتو است بطوری که

$$\frac{f(\lambda x)}{f(x)} \to \lambda^{-\frac{1}{\gamma}} \quad as \ x \to \infty \ , \forall \lambda > 0$$

خواص مجانبی $\hat{\theta}^{G}_{n}$ در قالب دو قضیه مطرح شده اند:

قضیه۲-۱: فرض کنید (A.۱) برقرار باشد . اگر $k_n \to \infty$, $\frac{k_n}{n} \to 0$ آنگاه $\hat{\theta}_n^G \xrightarrow{P} \theta$.

برای مطرح کردن قضیه۲-۲ در مورد توزیع حدی $\hat{\theta}_n^G$ لازم است که یک شرط دیگری در مورد تابع ℓ داشته باشیم:

یک $\eta \le 0$ و یک تابع $b(x)$ وجود داشته باشند ، بطوریکه وقتی $x \to \infty$ آنگاه $b(x) \to 0$ همچنین برای هر $1 < A < \infty$ بطوری که برای هر $\lambda \in [1, A]$

$$\log(\frac{\ell(\lambda x)}{\ell(x)}) \sim b(x) K_\eta(\lambda) \qquad (A.۲)$$

$$K_\eta(\lambda) = \int_1^\lambda u^{\eta-1} du = \frac{1}{\eta}(\lambda^\eta - 1)$$

(منظور از ~ هم ارزی می باشد)

پارامتر $\eta \le 0$ نرخ همگرایی $\frac{\ell(\lambda x)}{\ell(x)}$ به۱ را نشان می دهد. هر چه η به 0 نزدیکتر باشد نرخ همگرایی آهسته تر است ونیز داریم که $|b| \in R_\eta$.

قضیه ۲-۲: فرض کنید (A.۱) و (A.۲) برقرار باشند. آنگاه

$$k_n^{\frac{1}{2}}(\hat{\theta}^{G}_{n} - \theta) \xrightarrow{d} N(0, \theta^2)$$

برای هر دنباله k_n که

$$k_n \to \infty \ , \ k_n^{\frac{1}{2}} b(\log(\frac{n}{k_n})) \to 0 \quad , \quad \frac{k_n^{\frac{1}{2}}}{\log(\frac{n}{k_n})} \to 0$$

نکته۲-۳: از شرط $\frac{k_n^{\frac{1}{2}}}{\log(\frac{n}{k_n})} \to 0$ می توان $\frac{k_n}{n} \to 0$ را نشان داد . یعنی همواره در قضیه ۲- ۲ شرایط قضیه۲-۱ برقرار خواهد بود.

اثبات: چون $\frac{k_n^{\frac{1}{2}}}{\log(\frac{n}{k_n})} \to 0$ و داریم $k_n \to \infty$. پس مخرج با سرعت بیشتری باید به سمت

∞برود پس $\log(\frac{n}{k_n}) \to \infty$. چون $\log$ یک تابع صعودی است پس $\frac{n}{k_n} \to \infty$ آنگاه داریم $\frac{k_n}{n} \to 0$

□

نکته ۲-۴: از قضیه ۲- ۲ ، میانگین مربعات خطای مجانبی $\hat{\theta}_n^G$ برابر است با

$$AMSE(\hat{\theta}_n^G) = \frac{\theta^2}{k_n}$$

توجه داریم که با افزایش k_n , $AMSE(\hat{\theta}_n^G)$ کاهش می یابد و در بی نهایت به سمت 0 میل می کند.

نکته۲-۵: از قضیه ۲-۲ می توان فاصله اطمینان مجانبی در سطح $1-\alpha$ برای پارامتر θ را به صورت زیر ارائه داد:

$$\left(\frac{\sqrt{k_n}\,\hat{\theta}_n^G}{\sqrt{k_n} + z_{\frac{\alpha}{2}}}, \frac{\sqrt{k_n}\,\hat{\theta}_n^G}{\sqrt{k_n} - z_{\frac{\alpha}{2}}}\right)$$

$z_{\frac{\alpha}{2}}$ را از جدول نرمال استاندارد به دست می آوریم.

۲-۳- مثالهایی برای توزیع دمی – وایبل

در این بخش بعضی از توزیعهایی که در شرایط (A.۱) و (A.۲) صدق می کنند را مطرح می

کنیم.

- توزیع نرمال $N(\mu,\sigma^2)$

$$H^{\leftarrow}(x)=x^{\frac{1}{2}}\ell(x)$$

$$\ell(x)=\sqrt{2}\ \sigma-\frac{\sigma}{2^{\frac{3}{2}}}\ \frac{\log x}{x}+0(\frac{1}{x})$$

$$\theta=\frac{1}{2}\ ,\ \eta=-1\ \ ,\ b(x)=\frac{\log x}{4x}$$

- توزیع گاما $\Gamma(\beta,\alpha)$

$$f(x)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}\ x^{\alpha-1}\ e^{-\beta x}$$

$$H^{\leftarrow}(x)=x\,\ell(x)$$

$$\ell(x)=\frac{1}{\beta}+\frac{\alpha-1}{\beta}\ \frac{\log x}{x}+0(\frac{1}{x})$$

$$\theta=1\ ,\ \eta=-1\ ,\ b(x)=(1-\alpha)\frac{\log x}{x}$$

- توزیع وایبل $W(\alpha,\lambda)$

$$f(x)=\frac{\alpha}{\lambda^{\alpha}}\ x^{\alpha-1}e^{-\frac{x^{\alpha}}{\lambda^{\alpha}}}$$

$$H^{\leftarrow}(x)=\lambda\ x^{\frac{1}{\alpha}}$$

$$\ell(x)=\lambda$$

$$\theta=\frac{1}{\alpha}\ ,\ \eta=-\infty\ \ ,\ b(x)=0$$

در اینجا اگر برآورد پارامتر θ را بخواهیم در واقع باید برآورد پارامتر شکل توزیع وایبل را پیدا کنیم.

با بررسی (A.۲) چون ℓ یک تابع ثابت است $\log(\frac{\ell(\lambda x)}{\ell(x)})=0$ پس $b(x)=0$ و مقدار $\eta=-\infty$ را قرار داد می کنیم.

× توزیع وایبل توسیع یافته

$$EW(\alpha,\beta)\quad ;\ \alpha\in(0,1),\beta\in\mathrm{R}$$

$$F(x)=1-r(x)\ \exp(-x^{\alpha})\qquad r\in R_{\beta}$$

$$H^{\leftarrow}(x)=x^{\frac{1}{\alpha}}\ \ell(x)$$

$$\ell(x)=1+\frac{\beta}{\alpha^{2}}\ \frac{\log x}{x}+0(\frac{1}{x})$$