



۱۳۷۹ / ۸ / ۱

بسم الله الرحمن الرحيم

برآوردهای مقیاس در خانواده‌های توزیع مکان-مقیاس متقاضی توسعه برآوردهای
برای نمونه‌های کامل و OUAE و BLUE سانسور شده

بوسیله

مهدی خدادادی‌پور

پایان نامه

ارائه شده به دانشکده تحصیلات تکمیلی بعنوان بخش

از فعالیت‌های تحصیلی لازم برای اخذ

درجه کارشناسی ارشد

در رشتہ

آمار ریاضی

• ۸۷۷۱

از

دانشگاه شیراز

شیراز، ایران

ارزیابی و تصویب شده توسط کمیته پایان نامه با درجه: بسیار خوب

امضاء اعضاء کمیته پایان نامه

دکتر ناهید سنجری فارسی‌پور، دانشیار بخش آمار دانشگاه شیراز (رئیس کمیته)

دکتر احمد رضا سلطانی استاد بخش آمار دانشگاه شیراز

دکتر فریبرز حیدری استادیار بخش آمار دانشگاه شیراز

آبان ماه ۱۳۷۸

۳۱۸۱۰

تقدیم به:

پدر و مادر عزیزم و همسر مهربانم که همواره در
طول زندگی مدیون بزرگواریها یاشان خواهم بود

سپاسگزاری

اینک که به باری پروردگار متعال پایاننامه خود را به اتمام رسانده‌ام بر عهده من است که از زحمات و رهنماهای همیشه راهنمایم، استاد ارجمند جناب خانم دکتر ناهید سنجری فارسی‌پور تشکر کنم که اگر نبود زحمات بیلریغ و بزرگوارانه ایشان بی‌شک اتمام این رساله میسر نمی‌گشت و شایسته است که از زحمات بی‌شایه اعضای محترم کمیته پایاننامه جناب آقای دکتر احمد رضا سلطانی و همچنین از آقای دکتر فریبرز حیدری سپاسگزاری نمایم.

همچنین در پایان از پدر و مادر عزیزم که اولین معلم زندگی‌ام هستند، همه آنانی که در تحصیل علم و دانش مرا مشوق بوده‌اند و راه را بر من هموار ساختند صمیمانه تشکر می‌کنم.

چکیده

برآورد پارامتر مقیاس در خانواده‌های توزیع مکان- مقیاس توسط برآوردهای (OUAE) و (BLUE) برای نمونه‌های کامل و سانسور شده

توسط

مهدی خدادادی پور

یک برآوردکننده جدیدی برای برآورد پارامتر مقیاس، توزیع‌های مکان - مقیاس متقارن (با فرض مشخص بودن پارامتر مکان) معرفی می‌گردد که یک برآوردکننده بهینه و بصورت یک ترکیب خطی از قدر مطلق آماره‌های ترتیبی می‌باشد. این برآوردکننده را برآوردکننده بهینه ناریب خطی (OUAE) می‌نامیم. با محاسبه گشتاورهای قدر مطلق از آماره‌های ترتیبی، کارآئی نسبی این برآوردگر جدید را نسبت به بهترین برآوردکننده ناریب خطی (BLUE) برای توزیعهای مستطیلی - نرمال - وایل دوگانه - نمائی دوگانه محاسبه می‌گردد. به طوری که نشان داده می‌شود که در بیشترین حالات (برای نمونه‌های مختلف) برآوردکننده جدید کارتر از برآوردکننده BLUE می‌باشد.

فهرست مطالب

عنوان	صفحة
فهرست جداول
فصل اول: مقدمه و تاریخچه	۱
آمارهای ترتیبی	۲
سانسور نوع اول	۴
سانسور نوع دوم	۵
مدل‌های خطی و برآوردهای حداقل مربعات	۵
تعریف مدل‌های خطی	۶
برآوردهای حداقل مربعات	۸
کارآئی	۱۰
تابع ψ	۱۰
تاریخچه	۱۳

فصل دوم: برآوردهای OUAE پارامتر مقیاس در خانواده توزیعهای مکان- مقیاس

متقارن ۱۴

برآوردهای BLUE پارامترهای مکان- مقیاس با استفاده از آمارهای ترتیبی ۱۵

صفحة	عنوان
۱۶	وجود برآوردهای BLUE براوردهای BLUE پارامترهای مکان_مقیاس
۱۷	برآوردهای BLUE خاصیت تقارنی آمارههای ترتیبی از توزیعهای متقارن
۱۹	برآوردهای پارامترهای مکان_مقیاس برای توزیعهای متقارن براوردهای BLUE مکان_مقیاس برای نمونههای سانسور شده
۲۱	محاسبه معکوس ماتریس‌های الگو شده براوردهای BLUE توزیعهای یک پارامتری
۲۳	برآوردهای ناریب بهینه خطی پارامتر مقیاس خانوادهای مکان_مقیاس متقارن توسط قدر مطلق آمارههای ترتیبی
۲۴	ارتباط بین گشتاورهای دو مجموعه از آمارههای ترتیبی فصل سوم: برآوردهای BLUE و OUAE
۲۸	و نرمال براوردهای BLUE و OUEA گشتاورهای آمارههای ترتیبی استاندارد شده توزیع مستطیلی
۲۹	ارتباط بین گشتاورهای دو مجموعه از آمارههای ترتیبی گشتاورهای قدر مطلق آمارههای ترتیبی استاندارد شده توزیع مستطیلی
۳۲	گشتاورهای آمارههای ترتیبی استاندارد شده توزیع مستطیلی براوردهای BLUE و OUEA گشتاورهای آمارههای ترتیبی استاندارد شده توزیع نرمال
۴۱	برآوردهای BLUE و OUEA گشتاورهای آمارههای ترتیبی استاندارد شده توزیع نرمال براوردهای BLUE و OUEA گشتاورهای آمارههای ترتیبی استاندارد شده توزیع نرمال
۴۲	برآوردهای BLUE و OUEA گشتاورهای آمارههای ترتیبی استاندارد شده توزیع نرمال براوردهای BLUE و OUEA گشتاورهای آمارههای ترتیبی استاندارد شده توزیع نرمال
۴۴	برآوردهای BLUE و OUEA گشتاورهای آمارههای ترتیبی استاندارد شده توزیع نرمال براوردهای BLUE و OUEA گشتاورهای آمارههای ترتیبی استاندارد شده توزیع نرمال
۴۸	برآوردهای BLUE و OUEA گشتاورهای آمارههای ترتیبی استاندارد شده توزیع نرمال براوردهای BLUE و OUEA گشتاورهای آمارههای ترتیبی استاندارد شده توزیع نرمال

عنوان

صفحه

گشتاورهای قدر مطلق آماره‌های ترتیبی توزیع نرمال استاندارد ۵۱

**فصل چهارم: برآورده کننده OUEA و BLUE پارامتر مقیاس توزیع‌های نمائی
دوگانه و واپل دوگانه ۵۷**

برآورده کننده OUAE و BLUE پارامتر مقیاس توزیع نمائی دوگانه ۵۷

گشتاورهای قدر مطلق از آماره‌های ترتیبی استاندارد شده توزیع نمائی دوگانه ۵۸

گشتاورهای آماره‌های ترتیبی استاندارد شده توزیع نمائی دوگانه ۶۳

برآورده کننده OUAE و BLUE پارامتر مقیاس توزیع واپل دوگانه ۶۹

گشتاورهای قدر مطلق از آماره‌های ترتیبی استاندارد شده توزیع واپل دوگانه ۶۹

گشتاورهای آماره‌های ترتیبی استاندارد شده توزیع واپل دوگانه ۷۲

فصل پنجم: نتیجه‌گیری ۹۰

واژه نامه فارسی- انگلیسی ۹۵

واژه نامه انگلیسی- فارسی ۹۷

منابع ۹۹

صفحه چکیده و صفحه عنوان به زبان انگلیسی
.....

فهرست جداول

عنوان	صفحة
جدول ۱_۱-۳: ضرایب برآورده شده BLUE و OUAE و کارآئی نسبی برآورده شده	
نسبت به برآورده شده BLUE، پارامتر مقیاس در نمونه‌های کامل و سانسور شده متقابله در توزیع مستطیلی متقابله ۴۶	۴۶
جدول ۲_۱-۳: ضرایب برآورده شده BLUE و OUAE و کارآئی نسبی برآورده شده	
نسبت به برآورده شده BLUE، پارامتر مقیاس در نمونه‌های سانسور شده راست (چپ) در توزیع مستطیلی متقابله ۴۷	۴۷
جدول ۱_۳: میانگین آماره‌های ترتیبی استاندارد شده توزیع نرمال ۴۹	۴۹
جدول ۲_۳: واریانس، کوواریانس آماره‌های ترتیبی استاندارد شده توزیع نرمال ۵۰	۵۰
جدول ۳_۳: میانگین آماره‌های ترتیبی کای درجه ۵۲	۵۲
جدول ۴_۳: واریانس و کوواریانس آماره‌های ترتیبی توزیع کای درجه ۵۳	۵۳
جدول ۱_۲-۳: ضرایب برآورده شده BLUE و OUAE و کارآئی نسبی برآورده شده	
نسبت به برآورده شده BLUE، پارامتر مقیاس در نمونه‌های کامل و سانسور شده متقابله توزیع نرمال ۵۵	۵۵
جدول ۲_۲-۳: ضرایب برآورده شده BLUE و OUAE و کارآئی نسبی برآورده شده	
نسبت به برآورده شده BLUE، پارامتر مقیاس در نمونه‌های راست (چپ) در توزیع نرمال .. ۵۶	۵۶
جدول ۴_۱: میانگین و واریانس آماره‌های ترتیبی نمائی استاندارد شده ۶۱	۶۱

جدول ۲_۴: واریانس و کوواریانس آماره‌های ترتیبی استاندارد شده توزیع نمائی دو گانه ...	۶۴
جدول ۳_۴: میانگین آماره‌های ترتیبی استاندارد شده توزیع نمائی ...	۶۵
جدول ۲_۱_۴: ضرایب برآورده شده OUAE و کارآئی نسبی برآورده شده BLUE	
نسبت به برآورده شده BLUE، پارامتر مقیاس در نمونه‌های سانسور شده متقارن، توزیع نمائی	
دو گانه ...	۶۸
جدول ۴_۴: میانگی آماره‌های ترتیبی استاندارد شده توزیع وایل دو گانه ...	۷۴
جدول ۵_۴: واریانس و کوواریانس آماره‌های ترتیبی استاندارد شده توزیع وایل دو گانه ...	۷۹
ادامه جدول ۴_۵ ...	۸۰
ادامه جدول ۴_۵ ...	۸۱
جدول ۱_۲_۱: ضرایب برآورده شده OUAE و کارآئی نسبی برآورده شده BLUE	
نسبت به برآورده شده BLUE، پارامتر مقیاس در نمونه‌های کامل و سانسور شده متقارن در	
توزیع وایل دو گانه ...	۸۳
ادامه جدول ۱_۲_۱: ...	۸۴
ادامه جدول ۱_۲_۱: ...	۸۵
ادامه جدول ۱_۲_۱: ...	۸۶
جدول ۲_۲_۴: ضرایب برآورده شده OUAE و کارآئی نسبی برآورده شده BLUE	
نسبت به برآورده شده BLUE، پارامتر مقیاس در نمونه‌های سانسور شده راست (چپ) توزیع	
وایل دو گانه ...	۸۷
ادامه جدول ۲_۲_۴: ...	۸۸

فصل اول

مقدمه و تاریخچه

هدف اصلی از ارائه مطالب مندرج در این پایاننامه، معرفی برآورده کننده جدیدی بنام برآورده کننده ناریب بهینه خطی؛ (OUAE) پارامتر مقیاس (Scale) توسط قدر مطلق آماره‌های ترتیبی در خانواده توزیع‌های مکان - مقیاس (Location - Scale) (با مشخص بودن پارامتر مکان) می‌باشد.

همچنین محاسبه آن برای یک نمونه تصادفی به اندازه $n = 3$ برای توزیع‌های مستطیلی متقارن - نرمال - نمائی دوگانه (لاپلاس) - وایل دوگانه.

هدف دیگر مورد نظر بررسی کارائی نسبی برآورده کننده OUAE نسبت به برآورده BLUE (بهترین برآورده کننده ناریب خطی توسط آماره‌های ترتیبی) پارامتر مقیاس می‌باشد که نشان داده می‌شود عموماً برای نمونه‌های تصادفی به اندازه‌ها مختلف در توزیع‌های ذکر شده برآورده کننده OUAE کاراتر از برآورده کننده BLUE است.

برای محاسبه هر دو برآورده، نیاز به محاسبه گشتاور اول و دوم و گشتاور حاصل‌ضریبی از قدر مطلق آماره‌های ترتیبی استاندارد شده و همچنین آماره‌های ترتیبی استاندار شده توزیع‌های ذکر شده می‌باشد که چگونگی روش محاسبه گشتاورها هدف دیگر ارائه مطالب می‌باشد.

بنابراین برای بررسی اهداف فوق؛ ارائه مطالب لازم در ۳ فصل، دنبال خواهد شد.

فصل دوم: به معرفی برآورده کننده OUAE پارامتر مقیاس و برآورده کننده BLUE پارامترهای مکان - مقیاس - توزیع‌های مکان - مقیاس و روش محاسبه آن می‌پردازد.

فصل سوم: برآورده کننده‌های OUAE و BLUE پارامتر مقیاس در توزیع‌ها مثلثی متقارن و نرمال (با مشخص بودن پارامتر مکان) برای نمونه‌ها تصادفی به اندازه $n = 3$ محاسبه می‌گردد

و طبق جداولی، ضرایب برآوردها و واریانس آنها و کارانی نسبی OUAE نسبت به BLUE برای نمونه‌های فوق ارائه می‌گردد و نتایجی از این جداول بدست می‌آید.

فصل چهارم: همانند فصل سوم برآوردکننده OUAE و BLUE پارامتر مقیاس برای توزیع‌های نمائی دوگانه و واپل دوگانه برای نمونه‌های تصادفی به اندازه $n = 3(1)10$ محاسبه می‌گردد و جداولی از ضرایب برآوردها و واریانس آنها و کارانی نسبی OUAE نسبت به BLUE بدست می‌آید که نتایجی قابل توجه از این جدول‌ها ارائه می‌گردد.

قبل از پرداختن به مطالب فصل‌های ۲ تا ۴ نیاز به یادآوری مطالبی می‌باشیم که در فصلهای مذکور مورد استفاده قرار می‌گیرند که به اختصار مطالب یادآوری می‌شوند.

۱-۱- آماره‌های ترتیبی

فرض کنید X_1, \dots, X_n متغیرهای تصادفی و پیوسته باشند که هر یک دارای تابع توزیع $F(x)$ و تابع چگالی $f(x)$ باشند، متغیرهای تصادفی $X_{1:n}, \dots, X_{n:n}$ را در نظر بگیریم با این فرض که این نمونه از جایگشت مجموعه X_1, \dots, X_n به طوری که به صورت صعودی مرتب شده‌اند، حاصل شده است این نمونه را، آماره‌های ترتیبی گویند که متغیرهایی وابسته می‌باشند. مبحث آماره‌های ترتیبی، بیشتر در مورد خواص و کاربرد متغیرهای تصادفی مرتب شده و توابعی از آنها بحث و گفتگو می‌کند.

مثال ۱-۱: دستگاهی دارای n قسمت می‌باشد زمانهای خراب شدن این قسمتها عبارتند از X_1, \dots, X_n که دارای فرضهای ارائه شده در قسمت فوق هستند، آنگاه X_k عبارتست از زمانی که طی آن قسمت k از کار می‌افتد، اگر دستگاهی به مجرد اینکه یکی از قسمتها از کار بیفتد متوقف گردد، آنگاه $X_{1:n} = \min(X_1, \dots, X_n)$ عبارت است زمان خراب شدن دستگاه. اگر دستگاه تا زمانی که n امین قسمت خراب شود متوقف نشود آنگاه $X_{n:n} = \max(X_1, \dots, X_n)$ عبارت

است از زمان خراب شدن دستگاه.

قضیه: (۱-۱) فرض کنید X_1, \dots, X_n یک نمونه تصادفی به اندازه n از یک تابع پیوسته با تابع

چگالی $f(x)$ و توزیع $F(x)$ باشد، بنابراین تابع چگالی r امین آماره ترتیبی $X_{r:n}$ بصورت

$$f_r(x) = D_{r,n} [f(x)]^{r-1} [1 - f(x)]^{n-r} f(x)$$

است و تابع چگالی توأم r امین، s امین آماره ترتیبی بصورت

$$f_{r,s}(x, y) = D_{r,s,n} f(x) f(y) F(x)^{r-1} [F(y) - F(x)]^{s-r-1}$$

$$\times [1 - f(y)]^{n-s} ; x \leq y$$

می‌باشد، همچنین تابع چگالی توأم n آماره ترتیبی بصورت

$$f_{1,2,\dots,n}(x_1, \dots, x_n) = n! \pi_{i=1}^n f(x_i)$$

$$-\infty < x_1 < x_2 < \dots < x_n < +\infty$$

می‌باشد.

$$D_{r,n} = \frac{n!}{(r-1)!(n-r)!}$$

$$D_{r,s,n} = \frac{n!}{(r-1)!(s-n-r)!(n-s)!}$$

[۲] اثبات:

n قطعه را در نظر می‌گیریم، برای آزمون طول عمر، قطعات را مورد آزمایش قرار می‌دهیم تا

n قطعه از کار بیافتد یا در اصطلاح مشاهده شود که زمان شکست هر قطعه، زمان طول

عمر قطعه می‌باشد، بعنوان مثال در یک کاربرد نمونه‌ای پژوهشی یا طول عمر ممکن است شروع

یک بیماری مزمن تا لحظه مرگ باشد یا در قابلیت اعتماد صنعتی زمان شکست به عنوان طول عمر

قطعات دستگاهها در نظر گرفته می‌شود.

فرض کنید طول عمر هر قطعه متغیر تصادفی با تابع توزیع پیوسته $F(x; \theta)$ و تابع چگالی $f(x; \theta)$ باشند، که θ پارامتر مجهول می‌باشد، (طول عمر قطعات از هم مستقل می‌باشند) در نتیجه با توجه به مطالب فوق یک نمونه تصادفی مرتب شده $(X_{1:n}, \dots, X_{n:n})$ خواهیم داشت. حالی را در نظر بگیرید که ما مجبور شویم بنا به دلایلی آزمایش را قطع کنیم قبل از اینکه همه قطعات از کار بیافتد، در نتیجه زمان شکست یا طول عمر بعضی از قطعات در دسترس ما نیست، پس در این حالت نمونه تصادفی مرتب شده را سانسور شده گویند در اینجا به تعریف دو نوع از نمونه‌های سانسور شده می‌پردازیم.

۱-۱-۱ سانسور نوع I (سانسور زمانی)

فرض کنید که تصمیم بگیریم که آزمایش فوق را در یک زمان مشخص t متوقف کیم، که در این صورت زمانهای شکست قطعات، شکست خورده قبل از زمان t ثبت می‌شود، نمونه بدست آمده را در این حالت سانسور نوع I یا سانسور زمانی گویند که سانسور راست نیز می‌نامند زیرا داده‌های سمت راست نمونه مرتب شده (داده‌ها بزرگ) سانسور شده است. واضح است که در این حالت تعداد مشاهدات آماره‌های ترتیبی یک متغیر تصادفی است. هر گاه متغیر تصادفی را R تعریف کنیم، مقادیر زیر را می‌تواند بگیرد

$$R = 0, 1, \dots, n$$

که تابع درستنمایی آن بصورت

$$L(\theta | r, \underline{x}) = \begin{cases} \frac{n!}{(n-r)!} \pi_{i=1}^r f(x_i) [1 - F(t, \theta)]^{n-r} & 0 < r_1 < n \\ & x_1 < x_2 < \dots < x_r < t \\ \{1 - F(t, \theta)\}^n & r = 0 \quad t < x_1 \end{cases} \quad (1.1)$$

می‌باشد. (\underline{x} یک بردار از مشاهدات آماره‌ها ترتیبی سانسور شده است)

: [۲] اثبات

۱-۱-۲ - سانسور نوع II (سانسور شکستی)

فرض کنید که بنا به دلایلی تصمیم بگیریم که آزمایش را وقتیکه r امین شکست رخ داد متوقف کنیم، در این حالت نمونه مرتب شده $(X_{1:n}, \dots, X_{r:n})$ را نمونه سانسور شده نوع II یا سانسور شکستی نامند. در این حالت $X_{r:n}$ زمان شکست r امین قطعه، متغیر تصادفی می‌باشد با تابع درستنمایی به صورت

$$L(\theta|r, \underline{x}) = \frac{n!}{(n-r)!} \pi_{i=1}^r f(x_i) [1 - F(x_r, \theta)]^{n-r} \quad x_1 < \dots < x_r \quad (1.2)$$

: [۲] اثبات

همانطور که در سانسور نوع I اشاره گردید چون داده‌های بزرگ سانسور شده است در این حالت نمونه را نمونه سانسور راست نیز نامیده می‌گردد. همچنین در حالتی که داده‌ها کوچک سانسور شوند نمونه $(X_{r+1:n}, \dots, X_{n:n})$ را سانسور چپ گویند. حتی داده‌ها می‌توانند از طرف چپ و راست سانسور شوند که نمونه $(X_{r_1+1:n}, \dots, X_{n-r_2:n})$ را سانسور دو طرفه گویند که r_1 تعداد مشاهدات سانسور شده طرف چپ و r_2 تعداد مشاهدات سانسور شده طرف راست را نشان می‌دهد. در صورتیکه $r_1 = r_2 = r$ باشد نمونه را نمونه سانسور شده متقاضی نامند و هر گاه $0 = r_1 = r_2$ باشد نمونه تصادفی را نمونه کامل گویند.

لازم به ذکر است به طور کلی هر زیر مجموعه از آماره‌های ترتیبی $(X_{1:n}, \dots, X_{n:n})$ برای هر توزیعی، را سانسور نوع II نامند.

۲-۱ - مدل‌های خطی و برآوردهای حداقل مربعات

در این بخش مختصرآ در مورد مدل‌های خطی و برآوردهای حداقل مربعات - (Least

و نتایج حاصل از آن به بحث می‌پردازد که در فصل بعد مورد استفاده قرار می‌گیرد. پرداختن به مسائل در مورد مدل‌های خطی، بدون استفاده از ماتریس‌ها امکان‌پذیر نمی‌باشد.

اینکه به معرفی چند نماد و تعریف می‌پردازیم. ماتریس $(z_{ij}) = A$ را در نظر بگیریم که درایه j زام آن $z_{ij} = a_{ij}$ می‌باشد، ترانهاده ماتریس A را $A' = (a_{ji})$ یاف می‌کنیم که ماتریسی است که جای سطر و ستونش عوض شده باشد، همچنین هر گاه A یک ماتریس مربع و معکوس‌پذیر باشد، معکوس ماتریس A را با نماد A^{-1} نشان می‌دهیم.

ماتریس A را ماتریس متقارن گوئیم هر گاه $A' = A$ باشد. هر گاه ماتریس $A \in n \times n$ متقارن و برای

$$\forall y \neq 0 \in R^n, \quad y' A y > 0$$

ماتریس A را ماتریس معین مثبت (positive definit) گویند که چنین ماتریس معکوس‌پذیر می‌باشد.

۱_۲_۱ - تعریف مدل‌های خطی:

متغیرهای تصادفی وجود دارند که مشاهدات آنها در یک آزمایش اغلب وابسته به یک یا چند متغیر دیگر می‌باشد. بعنوان مثال در یک آزمایش شیمیایی، متغیر سرعت واکنش وابسته به متغیرهای درجه حرارت و زمان واکنش می‌باشد.

متغیر تصادفی y_i را در نظر می‌گیریم بطوریکه مشاهده، y_i وابسته به مقادیر مشاهده شده

$$a_{i0}, a_{i1}, \dots, a_{ip} \quad i = 1, \dots, n \quad n \geq p + 1$$

باشد با فرض اینکه

$$E(Y_i) = \sum_{k=0}^p a_{ik} B_k \quad (1.3)$$