بسم الله الرحمن الرحیم

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

دانشکده علوم پایه
گروه آمار

پایان‌نامه جهت اخذ درجه کارشناسی ارشد
رشته آمار ریاضی

عنوان :
پیش بینی مقادیر رکورد و آماره های مرتب

استاد راهنما:
دکتر بهاءالدین خالدی

نگارش:
الهام داوودی

بهمن ماه ۱۳۹۰

دانشکده علوم پایه
گروه آمار

پایان‌نامه جهت اخذ درجه کارشناسی ارشد
رشته آمار ریاضی

نام دانشجو:
الهام داوودی

تحت عنوان :
پیش بینی مقادیر رکورد و آماره های مرتب

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

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

استاد داور داخل گروه با مرتبه‌ی علمی امضاء:

استاد داور خارج گروه با مرتبه‌ی علمی امضاء:

... و در آغاز هیچ نبود کلمه بود و آن کلمه خدا بود...

مرا کسی نساخت، خدا ساخت، نه آنچنان که کسی می خواست، که من کسی نداشتم کسم خدا بود خدا را سپاس. الهی تو را سپاس که به برگیری قطره ای از اقیانوس بی کران علم و دانشت یاریمان دادی و در بیت زیر سکنایمان که:

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

اکنون که با لطف و یاری پروردگار انجام این رساله به اتمام رسیده است، بر خود لازم می دانم از کلیه عزیزانی که در این راه مرا یاری نموده اند صمیمانه قدردانی نمایم. از جناب آقای دکتر بهاء الدین خالدی استاد عزیزم که با نظرات و پیشنهادات ارزنده و زحمات بی دریغشان در طی مراحل اجرا، تدوین و ارائه رساله مرا یاری نموده اند، صمیمانه قدردانی می نمایم. از پدر و مادر فداکارم که در کلیه مراحل تحصیل مشوق و راهنمای من بودند و با نور شمع وجودشان، روشنی بخش راهم گردیدند، خالصانه سپاسگذارم و بوسه بر دستان پر مهرشان می نهم که هر چه دارم از وجود پاک و مهربان آنهاست و به یکایک تارهای سپید مویشان هزاران دین دارم. نهایت سپاس را به برادران و خواهر عزیزم پیشکش می نمایم که وجودشان مرا سراسر لطف بوده است و مهربانی.

هم چنین، سپاسگذار زحمات بی دریغ اساتید بزرگوار آقایان دکتر سیاره، دکتر هاشمی، دکتر نیاپرست، دکتر نبز اسماعیل زاده و خانم دکتر شرفی، هستم.

الهام داوودی

کرمانشاه - بهمن ماه ۱۳۹۰

تقدیم به

پدر و مادر عزیز

و خواهر و برادران مهربانم

چکیده

در این پایان نامه، رکوردها و آماره های مرتب معرفی شده است سپس بر اساس رکوردهای مشاهده شده از دنباله ای از متغیرهای تصادفی مستقل با توزیع مشترک، فواصل پیش بینی رکوردهای بالا و رکوردهای پایین برای رکوردهایی از دنباله ای مستقل و هم چنین مسئله پیش بینی آماره های مرتب بر اساس مقادیر رکوردی مشاهده شده و پیش بینی رکوردهای آینده بر اساس آماره های مرتب به دست آمده است. احتمال های پوشش این فواصل دقیق بوده و آزاد توزیع می باشند. فواصل پیش بینی درونی و بیرونی برای فواصل رکوردی آینده و فاصله رکوردی آنها بر اساس آماره های مرتب مشاهده شده با توزیع مشترک داده شده است. به علاوه، وقتی که یک رکورد جدید از هر نوع (بالا یا پایین) رخ بدهد، بر اساس بزرگترین و کوچکترین مشاهدات (رکوردهای جاری) پیش بینی آماره های مرتب آینده، مورد بحث قرار گرفته است. فواصل پیش بینی پیشنهاد شده، آزاد توزیع هستند. سرانجام پیش بینی بیرونی و درونی برای فواصل رکوردی جاری آینده بر اساس آماره های مرتب به دست آمده است.

کلمات کلیدی:

آماره های مرتب، رکوردها، فاصله رکوردی، برد نمونه ای، ضریب پیش بینی، رکوردهای جاری، فواصل پیش بینی ناپارامتری، پوشش رکوردی .

فهرست مطالب

پیشگفتار

به دلیل اهمیت رکوردها و کاربردشان در مسائل اجتماعی، اقتصادی، هواشناسی، ورزشی، ژئوفیزیک و صنعتی و غیره به استنباط در زمینه ی رکوردها و استفاده از آن ها برای استنباط درباره ی پارامترهای جامعه می پردازیم. یکی از مسائل استنباطی در زمینه رکوردها، مسئله پیش بینی رکوردهای آینده بر اساس آماره مرتب مشاهده شده و آماره های مرتب آینده بر اساس رکوردهای مشاهده شده می باشد. بر این اساس در این پایان نامه، ابتدا به معرفی رکوردها اعم از رکوردهای معمولی و رکوردهای جاری و هم چنین آماره های مرتب و بیان توابع توزیع آن ها، می پردازیم. در فصل دوم، نخست فواصل اطمینان ناپارامتری برای رکوردهای آینده بر اساس رکوردهای مشاهده شده از دو دنباله از متغیرهای تصادفی مستقل و هم توزیع را بیان می کنیم و در ادامه مسئله پیش بینی رکوردها بر اساس آماره های مرتب مشاهده شده و پیش بینی آماره های مرتب بر اساس رکوردهای مشاهده شده از دو دنباله مستقل را مورد بحث قرار می دهیم. احتمال های پوشش این فواصل محاسبه و جداول آن ها به ازای مقادیر مختلف را به دست می آوریم. در انتهای این فصل، با یک مثال شیوه ی استفاده از مقادیر به دست آمده را مورد ارزیابی قرار می دهیم. در فصل سوم، فواصل پیش بینی آزاد توزیع بیرونی و درونی برای مقادیر رکوردی از دنباله ی X بر اساس آماره های مرتب از دنباله مستقل Y را مورد مطالعه قرار می دهیم. در این فصل فواصل پیش بینی برای رکوردهای بالا، پایین و بالا و پایین به طور همزمان بر اساس آماره های مرتب و فواصل پیش بینی درونی رکوردها بر اساس آماره های مرتب و سرانجام، فواصل پیش بینی برای فاصله رکوردها را مورد توجه قرار می دهیم. در فصل چهارم به بررسی رکوردهای جاری خواهیم پرداخت به این صورت که فواصل پیش بینی آماره های مرتب بر اساس رکوردهای جاری را محاسبه می نماییم. این فواصل بر اساس رکوردهای جاری بالا و رکوردهای جاری پایین می باشند. در ادامه، فواصل پیش بینی بر اساس پوشش رکوردی جاری را به دست می آوریم. در انتها، پیش بینی رکوردهای جاری بر اساس آماره های مرتب شامل رکوردهای جاری بالا و رکوردهای جاری پایین را محاسبه می کنیم. سرانجام، در فصل پنجم فواصل پیش بینی درونی و بیرونی برای رکوردهای جاری بالا و پایین بر اساس آماره های مرتب را به دست خواهیم آورد. قابل ذکر است که منظور از رکوردها در این پایان نامه رکوردهای معمولی می باشد.

فهرست نشانه‌ها و نمادها

$Pr(.)$	اندازه احتمال
$E(.)$	امید ریاضی
$\mid$	به شرط اینکه
∞	بی نهایت
F	تابع توزیع
f	تابع چگالی احتمال
$\in$	عضویت مجموعه ای
Max	عملگر مجموعه ای ماکزیمم
Min	عملگر مجموعه ای می نیمم
R_n	مقدار رکورد n ام
$X_۱, X_۲, \cdots, X_n$	نمونه تصادفی به اندازه ی n
$F^{-۱}$	تابع معکوس تابع توزیع F
$\bar{F}$	تابع بقا توزیع F
iid	به طور مستقل و مشابه توزیع شده
$=^d$	برابری در توزیع (هم توزیعی)
$X_{۱:n}, X_{۲:n}, \cdots, X_{n:n}$	آماره های مرتب
PRR	پیش بینی فواصل برای فواصل رکوردی بر اساس مقادیر رکوردی
PRO	پیش بینی فواصل برای فواصل رکوردی بر اساس آماره های مرتب

فصل ۱

تعاریف و مفاهیم اولیه

۱-۱ رکوردها

در بیشتر مسائل آماری داده ها یا مشاهدات به منظور اطلاعاتی درباره جامعه مورد نظر جمع آوری می شوند. زمانی که به کل داده ها دسترسی نداشته باشیم، اینکه باقیمانده داده ها حاوی چه اطلاعاتی هستند حائز اهمیت است. این وضعیت به ویژه در آزمایش های دنباله ای رخ می دهد که در آن ها فقط مشاهداتی ثبت می شوند که از مقادیر قبلی خود کوچک تر یا بزرگ تر باشند. این داده ها به داده های رکوردی معروف هستند. برای مثال الوارهایی که برای ساختن قایق و یا بلوک هایی که برای سازه های عمرانی مورد استفاده قرار می گیرند، در معرض فشارهای فزاینده ای قرار داده می شوند، اما ما تنها به نقطه تخریب نمونه هایی دسترسی پیدا می کنیم که از نمونه های قبلی خود نامقاوم تر باشند. به این ترتیب در چنین آزمایش هایی اندازه گیری به گونه ای صورت می گیرد که تنها مقادیری که کوچک تر از مقادیر قبلی خود هستند ثبت می شوند. در بسیاری از موارد تنها داده های آماری ثبت می شوند که رکورد باشند، مانند رخدادهای ورزشی در مسابقات المپیک، تغییرات هواشناسی، زلزله نگاری و غیره. لذا این موضوع حائز اهمیت است که بتوان با استفاده از رکوردها، تحلیل آماری و استنباط مورد نظر درباره جامعه مورد بررسی را انجام داد. عموم مردم داده های رکوردی را مورد توجه قرار می دهند. سؤال هایی نظیر گرم ترین روز در ماه گذشته، بیشترین و کم ترین میزان بارندگی در سال جاری، پایین ترین نرخ سود سهام در بازار، بالاترین میزان تورم در چند سال اخیر، بیشترین مقدار مرگ و میر در نوزادان زیر یک سال و حتی تعداد پرتاب های ناموفق یک بسکتبالیست معروف، شامل داده های رکوردی هستند. در سال های اخیر محققین زیادی به بررسی مقادیر رکوردی پرداختند. اولین بار چندلر [۱] (۱۹۵۲) بر اساس یک مدل رکوردی شامل مشاهدات iid، مطالعه مقادیر رکوردی را معرفی کرد و خواص اساسی زیادی از رکوردها را به دست آورد.

در برخی از مراجع، مطالعه مقادیر رکوردی مورد توجه قرار گرفته است، از آن جمله می توان به مقاله گلیک [۲](۱۹۷۸) و نوزوروف [۳] (۱۹۸۷) اشاره نمود. کتاب های آرنولد [۴] و همکاران (۱۹۹۸) و آماره های رکوردی احسان اله [۵] (۲۰۰۴) نیز به معرفی مقادیر رکوردی پرداخته است.

به دلیل اهمیت رکوردها و کاربرد آن ها در مسائل اجتماعی، اقتصادی، هواشناسی، ورزشی، ژئوفیزیک و صنعتی و غیره به استنباط در زمینه رکوردها و هم چنین استفاده از آماره های رکوردی برای استنباط درباره ی پارامترهای جامعه می

[۱] Chandler

[۲] Glick

[۳] Nevzorov

[۴] Arnold

[۵] Ahsanullah

پردازیم. یکی از مسائل استنباطی در این زمینه مسئله پیش بینی رکوردها و آماره های مرتب می باشد که در این پایان نامه مورد مطالعه قرار می گیرد. از این رو می توان رکوردهای آینده را بر اساس رکوردهای مشاهده شده و یا آماره های مرتب مشاهده شده محاسبه نمود. هم چنین آماره های مرتب آینده را نیز می توان بر اساس رکوردهایی که در دسترس هستند به دست آورد. با توجه به اینکه رکوردها را می توان به صورت رکوردهای معمولی و رکوردهای جاری در نظر گرفت. در ادامه این فصل، در بخش دوم به معرفی رکوردهای معمولی می پردازیم، سپس در بخش سوم به معرفی رکوردهای جاری و سرانجام در بخش چهارم به تعریف آماره های مرتب خواهیم پرداخت.

۱-۲ رکوردهای معمولی

در این بخش به تعریف و خواص رکوردهای معمولی می پردازیم.

تعریف ۱-۲-۱. فرض کنید X_1 ، X_2، X_3 ، $\cdots$ دنباله ای نامتناهی از متغیرهای تصادفی پیوسته مستقل با تابع توزیع تجمعی مشترک $F(x)$ و تابع چگالی احتمال $f(x)$ باشد. در این صورت X_j را یک رکورد بالا گویند اگر از تمام مقادیر قبلی

X_1 ، X_2 ، X_3 ، $\cdots$ ، X_{j-1} بزرگ تر باشد، بنابراین X_j یک رکورد بالا است، اگر برای هر $i < j$ ، $X_i < X_j$.

تعریف ۱-۲-۲. فرض کنید X_1 ، X_2 ، X_3 ، $\cdots$ یک دنباله نامتناهی از متغیرهای تصادفی پیوسته مستقل با تابع توزیع تجمعی مشترک $F(x)$ و تابع چگالی احتمال $f(x)$ باشد. پس X_j را یک رکورد پایین گویند اگر از تمام مشاهدات قبلی X_1 ، X_2 ، X_3 ، $\cdots$ ، X_{j-1} کوچک تر باشد. بنابراین X_j یک رکورد پایین است، اگر برای هر $i < j$ ، $X_i > X_j$.

تعریف ۱-۲-۳. اگر X_j در زمان j مشاهده شده باشد، آن گاه دنباله زمانی رکورد که به صورت $\{T_n; n \geq 0\}$ نشان داده می شود، به شکل زیر تعریف می شود:

$T_0 = 1$ ، با احتمال ۱

و برای $n \geq 1$ ؛

$$T_n = min\left\{j : X_j > X_{T_{n-1}}\right\}$$

دنباله مقدار رکوردی $\{R_n\}$ نیز به این صورت تعریف می شود:

$$R_n = X_{T_n} \qquad n = 0, 1, 2, \ldots$$

R_0 مقدار رکورد بدیهی و بقیه رکوردها را رکوردهای غیر بدیهی می نامند.

۱-۲-۱ توابع چگالی احتمال رکوردهای معمولی

فرض کنید X_1 ، X_2 ، X_3 ، $\cdots$ یک دنباله نامتناهی از متغیرهای تصادفی پیوسته مستقل با تابع توزیع تجمعی مشترک $F(x)$ و تابع چگالی احتمال $f(x)$ باشد. هم چنین فرض کنید R_0 ، R_1 ، $\cdots$ ، R_n رکوردهای بالای معمولی متناظر با دنباله X باشند.

R_0 ، R_1 ، $\cdots$ ، R_n دارای تابع چگالی احتمال توأم زیر است:

$$\begin{aligned} f_{R_0,R_1,\ldots,R_n}(r_0,r_1,\ldots,r_n) &= \frac{\prod_{i=0}^{n} f(r_i)}{\prod_{i=0}^{n-1}[1-F(r_i)]} \\ &= f(r_n)\prod_{i=0}^{n-1} h(r_i) \qquad -\infty < r_0 < r_1 < \ldots < r_n < \infty \end{aligned}$$

که $h(r) = \dfrac{f(r)}{[1-F(r)]}$ تابع نرخ شکست F می باشد.

هم چنین تابع چگالی توأم R_n و R_m برای $m < n$ به صورت زیر به دست می آید:

$$f_{R_m,R_n}(r_m,r_n) = \frac{[-\log(1-F(r_m))]^m}{m!}\frac{[-\log(\frac{1-F(r_n)}{1-F(r_m)})]^{n-m-1}}{(n-m-1)!}\frac{f(r_m)f(r_n)}{[1-F(r_m)]}, \ -\infty < r_m < r_n < \infty$$

(۱-۱)

تابع چگالی احتمال R_n ، n امین رکورد بالا عبارت است از:

$$f_{R_n}(r) = f(r)\frac{[-\log(1-F(r))]^{n-1}}{(n-1)!} \tag{۱-۲}$$

و تابع چگالی احتمال $R_n^\star$ ، n امین رکورد پایین به صورت زیر است:

$$f_{R_n^\star}(r) = f(r)\frac{[-\log F(r)]^{n-1}}{(n-1)!} \tag{۱-۳}$$

برای اثبات این روابط می توان به آرنولد و همکاران (۱۹۹۸) مراجعه نمایید.

۱-۲-۲ توابع توزیع تجمعی رکوردهای معمولی

فرض کنید X_1، X_2، X_3، $\cdots$ یک دنباله از متغیرهای تصادفی پیوسته مستقل با تابع توزیع تجمعی مشترک $F(x)$ و تابع چگالی احتمال $f(x)$ باشد. اگر n امین مقدار رکوردی بالا را با R_n نشان دهیم، تابع بقای R_n، وقتی $n \geq 1$ باشد، در زیر آمده است:

$$\bar{F}_{R_n}(x) = \bar{F}(x)\sum_{k=0}^{n-1}\frac{\{-\log \bar{F}(x)\}^k}{k!}, \tag{۱-۴}$$

که $\bar{F}(x) = 1 - F(x)$ تابع بقای X است.

حال اگر $R_n^\star$، r امین مقدار رکوردی پایین از دنباله X باشد، تابع بقا L_r برای $r \geq 1$ به صورت زیر است:

$$\bar{F}_{R_n^{\star +}}(x) = 1 - F(x)\sum_{k=0}^{r-1}\frac{\{-\log F(x)\}^k}{k!}. \tag{۱-۵}$$

آرنولد و همکاران (۱۹۹۸) را ملاحظه نمایید.

آن چه تاکنون گفته شد، مربوط به رکوردهای معمولی بود که در فصل های دوم و سوم به بررسی فواصل پیش بینی آن ها بر اساس آماره های مرتب و هم چنین فواصل پیش بینی آماره های مرتب بر اساس رکوردها می پردازیم. رکوردهای جاری نیز در این پایان نامه مورد مطالعه قرار گرفته اند، که در فصل چهارم و پنجم به بررسی آن ها خواهیم پرداخت.

۱-۳ رکوردهای جاری

ابتدا تعریف رکوردهای جاری را بیان می کنیم.

تعریف ۱-۳-۱. فرض کنید X_1، X_2، X_3، $\cdots$ یک دنباله از متغیرهای تصادفی مستقل با تابع توزیع تجمعی مشترک $F(x)$ و تابع چگالی احتمال $f(x)$ باشد. هم چنین فرض کنید وقتی m امین رکورد رخ داده باشد، $\acute{L_m}$ رکورد جاری پایین و $\acute{U_m}$ رکورد جاری بالا باشد. بنابراین $\acute{L}_{m+1} = \acute{L_m}$ اگر و تنها اگر $\acute{U_m} < \acute{U}_{m+1}$ و $\acute{U_m} = \acute{U}_{m+1}$ اگر وتنها اگر $\acute{L_m} > \acute{L}_{m+1}$.

نکته ۱-۳-۲. یادآور می شویم که اگر U_n و L_n، به ترتیب n امین رکورد معمولی بالا و پایین باشد، آن گاه $U_1 =$ $L_1 \equiv X_1$. هم چنین اگر $\acute{U_n}$ و $\acute{L_n}$، به ترتیب n امین رکورد جاری بالا و پایین باشد، آن گاه $\acute{U_0} = \acute{L_0} \equiv X_1$.

نکته ۱-۳-۳. تابع توزیع تجمعی پواسن با تابع چگالی گامای ناقص (بریده شده) معادل است، یعنی؛

$$\int_t^{+\infty}\frac{\lambda^n x^{n-1}e^{-\lambda x}}{\Gamma(n)}dx = \sum_{k=0}^{n-1}e^{-\lambda t}\frac{(\lambda t)^k}{k!} \tag{۱-۶}$$

که $\Gamma(.)$ تابع گامای کامل است.

قضیه ۱-۳-۴. فرض کنید X_1 ، X_2 ، X_3 ، $\cdots$ یک دنباله از متغیرهای تصادفی مستقل با تابع توزیع تجمعی مشترک $F(x)$ و تابع چگالی احتمال $f(x)$ باشد. هم چنین فرض کنید $\acute{L_m}$ رکورد جاری پایین و $\acute{U_m}$ رکورد جاری بالا باشد. بنابراین تابع چگالی توأم و تابع توزیع تجمعی توأم $(\acute{L_m}, \acute{U_m})$ به صورت زیر به دست می آید:

$$f_m(x,y) = 2^m f(x) f(y) \frac{\{-\log\{1-F(y)+F(x)\}\}^{m-1}}{(m-1)!} \tag{۱-۷}$$

و؛

$$F_m(x,y) \quad = \quad 2^m F(x) + \sum_{k=0}^{m-1} 2^{m-1-k}(T_k(y) - U_k(x,y)), \tag{۱-۸}$$

هرگاه $-\infty < x < y < \infty$ باشد. که در آن:

$$T_k(y) = \{1-F(y)\}^2 \textstyle\sum_{j=0}^{k} \frac{\{-2\log(1-F(y))\}^j}{j!}$$

و؛

$$U_k(x,y) = \{1-F(y)+F(x)\}^2 \textstyle\sum_{j=0}^{k} \frac{\{-2\log(1-F(y)+F(x))\}^j}{j!}.$$

اثبات. بدون از دست دادن کلیت، فرض کنید رکوردها از یک دنباله از متغیرهای تصادفی مستقل از توزیع $U(0,1)$ به دست آمده باشند، در این صورت با استفاده از قضیه تبدیل انتگرال احتمال نتیجه برای هر تابع توزیع تجمعی پیوسته به دست می آید. بنابراین، فرض کنید $\acute{L_m^\star}$ ، m امین رکورد جاری پایین و $\acute{U_m^\star}$ ، m امین رکورد جاری بالای متناظر با دنباله ی متغیرهای تصادفی مستقل از توزیع $U(0,1)$ باشند. حال احتمال شرطی زیر را در نظر بگیرید.

$$Pr(\acute{L_m^\star} \leq x, \acute{U_m^\star} > y \mid \acute{L_{m-1}^\star} = t, \acute{U_{m-1}^\star} = s) = \begin{cases} I_1 & x \geq t, y \leq s \\ I_2 & x < t, y > s \\ I_3 & x < t, y \leq s \\ I_4 & x \geq t, y > s \end{cases} \tag{۱-۹}$$

که $0 < x < y < 1$ و $0 < t < s < 1$ طوری که $m = 1,2,3,\ldots$.

از این رو، اگر $x \geq t$ باشد آن گاه $x \geq \acute{L_{m-1}^\star}$ و از آنجائیکه $\acute{L_m^\star} < \acute{L_{m-1}^\star}$ بنابراین می توان نتیجه گرفت $\acute{L_m^\star} \leq x$ و نیز اگر $y < s$ یعنی $y < \acute{U_{m-1}^\star}$ ، با توجه به اینکه $\acute{U_{m-1}^\star} < \acute{U_m^\star}$ آن گاه $y < \acute{U_m^\star}$ می باشد. بنابراین هر دو پیشامد همواره وجود دارد از این رو $I_1 = 1$ به دست می آید. هم چنین اگر $x < t$ آن گاه $x < \acute{L_{m-1}^\star}$ و از آنجائیکه احتمال

دارد که $L^{\star'}_{m-۱}$ با $L'^{\star}_m$ مساوی باشد می توان نتیجه گرفت که $x < L'^{\star}_m$ و به طور مشابه $y > U'^{\star}_m$. بر همین اساس وقتی $y > s, x < t$، $I_۲ = ۰$ خواهد بود. اینک حالتی را در نظر می گیریم که $y \le s, x < t$، در این صورت فرض کنید $Z_۱$ نخستین مشاهده ای باشد که بعد از $(L^{\star'}_{m-۱} = t, U^{\star'}_{m-۱} = s)$ رخ داده است، سه حالت برای $Z_۱$ وجود دارد $Z_۱ < t$، $t < Z_۱ < s$ یا $Z_۱ > s$. در حالت اول می توان گفت پیشامد $U^{\star'}_{m-۱} = s$ و $L^{\star'}_{m-۱} = s$ و $Z_۱ < t$ با پیشامد $U'^{\star}_m = U^{\star'}_{m-۱}$ و $L'^{\star}_m < L^{\star'}_{m-۱}$ و هم چنین $L'^{\star}_m =^d U(۰,t)$ معادل است. اگر حالت دوم را در نظر بگیریم پیشامد $U^{\star'}_{m-۱} = s$ و $L^{\star'}_{m-۱} = t$ و $t < Z_۱ < s$ به این معنی است که رکورد جدیدی اتفاق نیفتاده است. هم چنین پیشامد $U^{\star'}_{m-۱} = s$ و $L^{\star'}_{m-۱} = t$ و $Z_۱ > s$ با پیشامد $L^{\star'}_{m-۱} = L'^{\star}_m$ و $U^{\star'}_{m-۱} < U'^{\star}_m$ معادل خواهد بود که $U'^{\star}_m =^d U(s,۱)$. بنابراین:

$$\begin{aligned}
I_۳ &= Pr(L'^{\star}_m \le x, U'^{\star}_m > y \mid Z_۱ < t, L^{\star'}_{m-۱} = t, U^{\star'}_{m-۱} = s) Pr(Z_۱ < t) \\
&\quad + Pr(L'^{\star}_m \le x, U'^{\star}_m > y \mid t < Z_۱ < s, L^{\star'}_{m-۱} = t, U^{\star'}_{m-۱} = s) Pr(t < Z_۱ < s) \\
&\quad + Pr(L'^{\star}_m \le x, U'^{\star}_m > y \mid Z_۱ > s, L^{\star'}_{m-۱} = t, U^{\star'}_{m-۱}) Pr(Z_۱ > s) \\
&= \frac{x}{t} t + I_۳(s-t) + ۰
\end{aligned}$$

بنابراین داریم:

$$I_۳(۱ - s + t) = x \Rightarrow I_۳ = \tfrac{x}{۱-s+t}$$

و به طور مشابه می توان به دست آورد که:

$$I_۴ = \tfrac{۱-y}{۱-s+t}.$$

با استفاده از قضیه احتمال کل داریم:

$$\begin{aligned}
Pr(L'^{\star}_m \le x, U'^{\star}_m > y) &= \int_۰^x \int_y^۱ f_{m-۱}(t,s) ds dt \\
&\quad + \int_y^۱ \int_x^s \frac{x}{t + ۱ - s} f_{m-۱}(t,s) dt ds \\
&\quad + \int_۰^x \int_t^y \frac{۱ - y}{t + ۱ - s} f_{m-۱}(t,s) ds dt. \qquad (۱-۱۰)
\end{aligned}$$

با توجه به عبارت بالا، اولین انتگرال $Pr(L^{\star'}_{m-۱} \le x, U^{\star'}_{m-۱} > y)$ می باشد. حال با توجه به اینکه:

$$F_k(x,y) = Pr(L'^{\star}_k \le x) - Pr(L'^{\star}_k < x, U'^{\star}_k > y)$$

وقتی که $k = m, m-1$ ، از این رو می توان نوشت:

$$Pr(L'^{\star}_{m-1} \leq x, U'^{\star}_{m-1} > y) = Pr(L'^{\star}_{m-1} \leq x) - F_{m-1}(x,y) \tag{۱۱-۱}$$

حال با استفاده از رابطه (۹.۱) و (۱۰.۱) به دست می آوریم:

$$\begin{aligned} F_m(x,y) &= F_{m-1}(x,y) \\ &- \int_y^1 \int_x^s \frac{x}{t+1-s} f_{m-1}(t,s)\,dtds \\ &- \int_0^x \int_t^y \frac{1-y}{t+1-s} f_{m-1}(t,s)\,dsdt \\ &+ Pr(L'^{\star}_m \leq x) - Pr(L'^{\star}_{m-1} \leq x). \end{aligned} \tag{۱۲-۱}$$

برای به دست آوردن تابع چگالی احتمال توأم m امین رکورد جاری بالا و پایین از توزیع یکنواخت استاندارد، با توجه به پیوسته بودن F می توان نسبت به x و y مشتق گرفت، بنابراین داریم:

$$f_m(x,y) = \int_x^y \frac{1}{t+1-y} f_{m-1}(t,y)\,dt + \int_x^y \frac{1}{x+1-s} f_{m-1}(x,s)\,ds. \tag{۱۳-۱}$$

برای $m = 1$ نمی توان با جایگذاری در فرمول فوق مقدار توزیع توأم را محاسبه کرد. اما می توان گفت (می دانیم که وقتی $m = 0$ ، $\acute{L}_0 = \acute{U}_0 \equiv X_1$) اولین رکورد جاری (یعنی وقتی $m = 1$) یا از x_1 کوچک تر و یا بزرگ تر است بنابراین دو مشاهده داریم به صورت $x_1 < x_2$ یا $x_2 < x_1$. توزیع توأم نخستین دو آماره مرتب از یک نمونه به حجم ۲ از توزیع یکنواخت استاندارد به صورت زیر تعریف می شود:

$$f_{X_1,X_2}(x,y) = \tfrac{2!}{0!0!}\{F(x)\}^0\{F(y)-F(x)\}^0 f(y)f(x) = 2.$$

بنابراین اگر $m = 1$ باشد، تابع چگالی احتمال توأم نخستین دو آماره مرتب از یک نمونه تصادفی از اندازه ۲ به دست می آید، یعنی:

$$f_1(x,y) = 2 \qquad 0 < x < y < 1.$$

بنابراین قضیه برای $m = 1$ برقرار است. با استفاده از استقرای ریاضی، فرض کنید قضیه برای $m = n$ نیز برقرار باشد، سپس درستی رابطه را به ازای $m = n+1$ اثبات می شود. تابع چگالی احتمال توأم بالا را برای $m = n+1$، می توان با به کار بردن رابطه (۱۳.۱) و هم چنین با استفاده از درستی فرض استقرای $m = n$ و جایگذاری رابطه (۷.۱) به جای تابع

چگالی احتمال زیر انتگرال، به صورت زیر به دست آورد:

$$\begin{aligned} f_{n+1}(x,y) &= \frac{2^n}{(n-1)!}\left\{\int_x^y \frac{1}{1-y+t}\{-\log(1-y+t)\}^{n-1}\,dt\right\} \\ &\quad + \frac{2^n}{(n-1)!}\left\{\int_x^y \frac{1}{1-s+x}\{-\log(1-s+x)\}^{n-1}\,ds\right\} \end{aligned}$$

برای حل انتگرال های عبارت بالا تغییر متغیرهای $w=-\log(1-y+t)$ و $u=-\log(1-s+x)$ را به کار می بریم:

$$\begin{aligned} f_{n+1}(x,y) &= \frac{2^n}{(n-1)!}\left[\int_0^{-\log(1-y+x)} w^{n-1}dw + \int_0^{-\log(1-y+x)} u^{n-1}du\right] \\ &= \frac{2^n}{(n-1)!}\left[\frac{\{-\log(1-y+x)\}^n}{n} + \frac{\{-\log(1-y+x)\}^n}{n}\right] \\ &= 2^{n+1}\frac{\{-\log(1-y+x)\}^n}{n!}. \end{aligned}$$

بنابراین قضیه برای $m=n+1$ برقرار است، با توجه به استقرای ریاضی در نتیجه برای $m=n$ نیز برقرار می باشد. سرانجام با استفاده از قضیه تبدیل انتگرال احتمال و در نظر گرفتن تبدیل های $x=L'_m=F^{-1}(L'^{\star}_m)$ و $y=U'_m=F^{-1}(U'^{\star}_m)$، اثبات (۷.۱) تمام است.

حال رابطه (۸.۱) را ثابت می کنیم. در اینجا نیز رابطه (۸.۱) برای متغیرهای تصادفی توزیع $U(0,1)$ اثبات شده است. رابطه (۸.۱) در واقع حل انتگرال زیر است:

$$F_{L'^{\star}_m,U'^{\star}_m}(t,s)=\frac{2^m}{(m-1)!}\int_0^t\int_x^s\{-\log(1-y+x)\}^{m-1}\,dydx$$