$k-$رکوردها، تعاریف و مفاهیم اساسی

محمد نوری [۱] حمید رضا نیلی ثانی [۲]

چکیده:

در این مقاله ضرورت توجه به داده‌های $k-$رکوردی توضیح داده می‌شود. در ادامه تعاریف مختلف $k-$رکوردها بیان و برخی از نظرات از جمله نظرات دمبنیسکا و بلازکیز (۲۰۰۵) در خصوص تفاوت بین این تعاریف شرح داده شدند. در پایان برخی از ویژگی‌های $k-$رکوردها مورد بحث قرار می‌گیرند.

واژه‌های کلیدی: متغیرهای تصادفی iid ، متغیرهای تصادفی مطلقا پیوسته، $k-$رکوردها.

۱ مقدمه

در محیط پیرامون زندگی مان، آزمایش‌ها و رخدادهای زیادي وجود دارند که در دوره‌های مختلف تکرار می‌شوند. المپیک‌های ورزشي، المپیادهاي علمي، آزمون‌های سراسري، تولیدات یک کارخانه یا دوره‌های مختلف تولید يك كالاي صنعتي و برخی از پدیده‌های طبیعی، به عنوان نمونه بارندگی‌های سالیانه، از جمله مثال‌های متنوعي هستند که تکرار آزمایش‌ها در آن‌ها به وضوح دیده مي شوند. در این آزمایش‌ها، ممکن است اندازه گیری‌ها به طور منظم انجام ولی تنها مقادیری که بزرگتر (یا کوچکتر) از مقادیر قبلی هستند ثبت شوند. این مقادیر، رکوردها می‌باشند. در دهه‌های اخیر این متغیرها مورد توجه بسیاری از محققین ومولفین قرار گرفته‌اند که از آن جمله می‌توان به ناگاراجا (۱۹۸۸)، کامپس (۱۹۹۵)، آرنولد و همکاران (۱۹۹۸)و نزروف (۲۰۰۲) اشاره نمود. گهگاهی، از جمله مواردی که احساس می‌شود در ثبت مقادیر رکورد ممکن است خطاهای صورت گرفته و یا داده‌ها شامل مقادیر پرت باشند، ترجیح داده می‌شود دومین یا سومین و یا در حالت کلی $k-$امین بزرگترین (یا کوچکترین) مقدار نسبت به مقادیر قبلي را ثبت و مطالعه کنیم. به چنین مقادیری $k-$رکوردهای بالا (یا پایین)گفته مي شود. به عنوان نمونه درخواست‌های خسارت در برخی از بیمه نامه‌ها (بجزبیمه عمر) از این گونه‌اند، کامپس (۱۹۹۵). در دو دهه اخیر $k-$رکوردها مورد توجه بیشتری قرار گرفته‌اند. دزبدزیلا و کوپنسکی ((b)۱۹۷۶) ، آرنولد و همکاران (۱۹۹۸)، نزروف (۲۰۰۲) ، هافمن و بالاکریشنان (۲۰۰۴)، دمبینسکا و بلازکیز (۲۰۰۵) و احمدي و همکاران (۲۰۰۹) از جمله افرادی هستند که ویژگی‌های $k-$رکوردها را مطالعه نموده‌اند.

در بخش دوم تعاریف مختلف $k-$رکوردها را ارائه و ضمن فراخوان نظرات دمبینسکا و بلاکی(۲۰۰۵) این تعاریف را با هم مقایسه می‌کنیم. بخش پایانی را به بیان برخی از ویژگی‌های اساسی $k-$رکوردها اختصاص داده‌ایم.

۲ مفاهیم مقدماتی $k-$رکوردها

همانگونه که در مقدمه نیز اشاره شد تعاریف مختلفی برای $k-$رکوردها ارائه شده‌اند که متعاقبا آن‌ها را بیان می‌کنیم. در سرتاسر مقاله k ثابت

[۱] کارشناس ارشد، دانشگاه بیرجند

[۲] گروه آمار دانشگاه بیرجند

صحیح مثبت و k' یک مقدار حقیقی است. فرض کنید $\{X_n, n \geq 1\}$ دنباله‌ای از متغیرهای تصادفی مستقل و هم توزیع با تابع توزیع مشترک F باشد. در این مقاله از نماد $X_{r:n}$ برای نمایش r امین آماره مرتب در نمونه تصادفی به حجم n استفاده می‌کنیم. در این صورت $X_{n-k+1:n}$ نشان دهنده k امین مشاهده از نظر بزرگی و $\{X_{n-k+1:n}\}$ نسبت به k، دنباله‌ای نزولی می‌باشد. به عنوان نمونه برای ۱۰ مشاهده

۵.۴، ۶، ۴.۲، ۳.۶، ۵.۲، ۷، ۳.۱، ۴.۳، ۹، ۷.۵

دنباله $k-$امین مشاهده از نظر بزرگی،

$X_{n-k+1:n}$: ۹،۷.۵،۶،۵.۴،۵.۲،۴.۳،۳.۶،۳.۱،۳.۶،۳.۱

دنباله‌ای نزولی می‌باشد.

تعریف ۱.۲. *فرض کنید $\{X_{n:n+k-1}\}$ دنباله‌ای از آماره‌های مرتب مشاهدات iid می‌باشد. با حذف مقادیر تکراری دنباله اکیدا صعودی از مقادیر $k-$ رکوردها حاصل می‌شود. دنباله $\{L_n^{(k)}\}$ را به طور بازگشتی و با ضابطه زیر تعریف می‌کنیم.*

$$L_1^{(k)} = 1,$$

$$L_{n+1}^{(k)} = \min\{j : j > L_n^{(k)}, X_{j:j+k-1} > X_{L_n^{(k)}:L_n^{(k)}+k-1}\},$$

$n \geq 1$. دنباله $\{L_n^{(k)}, n \geq 1\}$ را دنباله زمان‌های $k-$ رکورد بالای قوی و $\{S_n^{(k)}, n \geq 1\}$ که $S_n^{(k)} = X_{L_n^{(k)}:L_n^{(k)}+k-1}$ ، را مقادیر $k-$رکوردهای بالای قوی می‌نامیم.

تبصره۱: واژه قوی را به این دلیل به کار می‌بریم که بین تعاریف مختلف تمیز قائل شویم. همچنین از این پس از واژه "بالا" استفاده نمی‌کنیم و هر جا که صحبت از $k-$ رکورد باشد مقصودمان $k-$ رکورد بالاست، مگر آن که خلاف آن به صراحت قید گردد.
برای مشاهدات فوق و به ازاء $k = 2$ مقادیر ۲-امین داده‌ها در جدول ۱ آورده شده‌اند. همچنین برای این مشاهدات زمان‌ها و مقادیر $k-$رکوردها ($k = 2$) در جدول ۲ آورده شده‌اند.

تعریف ۲.۲. *تحت شرائط تعریف ۱ دنباله $\{T_n^{(k)}\}$ که به صورت بازگشتی با ضابطه*

$$T_1^{(k)} = k,$$

$$T_{n+1}^{(k)} = \min\{j : j > T_n^{(k)}, \quad X_j > X_{T_{n-k+1}^{(k)}:T_n^{(k)}}\}$$

تعریف می‌گردد را دنباله زمان‌های $k-$رکورد و دنباله $\{R_n^{(k)}\}$، $R_n^{(k)} = X_{T_{n-k+1}^{(k)}:T_n^{(k)}}$ ، را دنباله مقادیر $k-$رکورد می‌نامیم.

آرنولد و همکاران(۱۹۹۸)، $\{R_n^{(k)}\}$ را دنباله $k-$رکوردهای نوع ۲ نامیدند. برخی از مولفین تعریف دوم را به عنوان تعریف $k-$رکوردها پذیرفته‌اند اما در بررسی خواص آن‌ها در واقع خواص $k-$رکوردهای قوی را بدست آورند. پس از ارائه سایر تعاریف، با مثالی هم ارز نبودن این تعاریف تشریح می‌گردد.

تعریف ۳.۲. *دنباله $\{U_n^{(k)}\}$ به طور بازگشتی به صورت زیر تعریف می‌شود:*

$$U_1^{(k)} = k$$

$$U_{n+1}^{(k)} = \min\{\mathrm{j} : \mathrm{j} > U_n^{(k)} \;\;, X_j \geq X_{U_{n-k+1}^{(k)}:U_n^{(k)}},\}$$

دنباله $\{U_n^{(k)}\}$ را دنباله زمان‌های $k-$رکوردهای ضعیف و دنباله $W_n^{(k)} = X_{U_{n-k+1}^{(k)}:U_n^{(k)}}$ ، را دنباله مقادیر $k-$رکورد ضعیف می‌نامیم.

تعریف ۴.۲. *فرض کنید*

$$\rho_n = \#\{j \leq n, X_n \leq X_j\}$$

که $\#$ نشان دهنده "تعداد عناصر مجموعه" می‌باشد. گوییم X_n یک $k-$رکورد از نوع ۱ است اگر $\rho_n = k$ هنگامی که $n \geq k$. مقادیر $k-$رکوردهای نوع ۱ را به $n \geq k, V_n^{(k)}$ نمایش می‌دهیم.

در مثال زیر و برمبنای داده‌های دمبنیسکا و بلازکیز (۲۰۰۵) چهار تعریف را با هم مقایسه می‌کنیم.

مثال ۵.۲. *فرض کنید $k = 2$. سطر اول جدول ۳ مقادیر متغیر صحیح مثبت n، شماره مشاهده داده و سطر دوم مقادیر داده‌ها را نشان می‌دهد. سطرهای سوم، چهارم و پنجم به ترتیب به زمان‌ها و مقادیر $k-$رکوردهای قوی اختصاص یافته‌اند. زمان‌ها و مقادیر $k-$رکوردها را در سطرهای ۶ الی ۸ و زمان‌ها و مقادیر $k-$رکوردهای ضعیف را در سطرهای ۹، ۱۰ و ۱۱ گنجانده‌ایم. سطرهای دوازدهم، سیزده ام و چهارده ام نشان دهنده زمان‌ها و مقادیر $k-$رکوردهای نوع ۱ می‌باشند.*

نزروف(۲۰۰۲) تعاریف اول و دوم را معادل می‌داند(ص ۸۳ روابط ۳.۱۹ و ۴.۱۹). از اینکه اگر

$$X_{j:j+k-1} > X_{L_n^{(k)}:L_n^{(k)}+k-1}$$

آنگاه

$$X_j > X_{T_{n-k+1}^{(k)}:T_n^{(k)}}$$

لذا هر زمان که $k-$رکورد قوی مشاهده شود، $k-$رکورد نیز رخ می‌دهد. همچنین بدیهی است که هر زمان که یک $k-$رکورد مشاهده می‌شود یک

$k-$ ضعیف نیز مشاهده می‌گردد.

به تفاوت تعاریف مختلف توجه کنید. هنگامی که $\{X_n, n \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی پیوسته باشد تعاریف ۱، ۲ و ۳هم ارز هستند. در این حالت مقادیر رکوردها $k-$رکوردها (به ازاء $k =$ ۱ می‌باشند.

به سهولت می‌توان نشان داد که اگر φ تابعی اکیدا صعودی $Y_i = \varphi(X_i)$، $\{T_n\}$ و $\{N_n\}$ به ترتیب زمان‌ها و تعداد $k-$رکوردهای وابسته به دنباله X_i و $\{T'_n\}$ و $\{N'_n\}$ زمان ها و تعداد $k-$رکوردهای وابسته به دنباله Y_i باشند، آنگاه

$$T'_n \stackrel{d}{=} T_n, \qquad N'_n \stackrel{d}{=} N_n$$

گزاره ۶.۲. *تحت شرط پیوستگی توزیع F مقادیر $k-$رکوردها با مقادیر رکوردها از توزیع*

$$G(x) = (۱ - (۱ - F(x))^k)$$

هم توزیع می‌باشند.

اثبات. نخست توزیع $R_۱^{(k)}$، اولین $k-$رکورد از توزیع F را به دست می‌آوریم.

$$\begin{aligned} P(R_۱^{(k)} \leq x) &= ۱ - P(R_۱^{(k)} > x) \\ &= ۱ - P(X_۱ > x, ..., X_k > x) \\ &= ۱ - (P(X_۱ > x))^k \\ &= (۱ - (۱ - F(x))^k), \end{aligned}$$

یعنی مقدار اولین $k-$رکورداز توزیع F با مقدار اولین رکورد از توزیع G هم توزیع می‌باشد.

فرض کنید $H = F^{-۱}(۱ - e^{-x})$. از اینکه F اکیدا صعودی است نتیجه می‌شود که H نیز اکیدا صعودی است. بنابراین اگر $E_n^{(k)}$ نشان دهنده دنباله $k-$رکوردها از توزیع نمایی با میانگین یک باشد، آنگاه

$$(R_۱^{(k)}, ..., R_n^{(k)}) \stackrel{d}{=} (H(E_۱^{(k)}), ..., H(E_n^{(k)})) \qquad (۱)$$

همچنین از خاصیت فقدان حافظگی توزیع نمایی نتیجه می‌شود که

$$E_۱^{(۱)} \text{ ، } E_۲^{(۱)} - E_۱^{(۱)} \text{ ، } ... \text{ ، } E_n^{(۱)} - E_{n-۱}^{(۱)}$$

متغیرهای تصادفی مستقل و همتوزیع((iid)) با توزیع نمایی با میانگین یک می‌باشند.

فرض کنید $Y_i = min\{X_{(i-۱)k+۱}, ..., X_{ik}\}$ در این صورت

$$\begin{aligned} P(Y_i \leq x) &= ۱ - P(Y_i > x) \\ &= ۱ - \textstyle\prod_{j=(i-۱)k+۱}^{ik} P(X_j > x) \\ &= (۱ - (-F(x))^k) \end{aligned}$$

از (۱) نتیجه می‌شود

$$R_n^{(۱)} \stackrel{d}{=} H(E_n^{(۱)}) = H(W_۱ + ... + W_n), \qquad (۲)$$

که در آن $W_i \sim iid$ متغیرهای تصادفی دارای توزیع نمایی با میانگین یک هستند. دزبدزیلا و کوپنسکی ((a)۱۹۷۶) نتیجه فوق را تعمیم و نشان دادند که برای هر $k \geq ۱$

$$E_n^{(k)} \stackrel{d}{=} \frac{W_۱ + ... + W_n}{k}, \quad n \geq ۱.$$

از طرف دیگر و بنا به (۱) برای هر $k \geq ۱$

$$R_n^{(k)} \stackrel{d}{=} H(\frac{W_۱ + ... + W_n}{k})$$

روابط مشابهی برای دنباله Y_n برقرار و داریم

$$R_{Y_n}^{(۱)} \stackrel{d}{=} H^*(W_۱ + ... + W_n).$$

که در آن $R_{Y_n}^{(۱)}$ دنباله مقادیر رکوردها وابسته به Y_iها و $H^* = G^{-۱}(۱ - e^{-۱})$ می‌باشد. از (۱) و (۲) نتیجه می‌شود که

$$\begin{aligned} R_{Y_n}^{(۱)} &\stackrel{d}{=} H^*(W_۱ + ... + W_n) \\ &\stackrel{d}{=} H(\frac{W_۱ + ... + W_n}{k}) \\ &\stackrel{d}{=} E_n^{(k)} \stackrel{d}{=} R_n^{(k)} \end{aligned}$$

□

۳ رفتار حدی فرایند شمارش $k-$رکوردها

برای اثبات قضایای اصلی به نتایج مقدماتی زیر نیاز داریم. اثبات لم‌های ۱ و ۲ را می‌توانید در گات (۲۰۰۵) (ص ۲۸۷ و ۱۶۶) بیابید.

لم ۱.۳. *فرض کنید $X_۱, X_۲, ...$ دنباله‌ای از متغیرهای تصادفی مستقل و به طور یکنواخت کراندار باشد آنگاه:*

$$\sum_{n=۱}^{\infty}(X_n - E(X_n))$$

همگرا $a.s.$

$$\Longleftrightarrow \qquad \sum_{n=۱}^{\infty} Var(X_n) < \infty.$$

لم ۲.۳. *فرض کنید $X_۱, X_۲, ...$ و $Y_۱, Y_۲, ...$ دو دنباله از متغیرهای تصادفی باشند به قسمی که*

$$X_n \xrightarrow{a.s.} X \quad \text{و} \quad Y_n \xrightarrow{a.s.} Y$$

وقتی $n \to \infty$. آنگاه

$$X_n + Y_n \xrightarrow{a.s.} X + Y$$

وقتی $n \to \infty$.

تعریف ۳.۳. *فرض کنید $\{X_n, n \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی مستقل با واریانس‌های متناهی باشد. همچنین فرض کنید*

$$m_i = E(X_i) \qquad \sigma_i^۲ = Var(X_i)$$

$$M_n = \sum_{i=۱}^{n} EX_i, \qquad s_n^۲ = \sum_{i=۱}^{n} \sigma_i^۲$$

گوییم دنباله $\{X_n, n \geq ۱\}$ در شرائط لیاپانوف صدق می‌کند، اگر $\delta > ۰$ موجود باشد به قسمی که

$$\frac{\sum_{k=۱}^{n} E(|X_k - M_k|^{۲+\delta})}{s_n^{۲+\delta}} \longrightarrow ۰$$

شرح بیشتری در خصوص این تعریف و لم زیر را می‌توانید در کورالف و سینایی (۲۰۰۷)(ص ۱۳۴) ملاحظه کنید.

لم ۴.۳. *فرض کنید $\{X_n, n \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی مستقل با واریانس‌های متناهی باشد که در شرط لیاپانوف صدق می‌کند. آنگاه*

$$\frac{S_n - M_n}{s_n} \xrightarrow{d} N(۰, ۱),$$

$S_n = \sum_{k=۱}^{n} X_i$.

تعریف ۵.۳. *فرض کنید $\{X_n, n \geq ۱\}$ دنباله‌ای نامتناهی از متغیرهای تصادفی مستقل و همتوزیع مطلقا پیوسته باشد. فرآیند شمارشی $\{\mu_n^{(k)}, n \geq k\}$ به صورت زیر تعریف می‌شود:*

تعداد k−رکوردهای بین $X_n, ..., X_۱$

$$\mu_n^{(k)} = max\{m : T_m^{(k)} \leq n\}$$

مثال زیر می‌تواند به تشریح رفتارهای حدی k−رکوردها و تفاوت بین آن‌ها، به ازاء مقادیر مختلف k، کمک نماید.

مثال ۶.۳. *برای داده‌های زیر*

۰، ۱، ۲، ۱.۵، ۱.۶، ۱.۷، ۳، ۲.۵، ۲.۶، ۲.۷، ۳.۵، ...

مقادیر رکوردها را در سطر اول و مقادیر ۲ - رکوردها را درسطر دوم جدول ۴ نوشته‌ایم. این جدول ضرورت بررسی رفتار حدی $\mu_n^{(k)}$ را نشان می‌دهد.

گات(۲۰۰۲)رفتار حدی $\mu_n^{(k)}$ به ازاء $k = ۱$، را بررسی نمود. در این بخش برخی از روابط مورد استفاده و نتایج گات (۲۰۰۲) را برای k− رکوردها ثابت می‌کنیم.

لم ۷.۳. *فرض کنید*

$$I_{m,k} = \begin{cases} ۱, & \text{اگر در زمان مشاهده } m\text{ام یک} \\ & k\text{- رکورد رخ دهد} \\ ۰, & \text{در غیر این صورت} \end{cases}$$

یعنی

$$I_{m,k} = I_{(X_m > X_{m-k,m-۱})}$$

متغیرهای $m \geq k, I_{m,k}$ متغیرهای تصادفی گسسته با تابع احتمال

$$P(I_{m,k} = ۱) = \frac{k}{m}$$

می‌باشند

اثبات. فرض کنید $X_{۱:m-۱} < X_{۲:m-۱} < ... < X_{m-۱:m-۱}$ ، $m-۱$ آماره مرتب از مشاهدات می‌باشند. $X_m = y$ یک k−رکورد است اگر

$$X_{m-k:m-۱} < X_{m-k+۱:m-۱} < \cdots$$
$$< X_{m-۱:m-۱} < y$$

یا

$$X_{m-k:m-۱} < X_{m-k+۱:m-۱} < \cdots$$
$$< y < X_{m-۱:m-۱}$$
$$\vdots \qquad \vdots \qquad \vdots$$

یا

$$X_{m-k:m-۱} < y < X_{m-k+۱:m-۱} < \cdots$$
$$< X_{m-۱:m-۱}$$

پس به k حالت X_m می‌تواند یک k−رکورد باشد. این در حالی است که X_m به m حالت،

$$X_{۱:m-۱} < X_{۲:m-۱} < \cdots$$
$$< X_{m-۱:m-۱} < y$$
$$\vdots \qquad \vdots \qquad \vdots$$
$$y < X_{۱:m-۱} < \cdots$$
$$< X_{m-۱:m-۱}$$

مقدار خود را اختیار می‌کند. هر یک از حالت‌های فوق شانس برابری برای وقوع دارند، لذا

$$P(I_{m,k} = ۱) = \frac{k}{m}$$

برای اثبات مستقل بودن، فرض کنید $\{Y_m\}$ دنباله‌ای از متغیرهای تصادفی iid با تابع توزیع

$$G(x) = (۱ - (۱ - F(x))^k)$$

و I'_m متغیرهای تصادفی نشانگر زمان‌های رکورد متناظربا Y_m باشند، یعنی

$$I'_m = \begin{cases} ۱, & \text{اگر در زمان مشاهده } m\text{ام یک} \\ & k\text{- رکورد رخ دهد} \\ ۰, & \text{در غیر این صورت.} \end{cases}$$

اگر $k < k_۱ < k_۲ < \cdots < k_n$ مقادیر صحیح نامنفی باشند آنگاه بنا به گزاره ۱ مقادیر صحیح نامنفی $k < k'_۱ < k'_۲ < \cdots < k'_n$ وجود دارد به قسمی که

$$P(I_{k_۱,k} = ۱, \cdots I_{k_n,k} = ۱)$$

$$= P(I_{k'_۱} = ۱, \cdots I_{k'_n} = ۱)$$

$$= \prod_{j=۱}^{n} P(I_{k'_۱} = ۱) \qquad \text{رنی ۱۹۶۲}$$

$$= \prod_{j=۱}^{n} P(I_{k_j,k} = ۱) \qquad \text{بنا به گزاره ۱}$$

لذا $\{I_{m,k}\}$ دنباله‌ای از متغیرهای تصادفی مستقل می‌باشد

□

مثال ۸.۳. *فرض کنید ۴ $= m$. از نماد $X(i) = j$ برای نشان دادن این که j−امین مشاهده i−امین آماره ترتیبی می‌باشد استفاده کرده‌ایم. همچنین فرض کنید X_iها متغیرهای تصادفی پیوسته باشند. جدول ۵ نشان دهنده حالت‌های مختلفی است که $X(i)$ها می‌توانند داشته باشند. هنگامی که j−امین مشاهده امین ۲-رکورد باشد مقدار آن را به $R_i^{(۲)} = j$ نشان داده‌ایم. ملاحظه می‌شود که*

$$P(I_{۲,۲} = ۱) = \frac{n(I_{۲,۲} = ۱)}{n(S)} = ۱$$

قضیه ۹.۳. *برای فرآیند شمارشی $\mu_n^{(k)}$ داریم*

(الف) $m_{n,k} = E(\mu_n^{(k)})$

$$= k\log(n-k+۱) + k'\gamma + o(۱);$$

ب) $Var(\mu_n^k) = k\log(n-k+۱)$

$$+ k'\gamma + o(۱);$$

ج) $\frac{(\mu_n^{(k)})}{k\log(n-k+۱)} \xrightarrow{a.s.} ۱;$

د) $\frac{۱}{k\log(n-k+۱)} \sum_{m=k}^{n} (I_{m,k} - \frac{k}{m}) \xrightarrow{a.s.} ۰$

و) $\frac{(\mu_n^{(k)}) - E(\mu_n^{(k)})}{\sqrt{k\log(n-k+۱)}} \xrightarrow{d} N(۰, ۱).$

که k' مقداری حقیقی است که در مکان‌های مختلف ممکن است با هم متفاوت باشند و γ ثابت اویلر و برابر ۰/۵۷۷۲ است.

اثبات. (الف). با توجه به تعریف

$$\mu_n^{(k)} = \sum_{m=k}^{n} (I_{m,k}), \qquad m \ge k$$

به سادگی داریم

$$\begin{aligned} E(\mu_n^{(k)}) &= \textstyle\sum_{m=k}^{n} (\frac{k}{m}) \\ &= \textstyle\sum_{m=۱}^{n} (\frac{k}{m}) - \sum_{m=۱}^{k-۱} (\frac{k}{m}) \\ &= k\log(n-k+۱) + k'\gamma + o(۱) \end{aligned}$$

(آپوستل ۱۳۶۷ ص ۲۷۹).

(ب) با توجه به گزاره (۱) و واریانس متغیرهای برنولی و از اینکه $\sum_{m=۱}^{\infty} (\frac{۱}{m})^۲ = \frac{\pi^۲}{۶}$ ، داریم:

$$\begin{aligned} Var(\mu_n^{(k)}) &= \textstyle\sum_{m=k}^{n} \frac{k}{m}(۱ - \frac{k}{m}) \\ &= \textstyle\sum_{m=k}^{n} \frac{k}{m} - \sum_{m=k}^{n} (\frac{k}{m})^۲ \\ &= k\log(n-k+۱) + k'\gamma + o(۱) \end{aligned}$$

اثبات قسمت (ج) بدیهی است.

(د). می‌دانیم $\{I_{m,k} - \frac{k}{m}, m \ge k\}$ دنباله‌ای از متغیرهای تصادفی به طور یکنواخت کراندار و مستقل با میانگین صفر و واریانس‌های متناهی $\frac{k}{m}(۱ - \frac{k}{m})$ است، بنابراین

$$\{\frac{I_{m,k} - \frac{k}{m}}{k\log m}, m \ge k\}$$

نیز دنباله‌ای از متغیرهای تصادفی به طور یکنواخت کراندار و مستقل با واریانس‌های متناهی

$$(\frac{۱}{k\log m})^۲ (\frac{k}{m}(۱ - \frac{k}{m}))$$

و میانگین صفر بوده و

$$\begin{aligned} \sum_{m=k}^{\infty} Var(\frac{I_{m,k} - \frac{k}{m}}{k\log m}) &= \sum_{m=k}^{\infty} (\frac{\frac{k}{m}(۱ - \frac{k}{m})}{(k\log m)^۲} \\ &\le \sum_{m=۲}^{\infty} \frac{۱}{km(\log m)^۲} < \infty. \end{aligned}$$

یادآوری می‌کنیم که اگر $\{x_n\}$ دنباله‌ای از اعداد حقیقی و $\{a_n\}$ دنباله‌ای از مقادیر مثبت باشند به قسمی که $a_n \nearrow \infty$ آنگاه

$$\sum_{n=k}^{\infty} \frac{x_n}{a_n} < \infty \Longrightarrow \frac{۱}{a_n} \sum_{j=k}^{n} x_j \to ۰ \qquad (۳)$$

رابطه (۳) از لم شناخته شده کرونکر به ازاء $x_j = ۰$ ، برای $j < k$ ، نتیجه می‌شود. بنابراین و بنا به لم ۱، $\sum_{m=k}^{\infty} \frac{I_{m,k}-\frac{k}{m}}{k(\log m)}$ همگرا $a.s.$ است. لذا بنا به رابطه (۳)داریم

$$\frac{۱}{k(\log n)} \sum_{m=k}^{n} (I_{m,k} - \frac{k}{m}) \xrightarrow{a.s.} ۰$$

$$\Rightarrow \frac{۱}{k\log(n-k+۱)} \sum_{m=k}^{n} (I_{m,k} - \frac{k}{m}) \xrightarrow{a.s.} ۰$$

برای اثبات قسمت (ه) ابتدا نشان می‌دهیم که شرائط لیاپانوف برقرار است. از قسمت (ب) داریم

$$S_n^۳ = (k(\log(n-k+۱) + k'\gamma + o(۱))^{۳/۲}$$

از طرف دیگر

$$E|I_{m,k} - E(I_{m,k})|^۳ = E|I_{m,k} - \frac{k}{m}|^۳$$
$$= |۱ - \frac{k}{m}|^۳ \frac{k}{m} + (\frac{k}{m})^۳ (۱ - \frac{k}{m})$$
$$\leq \frac{k}{m} + (\frac{k}{m})^۳$$

بنابراین

$$\frac{\sum_{m=k}^{n} E|I_{m,k} - E(I_{m,k})|^۳}{S_n^۳}$$
$$\leq \frac{\sum_{m=k}^{n} (\frac{k}{m} + (\frac{k}{m})^۳)}{(k(\log(n-k+۱) + k'\gamma + o(۱))^{۳/۲}}$$
$$= \frac{k(\log(n) - \log(k)) + ۲\gamma + k'k + o(۱)}{(k\log(n-k+۱) + k'\gamma + o(۱))^{۳/۲}}$$
$$\longrightarrow ۰$$

بنابراین شرایط لیاپانوف برقرار است. پس بنا به لم ۳ برقراری قسمت (ه) ثابت می‌شود. □

تشکر و قدردانی: نویسندگان از داوران محترم به دلیل دقت نظر و راهنمایی‌هایی ارزنده شان کمال تشکر را دارند.

جدول ۱: مقادیر ۲-امین داده‌ها از نظر بزرگی

n	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰
X_n	۵.۲	۶	۹	۴.۳	۳.۱	۷	۳.۶	۲	۴	۷.۵
$X_{n:n+۲-۱}$		$X_{۱:۲}$	$X_{۲:۳}$	$X_{۳:۴}$	$X_{۴:۵}$	$X_{۵:۶}$	$X_{۶:۷}$	$X_{۷:۸}$	$X_{۸:۹}$	$X_{۹:۱۰}$
		۵.۲	۶	۶	۶	۷	۷	۷	۷	۷.۵

جدول ۲: زمان‌ها و مقادیر ۲-رکوردها

n	$L_n^{(۲)}$	$S_n^۲$
۱	$L_۱^{(۲)} = ۱$	۵.۲
۲	$L_۲^{(۲)} = min\{j : j > ۱, X_{j:j+۱} > ۵/۲\} = ۲$	۶
۳	$L_۳^{(۲)} = min\{j : j > ۱, X_{j:j+۱} > ۶\} = ۵$	۷
۴	$L_۴^{(۲)} = min\{j : j > ۵, X_{j:j+۱} > ۷\} = ۹$	۷.۵

جدول ۳: مقایسه‌ای بین ۲-رکوردها

۱	n	۱	۲	۳	۴	۵	۶	۷	۸
۲	X_n	۰	۱	۱	۱	۰	۴	۰	۳
۳	$L_n^{(2)}$	۱	۲	۷					
۴	$\{j > L_n^{(2)}, X_{j:j+1} > X_{L_n^{(2)}:L_n^{(2)}+1}\}$	$\{j > 1, X_{j:j+1} > X_{1:2}\} = \{2, \cdots, 7\}$	$\{j > 2, X_{\mathrm{j:j+1}} > 1\} = \{7\}$						
۵	$S_n^{(2)} = X_{L_n^2:L_n^2+1}$	$X_{1:2} = 0$	$X_{2:3} = 1$	$X_{7:8} = 3$					
۶	$T_n^{(2)}$	۲	۳	۶	۸				
۷	$\{j > T_n^{(2)}, X_j > X_{T_{n-1}^{(2)}:T_n^{(2)}}\}$	$\{j > 2, X_j > 0\} = \{3, 4, 6, 8\}$	$\{\mathrm{j} > 3, X_j > 1\} = \{6, 8\}$	$\{\mathrm{j} > 6, X_j > 1\} = \{8\}$					
۸	$R_n^{(2)} = X_{T_{n-1}^{(2)}:T_n^{(2)}}$	$X_{1:2} = 0$	$X_{2:3} = 1$	$X_{5:6} = 1$	$X_{7:8} = 3$				
۹	$U_n^{(2)}$	۲	۳	۴	۶	۸			
۱۰	$R_n^{(2)} = \{j > U_n^2, X_j \geq X_{U_{n-1}^{(2)}:U_n^{(2)}}\}$	$\{3, 4, 5, 6, 7, 8\}$	$\{4, 6, 8\}$	$\{6, 8\}$	$\{8\}$				
۱۱	$W_n^{(2)} = X_{U_{n-1}^{(2)}:U_n^{(2)}}$	$X_{1:2} = 0$	$X_{2:3} = 1$	$X_{3:4} = 1$	$X_{5:6} = 1$	$X_{7:8} = 3$			
۱۲	$\{j \leq n, X_n \leq X_j\}$	$\{j \leq 1\ 1 \leq X_j\}$	$\{j \leq 2, X_1 \leq X_1\}$	$\{j \leq 3, 1 \leq X_j\}$	$\{j \leq 4, 1 \leq X_j\}$	$\{j \leq 5, 0 \leq X_j\}$	$\{j \leq 6, 4 \leq X_j\}$	$\{j \leq 7, 0 \leq X_j\}$	$\{j \leq 8, 3 \leq X_j\}$
۱۳	$\{X_j \geq X_n\}$	$\{X_1\}$	$\{X_2\}$	$\{X_2, X_3\}$	$\{X_2, \ldots, X_4\}$	$\{X_1, \ldots, X_5\}$	$\{X_6\}$	$\{X_1, \ldots, X_7\}$	$\{X_6, X_8\}$
۱۴	ρ_n	۱	۱	۲	۳	۵	۱	۷	۲
۱۵	$V_n^{(2)}$	$X_3 = 1$	$X_8 = 3$						

جدول۴: مقایسه رفتار رکوردها و ۲- رکوردها

n	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	$\cdots$
$R_n^{(1)}$	۰	۱	۲	۳	۳.۵	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$R_n^{(2)}$	۰	۱	۱.۵	۱.۶	۱.۷	۲	۲.۵	۲.۶	۲.۷	۳	$\cdots$

جدول۵: مقایسه‌ای بین ۲- رکوردها

$R_{۳}^{(۲)}$	$R_{۲}^{(۲)}$	$R_{۱}^{(۲)}$	$X(۴)$	$X(۳)$	$X(۲)$	$X(۱)$	$R_{۳}^{(۲)}$	$R_{۲}^{(۲)}$	$R_{۱}^{(۲)}$	$X(۴)$	$X(۳)$	$X(۲)$	$X(۱)$
-	۳	۲	۴	۲	۱	۳	۳	۲	۱	۴	۳	۲	۱
-	-	۱	۲	۴	۱	۳	۴	۲	۱	۳	۴	۲	۱
-	۱	۲	۴	۱	۲	۳	۲	۳	۱	۴	۲	۳	۱
-	۱	۲	۱	۴	۲	۳	-	۲	۱	۲	۴	۳	۱
-	-	۱	۲	۱	۴	۳	۴	۳	۱	۳	۲	۴	۱
-	-	۱	۱	۲	۴	۳	-	۳	۱	۲	۳	۴	۱
۴	۳	۱	۳	۲	۱	۴	۳	۱	۲	۴	۳	۱	۲
-	۳	۱	۲	۳	۱	۴	۴	۱	۲	۳	۴	۱	۲
۳	۴	۲	۳	۱	۲	۴	-	۲	۱	۴	۱	۳	۲
-	۳	۲	۱	۳	۲	۴	-	۳	۲	۱	۴	۳	۲
-	-	۲	۲	۱	۳	۴	-	۴	۱	۳	۱	۴	۲
-	-	۲	۱	۲	۳	۴	-	۳	۱	۱	۳	۴	۲

مراجع

[۱] . آپوستل، تام م، ۱۳۶۹، آنالیز ریاضی، مترجم، علی اکبر عالم زاده، موسسه انتشارات علمی دانشگاه صنعتی شریف.

[2] Ahmadi, J.; Razmkhah, M. and Balakrishnan, N. (2009). Current k-records and their use in distribution-free confidence intervals, *Stat. and Prob. Lett.*, **79**, 29-37.

[3] Arnold, B.C.; Balakrishnan, N. and Nagaraja, H.N. (1998). *Records*. John Wiley Sons, New York.

[4] Dembinska, A. and Lopez-Blazquez, F. (2005). kth records from discrete distributions, *Stat. and Prob. Lett.*, **71,** 203–214.

[5] Dziubdziela, W. and Kopocinsky, B. (1976(a)). *Limit theorems for linear combinations of order statistics*, Lectures Notes in Math., Vol. 550, Springer-Verlag, PP. 63-79.

[6] Dziubdziela, W. and Kopocinsky, B. (1976(b)). Limiting properties of the k-th record values, *Zastos. Mat.***15,** 187–190.

[7] Gut, A. (2005). *Probability: A Graduate Course*. Springer texts in Statistics.

[8] Hofmann, G. and Balakrishnan, N. (2004). *Fisher Information in k-Records,* The Institute of Statistical Mathematics, Vol. 56, No. 2, 383-396.

[9] Kampas, U. (1995). *A Concept of Generalize Order Statistics.* B.G. Teubner Stuttgart.

[10] Koralov, L. B. and Sinai, Y G. (2007). *Theory of Probability and Random Processes*. Springer-Verlag, New York Inc.

[11] Nagaraga, H.N. (1988). Record values and related Statistics: A review. *Comm. Stat. Theory Meth.* **17,** 2223-2238

[12] Nevzorov, v.B. (2001). *Records: Mathematical Theory*. American Mathematical Society.

[13] Renyi,A., (1962). *On the extreme elements of observation, Mat III.* Oszt.Kostl.

[14] Resnick, S. I., (1999). *A Probability Path*. Birkhauser, Boston.