پیش‌بینی ناپارامتری داده‌های ترتیبی در مسائل دو نمونه‌ای

جعفر احمدی (مدعو)
گروه آمار، دانشگاه فردوسی مشهد

در این مقاله ضمن معرفی مختصری از انواع داده‌های ترتیبی مسئله پیش‌بینی در آنها مطرح می‌شود. پیش بینی حوادث آینده یکی از مسائل مهم در برنامه‌ریزی زندگی روزمره است، بنابراین مطالعه آن حائز اهمیت است، اکثر این پدیده‌ها در قالب مدل‌های آماری قابل بررسی هستند. در اینجا فرض بر این است که پدیده‌های مشاهده شده از نوع متغیرهای تصادفی پیوسته‌اند و ماهیت مسئله طوری است که یا ثبت داده‌ها بر اساس یک ترتیب منظم انجام شده و یا پژوهشگر علاقه‌مند است برای مطالعه رفتار احتمالی متغیرها آنها را بصورت صعودی یا نزولی مرتب نماید. تاکید این مطالعه بر پیش‌بینی دو نمونه‌ای است، که بصورت ناپارامتری انجام می‌شود. با استفاده از داده‌های ترتیبی مشاهده شده حاصل از نمونه اول فواصل پیش بینی آزاد توزیع برای داده‌های ترتیبی نمونه آینده (مشاهده نشده) ارائه می‌گردد. تفاوت کار با مجموعه تحقیقات انجام شده در این است که نوع داده‌های ترتیبی دو نمونه متفاوت و تنها با فرض پیوسته بودن توزیع آماری جامعه ضرایب پیش‌بینی بدست آمده دقیق و آزاد توزیع هستند.

واژه‌های کلیدی: آماره‌های مرتب، آماره‌های رکوردی، سانسور فزآینده، فاصله پیش‌بینی، پوشش رکوردی.

۱ مقدمه

فرض کنید $X_1, X_2, \cdots, X_n$ نمونه‌ای تصادفی به حجم n از متغیرهای تصادفی پیوسته از جامعه‌ای با تابع توزیع و تابع چگالی احتمال به ترتیب $F(x)$ و $f(x)$ باشد. اگر این متغیرها را بصورت صعودی مرتب شوند و مقادیر مرتب شده را با $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ نشان دهیم، در این صورت $X_{r:n}$ را rامین آماره مرتب نمونه فوق گویند و تابع توزیع آن بصورت

زیر داده می‌شود

$$F_{r:n}(x)=\sum_{i=r}^{n}\binom{n}{i}\{F(x)\}^{i}\{1-F(x)\}^{n-i}.$$

برای مطالعه خواص آماره‌های مرتب و کاربردهای آن می‌توان به آرنولد و همکاران (۱۹۹۲) و یا دیوید و ناگاراجا(۲۰۰۳) مراجعه نمود.
حال فرض کنید $\{X_n, n\geq 1\}$ دنباله‌ای از متغیرهای تصادفی پیوسته از جامعه‌ای با تابع توزیع احتمال و تابع چگالی احتمال به ترتیب $F(x)$ و $f(x)$ باشد، در این صورت X_k را یک رکورد بالا (پایین) گویند هر گاه $X_k>\max\{X_1,...,X_{k-1}\}$ ($X_k<\min\{X_1,...,X_{k-1}\}$).
در اینجا nامین رکورد بالا را با U_n نشان می‌دهیم،($U_1=X_1$) در اینصورت تابع توزیع U_n بصورت زیر داده می‌شود

$$F_{U_n}(x)=1-\bar{F}(x)\sum_{k=0}^{n-1}\frac{\{-\log\bar{F}(x)\}^{k}}{k!},$$

که در آن $\bar{F}(x)=1-F(x)$ تابع بقای X است. با جایگذاری $F(x)$ بجای $\bar{F}(x)$ در عبارت بالا، تابع بقای nامین رکورد پایین حاصل می‌شود. برای مطالعه خواص رکوردها و کاربردهای آن می‌توان به آرنولد و همکاران (۱۹۹۸) مراجعه نمود.

در مسایل پیش‌بینی علاقه‌مندیم با استفاده از مقادیر مشاهده شده بتوانیم حوادث آینده که هنوز مشاهده نشده‌اند را تخمین بزنیم. این مسئله می‌تواند به دو صورت انجام گیرد که هر دو روش توسط پژوهش‌گران زیادی مورد مطالعه قرار گرفته است:
الف) متغیرهای که باید پیش بینی شوند دنباله مقادیر مشاهده شده از همان نمونه هستند، در نتیجه به متغیرهای قبلی وابسته‌اند. این روش در مقالات آماری به**پیش بینی یک نمونه‌ای** معروف است.
ب) متغیرهای که باید پیش‌بینی شوند از یک نمونه آینده و مستقل از مقادیر مشاهده شده هستند. این روش در مقالات آماری به**پیش‌بینی دو نمونه‌ای** معروف است.

پیش‌بینی رکوردها برحسب رکوردها و آماره‌های مرتب برحسب آماره‌های مرتب توسط محققین زیادی انجام شده است. در این خصوص می‌توان به مجموعه تحقیقات انجام شده توسط احسان‌اله (۱۹۸۰)، دانسمر(۱۹۸۳)، راقب و ناگاراجا(۱۹۹۵)، کامینسکی و نلسن(۱۹۹۸)، راقب و بالاکریشنان (۲۰۰۸) و راقب (۲۰۰۹) اشاره نمود. مؤلفین تمام موارد فوق، داده‌های ترتیبی هر دو نمونه را از یک نوع در نظر گرفته‌اند.
اخیراً احمدی و بالاکریشنان(۲۰۱۰) ایده جدیدی در زمینه پیش‌بینی دو نمونه‌ای داده‌های ترتیبی مطرح نموده و فواصل پیش‌بینی ناپارامتری برای آماره‌های مرتب برحسب رکوردها و

همچنین برای رکوردها بر حسب آماره‌های مرتب بدست آورده‌اند، که در بخش بعدی برخی از نتایج آنها می‌آید. در همین راستا احمدی و میرمصطفائی(۲۰۰۹) مسئله را بصورت پارامتری در نظر گرفته و پیش بینی کننده‌های نقطه‌ای برای آماره‌های مرتب و رکوردها تحت مدل نمایی بدست آورده‌اند. در اینجا قصد داریم نتایج احمدی و بالاکریشنان(۲۰۱۰) را برای سایر داده‌های ترتیبی تعمیم دهیم، که بطور جداگانه در بخش‌های بعدی ارائه خواهند شد.

۲ پیش‌بینی آماره‌های مرتب و رکوردها

در این بخش برخی از نتایج بدست آمده توسط احمدی و بالاکریشنان(۲۰۱۰) بطور مختصر ارائه می‌شود.

الف) **پیش بینی آماره‌های مرتب:** اگر فرض کنیم $Y_{1:n} \leq Y_{2:n} \leq \cdots \leq Y_{n:n}$ آماره‌های مرتب حاصل از نمونه تصادفی آینده به حجم n از جامعه Y_iها باشند، در این صورت فاصله پیش بینی بر اساس رکوردهای مشاهده شده برای $Y_{i:n}$ها بصورت زیر داده می‌شود، بطوریکه پوشش احتمال آن به توزیع جامعه بستگی ندارد. برای $1 \leq i < j$

$$\begin{aligned}\alpha_1(i,j;k,n) &= P(U_i \leq Y_{k:n} \leq U_j)\\ &= k\binom{n}{k}\sum_{r=i}^{j-1}\sum_{s=0}^{k-1}\binom{k-1}{s}\frac{(-1)^s}{(n-k+s+2)^{r+1}}.\end{aligned}$$

اگر nامین رکورد پایین را با L_n نشان دهیم، $(L_1 = X_1)$ آنگاه برای $1 \leq i < j$،

$$P(L_j \leq Y_{k:n} \leq L_i) = P(U_i \leq Y_{n-k+1:n} \leq U_j).$$

در اینصورت برای $1 \leq i < j$، بر اساس نتایج احمدی و بالاکریشنان (۲۰۱۰) $(L_j,\ L_i)$ یک فاصله پیش بینی برای $Y_{k:n}$ با ضریب پیش‌بینی زیر است

$$\alpha_2(i,j;k,n) = k\binom{n}{k}\sum_{r=i}^{j-1}\sum_{s=0}^{n-k}\binom{n-k}{s}\frac{(-1)^s}{(k+s+1)^{r+1}}.$$

الف) **پیش‌بینی رکوردها:** حال فرض کنید $Y_{1:n} \leq Y_{2:n} \leq \cdots \leq Y_{n:n}$ آماره‌های مرتب نمونه تصادفی به حجم n از جامعه F باشند، و بخواهیم بر اساس $Y_{i:n}$های مشاهده شده رکوردهای

مستخرج از یک دنباله آینده از جامعه F را تخمین بزنیم. در این صورت $(Y_{i:n}, Y_{j:n})$ برای $1 \leq i \leq j \leq n$ یک فاصله پیش‌بینی برای kامین رکورد بالا، U_k، با ضریب پیش‌بینی زیر است که به F بستگی ندارد.

$$\alpha_3(k; i, j, n) = \sum_{t=i}^{j-1} \sum_{r=0}^{t} \binom{n}{t} \binom{t}{r} \frac{(-1)^r}{(n+r-t+1)^k}.$$

با مفروضات فوق $(Y_{i:n}, Y_{j:n})$ برای $1 \leq i \leq j \leq n$ یک فاصله پیش‌بینی برای kامین رکورد پایین، L_k، با ضریب پیش‌بینی زیر است.

$$\alpha_4(k; i, j, n) = \sum_{t=i}^{j-1} \sum_{r=0}^{n-t} \binom{n}{t} \binom{n-t}{r} \frac{(-1)^r}{(t+r+1)^k}.$$

با معلوم بودن ضریب پیش‌بینی از قبل داده شده می‌توان i و j مناسب را بدست آورد.

۳ سانسور فزاینده

در اینجا ابتدا بطور مختصر چگونگی طرح سانسور فزاینده را شرح می‌دهیم. فرض کنید در یک آزمایش طول عمر n مؤلفه در زمان $t = 0$ تحت آزمایش قرار گیرند. با از کار افتادن اولین مؤلفه تعداد R_1 واحد از واحدهای سالم بطور تصادفی از آزمایش خارج می‌شوند. در زمان از کارافتادگی مؤلفه دوم تعداد R_2 واحد از واحدهای سالم بطور تصادفی از آزمایش خارج می‌شوند و آزمایش به همین روال ادامه می‌یابد تا زمان خرابی mامین مؤلفه که در این لحظه تمام واحدهای سالم باقی مانده که برابر با $R_m = n - m - \sum_{i=1}^{n-1} R_i$ است از آزمایش کنار گذاشته می‌شوند. در اینجا فرض بر این است که مقادیر m و تمام R_iها از قبل معلوم می‌باشند. این شیوه آزمایش به طرح سانسور فزاینده نوع II با $\widetilde{R} = (R_1, \cdots, R_m),\ m \leq n,$ معروف است و زمان خرابی اولین مؤلفه تا mامین مؤلفه را به ترتیب با نماد $X^{\widetilde{R}}_{1:m:n}$، $X^{\widetilde{R}}_{2:m:n}$ و ... $X^{\widetilde{R}}_{m:m:n}$ نشان می‌دهند. بدیهی است که $X^{\widetilde{R}}_{1:m:n} \leq X^{\widetilde{R}}_{2:m:n} \leq \cdots \leq X^{\widetilde{R}}_{m:m:n}$. برای سادگی در اینجا rامین آماره مرتب سانسور فزاینده نوع II را با $X_{r:m:n}$ نشان می‌دهیم.

حال اگر فرض کنیم $X_1, X_2, \cdots, X_n$ نمونه‌ای تصادفی از متغیرهای تصادفی پیوسته از جامعه‌ای با تابع توزیع احتمال و تابع چگالی احتمال به ترتیب $F(x)$ و $f(x)$ باشد. دراین

صورت تابع توزیع احتمال $X_{r:m:n}$ بصورت زیر داده می‌شود

$$F_{X_{r:m:n}}(x) = 1 - c_{r-1} \sum_{i=1}^{r} \frac{a_{i,r}}{\gamma_i} [1 - F(x)]^{\gamma_i}, \quad r = 1, \cdots, m,$$

که در آن برای $r = 1, \cdots, m$ ، $\gamma_r = m - r + 1 + \sum_{i=r}^{m} R_i$ ، $c_{r-1} = \prod_{i=1}^{r} \gamma_i$ و $a_{i,r} = \prod_{j=1, j\neq i}^{r} \frac{1}{\gamma_j - \gamma_i}$.

با مشتق‌گیری نسبت به x، تابع چگالی احتمال $X_{r:m:n}$ بدست می آید که

$$f_{X_{r:m:n}}(x) = c_{r-1} \sum_{i=1}^{r} a_{i,r} f(x) [1 - F(x)]^{\gamma_i - 1}, \quad r = 1, \cdots, m.$$

برای مطالعه در باره خواص آماره مرتب سانسور فزاینده و کاربردهای آن به بالاکریشنان و آگراوالا (۲۰۰۰) و بالاکریشنان (۲۰۰۷) مراجعه شود.

۴ k-رکورد ها

گاهی مواقع دومین ماکزیمم یا سومین ماکزیمم همانند رتبه‌ها در مسابقات علمی نیز حایز اهمیت است، در نتیجه آماره‌های k-رکوردی مطرح شده که به صورت‌های متفاوت توسط پژوهش‌گران تعریف شده است. در اینجا ما صورتی که در صفحه ۴۳ کتاب آرنولد و همکاران (۱۹۹۸) آمده است را در نظر می‌گیریم. با فرض $T_{1,k} = k$ و برای $n \geq 2$

$$T_{n,k} = \min\{j : X_j > X_{T_{n-1,k}-k+1:T_{n-1,k}}\}$$

آنگاه $U_{n,k} = X_{T_{n,k}-k+1:T_{n,k}}$، nامین k-رکورد بالا در دنباله X_iها است. بعنوان مثال فرض کنید دنباله زیر را داشته باشیم

$$71, 23, 57, 94, 84, 17, 67, 54, 45, 87, 74, 36, \ldots$$

در اینصورت دنباله ۲- رکوردها عبارتند از: ۲۳, ۵۷, ۷۱, ۸۴, ۸۷.

با فرض اینکه X_iها یک دنباله از متغیرهای تصادفی از جامعه‌ای با تابع توزیع $F(x)$ و تابع چگالی احتمال $f(x)$ باشد، آنگاه تابع چگالی احتمال $U_{n,k}$ بصورت زیر است

$$f_{U_{n,k}}(x) = \frac{k^n [-\log(1 - F(x))]^{n-1}}{(n-1)!} (1 - F(x))^{k-1} f(x).$$

۵ پیش بینی k-رکوردها و آماره‌های مرتب سانسور فزاینده

در این بخش نتایج احمدی و بالاکریشنان (۲۰۱۰) را برای داده‌های سانسور فزاینده و k-رکوردها تعمیم می‌دهیم.

الف) پیش بینی آماره‌های مرتب سانسور فزاینده بر حسب k-رکوردها: با توجه به فرضیات مقاله، برای $n \geq ۱$ متغیر تصادفی $U_{n,k}$ پیوسته و برای $۱ \leq i < j$ آماره $U_{i,k}$ بطور تصادفی کوچکتر از $U_{j,k}$ است، بنابراین

$$P(U_{i,k} \leq Y_{r:m:n} \leq U_{j,k}) = P(U_{j,k} \geq Y_{r:m:n}) - P(U_{i,k} \geq Y_{r:m:n}).$$

از طرفی داریم

$$\begin{aligned} P(U_{i,k} \geq Y_{r:m:n} &= \int_{-\infty}^{+\infty} P(U_{i,k} \geq y) dF_{Y_{r:M:n}}(y) \\ &= \int_{-\infty}^{+\infty} P(U_{i,k} \geq y) c_{r-۱} \sum_{t=۱}^{r} a_{t,r} f(y)[\bar{F}(y)]^{\gamma_t - ۱} dy \\ &= \int_{-\infty}^{+\infty} [\bar{F}(x)]^k \sum_{s=۰}^{i-۱} \frac{\{-k \log \bar{F}(x)\}^s}{s!} c_{r-۱} \sum_{t=۱}^{r} a_{t,r} f(y)[\bar{F}(y)]^{\gamma_t - ۱} dy \\ &= c_{r-۱} \sum_{t=۱}^{r} \sum_{s=۰}^{i-۱} a_{t,r} k^s \int_{۰}^{۱} (۱-y)^{\gamma_t + k - ۱} \frac{[-\log(۱-y)]^s}{s!} dy \\ &= c_{r-۱} \sum_{t=۱}^{r} \sum_{s=۰}^{i-۱} \frac{a_{t,r} k^s}{(\gamma_t + k)^{s+۱}} \\ &= ۱ - c_{r-۱} \sum_{t=۱}^{r} \frac{a_{t,r}}{\gamma_t} \left(\frac{k}{\gamma_t + k} \right)^i . \end{aligned} \tag{۱}$$

بنابر این نتایج زیر را داریم.

قضیه ۱ فرض کنید $\{X_n, n \geq ۱\}$ دنباله‌ای از متغیرهای تصادفی پیوسته از جامعه‌ای با تابع توزیع و تابع چگالی احتمال به ترتیب $F(x)$ و $f(x)$ باشد. بعلاوه فرض کنیم $Y_{۱:m:n} \leq Y_{۲:m:n} \leq \cdots \leq Y_{m:n:n}$ آماره‌های مرتب سانسور فزآینده نوع II حاصل از نمونه تصادفی آینده به حجم n از جامعه X_iها بر اساس طرح سانسور $\widetilde{R} = (R_۱, \cdots, R_m)$,

باشند. اگر sامین kـرکورد بالای مربوط به دنباله اول را با $U_{s,k}$ نشان دهیم، آنگاه $[U_{i,k},\ U_{j,k}],\quad j > i \geq ۱$ یک فاصله پیش‌بینی دو طرفه برای $Y_{r:m:n}$، $۱ \leq r \leq m$ با پوشش احتمال زیر است که به F بستگی ندارد:

$$\alpha_۵(r,m,n;i,j,k) = c_{r-۱} \sum_{t=۱}^{r} \frac{a_{t,r}}{\gamma_t} \left(\left(\frac{k}{\gamma_t + k}\right)^i - \left(\frac{k}{\gamma_t + k}\right)^j \right),$$

که در آن γ_t، $a_{t,r}$ و $c_{r-۱}$ بر حسب طرح سانسور $\widetilde{R} = (R_۱, \cdots, R_m)$ داده می‌شوند.

ب) **پیش بینی kـرکوردها بر حسب آماره‌های مرتب سانسور فزاینده:** با استفاده از (۱) نتیجه زیر حاصل می‌شود.

قضیه ۲ تحت مفروضات قضیه ۱، فرض کنید $Y_{۱:m:n} \leq Y_{۲:m:n} \leq \cdots \leq Y_{m:n:n}$ آماره‌های مرتب سانسور فزآینده نوع II حاصل از نمونه تصادفی به حجم n از جامعه‌ای با تابع توزیع احتمال و تابع چگالی احتمال به ترتیب $F(x)$ و $f(x)$ بر اساس طرح سانسور $\widetilde{R} = (R_۱, \cdots, R_m)$، باشند. بعلاوه اگر iامین kـرکورد بالای مربوط به دنباله آینده از همان جامعه را با $U_{i,k}$ نشان دهیم، آنگاه $[Y_{r:m:n},\ Y_{s:m:n}],\quad m > s > r \geq ۱$ یک فاصله پیش بینی دو طرفه برای $U_{i,k}$، $i \geq ۱$ با پوشش احتمال زیر است که به F بستگی ندارد:

$$\alpha_۶(i,k;r,s,m,n) = c_{s-۱} \sum_{t=۱}^{s} \frac{a_{t,s}}{\gamma_t} \left(\frac{k}{\gamma_t + k}\right)^i - c_{r-۱} \sum_{t=۱}^{r} \frac{a_{t,r}}{\gamma_t} \left(\frac{k}{\gamma_t + k}\right)^i.$$

۶ رکوردهای جاری

فرض کنید R_n^l و R_n^s به ترتیب بزرگترین و کوچکترین مشاهده‌ای در دنباله X_iها باشند، زمانیکه nامین رکورد از هر نوع بالا یا پایین رخ داده باشد، در این صورت R_n^l و R_n^s را به ترتیب nامین رکورد جاری بالا و پایین گویند. بعلاوه فاصله (R_n^s, R_n^l) را پوشش رکوردی و تفاضل $R_n^l - R_n^s$ را دامنه رکوردها گویند. بعنوان فرض کنید دنباله زیر را داشته باشیم

$$۳,\ ۲,\ ۲/۵,\ ۲/۶,\ ۱/۷,\ ۳/۷,\ ۲/۲,\ ۱/۵,\ ۲/۷,\ ۲/۳,\ ۱/۲,\ ۴,\ ۲/۵,\ ۴/۷,\ ۴/۱,\ \cdots.$$

دراین صورت رکوردهای جاری مستخرج از دنباله فوق عبارتند از:

m	۱	۲	۳	۴	۵	۶	۷	۸
R_m^l	۳	۳	۳	۳/۷	۳/۷	۳/۷	۴	۴/۷
R_m^s	۳	۲	۱/۷	۱/۷	۱/۵	۱/۲	۱/۲	۱/۲

حال فرض کنید X_iها دنباله‌ای از متغیرهای تصادفی مستقل و هم توزیع از جامعه پیوسته با تابع چگالی $f(x)$ و تابع توزیع احتمال $F(x)$ باشد. در اینصورت تابع چگالی احتمال R_n^l عبارت است از

$$f_{R_n^l}(x) = ۲^n f(x)\{۱ - \bar{F}(x)\sum_{j=۰}^{n-۱}\frac{[-\log \bar{F}(x)]^j}{j!}\}. \qquad (۲)$$

بطور مشابه تابع چگالی احتمال R_n^s بدست می‌آید، که کافی است در عبارت بالا به جای $\bar{F}(x)$ تابع $F(x)$ را قرار دهیم. برای مطالعه بیشتر می‌توان به راقب(۲۰۰۹) مراجعه نمود.

۷ پیش‌بینی آماره‌های مرتب و رکوردهای جاری

با استفاده از ایده ارائه شده در بخش‌های قبل، در اینجا فواصل پیش بینی ناپارامتری برای آماره‌های مرتب نمونه آینده بر حسب پوشش رکوردی بدست می‌آوریم. همچنین برحسب آماره‌های مرتب مشاهده شده پیش‌بینی کننده‌های برای رکوردهای جاری دنباله آینده ارائه می‌شود.

الف) **پیش بینی آماره‌های مرتب بر حسب پوشش رکوردی:** در این خصوص نتایج زیر بدست آمده است.

قضیه ۱ فرض کنید R_n^l و R_n^s رکوردهای جاری مستخرج از دنباله $\{X_n, n \geq ۱\}$ از متغیرهای تصادفی پیوسته از جامعه‌ای با تابع توزیع و تابع چگالی احتمال به ترتیب $F(x)$ و $f(x)$ باشند. بعلاوه فرض کنید $Y_{۱:n} \leq Y_{۲:n} \leq \cdots \leq Y_{n:n}$ آماره‌های مرتب حاصل از نمونه تصادفی آینده به حجم n از جامعه X_iها باشند. در این صورت mامین پوشش رکورد $[R_m^s,\ R_m^l],\ \ m \geq ۲$ یک فاصله پیش بینی دو طرفه برای $Y_{k:n}$، $۱ \leq k \leq n$ با پوشش احتمال زیر است که به F بستگی ندارد:

$$\alpha_۷(m;k,n) \quad = \quad ۲^m - ۱ + k\binom{n}{k}\sum_{r=۰}^{m-۱}(۲^r - ۲^m)$$

$$\times\left\{\sum_{s=0}^{n-k}\binom{n-k}{s}\frac{(-1)^s}{(k+s+2)^{r+1}}+\sum_{s=0}^{k-1}\binom{k-1}{s}\frac{(-1)^s}{(n-k+s+3)^{r+1}}\right\}.$$

حالت‌های خاص برای آماره‌های مرتب فرین با توجه به تقارن $\alpha_7(m;k,n)$ نسبت k، عبارت ساده زیر بدست می آید.

$$P(R_m^s\le Y_{n:n}\le R_m^l)=P(R_m^s\le Y_{1:n}\le R_m^l)$$
$$=1-2^m\left\{\frac{(n+2)^{-m}}{(n+1)}+n\sum_{s=0}^{n-1}\binom{n-1}{s}(-1)^s\frac{(s+3)^{-m}}{(s+1)(s+2)}\right\}.$$

ب) **پیش‌بینی رکوردهای جاری بر حسب آماره‌های مرتب:** بطور مشابه نتایج زیر برای پیش‌بینی رکوردهای جاری دنباله آینده بدست می آید.

قضیه ۲ تحت مفروضات قضیه ۱، $[Y_{i:n},Y_{j:n}]$ یک فاصله پیش بینی دو طرفه برای R_k^l، با پوشش احتمال زیر است:

$$\alpha_8(k;i,j,n)=2^k\sum_{t=i}^{j-1}\sum_{s=0}^{t}\binom{n}{t}\binom{t}{s}\frac{(-1)^s}{(n-t+s+1)(n-t+s+2)^k}.$$

۸ جمع بندی نتایج

دراین مقاله تاکید اصلی مطالعه پیش‌بینی فاصله‌ای ناپارامتری دو نمونه‌ای در حیطه داده‌های ترتیبی بود، بطوریکه نوع داده‌های ترتیبی دونمونه متفاوت در نظر گرفته شده است. نتایج بدست آمده را می‌توان به حالت‌های زیر تعمیم داد:

- با در نظر گرفتن مدل‌های پارامتری در حالت کلاسیک و بیزی نظیر کار انجام شده توسط احمدی و میر مصطفائی (۲۰۰۹) برای سایر توزیع‌های آماری.
- تعیین پیش‌بینی برای فواصل، تفاضل و بطور کلی توابعی از داده‌های ترتیبی.
- با در نظر گرفتن آماره‌های مرتب تعمیم یافته می‌توان نتایج این بحث را تعمیم داد. این آماره‌ها دارای یک ساختار کلی هستند بطوریکه برخی از مدل‌های داده‌های

ترتیبی نظیر رکوردها و سانسور فزاینده با جایگذاری مقادیر خاص در پارامترهای مدل آماره‌های مرتب تعمیم یافته حاصل می‌شوند. برای آشنایی با آماره‌های مرتب تعمیم یافته می‌توان به کمس(۱۹۹۵) مراجعه نمود.

مطالعه موارد فوق توسط نویسنده و تعدادی از همکاران در حال انجام است، امیداواریم نتایج مربوطه بزودی در اختیار علاقه‌مندان قرار گیرد.

مراجع

[1] Ahmadi, J. and Balakrishnan, N. (2010). Prediction of order statistics and record values from two independent sequences, *Statistics* (to appear).

[2] Ahmadi, J. and MirMostafaee, S. M. T. K. (2009). Prediction intervals for future records and order statistics coming from two parameter exponential distribution, *Statistics & Probability Letters*, **79**, 977-983.

[3] Ahsanullah, M. (1980). Linear prediction of record values for the two parameter exponential distribution, *Ann. Inst. Statist. Math.* **32**, 363-ت368.

[4] Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1992), *A First Course in Order Statistics*, John Wiley & Sons, New York.

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

[6] Balakrishnan, N. and Aggarwala, R. (2000). *Progressive Censoring: Theory, Methods, and Applications*, Birkhäuser, Boston.

[7] Balakrishnan, N. (2007). Progressive censoring methodology: An appraisal. *Test*, **16**, 211–256.

[8] David, H. A. and Nagaraja, H. N. (2003). *Order Statistics*, Third edition, John Wiley & Sons, Hoboken, New Jersey.

[9] Dunsmore, I. R. (1983). The future occurrence of records, *Ann. Inst. Statist. Math.* **35**, 267–277.

[10] Kaminsky, K. S. and Nelson, P. I. (1998). Prediction of order statistics, In *Handbook of Statistics – 17: Order Statistics: Applications* (Eds., N. Balakrishnan and C.R. Rao), pp. 431–450, North-Holland, Amsterdam.

[11] Kamps, U. (1995) *A Concept of Generalized Order Statistics.* Teubner, Stuttgart.

[12] Raqab, M. Z. (2009). Distribution-free prediction intervals for the future current record statistics, *Statist. Pap.*, **53**, 429–439.

[13] Raqab, M. and Balakrishnan, N. (2008). Prediction intervals for future records, *Statist. Probab. Lett.* **78**, 1955–1963.

[14] Raqab, M. Z. and Nagaraja, H. N. (1995). On some predictions of future order statistic, *Metron*, **53**, 185–204.