رکوردها

هيچ‌کس نمي‌تواند علاقه‌ي خود را به رکوردها انکار کند. گرمترين روز سال, بالاترين قيمت يک کالاي معين, کمترين زمان بدست آمده در رشته‌هاي مختلف ورزشي و ... نمونه‌هايي از رکوردها هستند که همواره مورد علاقه متخصصان و عموم بوده‌اند. در مبحث رکوردها مي‌توان مطالب جالبي را مورد مطالعه قرار داد, همانند مدت زمان شکسته شدن يک رکورد. بنابراين ديده مي‌شود که کار روي رکوردها وسعت زيادي دارد. اما ساده نيست, به دو دليل: اولاً مدل رکورد پايه با مشاهدات $i.i.d$ که اولين بار توسط چندلر(Chandler) بيان شد بسيار ساده است و وقتي چند شرط به آن افزوده مي‌شود تا به واقعيت نزديکتر گردد, بسيار مشکل مي‌شود. دوماً رکوردها نادر هستند, يعني مدت زمان زيادي طول مي‌کشد تا در يک دنباله يک رکورد رخ دهد. بنابراين براي بررسي رکوردها, نمونه هاي کمي در دسترس است. تا به حال افراد کمي روي رکوردها کار کرده‌اند, اما همان افراد به نتايج زيادي دست يافته‌اند. چندلر را مي‌توان پدر اين مبحث آماري ناميد. وي براي اولين بار مقادير رکوردها را در رابطه با شرايط آب و هوا در نظر گرفت. به دنبال او, روابط ميان رکوردها و مباحث مربوطه, شناخته و بسط داده‌شدند.

فرآيند رکورد استاندارد

دنباله‌اي نامتناهي از متغيرهاي تصادفي مستقل و همتوزيع $X_1, X_2, X_3, \ldots$ را در نظر بگيريد. تابع توزيع مشترک اين دنباله را با F نشان مي‌دهيم. هيچ شرطي روي F نمي‌گذاريم, اما واضح است که F نبايد تباهيده باشد. اغلب F پيوسته فرض مي‌شود تا بهتر بتوان روي آن کار کرد.

تعريف مقدار رکورد بالايي

مشاهده X_j يک مقدار رکورد بالايي ناميده مي‌شود, اگر مقدار آن از تمامي مشاهدات قبل از خود بيشتر باشد. بنابراين تعريف, X_j يک رکورد است اگر به ازاي همه i ها $(i<j)$ ؛ $X_j > X_i$.

- مقدار رکورد پاييني نيز به طريق مشابه تعريف مي‌شود کافي فرض شود: $Y_n = -X_n$.

تعريف زمان رخداد رکورد

زمان‌هايي که در آن رکوردها رخ مي‌دهند, نيز بسيار مورد علاقه مي‌باشند. X_j را رکورد رخ داده در زمان j در نظر بگيريد, آنگاه دنباله‌ي زمان رکورد $\{T_n; n \geq 0\}$ به شکل زير تعريف مي شود:

$$T_0 = 1 \qquad \text{با احتمال ۱}$$

$$T_n = \min\{j : X_j > X_{T_{n-1}}\} \qquad \text{و براي هر } n \geq 1$$

تعريف دنباله‌ي مقدار رکورد

دنباله‌ي $\{R_n\}$ به شکل زير تعريف مي‌شود:

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

R_0 را بعنوان مقدار مرجع يا رکورد کم‌اهميت در نظر مي‌گيريم و همواره بررسي ساير رکوردها مورد نظر است.

در تعريف بالا, از دنباله رکوردها, به طور قطع فرض مي‌شود F رکورد غير قابل شکست توليد نخواهد کرد. بنابراين بايد مقداري مثل x_0 وجود داشته باشد که $F(x_0) - F(x_0^-) > 0$, $F(x_0) = 1$

يعني بزرگترين مقدار واقعي ممکن وجود دارد که مي‌تواند با احتمال مثبت توسط X_j ها بدست‌آيد.

به عنوان مثال, اگر مکرراً تاسي را پرتاب کنيم و X_j نشان دهنده رقم بدست آمده در پرتاب j ام باشد, مقدار $X_j = 6$, يک رکورد غيرقابل شکست است. در اينجا $\{R_n\}$ را مي‌توان يک زنجير ماركوف در نظر گرفت که "۶" يک حالت جذب آن باشد.

تعريف دنباله نموهاي رکورد

دنباله $\{J_n; n \geq 0\}$ را در نظر بگيريد که به شکل زير تعريف مي‌شود:

$$J_0 = R_0$$

$$J_n = R_n - R_{n-1} \qquad , \text{و براي هر } n > 1$$

- R_0 را رکورد کم اهميت و بقيه را رکوردهاي بااهميت مي‌ناميم.

تعريف زمان بين رکوردها

Δ_n را در نظر بگيريد که به شکل زير تعريف مي‌شود:

$$\Delta_n = T_n - T_{n-1}; n = 1,2,3,...$$

تعريف فرآيند شمارشي رکوردها

دنباله $\{N_n; n \geq 1\}$ را در نظر بگيريد که در آن $N_n = \{$تعداد رکوردهاي بين $X_1, X_2, ..., X_n\}$.

- واضح است که $N_1 = 1$, زيرا که X_1 هميشه يک رکورد است.

مقادير مدل از رکورد کلاسيک

در اين مدل X_i ها متغيرهاي تصادفي مستقي, همتوزيع و پيوسته هستند. جهت راحتي کار, ابتدا متغير نمايي را بررسي مي‌کنيم. فرض کنيد $\{X_j'; j \geq 1\}$ نشان‌دهنده‌ي دنباله‌اي از متغيرهاي تصادفي مستقل نمايي با پارامتر ۱ باشد.

مي‌دانيم توزيع نمايي داراي خاصيت کم‌حافظگي است و تفاضل ميان رکوردهاي پياپي يعني $\{J_n'\} = \{R_n' - R_{n-1}'; n \geq 1\}$ نيز متغيرهاي تصادفي مستقل و همتوزيع داراي توزيع نمايي استاندارد هستند, پس: $\sum J_i' = R_n' - R_0' \sim Gamma(n,1)$

آن‌گاه براي nامين رکورد, R_n', بدست آمده از دنباله‌اي $i.i.d$ و $\exp(1)$ داريم:

$$R_n' \sim Gamma(n+1,1); n = 0,1,2,... \quad (۱)$$

حال با استفاده از اين نتيجه, توزيع n امين رکورد متناظر با دنباله‌اي از متغيرهاي تصادفي $i.i.d$, $\{X_j\}$ با توزيع مشترک پيوسته F را بدست مي‌آوريم.

تابع توزيع و چگالي R_n

مي‌دانيم اگر X داراي توزيع پيوسته F باشد, آن‌گاه $H(X) = -Log(1-F(X))$ داراي توزيع نمايي استاندارد است. پس مي‌توان نتايج مربوط به F نمايي را به F پيوسته تعميم داد.

$X \overset{d}{=} F^{-1}(1-e^{-X'})$ که در آن همانند قبل: $X' \sim \exp(1)$. X تابعي يکنوا از X' است, پس n امين رکورد از دنباله $\{X_j\}$ تابعي از n امين رکورد دنباله X' مي‌باشد, بنابراين

$$R_n \overset{d}{=} F^{-1}(1-e^{-R_n'}); n = 0,1,2,... \quad (۲)$$

با انتگرال‌گيري جزء به جزء و با استفاده از (۱) رابطه‌ي زير بدست مي‌آيد:

$$P(R_n' > r') = \int_{r'}^{\infty} \frac{1}{n!}(R_n')^n e^{-R_n'} dR_n' = e^{-r'} \sum_{k=0}^{n} \frac{(r')^k}{k!} \quad ; r' > 0 \quad (۳)$$

حال با جايگذاري (۲) در (۳) خواهيم داشت:

$$P(R_n > r) = \frac{[1-F(r)]\sum_{k=0}^{n}[-\log(1-F(r))]^k}{k!}$$

و يا معادل با آن داريم:

$$P(R_n \leq r) = \int_0^{[-Log(1-F(r)]} \frac{w^n e^{-w}}{n!} dw \quad (۴)$$

اگر F مطلقاً پيوسته و f تابع چگالي متناظر با آن باشد, مي‌توان با مشتق‌گيري تابع چگالي متناظر با (۴) را بدست آورد, داريم:

$$F_{R_n}(r) = \frac{f(r)[-Log(1-F(r)]^n}{n!} \quad (۵)$$