فهرست مندرجات

پیشگفتار

فصل ۱ را با مروری کوتاه بر آزمون نسبت درستنمایی (LRT) آغاز می‌کنیم. سپس به معرفی مفهوم نقطه تغییر و فرضیات مرتبط با آن می‌پردازیم. همچنین در این فصل نظریه مقادیر فرین و چند مفهوم احتمالی بیان می‌شود. در فصل ۲ به مطالعه رفتار مجانبی آماره لگاریتم نسبت درستنمایی برای آزمون یک تغییر در توزیع وایبل ۳ پارامتری می‌پردازیم. همچنین نشان می‌دهیم اگر پارامتر شکل (α) بیشتر از ۲ باشد، قانون لگاریتم مکرر برای برآوردگرهای درستنمایی ماکسیمم برقرار است و آماره آزمون لگاریتم درستنمایی به طور مجانبی دارای توزیع گامبل است. در فصل ۳، توزیع مجانبی آماره آزمون نسبت درستنمایی ماکسیمم را زمانی که پارامترهای توزیع نرمال در یک نقطه نامعلوم تغییر کرده‌اند محاسبه می‌کنیم. نظر به این که اثبات‌ها براساس فرآیند ارنشتین—اهلنبک و فرآیند وینر است، در این فصل به معرفی مختصر این فرآیندها پرداخته می‌شود. در انتهای فصل، ۲ مثال کاربردی نیز ارائه می‌شود. با توجه به این که نقطه تغییر در شاخه‌های مختلف استنباط آماری کاربرد دارد، در فصل ۴ مختصری از این روش را در مبحث رگرسیون خطی مورد بررسی قرار می‌دهیم. لازم به ذکر است که قسمت‌هایی که در پایان‌نامه با علامت $*$ مشخص شده‌اند، توسط نگارنده اثبات شده‌اند.

فصل ۱

تعاریف و مقدمات

۱.۱ مقدمه

در این فصل مروری کوتاه بر آزمون نسبت درستنمایی (LRT[۱]) انجام می‌شود. سپس به معرفی نقطه تغییر و فرضیات مرتبط با آن می‌پردازیم. همچنین در این فصل نظریه مقادیر فرین و چند مفهوم احتمالی بیان می‌شود.

کلمه آمار در ابتدا در جمع‌آوری اعداد و ارقام به کار می‌رفت، اعداد جمع‌آوری شده اینک داده[۲] نامیده می‌شوند و آمار امروز دارای معنی وسیع‌تری شده است و به معنی علم تصمیم‌گیری به کار می‌رود و روز به روز بر دامنه علم آمار افزوده می‌گردد. علم آمار به دو شاخه تقسیم می‌شود: آمار توصیفی و آمار استنباطی.

هدف از آمار توصیفی، خلاصه کردن کمی اطلاعاتی است که در مورد یک جامعه ملموس به وسیله یک بررسی فراگیر جمع‌آوری شده است، نظیر جمعیت یک کشور که از خلال یک سرشماری عمومی مورد بررسی است. هدف آمار توصیفی توجیه نیست بلکه توصیف، استخراج نکات اساسی، تحقق بخشیدن به ترکیب اطلاعات به کمک زبان اعداد است. در واقع در آمار توصیفی، داده‌ها، یعنی اطلاعات عددی درباره امری طبق قواعدی خلاصه و سپس جدول‌های فراوانی و نمودارهای آماری ارائه می شود.

هدف از آمار استنباطی، بررسی استنتاج‌های آماری یعنی تجزیه و تحلیل اطلاعات به دست آمده از یک روش تصادفی است. استنباط آماری دارای دو قسمت عمده است: برآورد پارامترها و آزمون فرضیه‌ها. همان طور که می‌دانیم چندین روش در آزمون فرضیه‌ها وجود دارد، از جمله روش نیمن پیرسونی، نسبت درستنمایی، نسبت درستنمایی تعمیم یافته و .... در

[۱] Likelihood ratio test

[۲] Data

این قسمت روش نسبت درستنمایی مورد بررسی قرار می‌گیرد.

۲.۱ آزمون نسبت درستنمایی

روش LRT در سال ۱۹۲۸ توسط نیمن و پیرسون مطرح شد. نقشی که این روش در نظریه آزمون فرضیه‌ها داراست، همانند نقش روش درستنمایی ماکسیمم (ML[۳]) در نظریه برآوردهاست. روش LRT با وجود این که به طور شهودی قابل توجه است، همواره بهترین روش نیست و نباید همیشه مورد استفاده قرار گیرد. در حقیقت قضیه‌ای که بیانگر این باشد که روش LRT جز در موارد مجانبی دارای خواص خوبی است، وجود ندارد. به طور خلاصه روش LRT در حالت کلی قابل توجیه‌تر از روش برآورد ML نیست.

در حقیقت روش LRT تعمیم آزمون نیمن—پیرسون است که عموما نتایج قطعی و خوبی در آزمون بسیاری از فرضیه‌ها ارائه می‌دهد. فرض کنید $X_۱$، ...، X_n یک نمونه تصادفی nتایی از توزیعی با تابع چگالی احتمال (یا تابع احتمال) $f_\theta(x)$، $\theta \in \Theta$ باشد. همچنین فرض کنید $\Theta_۰$ و $\Theta_۱$ زیرمجموعه‌هایی از Θ باشند، به طوری که

$$\Theta_۰ \cup \Theta_۱ = \Theta, \quad \Theta_۰ \cap \Theta_۱ = \varnothing$$

در این جا علاقه‌مند به ارائه روشی برای آزمون فرضیه زیر هستیم.

$$H_۰ : \theta \in \Theta_۰ \quad \text{vs} \quad H_۱ : \theta \in \Theta_۱$$

[۳] Maximum likelihood

می‌دانیم که قانون درستنمایی ماکسیمم مبتنی بر مقداری از θ، یعنی $\hat{\theta}$ است که تابع درستنمایی $L(\theta)$ را ماکسیمم می‌کند. یعنی

$$L(\hat{\theta}) = \sup_{\theta} L(\theta)$$

در حقیقت چنین مقداری از $\hat{\theta}$، بیشترین حمایت را از مشاهدات $\underline{x} = (x_1, \ldots, x_n)$ می‌کند. به عبارت دیگر، مشاهدات $\underline{x} = (x_1, \ldots, x_n)$ شانس بیشتری به نقطه‌ای از فضای پارامتر می‌دهد که آن نقطه را با $\hat{\theta}$ نشان می‌دهیم. بنابراین در انتخاب بین H_0 و H_1، طبیعی به نظر می‌رسد که مقایسه بین بهترین تبیین کننده موجود در H_0 و بهترین تبیین کننده موجود در H_1 انجام پذیرد. پس وقتی که $X = x$ مشاهده شود. فرضیه H_0 را رد می‌کنیم، اگر و فقط اگر

$$\sup_{\Theta_0} L(\theta) < \sup_{\Theta_1} L(\theta)$$

تعمیم ایده فوق برای دستیابی به انعطاف بیشتر در مورد سطح آزمون، ما را بر آن می‌دارد که ناحیه ردی به شکل زیر داشته باشیم

$$\lambda^* = \frac{\sup_{\Theta_0} L(\theta)}{\sup_{\Theta_1} L(\theta)} < \lambda_0^*$$

که در آن λ_0^* عددی ثابت است. چنین ناحیه ردی، روش نسبت درستنمایی را تعریف می‌کند. در بسیاری از آزمون‌ها، راحت‌تر است که شکل اصلاح شده‌ای (تعمیم یافته) از نسبت درستنمایی را (که گاهی نسبت درستنمایی تعمیم یافته نیز نامیده می‌شود) مورد استفاده قرار دهیم. در شکل تعمیم یافته، وقتی که $X = x$ مشاهده شود، بهترین تبیین کننده θ در Θ_0 با بهترین تبیین کننده θ در Θ مقایسه می‌شود، یعنی

$$\lambda(x) = \frac{\sup_{\Theta_0} L(\theta)}{\sup_{\Theta} L(\theta)} = \frac{L(\hat{\theta}_0)}{L(\hat{\theta})}$$

اگر $f_\theta(x)$ یک تابع احتمال در نظر گرفته شود، معقول بودن LRT بیشتر تفهیم خواهد شد. آزمون نسبت درستنمایی وقتی $\lambda(x) < \lambda_0$ باشد فرضیه H_0 را رد می‌کند، که در آن λ_0 عددی ثابت در فاصله‌ی $[0, 1]$ است و با توجه به سطح آزمون مشخص می‌شود.
بنابراین تابع آزمون زیر را برای آزمون نسبت درستنمایی خواهیم داشت.

$$\phi(x) = \begin{cases} 1 & \lambda(x) < \lambda_0 \\ \gamma & \lambda(x) = \lambda_0 \\ 0 & \lambda(x) > \lambda_0 \end{cases}$$

که در آن α از رابطه زیر به دست می آید.

$$\alpha = \sup_{\theta \in \Theta_0} E_\theta[\phi(x)]$$

$\lambda(x)$ آماره آزمون نسبت درستنمایی نامیده می‌شود. اگر بتوانیم نشان دهیم که آماره آزمون LR با یک آماره مناسب (که توزیع آن کاملا مشخص یا قابل حصول است) رابطه یک به یک دارد، نه تنها مشکلی درارائه آزمونی معادل برای H_0 در مقابل H_1 نخواهیم داشت، بلکه این کار عملا توصیه می‌شود. هر چند، روش LRT همیشه راحت و ساده نیست، به خصوص وقتی که آماره آزمون تابعی پیچیده از داده‌ها باشد و نتوان توزیع دقیق آن را به دست آورد. در چنین مواردی باید خود را مقید به توزیع مجانبی آماره آزمون کنیم، به شرط آن که «شرایط مطلوب» وجود داشته باشد.

۳.۱ شرایط مطلوب برای توزیع مجانبی آماره آزمون LR چیست؟

اگر X یک متغیر تصادفی با توزیعی از خانواده چگالی‌های $\{P_\theta(x); \theta \in \Theta\}$ باشد، در این صورت شرایط مطلوب عبارت است از:

الف— Θ یک زیرفاصله باز از اعداد حقیقی است یعنی $\Theta \subseteq \mathbb{R}$.

ب— مشتق تابع چگالی نسبت به θ، یعنی $\frac{\partial}{\partial\theta}P_\theta(x)$ وجود دارد.

ج— جابه‌جایی عملگرهای مشتق و انتگرال مجاز است، یعنی

$$\frac{\partial}{\partial\theta}\int P_\theta(x)d\mu(x) = \int \frac{\partial}{\partial\theta}P_\theta(x)d\mu(x)$$

د— برای هر $\theta \in \Theta$،

$$E[\frac{\partial}{\partial\theta}LnP_\theta(x)]^۲ > ۰$$

ه— برای هر برآوردگر T پارامتر $\gamma(\theta)$،

$$\begin{aligned}\frac{\partial}{\partial\theta}E_\theta[T(X)] &= \frac{\partial}{\partial\theta}\int t(x)P_\theta(x)d\mu(x) \\ &= \int t(x)\frac{\partial}{\partial\theta}P_\theta(x)d\mu(x)\end{aligned}$$

۴.۱ توزیع مجانبی آماره LRT

در مواردی که آماره LRT قابل تبدیل به یک آماره ساده نباشد، برای به دست آوردن توزیع آماره LRT باید حالتی را که حجم نمونه، بزرگ است، در نظر بگیریم. با توجه به این که $\lambda(x)$ روی فاصله $(۰, ۱)$ توزیع شده است، برای هر مقدار ثابت و مثبت c، آماره $W = -۲c\ln\lambda(X)$ روی فاصله $(۰, \infty)$ توزیع خواهد شد. بنابراین طبیعی به نظر می‌رسد که c را به گونه‌ای انتخاب کنیم که W دارای توزیع تقریبی $\chi^۲$ باشد.

قضیه. ۱.۱ فرض کنید $X_۱, \dots, X_n$ یک نمونه تصادفی از توزیعی با خانواده چگالی‌های $\{f_\theta; \theta \in \Theta \subseteq \mathbb{R}^k\}$ باشد، به طوری که در شرایط مطلوب صدق می‌کنند.

در آزمون زیر با استفاده از روش LRT

$$H_۰ : \theta_i = \theta_{i_۰}, i = ۱, \dots, r \quad \text{vs} \quad H_۱ : \theta_i \neq \theta_{i_۰}, ۱ \leq i \leq r$$ ،i حداقل برای یک

که در آن $\theta_{i_۰}$، $i = ۱, \dots, r$ معلوم و $۱ \leq r \leq k$ است، تحت فرضیه $H_۰$ (وقتی که اندازه نمونه بزرگ باشد) آماره $-۲\ln\lambda(x)$ دارای توزیع تقریبی $\chi^۲$ با $r = \dim\Theta - \dim\Theta_۰$ درجه آزادی است که Θ فضای پارامتر در حالت کلی و $\Theta_۰$ فضای پارامتر تحت فرضیه $H_۰$ است.

توجه داریم که اگر دامنه توزیع متغیرهای تصادفی به پارامتر نامعلوم θ بستگی داشته باشد، توزیع چنین متغیرهای تصادفی در شرایط مطلوب صدق نمی‌کند.

حال سوالی که در این قسمت پیش می‌آید این است که اگر شرایط مجانبی آماره LRT عوض شود، برای توزیع مجانبی آماره نسبت درستنمایی چه اتفاقی می‌افتد؟ مثلا زمانی که از

نمونه‌ای استفاده می‌شود که شرایط هم‌توزیعی برای آن برقرار نیست. مانند نمونه ای که با استفاده از نقطه تغییر[۴] بدست می آید و در بخش بعد به تعریف آن می‌پردازیم.

۵.۱ نقطه تغییر چیست؟

فرض کنید X_1، X_2، ...و X_n بردارهای تصادفی مستقل در $\mathbb{R}^m$ با توابع توزیع $F(X;\theta_1,\eta_1)$، ...و $F(X;\theta_n,\eta_n)$ باشند که $\theta_i \in H^{(1)} \subseteq \mathbb{R}^d$ و $\eta_i \in H^{(2)} \subseteq \mathbb{R}^p$.

می‌خواهیم فرضیه H_0 را در برابر فرضیه H_A آزمون کنیم:

$$H_0: \quad \theta_1 = \ldots = \theta_n, \quad \eta_1 = \ldots = \eta_n,$$

$$H_A: \quad \eta_1 = \ldots = \eta_n,$$

k^*ای وجود دارد که $1 \leq k^* \leq n$

$$\theta_1 = \ldots = \theta_{k^*} \neq \theta_{k^*+1} = \ldots = \theta_n$$

به این حالت یک مدل دارای حداکثر یک نقطه تغییر می‌گوییم (AMOC[۵]).

فرض می‌کنیم ηها پارامترهای مزاحم و ثابتند در حالی که θ تحت فرضیه مقابل تغییر می‌کند. همه پارامترها تحت فرضیه H_0 و تحت فرضیه مقابل نامعلومند.

فرض کنید که زمان تغییر k^* معلوم باشد. می‌خواهیم آزمون نسبت درستنمایی را به کار ببریم. به طور معمول فرضیات زیر را در نظر می‌گیریم:

---

[۴] Change point

[۵] At most one change-point

C.۱— X_1، ... و X_n دارای توابع چگالی $f(X;\theta_1,\eta_1)$، ... و $f(X;\theta_n,\eta_n)$ نسبت به v هستند که v یک اندازه σ متناهی روی $\mathbb{R}^m$ است.

C.۲— $F(X;\theta,\eta)$ اندازه‌های متفاوت تولید می‌کنند اگر $(\theta,\eta)\in\Theta=\Theta^{(1)}\times\Theta^{(2)}$

اگر C.۲ برقرار نباشد، در این صورت پارامترها به طور سازگار قابل برآورد نیستند (لهمن (۱۹۹۹) صفحه ۴۰۹).

اگر تغییر در $k^*=k$ رخ دهد، در این صورت H_0 را برای مقادیر کوچک Λ_k رد می‌کنیم:

$$\Lambda_k=\frac{\sup\limits_{(\theta,\eta)\in\Theta^{(1)}\times\Theta^{(2)}}\prod\limits_{1\le i\le n}f(x_i;\theta,\eta)}{\sup\limits_{(\theta,\tau,\eta)\in\Theta^{(1)}\times\Theta^{(1)}\times\Theta^{(2)}}\prod\limits_{1\le i\le k}f(x_i;\theta,\eta)\prod\limits_{k<i\le n}f(x_i;\tau,\eta)} \qquad (۱-۱)$$

با استفاده از برآوردگرهای درستنمایی ماکسیمم، می‌توان یک عبارت ساده‌تر برای Λ_k به دست آورد. قرار می‌دهیم:

$$g(x;y)=\log f(x;y) \qquad (۲-۱)$$

$$g_i(x;y)=\frac{\partial}{\partial y_i}g(x;y) \qquad (۳-۱)$$

برای هر $X\in\mathbb{R}^m$ و $Y=(Y_1,\ldots,Y_{d+p})\in\Theta^{(1)}\times\Theta^{(2)}$.

فرضیه زیر می‌گوید که بدون توجه به این که نقطه تغییر چه زمانی رخ می‌دهد (یا حتی اگر رخ ندهد) می‌توانیم پارامترها را با استفاده از معادلات درستنمایی برآورد کنیم.

C.۳— برای هر k که $1\le k\le n$ می‌توانیم برآوردگرهای یکتای $\hat{\theta}_k$، θ_k^* و $\hat{\eta}_k$ را بیابیم، به قسمی که

$$\sum_{1\le j\le k}g_i(x_j;\hat{\theta}_k,\hat{\eta}_k)=0,\quad 1\le i\le d \qquad (۴-۱)$$

$$\sum_{k<j\leq n} g_i(x_j;\theta_k^*,\hat{\eta}_k)=0, \quad 1\leq i\leq d \tag{۱—۵}$$

$$\sum_{1\leq j\leq k} g_i(x_j;\hat{\theta}_k,\hat{\eta}_k)+\sum_{k<j\leq n} g_i(x_j;\theta_k^*,\hat{\eta}_k)=0, \quad d+1\leq i\leq d+p \tag{۱—۶}$$

(توجه داریم که اگر $k=n$، در این صورت فقط (۱—۴) و (۱—۵) را حل می‌کنیم و θ_n^* قابل تعریف نیست. این حالتی است که هیچ تغییری رخ نمی‌دهد و بنابراین برآوردی برای پارامتر بعد از تغییر نداریم.)

البته C.۳ می‌تواند با شرط زیر نیز جایگزین شود:

C.۳* ثابت k_0 وجود دارد به قسمی که می‌توان $\hat{\theta}_k$، θ_k^* و $\hat{\eta}_k$ یکتا برای هر k، که $k_0\leq k\leq n-k_0$ یافت به طوری که (۱—۴)—(۱—۶) برقرار باشند. همچنین می‌توان $\hat{\theta}_n$ و $\hat{\eta}_n$ یکتا را به قسمی یافت که

$$\sum_{1\leq j\leq n} g_i(x_j;\hat{\theta}_n,\hat{\eta}_n)=0, \quad 1\leq i\leq d$$

نسبت درستنمایی را می‌توان به صورت زیر نوشت:

$$-2\log\Lambda_k=2\{L_k(\hat{\theta}_k,\hat{\eta}_k)+L_k^*(\theta_k^*,\hat{\eta}_k)-L_n(\hat{\theta}_n,\hat{\eta}_n)\} \tag{۱—۷}$$

که

$$L_k(Y)=\sum_{1\leq j\leq k} g(X_j;Y), \tag{۱—۸}$$

$$L_k^*(Y)=\sum_{k<j\leq n} g(X_j;Y) \tag{۱—۹}$$

آماره آزمون را به صورت زیر در نظر می‌گیریم

$$Z_n=\max_{1\leq k<n}(-2\log\Lambda_k). \tag{۱—۱۰}$$

و H_0 را برای مقادیر بزرگ Z_n رد می‌کنیم.

(آماره Z_n تابعی از فرایند $\{-2\log\Lambda_k, 1 \le k < n\}$ می‌باشد. بنابراین می‌توانیم توزیع حدی Z_n را به عنوان تقریبی برای فرآیند $\{-2\log\Lambda_k, 1 \le k < n\}$ که بر حسب k یکنواخت است (برای nهای بزرگ) در نظر بگیریم).

قرار می‌دهیم

$$g_{i_1,\ldots,i_r}(x;y) = \frac{\partial^r g(x;y)}{\partial y_{i_1}\ldots\partial y_{i_r}}, \quad 1 \le i_1,\ldots,i_r \le d+p \tag{1-11}$$

که $x \in \mathbb{R}^m$ و $y = (y_1,\ldots,y_{d+p}) \in \Theta^{(1)} \times \Theta^{(2)}$.

مقادیر صحیح پارامترها تحت H_0، (θ_0, η_0) هستند.

C.4– بازه باز $\Theta_0 \subseteq \Theta \subseteq \mathbb{R}^{d+p}$ شامل (θ_0, η_0) وجود دارد به قسمی که برای هر $y \in \Theta_0$ و $x \in \mathbb{R}^m$، $g_{i,j,k}(x;y)$، $g_{i,j}(x;y)$ و $g_i(x;y)$، $1 \le i,j,k \le d+p$ موجود و در y پیوسته‌اند.

C.5– توابع $M_1(x)$ و $M_2(x)$ موجودند به قسمی که برای هر $x \in \mathbb{R}^m$ و $y \in \Theta_0$

$$|g_i(x;y)| \le M_1(x), |g_{i,j}(x;y)| \le M_2(x), \tag{1-12}$$

$$|g_{i,j,k}(x;y)| \le M_2(x), \quad 1 \le i,j,k \le d+p \tag{1-13}$$

توابع M_1 و M_2 در

$$\int_{\mathbb{R}^m} M_1(x)v(dx) < \infty \tag{1-14}$$

$$E_{(\theta_0,\eta_0)} M_2(X_1) < \infty \tag{1-15}$$

صدق می‌کنند

C.۶— برای هر $y \in \Theta_\circ$ و برای هر $1 \le i \le d+p$، $E_Y g_i(X_1;Y) = 0$.

C.۷— برای هر $y \in \Theta_\circ$ و برای هر $1 \le i, j \le d+p$،

$$J_{ij}(y) = E_Y g_i(X_1;Y) g_j(X_1;Y) = -E_Y^2 g_{i,j}(X_1;Y).$$

و $J^{-1}(y)$ها موجود و پیوسته هستند که

$$J(y) = \{J_{i,j}(y), 1 \le i, j \le d+p\}.$$

ماتریس اطلاع است.

C.۸— برای هر $1 \le i, j \le d+p$، $\underset{(\theta_\circ,\eta_\circ)}{\operatorname{var}} g_{i,j}(x_1;\theta_\circ,\eta_\circ) < \infty$.

C.۹— برای هر $1 \le i \le d+p$ و برای بعضی مقادیر $\mu > 2$،

$$E_{(\theta_\circ,\eta_\circ)} |g_i(x_1;\theta_\circ,\eta_\circ)|^\mu < \infty$$

همانگونه که در فصل دوم خواهیم دید نقطه تغییر و اثر آن را بر آزمون لگاریتم نسبت درستنمایی در توزیع وایبل سه پارامتری بررسی می‌کنیم. نظر به این که توزیع وایبل، توزیعی فرین است ابتدا مختصری در مورد توزیع‌های فرین و سپس کاربردهای توزیع وایبل بیان می‌شود.

۶.۱ نظریه مقادیر فرین

نظریه مقادیر فرین[۶] توسط ریاضیدان آلمانی، امیل ژولیس گامبل[۷] با معرفی توزیع گامبل گسترش یافت. این نظریه یکی از شاخه‌های آمار است که در ارتباط با تغییرات فرین از میانه توزیع احتمال است. از این نظریه برای ارزیابی نوع توزیع احتمالی تولید شده توسط فرآیندها استفاده می‌شود. همچنین این نظریه در ارزیابی خطرات در حوادث غیر مترقبه نظیر طوفان‌های خیلی شدید که در طی ۱۰۰ سال رخ می‌دهد، کاربرد دارد. توزیع‌های مقادیر فرین به توزیع‌های مرتبط با می‌نیمم و ماکسیمم از یک مجموعه بزرگ از مشاهدات همتوزیع، محدود است.

توزیع‌های فرین سه نوع هستند:

۱) **نوع اول:** توزیع گامبل (توزیع فیشر—تیپت[۸] یا توزیع لگ وایبل[۹])

۲) **نوع دوم:** توزیع فرچه[۱۰]

۳) **نوع سوم:** توزیع وایبل (توزیع رزین—راملر[۱۱])

حال مختصری از کاربردهای توزیع وایبل را که از منابع اینترنتی اخذ شده است بیان می‌کنیم:

۱) در تحلیل بقا

[۶] Extreme value theory
[۷] Emil Julius Gumbel
[۸] Fisher-Tippet
[۹] Log-Weibull
[۱۰] Frechet
[۱۱] Rosin-Rammler

۲) برای نمایش زمان تحویل در مهندسی صنایع

۳) در نظریه مقادیر فرین

۴) در پیش‌بینی آب و هوا

۵) در قابلیت اعتماد و تحلیل خطر

۶) در سیستم‌های رادار برای مدل‌سازی سطح سیگنال‌های دریافتی توسط چند نوع بی‌نظمی

۷) برای مدل‌سازی کانال‌های محو شده در سیستم ارتباط بدون سیم

۸) در بیمه برای مدل‌سازی حجم بیمه اتکایی

۹) در پیش‌بینی تغییرات تکنولوژی

۱۰) برای توصیف توزیع سرعت باد

در ادامه به معرفی چند مفهوم در احتمال و قضایای مورد نیاز می‌پردازیم:

۷.۱ چند مفهوم احتمالی

فرض کنید $\{a_n\}_{n=1}^{n=+\infty}$ و $\{b_n\}_{n=1}^{n=+\infty}$، دو دنباله از اعداد حقیقی و $\{X_n\}_{n=1}^{\infty}$ دنباله‌ای از متغیرهای تصادفی باشند.

تعریف. ۲.۱ اگر نسبت $|\frac{a_n}{b_n}|$ برای هر n کراندار باشد آن‌گاه $a_n = O(b_n)$. یعنی

$$\exists k \in \mathbb{R}, n(k) \ni n > n(k) \Rightarrow |a_n| < k|b_n|.$$

تعریف. ۳.۱ اگر نسبت $|\frac{a_n}{b_n}|$ همگرا به صفر باشد، آن‌گاه $a_n = o(b_n)$. یعنی

$$\forall \varepsilon > 0, \exists n(\varepsilon) \ni n > n(\varepsilon) \Rightarrow |a_n| < \varepsilon|b_n|.$$

خواص

۱) $a_n = O(a_n)$

۲) $a_n = o(1) \Rightarrow \lim\limits_{n\to\infty} a_n = 0$

۳) برای $n > n(k)$، $a_n = O(1) \Rightarrow |a_n| < k$

۴) $a_n o(1) = o(a_n)$

۵) $a_n O(1) = O(a_n)$

۶) $O(a_n)O(b_n) = O(a_n b_n)$

۷) $O(a_n)o(b_n) = o(a_n b_n)$

۸) $o(a_n)o(b_n) = o(a_n b_n)$

۹) $O(o(a_n)) = o(a_n)$

۱۰) $o(o(a_n)) = o(a_n)$

۱۱) $o(O(a_n)) = o(a_n)$

۱۲) $O(a_n) + O(b_n) = \max(O(a_n), O(b_n)) = O(\max(|a_n|, |b_n|))$

۱۳)

$$O(a_n) + o(b_n) = \begin{cases} O(a_n), & b_n = o(a_n) \text{ یا } b_n = O(a_n) \\ O(b_n), & a_n = O(b_n) \\ o(b_n), & a_n = o(b_n) \end{cases}$$

توجه کنید اگر $b_n = o(1)$، آن‌گاه $O(b_n) = o(1)$.

۱۴) $o(a_n) + o(b_n) = o(\max(|a_n|, |b_n|))$

تعریف. ۴.۱ اگر $X_n \xrightarrow{p} 0$، آن‌گاه $X_n = o_p(1)$. یعنی

$$\forall \varepsilon > 0, \lim_{n\to\infty} p(|X_n| < \varepsilon) = 1$$

یا به طور معادل

$$\forall \varepsilon > 0, \eta > 0 \exists n(\varepsilon, \eta) \ni n > n(\varepsilon, \eta) \Rightarrow p(|X_n| < \varepsilon) \geq 1 - \eta.$$

بنابراین می‌توان گفت اگر $X_n = o(1)$، آن‌گاه $X_n = o_p(1)$.