۴-۱ بردار تصادفی

بردار تصادفی مانند $\mathbb{X}$ برداری است که یکایک عناصر و مؤلفه های آن مانند x_i یک متغیر تصادفی باشد .

$$\mathbb{X} = [X_1, \ldots\ldots X_n] \quad (۴-۱)$$

احتمال این که $\mathbb{X}$ در ناحیه D از فضای n بعدی قرار داشته باشد ،

با جرم‌های احتمال در D برابر است یعنی

$$P\{\mathbb{X} \in D\} = \int_D f(\mathbb{X})\, d\mathbb{X} \quad , \quad \mathbb{X} = [X_1, \ldots, X_n] \quad (۴-۲)$$

که در آن

$$f(\mathbb{X}) = f(x_1, \ldots, x_n) = \frac{\partial^n F(x_1, \ldots, x_n)}{\partial x_1 \ldots\ldots \partial x_n} \quad (۴-۳)$$

تابع چگالی احتمال توأم (یا چند متغیره) متغیرهای تصادفی X_i
(یا تابع چگالی احتمال بردار $\mathbb{X}$) بوده و

$$F(\mathbb{X}) = F(x_1, \ldots, x_n) = P\{X_1 \le x_1, \ldots, X_n \le x_n\} \quad (۴-۴)$$

تابع توزیع توأم متغیرهای تصادفی X_i (یا تابع توزیع بردار $\mathbb{X}$) می باشد .
دو خاصیت زیر به آسانی قابل مشاهده و استنتاج است

$$F(x_1, x_3) = F(x_1, \infty, x_3, \infty)$$

(۴-۵)

$$f(x_1, x_3) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1, x_2, x_3, x_4)\, dx_2\, dx_4$$

اکنون k تابع زیر را در نظر بگیرید :

$$g_1(\mathbb{X}), \ldots\ldots g_k(\mathbb{X}) \quad , \quad \mathbb{X} = [X_1, \ldots\ldots, X_n]$$

متغیرهای تصادفی جدیدی را تشکیل می دهیم .

$$Y_1 = g_1(\mathbb{X}), \ldots\ldots Y_k = g_k(\mathbb{X}) \qquad (۴-۶)$$

$$k = n$$

بوده و سیستم معادلات زیر را حل می کنیم

$$g_1(\mathbb{X}) = Y_1, \ldots\ldots\ldots, g_n(\mathbb{X}) = Y_n \qquad (۴-۷)$$

اگر سیستم معادلات فوق جواب نداشته باشد در آن صورت $f_y(y_1, \ldots, y_n) = ۰$

است . اگر تنها جواب این سیستم عبارت باشد از $\mathbb{X} = [x_1, \ldots, x_n]$ **، در آن صورت**

$$f_y(y_1, \ldots\ldots, y_n) = \frac{f_X(x_1, \ldots\ldots, x_n)}{|j(x_1, \ldots, x_n)|} \qquad (۴-۸)$$

که در آن

$$j(x_1, \ldots, x_n) = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \cdots\cdots\cdots & \frac{\partial g_1}{\partial x_n} \\ \cdots\cdots & \cdots\cdots\cdots & \cdots\cdots \\ \frac{\partial g_n}{\partial x_1} & \cdots\cdots\cdots & \frac{\partial g_n}{\partial x_n} \end{vmatrix}$$

ژاکوبین تبدیل (۷-۴) است . اگر سیستم معادلات چند
جواب داشته باشد در آن صورت باید توابع چگالی
حاصل از این جواب ها را با هم جمع کرد .

متغیرهای تصادفی $X_1 , ... , X_n$
را هنگامی (متقابلاً) مستقل می‌نامند که پیشامدهای

$$\{ X_1 \leq x_1 \} , ... , \{ X_n \leq x_n \}$$

از یکدیگر مستقل باشند . بنابراین می‌توان گفت :

$$F (x_1 , ... , x_n) = F (x_1) ..... F (x_n) \quad (۴-۱۰)$$
$$f (x_1 , ... , x_n) = f (x_1) ..... f (x_n)$$

می‌توان مفهوم استقلال را تعمیم داده و استقلال گروهی را تعریف کرد .

گروه G_x متشکل از متغیرهای تصادفی $X_1 , ... , X_n$

را هنگامی مستقل از گروه G_y متشکل از متغیرهای تصادفی $Y_1 , ... , Y_k$
می نامیم که رابطه زیر برقرار باشد :

$$f (x_1 , ... ,x_n , y_1 , ... ,y_k) = f (x_1 , ... ,x_n) f (y_1 , ... ,y_k)$$
(۴-۱۱)

متوسط و کواریانس

با تعمیم رابطه (۳۶-۳) به n متغیر تصادفی می‌توان نتیجه گرفت که متوسط

$g(X_1,...,X_n)$ عبارت است از

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g(x_1,...,x_n)\, f(x_1,...,x_n)\, dx_1 \ldots dx_n \qquad (۴-۱۲)$$

اگر متغیرهای تصادفی $Z_i = X_i + jY_i$ مختلط باشند در آن صورت متوسط

$g(Z_1,...,Z_n)$ برابر است با

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g(z_1,...,z_n)\, f(x_1,...,x_n, y_1,...,y_n)\, dx_1 \ldots dy_n$$

و به ازاء هر بردار تصادفی حقیقی یا مختلط $\mathbb{X}$ می‌توان گفت :

$$E\{a_1 g_1(\mathbb{X}) + \ldots + a_m g_m(\mathbb{X})\} = a_1 E\{g_1(\mathbb{X})\} + \ldots + a_m E\{g_m(\mathbb{X})\}$$

مثال ۱-۴ :

متغیرهای تصادفی

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad , \quad \bar{V} = \frac{1}{n-1}\sum_{i=1}^{n} (X_i - \bar{X})^2$$

را بنا به تعریف به ترتیب متوسط نمونه و واریانس نمونه X_i می نامند.

نشان دهید که اگر متغیرهای تصادفی X_i

ناهم بسته با متوسط یکسان $E(X_i) = \eta$ و واریانس یکسان $\sigma_i^2 = \sigma^2$

باشند ، در آن صورت $E\{\bar{X}\} = \eta$, $\sigma_{\bar{X}}^2 = \sigma^2/n$, $E\{\bar{V}\} = \sigma^2$

است . برای حل با استناد به خاصیت خطی امید ریاضی می‌توان نوشت :

$$E\{\bar{X}\} = E\left\{\frac{1}{n}\sum_{i=1}^{n} X_i\right\} = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \eta = \eta$$

$$\sigma_{\bar{X}}^{2} = E\{\ |\ \bar{X} - E(\bar{X})\ |^{2}\}$$

$$= E\{\ |\ \frac{1}{n}\sum_{i=1}^{n} X_i - \ \eta\ |^{2}\} = E\{\ |\ \frac{1}{n}\sum_{i=1}^{n} (X_i - \eta)\ |^{2}\}$$

$$= \frac{1}{n^{2}}\sum_{i=1}^{n} E\left\{X_i - \eta\right\}^{2} = \frac{1}{n^{2}}\sum_{i=1}^{n} \sigma_i^{2} = \frac{1}{n^{2}}\sum_{i=1}^{n} \sigma^{2} = \sigma^{2}/n$$

هم چنین می توان نوشت :

$$E\{(X_i - \eta)(\bar{X} - \eta)\} = \frac{1}{n}\ E\{(X_i - \eta)\ [(X_1 - \eta)+....+(X_n - \eta)]\}$$

$$= \frac{1}{n} E\{(X_i - \eta)(X_i - \eta)\} = \sigma^{2}/n$$

چون بنا به فرض متغیرهای تصادفی X_i و X_j ناهم بسته هستند . بنابراین

$$E\{(X_i - \bar{X})^{2}\} = E\{[(X_i - \eta) - (\bar{X} - \eta)]^{2}\} =$$

$$\sigma^{2} + \frac{\sigma^{2}}{n} - \frac{2\sigma^{2}}{n} = \frac{n-1}{n}\sigma^{2}$$

پس

$$E\{\bar{V}\} = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^{2} = \frac{n}{n-1} \times \frac{n-1}{n}\sigma^{2} = \sigma^{2}$$

ماتریس همبستگی بردار تصادفی

$$\mathbb{X} = [\, X_1 \,, \,\ldots\ldots\, , X_n \,]$$

و نیز ماتریس کوواریانس این بردار بنا به تعریف عبارت است از :

$$R_n = \begin{bmatrix} R_{11} & \ldots\ldots\ldots & R_{1n} \\ \ldots & \ldots\ldots\ldots\ldots & \ldots \\ R_{n1} & \ldots\ldots\ldots & R_{nn} \end{bmatrix} \qquad C_n = \begin{bmatrix} C_{11} & \ldots\ldots\ldots & C_{1n} \\ \ldots & \ldots\ldots\ldots\ldots & \ldots \\ C_{n1} & \ldots\ldots\ldots & C_{nn} \end{bmatrix}$$

به طوری که

$$R_{ij} = E\,\{\, X_i\, X_j^* \} = R_{ji}^* \qquad , \qquad C_{ij} = R_{ij} - \eta_i \eta_j^* = C_{ji}^*$$

بدیهی است که

$$R_n = E\,\{\, \mathbb{X}^t\, \mathbb{X}^* \}$$

بوده و در آن $\mathbb{X}^t$ ترانسپوزه بردار $\mathbb{X}$ (بردار ستونی) می باشد . در سطور زیرین خواص ماتریس R_n و دترمینان آن Δ_n را مورد بررسی قرار می دهیم .

باید توجه کرد که خواص ماتریس کوواریانس C_n مشابه خواص ماتریس R_n می باشد ، چون C_n در واقع ماتریس همبستگی متغیرهای تصادفی " تمرکز یافته " $X_i - \eta_i$ هستند .

۱ – ماتریس R_n یک ماتریس معین غیر منفی است یعنی :

$$Q = \sum_{i,j} a_i\, a_j^* R_{ij} = \mathbb{A} R_n \mathbb{A}^h \geq 0 \qquad (۴-۱۳)$$

که در آن $\mathbb{A}^h$ بردار ترانسپوزه مزدوج مختلط بردار $\mathbb{A} = [\, a_1 \,, \ldots\ldots , a_n \,]$ است

برای اثبات این خاصیت می توان به ویژگی خطی بودن مقادیر امید ریاضی استناد کرد

$$E\{|a_1 X_1 + ... + a_n X_n^2|^2\} = \sum_{i,j} a_i a_j^* \; E[X_i X_j^*] \quad (۴-۱۴)$$

اگر به ازاء هر $\mathbb{A} \neq 0$ تابع Q مطلقاً مثبت باشد یعنی اگر

$Q > 0$ باشد در آن صورت R_n را معین مثبت می نامند .

تفاوت بین $Q > 0$ و $Q \geq 0$ با مستقل خطی بودن مرتبط است .

۲- متغیرهای تصادفی X_i را هنگامی مستقل خطی می نامند

که به ازاء هر $\mathbb{A} \neq 0$ رابطهٔ زیر برقرار باشد.

$$E\{|a_1 X_1 + ... + a_n X_n|^2\} > 0 \quad (۴-۱۵)$$

در این حالت ماتریس هم بستگی آنها معین مثبت خواهد بود.

متغیرهای تصادفی X_i را هنگامی وابسته خطی می نامیم

که به ازاء برخی $\mathbb{A} \neq 0$ رابطهٔ زیر برقرار باشد.

$$a_1 X_1 + ... + a_n X_n = 0 \qquad (۴-۱۶)$$

در این مورد ،Q مربوطه برابر صفر بوده و ماتریس R_n تکین (Singular)

خواهد بود . با توجه به تعریف فوق می‌توان نتیجه گرفت که اگر متغیرهای تصادفی X_i

مستقل خطی باشند در آن صورت هر زیر مجموعه نیز مستقل خطی خواهد بود.

دترمینان Δ_n ماتریس هم بستگی حقیقی است چون $R_{ij} = R_{ji}^*$ است .

هم چنین می توان نشان داد که $\Delta_n \geq 0$ بوده

و حالت تساوی هنگامی رخ می دهد که متغیرهای تصادفی X_i

با یکدیگر رابطه خطی داشته باشند (وابسته خطی) .

در خاتمه باید اضافه کرد که

$$\Delta_n \leq R_{11} R_{22} \dots R_{nn}$$

و حالت تساوی هنگامی برقرار خواهد بود که

متغیرهای تصادفی X_i (متقابلاً) متعامد باشند یعنی ماتریس R_n قطری باشد

توابع چگالی احتمالی شرطی و توابع مشخصه

مشابه مورد دو متغیر تصادفی ،

می‌توان تابع چگالی احتمال شرطی متغیرهای تصادفی

$X_{k+1}, \dots\dots, X_n$ مشروط به $X_1, \dots\dots, X_k$ را به صورت زیر تعریف کرد.

$$f(x_n, \dots, x_{k+1} / x_k, \dots, x_1) = \frac{f(x_1, \dots, x_k, \dots, x_n)}{f(x_1, \dots, x_k)} \qquad (۴-۱۷)$$

تابع توزیع احتمال مربوطه نیز بر اساس رابطهٔ زیر قابل تعیین است

$$F(x_n,\dots,x_{k+1}/x_k,\dots,x_1)= \qquad (۴-۱۸)$$
$$=\int_{-\infty}^{x_n}\dots\int_{-\infty}^{x_{k+1}} f(\alpha_n,\dots,\alpha_{k+1}/x_k,\dots,x_1)\,d\alpha_{k+1}\dots d\alpha_n$$

برای مثال می‌توان نوشت

$$f(x_1/x_2,x_3)=\frac{f(x_1,x_2,x_3)}{f(x_2,x_3)}=\frac{dF(x_1/x_2,x_3)}{dx_1}$$

قاعده زنجیره ای زیر نیز رابطهٔ مفیدی است

$$f(x_1,\dots,x_n)=f(x_n/x_{n-1},\dots,x_1)\dots\dots f(x_2/x_1)f(x_1) \qquad (۴-۱۹)$$

هم چنین به رابطهٔ زیر که به طور گسترده‌ای به کار می‌رود باید توجه کافی کرده و آن را به موارد مشابه دیگر تعمیم داد.

$$f(x_1/x_3)=\int_{-\infty}^{\infty} f(x_1/x_2,x_3)f(x_2/x_3)dx_2 \qquad (۴-۲۰)$$

قاعده فوق به متغیرهای تصادفی گسسته نیز قابل اعمال است . برای مثال اگر متغیر تصادفی X_1,X_2,X_3 به ترتیب مقادیر c_r,b_k,a_i را اختیار کنند ، در آن صورت

$$P[X_1=a_i/X_3=c_k]=\sum_k P[X_1=a_i/b_k,c_r]P[X_2=b_k/c_r]$$

بدیهی است به منظور تعیین متوسط شرطی یک متغیر تصادفی از تابع چگالی احتمال شرطی آن متغیر باید استفاده کرد یعنی

$$E\{X_1/x_2,\ldots,x_n\}=\int_{-\infty}^{\infty}x_1 f(x_1/x_2,\ldots,x_n)dx_1 \qquad (۴-۲۱)$$

اگر بخواهیم یکی از متغیرهای موجود در شرط تابع چگالی احتمال شرطی را حذف کنیم باید از مفهوم زیر استفاده کرد

$$E\{X_1/x_3\}=\int_{-\infty}^{\infty}E(X_1/x_2,x_3)f(x_2/x_3)dx_2 \qquad (۴-۲۲)$$

و در مورد حالت گسسته داریم

$$E\{X_1/c_r\}=\sum_k E(X_1/b_k,c_r)P(X_2=b_k/c_r) \qquad (۴-۲۳)$$

تابع مشخصه یک بردار تصادفی بنا به تعریف عبارت است از

$$\Phi(\Omega)=E\{e^{j\Omega\mathbb{X}^t}\}=E\{e^{j(\omega_1X_1+\ldots+\omega_nX_n)}\}=\Phi(j\Omega) \qquad (۴-۲۳)$$

که در آن $\mathbb{X}=[X_1,\ldots,X_n]$ و $\Omega=[\omega_1,\ldots,\omega_n]$ می باشد.

به عنوان یک کاربرد از تعریف و مفهوم فوق فرض کنید که $Z=X_1+\ldots+X_n$ بوده و متغیرهای تصادفی X_i مستقل از یکدیگر می‌باشند .

می‌توان گفت

$$E\{e^{j(\omega_1 X_1+\ \ldots\ +\omega_n X_n)}\}=E\{e^{j\omega_1 X_1}\}\cdots\cdots E\{e^{j\omega_n X_n}\}$$

پس

$$\Phi_z(\omega)=E\{e^{j\omega(X_1+\cdots\cdots+X_n)}\}=\Phi_1(\omega).\ .\ .\ .\ .\ \Phi_n(\omega)$$

که $\Phi_i(\omega)$ تابع مشخصه X_i می باشد. اگر از رابطهٔ فوق تبدیل فوریه عکس گرفته و همراه آن از خاصیت کانوولوشن - ضرب این تبدیل استفاده نماییم نتیجه عبارت خواهد بود :

$$f_z(3)=f_1(3)*f_{2}(3)*\ \ \cdots\cdots\ \ *f_n(3)$$

مثال ۲-۴

تابع چگالی احتمال و تابع مشخصهٔ یک بردار نرمال با متوسط صفر را تعیین کنید.

بردار نرمال به طول n برداری است که مولفه های آن n متغیر تصادفی توأماً نرمال باشند.

به عبارت دیگر متغیرهای تصادفی X_i هنگامی توأماً نرمال هستند که هرگونه ترکیب خطی آنها نیز خودیک متغیر نرمال باشد. یعنی به ازاء هر $\mathbb{A}$

$$a_1X_1+\ \cdots\cdots\ +a_nX_n=\mathbb{A}\mathbb{X}^t$$

خود نیز یک متغیر نرمال باشد .

با توجه به تعریف توأماً نرمال بودن می‌توان گفت که

W نیز یک متغیر نرمال است .

$$W = \omega_1 X_1 + ... + \omega_n X_n = \Omega \mathbb{X}^t$$

با توجه به فرض $E[X_i] = ۰$

می‌توان نتیجه گرفت که

$$E\{W\} = ۰ \quad , \quad E\{W^۲\} = \sum_{i,j} w_i w_j C_{ij} = \sigma_W^۲$$

با استفاده از تابع مشخصه یک متغیر تصادفی نرمال که در آن

$\omega = ۱ \quad , \quad \eta = ۰$ است ، می توان نوشت :

$$E\{e^{jw}\} = \exp\left[-\frac{\sigma_W^۲}{۲}\right]$$

و بنابراین

$$\Phi(\Omega) = E\{e^{j\Omega \mathbb{X}^t}\} = \exp\left\{-\frac{۱}{۲}\sum_{i,j} \omega_i \omega_j C_{ij}\right\} \qquad (۴\text{-}۲۴)$$

حال اگر از رابطه فوق تبدیل فوریه عکس بگیریم . داریم

$$f(\mathbb{X}) = \frac{۱}{\sqrt{(۲\pi)^n \Delta}} \exp\left\{-\frac{۱}{۲}\mathbb{X} C^{-۱} \mathbb{X}^t\right\} \qquad (۴\text{-}۲۵)$$

که Δ دترمینان ماتریس C است .

باید توجه کرد که در حالت خاص ، اگر متغیرهای تصادفی X_i

توأماً نرمال و ناهم بسته باشند ، این متغیرها مستقل بوده

و در نتیجه ماتریس کوواریانس آنها قطری با عناصر قطری $\sigma_i^۲$ خواهد بود .

در این حالت ماتریس C^{-1} نیز قطری با عناصر قطری $\frac{1}{\sigma_i^2}$

بوده و تابع چگالی احتمال این بردار نرمال عبارت خواهد بود از

$$f(x_1, \ldots\ldots, x_n) =$$

$$= \frac{1}{\sigma_1 \ldots\ldots \sigma_n \sqrt{(2\pi)^n}} \exp\left\{-\frac{1}{2}\left(\frac{x_1^2}{\sigma_1^2} + \ldots\ldots + \frac{x_n^2}{\sigma_n^2}\right)\right\}$$

مثال ۳-۴

اگر X_i متغیرهای تصادفی توأماً نرمال با متوسط صفر بوده و

$E\{X_i X_j\} = C_{ij}$ باشد ثابت کنید که

$$E\{X_1 X_2 X_3 X_4\} = C_{12}C_{34} + C_{13}C_{24} + C_{14}C_{23}$$

توابع نمایی دو طرف راست و چپ رابطه (۲۴-۴) را بسط داده

و فقط جملات شامل عبارت $\omega_1\omega_2\omega_3\omega_4$

را نشان می دهیم .

$$E\{e^{j(\omega_1X_1 + \ldots + \omega_4X_4)}\} = \ldots\ldots + \frac{1}{4!}E\{(\omega_1X_1 + \ldots + \omega_4X_4)^4\} + \ldots\ldots$$

$$= \ldots\ldots + \frac{24}{4!}E\{X_1X_2X_3X_4\}\,\omega_1\omega_2\omega_3\omega_4$$

از طرف دیگر

$$\exp\{-\frac{1}{2}\sum_{i,j}\omega_i\omega_j C_{ij}\} = +\frac{1}{2}(\frac{1}{2}\sum_{i,j}\omega_i\omega_j C_{ij})^2 + \dots\dots$$

$$= \dots\dots + \frac{8}{8}(C_{12}C_{34} + C_{13}C_{24} + C_{14}C_{23})\,\omega_1\omega_2\omega_3\omega_4$$

با مساوی قراردادن ضرایب $\omega_1\omega_2\omega_3\omega_4$

در دو بسط فوق رابطه مورد نظر اثبات می شود .

۲-۴- تخمین متوسط مربع

موضوع تخمین یکی از مفاهیم اساسی در کاربردهای احتمال بوده و در فصول آتی با جزئیات بیشتر مورد بررسی و بحث قرار خواهد گرفت . در این بخش ایده اصلی به صورت تخمین یک متغیر تصادفی مانند Y بر حسب متغیر تصادفی دیگر مانند X مطرح و معرفی می گردد .

در سراسر این تحلیل ، معیار بهینه بودن ، حداقل سازی مقدار متوسط مربع خطای (MSE)(Mean Square Error) تخمین است .
در تخمین با معیار MSE می‌توان از سه روش زیر استفاده کرد .

i) تخمین با مقدار ثابت

در این نوع تخمین ، می خواهیم متغیر تصادفی مجهول Y

را با مقدار ثابت C

به نحوی تخمین بزنیم که گشتاور دوم تفاضل (خطا)

$Y - C$ حداقل شود .

$$e = E\{(y - C)^{۲}\} = \int_{-\infty}^{\infty} (y - C)^{۲} f(y)\, dy \qquad (۴-۲۶)$$

بدیهی است e به مقدار ثابت C وابسته بوده

و حداقل آن عبارت است از :

$$\frac{de}{dC} = \int_{-\infty}^{\infty} ۲(y - C) f(y) dy = ۰$$

یعنی اگر

$$C = \int_{-\infty}^{\infty} y\, f(y) dy = E(y) \qquad (۴-۲۷)$$

در آن صورت متوسط مربع خطای تخمین حداقل خواهد بود .

این نتیجه در واقع همان اصل معروف در مکانیک است ،

که بر اساس آن گشتاور اینرسی (اجرام) نسبت به نقطه C

هنگامی حداقل است که C مرکز ثقل آن جسم (اجرام) باشد .

هم چنین قابل توجه است که متوسط مربع خطای

حداقل در واقع واریانس متغیر تصادفی Y خواهد بود .

$$\text{Min } E\{(Y-C)^{۲}\} = E\{[Y-E(Y)]^{۲}\} = \sigma_y^{۲}$$

ii) تخمین غیرخطی

در این حالت متغیر تصادفی Y را نه با مقدار ثابت بلکه به صورت تابعی از متغیر تصادفی X (داده ها) یعنی $C(X)$ تخمین می زنیم . در اینجا نیز تابع $C(X)$ را باید به نحوی تعیین کرد که متوسط مربع خطای تخمین حداقل باشد .

$$e = E\{[Y - C(X)]^2\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [y - c(x)]^2 f(x, y)\,dx\,dy$$

رابطه فوق را می‌توان به صورت زیر نوشت

$$e = \int_{-\infty}^{\infty} f(x) \int_{-\infty}^{\infty} [y - c(x)]^2 f(y/x)\,dy\,dx$$

از آنجا که انتگراندها (توابع زیر انتگرال) مثبت هستند می‌توان ادعا کرد ، که برای حداقل شدن متوسط مربع خطا یا e کافی است ، انتگرال داخلی به ازاء هر مقدار x حداقل شود . این انتگرال به شکل انتگرال رابطه (۴-۲۶) می باشد که در آن C به $C(x)$ و $f(y)$ به $f(y/x)$ تبدیل شده است ، بنابراین می‌توان نتیجه گرفت که انتگرال مذکور هنگامی حداقل است که $C(x)$ به شکل (۴-۲۷) تعریف شود . مشروط بر آن که $f(y)$ به $f(y/x)$ تغییر یابد یعنی

$$C(x) = E\{y/x\} = \int_{-\infty}^{\infty} y\, f(y/x)\,dy \qquad (۴-۲۸)$$

بدیهی است که در این حالت متوسط مربع خطاهای حداقل واریانس شرطی Y یعنی $\sigma^2 Y/x$ خواهد بود . باید توجه کرد که در حالت کلی تابع $C(X)$

یک تابع غیر خطی از متغیر تصادفی داده های X می باشد .
نکته دیگری که باید روی آن تاکید شود عبارت است از این که
اگر Y,X مستقل از یکدیگر باشند در آن صورت

مقدار ثابت $= E\{Y/x\} = E\{Y\}$

بوده و در این حالت اطلاع از X هیچ گونه
نقشی در تخمین Y ندارد .

iii) تخمین خطی

همانگونه که ملاحظه شد در تخمین MSE غیرخطی به توابع چگالی احتمال توأم یا شرطی نیاز است . در بسیاری از موارد و کاربردها این توابع احتمال نامعلوم بوده و در دسترس نیستند و بدین سبب آسان تر خواهد بود
اگر بتوان به جای توابع احتمال از گشتاورهای اول و دوم استفاده نمود .
در واقع چنین شرایطی در تخمین خطی وجود داشته و باعث می‌گردد ،
که این نوع تخمین به خاطر سهولت به طرز گسترده ای به کار رود ،
اگر چه فاقد دقت تخمین غیرخطی می‌باشد .
در این نوع تخمین می‌خواهیم متغیر تصادفی مجهول Y را بر حسب تابع خطی $AX+B$ تخمین بزنیم . ضرایب B , A
باید به نحوی تعیین شوند که متوسط مربع خطای تخمین زیر حداقل گردد .

$$e = E\{[Y-(AX+B)]^2\} \qquad (4-29)$$

با گرفتن مشتق از e نسبت ضرایب A , B و مساوی صفر قرار دادن آنها مقادیر زیر به دست می‌آید

$$A = \frac{\mu_{11}}{\mu_{20}} = \frac{r\sigma_y}{\sigma_x} \quad , \quad B = \eta_y - A\eta_x \qquad (4-30)$$

و بالطبع متوسط مربع خطای حداقل در تخمین خطی برابر است با

$$e = e_m = \mu_{02} - \frac{\mu_{11}^2}{\mu_{20}} = \sigma_y^2 (1 - r^2)$$

که در روابط فوق $r = \rho$ ضریب هم بستگی بین دو متغیر تصادفی Y, X می باشد .

در تخمین خطی فوق یعنی $C(X)=AX+B$ را تخمین خطی غیر همگن (Nonhomogeneous) Y بر حسب X نامند .

اگر تخمین خطی Y برحسب X
بدون مقدار ثابت باشد یعنی $C(X)=AX$ ($B=0$) ، در آن صورت تخمین خطی را تخمین خطی همگن (Homogeneous) Y برحسب X گویند .
نکته شایان توجه در این رابطه موردی است که دو متغیر تصادفی Y, X متغیرهای تصادفی توأماً نرمال باشند . در این حالت تخمین غیرخطی بر اساس معیار متوسط مربع خطای حداقل برابر با تخمین خطی با همان معیار است . به عبارت دیگر تخمین خطی و غیرخطی (MSE) دو متغیر تصادفی توأماً نرمال برابر بوده و به همین علت می‌توان گفت که
بهترین تخمین دو متغیر تصادفی توأماً نرمال همان تخمین خطی آنها است .

اصل تعامد Orthogonality

بر اساس این اصل ، متوسط مربع خطای تخمین خطی هنگامی حداقل است

که خطای تخمین $Y-(AX+B)$ بر داده های X متعامد باشد یعنی

$$E\{[Y-(AX+B)]X\}=0 \qquad (۴-۳۱)$$

اصل مذکور در واقع پایه و اساس تخمین MS بوده و بطور گسترده ای به کار می‌رود.

در این بخش ما آن را در حالت تخمین خطی همگن اثبات می‌کنیم .

در چنین تخمینی ، یعنی aX فرض کنید که خطای تخمین MS

$(Y-aX)$ بر X متعامد است یعنی

$$E\{(Y-aX)X\}=0$$

است (متغیر تصادفی داده های X حقیقی فرض شده است) .

حال تخمین خطی دیگری مانند bX را در نظر گرفته

و نشان می‌دهیم که خطای تخمین MS آن از خطای تخمین فوق بیشتر است

(به ازاء هر مقدار b) . بدین منظور می‌توان نوشت

$$E\{(Y-bX)^2\}=E\{(Y-aX+aX-bX)^2\}=E\{[(Y-aX)+(a-b)X]^2\}$$
$$=E\{[Y-aX)]^2\}+(a-b)^2E[X^2]+2(a-b)E\{[Y-aX)X]\}$$

جمله آخر برابر صفر بوده و جمله دوم همیشه مثبت است پس می‌توان نتیجه گرفت

که به ازاء جمیع مقادیر b

$$E\{[(Y-bX)]^2\}\geq E\{[(Y-aX)]^2\}$$

پس باید گفت که اگر خطای تخمین بر داده ها متعامد باشد متوسط مربع خطای

تخمین حداقل خواهد بود .

تخمین خطی MS متغیر تصادفی Y بر حسب X را به شکل $\hat{E}\{Y/X\}$ نشان می‌دهند .

بنابراین

$$\hat{E}\{Y/X\}=aX \qquad , \qquad a=\frac{E\{XY\}}{E\{X^2\}} \qquad (۴-۳۲)$$

و خطای MS برابر است .

$$e = E\{(Y-aX)Y\} - E\{(Y-aX)aX\}=E\{Y^2\}-E\{(aX)^2\} \qquad (۴-۳۳)$$

در رابطه فوق عبارت دوم بنا به اصل تعامد صفر بوده و به جای Y تخمین خطی aX جایگزین شده است .

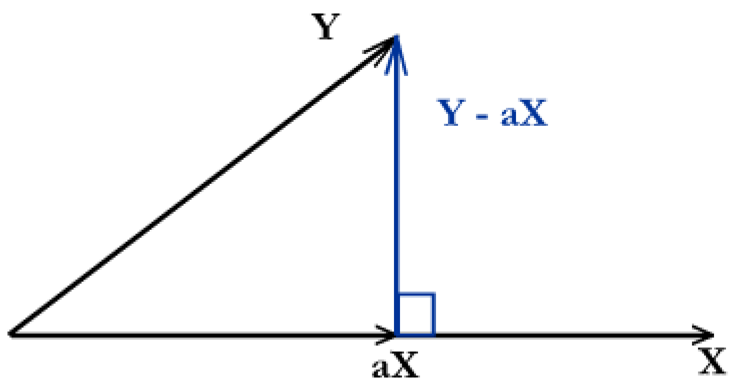
تعبیر مفید دیگر برای اصل تعامد در واقع تعبیر هندسی است .

اگر متغیرهای تصادفی را بر طبق شکل ۱-۴ به صورت بردار نمایش دهیم

بردار تفاضل $Y- aX$ برداری است که از نقطه aX روی خط X به Y

وصل شده و طول این بردار برابر $\sqrt{e}$ است .

بدیهی است که این طول هنگامی حداقل خواهد بود که $Y-aX$ بر X عمود باشد .



شکل ۱-۴

تخمین خطی MS یک متغیر تصادفی مجهول مانند S را می‌توان برحسب متغیرهای تصادفی X_i (بردار داده ها) بیان کرد . به عبارت دیگر $\hat{S}$ یا تخمین خطی S عبارت است از

$$\hat{S} = a_1 X_1 + \dots\dots + a_n X_n \qquad (۴-۳۴)$$

که در آن $a_1, \dots, a_n$ ضرایب ثابت بوده و به نحوی تعیین می‌شوند ، که مقدار متوسط مربع خطای زیرحداقل گردد

$$P = E\{(S - \hat{S})^۲\} = E\{[S - (a_1 X_1 + \dots + a_n X_n)]^۲\} \qquad (۴-۳۵)$$

با توجه به اصل تعامد ، P هنگامی حداقل است که خطای تخمین $S - \hat{S}$ بر داده های X_i متعامد باشد چون P هنگامی حداقل خواهد بود که

$$\frac{\partial P}{\partial a_i} = E\{-۲[S - (a_1 X_1 + \dots + a_n X_n)]X_i\} = ۰$$

و از رابطه فوق می‌توان رابطه تعامد را نتیجه گرفت

$$E\{[S - (a_1 X_1 + \dots + a_n X_n)]X_i\} = ۰ \quad ، \quad i = ۱,۲,\dots\dots,n \qquad (۴-۳۶)$$

مفاهیم فوق را می‌توان به صورت برداری بیان کرد . ابتدا بردارهای سطری را تعریف می‌کنیم .

$$\mathbb{X} = [X_1, \dots\dots\dots\dots, X_n] \quad , \quad \mathbb{A} = [a_1, \dots\dots\dots\dots, a_n]$$

http://vc.iust.ac.ir/file.php/99/Chapter%204/session2/4-asle%20taamod-edameh/index.swf
Screen clipping taken: ۲۶/۰۲/۲۰۰۸; ۱۰:۱۵ ب.ظ

اگر $R_{ij}=E\{X_iX_j\}$ ، $R_{\circ j}=E\{SX_j\}$ بوده و ماتریس هم بستگی داده ها $R=E\{\mathbb{X}^t\mathbb{X}\}$ و بردار هم بستگی مجهول و داده‌ها یعنی $R_\circ=[R_{\circ 1},........,R_{\circ n}]$ را در نظر بگیریم با استناد به اصل تعامد می‌توان نوشت:

$$E\{[S-\mathbb{A}\mathbb{X}^t]\mathbb{X}\}=\circ \quad , R_\circ-\mathbb{A}R=\circ$$
$$\mathbb{A}=R_\circ R^{-1} \qquad (۴-۳۷)$$

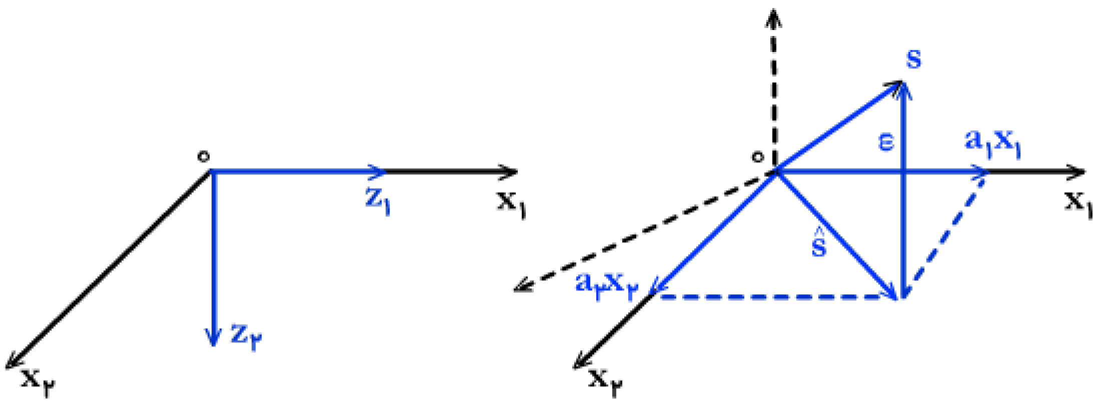
بنابراین ضرایب a_i از رابطه فوق به دست می‌آید . متوسط مربع خطای حداقل نیز از رابطه زیر قابل تعیین است . این رابطه با توجه به اصل تعامد و $S-\hat{S}\perp\hat{S}$ نتیجه می شود .

$$P=E\{(S-\hat{S})S\}=E\{S^۲\}-\mathbb{A}R_\circ^t \qquad (۴-۳۸)$$

تعبیر هندسی اصل تعامد نیز با توجه به شکل ۲-۴ قابل توصیف و بررسی می باشد .

بردار تخمین S برداری در زیر فضای S_n با ابعاد داده های X_i بوده و بردار خطای $\varepsilon=S-\hat{S}$ برداری است که از $\hat{S}$ به S وصل می شود .

قضیه تصویر (Projection Theorem) بیان می‌کند که طول ε هنگامی حداقل است که ε بر X_i عمود باشد یعنی بر زیر فضای S_n داده ها متعامد گردد .



(ب) شکل ۲-۴ (الف)

بنابراین می‌توان گفت که تخمین $\hat{S}$ "تصویر" S بر زیر فضای S_n است . اگر S برداری در S_n باشد در آن صورت $\hat{S}=S$ بوده و $P=0$ خواهد بود . در این مورد $n+1$ متغیر های تصادفی $S, X_1, ..., X_n$ به طور خطی وابسته بوده و دترمینان Δ_{n+1} ماتریس هم بستگی آنها صفر است . اگر S بر S_n متعامد باشد ، در آن صورت $\hat{S}=0$ ، $P=E\{|S|^2\}$ خواهد بود . در این مورد S بر تمام داده های X_i عمود بوده و به ازاء $j\neq 0$ ، $R_{0j}=E\{SX_j\}=0$ است

تخمین غیر خطی MS شامل تعیین تابع داده ها $g(X_1,.....,X_n)=g(\mathbb{X})$ به نحوی است که خطای MS

$$P=E\{[S - g(\mathbb{X})]^2\}$$

حداقل شود . به سهولت می توان نشان داد که برای حداقل شدن P کافی است که

$$g(X)=E(S/X)=\int_{-\infty}^{\infty} Sf_s(S/X)dS \qquad (۴-۳۹)$$

تابع $E(S/X)$ متوسط شرطی متغیر تصادفی مجهول S با فرض $\mathbb{X}=X$ است

مثال ۴-۴

متغیرهای تصادفی $X_۲, X_۱$ متغیرهای تصادفی توأماً نرمال با متوسط صفر می باشند

چگالی طیفی شرطی $f(x_۲/x_۱)$ را تعیین کنید .

با توجه به رابطه (۳۲-۴) می‌توان نوشت

$$E\{x_۲/x_۱\}=ax_۱ \quad , a=\frac{R_{۱۲}}{R_{۱۱}}$$

$$\sigma^۲_{x_۲/x_۱}=P=E\{(X_۲-aX_۱)X_۲\}=R_{۲۲}-aR_{۱۲}$$

بنابراین با توجه به اینکه چگالی احتمال شرطی مورد نظر نرمال است

می توان نتیجه گرفت

$$f(x_۲/x_۱)=\frac{۱}{\sqrt{۲\pi p}}e^{-(x_۲-ax_۱)^۲/۲p}$$

مثال ۵-۴

در مثال ۴-۴ فرض کنید متغیرهای $X_۲, X_۱$ معلوم بوده و متغیر $X_۳$

را بر اساس تخمین خطی MS بر حسب $X_۲, X_۱$ تعیین می‌نماییم

اکنون تابع چگالی احتمال $f(x_۳/x_۱,x_۲)$ را به دست آورید .

در این حالت

$$E\{X_۳/X_۱,X_۲\}=a_۱X_۱+a_۲X_۲$$

بوده و ضرایب $a_۲,a_۱$ با استفاده از اصل تعامد ، در واقع جواب دستگاه معادلات

زیر می‌باشد .

$$R_{۱۱}a_۱+R_{۱۲}a_۲=R_{۱۳} \quad , \quad R_{۱۲}a_۱+R_{۲۲}a_۲=R_{۲۳}$$

علاوه بر این

$$\sigma^{2}_{x_3/x_1,x_2} = P = R_{33} - (R_{13}a_1 + R_{23}a_2)$$

و تابع چگالی احتمال شرطی عبارت است از

$$f(x_3/x_1,x_2) = \frac{1}{\sqrt{2\pi p}}\; e^{-(x_3 - a_1x_1 - a_2x_2)/2p}$$

تبدیل ارتونرمال داده ها

اگر داده های X_i متعامد باشند یعنی اگر به ازاء $R_{ij} = 0$ ، $i \neq j$

باشد در آن صورت ماتریس R یک ماتریس قطری بوده و حل معادلهٔ

حاصل از اصل تعامد و محاسبهٔ ضرایب a_i

ساده تر شده و به صورت زیر قابل تعیین است

$$a_i = \frac{R_{0i}}{R_{ii}} = \frac{E\{SX_i\}}{E\{X_i^2\}} \qquad (۴-۴۰)$$

پس تعیین تصویر $\hat{S}$ از S ساده تر خواهد شد مشروط بر آن که

بتوان داده های X_i را بر حسب مجموعهٔ اورتونرمال از بردارها بیان کرد

بدین منظور می‌خواهیم مجموعهٔ $\{Z_k\}$ از n متغیر تصادفی اورتونرمال Z_k

را که معادل خطی مجموعهٔ داده های $\{X_k\}$ است تعیین کنیم.

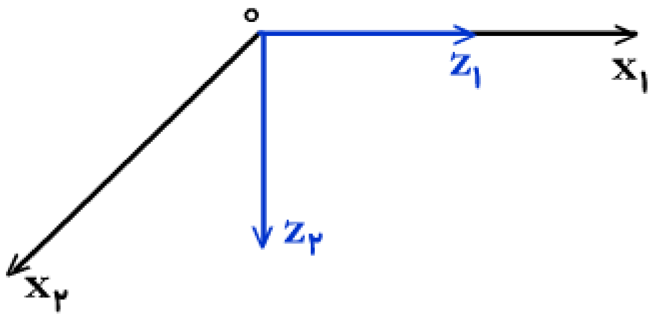
هدف از عبارت معادل خطی آن است که هر Z_k تابعی خطی از عناصر مجموعهٔ $\{X_k\}$

بوده و هر X_k تابع خطی از عناصر مجموعهٔ $\{Z_k\}$ است.

مجموعهٔ $\{X_k\}$ منحصر به فرد نیست و ما با استفاده از روش گرام-

اشمیت (Gram-Schmidt) آن را تعیین خواهیم کرد (شکل ۲-۴ ب).

در این روش، هر Z_k فقط به k داده های اول $X_k , \ldots\ldots , X_1$ وابسته است.



شکل (۲-۴ ب)

بنابراین

$$
\begin{aligned}
Z_1 &= \gamma_1^1 X_1 \\
Z_2 &= \gamma_1^2 X_1 + \gamma_2^2 X_2 \\
&\dots\dots\dots\dots\dots\dots \\
Z_n &= \gamma_1^n X_1 + \gamma_2^n X_2 + \dots\dots + \gamma_n^n X_n
\end{aligned}
\qquad (۴-۴۱)
$$

در نماد γ_r^k ، k اندیس نشان دهندهٔ معادلهٔ k ام و r اندیسی است که مقادیر ۱ تا k را اختیار می کند . ضریب γ_1^1 از شرط نرمالیزه کردن به دست می‌آیند.

$$E\{Z_1^2\} = (\gamma_1^1)^2 R_{11} = 1$$

برای تعیین ضرایب γ_1^2 و γ_2^2 مشاهده می کنیم که $Z_2 \perp X_1$ است .
چون بنا به تعریف $Z_2 \perp Z_1$ می‌باشد ، پس

$$E\{Z_2 X_1\} = 0 = \gamma_1^2 R_{11} + \gamma_2^2 R_{21}$$

شرط $E\{Z_2^2\} = 1$ معادله دوم را بیان می کند . مشابهاً از آنجا که $Z_k \perp Z_r$ به ازاء $r < k$ است نتیجه می گیریم که $Z_k \perp X_r$ به ازاء $r < k$ است.
اگر طرفین معادلهٔ k ام در رابطهٔ (۴-۴۱) را در X_r ضرب کرده و از مفاهیم فوق استفاده کنیم . می‌توان نوشت

$$E\{Z_k X_r\} = 0 = \gamma_1^k R_{1r} + \dots\dots + \gamma_k^k R_{kr} \quad , \quad 1 \le r \le k-1 \qquad (۴-۴۲)$$

سیستم معادلات فوق شامل $k-1$ معادله برای k

مجهول $\gamma_k^k \ldots\ldots \gamma_1^k$ می باشد. شرط $E\{Z_k^2\}=1$ معادلهٔ

دیگر خواهد بود . دستگاه معادلات (۴-۴۱) را به صورت

برداری می‌توان بازنویسی کرد.

$$\mathbb{Z} = \mathbb{X}\,\boldsymbol{\Gamma} \qquad (۴-۴۲)$$

که در آن $\mathbb{Z}$ یک بردار سطری با عناصر Z_k می باشد.

با حل معادلات فوق، جواب $\mathbb{X}$ را می توان به دست آورد.

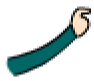
$$\mathbb{X} = \mathbb{Z}\boldsymbol{\Gamma}^{-1} = \mathbb{Z}\mathbb{L}$$

$$\begin{aligned}
X_1 &= \ell_1^1 Z_1 \\
X_2 &= \ell_1^2 Z_1 + \ell_2^2 Z_2 \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
X_n &= \ell_1^n Z_1 + \ell_2^n Z_2 + \ldots\ldots + \ell_n^n Z_n
\end{aligned} \qquad (۴-۴۳)$$

در روابط فوق ماتریس $\boldsymbol{\Gamma}$ و ماتریس عکس آن $\mathbb{L}$

هر دو ماتریس های مثلثی بالا هستند

$$\boldsymbol{\Gamma} = \begin{bmatrix} \gamma_1^1 & \gamma_1^2 \ldots\ldots & \gamma_1^n \\ & \gamma_2^2 \ldots\ldots & \gamma_2^n \\ & 0 \ldots\ldots & \gamma_n^n \end{bmatrix}$$



$$\mathbb{L} = \begin{bmatrix} \ell_1^1 & \ell_1^2 \ldots\ldots & \ell_1^n \\ & \ell_2^2 \ldots\ldots & \ell_2^n \\ & 0 \ldots\ldots & \ell_n^n \end{bmatrix}$$

از آنجا که بنا به تعریف $E\{Z_i Z_j\}=\delta(i-j)$ است نتیجه می‌گیریم که

$$E\{\mathbb{Z}^t\mathbb{Z}\}=1_n=E\{\Gamma^t\mathbb{X}^t\mathbb{X}\Gamma\}=\Gamma^t E\{\mathbb{X}^t\mathbb{X}\}\Gamma \qquad (۴-۴۴)$$

که در آن 1_n ماتریس واحد می باشد . بنابراین

$$\Gamma^t R\Gamma=1_n \quad , \quad R=\mathbb{L}^t\mathbb{L} \quad , \quad R^{-1}=\Gamma^t\Gamma \qquad (۴-۴۵)$$

پس همانگونه که ملاحظه می شود ما ماتریس R و معکوس آن R^{-1} را به صورت حاصلضرب ماتریس های مثلثی بالا و مثلثی پایین بیان کرده ایم.

مبنای اورتونرمال $\{Z_n\}$ در (۴-۴۱) در واقع نسخه محدود فرآیند های ابداع (Innovations) است که در فصول آتی مورد بررسی قرار می گیرد.

ماتریس های Γ , L به ترتیب به فیلتر سفید کننده و فیلتر ابداع مربوط بوده و رابطه (۴-۴۴) را تجزیه به عوامل طیفی (Spectral Factorization) می نامند

با توجه به معادل خطی بودن دو مجموعه $\{X_k\}, \{Z_k\}$ باید گفت که تخمین متغیر تصادفی S بر حسب داده های $\{X_k\}$ را می توان به صورت تخمین S برحسب مجموعه $\{Z_k\}$ نیز بیان کرد .

$$\hat{S}=b_1Z_1+\dots+b_nZ_n=\mathbb{B}\,\mathbb{Z}^t$$

که مجدداً ضرایب b_k با استفاده از اصل تعامد تعیین می شوند

$$S-\hat{S}\perp Z_k \qquad 1\le k\le n$$

یا

$$E\{(S-\mathbb{B}\mathbb{Z}^t)\mathbb{Z}\}=0=E\{S\mathbb{Z}\}-\mathbb{B}$$

و از آن می‌توان نتیجه گرفت که

$$\mathbb{B}=E\{S\mathbb{Z}\}=E\{S\mathbb{X}\Gamma\}=R_0\Gamma$$

بنابراین تخمین $\mathbb{S}$ عبارت است از

$$\mathbb{S} = \mathbb{B}\mathbb{Z}^t = \mathbb{B}\Gamma^t\mathbb{X}^t = \mathbb{A}\mathbb{X}^t \quad , \quad \mathbb{A} = \mathbb{B}\Gamma^t$$

در این حالت اگر ماتریس Γ معلوم و در دست باشد تعیین بردار $\mathbb{A}$ ساده تر می‌شود .

۳-۴- همگرایی تصادفی

یک مساله اساسی در نظریه احتمال تعیین خواص مجانبی دنباله های تصادفی است . در این بخش ضمن بررسی موضوع ، مفاهیم مربوطه را طبقه بندی کرده و بحث را با یک مثال ساده آغاز می‌کنیم . فرض کنید می خواهیم طول a یک جسم را اندازه بگیریم .

به علت خطا های اندازه گیری مقدار اندازه گرفته شده به صورت $X=a+\gamma$ خواهد بود که در آن γ عبارت خطا است .

اگر خطا های سیستماتیک رخ ندهد در این صورت γ یک متغیر تصادفی با متوسط صفر است . در این حالت اگر انحراف معیار σ مربوط به خطای γ نسبت به a کوچک باشد در آن صورت مقدار مشاهده شده $X(\xi)$ در یک اندازه گیری معین تخمین رضایت بخشی از طول مجهول a است . از دید مفاهیم احتمال این نتیجه را به این صورت می‌توان بیان کرد .

متوسط متغیر تصادفی X برابر a و واریانس آن برابر σ^2 می باشد .
با استناد به نامساوی چیی چف می توان نتیجه گرفت که

$$P\{\ |X - a| < \varepsilon\ \} > 1 - \frac{\sigma^2}{\varepsilon^2} \qquad (۴-۴۶)$$

بنابراین اگر $\varepsilon \gg \sigma$ باشد در این صورت احتمال این که $|X - a|$
کمتر از ε باشد به یک نزدیک است .
با توجه به مطلب فوق ، می توان گفت " تقریباً اطمینان "
داریم که مقدار مشاهده شده $X(\xi)$ بین $a - \varepsilon$ و $a + \varepsilon$
قرار دارد یا معادلاً مجهول a بین $X(\xi) - \varepsilon$ و $X(\xi) + \varepsilon$
می باشد . به عبارت دیگر ، قرائت $X(\xi)$
در یک اندازه گیری معین " تقریباً با اطمینان " تخمین رضایت بخشی از طول a

خواهد بود مشروط بر آن که $\sigma \ll a$
باشد . اگر σ در مقایسه با a کوچک نباشد ،
در آن صورت یک اندازه گیری به تنهایی نمی تواند تخمین کافی از a
را ارائه دهد . به منظور بهبود دقت اندازه گیری را به دفعات زیاد
انجام داده و از نتایج حاصله متوسط می گیریم .
مدل مبتنی بر احتمالات ، اکنون در فضای حاصلضربی

$$S^n = S \times \ldots\ldots\ldots\ldots \times S$$

که از n بار تکرار آزمایش یک اندازه گیری تنها
تشکیل یافته است قرار دارد .
اگر اندازه گیری ها مستقل باشند در آن صورت قرائت اندازه گیری i ام
عبارت است از مجموع $X_i = a + \gamma_i$

که در آن مؤلفه های نویز γ_i متغیرهای تصادفی مستقل با متوسط صفر و واریانس σ^2 است .

نکته مذکور منجر به این نتیجه می شود که متوسط نمونه اندازه گیری ها

$$\bar{X} = \frac{X_1 + \dots\dots\dots + X_n}{n}$$

خود نیز یک متغیر تصادفی با متوسط a و واریانس $\frac{\sigma^2}{n}$ می باشد . بنابراین اگر n به اندازه ای بزرگ باشد که $\sigma^2 \ll na^2$ باشد ، در آن صورت مقدار $\bar{X}(\xi)$ متوسط نمونه $\bar{X}$ در یک عملکرد تنها از آزمایش S^n (مرکب از n اندازه گیری مستقل) تخمین رضایت بخشی از مجهول a خواهد بود .

به منظور تعیین کران خطا برای تخمین a توسط $\bar{X}$، از رابطهٔ (۴-۴۶) استفاده

$$P\{|X - a| < \varepsilon\} > 1 - \frac{\sigma^2}{\varepsilon^2} \qquad (۴\text{-}۴۶)$$

کرده و فرض می‌کنیم که n به قدری بزرگ می باشد که $\frac{\sigma^2}{na^2} = 10^{-4}$ بوده و می‌خواهیم احتمال اینکه $0/9a < X < 1/1a$ قرار گیرد را حساب کنیم . جواب بر اساس رابطهٔ (۴-۴۶) به ازاء $\varepsilon = 0/1a$ برابر است با

$$P\{0/9a < \bar{X} < 1/1a\} \geq 1 - \frac{100\sigma^2}{na^2} = 0/99$$

بنابراین ، اگر آزمایش به تعداد $n = 10^4\ \sigma^2/a^2$ تکرار شود در آن صورت "تقریباً اطمینان" داریم که در ۹۹ درصد موارد تخمین $\bar{X}$ از a بین $0/9a$ و $1/1a$ خواهد بود. با توجه به بحث فوق، حالات همگرائی مختلف دنباله های متغیرهای تصادفی را می توان به شکل زیر مطرح کرد.

بنا به تعریف ، دنباله تصادفی یا فرآیند تصادفی

گسسته زمان در واقع دنباله ای

از متغیرهای تصادفی

$$X_1, \ldots\ldots, X_n, \ldots$$

می باشد . به ازاء یک ξ معین ، $X_n(\xi)$

دنباله ای از اعداد است که امکان دارد همگرا شود یا همگرا نشود. این نکته

نشان می‌دهد که همگرائی دنباله‌های تصادفی را می‌توان

به صور مختلف تعبیر و تفسیرنمود .

همگرایی در همه جا (Every where)(e)

دنباله ای از اعداد x_n هنگامی به حد x میل می‌کند که به ازاء $\varepsilon > 0$ مفروض

بتوان عددی مانند n_0 را تعیین کرد . و

(۴-۴۷)	به ازاء هر	$n > n_0$	$\lvert x_n - x \rvert < \varepsilon$

اگر دنبالهٔ اعداد $X_n(\xi)$ مطابق رابطهٔ (۴-۴۷) به ازاء هر ξ همگرا شود

در آن صورت گویند که دنبالهٔ تصادفی X_n در همه جا همگرا می‌شود . حد

همگرایی عددی است که به طور کلی به ξ بستگی دارد ، به عبارت دیگر حدّ

دنباله های تصادفی X_n متغیر تصادفی X است .

$$\text{به ازاء} \qquad n \to \infty \quad , \quad X_n \xrightarrow{e} X$$

همگرایی تقریباً در همه جا (Almost Everywhere)(ae)

اگر مجموعهٔ نتایج ξ به نحوی باشد که

$$\lim X_n(\xi) = X(\xi) \quad , \quad n \to \infty \text{ به ازاء} \qquad (۴-۴۸)$$

بوده و احتمال آن برابر یک باشد در آن صورت گوییم که دنبالهٔ X_n تقریباً در همه جا (یا با احتمال ۱) همگرا می‌شود.

این نوع همگرایی به شکل زیر توصیف می‌شود.

$$X_n \xrightarrow{ae} X \qquad P\{X_n \to X\} = ۱ \qquad n \to \infty \text{ به ازاء}$$

که در رابطهٔ فوق $\{X_n \to X\}$ پیشامدی شامل تمام نتایج ξ است که به ازاء آنها $X_n(\xi) \to X(\xi)$ می‌باشد

همگرایی به مفهوم متوسط مربع (Mean Square)(ms)

دنباله X_n هنگامی به متغیر تصادفی X از نظر مفهوم MS میل می‌کند که

$$E\{|X_n - X|^۲\} \to ۰ \qquad n \to \infty \text{ به ازاء} \qquad (۴-۵۰)$$

$$X_n \xrightarrow{MS} X$$

این نوع همگرایی را حد در متوسط (Limit in Time) نامیده و آن را اغلب به صورت زیر معرفی می کنند. $\text{l.i.m. } X_n = X \ , \ n \to \infty$

همگرایی در احتمال (P)

احتمال $P\{|X - X_n| > \varepsilon\}$ مربوط به پیشامد $\{|X - X_n| > \varepsilon\}$ دنباله ای از اعداد است که به ε بستگی دارند . اگر این دنباله به ازاء هر $\varepsilon > ۰$ به مقدار صفر (۰) میل می‌کند.

$$P\{|X - X_n| > \varepsilon\} \to 0 \quad n \to \infty \quad \text{به ازاء} \quad (۴-۵۱)$$

در آن صورت گوییم که دنباله X_n از نظر احتمال به متغیر تصادفی X همگرا می‌شود . این نوع همگرایی را همگرایی تصادفی (Stochastic Convergence) نامند .

$$X_n \xrightarrow{P} X$$

همگرایی در توزیع (d)

اگر توابع توزیع متغیرهای تصادفی X_n و X را به ترتیب با $F_n(x)$ و $F(x)$ نشان داده و اگر به ازاء هر نقطهٔ پیوسته ای از $F(x)$

$$F_n(x) \to F(x) \quad n \to \infty \quad \text{به ازاء} \quad (۴-۵۲)$$

در آن صورت گوییم که دنباله X_n از نظر توزیع به متغیر تصادفی X همگرا می‌شود . شایان توجه است که در این مورد ، لزومی ندارد ، دنبالهٔ $X_n(\xi)$ به ازاء هر ξ همگرا شود .

$$X_n \xrightarrow{d} X$$

همانطور که گفته شد ، دنباله یقینی x_n اگر در رابطهٔ (۴-۴۷) صدق کند ، همگرا می‌شود . این تعریف شامل مقدار حدّی x_n یعنی x می‌باشد .

قضیهٔ زیر که به معیار کوشی معروف است ، شرایطی را برای همگرائی x_n توصیف می کند که در آن از x استفاده نمی شود . اگر به ازاء هر $m > 0$

$$|x_{n+m} - x_n| \to 0 \qquad n \to \infty \text{ به ازاء } \qquad (۴-۵۳)$$

در آن صورت دنبالهٔ x_n همگرا می شود .

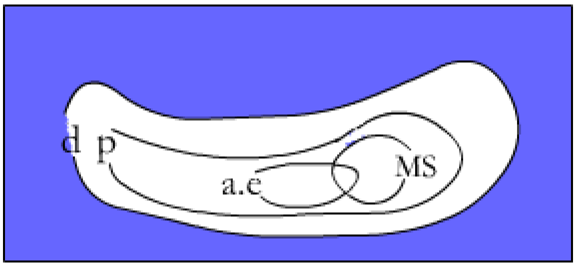
قضیهٔ فوق هم چنین برای دنباله تصادفی قابل اعمال است. در این مورد، مقدارحدّی باید بر طبق مورد مربوطه تعبیر و تعیین گردد.

برای مثال اگر به ازاء هر $m > 0$

$$E\{|X_{n+m} - X_n|^2\} \to 0 \quad , \qquad n \to \infty \text{ به ازاء }$$

باشد در آن صورت دنباله تصادفی X_n از نظر MS همگرا می شود.

در شکل (۳ -۴) رابطهٔ مفاهیم مختلف همگرایی نشان داده شده است . هر نقطه داخل مستطیل نمایندهٔ یک دنبالهٔ تصادفی می‌باشد . حرف مندرج روی هر منحنی بیان می‌کند که تمام دنباله‌های داخل منحنی از نظر مفهوم آن حرف ، همگرا می‌شوند . ناحیه سایه دار شامل تمام دنباله هایی است که از نقطه نظر هر مفهومی همگرا نمی‌شوند . حرف d در منحنی بیرونی نشان می‌دهد که اگر دنباله ای همگرا شود در آن صورت از نظر توزیع نیز همگرا خواهد شد . نکتهٔ مهم دیگر آن است که اگر دنباله ای از نظر MS همگرا شود از نظر احتمال نیز همگرا خواهد شد .



شکل (۳ -۴)

در حقیقت بنا به نامساوی چپی چف می‌توان گفت :

$$P\{|X_n-X|>\varepsilon\} \leq \frac{E\{|X_n-X|^2\}}{\varepsilon^2}$$

اگر از نظر MS $X_n \to X$ میل کند ، در آن صورت به ازاء $\varepsilon > 0$ ثابت ، طرف راست نامساوی به مقدار صفر میل می‌کند و در نتیجه به ازاء $n \to \infty$ طرف چپ نامساوی به صفر میل نموده و رابطهٔ (۵۱-۴) اقناع می‌شود .

به هر حال عکس این مطلب صحت ندارد . اگر X_n محدود کراندار نباشد ، در آن صورت $P\{|X_n-X|>\varepsilon\}$ ممکن است به صفر میل کرده ولی

$$E\{|X_n-X|^2\}$$

به صفر میل نکند .

در هر صورت ، اگر X_n در خارج بازه ای مانند $(-c,c)$ به ازاء

$$n > n_0$$

میرا و صفر گردد ، در آن صورت همگرایی P و همگرایی MS معادل هستند .

بدیهی است که همگرایی به مفهوم تقریباً در همه جا بر اساس رابطهٔ (۴۸-۴) همگرایی به مفهوم احتمال (P) نیز می‌باشد. به طور شهودی می‌توان نشان داد که عکس این امر صحّت ندارد.

مثال ۶-۴

فرض کنید که ξ به طور تصادفی از بازهٔ $S=[0,1]$ انتخاب می‌شود . احتمال این که ξ در زیر بازهٔ S قرار داشته باشد با طول آن زیر بازه برابر است . به ازاء .... ،۱،۲ = n پنج دنبالهٔ تصادفی زیر را تعریف می کنیم .

$$U_n(\xi)=\frac{\xi}{n} \quad , \quad V_n(\xi)=\xi\left(1-\frac{1}{n}\right) \qquad W_n(\xi)=\xi e^n$$

$$Y_n(\xi)=\mathrm{Cos}\,2\pi n\xi \qquad Z_n(\xi)=e^{-n(n\xi-1)}$$

کدام یک از این دنباله ها همه جا همگرا می‌شوند و کدام یک تقریباً در همه جا همگرا می گردند؟ متغیر تصادفی حدّی را که دنباله ها به آن میل می کنند، تعیین کنید.

دنباله $U_n(\xi)$ به ازاء جمیع مقادیر ξ به صفر میل می‌کند پس از نظر همه جا همگرا می‌شود .

$$\text{به ازاء تمام مقادیر} \quad \xi\in S \quad , \quad n\to\infty \quad , \quad U_n(\xi)\to U(\xi)=0$$

توجه کنید که در این مورد تمام دنباله های نمونه به همان مقدار حدّی یعنی صفر همگرا می‌شوند.

دنباله $V_n(\xi)$ به ازاء تمام مقادیر ξ به ξ همگرا می شود و در نتیجه می‌توان گفت که از نظر همه جا همگرا می‌شود.

$$\text{به ازاء تمام مقادیر} \quad \xi\in S \quad , \quad n\to\infty \quad , \quad V_n(\xi)\to V(\xi)=\xi$$

در این حالت تمام دنباله های نمونه به مقادیر متفاوت همگرا شده و متغیر تصادفی حدّی $V(\xi)$ در واقع متغیر تصادفی با توزیع یکنواخت در بازهٔ $[0,1]$ است .

دنباله $W_n(\xi)$ به ازاء $\xi = 0$ به صفر همگرا می‌شود ولی به ازاء سایر مقادیر ξ به بی نهایت واگرا می‌شود . بنابراین این دنبالهٔ تصادفی همگرا نمی‌شود .

دنباله $Y_n(\xi)$ به ازاء $\xi = 0$ و $\xi = 1$ به مقدار ۱ همگرا می‌شود ولی به ازاء سایر مقادیر ξ بین ۱– و ۱ نوسان می‌کند . پس این دنبالهٔ متغیرهای تصادفی همگرا نمی‌شود.

دنباله $Z_n(\xi)$ مورد جالب توجهی است . به ازاء $\xi = 0$ می‌توان نوشت

$$Z(0)=e^n \to \infty \qquad , \qquad n \to \infty$$

از طرف دیگر ، به ازاء $\xi > 0$ و مقادیر $n > \frac{1}{\xi}$ ، دنباله $Z_n(\xi)$ به صورت نمایی کاهش یافته و به مقدار صفر میل می‌کند. پس

به ازاء تمام مقادیر $\xi > 0 \quad , \quad Z_n(\xi) \to 0$

ولی $P[\xi > 0]=1$ پس $Z_n(\xi)$ به مقدار صفر به مفهوم تقریباً در همه جا همگرا می‌شود. به هر حال $Z_n(\xi)$ به مقدار صفر به مفهوم همه جا همگرا نمی‌شود.

مثال ۷–۴

در مثال ۶–۴ آیا دنباله های تصادفی $V_n(\xi)$ و $Z_n(\xi)$ از نظر متوسط مربع (MS) همگرا می‌شوند؟ در مثال ۶–۴ ملاحظه کردیم که دنباله $V_n(\xi)$ به متغیر تصادفی ξ همه جا همگرا می‌شود .

بنابراین $$E\left[(V_n(\xi) - \xi)^2\right] = E\left[\left(\frac{\xi}{n}\right)^2\right] = \int_0^1 \left(\frac{\xi}{n}\right)^2 d\xi = \frac{1}{3n^2}$$

که در رابطه فوق توزیع یکنواخت در بازهٔ $[0,1]$ برای متغیرهای تصادفی ξ منظور شده است. حال با میل کردن n به بی‌نهایت ، متوسط مربع خطا به صفر همگرا شده و می‌توان گفت که دنبالهٔ $V_n(\xi)$ از نظر متوسط مربع همگرا می‌شود.

هم چنین در مثال ۶–۴ ، مشاهده شد که $Z_n(\xi)$ تقریباً در همه جا به صفر همگرا می شود.

بنابراین

$$E\left[(Z_n(\xi)-0)^2\right]=E\left[e^{2n(n\xi-1)}\right]=e^{2n}\int_0^1 e^{-2n^2\xi}\,d\xi=\frac{e^{2n}}{2n^2}\,(1-e^{-2n^2})$$

با میل کردن n به بی نهایت ، طرف راست رابطهٔ فوق نامحدود و بی‌نهایت شده و در نتیجه می‌توان گفت که دنبالهٔ تصادفی $Z_n(\xi)$ از نظر مفهوم متوسط مربع همگرا نمی‌شود اگر چه این دنباله از نظر مفهوم تقریباً در همه جا همگرا می‌شود .

می‌توان نشان داد که در یک آزمایش مفروض ، اگر احتمال وقوع پیشآمد A برابر p بوده و تعداد رخداد A در n بار تکرار آزمایش برابر k باشد، در آن صورت

$$P\left\{\left|\frac{k}{n}-p\right|<\varepsilon\right\}\to 1 \quad , \quad n\to\infty \qquad (4-54)$$

بدین منظور ، متغیرهای تصادفی زیر را در نظر بگیرید

$$X_i=\begin{cases}1\ , & \text{اگر } A \text{ در آزمایش } i \text{ ام رخ دهد}\\ 0\ , & \text{در غیر این صورت}\end{cases}$$

حال متوسط نمونه این متغیرهای تصادفی عبارت است از

$$\overline{X}_n=\frac{X_1+\dots+X_n}{n}$$

و ضمناً می‌دانیم که

$$E\{X_i\}=E\{\overline{X}_n\}=p \quad , \quad \sigma^2_{x_i}=pq \ , \ \sigma^2_{\overline{x}_n}=\frac{pq}{n}$$

علاوه بر آن

$$pq=p(1-p)\le\frac{1}{4}$$

بنابراین با استناد به نامساوی چپی چف می‌توان گفت :

$$P\{|\overline{X}_n - p| < \varepsilon\} \geq 1 - \frac{pq}{n\varepsilon^2} \geq 1 - \frac{1}{4n\varepsilon^2} \xrightarrow[n\to\infty]{} 1$$

نتیجه فوق در واقع رابطه (۴-۵۴) را اثبات می‌کند چون اگر A به تعداد k رخ دهد

، در آن صورت $\overline{X}_n(\xi) = \frac{k}{n}$ خواهد بود .

بنابراین با توجه به بحث فوق نشان دادیم که $\overline{X}_n$ یا متوسط نمونه از نظر

احتمال به p همگرا می‌شود . می‌دانیم که

$$\eta = E\{\overline{X}_n\} = p$$

پس

$$P\{|\overline{X}_n - \eta| > \varepsilon\} \to 0 \qquad , \qquad n \to \infty$$

و به عبارت دیگر متوسط نمونه به متوسط آماری

$$\overline{X}_n \xrightarrow{p} \eta$$

از نظر احتمال میل کرده و همگرا می‌شود . این مفهوم که در شرایط $n \to \infty$

اعتبار و صحّت دارد به قانون ضعیف اعداد بزرگ (برنولی) شهرت دارد . می‌توان

اثبات کرد که متوسط نمونه $\overline{X}_n$ به p یا متوسط آماری متغیرهای تصادفی X_i

از نظر مفهوم تقریباً در همه جا نیز همگرا می شود.

$$\overline{X}_n \xrightarrow{ac} \eta$$

این مفهوم را که نیز در شرایط $n \to \infty$ اعتبار دارد قانون قوی اعداد بزرگ

(بورل) می‌نامند . در این مرحله از بررسی ، بهتر است نگاهی به قضیهٔ مارکوف

انداخته و با آن آشنا شویم . دنباله X_i متشکل از متغیرهای تصادفی را در نظر گرفته و متوسط نمونه آن

$$\overline{X}_n = \frac{X_1 + \dots + X_n}{n}$$

می‌باشد . بدیهی است $\overline{X}_n$ خود نیز یک متغیر تصادفی است که مقادیر $\overline{X}_n(\xi)$ به نتیجه آزمایشی ξ بستگی دارد . بنا به این قضیه اگرمتغیرهای تصادفی X_i به نحوی باشند که به ازاء $n \to \infty$ متوسط $\overline{\eta}_n$ متغیرهای $\overline{X}_n$ به حد η میل کرده

و نیز به ازاء $n \to \infty$ واریانس متغیرهای $\overline{X}_n$ یعنی $\overline{\sigma}_n^2$ به حدّ صفر میل کند ، یعنی

$$E\{\overline{X}_n\} = \overline{\eta}_n \xrightarrow[n \to \infty]{} \eta \quad , \quad \overline{\sigma}_n^2 = E\{(\overline{X}_n - \overline{\eta}_n)^2\} \xrightarrow[n \to \infty]{} 0$$

در آن صورت ، متغیر تصادفی $\overline{X}_n$ از نظر مفهوم MS به η همگرا می‌شود . یعنی

$$\overline{X}_n \xrightarrow{MS} \eta \qquad \text{یا} \qquad E\{(\overline{X}_n - \eta)^2\} \xrightarrow[n \to \infty]{} 0$$

۴-۴ قضیهٔ حدّ مرکزی

n متغیر تصادفی مستقل X_i را در نظر گرفته و مجموع آنها را تشکیل می‌دهیم .

$$X = X_1 + \dots + X_n$$

مجموع فوق یک متغیر تصادفی با متوسط $\eta = \eta_1 + \dots + \eta_n$ و واریانس

$$\sigma^2 = \sigma_1^2 + \dots + \sigma_n^2$$

می باشد.

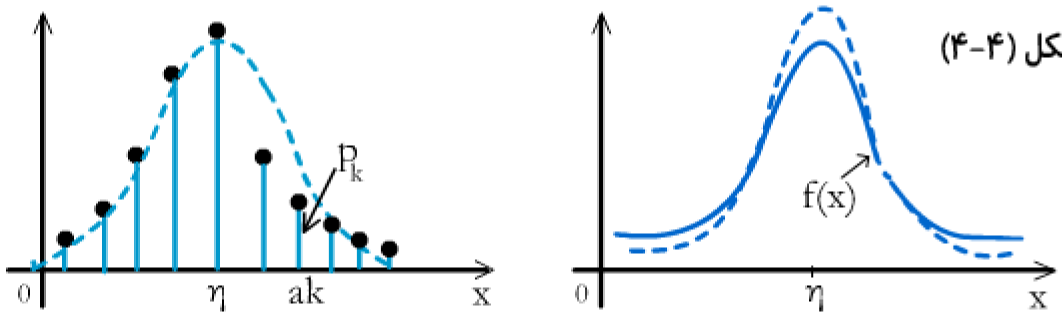
بنا به قضیهٔ حدّ مرکزی (Central Limit Theorem) تحت شرایط عمومی معینی ، توزیع $F(x)$ متغیر مجموع X با افزایش n به توزیع نرمال با همان متوسط و واریانس میل می‌کند .

$$F(x) \simeq G\left(\frac{x-\eta}{\sigma}\right) \qquad (۴-۵۵)$$

علاوه بر این ، اگر متغیرهای تصادفی X_i از نوع پیوسته باشند در آن صورت تابع چگالي احتمال $f(x)$ به چگالی نرمال همگرا می‌شود . (شکل ۴-۴ الف)

$$f(x) \simeq \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\eta)^2/2\sigma^2} \qquad (۴-۵۶)$$

شکل (۴-۴)



(ب) (الف)

این قضیهٔ مهم را می‌توان به صورت حدّی نیز بیان کرد . اگر $Z=(X-\eta)/\sigma$ فرض شود ، در این‌صورت برای متغیر تصادفی کلی یا پیوسته به ترتیب می‌توان نوشت :

$$F_z(3) \xrightarrow[n\to\infty]{} G(3) , \quad f_z(3) \xrightarrow[n\to\infty]{} \frac{۱}{\sqrt{۲\pi}} e^{-3^{۲}/۲}$$

قبل از آن که به اثبات این قضیه بپردازیم باید گفت که CLT را می‌توان به عنوان خاصیتی از کانوولوشن تعبیر نمود . کانوولوشن تعداد زیادی از توابع مثبت تقریباً یک تابع نرمال خواهد شد .

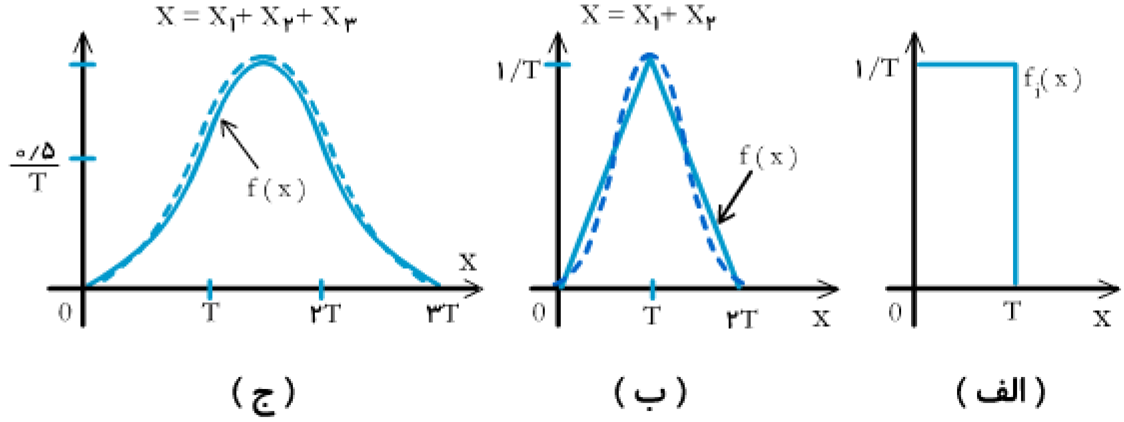
ماهیت تقریب CLT و مقدار مورد لزوم n به ازاء یک کران خطای مشخص به شکل توابع چگالی احتمال $f_i(x)$ بستگی دارد . اگر متغیرهای تصادفی X_i متغیرهای مستقل با توزیع یکسان (i.i.d) باشند، مقدار $n = ۳۰$ در اکثر کاربردها کافی است . در حقیقت ، اگر توابع چگالی $f_i(x)$ هموار باشند مقدار n تا حد ۵ را می‌توان به کار برد . مثال زیر این نکته را به خوبی نشان می‌دهد.

مثال ۸–۴

متغیرهای تصادفی X_i متغیرهای (i.i.d) با توزیع یکنواخت در بازهٔ (۰,۱) هستند . تابع چگالی $f_x(x)$ مجموع آنها X را با تقریب نرمال به ازاء $n = ۲$ و $n = ۳$ می‌خواهیم مقایسه کنیم . در این مسئله

$$\eta_i = \frac{T}{۲} , \quad \sigma_i^{۲} = \frac{T^{۲}}{۱۲} , \quad \eta = n\frac{T}{۲} , \quad \sigma^{۲} = n\frac{T^{۲}}{۱۲}$$

شکل (۵-۴)



به ازاء $n = 2$ ، تابع چگالی $f(x)$ یک مثلث است که از کانوولوشن یک پالس چهار گوش با خود به وجود می آید (شکل ۵-۴) .

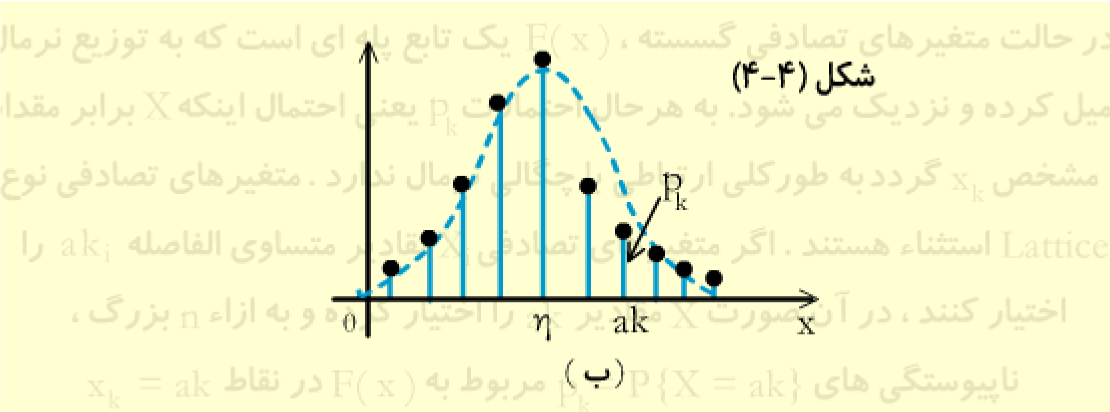
$$\eta = T \quad , \quad \sigma^2 = \frac{T^2}{6} \quad , \quad f(x) \simeq \frac{1}{T}\sqrt{\frac{3}{\pi}}\; e^{-3(x-T)^2/T^2}$$

به ازاء $n = 3$ ، $f(x)$ از سه قطعه سهمی شکل تشکیل شده که از کانوولوشن یک مثلث با پالس چهارگوش ایجاد گردیده است .

$$\eta = \frac{3T}{2} \quad , \quad \sigma^2 = \frac{T^2}{4} \quad , \quad f(x) \approx \frac{1}{T}\sqrt{\frac{2}{\pi}}\; e^{-2(x-1/5T)^2/T^2}$$

همان گونه که در شکل (۵-۴) مشاهده می‌شود حتی برای چنین مقادیر کوچک n خطای تقریب کوچک و ناچیز می باشد .

در حالت متغیرهای تصادفی گسسته ، $F(x)$ یک تابع پله ای است که به توزیع نرمال میل کرده و نزدیک می شود. به هرحال احتمالات p_k یعنی احتمال اینکه X برابر مقدار مشخص x_k گردد به طورکلی ارتباطی با چگالی نرمال ندارد . متغیرهای تصادفی نوع Lattice استثناء هستند . اگر متغیرهای تصادفی X_i مقادیر متساوی الفاصله ak_i را اختیار کنند ، در آن صورت X مقادیر ak را اختیار کرده و به ازاء n بزرگ ، ناپیوستگی های $p_k = P\{X = ak\}$ مربوط به $F(x)$ در نقاط $x_k = ak$



برابر نمونه های چگالی نرمال (شکل ۴-۴ ب) خواهد بود .

$$P\{X = ak\} \simeq \frac{۱}{\sigma\sqrt{۲\pi}} e^{-(ak-\eta)^{۲}/۲\sigma^{۲}} \qquad (۴\text{-}۵۷)$$

نکته قابل توجه دیگر ، مورد متغیرهای تصادفی X_i می‌باشد که مستقل با توزیع یکسان بوده و مقادیر ۱ ، ۰ را با احتمالات به ترتیب p , q اختیار می کنند . تحت این شرایط مجموع آنها یا X متغیر تصادفی از نوع Lattice بوده و مقادیر $k = 0, \ldots, n$ را اختیار می کند .در این حالت :

$$E\{X\} = nE\{X_i\} = np \ , \ \sigma_x^2 = n\sigma_1^2 = npq$$

با قرار دادن مقادیر فوق در رابطه (۴-۵۷) تقریب زیر حاصل می شود

$$P\{X = ak\} = \binom{n}{k} p^k q^{n-k} \simeq \frac{1}{\sqrt{2\pi npq}} e^{-(k-np)^2/2npq}$$

این رابطه نشان می‌دهد که قضیه دوموار-لاپلاس (تقریب نرمال) مورد خاصی از نوع Lattice قضیه حد مرکزی می‌باشد . به منظور اثبات قضیه حد مرکزی ، صحت رابطه (۴-۵۶) را می‌توان با استفاده از توابع مشخصه نشان داد .

برای سهولت محاسبات ، فرض می‌کنیم که $\eta_i = 0$ بوده و توابع مشخصه متغیرهای تصادفی X_i , $X = X_1 + \ldots + X_n$ را به ترتیب با $\Phi_i(\omega)$ و $\Phi(\omega)$ نشان می‌دهیم .

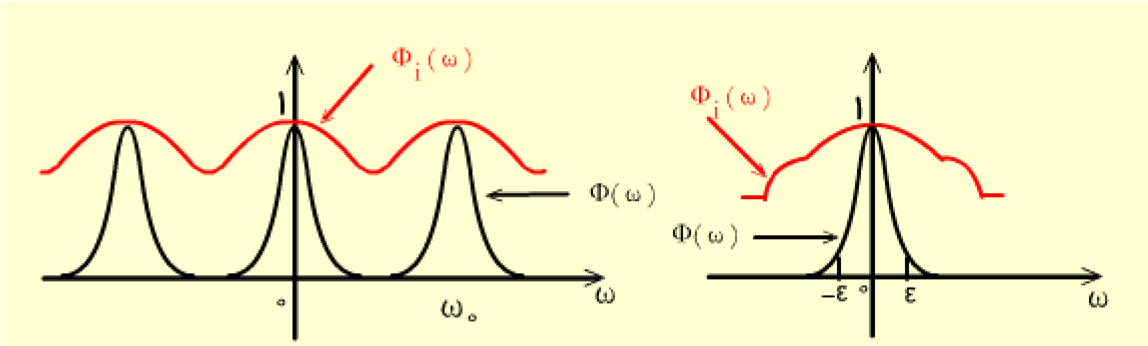
با توجه به استقلال X_i ها می‌توان گفت :

$$\Phi(\omega) = \Phi_1(\omega) \ldots \Phi_n(\omega)$$

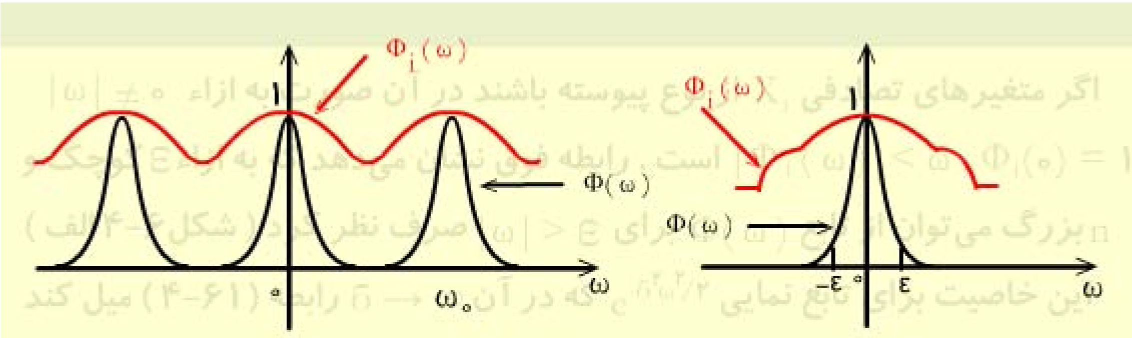
در مجاورت مبداء ، توابع $\psi_i(\omega) = \mathrm{Ln}\,\Phi_i(\omega)$ را می‌توان با یک سهمی تقریب زد.

$$\psi_i(\omega) \simeq -\frac{1}{2}\sigma_i^2\omega^2 \ , \ \Phi_i(\omega) = e^{-\sigma_i^2\omega^2/2} \quad |\omega| < \varepsilon \ \text{به ازاء} \quad (۴\text{-}۵۸)$$

اگر متغیرهای تصادفی X_i از نوع پیوسته باشند در آن صورت به ازاء $|\omega| \neq ۰$ ، $\Phi_i(۰) = ۱$, $|\Phi_i(\omega)| < \omega$ است . رابطه فوق نشان می‌دهد که به ازاء ε کوچک و n بزرگ می‌توان از تابع $\Phi(\omega)$ برای $|\omega| > \varepsilon$ صرف نظر کرد (شکل ۶-۴ الف) این خاصیت برای تابع نمایی $e^{-\sigma^2\omega^2/2}$ که در آن $\sigma \to \infty$ رابطه (۶۱-۴) میل کند نیز صادق است . با توجه به بحث فوق می‌توان نتیجه گرفت که



(ب) شکل ۶-۴ (الف)



(ب) شکل ۶-۴ (الف)

(۴-۵۹) به ازاء تمام مقادیرها $\Phi(\omega) \simeq e^{-\sigma_1^2\omega^2/2} \cdots e^{-\sigma_n^2\omega^2/2} = e^{-\sigma^2\omega^2/2}$

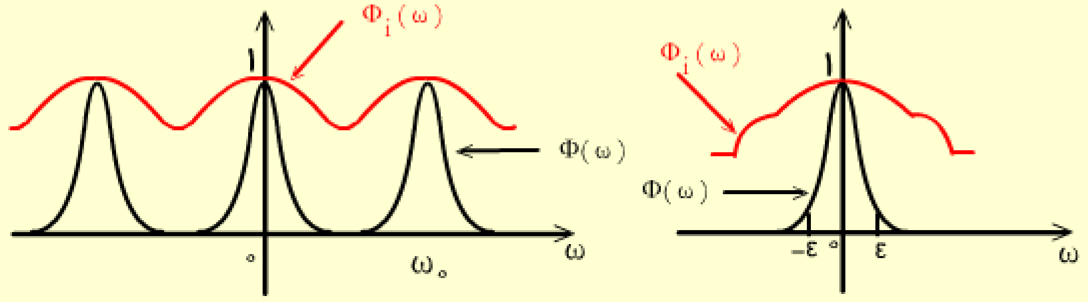
و این نتیجه کاملاً هماهنگ با رابطه (۵۶-۴) است . شکل دقیق قضیه بیان می‌کند که متغیر تصادفی نرمالیزه شده

$$Z = \frac{X_1 + \dots + X_n}{\sigma} \quad , \quad \sigma^2 = \sigma_1^2 + \dots + \sigma_n^2$$

به ازاء $\infty \rightarrow n$ به متغیر تصادفی $N(۰ , ۱)$ میل می‌کند .

$$f_z(3) \xrightarrow[n\to\infty]{} \frac{۱}{\sqrt{۲\pi}} e^{-3^۲/۲} \qquad (۴-۶۰)$$

برای اثبات فرض می‌کنیم که متغیرهای تصادفی X_i از نوع $i.i.d$ هستند و در این حالت : $\Phi_1(\omega) = \ldots = \Phi_n(\omega)$ ، $\sigma = \sigma_i \sqrt{n}$



(ب) شکل ۶-۴ (الف)

بنابراین

$$\Phi_z(\omega) = \Phi_i^n\left(\frac{\omega}{\sigma_i\sqrt{n}}\right)$$

حال با بسط توابع $\psi_i(\omega) = Ln\,\Phi_i(\omega)$ حول مبدأ می‌توان گفت :

$$\psi_i(\omega) = -\frac{\sigma_i^۲\omega^۲}{۲} + O(\omega^۳)$$

بنابراین

$$\psi_z(\omega) = n\psi_i\left(\frac{\omega}{\sigma_i\sqrt{n}}\right) = -\frac{\omega^۲}{۲} + O\left(\frac{۱}{\sqrt{n}}\right) \xrightarrow[n\to\infty]{} \frac{\omega^۲}{۲}$$

رابطه بالا نشان می‌دهد که به ازاء $n\to\infty$ تابع $\Phi_z(\omega)$ به $e^{-\omega^۲/۲}$ میل کرده و در نتیجه رابطه (۴-۶۰) حاصل می‌شود .

همان گونه که ملاحظه شد ، قضیه همیشه صادق نبوده و واقعیت ندارد .

مجموعه ای از شرایط کافی برای صحت قضیه حد مرکزی

عبارت است از :

الف) (۴-۶۱)
$$\sigma_1^2 + \ldots + \sigma_n^2 \xrightarrow[n \to \infty]{} \infty$$

ب) عددی مانند ۲ $< \alpha$ و ثابت محدودی مانند k را می‌توان یافت که

(۴-۶۲)
$$\int_{-\infty}^{+\infty} x^{\alpha} f_i(x)\, dx < k < \infty$$

البته این شرایط کلی‌ترین شرایط نیستند . به هرحال آنها گستره وسیعی از کاربردها را در برمی‌گیرند . برای مثال اگر بتوان ثابت $\varepsilon > 0$ را به نحوی تعیین کرد که $\sigma_i > \varepsilon$ به ازاء تمام مقادیر i باشد در آن صورت رابطه (۴-۶۱) اقناع خواهد شد .

هم چنین اگر تمام توابع چگالی احتمال $f_i(x)$ در خارج بازه محدود $(-C, C)$ صفر باشند ، (مستقل از طول کم یا زیاد بازه) در آن صورت رابطه (۴-۶۲) ارضاء خواهد شد . قضیه حد مرکزی را می‌توان برای حاصل ضرب متغیرهای تصادفی نیز مطرح کرد . فرض کنید n متغیر تصادفی مستقل مثبت X_i در دست بوده و حاصل ضرب آنها را در نظر می‌گیریم .

$$Y = X_1 X_2 \ldots X_n, \quad X_i > 0$$

می‌توان به سهولت نشان داد که برای n بزرگ، تابع چگالی احتمال Y تقریباً تابع لوگ نرمال (Lognormal) یا نرمال لگاریتمی خواهد بود

$$f_y(y) = \frac{۱}{y\sigma\sqrt{۲\pi}} \exp\{-\frac{۱}{۲\sigma^۲}(\text{Ln } y - \eta)^۲\}\, U(y)$$

$$\eta = \sum_{i=۱}^{n} E\{\text{Ln } X_i\} \quad , \quad \sigma^۲ = \sum_{i=۱}^{n} \text{var}(\text{Ln } X_i) \qquad (۴-۶۳)$$

برای اثبات می توان متغیر تصادفی $Z = \text{Ln } Y = \text{Ln } X_۱ + \ldots\ldots + \text{Ln } X_n$ را تعریف کرده و با استناد به قضیه حد مرکزی (برای مجموع متغیرهای تصادفی) نتیجه گرفت که متغیر تصادفی Z به ازاء n بزرگ یک متغیر تصادفی تقریباً نرمال با متوسط η و واریانس $\sigma^۲$ است . از طرف دیگر می‌دانیم که

$$Y = e^Z \quad , \quad g'(3) = e^3$$

اگر $Y > ۰$ باشد معادله $Y = e^3$ فقط یک جواب $3 = \text{Ln } y$ دارد . پس

$$f_y(y) = \frac{۱}{y} f_Z(\text{Ln } y) \quad , \quad y > ۰$$

اگر $Y < ۰$ باشد ، در آن صورت $f_y(y) = ۰$ خواهد بود . حال با توجه به اینکه Z متغیر تصادفی $N(\eta , \sigma^۲)$ است می‌توان نتیجه گرفت که

$$f_y(y) \approx \frac{۱}{y\sigma\sqrt{۲\pi}} e^{-(\text{Ln } y - \eta)^۲ / ۲\sigma^۲}$$

یعنی Y متغیر تصادفی با توزیع لوگ نرمال است بدیهی است که نتیجه فوق هنگامی اعتبار دارد که متغیرهای تصادفی $\text{Ln } X_i$ شرایط لازم برای اعتبار قضیه حد مرکزی را اقناع کنند .

مثال ۹-۴

فرض کنید که متغیرهای تصادفی X_i در بازه (۱ , ۰) یکنواخت هستند .

در این حالت :

$$E\{\ln X_i\} = \int_0^1 \ln x\, dx = -1$$

$$E\{(\ln X_i)^2\} = \int_0^1 (\ln x)^2 dx = 2$$

بنابراین چون $\eta = -n$, $\sigma^2 = n$ می‌باشد ، با استناد به رابطه (۶۳-۴) تابع چگالی احتمال $Y = X_1 \ldots X_n$ برابر است با :

$$f_y(y) = \frac{1}{y\sqrt{2\pi n}} \exp\left\{-\frac{1}{2n}(\ln y + n)^2\right\} U(y)$$

مسایل فصل ۴

۱-۴ نشان دهید که اگر $F(x, y, z)$ یک توزیع مشترک باشد، آن گاه برای هر $x_1 \le x_2$ ، $y_1 \le y_2$ و $z_1 \le z_2$ داریم:

$$F(x_2, y_2, z_2)+F(x_1, y_1, z_1)+F(x_1, y_2, z_1)+F(x_2, y_1, z_1)$$
$$-F(x_1, y_2, z_2)-F(x_2, y_1, z_2)-F(x_2, y_2, z_1)-F(x_1, y_1, z_1) \ge 0$$

۲-۴ نشان دهید که اگر متغیرهای تصادفی Z,Y,X به صورت مشترک نرمال و جفت، جفت مستقل باشند، آن گاه مستقل هستند.

۳-۴ متغیرهای تصادفی X_i دارای چگالی یکسان و مستقل و در فاصلهٔ (۰/۵ , ۰/۵-) یکنواخت هستند. نشان دهید که داریم:

$$E\{(X_1+X_2+X_3)^4\} = \frac{13}{80}$$

۴-۴ نشان دهید که اگر متغیرهای تصادفی Z,Y,X مستقل و چگالی مشترک آنها دارای تقارن کروی :

$$f(x, y, z)=f\sqrt{(x^2+y^2+z^2)}$$

باشند، آن گاه آن ها نرمال هستند.

۴-۵ نشان دهید که اگر متغیرهای تصادفی Z,Y,X طوری باشند که $r_{xy}=r_{yz}=1$ برقرار باشد، آن گاه $r_{xz}=1$ است.

۴-۶ نشان دهید که

$$E\{X_1X_2|X_3\}=E\{E\{X_1X_2|X_2, X_3\}|X_3\}$$
$$=E\{X_2E\{X_1X_2|X_2, X_3\}|X_3\}$$

۴-۷ نشان دهید که

$$\hat{E}\{Y|X_1\}=\hat{E}\{\hat{E}\{Y|X_1, X_2\}|X_1\}$$

است، در جایی که

$$\hat{E}\{Y|X_1, X_2\}=a_1X_1+a_2X_2$$

یک تخمین MS خطی از Y بر حسب X_1 و X_2 است.

۴-۸ نشان دهیدکه اگر داشته باشیم:

$$X_i\geq 0 \quad , \quad E\{X_i^2\}=M \quad و \quad S=\sum_{i=1}^{n}X_i$$

آن گاه:

$$E\{s^2\}\leq ME\{N^2\}$$

۴-۹ تعداد تصادفات در یک روز یک متغیر تصادفی N با پارامتر a می باشد.

احتمال این که یک تصادف کشنده رخ دهد برابر p است.

نشان دهید که تعداد M تصادف خطرناک در یک روز یک متغیر تصادفی

پواسون با پارامتر ap است.

راهنمایی:

$$E\{e^{j\omega M}|N=n\}=\sum_{k=0}^{n} e^{j\omega k}\binom{n}{k} p^{k} q^{n-k}=(pe^{j\omega}+q)^{n}$$

۴-۱۰ متغیرهای تصادفی X_k مستقل با چگالی های $f_k(x)$ و متغیر تصادفی

N مستقل از X_k با $P=\{N=k\}=P_k$ می باشد. نشان دهید که اگر

$$S=\sum_{k=1}^{n} X_k$$

باشد آنگاه داریم:

$$f_s(S)=\sum_{k=1}^{\infty} P_k\,[f_1(S)*\dots*f_k(S)]$$

۴-۱۱ متغیرهای تصادفی X_i دارای چگالی یکسان و مستقل با تابع گشتاور

$\Phi_x(S)=E\{e^{sX_i}\}$ می باشند. متغیر تصادفی N مقادیر ... ,۱ ,۰ را می گیرند.

و تابع گشتاور آن برابر $\Gamma_n(z)=E\{Z^n\}$ است.

نشان دهید که اگر

$$Y=\sum_{i=1}^{n} X_i$$

باشد آنگاه داریم:

$$\Phi_x(S)=E\{e^{sY}\}=\Gamma_n[\Phi_x(S)]$$

راهنمایی:

$$E\{e^{sY}|N=k\}=E\{e^{s(X_1+...+X_k)}\}=\Phi_x^k(S)$$

حالت خاص اگر N پواسون با پارامتر a باشد،

آن گاه

$$\Phi_y(S)=e^{a\Phi_x(S)-a}$$

می باشد.

۴-۱۲ متغیرهای تصادفی X_i دارای چگالی یکسان و مستقل و در فاصلهٔ (۰,۱) یکنواخت می‌باشند. نشان دهید که اگر $Y=\max X_i$ باشد، آن گاه برای $0 \le y \le 1$ ، $F(y)=y^n$ است.

۴-۱۳ تعداد N متغیر تصادفی Z_i با توزیع $N(\eta_i, 1)$ داده شده اند. متغیر تصادفی $W=Z_1^2+...+Z_n^2$ را تشکیل می دهیم.

این متغیر تصادفی را مربع - کای غیر متمرکز با n درجه آزادی با خروج از مرکز $e=\eta_1^2+...+\eta_n^2$ گویند.

نشان دهید که تابع مولد گشتاور آن

$$\Phi_w(s)=\frac{1}{\sqrt{(1-2s)^n}}\exp\left\{\frac{es}{1-2s}\right\}$$

است.

۴-۱۴ نشان دهید که اگر $\alpha_0+\alpha_1X_1+\alpha_2X_2$ تخمین MS خطی غیر همگن از S برحسب X_1 و X_2 باشد، آن گاه

$$\hat{E}\{S-\eta_s | X_1-\eta_1, X_2-\eta_2\}=\alpha(X_1-\eta_1)+\alpha_2(X_2-\eta_2)$$

۴-۱۵ نشان دهید که

$$\hat{E}\{Y|X_1\}=\hat{E}\{\hat{E}\{Y|X_1, X_2\}|X_1\}$$

۴-۱۶ به طور تصادفی n نقطه را در فاصلهٔ $(0, 1)$ قرار می دهیم.
با X و Y به ترتیب فاصلهٔ اولین و آخرین نقطه را مبدأ نمایش می دهیم.
$F(x)$، $F(y)$ و $F(x,y)$ را بیابید.

۴-۱۷ متغیرهای تصادفی X_i به صورت $N(0; \sigma)$ و مستقل می باشند.
نشان دهید که اگر

$$Z=\frac{\sqrt{\pi}}{2n}\sum_{i=1}^{n}|X_{2i}-X_{2i-1}|$$

باشد آن گاه داریم:

$$E\{Z\}=\sigma \qquad \sigma_z^2=\frac{\pi-2}{2n}\sigma^2$$

۴-۱۸ نشان دهید که اگر R ماتریس همبستگی بردار تصادفی
$\mathbb{X}:[X_1,...,X_n]$ و R^{-1} معکوس آن باشد، آن گاه داریم:

$$E\{\mathbb{X}R^{-1}\mathbb{X}^t\}=n$$

۴-۱۹ نشان دهید که اگرمتغیرهای تصادفی X_i از نوع پیوسته و مستقل باشند،
آن گاه برای n به قدر کافی بزرگ چگالی $\mathrm{Sin}(X_1+...+X_n)$ تقریباً مساوی
با چگالی $\mathrm{Sin}\ X$ است، که X متغیر تصادفی یکنواخت در فاصلهٔ $(-\pi, \pi)$
می باشد.

۴-۲۰ نشان دهید که اگر

$$E\{|X_n - a_n|^2\} \longrightarrow 0 \quad \text{و} \quad a_n \longrightarrow a$$

باشد، آن گاه همچنان که $n \longrightarrow \infty$ می رود، از دید MS، $X_n \longrightarrow a$ می رود.

۴-۲۱ یک مجموع نامحدود به سبب تعریف حد زیر است:

$$\sum_{k=1}^{n} X_k = \lim_{n\to\infty} Y_n \qquad Y_n = \sum_{k=1}^{n} X_k$$

نشان دهید که اگر متغیرهای تصادفی X_k مستقل خطی با میانگین صفر و واریانس σ_k^2 باشند، آن گاه مجموع از دید MS وجود دارد، اگر و فقط اگر رابطهٔ زیر را داشته باشیم:

$$\sum_{k=1}^{n} \sigma_k^2 < \infty$$

راهنمایی:

$$E\{(Y_{n+m} - Y_n)^2\} = \sum_{k=n+1}^{n+m} \sigma_k^2$$

۴-۲۲ متغیرهای تصادفی X_i دارای چگالی یکسان و مستقل با چگالی $ce^{-cx}U(x)$ می‌باشند. نشان دهید که اگر $X = X_1, ..., X_n$ باشد، آن گاه $f_x(x)$ یک چگالی ارلنگ است.

۴-۲۳ مقاومتهای r_1, r_2, r_3, r_4 متغیرهای تصادفی مستقل و هر یک در فاصلهٔ (۴۵۰,۵۵۰) یکنواخت می باشند. با استفاده از قضیهٔ حد مرکزی $P\{1900 \le r_1 + r_2 + r_3 + r_4 \le 2100\}$ را بیابید.

۴-۲۴ نشان دهید که اگرمتغیرهای تصادفی X_i دارای چگالی کوشی باشند، آن گاه قضیه حد مرکزی برقرار نمی باشد.