فصل هفتم

در فصل قبل برخی توزیع‌های معروف و کاربردی را برای متغیر تصادفی گسسته ارایه نمودیم و خواص و ویژگیهای هر یک را بررسی نمودیم. در این فصل به بررسی توزیعهای مبتنی بر متغیر تصادفی پیوسته می‌پردازیم.

۷.۱ توزیع یکنواخت پیوسته

ساده‌ترین توزیع پیوسته، توزیع یکنواخت می‌باشد. فرض کنید تمام نقاط در بازه (a,b) دارای امکان وقوع یکسان باشند، در این صورت متغیر تصادفی X را که برد آن مقادیر موجود در بازه (a,b) می‌باشد، متغیر تصادفی یکنواخت می‌نامند. که با نماد $X \sim U(a,b)$ نشان داده می‌شود.

تابع چگالی احتمال متغیر تصادفی یکنواخت برابر است با:

$$f_X(x) = \begin{cases} \dfrac{۱}{b-a} & a < x < b \\ ۰ & \text{سایر مقادیر} \end{cases}$$

و به این ترتیب بدست می‌آید که چون X یکنواخت می‌باشد در نظر می‌گیریم:

$$f_X(x) = c \qquad a < x < b$$

در نتیجه:

$$\int_{-\infty}^{+\infty} f_X(x)\,dx = ۱ \quad \Rightarrow \quad \int_a^b c\,dx = c\,x\Big|_a^b = c\,(b-a) = ۱ \quad \Rightarrow \quad c = \frac{۱}{b-a}$$

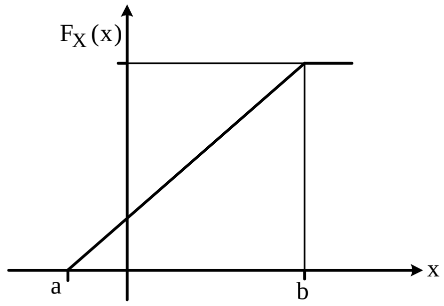
تابع توزیع متغیر تصادفی یکنواخت X برابر است با:

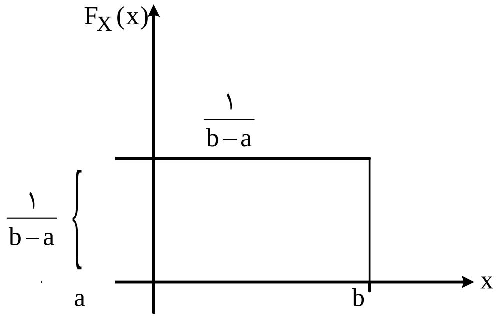
$$f_X(x) = \int_{-\infty}^{t} f_X(x) = \int_a^t \frac{۱}{b-a}\,dx = \frac{t-a}{b-a}$$

پس:

$$f_X(x) = \begin{cases} ۰ & t<a \\ \dfrac{t-a}{b-a} & a \le t \le b \\ ۱ & t > b \end{cases}$$

۷-۲ در شکل زیر نمودار تابع چگالی و تابع توزیع متغیر تصادفی یکنواخت نشان داده شده است:





مقادیر امید ریاضی واریانس و تابع مولد گشتاور را برای متغیر تصادفی یکنواخت X محاسبه می‌کنیم:

$$E[X] = \int_a^b x\,\frac{۱}{b-a}\,dx = \frac{۱}{b-a}\,\frac{x^۲}{۲}\Big|_a^b =$$

$$= \frac{۱}{b-a}\left(\frac{b^۲}{۲} - \frac{a^۲}{۲}\right) = \frac{(b-a)\,(a+b)}{۲(b-a)} = \frac{a+b}{۲}$$

$$E[X^{۲}] = \int_a^b x^{۲} \frac{۱}{b-a} dx = \frac{۱}{b-a} \left(\frac{x^{۳}}{۳}\right)\Big|_a^b$$
$$= \frac{(b-a)(b^{۲}-ab+a^{۲})}{۳(b-a)} = \frac{a^{۲}+ab+b^{۲}}{۳}$$

پس:

$$var(X) = E[X^{۲}] - E^{۲}[X] = \frac{a^{۲}+ab+b^{۲}}{۳} - \frac{a^{۲}+۲ab+b^{۲}}{۴} = \frac{(b-a)^{۲}}{۱۲}$$

$$m_X(t) = E[e^{tX}] = \int_a^b e^{tX} \frac{۱}{b-a} dx = \frac{۱}{b-a} \left(\frac{e^{tX}}{t}\right)\Big|_a^b = \frac{e^{tb}-e^{ta}}{t(b-a)}$$

با فرض $a \le c < d \le b$ مقدار احتمال $c \le x \le d$ برابر است با:

$$p(c \le x \le d) = \int_c^d \frac{۱}{b-a} dx = \frac{d-c}{b-a}$$

ملاحظه می‌کنید که مطابق با تعریف متغیر تصادفی یکنواخت مقدار احتمال تنها به طول بازه (c,d) وابسته است و نه به مقادیر c و d.

۳-۷ مثال ۱: نمرات دانشجویان یک کلاس در درس آمار به طور یکنواخت در فاصله ۱۲ الی ۲۰ توزیع شده است مطلوبست:

الف) تابع توزیع احتمال:

ب) احتمال قرارگیری نمرات در بازه $(۲۰-۱۸)$, $(۱۵-۱۳)$.

ج) میانگین نمرات کلاس، واریانس.

حل: الف) متغیر تصادفی X دارای توزیع یکنواخت است بنابراین:

$$f_X(x) = \begin{cases} \frac{۱}{۲۰-۱۲} = \frac{۱}{۸} & ۱۲ \le X \le ۲۰ \\ ۰ & \text{سایر مقادیر} \end{cases}$$

ب):

$$P(۱۸ \le X \le ۲۰) = \int_{۱۸}^{۲۰} \frac{۱}{۸} dx \; \frac{۲۰-۱۸}{۸} = \frac{۲}{۸} = \frac{۱}{۴}$$
$$P(۱۳ \le X \le ۱۵) = \int_{۱۳}^{۱۵} \frac{۱}{۸} dx \; \frac{۱۵-۱۳}{۸} = \frac{۲}{۸} = \frac{۱}{۴}$$

ج):

$$E[X] = \frac{۲۰+۱۲}{۲} = ۱۶$$
$$var(X) = \frac{(۲۰-۱۲)^{۲}}{۱۲} = \frac{۶۴}{۱۲} = \frac{۱۶}{۳}$$

۴-۷ ۶ . ۲ متغیر تصادفی نمایی

یک فرایند پواسون با پارامتر λ که برای یک واحد زمان نظاره می‌شود، را در نظر بگیرید اگر زمان شروع فرایند صفر $(t=۰)$ باشد و T مدت زمانی باشد که باید بگذرد تا اولین پیشامد رخ دهد در این صورت T یک متغیر تصادفی نمایی با پارامتر λ نامیده می‌شود.

تابع توزیع نمایی را با استفاده از تعریف بدست می‌آوریم:

می‌دانیم $f_X(t) = p(X \le t)$ حال اگر $t < ۰$ باشد بنا به تعریف زمان منفی معنا ندارد. بنابراین اگر $t < ۰$ داریم: $p(X \le t) = ۰$ یعنی:

$$f_X(t) = ۰ \qquad t<۰$$

حال فرض می‌کنیم $t \geq 0$ باشد، احتمال اینکه هیچ پیشامدی در بازه $(0\,,\,t)$ رخ ندهد بنابه تابع چگالی احتمال پواسون برابر است با:

$$f_X(x=0) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}\Big|_{x=0} = e^{-\lambda t}$$

بنابراین $p(X > t) = e^{-\lambda t}$ اما $p(X \leq t) = 1 - p(X > t)$ و داریم:

$$F_X(t) = p(X \leq t) = 1 - p(X > t) = 1 - e^{-\lambda t}$$

در نتیجه تابه توزیع متغیر تصادفی نمایی بصورت زیر بدست می‌آید:

$$F_X(t) = \begin{cases} 1 - e^{-\lambda t} & t > 0 \\ 0 & t < 0 \end{cases}$$

با مشتق‌گیری از $F_X(t)$ تابع چگالی را بدست می‌آوریم:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

۵-۷ برای بدست آوردن مقادیر امید و واریانس از تابع مولد گشتاور استفاده می‌کنیم:

$$m_X(t) = E\left[e^{tx}\right] = \int_0^\infty e^{tx}\ \lambda e^{-\lambda x}\ dx = \lambda \int_0^\infty e^{tx}\ \lambda e^{-x(\lambda - t)}\ dx$$

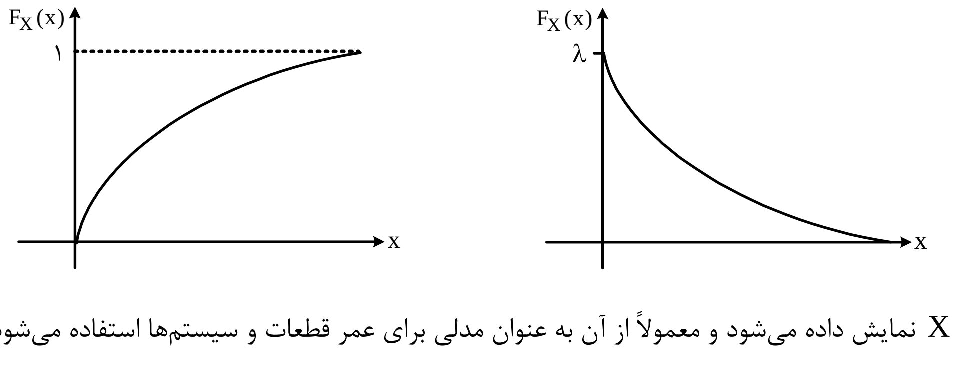
$$= \lambda\ \left(\frac{1}{t-\lambda}\ e^{-x(\lambda-t)}\right)\Big|_0^\infty = \frac{-\lambda}{t-\lambda} = \frac{\lambda}{\lambda - t} \quad (t < \lambda)$$

$$E[X] = m'_X(t=0) = \frac{\lambda}{(\lambda - t)^2}\Big|_{t=0} = \frac{1}{\lambda}$$

$$E[X^2] = m''_X(t=0) = \frac{2\lambda}{(\lambda - t)^3}\Big|_{t=0} = \frac{2}{\lambda^2}$$

$$var(X) = E[X^2] - E^2[X] = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

شکل زیر نمودار تابع توزیع و چگالی را برای متغیر تصادفی نمایش می‌دهد:



متغیر تصادفی نمایی با نماد $X \sim E(\lambda)$ نمایش داده می‌شود و معمولاً از آن به عنوان مدلی برای عمر قطعات و سیستم‌ها استفاده می‌شود.

۶-۷ مثال ۲: اگر لامپهای تولید شده توسط یک کارخانه به طور متوسط هر ۶ ماه یکبار بسوزند مطلوبست:

الف) تابع چگالی احتمال برای مدت زمان کارکرد لامپها.

ب) احتمال اینکه از زمان خرید، یک لامپ حداقل ۲ ماه کار کند؟

ج) احتمال اینکه یک لامپ قبل از ۴ ماه بسوزد؟

د) اگر ۱۰ عدد لامپ خریداری کنیم احتمال اینکه حداقل ۲ عدد از لامپها قبل از ۴ ماه بسوزند؟

ه) میانگین طول عمر، واریانس و تابع مولد گشتاور را برای متغیر تصادفی X که مدت زمان کارکرد لامپها می‌باشد بدست بیاورید.

۷-۷ حل: الف) متغیر تصادفی X را طول عمر لامپهای خریداری شده در نظر می‌گیریم در این صورت X یک متغیر تصادفی نمایی با پارامتر λ می‌باشد. برای بدست آوردن پارامتر λ ابتدا هر واحد زمانی را یک ماه در نظر می‌گیریم، λ برابر است با تعداد پیشامدها در یک واحد زمانی که در اینجا چون هر ۶ ماه یک لامپ می‌سوزد معادل است با اینکه $\frac{۱}{۶}$ لامپ در هر یک ماه می‌سوزد یعنی $\lambda = \frac{۱}{۶}$ و تابع چگالی برابر است با:

$$f_X(x) = \frac{۱}{۶}\ e^{-\frac{۱}{۶}x} \qquad x > ۰$$

$$= ۰ \qquad \text{سایر مقادیر}$$

ب) احتمال اینکه حداقل ۲ ماه طول بکشد تا یک لامپ بسوزد برابر است با:

$$p(X > ۲) = \int_۲^\infty \frac{۱}{۶}\ e^{-\frac{x}{۶}}\,dx = \frac{۱}{۶}(-۶\,e^{-\frac{x}{۶}})\Big|_۲^\infty$$

$$= \frac{۱}{۶}(۰ - (-۶\,e^{-\frac{۲}{۶}})) = e^{-\frac{۱}{۳}} = ۰/۷۱۶$$

ج) احتمال اینکه یک لامپ قبل از ۴ ماه بسوزد برابر است با:

$$p(X \le ۴) = \int_۰^۴ \frac{۱}{۶}\ e^{-\frac{x}{۶}}\,dx = \frac{۱}{۶}(-۶\,e^{-\frac{x}{۶}})\Big|_۰^۴ = ۱-۶\,e^{-\frac{۴}{۶}} = ۰/۴۸۶$$

۷-۸ د) متغیر Y را بصورت زیر تعریف می‌کنیم:

$Y =$ تعداد لامپها از بین ۱۰ لامپ که در کمتر از ۴ ماه می‌سوزند.

طبق تعریف Y یک متغیر تصادفی دو جمله‌ای با پارامتر $n = ۱۰$ و $p = ۰/۴۸۶$ می‌باشد. پارامتر p از بند (ج) بدست می‌آید. احتمال اینکه هر لامپ در کمتر از ۴ ماه بسوزد برابر ۰/۴۸۶ می‌باشد، بنابراین می‌بایستی پارامتر p برابر ۰/۴۸۶ انتخاب شود. حال با احتمال اینکه حداقل ۲ عدد از لامپها قبل از ۴ ماه بسوزند برابر است با:

$$p(Y \ge ۲) = ۱-p(Y < ۲) = ۱-[p(Y=۱)+p(Y=۰)]$$

$$=۱-\left[\binom{۱۰}{۱}(۰/۴۸۶)^۱(۰/۵۱۴)^۹+\binom{۱۰}{۰}(۰/۴۸۶)^۰(۰/۵۱۴)^{۱۰}\right]$$

$$=۱-(۰/۰۱۲+۰/۰۰۱۲) = ۰/۹۸۶$$

پس:

$$f_X(x) = \binom{۱۰}{y}(۰/۴۸۶)^y(۰/۵۱۴)^{۱۰-y}$$

اگر امید ریاضی $f_Y(y)$ را محاسبه کنیم داریم: $E[X] = np = ۱۰\times۰/۴۸۶ = ۴/۸۶$ یعنی تقریباً از هر ۱۰ عدد لامپ خریداری شده بطور متوسط ۵ لامپ در کمتر از ۴ ماه می‌سوزند.

ه) میانگین طول عمر لامپ‌ها عبارتست از: $E[X] = \frac{۱}{\lambda} = \frac{۱}{\frac{۱}{۶}} = ۶$ که با توجه به اینکه واحد زمان را هر یک ماه در نظر گرفتیم بنابراین میانگین طول عمر هر لامپ ۶ ماه می‌باشد که از صورت مثال نیز همین انتظار می‌رفت.

$$var(X) = \frac{۱}{\lambda^۲} = \frac{۱}{(\frac{۱}{۶})^۲} = ۳۶$$

$$m_X(t) = \frac{\lambda}{\lambda - t} = \frac{\frac{۱}{۶}}{\frac{۱}{۶} - t} = \frac{۱}{۱ - ۶t}$$

۷-۹ ۶ . ۲ . ۱ رابطه توزیع هندسی و توزیع نمایی

توزیع هندسی طبق تعریف عبارتست از تعداد آزمایشها قبل از رسیدن به اولین پیروزی توزیع نمایی نیز از جهت تعریف تشابه زیادی با توزیع هندسی دارد. توزیع نمایی هم عبارتست از مدت زمانی که در یک فرایند پواسون با پارامتر λ باید سپری شود تا اولین پیشامد رخ دهد توجه کنید که میانگین توزیع هندسی $\frac{۱}{p}$ و توزیع نمایی $\frac{۱}{\lambda}$ می‌باشد. متغیرهای تصادفی هندسی و نمایی در یک خاصیت ویژه مشترک می‌باشند که هیچ متغیر تصادفی گسسته یا پیوسته دیگری این حالت را ندارد. ویژگی فوق به بی‌حافظگی معروف است و به صورت زیر می‌باشد:

برای توزیع نمایی: $p(X > a + b \mid X > a) = p(X > b)$

برای توزیع هندسی: $p(X = a + b \mid X \geq a) = p(X = b)$

اثبات برای توزیع نمایی: فرض می‌نیم A برابر با پیشامد $X > a$ و B برابر پیشامد $X > b$ و C برابر پیشامد $X > a+b$ باشد در این صورت باید ثابت کنیم: $p(C|A) = p(B)$ می‌دانیم:

$$p(C|A) = \frac{p(C \cap A)}{p(A)}$$

از آنجا که پیشامد C زیر مجموعه‌ای از پیشامد A است $(\{X > a+b\} \subset \{X > a\})$ بنابراین $C \cap A = C$ و داریم:

$$p(C|A) = \frac{p(C \cap A)}{p(A)} = \frac{p(C)}{p(A)}$$

۷-۱۰ حال مقادیر $p(A)$ و $p(B)$ و $p(C)$ را بدست می‌آوریم:

$$p(A) = p(X > a) = \int_a^{\infty} \lambda e^{-\lambda x}\, dx = e^{-\lambda a}$$

$$p(B) = p(X > b) = \int_b^{\infty} \lambda e^{-\lambda x}\, dx = e^{-\lambda b}$$

$$p(C) = p(X > a+b) = \int_{a+b}^{\infty} \lambda e^{-\lambda x}\, dx = e^{-\lambda (a+b)}$$

$$p(C|A) = \frac{p(C)}{p(A)} = \frac{e^{-\lambda(a+b)}}{e^{-\lambda a}} = e^{-\lambda b} = p(B)$$

بی‌حافظگی به این مفهوم است که اگر قطعه‌ای یا سیستمی به مدت a واحد زمان کار کرده باشد، احتمال اینکه حداقل b واحد زمان دیگر نیز کار کند $(a+b)$ برابر است با احتمال اینکه قطعه یا سیستم از لحظه صفر بخواهد حداقل b واحد زمان کار کند.

۷-۱۱ مثال ۳: در یک کارخانه ماشین‌های تولیدی هر یک ماه نیازمند سرویس تعمیرات باشند. اگر یک ماشین تولیدی ۶ ماه بدون تعمیرات کار کرده باشد احتمال اینکه در طول ماه بعد نیازمند تعمیات باشد چقدر است؟

حل: متغیر تصادفی X را مدت زمان لازم قبل از اولین تعمیر در نظر می‌گیریم در این صورت X یک متغیر تصادفی نمایی با پارامتر $\lambda = ۱$ می‌باشد. می‌بایستی مقدار احتمال زیر را محاسبه کنیم:

$$p(X > ۶ + ۱ \mid X > ۶) = p(X > ۷ \mid X > ۶)$$

$$= p(X > ۱) = \int_۱^{+\infty} e^{-x}\, dx = e^{-۱} = ۰/۳۶$$

۷-۱۲ ۶ . ۳ توزیع گاما

اگر متغیرهای تصادفی $X_۱ , X_۲ , \cdots , X_n$ دارای توزیع نمایی با پارامتر λ باشند در این صورت متغیر تصادفی Y را که برابر است با مجموع متغیرهای تصادفی $X_۱$ تا X_n، یک متغیر تصادفی گاما می‌نامیم.
$$Y = X_۱ + X_۲ + \cdots + X_n$$
تابع چگالی متغیر تصادفی گاما برابر است با:

$$f_X(x) = \frac{(\lambda x)^{n-۱}}{T(n)} \lambda e^{-\lambda x} \qquad x \geq ۰$$
$$= ۰ \qquad \text{سایر مقادیر}$$

که در آن $T(n)$ تابع گاما می‌باشد و به صورت زیر تعریف می‌شود:

$$T(n) = \int_۰^{\infty} x^{n-۱} e^{-x}\, dx$$

$T(n)$ به ازای هر $n > ۰$ موجود است و روابط زیر برای تابع گاما برقرار است:

۱- $T(n+۱) = n\,T(n)$

۲- $T(n) = (n-۱)!$ اگر n عدد طبیعی باشد

۳- $T(\frac{۱}{۲}) = \sqrt{\pi}$

از آنجا که متغیر تصادفی گاما مجموع n متغیر تصادفی نمایی است پس می‌توان مقادیر امید، واریانس و تابع مولد گشتاور را با استفاده از این خاصیت بدست آورد:

$$E[X] = E\left[\sum_{i=۱}^{n} X_i\right] = \sum_{i=۱}^{n} E[X_i] = \sum_{i=۱}^{n} \frac{۱}{\lambda} = \frac{n}{\lambda}$$

$$\operatorname{var}(X) = \operatorname{var}\left(\sum_{i=۱}^{n} X_i\right) = \sum_{i=۱}^{n} \operatorname{var}(X_i) = \sum_{۱}^{n} \frac{۱}{\lambda^۲} = \frac{n}{\lambda^۲}$$

$$m_X(t) = E\left[e^{tX}\right] = E\left[e^{t\sum_{۱}^{n} X_i}\right] = E\left[e^{tX_۱}\right] E\left[e^{tX_۲}\right] \cdots E\left[e^{tX_n}\right]$$
$$= \left(\frac{\lambda}{\lambda - t}\right)\left(\frac{\lambda}{\lambda - t}\right) \cdots \left(\frac{\lambda}{\lambda - t}\right) = \left(\frac{\lambda}{\lambda - t}\right)^n$$

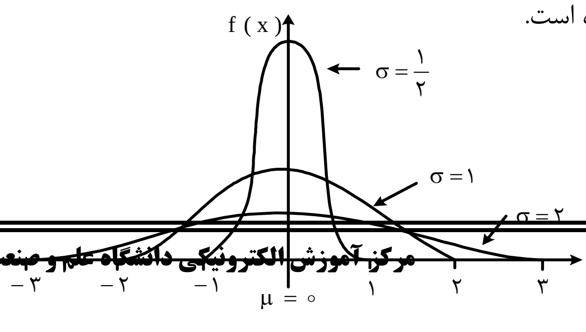
۷-۱۳ توزیع نرمال

مهمترین توزیع پیوسته که آنرا بررسی می‌کنیم توزیع نرمال می‌باشد، بسیاری از پدیده‌های طبیعی مثل قد و وزن افراد، نمرات درسی، میزان محصول در طول سال از توزیع نرمال پیروی می‌کنند.

متغیر تصادفی X دارای توزیع نرمال است و آنرا بصورت $X \sim N(\mu , \sigma^۲)$ نمایش می‌دهیم اگر تابع چگالی احتمال بفرم زیر باشد:

$$f_X(x) = \frac{۱}{\sigma\sqrt{۲\pi}} e^{\frac{-(x-\mu)^۲}{۲\sigma^۲}} \qquad -\infty < x < +\infty , \mu \in R , \sigma > ۰$$

همانطور که ملاحظه می‌نید دو پارامتر توزیع نرمال همان میانگین و واریانس متغیر تصادفی X می‌باشند. مد، میانه و میانگین توزیع نرمال با یکدیگر برابر می‌باشند. در شکل زیر توزیع نرمال با مقادیر واریانس متفاوت نشان داده شده است.



تغیر میانگین در توزیع نرمال تنها نمودار تابع را به سمت راست یا چپ منتقل می‌کند. منحنی نرمال دارای خواص زیر می‌باشد:

۱- نقاط عطف منحنی عبارتند از $x_۱ = \mu + \sigma \ , \ x_۲ = \mu - \sigma$.

۲- منحنی نسبت به خط $x = \mu$ متقارن است بنابراین:

$$۱- \quad f_X(\mu - a) = f_X(\mu + a) \qquad\qquad (a > ۰)$$

$$۲- \begin{cases} p(X > a) = p(X < -a) \\ ۱ - F_X(x) = F_X(-x) \end{cases}$$

۳- مساحت سطح زیر منحنی نمودار نرمال برابر واحد می‌باشد. زیرا بوضوح داریم: $\int_{-\infty}^{+\infty} f(x)\, dx = ۱$

۱۴-۷ برای بدست آوردن میانگین و واریانس از تابع مولد گشتاور استفاده می‌کنیم که بصورت زیر بدست می‌آید:

$$m_X(t) = E\left[e^{tX}\right] = \int_{-\infty}^{+\infty} e^{tx} \frac{۱}{\sigma\sqrt{۲\pi}} e^{\frac{-(x-\mu)^۲}{۲\sigma^۲}}\, dx$$

$$= \frac{۱}{\sigma\sqrt{۲\pi}} \int_{-\infty}^{+\infty} e^{tx - \frac{(x-\mu)^۲}{۲\sigma^۲}}\, dx$$

با تبدیل توان e به صورت مربع کامل داریم:

$$tx - \frac{(x-\mu)^۲}{۲\sigma^۲} = \frac{۱}{۲\sigma^۲}\left[x^۲ - ۲\mu x + \mu^۲ - ۲\sigma^۲ + x\right]$$

$$= \frac{۱}{۲\sigma^۲}\left[(x - \mu x - \sigma^۲ t)^۲ - ۲\sigma^۲ t\mu - \sigma^۴ t^۲\right]$$

$$= \frac{۱}{۲\sigma^۲}\left[(x - \mu - \sigma^۲ t)^۲\right] + t\mu + \frac{\sigma^۲ t^۲}{۲}$$

به این ترتیب:

$$m_X(t) = e^{\mu t + \frac{(\sigma t)^۲}{۲}} \int_{-\infty}^{+\infty} \underbrace{\frac{۱}{\sigma\sqrt{۲\pi}} e^{\frac{-(x-\mu-\sigma^۲ t)^۲}{۲\sigma^۲}}}_{g_X(x)}\, dx$$

مقدار انتگرال $\int_{-\infty}^{+\infty} g(n)\, dx$ برابر یک می‌باشد زیرا $g_X(x)$ خود یک تابع نرمال با پارامترهای $N(\mu + \delta^۲ t \ , \ \delta^۲)$ می‌باشد. بنابراین مقدار تابع مولد گشتاور برابر است با:

$$m_X(t) = e^{\mu t + \frac{\sigma^۲ t^۲}{۲}}$$

حال داریم:

$$E[X] = m'_X(t)\Big|_{t=۰} = \left(\mu + \frac{\sigma^۲}{۲}(۲t)\right) e^{\mu t + \frac{\sigma^۲ t^۲}{۲}}\Big|_{t=۰} = \mu$$

$$E[X^۲] = m''_X(t)\Big|_{t=۰} = \sigma^۲ (e^{\mu t+\frac{\sigma^۲ t^۲}{۲}})\ (\mu + \frac{\sigma^۲}{۲}(۲t))^۲\ e^{\mu t+\frac{\sigma^۲ t^۲}{۲}}\ \Big|_{t=۰} = \sigma^۲ + \mu^۲$$

$$\Rightarrow \operatorname{var}(X) = E[X^۲] - E^۲[X] = \sigma^۲ + \mu^۲ = \sigma^۲$$

۷-۱۵ تابع توزیع متغیر تصادفی نرمال X عبارتست از:

$$F_X(t) = \int_{-\infty}^{t} \frac{۱}{\delta\sqrt{۲\pi}}\ e^{\frac{-(x-\mu)^۲}{۲\delta^۲}}\ dx$$

برای انتگرال فوق یک تابع اثلیه نمی‌توان بدست آورد به همین دلیل مجبور هستیم از روشهای عددی یک جدول برای مقادیر متفاوت t بدست بیاوریم. اما از آنجا که دو پارامتر μ و $\delta^۲$ نیز متغیر می‌باشند می‌بایستی روشی بدست بیاوریم که بتوان مقدار احتمال را بدون وابستگی به μ و $\delta^۲$ بدست آورد. متغیر تصادفی Z را بصورت زیر تعریف می‌کنیم:

$$Z = \frac{X-\mu}{\sigma}$$

بوضوح Z متغیر تصادفی نرمال با میانگین ۰ و واریانس ۱ می‌باشد. $Z \sim N(۰,۱)$ متغیر تصادفی Z را نرمال استاندارد می‌نامند. (که در فصل اول نیز برای نمونه‌های گرفته شده از یک جامعه معرفی شد).

$$E[Z] = E\left[\frac{X-\mu}{\sigma}\right] = \frac{۱}{\sigma}\left[E[X]-\mu\right] = \frac{۱}{\sigma}(\mu-\mu) = ۰$$

$$\operatorname{var}(Z) = \operatorname{var}\left(\frac{X-\mu}{\sigma}\right) = \frac{۱}{\sigma^۲}\operatorname{var}(X-\mu) = \frac{\sigma^۲}{\sigma^۲} = ۱$$

$$m_Z(t) = e^{\frac{t^۲}{۲}}$$

حال با استفاده از تغیر متغیر $Z = \dfrac{X-\mu}{\sigma}$ می‌توان به سادگی مقادیر احتمال را از روی جدول توزیع متغیر تصادفی نرمال استاندارد بدست آورد.

تابع چگالی متغیر تصادفی نرمال استاندارد که آنرا با $\eta_Z(z)$ یا $\varphi(z)$ نمایش می‌دهیم برابر است با:

$$\eta_Z(z) = \varphi(z) = \frac{۱}{\sqrt{۲\pi}}\ e^{-\frac{Z^۲}{۲}} \qquad -\infty < Z < +\infty$$

تابع توزیع متغیر تصادفی نرمال استاندارد برابر است با:

$$\varphi(Z) = N_Z(z) = p_z(Z \le z) = \int_{-\infty}^{z} \frac{۱}{\sqrt{۲\pi}}\ e^{-\frac{Z^۲}{۲}}\ dx$$

جدول ضمیمه مقادیر تابع توزیع را برای $-۳/۵ \le Z \le ۳/۵$ نشان می‌دهد.

۷-۱۶ مثال ۴: با استفاده از جدول مقادیر احتمالات زیر را بدست بیاورید؟

الف) $p(Z < ۰) \quad , \quad p(Z > ۰)$

ب) $p(Z < ۱) \quad , \quad p(|Z| < \frac{۳}{۲})$

ج) $p(-۱ < Z < ۰) \quad , \quad p(۱ < Z < ۳)$

حل: الف)

$$p(Z < ۰) = \varphi(۰) = \frac{۱}{۲}$$

$$p(Z > ۰) = ۱ - \varphi(۰) = ۱ - \frac{۱}{۲} = \frac{۱}{۲}$$

ب)

$$p(Z < ۱) = \varphi(۱) = ۰/۸۴۱۳$$

$$p(|Z| < \frac{۳}{۲}) = p(-\frac{۳}{۲} < Z < \frac{۳}{۲}) = \varphi(\frac{۳}{۲}) - \varphi(\frac{۳}{۲})$$

$$= ۰/۹۳۳۲ - ۰/۰۶۶۸ = ۰/۸۶۶۴$$

ج)

$$p(-۱ < Z < ۰) = \varphi(۰) - \varphi(۱) = \frac{۱}{۲} - ۰/۱۵۸۷ = ۰/۳۴۱۳$$

$$p(۱ < Z < ۳) = \varphi(۳) - \varphi(۱) = ۰/۹۹۸۷ - ۰/۸۴۱۳ = ۰/۱۵۷۴$$

با توجه به مثال فوق می‌وان خواص زیر را برای متغیر تصادفی نرمال استاندارد به دست آورد:

۱. $\varphi(-\infty) = ۰$
۲. $\varphi(+\infty) = ۱$
۳. $\varphi(a) = ۱ - \varphi(-a)$

۱۷-۷ ۶ . ۴ . ۱ محاسبه مقادیر احتمال متغیر تصادفی نرمال

فرض کنید X یک متغیر تصادفی نرمال با میانگین μ و واریانس $\sigma^۲$ باشد $(X \sim N(\mu, \sigma^۲))$ برای بدست آوردن احتمال $p(a < x < b)$ داریم:

$$p(a < x < b) = p(\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma})$$

$$= P(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma})$$

$$= \varphi(\frac{b-\mu}{\sigma}) - \varphi(\frac{a-\mu}{\sigma})$$

۱۸-۷ مثال ۵: اگر $X \sim N(۵, ۱۶)$ مطلوبست محاسبه $p(۲ \le X < ۷)$.

حل: داریم $\mu = ۵$, $\sigma^۲ = ۱۶$ پس:

$$p(۲ \le X < ۷) = p(\frac{۲-۵}{۴} \le \frac{X-۵}{۴} < \frac{۷-۵}{۴})$$

$$= p(-\frac{۳}{۴} \le Z < \frac{۱}{۲})$$

$$= \varphi(\frac{۱}{۲}) - \varphi(-\frac{۳}{۴}) = ۰/۶۹۱۵ = ۰/۲۲۶۶ = ۰/۴۶۴۹$$

۱۹-۷ مثال ۶: نمرات دانشجویان یک کلاس از توزیع نرمال با میانگین ۱۵ و واریانس ۴ پیروی می‌کند اگر بدانیم تمامی نمرات از ۱۰ بیشتر هستند احتمال اینکه نمرات بین ۱۲ تا ۱۶ باشد چقدر است؟

حل: متغیر تصادفی X نرمال می‌باشد $(X \sim N(۱۵, ۴))$ می‌بایستی احتمال شرطی زیر را محاسبه کنیم:

$$p(۱۲ < X < ۱۶) \mid X > ۱۰) = \frac{p(۱۲ < X < ۱۶\,,\, X > ۱۰)}{p(X > ۱۰)}$$

$$= \frac{p(۱۲ < X < ۱۶)}{p(X > ۱۰)} = \frac{p(\frac{۱۲-۱۵}{۲} < \frac{X-۱۵}{۲} < \frac{۱۶-۱۵}{۲})}{p(\frac{X-۱۵}{۲} > \frac{۱۰-۱۵}{۲})}$$

$$= \frac{p(-\frac{۳}{۲} < Z < \frac{۱}{۲})}{p(Z > -\frac{۵}{۲})} = \frac{\varphi(\frac{۱}{۲}) - \varphi(-\frac{۳}{۲})}{\varphi(\frac{۵}{۲})} = \frac{۰/۶۹۱۵ - ۰/۰۶۶۸}{۰/۹۹۳۸} = ۰/۶۲۸۵$$

۲۰-۷ ۶ . ۴ . ۲ تقریب توزیع دو جمله‌ای به توزیع نرمال

اگر در یک توزیع دو جمله‌ای تعداد آزمایشها بسیار زیاد باشد یعنی $n \to +\infty$ و در عین حال احتمال پیروزی و شکست تقریباً برابر باشند یعنی $p \approx q \approx \frac{۱}{۲}$ در این صورت می‌توان مقادیر احتمال دو جمله‌ای را با استفاده از توزیع نرمال بدست آورد.

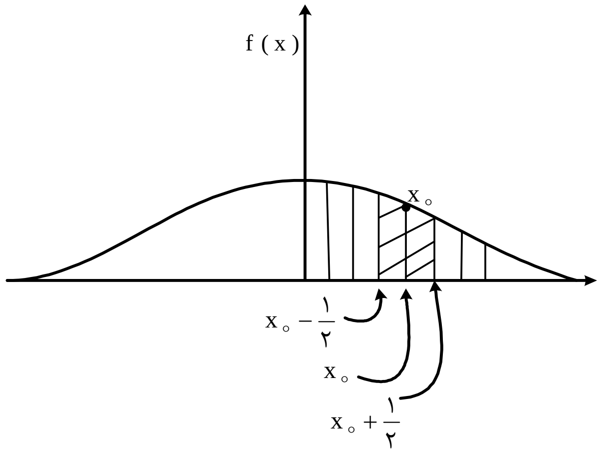
برای تقریب توزیع دو جمله‌ای به نرمال مقادیر امید و واریانس را برابر قرار می‌دهیم یعنی:

$$\mu = np \quad , \quad \sigma^۲ = npq$$

حال به شکل زیر توجه کنید:

برای بدست آوردن $p(X = x_\circ)$

در توزیع دو جمله‌ای کافیست.



در توزیع نرمال مساحت مشخص شده در شکل را بدست بیاوریم زیرا می‌دانیم $p(X = x_\circ) = f(x_\circ)$. و از طرفی مساحت حاشور خورده تقریباً با مساحت یک ذوذنقه برابر است:

$$\text{مساحت ذوذنقه حاشور خورده} = \text{ارتفاع} \times \frac{\text{مجموع دو قاعده}}{۲} = (x_\circ + \frac{۱}{۲} - (x_\circ - \frac{۱}{۲}) \times \frac{f(x_\circ + \frac{۱}{۲}) + f(x_\circ - \frac{۱}{۲})}{۲}$$

$$= \frac{f(x_\circ + \frac{۱}{۲}) + f(x_\circ - \frac{۱}{۲})}{۲} \approx f(x_\circ)$$

حال داریم:

$$p(X \subset x_\circ) = p(x_\circ - \frac{۱}{۲} < X < x_\circ + \frac{۱}{۲})$$

$$= P\, \frac{x_\circ - \frac{۱}{۲} - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{x_\circ + \frac{۱}{۲} - \mu}{\sigma}$$

$$= \varphi(\frac{x_\circ + \frac{۱}{۲} - \mu}{\sigma}) - \varphi(\frac{x_\circ - \frac{۱}{۲} - \mu}{\sigma})$$

و با جاگذاری مقادیر $\mu = np$, $\delta^{۲} = npq$ بدست می‌آوریم:

$$p(X=x_{\circ}) = \varphi\left(\frac{x_{\circ}+\frac{۱}{۲}-np}{\sqrt{npq}}\right) - \varphi\left(\frac{x_{\circ}-\frac{۱}{۲}-np}{\sqrt{npq}}\right)$$

به همین ترتیب:

$$p(X \le x_{\circ}) = \varphi\left(\frac{x_{\circ}+\frac{۱}{۲}-np}{\sqrt{npq}}\right)$$

$$p(a < X \le b) = \varphi\left(\frac{b+\frac{۱}{۲}-np}{\sqrt{npq}}\right) - \varphi\left(\frac{a-\frac{۱}{۲}-np}{\sqrt{npq}}\right)$$

۷-۲۱ مثال ۷: احتمال بارندگی در طول یک ماه در یک شهر ۰/۳ . مطلوبست احتمال اینکه حداقل در ۵۰ شهر از ۲۰۰ شهر کشور در طول یک ماه آینده باران ببارد؟

حل: در این مثال X یک متغیر تصادفی دو جمله‌ای با پارامترهای $X \sim B(۲۰۰\,,\,۰/۳)$ می‌باشد می‌خواهیم احتمال $p(X > ۵۰)$ را محاسبه کنیم:

$$p(X > ۵۰) = ۱-p(X \le ۵۰) = ۱-\varphi\left(\frac{۵۰+\frac{۱}{۲}-(۲۰۰\times ۰/۳}{\sqrt{۲۰۰\times ۰/۳\times ۰/۷}}\right)$$

$$=۱-\varphi\left(\frac{-۹/۵}{۶/۴}\right) = ۱-\varphi(-۱/۴۸) = ۱-۰/۰۶۹۴ = ۰/۹۳۰۶$$

۷-۲۲ ۳ . ۴ . ۶ تقریب توزیع پواسون به توزیع نرمال

برای توزیعهای پواسون با λ بزرگ (بزرگتر یا مساوی ۵) می‌توان مقادیر احتمال توزیع پواسون را با استفاده از توزیع نرمال بدست آورد در این حالت روابط دقیقاً مشابه روابط تقریب دو جمله‌ای به نرمال می‌باشند با این تفاوت که به جای μ , $\delta^{۲}$ داریم: $\mu = \lambda$, $\sigma^{۲} = \lambda$

$$p(X=x_{\circ}) = \varphi\left(\frac{x_{\circ}+\frac{۱}{۲}-\lambda}{\lambda}\right) - \varphi\left(\frac{x_{\circ}-\frac{۱}{۲}-\lambda}{\lambda}\right)$$

$$p(X \le x_{\circ}) = \varphi\left(\frac{x_{\circ}+\frac{۱}{۲}-\lambda}{\lambda}\right)$$

$$p(a < X \le b) = \varphi\left(\frac{b+\frac{۱}{۲}-\lambda}{\lambda}\right) - \varphi\left(\frac{a-\frac{۱}{۲}-\lambda}{\lambda}\right)$$

۷-۲۳ مثال ۸: تلفن یک شرکت در هر ساعت تقریباً ۲۰ مرتبه زنگ می‌زند مطلوبست احتمال اینکه تلفن در طول ۲ ساعت آینده حداقل ۲۰ و حداکثر ۴۰ بار زنگ بزند؟

حل: X یک متغیر تصادفی پواسون می‌باشد واحد زمانی را هر یک ساعت در نظر می‌گیریم در این صورت $\lambda = ۱۰$ می‌باشد و می‌بایستی احتمال $p(۲۰ > X < ۴۰)$ را محاسبه کنیم:

$$p(۲۰ > X < ۴۰) = \varphi\left(\frac{۴۰+\frac{۱}{۲}-۱۰}{۱۰}\right) - \varphi\left(\frac{۲۰-\frac{۱}{۲}-۱۰}{۱۰}\right)$$

$$= \varphi(۳/۰۵) - \varphi(۰/۹۵) = ۰/۹۹۸۹ - ۰/۸۲۸۹ = ۰/۱۷$$