برخی از توزیع‌های پیوسته

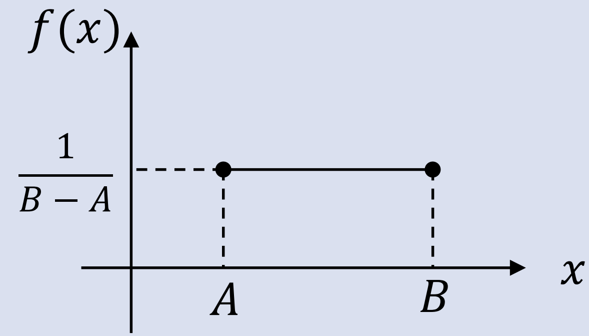
فردوس گرجی

چند توزیع احتمال پیوسته

- **توزیع یکنواخت پیوسته :**

تابع چگالی متغیر تصادفی یکنواخت پیوسته X در فاصله $[A,B]$ عبارت است از:

$$f(x;k)=\begin{cases}\dfrac{1}{B-A} & ; \quad A\le x\le B\\ 0 & ; \quad o.w\end{cases}$$



* توزیع مستطیلی (مستطیلی با قاعده $B-A$ و ارتفاع $\frac{1}{B-A}$)

- **قضیه :** میانگین و واریانس توزیع یکنواخت پیوسته عبارت است از:

$$\mu = \frac{A+B}{2} \qquad \sigma^2 = \frac{(B-A)^2}{12}$$

- **مثال۱:** فرض کنید می‌خواهیم عددی تصادفی در بازه $[0,1]$ به تصادف انتخاب کنیم. تابع چگالی آن را بیابید. میانگین عدد انتخابی چند است؟ احتمال اینکه عدد انتخابی بزرگتر از ۰/۷ باشد چقدر است؟

$$f(x) = \begin{cases} 1 & ; \quad 0 \le x \le 1 \\ 0 & ; \quad o.w \end{cases} \qquad \mu_X = \frac{1+0}{2} = \frac{1}{2} = \mathbf{0.5}$$

$$P(X \ge 0.7) = \int_{0.7}^{1} 1\,dx = x \Big|_{0.7}^{1} = 1 - 0.7 = \mathbf{0.3}$$

مثال۲ : فرض کنید دانش‌آموزی مسیر خانه تا مدرسه را به طور کاملا تصادفی بین ۱۵ تا ۲۰ دقیقه طی می‌کند. توزیع زمان حرکت او را محاسبه کنید. میانگین فاصله زمانی خانه او تا مدرسه چقدر است؟

$$f(x)=\begin{cases}\frac{1}{5} & ; \quad 15\le x\le 20\\ 0 & ; \quad o.w\end{cases}$$

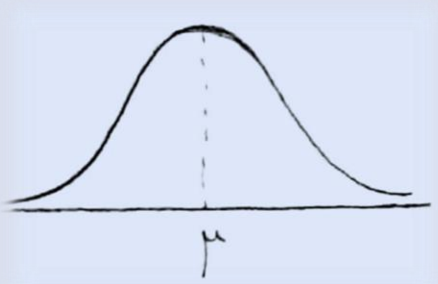
$$\mu=\frac{15+20}{2}=\frac{35}{2}=\mathbf{17.5}$$ دقیقه

❋ احتمال اینکه کمتر از ۱۷ دقیقه تا مدرسه راه طی کند چقدر است؟

$$P(X\le 17)=\int_{15}^{17}\frac{1}{5}\,dx=\frac{1}{5}x\Big|_{15}^{17}=\frac{17}{5}-\frac{15}{5}=\frac{2}{5}=\mathbf{0.4}$$

➢ توزیع نرمال یا گاوسی :

❋ مهمترین توزیع احتمال پیوسته در آمار
❋ بررسی پدیده‌های طبیعی (علت نام‌گذاری) و فیزیکی مانند نوسان فیزیکی حول یک مقدار خاص، مدل‌سازی خطاهای اندازه‌گیری، پردازش سیگنال‌ها و تصاویر، تقریب توزیع‌های دیگر، استفاده در توزیع‌های نمونه‌ای، اندازه قطعات ساخته شده در کارخانه، نمرات یک کلاس، قد و وزن و فشار خون افراد و ...
❋ به احترام کارل گاوس که نخستین بار این توزیع را پیشنهاد داد نامگذاری شده است. (برای محاسبه خطا)
❋ تابع چگالی آن با دو پارامتر میانگین (μ) و انحراف معیار (σ) توزیع مشخص می‌شود.
❋ شکل زنگوله‌ای شکل

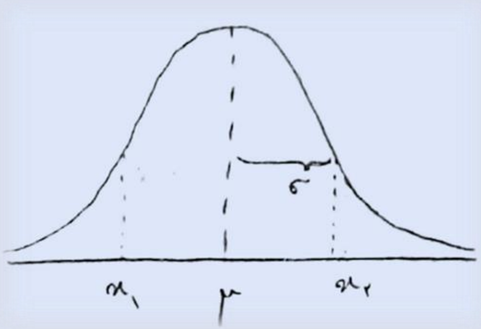


➢ توزیع نرمال یا گاوسی :

❋ تابع چگالی متغیر تصادفی نرمال X ، با میانگین μ و واریانس σ^2 عبارت است از:

$$n(x\,;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \qquad -\infty < x < \infty$$

که $e = 2.7182 \ldots$ و $\pi = 3.1415 \ldots$



❋ $\mu =$ میانگین ، میانه ، مد

❋ متقارن حول μ

❋ نقاط عطف $x_2 = \mu + \sigma$ و $x_2 = \mu - \sigma$

❋ $\lim\limits_{x\to\pm\infty} n(x\,;\mu,\sigma) = 0$

❋ در این فصل با دانستن μ و σ^2 مثال‌ها را حل می‌کنیم. در بسیاری از کاربردها (فصل‌های آینده)، داده‌ها در مسئله رفتاری شبیه توزیع نرمال دارند و ما به دنبال یافتن μ و σ^2 هستیم تا داده‌ها را تحلیل کنیم.

➢ محاسبه میانگین توزیع نرمال:

$$E(X) = \int_{-\infty}^{\infty} x \, . \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx \qquad \begin{cases} u = \frac{x-\mu}{\sigma} \\ u' = \frac{1}{\sigma} dx \end{cases}$$

$$= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (u + \frac{\mu}{\sigma}) e^{-\frac{1}{2}u^2} du$$

$$= \underbrace{\frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u e^{-\frac{u^2}{2}} du}_{\underbrace{-e^{-\frac{u^2}{2}} \Big|_{-\infty}^{\infty}}_{0}} + \mu \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du}_{\underbrace{n(x; 0,1)}_{1}} = \mu$$

➢ محاسبه واریانس توزیع نرمال:

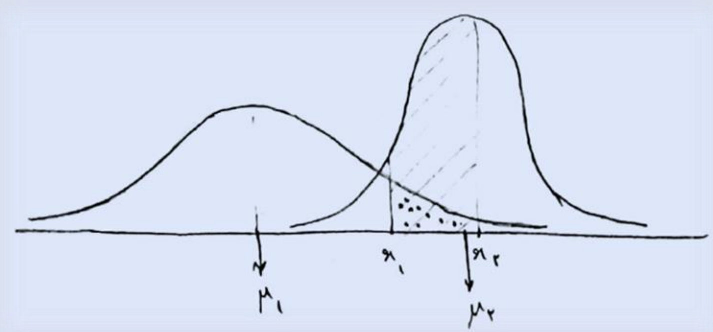
$$E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} dx \qquad \begin{cases} y = \frac{x-\mu}{\sigma} \\ dy = \frac{1}{\sigma} dx \end{cases}$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-\frac{y^2}{2}} dy \qquad \begin{cases} u = y \\ dv = y e^{-\frac{y^2}{2}} dy \rightarrow v = -e^{-\frac{y^2}{2}} \end{cases}$$

$$\frac{\sigma^2}{\sqrt{2\pi}} \left(\underbrace{\left(-y e^{-\frac{y^2}{2}}\right) \Big|_{-\infty}^{\infty}}_{0} - \int_{-\infty}^{\infty} -e^{-\frac{y^2}{2}} dy \right) = \sigma^2 \underbrace{\underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy}_{n(y;0,1)}}_{1} = \sigma^2$$

➢ سطح زیر منحنی توزیع نرمال :

مساحت زیر منحنی بین دو مقدار x_1 و x_1 = $P(x_1 \leq X \leq x_2)$



❋ محاسبه انتگرال چگالی نرمال سخت است.

❋ می‌توان به صورت عددی محاسبه کرد و به صورت جدول‌های تابع توزیع تجمعی ارائه کرد. (مثل جداول پوآسن)

❋ یا باید به ازای همه μ ها و همه σ های ممکن جدول تهیه کرد(بی‌شمار حالت) و یا از استاندارد سازی استفاده کرد.

❋ مقایسه داده‌ها و جوامع و تصمیم‌گیری نیز از دیگر فواید استاندارد سازی است.

❋ استاندارد سازی $\equiv$ تغییر متغیر $z = \frac{X-\mu}{\sigma}$

➢ استاندارد سازی:

$$z = \frac{X-\mu}{\sigma} \qquad \Rightarrow \qquad \boxed{x_1 \le X \le x_2 \quad \Leftrightarrow \quad \frac{x_1-\mu}{\sigma} \le Z \le \frac{x_2-\mu}{\sigma}}$$

$$\Downarrow$$

$$dZ = \frac{1}{\sigma}dx$$

$$P(x_1 \le X \le x_2) = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \int_{z_1}^{z_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz = P(z_1 \le Z \le z_2)$$

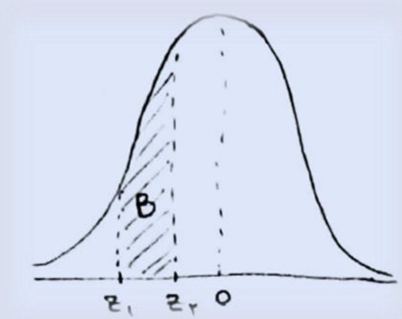
$$\boxed{n(z\,;0\,,1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}}$$ توزیع نرمال استاندارد

➢ تعریف:

توزیع متغیر تصادفی نرمال با میانگین صفر و واریانس یک را توزیع نرمال استاندارد گوییم.

A مساحت ناحیه $=$ B مساحت ناحیه

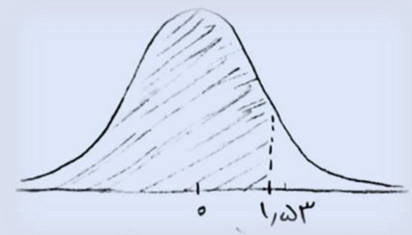
$$P(x_1 \le X \le x_2) = P(z_1 \le Z \le z_2)$$



➢ **مثال۳:** فرض کنید Z دارای توزیع نرمال استاندارد است. مقدار احتمالات زیر را بیابید:

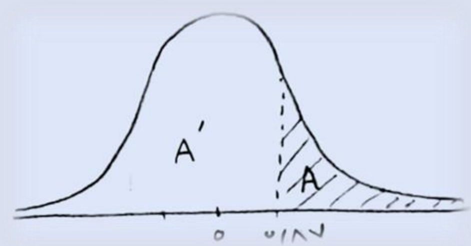
الف) $P(Z \leq 1.53)$ ب) $P(Z > 0.87)$

ج) $P(-1.94 \leq Z \leq 0.12)$



الف) $P(Z \leq 1.53)$

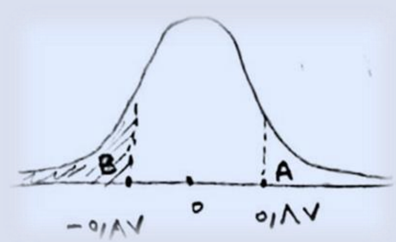
(با استفاده از جدول توزیع تجمعی Z)



ب) $P(Z > 0.87) = A$ مساحت

$= 1 - P(Z \leq 0.87)$

$= 1 - 0.8078 = 0.1922$



راه دوم با استفاده از تقارن توزیع

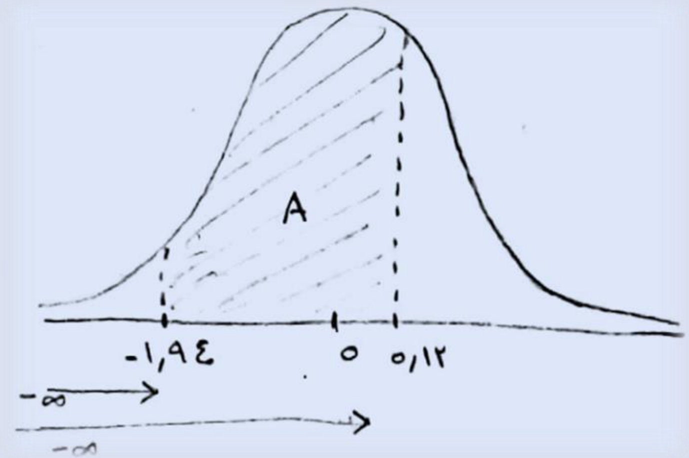
$P(Z > 0.87) = P(Z < -0.87)$

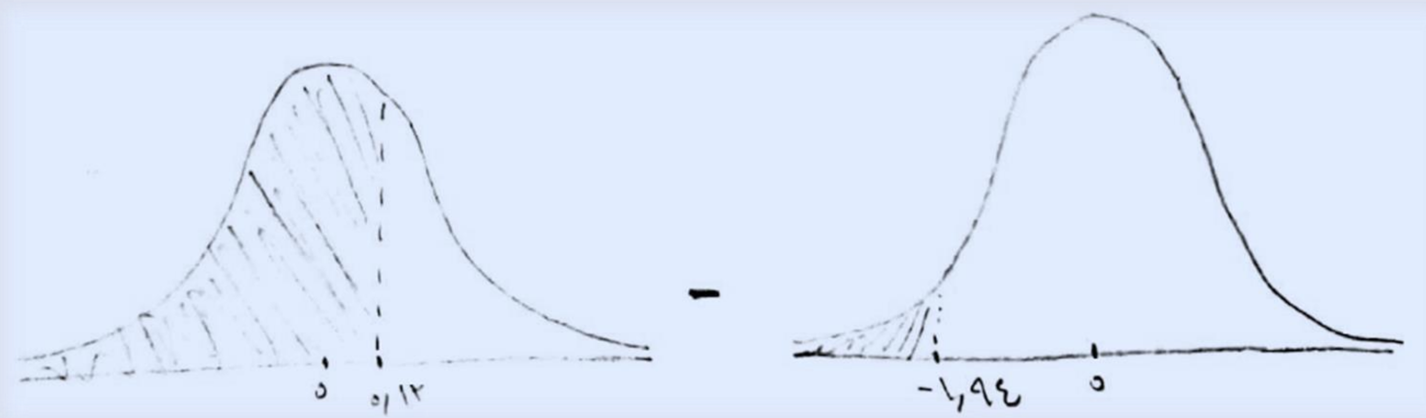
$= B$ مساحت $= 0.1922$

ج) $P(-1.94 \le Z \le 0.12) = A$ مساحت

$= P(Z \le 0.12) - P(Z < -1.94)$

$0.5478 - 0.0262 = 0.5216$





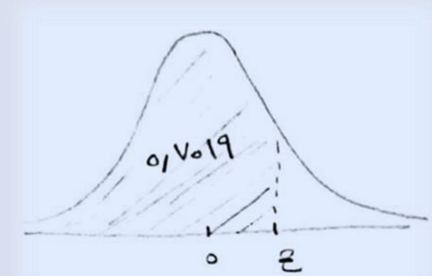
➢ **مثال۴:** فرض کنید Z دارای توزیع نرمال استاندارد است. مقدار z را به نحوی بیابید که:

الف) $P(Z \le z) = 0.7019$ ب) $P(Z \ge z) = 0.7517$

ج) $P(z \le Z \le 1.5) = 0.3$

الف) $P(Z \le z) = 0.7019$

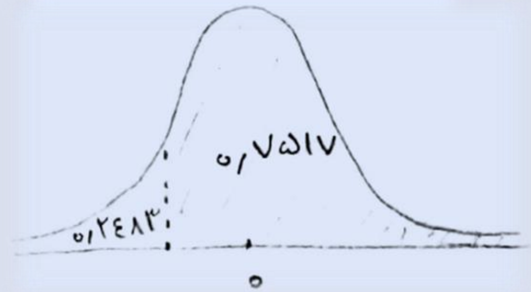
$\xrightarrow{\text{با مشاهده جدول}} \quad z = 0.53$



ب) $P(Z \ge z) = 0.7517$

$P(Z < z) = 1 - 0.7517 = 0.2483$

$\xrightarrow{\text{با مشاهده جدول}} \quad z = -0.68$

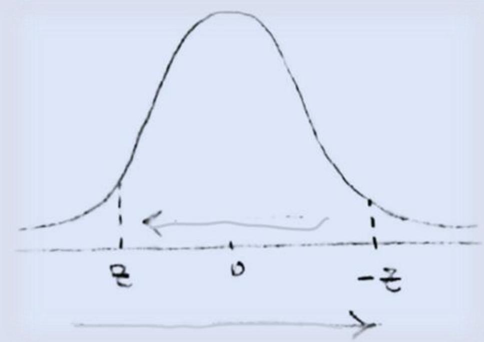


راه حل دوم

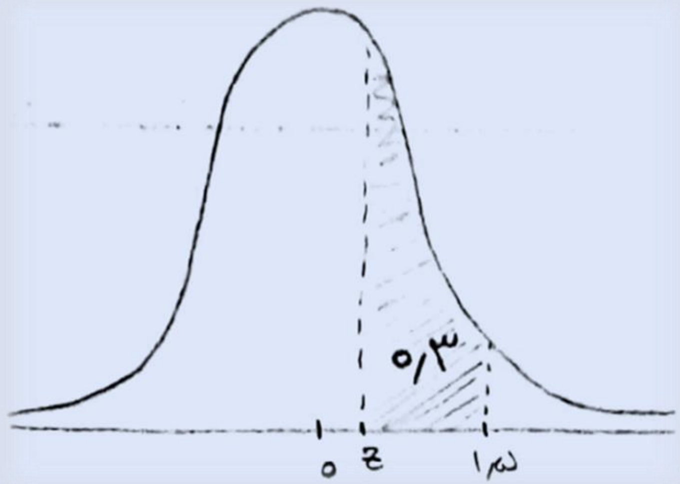
$P(Z \ge z) = 0.7517$

$= P(Z \le -z) \xrightarrow{\text{با مشاهده جدول}} -z = 0.68$

$\Longrightarrow z = -0.68$



ج) $P(z \le Z \le 1.5) = 0.3$



$$\Longrightarrow \underbrace{P(Z \le 1.5)}_{0.9332\ \text{(با مشاهده جدول)}} - (Z < z) = 0.3$$

$$\Longrightarrow P(Z < z) = 0.9332 - 0.3 = 0.6332$$

$$\xrightarrow{\text{با مشاهده جدول}} \quad z \approx 0.3$$

➢ **مثال۵ :** فرض کنید Z دارای توزیع نرمال استاندارد است. مقادیر z را در هر حالت بیابید:

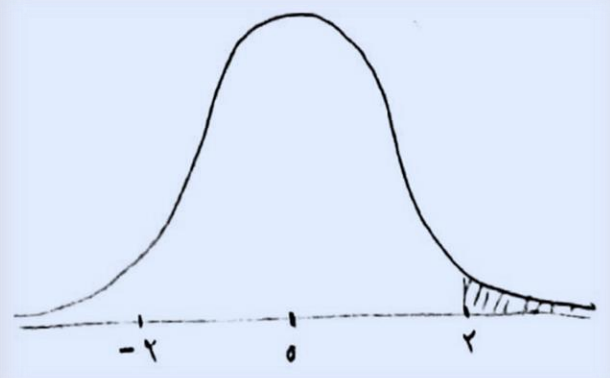
$$P(Z \le z) = 0.8944 \qquad \longrightarrow \qquad z = 1.25$$
$$P(Z < z) = 0.8962 \qquad \longrightarrow \qquad z = 1.26$$
$$P(Z < z) = 0.8950 \qquad \longrightarrow \qquad z \approx 1.25$$
$$P(Z < z) = 0.8959 \qquad \longrightarrow \qquad z \approx 1.26$$
$$P(Z < z) = 0.8953 \qquad \longrightarrow \qquad z \approx 1.255$$

- **مثال۶:** فرض کنید وزن بسته‌های شکلات تولیدشده توسط یک کارخانه دارای توزیع نرمال بوده و به طور متوسط ۵۰۰ گرم با انحراف معیار ۱۰ گرم باشد. احتمال اینکه بسته‌ای بیش از ۵۲۰ گرم وزن داشته باشد چقدر است؟

$\mu = 500 \qquad \sigma = 10$

$$P(X > 520) = P\left(\frac{X-500}{10} > \frac{520-500}{10}\right) = P(Z > 2)$$

$$\begin{cases} = 1 - P(Z \le 2) = 1 - 0.9772 = 0.0228 \\ = P(Z \le -2) = 0.0228 \end{cases}$$



➢ **مثال۷ :** قطر نوعی میلگرد خریداری شده دارای توزیع نرمال با میانگین ۱۶ میلی‌متر و انحراف معیار ۰/۲ میلی‌متر است. با احتمال ۹۵٪ قطر میلگردها از چه عددی بزرگتر است؟

$$P(X > x) = 0.95 \qquad x = ?$$

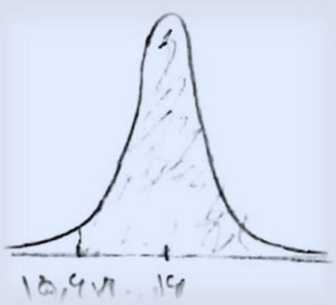
$$0.95 = P(X > x) = P\left(\frac{X-16}{0.2} > \frac{x-16}{0.2}\right) = P(Z > z) = 1 - P(Z \le z)$$

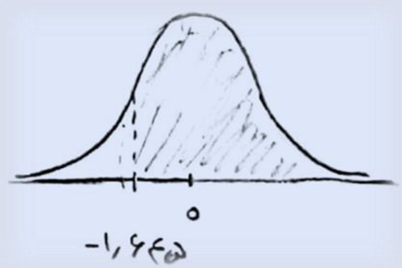
$$\Longrightarrow P(Z \le z) = 0.05 \qquad z = -1.645$$

$$\boxed{\frac{\boldsymbol{x}-\boldsymbol{\mu}}{\boldsymbol{\sigma}} = \boldsymbol{z} \qquad \Leftrightarrow \qquad \boldsymbol{x} = \boldsymbol{\sigma z} + \boldsymbol{\mu}}$$

$$\Longrightarrow x = 0.2(-1.654) + 16 = 15.671$$

با احتمال ۹۵٪ قطر میلگردها از ۱۵/۶۷۱ میلی‌متر بیشتر است.



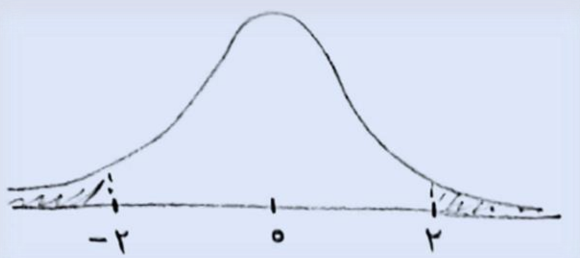


مثال۸: یک مهندس کنترل کیفیت وزن قطعات ساخته شده در مرحله‌ای از خط تولید یک کارخانه را بررسی می‌کند. طبق استاندارد، قطعاتی که وزن معادل ۱±۱۰ گرم دارند قابل قبول بوده و بقیه قطعات باید از فرآیند تولید خارج شوند. فرض کنید وزن قطعات ساخته شده دارای توزیع نرمال با میانگین ۱۰ گرم و واریانس ۰/۲۵ باشد. انتظار می‌رود چند درصد قطعات ساخته شده از فرآیند تولید خارج شوند؟

$$P\left(11 < X \quad \text{یا} \quad X < 9\right) = 1 - P(9 \leq X \leq 11)$$

$$= 1 - P\left(\frac{9-10}{0.5} < \frac{X-10}{0.5} < \frac{11-10}{0.5}\right) = 1 - P(-2 \leq Z \leq 2)$$

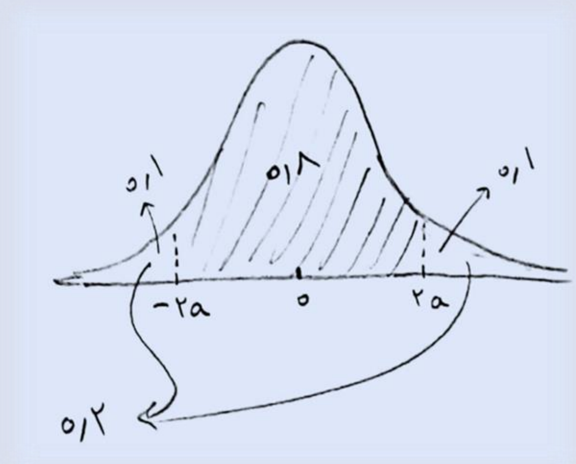
$$= 1 - \left(P(Z \leq 2) - P(Z \leq -2)\right) = 1 - (0.9772 - 0.0228) = \mathbf{0.0456}$$



بین ۴ تا ۵ درصد

- **مثال۹ :** در مثال ۷ ، اگر ۸۰٪ قطعات تولیدشده وزنی معادل ۱۰$\pm a$ گرم داشته باشند، مقدار a را بیابید.

$$P(10 - a < X < 10 + a) = 0.8$$
$$\Rightarrow P\left(\frac{10 - a - 10}{0.5} < \frac{X - \mu}{\sigma} < \frac{10 + a - 10}{0.5}\right) = 0.8$$
$$\Rightarrow P(-2a < Z < 2a) = 0.8$$
$$\Rightarrow P(Z \le 2a) = 0.1$$
$$\Rightarrow P(Z \le 2a) = 0.1$$
$$\Rightarrow z = -2a = -1.28$$
$$\Rightarrow \boldsymbol{a = 0.64}$$



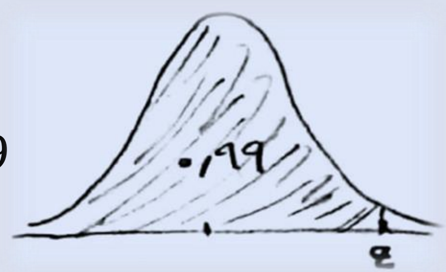
$$P(9.36 < X < 10.64) = 0.8$$

$$❈\ \boldsymbol{P(\underbrace{10 - a}_{x_1} < X < \underbrace{10 + a}_{x_2}) = P(\underbrace{-2a}_{z_1} < Z < \underbrace{2a}_{z_2})}$$

- **مثال ۱۰:** ولتاژی نویزی دارای توزیع نرمال با میانگین $\mu = 0$ و انحراف معیار $\sigma = 0.75$ ولت است. اگر X نشان‌دهنده‌ی ولتاژ نویز باشد، الف) احتمال اینکه $|x| \leq 1.5$ چقدر است؟ ب) ولتاژ نویز X از چه سطح ولتاژی در ۹۹٪ موارد تجاوز نمی‌کند؟ ج) $|x|$ از چه سطح ولتاژی در ۹۹٪ موارد تجاوز نمی‌کند؟

الف) $P(|x| \leq 1.5) = P(-1.5 \leq X \leq 1.5) = P\left(\frac{-1.5-0}{0.75} \leq \frac{X-0}{0.75} \leq \frac{1.5-0}{0.75}\right)$
$= P(-2 \leq Z \leq 2) = P(Z \leq 2) - P(Z < -2) = 0.9772 - 0.0228 = 0.9544$

ب) $P(X \leq x) = 0.99 \quad \Rightarrow \quad P\left(\frac{X-0}{0.75} \leq \frac{x-0}{0.75}\right) = 0.99$



$\Rightarrow \quad P(Z \leq z) = 0.99 \quad \Rightarrow z = 2.33 \quad \Rightarrow \quad x = \sigma z + \mu = 1.7475$ ولت

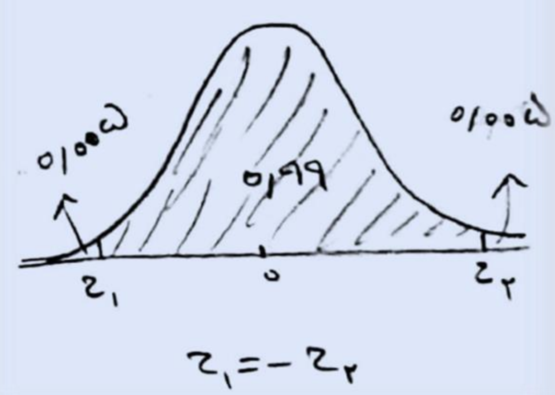
ج) $P(|X| \le x) = 0.99 \Rightarrow P(-x \le X \le x) = P\left(\frac{-x-0}{0.75} \le X \le \frac{x-0}{0.75}\right) = 0.99$

$\Rightarrow \quad P(z_1 \le Z \le z_2) = 0.99 \quad \Rightarrow \quad P(Z \le z_1) = 0.005$

$\Rightarrow z_1 = -2.575 \qquad z_2 = 2.575$

$$\Rightarrow \underbrace{-2.575 \times 0.75 + 0}_{x_1} \le X \le \underbrace{2.575 \times 0.75 + 0}_{x_2}$$

$\Rightarrow -1.931 \le X \le 1.931$



تقریب توزیع دوجمله‌ای با توزیع نرمال :

$$\text{تقریب خوب}\begin{cases} p \nrightarrow 0 & \text{اگر } n \text{ بزرگ باشد و} \\ p \nrightarrow 1 & \\ p \to {}^{1}/_{2} & \text{اگر } n \text{ کوچک باشد و} \end{cases}$$

قضیه: اگر X متغیر تصادفی دوجمله‌ای با میانگین $\mu = np$ و واریانس $\sigma^2 = npq$ باشد، آنگاه شکل حدی توزیع $z = \frac{X-np}{\sqrt{npq}}$ وقتی $n \to \infty$ توزیع نرمال استاندارد $n(z\,;0\,,1)$ است.

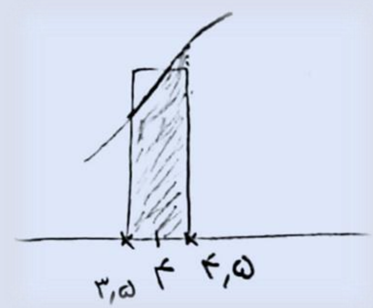
نکته:

$$X \sim b(x;n,p) \quad \Rightarrow \quad P(X=x) = b(x;n,p) \approx P(z_1 \le Z \le z_2)$$

گسسته‌سازی کمیت‌های پیوسته

$$z_1 = \frac{(x-0.5)-np}{\sqrt{npq}}$$

$$z_2 = \frac{(x+0.5)-np}{\sqrt{npq}}$$



مثال ۱۱: امتحانی شامل ۸۰ سوال تستی ۴ گزینه‌ای است. احتمال اینکه پاسخ‌های کاملا حدسی دانشجویی منجر به ۲۵ الی ۳۰ پاسخ درست شود چقدر است؟

$$p = \frac{1}{4} \quad , \quad n = 80 \quad , \quad \mu = np = 20$$

$$\sigma = \sqrt{npq} = \sqrt{\frac{1}{4} \times 80 \times \frac{3}{4}} = 3.873$$

$$P(25 \leq X \leq 30) = \sum_{x=25}^{30} b(x\,; 80\,, \frac{1}{4})$$

در منحنی نرمال با میانگین μ و واریانس σ^2 ، باید مساحت بین $x_1 = 24.5$ و $x_2 = 30.5$ را حساب کنیم.

$$P(24.5 \leq X \leq 30.5) = P\left(\frac{24.5 - 20}{3.873} \leq Z \leq \frac{30.5 - 20}{3.873}\right)$$

$$= P(1.16 < Z < 2.71) = P(Z < 2.71) - P(Z \leq 1.116)$$

$$= 0.9966 - 0.8770 = \mathbf{0.1196}$$

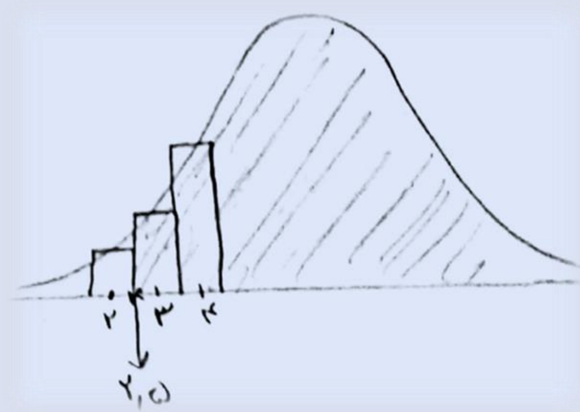
➢ **مثال۱۲:** شرکتی قطعاتی تولید می‌کند که ادعا دارد ۹۵٪ آن‌ها کیفیت مطلوب را دارد. در یک محموله ۱۰۰ تایی از قطعات مورد نظر، احتمال اینکه بیش از دو قطعه معیوب باشد چقدر است؟

پیروزی = معیوب بودن $\quad p = 0.05 \quad , \quad \mu = np = 5 \quad , \quad \sigma = \sqrt{npq} = 2.179$

$$P(X > 2) \approx P\left(\frac{X-5}{2.179} > \frac{2.5-5}{2.179}\right) = P(Z > -1.15) = 1 - P(Z < -1.15)$$

$= 1 - 0.1251 = 0.8749$ تقریب توزیع نرمال

0.8753 تقریب توزیع پوآسن



➢ تابع گاما :

تعریف: تابع گاما برای عدد حقیقی مثبت α عبارت است از:

$$\Gamma(\alpha)=\int_0^{\infty} x^{\alpha-1}\,e^{-x}dx \qquad ; \qquad \alpha>0$$

نکته:

$$\Gamma(\alpha)=\int_0^{\infty} x^{\alpha-1}\,e^{-x}dx=-x^{\alpha-1}e^{-x}\Big|_0^{\infty}-\int_0^{\infty}-(\alpha-1)\,x^{\alpha-2}e^{-x}dx$$

$$=(\alpha-1)\Gamma(\alpha-1)$$

$$\Rightarrow \Gamma(\alpha)=(\alpha-1)\Gamma(\alpha-1)=(\alpha-1)(\alpha-2)\Gamma(\alpha-2)\ldots$$

نکته: اگر $\alpha=\mathrm{n}$ آنگاه:

$$\Gamma(n)=(n-1)\Gamma(n-1)=(n-1)(n-2)\Gamma(n-2)=(n-1)(n-2)\ldots\Gamma(1)$$

$$\Gamma(1)=\int_0^{\infty} e^{-x}\,dx=-e^{-x}\Big|_0^{\infty}=\mathbf{1}$$

$$\Rightarrow \Gamma(n)=(n-1)!$$

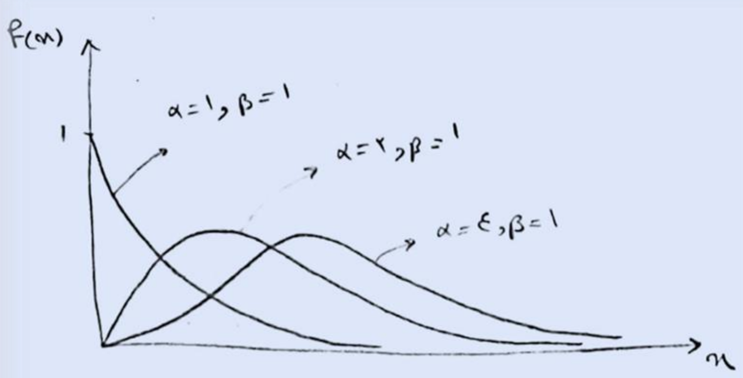
نکته: $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$

➢ توزیع گاما (Gamma distribution):

متغیر تصادفی پیوسته X دارای توزیع گاما با پارامترهای α و β است در صورتی که تابع چگالی آن به صورت زیر باشد:

$$f(x) = \begin{cases} \dfrac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} & x > 0 \\ 0 & o.w \end{cases}$$

که در آن $\alpha > 0$ و $\beta > 0$ و $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$



❈ تعریف خانواده‌ای از توزیع
❈ مسائل نظریه قابلیت اعتماد
❈ مسائل نظریه صف
❈ کاربردهای دیگر

➢ **قضیه:** میانگین و واریانس توزیع گاما عبارتند از:

$$\mu = \alpha\beta \qquad ; \qquad \sigma^2 = \alpha\beta^2$$

اثبات:

$$E(X) = \int_0^\infty \frac{x}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} dx = \frac{\beta}{\Gamma(\alpha)} \int_0^\infty \frac{x^\alpha}{\beta^\alpha} . e^{-\frac{x}{\beta}} . \frac{1}{\beta} dx \qquad \begin{cases} u = \frac{x}{\beta} \\ du = \frac{1}{\beta} dx \end{cases}$$

$$= \frac{\beta}{\Gamma(\alpha)} \int_0^\infty u^\alpha e^{-u} du = \frac{\beta}{\Gamma(\alpha)} \Gamma(\alpha + 1) = \frac{\beta}{\Gamma(\alpha)} \alpha\Gamma(\alpha) = \boldsymbol{\alpha\beta}$$

$$E(X^2) = \int_0^\infty \frac{x^2}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} dx = \frac{\beta^2}{\Gamma(\alpha)} \int_0^\infty \frac{x^{\alpha+1}}{\beta^{\alpha+1}} . e^{-\frac{x}{\beta}} . \frac{1}{\beta} dx$$

$$= \frac{\beta^2}{\Gamma(\alpha)} \int_0^\infty u^{\alpha+1} e^{-u} du = \frac{\beta^2}{\Gamma(\alpha)} \Gamma(\alpha + 2) = \frac{\beta^2}{\Gamma(\alpha)} (\alpha + 1)\alpha\Gamma(\alpha) = \beta^2(\alpha^2 + \alpha)$$

$$\sigma_X^2 = E(X^2) - \big(E(X)\big)^2 = \beta^2(\alpha^2 + \alpha) - \alpha^2\beta^2 = \beta^2(\alpha^2 + \alpha - \alpha^2) = \boldsymbol{\alpha\beta^2}$$

➢ **مثال۲:** زمان باردهی نوعی نهال میوه برحسب سال، دارای توزیع گاما با پارامترهای $\alpha = 2$ و $\beta = 3$ است. الف) احتمال اینکه نهالی خریداری شده از این نوع، زودتر از ۳ سال بار دهد چقدر است؟ ب) احتمال اینکه پس از ۶ سال بار دهد چقدر است؟ ج) زمان مورد انتظار باردهی نهال چقدر است؟

الف) $P(X<3) = \int_0^3 \frac{1}{3^2\Gamma(2)} x^{2-1} e^{-\frac{x}{3}}\,dx = \int_0^3 \frac{1}{9\Gamma(2)} x e^{-\frac{x}{3}}\,dx$ $\quad \begin{cases} u = \frac{x}{3} \\ du = \frac{1}{3}dx \end{cases}$

$= \int_0^1 \frac{1}{\Gamma(2)} u e^{-u}\,du = \mathbf{0.2640}$ $\rightarrow$ محاسبه تابع گامای ناقص

(با استفاده از جدول صفحه ۷۶۷ کتاب انگلیسی)

$x = 1\ ;\ \ \alpha = 2$

ب) $P(X>6) = \int_6^\infty \frac{1}{3^2\Gamma(2)} x^{2-1} e^{-\frac{x}{3}}\,dx = \int_6^\infty \frac{1}{9\Gamma(2)} x e^{-\frac{x}{3}}\,dx$

$= \int_2^\infty \frac{1}{\Gamma(2)} u e^{-u}\,du = 1 - \int_0^2 \frac{1}{\Gamma(2)} u e^{-u} du = 1 - 0.5940 = \mathbf{0.4060}$

(با استفاده از جدول)

ج) $E(X) = \alpha\beta = 2\times 3 = 6$ انتظار داریم ۶ سال پس از کاشت بار دهد

* در توزیع گاما اگر $\alpha = 1$ باشد، توزیع نمایی بدست می‌آید.

➢ توزیع نمایی (Exponential distribution) :

متغیر تصادفی پیوسته X دارای توزیع نمایی با پارامتر β است اگر تابع چگالی آن به صورت زیر باشد:

$$f(x) = \begin{cases} \frac{1}{\beta} e^{-\frac{x}{\beta}} & x > 0 \\ 0 & o.w \end{cases}$$

<u>**نتیجه**</u>: میانگین و واریانس توزیع نمایی عبارتند از:

$$\mu = \beta \qquad ; \qquad \sigma^2 = \beta^2$$

➢ **تمرین ۱:** زمان پاسخ دادن کامپیوتری معین برحسب ثانیه، دارای توزیع نمایی با میانگین ۳ است. الف) احتمال اینکه زمان پاسخ بیشتر از ۵ ثانیه باشد چقدر است؟ ب) احتمال اینکه زمان پاسخ بیشتر از ۱۰ ثانیه باشد چقدر است؟

$$f(x) = \frac{1}{3} e^{-\frac{x}{3}} \qquad ; \qquad x > 0$$

الف) $P(X > 5) = \int_5^\infty \frac{1}{3} e^{-\frac{x}{3}} dx = -e^{-\frac{x}{3}} \Big|_5^\infty = 0 + e^{-\frac{5}{3}} = 0.19$

ب) $P(X > 10) = \int_{10}^\infty \frac{1}{3} e^{-\frac{x}{3}} dx = -e^{-\frac{x}{3}} \Big|_{10}^\infty = 0 + e^{-\frac{10}{3}} = 0.04$

- **مثال ۳:** طول عمر یک نوع لامپ با توزیع نمایی با میانگین $\beta = 2$ سال مدل‌سازی می‌شود. الف) احتمال اینکه این نوع لامپ بیش از سه سال کار کند چقدر است؟ ب) اگر لوستری شامل ۵ عدد از این نوع لامپ باشد، احتمال اینکه این لوستر پس از سه سال، حداقل دو لامپ سالم داشته باشد چقدر است؟

$$\beta = 2 \quad \rightarrow \qquad f(x) = \frac{1}{2} e^{-\frac{x}{2}} \qquad ; \qquad x > 0$$

الف) $P(X > 3) = \int_3^\infty \frac{1}{2} e^{-\frac{x}{2}} dx = -e^{-\frac{x}{2}} \Big|_3^\infty = 0 + e^{-\frac{3}{2}} = \mathbf{0.22}$

ب) $P(Y \geq 2) = \sum_{y=2}^{5} b(x\,; 5\,, 0.22) = 1 - \sum_{y=0}^{1} b(x\,; 5\,, 0.22) = \mathbf{0.3047}$

➢ ارتباط توزیع پوآسن و نمایی و گاما :

پوآسن $\longleftarrow$ تعداد رخداد پیشامدها در زمان

(میانگین $= \lambda \leftarrow$ متوسط تعداد پیشامدها در واحد زمان)

نمایی $\longleftarrow$ مدت زمان لازم برای رخداد اولین پیشامد (یا زمان بین دو رخداد)

(میانگین $= \beta = \frac{1}{\lambda} \leftarrow$ متوسط زمان لازم برای رخداد اولین پیشامد)

گاما $\longleftarrow$ مدت زمان لازم برای رخداد α پیشامد (یا زمان بین یک پیشامد تا α پیشامد بعد)

(میانگین $= \alpha\beta \leftarrow$ متوسط زمان لازم برای رخداد α امین پیشامد)

(یکی از کاربردها ؛ $\alpha = n \leftarrow$ توزیع ارلانگ یا حاصل جمع چند توزیع نمایی)

X : زمان لازم برای وقوع اولین پیشامد پوآسن
Y : متغیر تصادفی پوآسن (تعداد رخداد پیشامد)

$P(X \geq x) = P($تا زمان x پیشامد رخ ندهد$) = poisson(y = 0; \mu = \lambda x)$

$$= \frac{(\lambda x)^0 e^{-\lambda x}}{0!} = e^{-\lambda x}$$

$\Longrightarrow$ X تابع توزیع تجمعی $= F(X) = P(X \leq x) = 1 - P(X \geq x) = 1 - e^{-\lambda x}$

$\Longrightarrow$ X تابع چگالی $= f(X) = \frac{dF}{dx} = \lambda e^{-\lambda x} \underset{\Uparrow}{=} \boldsymbol{\frac{1}{\beta} e^{-\frac{x}{\beta}}}$ چگالی توزیع نمایی

(فرکانس،تعداد رخداد : λ) $\quad \beta = \frac{1}{\lambda} \quad$ (زمان : β)

➢ **مثال ۴:** تعداد دفعات نقص فنی در یک سایت کامپیوتری دارای توزیع پوآسن با میانگین سه‌بار در یک ماه است. الف) احتمال اینکه این سایت در یک ماه بیش از ۴ بار دچار نقص فنی شود چقدر است؟ ب) احتمال اینکه سایت در کمتر از پانزده روز دچار نقص فنی شود چقدر است؟

الف) $X \sim Poisson(3) \quad \begin{matrix} \lambda = 3 \\ \tau = 1 \end{matrix}$

$$P(X > 4) = 1 - P(X \le 4) = 1 - \sum_{x=0}^{4} p(x;3) = 1 - 0.8153 = 0.1847$$

ب) زمان لازم تا وقوع نقص فنی: $Y \sim Exponential\left(\beta = \frac{1}{\lambda} = \frac{1}{3}\right)$

$$P\left(Y < \frac{1}{2}\right) = \int_0^{\frac{1}{2}} \frac{1}{\beta} e^{-\frac{x}{\beta}} dx = \int_0^{\frac{1}{2}} 3e^{-3x}\, dx = -e^{-3x} \Big|_0^{1/2} = -e^{-\frac{3}{2}} + 1 = 0.78$$

➢ **مثال ۵ :** در مثال ۴، الف) احتمال اینکه سایت در طول سه‌ماه، هفت بار نقص فنی پیدا کند چقدر است؟ ب) احتمال اینکه کمتر از ۳ ماه طول بکشد تا ۷ بار نقص فنی پیدا کند چقدر است؟

الف) $\begin{matrix} \lambda = 3 \\ \tau = 3 \end{matrix} \quad \rightarrow \quad X \sim Poisson(9)$

$$P(X=7) = P(7,9) = \sum_{x=0}^{7} p(x;9) - \sum_{x=0}^{6} p(x;9) = 0.3239 - 0.2068 = 0.1171$$

ب) زمان لازم برای رخداد ۷ نقص فنی: $Y \sim Gamma\left(\alpha = 7, \beta = \frac{1}{3}\right)$

$$P(Y<3) = \int_0^3 \frac{1}{\left(\frac{1}{3}\right)^7 \Gamma(7)} x^{7-1} e^{-\frac{x}{\frac{1}{3}}} dx = \int_0^3 \frac{1}{\Gamma(7)} (3)^6 x^6 e^{-3x} \times 3\, dx \xrightarrow{u=3x\,,\, du=3dx}$$

$$= \int_0^9 \frac{1}{\Gamma(7)} u^6 e^{-u}\, du = 0.7930$$ (با استفاده از جدول)

β

➢ **مثال ۶:** زمان میان رسیدن دو مراجع به دستگاه خودپرداز دارای توزیع نمایی با میانگین ۵ دقیقه است. الف) احتمال اینکه بیش از چهار مشتری در ده دقیقه برسند چقدر است؟ ب) احتمال اینکه زمان لازم برای رسیدن مشتری پنجم کمتر از ۱۵ دقیقه باشد چقدر است؟

الف) $X \sim Exponential(\beta = 5)$ زمان رسیدن مشتری بعدی

تعداد مشتری‌ها در ۱۰ دقیقه: $Y \sim Poisson(\lambda\tau) = Poisson(2)$

$\lambda = \frac{1}{5}$ یک دقیقه = واحد زمان $\tau = 10$

$$P(X > 4) = 1 - P(X \leq 4) = 1 - \sum_{x=0}^{4} p(x; 3) = 1 - 0.8153 = 0.1847$$

ب) زمان لازم برای رسیدن پنجمین مشتری: $Z \sim Gamma(\alpha = 5, \beta = 5)$

$$P(Z \leq 15) = \int_0^{15} \frac{1}{5^5 \Gamma(5)} x^{5-1} e^{-\frac{x}{5}} dx = \int_0^{15} \frac{1}{\Gamma(5)} \frac{x^4}{5^4} e^{-\frac{x}{5}} . \frac{1}{5} dx$$

$$\xrightarrow{u=\frac{x}{5}\ ,\ du=\frac{1}{5}dx} \int_0^{3} \frac{1}{\Gamma(5)} u^4 e^{-u} du = 0.1850$$

➢ **مثال ۷ :** در مطالعه درجه خلوص مواد خام، فرض کنید تعداد ذرات ناخالصی در یک گرم از ماده‌ای دارای توزیع پوآسن با میانگین ۰/۰۲ ذره در هر گرم باشد.
الف) مقدار مورد انتظار وزن ماده لازم جهت یافتن ۱۵ ذره ناخالصی چقدر است؟
ب) انحراف معیار وزن ماده لازم جهت یافتن ۱۵ ذره ناخالصی چقدر است؟

$\lambda = 0.02$ یک گرم = واحد وزن (جرم)

مقدار وزن لازم جهت یافتن ۱۵ ذره: $X \sim Gamma\left(\alpha = 15, \beta = \frac{1}{\lambda}\right)$

$\Longrightarrow \mu_X = \alpha\beta = 15 \times 50 = 750\ g$

$\sigma_X^2 = \alpha\beta^2 = 15 \times 50^2 = 37500\ g^2$

$\sigma_X = \sqrt{37500} \approx 193.65\ g$

تمرین ۶.۵۸ کتاب (ورژن انگلیسی) در صفحه ۲۰۷ را حل کنید.

در توزیع گاما: اگر $\beta = 2$ و $\alpha = \frac{\nu}{2}$ $\longleftarrow$ توزیع خی- دو با ν درجه آزادی

توزیع‌های نمونه‌ای
تحلیل واریانس
آمار ناپارامتری

➢ توزیع خی- دو Chi-squared distribution (χ^2) :

متغیر تصادفی پیوسته X ، دارای توزیع خی- دو با ν درجه آزادی است، اگر تابع چگالی آن به صورت زیر باشد:

$$f(x) = \begin{cases} \dfrac{1}{2^{\frac{\nu}{2}}\Gamma(\frac{\nu}{2})} x^{\frac{\nu}{2}-1} e^{-\frac{x}{2}} & x > 0 \\ 0 & o.w \end{cases}$$

که در آن ν یک عدد صحیح مثبت است.

نتیجه: میانگین و واریانس توزیع خی- دو عبارتند از:

$$\mu = \nu \qquad ; \qquad \sigma^2 = 2\nu$$

* در بسیاری از مسائل به جای کار با خود داده‌ها، از لگاریتم طبیعی (Ln) آن‌ها استفاده می‌کنیم. اگر لگاریتم طبیعی داده‌ها دارای توزیع نرمال باشد، می‌گوییم داده‌ها دارای توزیع نرمال لگاریتمی هستند.

- **توزیع نرمال لگاریتمی (Lognormal distribution) :**

متغیر تصادفی پیوسته X دارای توزیع نرمال لگاریتمی است اگر متغیر تصادفی $Y = Ln(X)$ دارای توزیع نرمال با میانگین μ و انحراف معیار σ باشد. تابع چگالی حاصل برحسب متغیر تصادفی X به صورت زیر است:

$$f(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{(lnx-\mu)^2}{2\sigma^2}} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

نتیجه: میانگین و واریانس توزیع نرمال لگاریتمی برابر است با :

$$E(X) = e^{\mu+\frac{\sigma^2}{2}} \quad ; \quad \sigma^2 = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$$

(برحسب گالن در ساعت)

- **تمرین ۲:** معلوم شده است که نرخ متوسط مصرف آب در شهری دارای توزیع نرمال لگاریتمی با پارامترهای $\mu = 5$ و $\sigma = 2$ است. احتمال اینکه در ساعت مفروضی بیش از ۵۰۰۰۰ گالن مصرف شود چقدر است؟

$$P(X > 50000) = 1 - P(X \leq 50000) = 1 - P(ln\, x \leq ln\, 50000)$$

$$1 - P\left(\frac{Ln\, X\ -\ \mu}{\sigma} \leq \frac{10.85\ -\ \mu}{\sigma}\right) = 1 - P(Z \leq 2.91) = 1 - 0.9982 = 0.0018$$