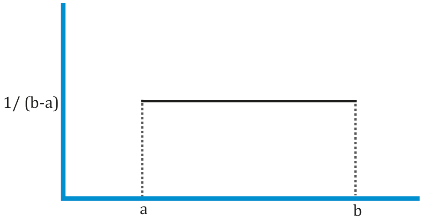
توزیع‌های شناخته شده‌ی پیوسته
-آمار و احتمالات مهندسی-

مدرس: مشکانی فراهانی

دانشگاه صنعتی امیرکبیر

۲۵ آبان ۱۳۹۹

توزیع یکنواخت پیوسته



توزیع یکنواخت پیوسته

ساده‌ترین توزیع در بین توزیع‌های احتمالی پیوسته که شبیه به همتای گسسته‌ی آن است.

متغیر تصادفی X دارای توزیع یکنواخت در فاصله‌ی $(a\ ,\ b)$ است، هرگاه دارای تابع چگالی زیر باشد:

$$f(x) = \frac{۱}{b-a}, \qquad a < x < b \tag{۱}$$

نمادی که برای توزیع یکنواخت پیوسته استفاده می‌شود: $X \sim U(a\ ,\ b)$

میانگین و واریانس توزیع یکنواخت پیوسته عبارت است از:

$$\mu = E(X) = \frac{a+b}{۲} \qquad\qquad \sigma^۲ = Var(X) = \frac{(b-a)^۲}{۱۲}$$

مثال ۱

فرض کنید B یک عدد تصادفی از بازه‌ی $[-۳,۳]$ باشد. احتمال این‌که معادله‌ی درجه دوم $x^۲ + Bx + ۱ = ۰$ حداقل یک ریشه‌ی حقیقی داشته باشد، چه‌قدر است؟

راه‌حل:

$$B \sim U(-۳,۳) \qquad\qquad f(b) = \frac{۱}{۶}$$

$$\begin{aligned} P\left(B^۲ - ۴ \geq ۰\right) &= P\left(B^۲ \geq ۴\right) = P\left(|B| \geq ۲\right) \\ &= ۱ - P(-۲ < B < ۲) = ۱ - \int_{-۲}^{۲} \frac{۱}{۶} db \\ &= ۱ - \frac{۴}{۶} = \frac{۱}{۳} \end{aligned}$$

مثال ۲

تابع چگالی احتمال مدت زمان لازم برای تکمیل مونتاژ وسیله‌ای $f(x) = ۰/۱$ برای $۳۰ < x < ۴۰$ ثانیه است. میانگین و واریانس مدت زمان مونتاژ چه‌قدر است؟

راه‌حل:

$$\mu = \frac{۳۰ + ۴۰}{۲} = ۳۵$$

$$\sigma^۲ = \frac{(۴۰ - ۳۰)^۲}{۱۲} = ۸/۳$$

وزن خالص یک علف‌کش بسته‌بندی شده بر حسب پوند دارای توزیع یکنواخت در بازه‌ی $۴۹/۷۵ < x < ۵۰/۲۵$ پوند است. چند درصد از آن‌ها وزنی بیشتر از ۵۰ پوند دارند؟

راه‌حل:

$$X \sim U(۴۹/۷۵, ۵۰/۲۵) \qquad \Rightarrow \qquad f(x) = \frac{۱}{۵۰/۲۵ - ۴۹/۷۵} = ۲$$

$$P(X > ۵۰) = \int_{۵۰}^{۵۰/۲۵} ۲dx = ۰/۵$$

$$۰/۵ \times ۱۰۰ = ۵۰\%$$

مثال ۴

مدت زمانی که طول می‌کشد تا شخصی فاصله‌ی بین خانه تا ایستگاه مترو را پیاده طی کند، دارای توزیع یکنواخت بین ۱۵ تا ۲۰ دقیقه است. اگر شخصی ساعت ۷:۳۰ از منزل خارج شود، احتمال این‌که او سوار قطاری شود که ساعت ۷:۴۸ به ایستگاه می‌رسد.

راه‌حل:

X: زمان طی کردن فاصله‌ی خانه تا مترو $\qquad X \sim U(۱۵, ۲۰)$

$$P(X < ۱۸) = \int_{۱۵}^{۱۸} \frac{۱}{۲۰ - ۱۵} dx = \frac{۳}{۵}$$

توزیع‌های
نمایی و گاما و کای-دو

توزیع نمایی

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

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

- این توزیع را با نماد $X \sim Exp(\beta)$ نمایش می‌دهیم.
- توزیع نمایی مدت زمان لازم تا اولین رخداد یا زمان بین دو رخداد متوالی را می‌تواند محاسبه نماید.
- از جمله کاربردهای توزیع نمایی عبارتند از:
  - خطوط انتظار یا صف‌ها
  - زمان ورود به عوارضی جاده‌ها
  - زمان خرابی قطعات با نرخ خرابی ثابت
  - زمان متلاشی شدن یک ذره‌ی رادیواکتیو
  - زمان ورود یک آمبولانس به صحنه‌ی تصادف

توزیع نمایی

- پارامتر β در توزیع نمایی متوسط زمان لازم برای وقوع اولین رخداد یا بین دو رخداد متوالی تعریف می‌شود.

- میانگین و واریانس توزیع نمایی به صورت زیر محاسبه می‌شود:

$$E(X) = \beta \qquad\qquad \sigma^2 = \beta^2$$

مثال ۵

مدت زمان لازم برای تعمیر یک اتومبیل در یک مرکز خدمات اتومبیل دارای توزیع نمایی با پارامتر ۳ ساعت است. مطلوبست محاسبه احتمال اینکه مدت زمان تعمیر بیش از ۳ ساعت به طول انجامد.

راه‌حل:

X: زمان لازم برای تعمیر اتومبیل

$$\begin{aligned} P\left(X > ۳\right) &= \int_{۳}^{\infty} \frac{۱}{۳} e^{-\frac{x}{۳}} dx \\ &= -e^{-\frac{x}{۳}}\Big]_{۳}^{\infty} \\ &= e^{-۱} \end{aligned}$$

مثال ۶

مدت زمانی که یک ساعت بدون وقفه کار می‌کند، متغیر تصادفی نمایی با پارامتر ۵۰ روز است. احتمال اینکه ساعتی کمتر از ۲۰ روز بدون وقفه کار کند را بیابید.

راه‌حل:
X: زمان کار کردن ساعت تا خرابی آن

$$P(X < ۲۰) = \int_{۰}^{۲۰} \frac{۱}{۵۰} e^{-\frac{x}{۵۰}} dx$$
$$= -e^{-\frac{x}{۵۰}}\Big]_{۰}^{۲۰}$$
$$= ۱ - e^{-\frac{۲}{۵}}$$

مثال ۷

فرض کنید متغیر تصادفی X نشانگر طول عمر نوعی لاستیک بر حسب ساعت باشد که دارای توزیع نمایی با میانگین ۵۰۰ ساعت است. اگر بدانیم که طول عمر یک لاستیک از این نوع از ۷۰۰ ساعت کمتر بوده است، احتمال اینکه طول عمر آن از ۳۰۰ ساعت بیشتر باشد را بیابید.

راه‌حل:

$$
\begin{aligned}
P(X > ۳۰۰ | X < ۷۰۰) &= \frac{P(X > ۳۰۰, X < ۷۰۰)}{P(X < ۷۰۰)} \\
&= \frac{P(۳۰۰ < X < ۷۰۰)}{P(X < ۷۰۰)} \\
&= \frac{\int_{۳۰۰}^{۷۰۰} \frac{۱}{۵۰۰} e^{-\frac{x}{۵۰۰}} dx}{\int_{۰}^{۷۰۰} \frac{۱}{۵۰۰} e^{-\frac{x}{۵۰۰}} dx} \\
&= \frac{e^{-\frac{۳}{۵}} - e^{-\frac{۷}{۵}}}{۱ - e^{-\frac{۷}{۵}}}
\end{aligned}
$$

مثال ۸

طول عمر هر دستگاه کامپیوتر دارای توزیع نمایی با میانگین ۱۷۰۰ ساعت می‌باشد. اگر آزمایشگاهی ۲۰ دستگاه کامپیوتر داشته باشد، احتمال اینکه حداقل یک دستگاه از آنها قبل از ۱۷۰۰ ساعت خراب شوند، را بیابید.

راه‌حل:

X: طول عمر هر دستگاه کامپیوتر $\qquad X \sim Exp\left(E(X) = \beta = ۱۷۰۰\right)$

Y: تعداد دستگاه‌هایی که قبل از ۱۷۰۰ ساعت خراب می‌شوند $\qquad Y \sim Bin(n = ۲۰, p = ?)$

$$p = P\left(\text{پیروزی}\right) = P\left(\text{طول عمر} < ۱۷۰۰\right) = P\left(X < ۱۷۰۰\right)$$
$$= \int_{۰}^{۱۷۰۰} \frac{۱}{۱۷۰۰} e^{-\frac{x}{۱۷۰۰}} dx$$
$$= -e^{-\frac{x}{۱۷۰۰}}\Big]_{۰}^{۱۷۰۰} = ۱ - e^{-۱}$$

$$P\left(Y \geq ۱\right) = ۱ - P\left(Y = ۰\right) = ۱ - \binom{۲۰}{۰} \left(۱ - e^{-۱}\right)^{۰} \left(e^{-۱}\right)^{۲۰} = ۱ - e^{-۲۰}$$

قضیه: اگر X نشان‌دهنده‌ی مدت زمان طی شده تا وقوع اولین رخداد در یک آزمایش پواسون با پارامتر λ باشد، آنگاه X یک متغیر تصادفی پیوسته با تابع چگالی $f(x) = \lambda e^{-\lambda x}$ است.

اثبات:
فرض کنید t یک عدد حقیقی مثبت و متغیر تصادفی N نشان‌دهنده‌ی تعداد رخدادها در زمان $[۰, t]$ باشد. زمان اولین رخداد فقط در صورتی بعد از زمان t است که در فاصله‌ی $[۰, t]$ هیچ رخدادی روی ندهد. پس $P(X > t) = P(N = ۰)$.

چون $N \sim P(\lambda t)$، پس $P(N = ۰) = e^{-\lambda t}$. در نتیجه

$$F_X(t) = P(X \leq t) = ۱ - P(X > t) = ۱ - e^{-\lambda t}, \qquad t > ۰$$

$$\Rightarrow \qquad f_X(t) = F'(t) = \lambda e^{-\lambda t}, \qquad t > ۰$$

رابطه‌ی توزیع نمایی و پواسون

در صورتی که تعداد اتفاقات در واحد زمان دارای توزیع پواسون با میانگین λ اتفاق باشد. آن‌گاه، متغیر تصادفی $X \geq 0$ زمان بین دو اتفاق متوالی یا زمان لازم برای اولین اتفاق دارای توزیع نمایی با میانگین زمان $\frac{1}{\lambda}$ است.

$$X \sim Exp\left(\beta = \frac{1}{\lambda}\right)$$

مثال ۹

اگر تعداد رانندگان با سرعت غیرمجاز که یک واحد رادار در مکان معینی از جاده‌ای در یک ساعت ثبت می‌کند، متغیری پواسون با $\lambda = ۸$ باشد، احتمال زمان انتظار کمتر از ۱۰ دقیقه بین مشاهدات متوالی رانندگان با سرعت غیرمجاز چه‌قدر است؟

راه‌حل:

$$\begin{aligned} P\left(\text{زمان انتظار} < ۱۰ \text{ دقیقه}\right) &= P\left(X < \frac{۱}{۶}\text{ ساعت}\right) \\ &= \int_{۰}^{\frac{۱}{۶}} ۸e^{-۸x}dx \\ &= -e^{-۸x}\Big]_{۰}^{\frac{۱}{۶}} \\ &= ۱ - e^{-\frac{۴}{۳}} \end{aligned}$$

مثال ۱۰

به طور متوسط تعداد ۵ تلفن در یک ساعت به تلفن‌خانه‌ی یک شرکت زده می‌شود. مطلوب است احتمال این‌که
الف- در یک ساعت حداقل ۲ تلفن زده شود.
ب- تلفن بعدی حداقل بعد از ۱۵ دقیقه زده شود.

راه‌حل:

X: تعداد تماس‌ها در یک ساعت $\qquad X \sim P(\lambda = ۵)$

Y: زمان بین دو تماس متوالی $\qquad Y \sim Exp(\beta = \frac{۱}{\lambda} = \frac{۱}{۵})$

الف $-$ $\quad P\left(X \geq ۲\right) = ۱ - P(X = ۰) - P(X = ۱) = ۱ - e^{-۵} - \frac{e^{-۵} \times ۵^{۱}}{۱}$

ب $-$ $\quad P\left(Y \geq ۱۵ \text{ دقیقه}\right) = P\left(Y \geq \frac{۱}{۴} \text{ ساعت}\right) = \int_{\frac{۱}{۴}}^{\infty} ۵e^{-۵y} dy = e^{-\frac{۵}{۴}}$

مثال ۱۱

فاصله بین دست‌اندازهای بزرگ در یک بزرگراه از توزیع نمایی با میانگین ۵ مایل پیروی می کند.
الف- احتمال عدم وجود دست‌اندازهای بزرگ در امتداد ۱۰ مایلی بزرگراه چقدر است؟
ب- احتمال اینکه دو دست‌انداز بزرگ در یک فاصله‌ی ۱۰ مایلی بزرگراه وجود داشته باشد، چقدر است؟
ج- انحراف استاندارد فاصله بین دست‌اندازهای بزرگ چه‌قدر است؟

راه‌حل:

X: فاصله‌ی بین دو دست‌انداز بزرگ $\qquad X \sim Exp(E(X) = \beta = ۵)$

Y: تعداد دست‌اندازها در ۱۰ مایل $\qquad Y \sim P(۱۰ \times \frac{۱}{۵} = ۲)$

الف − $\quad P(X > ۱۰) = \int_{۱۰}^{\infty} \frac{۱}{۵} e^{-\frac{x}{۵}} = -e^{-\frac{x}{۵}} \,\Big]_{۱۰}^{\infty} = e^{-۲}$

ب − $\quad P(Y = ۲) = \dfrac{e^{-۲} \times ۲^{۲}}{۲!} = ۰/۲۷$

ج − $\quad \sigma_X = \beta = ۵$

خاصیت بی‌حافظگی توزیع نمایی

در بین توزیع‌های پیوسته فقط توزیع نمایی دارای این خاصیت است:

$$P\left(X > a + b | X > a\right) = P\left(X > b\right), \qquad \forall a, b > ۰$$

اگر X طول عمر نوعی ابزار باشد، طبق رابطه‌ی فوق عمر ابزار با زمان کارکرد آن ارتباطی ندارد.

احتمال این‌که ابزار نویی بیش از b سال کار کند برابر است با احتمال این‌که ابزار کهنه‌ای که مثلاً بیش از a سال کار کرده، حداقل b سالِ دیگر کار کند.

به عبارت دیگر، احتمال این‌که این ابزار b سال آینده از کار باز بماند، به مدت زمانی که تا کنون کار کرده بستگی ندارد.

مثال ۱۲

فرض کنید مدت زمان مکالمه با تلفن همگانی از توزیع نمایی با میانگین ۴ دقیقه تبعیت کند. احتمال این‌که مدت مکالمه‌ی بیش از ۲۰ دقیقه طول بکشد، در حالی‌که حداقل ۴ دقیقه مکالمه کرده باشد را به دست آورید.

راه‌حل:

$$\begin{aligned} P\left(X > ۲۰ | X > ۴\right) &= P\left(X > ۱۶\right) \\ &= \int_{۱۶}^{\infty} \frac{۱}{۴} e^{-\frac{x}{۴}} dx \\ &= -e^{-\frac{x}{۴}}]_{۱۶}^{\infty} \\ &= e^{-۴} \end{aligned}$$

توزیع گاما

یک متغیر تصادفی نمایی زمان را تا اولین رخداد در یک فرآیند پواسون توصیف می‌کند.
تعمیم توزیع نمایی به صورت زمان تا وقتی است که r رخداد در یک فرآیند پواسون روی دهد.
بدین صورت که اگر اتفاقات در طول زمان رخ دهند، آن‌گاه مدت زمانی که فرد بایستی منتظر بماند تا r اتفاق در یک فرایند پواسون رخ دهد، دارای توزیع گاما است.

تعریف: متغیر تصادفی X که نشان‌دهنده‌ی زمان لازم تا رخداد r اتفاق در یک فرآیند پواسون با میانگین λ است، دارای توزیع گاما با پارامترهای (r, β) است هرگاه تابع چگالی احتمال X به صورت زیر باشد:

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

$$E(X) = r\beta \qquad\qquad \sigma^2 = r\beta^2$$

نکته: اگر در توزیع گاما $r = 1$ قرار دهیم، توزیع نمایی به دست می‌آید.

تابع گاما به صورت زیر تعریف می‌شود:

$$\Gamma(n) = \int_{0}^{\infty} x^{n-1} e^{-x} dx$$

با استفاده از انتگرال‌گیری جزءبه‌جزء با قرار دادن $u = x^{n-1}$ و $dv = e^{-x}dx$ به دست می‌آوریم:

$$\Gamma(n) = (n-1)\int_{0}^{\infty} x^{n-2} e^{-x} dx = (n-1)\Gamma(n-1)$$

و به همین ترتیب با محاسبه‌ی انتگرال‌های جزءبه‌جزء متوالی خواهیم داشت:

$$\Gamma(n) = (n-1)(n-2)(n-3)\cdots\Gamma(1) = (n-1)!$$

مثال ۱۳

در بیمارستانی نوزادان با نرخ ۱۲ نفر در روز متولد می‌شوند. احتمال این‌که حداقل ۷ ساعت طول بکشد تا ۳ نوزاد متولد شوند، چه‌قدر است؟

راه‌حل:

X: زمان لازم تا تولد ۳ نوزاد $\quad X \sim G(r = ۳, \beta = \frac{۱}{\lambda} = \frac{۱}{۱۲})$

$$P\left(X > ۷\ \text{ساعت}\right) = P\left(X > \frac{۷}{۲۴}\ \text{روز}\right)$$
$$= \int_{\frac{۷}{۲۴}}^{\infty} \frac{۱۲^{۳}}{\Gamma(۳)} x^{۲} e^{-۱۲x} dx$$
$$= \frac{۱}{\Gamma(۳)} \times e^{-۱۲x} \left[-۱۲^{۲}x^{۲} - ۲۴x - ۲\right]_{\frac{۷}{۲۴}}^{\infty}$$
$$= \frac{e^{-\frac{۷}{۲}}}{۲} left[\frac{۴۹}{۴} + ۷ + ۲ = ۰/۳۲$$

مثال ۱۴

فرض کنید به طور متوسط در هر ثانیه از یک ماده‌ی رادیواکتیو چهار ذره‌ی بتا منتشر می‌شود. احتمال این‌که برای انتشار دو ذره‌ی بتا حداقل دو ثانیه زمان لازم باشد، چه‌قدر است؟

راه‌حل:

X: زمان لازم از حال تا انتشار ۲ ذره‌ی بتا $\qquad X \sim G(r = ۲, \beta = \frac{۱}{\lambda} = \frac{۱}{۴})$

$$P\left(X \geq ۲\right) = \int_{۲}^{\infty} \frac{۴^{۲}}{\Gamma(۲)} x^{۱} e^{-۴x} dx$$
$$= e^{-۴x}\left[-۴x - ۱\right]_{۲}^{\infty}$$
$$= e^{-۸}\left[۸ + ۱\right] = ۰/۰۰۳$$

توزیع کای-دو

یک حالت خاص و مهم از توزیع گاما با قرار دادن $r = \frac{\nu}{۲}$ و $\beta = ۲$ به دست می‌آید. توزیع حاصل را توزیع کای-دو می‌نامند.

متغیر تصادفی X دارای توزیع کای-دو با ν درجه‌ی آزادی است، اگر تابع چگالی آن به صورت

$$f(x) = \frac{۱}{۲^{\frac{\nu}{۲}}\Gamma(\frac{\nu}{۲})} x^{\frac{\nu}{۲}-۱} e^{-\frac{x}{۲}}, \qquad x > ۰$$

باشد. ν یک عدد صحیح مثبت است.

میانگین و واریانس توزیع کای-دو برابر است با:

$$\mu = \nu \qquad\qquad \sigma^۲ = ۲\nu$$

توزیع نرمال

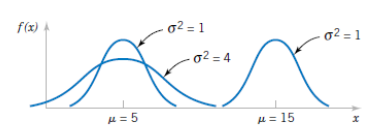
توزیع نرمال

- متغیر تصادفی X دارای توزیع نرمال است، اگر تابع چگالی آن به فرم زیر باشد:

$$f(x) = \frac{۱}{\sqrt{۲\pi\sigma^۲}} e^{-\frac{۱}{۲\sigma^۲}(x-\mu)^۲} \qquad x \in \mathbb{R}$$

- در این تابع $\mu = E(X)$ و $\sigma^۲ = Var(X)$ است: $\sigma^۲ > ۰ \quad , \quad -\infty < \mu < \infty$

- در چنین حالتی می‌گویند که متغیر تصادفی X دارای توزیع نرمال با میانگین μ و واریانس $\sigma^۲$ است و آن را با این نماد نمایش می‌دهیم: $X \sim N(\mu, \sigma^۲)$.



برخی ویژگی‌های توزیع

- نمودار تابع چگالی نرمال زنگی شکل و حول μ متقارن است.

- با افزایش σ^2 پراکندگی توزیع افزایش یافته (نمودار پهن‌تر) و با افزایش μ منحنی به سمت راست انتقال پیدا می‌کند.

- منحنی تنها دارای یک ماکسیمم در نقطه $x = \mu$ است.

- منحنی نسبت به خط $x = \mu$ متقارن است؛ یعنی $f(\mu + a) = f(\mu - a)$. پس $P(X < -a) = P(X > a)$

- در توزیع نرمال میانگین، میانه و مد با هم برابر هستند.

- سطح محصور بین منحنی و محور طول‌ها برابر یک واحد است.

توزیع نرمال استاندارد

توزیع نرمالی که میانگین آن صفر $\mu = ۰$ و واریانس آن یک $\sigma^۲ = ۱$ باشد را توزیع نرمال استاندارد می‌نامند.

متغیر تصادفی با توزیع نرمال استاندارد را با Z نمایش می‌دهند: $Z \sim N(۰, ۱)$

تابع توزیع تجمعی نرمال استاندارد را با Φ نمایش می‌دهند: $\Phi(z) = P(Z \leq z)$

محاسبه‌ی احتمال $\Phi(z) = P(Z \leq z)$ نیاز به حل انتگرالی دارد که بسیار وقت‌گیر است.

به همین دلیل، به ازای مقادیر مختلف a، مقدار احتمال $\Phi(z) = P(Z \leq z)$ در جدول ”نرمال استاندارد“ در پیوست کتاب آمده است.

برای استفاده از جدول دو شرط زیر باید برقرار باشد:
۱- متغیر مورد نظر دارای توزیع نرمال استاندارد باشد: $Z \sim N(۰, ۱)$
۲- احتمال خواسته شده به صورت $\Phi(z) = P(Z \leq z)$ باشد.

مثال ۱۵

مطلوبست محاسبه احتمال آنکه متغیری تصادفی با توزیع نرمال استاندارد مقداری
الف- کمتر از $۱/۵۳$
ب- بیشتر از $۱/۵۳$ را اختیار کند.

راه‌حل:

الف- $$P(Z < ۱/۵۳) = ۰/۹۳۷۰$$

ب- $$P(Z > ۱/۵۳) = ۱ - P(Z \leq ۱/۵۳) = ۱ - ۰/۹۳۷۰$$

مثال ۱۶

مطلوبست محاسبه احتمال پیشامدهای زیر به قسمی که $Z \sim N(۰, ۱)$

الف- $P\left(Z < ۱/۷۲\right) = ۰/۹۵۷۳$

ب- $P\left(۱/۳ < Z < ۱/۸\right) = ۰/۹۶۴۱ - ۰/۹۰۳۲$

ج- $P\left(-۰/۲۵ < Z < ۰/۴۵\right) = ۰/۶۷۳۶ - ۰/۴۰۱۳$

د- $P\left(Z > ۲/۱\right) = ۱ - P\left(Z \leq ۲/۱\right) = ۱ - ۰/۹۸۲۱$

نکته

هر گاه $X \sim N(\mu, \sigma^2)$، آنگاه هر ترکیب خطی از آن به صورت $Y = aX + b$ نیز دارای توزیع نرمال است.

مثال: متغیر تصادفی X دارای توزیع نرمال با میانگین ۱۵۰ و واریانس ۶۴ است. اگر متغیر تصادفی Y بر اساس $Y = \frac{1}{2}X + 25$ از X تبعیت کند، تابع چگالی Y را به دست آورید.

$$E(Y) = E\left(\frac{1}{2}X + 25\right) = \frac{1}{2}E(X) + 25 = \frac{150}{2} + 25 = 100$$

$$Var(Y) = Var\left(\frac{1}{2}X + 25\right) = \frac{1}{4}Var(X) = \frac{64}{4} = 16$$

$$\Rightarrow \qquad Y \sim N(100\ ,\ 16)$$

قضیه

تبدیل نرمال $N(\mu, \sigma^۲)$ به نرمال استاندارد $N(۰, ۱)$

۱- اگر X دارای توزیع نرمال با میانگین μ و واریانس $\sigma^۲$ باشد آنگاه $Z = \frac{X-\mu}{\sigma}$ توزیع نرمال استاندارد دارد.

$$E(Z) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{\mu-\mu}{\sigma} = ۰ \qquad Var(Z) = Var\left(\frac{X-\mu}{\sigma}\right) = \frac{Var(X)}{\sigma^۲} = ۱$$

۲- اگر $X \sim N(\mu, \sigma^۲)$ باشد آنگاه:

$$P(X \leq b) = P\left(\frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right) = P\left(Z \leq \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right)$$

$$P(a \leq X \leq b) = P\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

یاداوری: جذر واریانس را انحراف معیار یا انحراف استاندارد می‌نامند و آن را با σ نشان می‌دهند.

مثال ۱۷

اگر متغیر تصادفی X دارای توزیع نرمال با میانگین ۳ و واریانس ۴ باشد، مطلوبست احتمال $P(\text{۰} \leq X \leq \text{۴})$.

راه‌حل:

$$
\begin{aligned}
P\left(\text{۰} \leq X \leq \text{۴}\right) &= P\left(\frac{\text{۰}-\text{۳}}{\text{۲}} \leq \frac{X-\mu}{\sigma} \leq \frac{\text{۴}-\text{۳}}{\text{۲}}\right) \\
&= P\left(-\text{۱/۵} \leq Z \leq \text{۰/۵}\right) \\
&= \text{۰/۶۹۱۵} - \text{۰/۰۶۶۸}
\end{aligned}
$$

مثال ۱۸

قد جوانان یک شهر دارای توزیع نرمال با میانگین ۱۶۸ سانتی‌متر و واریانس ۴۰ می‌باشد. چند درصد از جوانان این شهر قدی بین ۱۶۰ تا ۱۷۶ سانتی‌متر دارند؟

راه‌حل:

X: قد جوانان $\qquad X \sim N(۱۶۸\ ,\ ۴۰)$

$$\begin{aligned} P\left(۱۶۰ < X < ۱۷۶\right) &= P\left(\frac{۱۶۰ - ۱۶۸}{\sqrt{۴۰}} < \frac{X - \mu}{\sigma} < \frac{۱۷۶ - ۱۶۸}{\sqrt{۴۰}}\right) \\ &= P\left(-۱/۲۷ < Z < ۱/۲۷\right) \\ &= ۰/۸۹۸۰ - ۰/۱۰۲۰ = ۰/۷۹۶ \times ۱۰۰ = ۷۹/۶\,\% \end{aligned}$$

مثال ۱۹

توزیع نمره‌های پایان ترم درس آمار و احتمال تقریباً $N(۱۴, ۳)$ است. چند درصد از دانشجویان از این درس نمره‌ی قبولی نخواهند گرفت؟

راه‌حل:

X: نمره دانشجویان $\qquad X \sim N(۱۴\,,\ ۳)$

$$P\left(X < ۱۰\right) = P\left(\frac{X-\mu}{\sigma} < \frac{۱۰ - ۱۴}{\sqrt{۳}}\right)$$
$$= P\left(Z < -۲/۳۱\right)$$
$$= ۰/۰۱۰۴ \quad \times ۱۰۰ = ۱/۰۴\,\%$$

مثال ۲۰

فرض کنید متغیر تصادفی Z دارای توزیع نرمال استاندارد باشد. مقدار z را تعیین کنید.

الف– $P\left(Z < z\right) = ۰/۹ \qquad \Rightarrow \qquad z_{۰/۹} = ۱/۲۸$

ب– $P\left(Z > z\right) = ۰/۹ \qquad \Rightarrow \qquad P\left(Z \leq z\right) = ۰/۱ \qquad \Rightarrow \qquad z_{۰/۱} = -۱/۲۸$

ج– $P\left(-z < Z < z\right) = ۰/۶۸$

$$\begin{aligned} P\left(-z < Z < z\right) &= P\left(Z < z\right) - P\left(Z < -z\right) \\ &= P\left(Z < z\right) - \left[۱ - P\left(Z < z\right)\right] \\ &= ۲P\left(Z < z\right) - ۱ = ۰/۶۸ \\ &\Rightarrow P\left(Z < z\right) = ۰/۸۴ \qquad \Rightarrow \qquad z_{۰/۸۴} = ۰/۹۹ \end{aligned}$$

مثال ۲۱

در یک آزمون بزرگ میانگین نمرات ۶۰ و انحراف معیار ۲۰ با توزیع نرمال است. اگر ۱۰ درصد از شرکت‌کنندگان بتوانند نمره قبولی بگیرند، حداقل نمره قبولی چقدر خواهد بود؟

راه‌حل:

$$
\begin{aligned}
P(X > a) = ۰/۱ \quad & \Rightarrow \quad P\left(\frac{X-\mu}{\sigma} > \frac{a-۶۰}{۲۰}\right) = ۰/۱ \\
& \Rightarrow \quad ۱ - P\left(Z > \frac{a-۶۰}{۲۰}\right) = ۱ - ۰/۱ \\
& \Rightarrow \quad P\left(Z \leq \frac{a-۶۰}{۲۰}\right) = ۰/۹ \\
& \Rightarrow \quad \frac{a-۶۰}{۲۰} = ۱/۲۸ \\
& \Rightarrow \quad a = ۸۵/۶
\end{aligned}
$$

مثال ۲۲

یک ماشین سازنده نوشابه طوری ساخته شده است که در هر لیوان به طور متوسط ۲۰۷ میلی‌لیتر نوشابه می‌ریزد. اگر مقدار نوشابه ریخته شده به‌وسیله دستگاه دارای توزیع نرمال با انحراف معیار ۱۵ میلی‌لیتر باشد.

الف- چند درصد لیوان‌ها شامل بیش از ۲۳۱ میلی‌لیتر نوشابه هستند؟

ب- ۲۵ درصد از لیوان‌ها پایین‌تر از چه مقداری از نوشابه را خواهند داشت؟

راه‌حل:

X: مقدار نوشابه‌ی ریخته شده داخل لیوان $\qquad X \sim N(\mu = ۲۰۷\,,\ \sigma = ۱۵)$

$$\text{الف-}\quad P(X > ۲۳۱) = P\left(\frac{X-\mu}{\sigma} > \frac{۲۳۱ - ۲۰۷}{۱۵}\right)$$
$$= P(Z > ۱/۶) = ۱ - P(Z \leq ۱/۶)$$
$$= ۱ - ۰/۹۴۵۲ = ۰/۰۵۴۸ \quad \times ۱۰۰ = ۵/۴۸\ \%$$

$$\text{ب-}\quad P(X < a) = ۰/۲۵ \qquad \Rightarrow \qquad P\left(\frac{X-\mu}{\sigma} < \frac{a - ۲۰۷}{۱۵}\right) = ۰/۲۵$$
$$\Rightarrow \qquad \frac{a-۲۰۷}{۱۵} = -۰/۶۷ \qquad \Rightarrow \qquad a = ۱۹۶/۹۵$$

مثال ۲۳

وقتی میله‌های پلاستیکی از قالب در می‌آیند، به طور خودکار به طول‌های ظاهری ۶ اینچ بریده می‌شوند. طول‌های واقعی به طور نرمال با میانگین ۶ اینچ و انحراف معیار ۰/۰۶ توزیع شده‌اند. اگر ۵۰ میله بریده شود، چه تعدادی از آن‌ها دارای طولی بیش از ۵/۸۵ اینچ هستند؟

راه‌حل:

X: طول میله‌های بریده شده $\qquad X \sim N(\mu = ۶\ ,\ \sigma = ۰/۰۶)$

$$E(X) = np = ۵۰ \times ۰/۹۹۳۸ = ۴۹/۶۹$$

$$\begin{aligned} p = P\left(X > ۵/۸۵\right) &= P\left(\frac{X-\mu}{\sigma} > \frac{۵/۸۵ - ۶}{۰/۰۶}\right) \\ &= P\left(Z > -۲/۵\right) = ۱ - P\left(Z \leq -۲/۵\right) \\ &= ۱ - ۰/۰۰۶۲ = ۰/۹۹۳۸ \end{aligned}$$

تقریب توزیع دوجمله‌ای بوسیله‌ی توزیع نرمال

اگر $X \sim Bin(n,p)$ با میانگین $E(X)=np$ و واریانس $Var(X)=np(۱-p)$ باشد،
زمانی که n بسیار بزرگ $(n \geq ۳۰)$ و p (احتمال پیروزی) به $۰/۵$ نزدیک باشد،
به طوری که $np > ۵$ **و** $n(۱-p) > ۵$ باشد، آنگاه

$$Z = \frac{X - np}{\sqrt{np(۱-p)}} \sim N(۰, ۱)$$

از طرفی چون یک توزیع گسسته را با یک توزیع پیوسته تقریب می‌زنیم، لازم است از تصحیح پیوستگی استفاده کنیم:

$$P(X = k) = P(k - ۰/۵ < X < k + ۰/۵)$$
$$P(X \leq k) = P(X < k + ۰/۵)$$
$$P(X \geq k) = P(X > k - ۰/۵)$$

مثال ۲۴

در جامعه‌ای ۵۶٪ از رأی‌دهندگان مرد هستند. احتمال اینکه از ۵۰ نفر رأی‌دهنده حداقل ۳۰ نفر مرد باشند چقدر است؟

راه‌حل:

X: تعداد رأی‌دهندگان مرد در بین ۵۰ نفر $\qquad X \sim Bin(n = ۵۰ \ , \ p = ۰/۵۶)$

$$\mu = E(X) = np = ۵۰ \times ۰/۵۶ = ۲۸$$

$$\sigma^۲ = Var(X) = np(۱ - p) = ۵۰ \times ۰/۵۶ \times ۰/۴۴ = ۱۲/۳۲$$

$$P\left(X \geq ۳۰\right) = P\left(\frac{X - np}{\sqrt{np(۱ - p)}} > \frac{۳۰ - ۰/۵ - ۲۸}{\sqrt{۱۲/۳۲}}\right)$$

$$= P\left(Z > ۰/۴۳\right) = ۱ - P\left(Z \leq ۰/۴۳\right)$$

$$= ۱ - ۰/۶۶۶۴$$

مثال ۲۵

در شهری ۳۰ درصد از افراد بالغ جامعه سیگاری هستند. مطلوب است احتمال اینکه در نمونه‌ای به اندازه ۱۰۰۰ فرد بالغ، کمتر از ۲۸۰ فرد سیگاری وجود داشته باشد.

راه‌حل:

X: تعداد افراد سیگاری در بین ۱۰۰۰ نفر $\qquad X \sim Bin(n = 1000\ ,\ p = 0/3)$

$$\mu = E(X) = np = 1000 \times 0/3 = 300$$
$$\sigma^2 = Var(X) = np(1-p) = 1000 \times 0/3 \times 0/7 = 210$$
$$\begin{aligned} P(X < 280) &= P(X \leq 279) \\ &= P\left(\frac{X - np}{\sqrt{np(1-p)}} < \frac{279 + 0/5 - 300}{\sqrt{210}}\right) \\ &= P(Z < -1/42) \\ &= 0/0778 \end{aligned}$$

مثال ۲۶

درصد افرادی که با قرار گرفتن در معرض یک نوع باکتری به بیماری مبتلا می‌شوند، ۲۰٪ است. فرض کنید ۵۰۰ نفر در معرض باکتری قرار گرفته‌اند. احتمال این‌که ۱۲۰ نفر از آن‌ها به بیماری مبتلا شوند، چه‌قدر است؟

راه‌حل:

X: تعداد افرادی که به بیماری مبتلا می‌شوند، در بین ۵۰۰ نفر $\qquad X \sim Bin(n = ۵۰۰\ ,\ p = ۰/۲)$

$$\begin{aligned}
\mu &= E(X) = np = ۵۰۰ \times ۰/۲ = ۱۰۰ \\
\sigma^۲ &= Var(X) = np(۱-p) = ۵۰۰ \times ۰/۲ \times ۰/۸ = ۸۰ \\
P\left(X = ۱۲۰\right) &= P\left(۱۲۰ - ۰/۵ < X < ۱۲۰ + ۰/۵\right) \\
&= P\left(\frac{۱۱۹/۵ - ۱۰۰}{\sqrt{۸۰}} < \frac{X - np}{\sqrt{np(۱-p)}} < \frac{۱۲۰/۵ - ۱۰۰}{\sqrt{۸۰}}\right) \\
&= P\left(۲/۱۸ < Z < ۲/۲۹\right) \\
&= ۰/۹۸۹۰ - ۰/۹۸۵۴
\end{aligned}$$

توزیع مجموع متغیرهای تصادفی نرمال

اگر $X_۱, X_۲, \ldots, X_n$ متغیرهای تصادفی مستقل باشند و $X_i \sim N(\mu_i, \sigma_i^۲)$ و قرار دهیم

$$Y = a_۱X_۱ + a_۲X_۲ + \ldots + a_nX_n = \sum_{i=۱}^{n} a_iX_i$$

در این صورت

$$Y \sim N\left(\sum_{i=۱}^{n} a_i\mu_i, \sum_{i=۱}^{n} a_i^۲\sigma_i^۲\right)$$

مثال ۲۷

یک دستگاه برش طولی و عرضی صفحات فلزی، صفحات مستطیل شکلی را تولید می‌کند. برش‌های طولی دارای توزیع نرمال با میانگین ۲ و واریانس ۴ و برش‌های عرضی دارای توزیع نرمال با میانگین ۲ و واریانس ۵ است. اگر این برش‌ها از هم مستقل باشند، احتمال اینکه محیط یک صفحه بریده شده از ۹/۲ کمتر باشد، چقدر است؟

راه‌حل:

$$X_۱ \sim N(۲,۴) \qquad , \qquad X_۲ \sim N(۲,۵)$$

$$Y = ۲X_۱ + ۲X_۲ \sim N(۸,۳۶)$$

$$E(Y) = E(۲X_۱ + ۲X_۲) = ۲E(X_۱) + ۲E(X_۲) = (۲\times ۲) + (۲ \times ۲) = ۸$$

$$Var(Y) = Var(۲X_۱ + ۲X_۲) = ۴Var(X_۱) + ۴Var(X_۲) = (۴\times ۴) + (۴\times ۵) = ۳۶$$

$$P(\text{محیط} < ۹/۲) = P(Y < ۹/۲) = P\left(\frac{Y-\mu}{\sigma} < \frac{۹/۲ - ۸}{\sqrt{۳۶}}\right) = P(Z < ۰/۲) = ۰/۵۷۹۳$$

مثال ۲۸

نمرات درس‌های ۴ واحدی ریاضی، ۳ واحدی آمار و ۲ واحدی زبان به ترتیب دارای توزیع نرمال با میانگین‌های ۱۲، ۱۴ و ۱۵ و انحراف معیارهای ۲، $\sqrt{۲}$ و $\sqrt{۲}$ هستند.
الف- اگر ۵۰ نفر این سه درس را انتخاب کرده باشند، معدل چند نفرشان در این سه درس از ۱۴/۵ بیشتر است؟

راه‌حل:

$$X_۱ \sim N(۱۲, \sigma_۱ = ۲) \quad , \quad X_۲ \sim N(۱۴, \sigma_۲ = \sqrt{۲}) \quad , \quad X_۳ \sim N(۱۵, \sigma_۳ = \sqrt{۲})$$

$$Y = \frac{۴X_۱ + ۳X_۲ + ۲X_۳}{۹} \sim N\left(\frac{۴۰}{۳}, \frac{۱۰}{۹}\right)$$

$$E(Y) = E\left(\frac{۴X_۱ + ۳X_۲ + ۲X_۳}{۹}\right) = \frac{۴}{۹}E(X_۱) + \frac{۳}{۹}E(X_۲) + \frac{۲}{۹}E(X_۳) = \left(\frac{۴}{۹} \times ۱۲\right) + \left(\frac{۳}{۹} \times ۱۴\right) + \left(\frac{۲}{۹} \times ۱۵\right) = \frac{۴۰}{۳}$$

$$Var(Y) = Var\left(\frac{۴X_۱ + ۳X_۲ + ۲X_۳}{۹}\right) = \left(\frac{۴}{۹}\right)^۲ Var(X_۱) + \left(\frac{۳}{۹}\right)^۲ Var(X_۲) + \left(\frac{۲}{۹}\right)^۲ Var(X_۳)$$
$$= \left(\frac{۱۶}{۸۱} \times ۴\right) + \left(\frac{۱}{۹} \times ۲\right) + \left(\frac{۴}{۸۱} \times ۲\right) = \frac{۱۰}{۹}$$

$$P(\text{معدل} > ۱۴/۵) = P(Y > ۱۴/۵) = P\left(\frac{Y-\mu}{\sigma} > \frac{۱۴/۵ - ۱۳/۳۳}{\sqrt{۱/۱۱}}\right) = P(Z > ۱/۱۱) = ۱ - ۰/۸۶۶۵ = ۰/۱۳۳۵$$

$$\text{میانگین تعداد نفرات} = np = ۵۰ \times ۰/۱۳۳۵ = ۶/۶۷۵$$

مثال ۲۸

ب- اگر ۴۰ نفر از دانشجویانی که این سه درس را انتخاب کرده‌اند به طور مستقل انتخاب کنیم، احتمال اینکه حداقل ۱۰ نفر از آنها دارای معدلی بیشتر از ۱۴/۵ در این سه درس باشند را بیابید.

راه‌حل:

K: تعداد دانشجویانی که معدل‌شان بیشتر از ۱۴/۵ شده است $\qquad K \sim Bin(n = ۴۰\ ,\ p = ۰/۱۳۳۵)$

$$\mu = E(K) = np = ۴۰ \times ۰/۱۳۳۵ = ۵/۳۴$$

$$\sigma^۲ = Var(K) = npq = ۴۰ \times ۰/۱۳۳۵ \times ۰/۸۶۶۵ = ۴/۶۳$$

$$P\left(K \geq ۱۰\right) = P\left(K > ۹/۵\right) = P\left(\frac{K-\mu}{\sigma} > \frac{۹/۵ - ۵/۳۴}{\sqrt{۴/۶۳}}\right)$$

$$= P\left(Z > ۱/۹۳\right)$$

$$= ۱ - ۰/۹۷۳۲ = ۰/۰۲۶۸$$

توزیع لگ نرمال

توزیع لگ نرمال

متغیر تصادفی پیوسته X دارای توزیع لگ-نرمال است، اگر متغیر تصادفی $Y = ln(X)$ دارای توزیع نرمال با میانگین μ و انحراف معیار σ باشد.

تابع چگالی حاصل بر حسب متغیر تصادفی X به صورت زیر است:

$$f(x) = \frac{۱}{x\sqrt{۲\pi\sigma^۲}} e^{\frac{-۱}{۲}\left(\frac{ln(x)-\mu}{\sigma}\right)^۲}, \qquad x > ۰$$

میانگین و واریانس توزیع لگ-نرمال برابر است با:

$$E(X) = e^{\mu+\frac{\sigma^۲}{۲}} \qquad\qquad Var(X) = \left(e^{\sigma^۲} - ۱\right) e^{۲\mu+\sigma^۲}$$

- توزیع لگ نرمال برای مدل‌بندی قیمت سهام، نرخ ارز، اندازه‌ی بافت‌های زنده بدن (مثل سطح پوست)، اندازه‌های فیزیولوژی مثل فشار خون افراد بزرگ‌سال، ... مورد استفاده قرار می‌گیرد.

مثال ۲۹

از روی تجربه معلوم شده که غلظت آلاینده‌های تولیدی کارخانه‌ی مواد شیمیایی رفتاری از خود نشان می‌دهد که قابل توصیف با توزیع لگ نرمال است. فرض کنید غلظت آلاینده بخصوصی، بر حسب واحد در هر میلیون، دارای توزیع لگ نرمال با پارامترهای $\mu = ۳/۲$ و $\sigma = ۱$ است. احتمال این‌که غلظت آلاینده از ۸ واحد در میلیون بیشتر شود، چه‌قدر است؟

راه‌حل:

$$\begin{aligned} P(X > ۸) &= P\left(ln(X) > ln(۸)\right) \\ &= P\left(\frac{ln(X) - \mu}{\sigma} > \frac{ln(۸) - ۳/۲}{۱}\right) \\ &= P\left(Z > -۱/۱۲\right) = ۱ - ۰/۱۳۱۴ \end{aligned}$$

توزیع وایبل