4.3 连续型随机变量的期望和方差

本节研究连续型随机变量的期望和方差, 有利于了解这些变量的整体数字特征.

定义 4.3 设连续随机变量 X 的概率密度函数为 $f(x)$, 若积分 $\int_{-\infty}^{+\infty}|x|f(x)dx$ 收敛, 称 $\int_{-\infty}^{+\infty}xf(x)dx$ 为随机变量 X 的 **期望**, 记为 $E(X)$, 即

$$E(X)=\int_{-\infty}^{+\infty}tf(t)dt.$$

类似于离散型随机变量, 连续型随机变量的期望具有以下一些性质:

引理 4.4 (线性关系) 对任意任意常数 a,b 和连续随机变量 X, 有 $E(aX+b)=aE(X)+b$.

引理 4.5 (Jensen 不等式) 对连续随机变量 X 和函数 $g(x)$,

- 若 $g(x)$ 是凸函数, 则有 $g(E(X))\leqslant E[g(X)]$;
- 若 $g(x)$ 是凹函数, 则有 $g(E(X))\geqslant E[g(X)]$.

对于非负的连续型随机变量, 也可以利用 $P(X>t)$ 来直接计算期望:

引理 4.6 若连续型随机变量 $X\geqslant 0$, 则有

$$E[X]=\int_0^{\infty}P(X>t)dt.$$

该定理对随机变量函数 $Y=g(X)\geqslant 0$ 也成立, 即 $E[g(X)]=\int_0^{\infty}P(g(X)>t)dt$.

证明 设随机变量 X 的概率密度为 $f(x)$, 首先观察得到

$$X=\int_0^X 1dt=\int_0^{+\infty}\mathbb{I}[t<X]dt=\int_0^{+\infty}\mathbb{I}[X>t]dt,$$

这里 $\mathbb{I}[\cdot]$ 表示指示函数, 如果论断为真, 其值为 1, 否则为 0. 两边同时取期望有

$$\begin{aligned}
E[X] &= E\left[\int_0^{+\infty}\mathbb{I}[X>t]dt\right] \\
&= \int_0^{+\infty}\left[\int_0^{+\infty}\mathbb{I}[x>t]f(x)dt\right]dx \qquad (\text{积分换序}) \\
&= \int_0^{+\infty}\left[\int_0^{+\infty}\mathbb{I}[x>t]f(x)dx\right]dt \\
&= \int_0^{+\infty}\left[\int_0^{t}\mathbb{I}[x>t]f(x)dx+\int_t^{+\infty}\mathbb{I}[x>t]f(x)dx\right]dt \\
&= \int_0^{+\infty}\left[\int_t^{+\infty}f(x)dx\right]dt=\int_0^{+\infty}P(X>t)dt.
\end{aligned}$$

对于连续随机变量函数的期望有

定理 4.2 设随机变量 X 的密度函数为 $f(x)$, 且 $\int_{-\infty}^{+\infty} g(t)f(t)dt$ 绝对可积, 则

$$E(g(X)) = \int_{-\infty}^{+\infty} g(t)f(t)dt.$$

该定理表明, 若已知随机变量 X 的密度函数为 $f(x)$, 以及随机变量函数 $Y = g(X)$, 可以直接利用随机变量 X 的密度函数来计算 Y 的期望, 而不需要知道随机变量 Y 的密度函数.

证明 该定理对一般的可积函数 $g(x)$ 均成立, 但证明过程却非常复杂, 这里仅给出非负随机变量函数 $g(x) \geqslant 0$ 的证明. 根据引理 4.6 有

$$\begin{aligned} E[g(X)] &= \int_0^{+\infty} P(g(X) \geqslant t)dt = \int_0^{+\infty} \int_{x:\ g(x)\geqslant t} f(x)dxdt \\ &= \int_{x:\ g(x)\geqslant 0} \int_0^{g(x)} f(x)dtdx = \int_{x:\ g(x)\geqslant 0} g(x)f(x)dx = \int_{-\infty}^{+\infty} g(x)f(x)dx\ , \end{aligned}$$

由此完成证明.

下面介绍物理学中用到的柯西分布 (Cauchy distribution), 它的期望不存在.

例 4.7 设连续随机变量 X 的密度函数为 $f(x) = 1/\pi(1+x^2)$ $(x \in \mathbb{R})$, 求期望 $E(X)$.

因为积分

$$\int_{-\infty}^{+\infty} \frac{|x|}{\pi(1+x^2)}dx = 2\int_0^{+\infty} \frac{x}{\pi(1+x^2)}dx = \frac{1}{\pi}\left[\ln(1+x^2)\right]_0^{+\infty} = +\infty,$$

由此可知期望 $E(X)$ 不存在.

例 4.8 古人运送粮草, 如果早到每天需要的存储费用 c 元, 如果晚到每天需要的延期费用为 C 元. 粮草在运送过程中存在天气、路况等不确定因素, 因此运送需要的天数是随机的, 其概率密度函数为 $f(x)$, 问什么时候出发才能使费用的期望值最小?

解 用随机变量 X 表示实际的运送天数, 分布函数为 $F(x)$. 不妨假设提前了 t 天出发, 那么所需费用为

$$\ell_t(X) = \begin{cases} c(t-X) & X \leqslant t, \\ C(X-t) & X > t. \end{cases}$$

因此可得

$$\begin{aligned} E[\ell_t(X)] &= \int_0^{+\infty} \ell_t(x)f(x)dx = \int_0^t c(t-x)f(x)dx + \int_t^{+\infty} C(x-t)f(x)dx \\ &= ctF(t) - c\int_0^t xf(x)dx + C\int_t^{+\infty} xf(x)dx - Ct(1-F(t))\ . \end{aligned}$$

对上式中的 t 求导、并令导数为零可得

$$\frac{d}{dt}E[\ell_t(X)] = cF(t) - C(1-F(t)) = (c+C)F(t) - C \ .$$

求解可得期望最小的天数 t^* 满足

$$F(t^*) = C/(C+c) \ .$$

定义 4.4 设连续随机变量 X 的概率密度为 $f(x)$, 若 $\int_{-\infty}^{+\infty}(t-E(X))^2 f(t)dt$ 收敛, 称为随机变量 X 的 **方差**, 记为 $\mathrm{Var}(X)$, 即

$$\mathrm{Var}(X) = E(X-E(X))^2 = \int_{-\infty}^{+\infty}(t-E(X))^2 f(t)dt \ .$$

其等价性定义为

$$\mathrm{Var}(X) = E(X-E(X))^2 = E(X^2) - (E(X))^2 = \int_{-\infty}^{+\infty} t^2 f(t)dt - \left(\int_{-\infty}^{+\infty} tf(t)dt\right)^2 .$$

类似于离散型随机变量, 连续型随机变量的方差具有如下性质:

- 对连续型随机变量 X 和常数 a,b, 有 $\mathrm{Var}(aX+b) = a^2\mathrm{Var}(X)$;
- 对连续型随机变量 X 和常数 a, 有 $\mathrm{Var}(X) = E(X-E(X))^2 \leqslant E(X-a)^2$;
- 对连续型随机变量 $X \in [a,b]$, 有 $\mathrm{Var}(X) = (b-E(X))(E(X)-1) \leqslant (b-a)^2/4$.

4.4 常用连续型随机变量

下面介绍几种常用的连续型随机变量.

4.4.1 均匀分布(uniform distribution)

定义 4.5 若随机变量 X 的概率密度函数为

$$f(x) = \begin{cases} 1/(b-a) & x \in [a,b] \\ 0 & \text{其它}, \end{cases}$$

称 X 在区间 $[a,b]$ 上服从 **均匀分布**, 记 $X \sim U(a,b)$.

根据上面的定义很容易发现服从均匀分布的随机变量落入区间任何一点的概率相同. 对任意实数 $x \in \mathbb{R}$ 有 $f(x) \geqslant 0$ 且

$$\int_{-\infty}^{+\infty} f(t)dt = \int_{-\infty}^{a} f(t)dt + \int_{a}^{b} f(t)dt + \int_{b}^{+\infty} f(t)dt = \int_{a}^{b} \frac{1}{b-a}dt = 1 \ .$$

若随机变量 $X \sim U(a,b)$, 则 X 落入内任一子区间 $[x, x+\Delta]$ 的概率

$$P(x \leqslant X \leqslant x+\Delta) = \int_x^{x+\Delta} \frac{1}{b-a} dt = \frac{\Delta}{b-a} \ .$$

该概率与子区间的具体位置 x 无关, 而与子区间长度 Δ 成正比, 由此给出了均匀分布的几何解释: 若随机变量 $X \ \sim U(a,b)$, 则 X 落入 $[a,b]$ 内任一子区间的概率与该区间的长度成正比, 与位置无关.

根据分布函数的定义可知 $X \sim U(a,b)$ 的分布函数为

$$F(x) = \begin{cases} 0 & x \leqslant a \\ (x-a)/(b-a) & a < x < b \\ 1 & x \geqslant b \end{cases}$$

随机变量 $X \sim U(a,b)$ 的密度函数和分布函数的示意图如下:

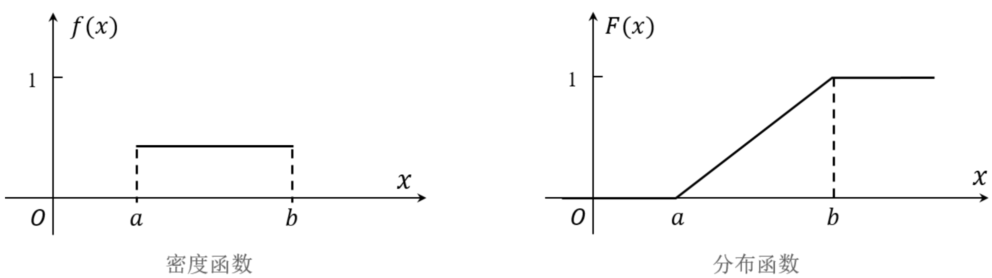


图 4.4 随机变量 $X \sim U(a,b)$ 的密度函数和分布函数

定理 4.3 若随机变量 $X \sim U(a,b)$, 则有

$$E(X) = (a+b)/2 \qquad \text{和} \qquad \mathrm{Var}(X) = (b-a)^2/12 \ .$$

证明 根据期望的定义有

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} tf(t)dt = \frac{1}{b-a}\int_a^b tdt = \frac{a+b}{2} \ , \\ E(X^2) &= \int_{-\infty}^{+\infty} t^2 f(t)dt = \frac{1}{b-a}\int_a^b t^2 dt = \frac{a^2+ab+b^2}{3} \ , \end{aligned}$$

从而得到方差

$$\mathrm{Var}(X) = E(X^2) - (E(X))^2 = \frac{a^2+ab+b^2}{3} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12} \ .$$

例 4.9 已知随机变量 $X \sim U(a,b)$, 对 $a<c<d<b$, 求 $P(X \leqslant c|X \leqslant d]$.

解 根据条件概率的定义有

$$P(X \leqslant c|X \leqslant d] = \frac{P(\{X \leqslant d\} \cap \{X \leqslant c\})}{P(X \leqslant d)} = \frac{P(X \leqslant c)}{P(X \leqslant d)} = \frac{c-a}{d-a} \ ,$$

即在 $X \leqslant d$ 的条件下, 随机变量 X 服从 $U(a,d)$.

例 4.10 设随机变量 $\xi \sim U(-3,6)$, 试求方程 $4x^2+4\xi x+(\xi+2)=0$ 有实根的概率.

解 易知随机变量 ξ 的概率密度函数

$$f(t) = \begin{cases} 1/9 & x \in [-3,6] \\ 0 & \text{其它}. \end{cases}$$

设事件 A 表示方程有实根, 于是有

$$\begin{aligned} P(A) &= P((4\xi)^2 - 4\times 4\times(\xi+2) \geqslant 0) = P((\xi+1)(\xi-2) \geqslant 0) \\ &= P\left(\{\xi \geqslant -1\} \cap \{\xi \geqslant 2\} \geqslant 0\right) + P\left(\{\xi \leqslant -1\} \cap \{\xi \leqslant 2\} \geqslant 0\right) \\ &= P(\xi \leqslant -1) + P(\xi \geqslant 2) = \int_{-3}^{-1} \frac{1}{9} dt + \int_{2}^{6} \frac{1}{9} dt = \frac{2}{3}. \end{aligned}$$

4.4.2 指数分布

指数分布常用于电话的通话时间和银行的服务等待时间, 也可以用于描述动物和电子元件的寿命, 在可靠性理论和排队论中具有广泛的应用.

定义 4.6 给定常数 $\lambda > 0$, 若随机变量 X 的密度函数

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geqslant 0 \\ 0 & \text{其它}, \end{cases}$$

称 X 服从 **参数为 λ 的指数分布**, 记 $X \sim e(\lambda)$.

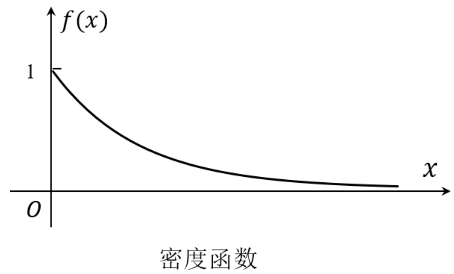
对任意实数 x 有密度函数 $f(x) \geqslant 0$, 以及

$$\int_{-\infty}^{+\infty} f(t) dt = \int_{0}^{+\infty} \lambda e^{-\lambda t} dt = \left[-e^{-\lambda t}\right]_{0}^{+\infty} = 1.$$

根据指数分布的密度函数很容易得到分布函数, 即当 $x \leqslant 0$ 时, 分布函数 $F(x)=0$; 当 $x>0$ 时, 分布函数

$$F(x) = \int_{0}^{x} \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}.$$

指数分布的密度函数和分布函数如图 4.5 所示.



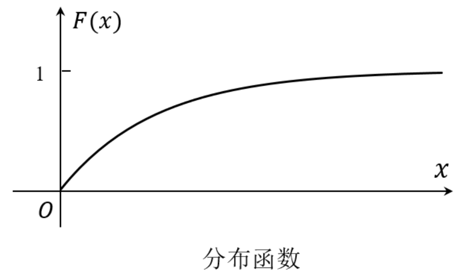


图 4.5 指数分布的密度函数和分布函数

引理 4.7 若随机变量 $X \sim e(\lambda)$, 则 $E(X) = 1/\lambda$ 和 $\mathrm{Var}(X) = 1/\lambda^2$.

证明 根据连续函数的定义有

$$E(X) = \int_0^{+\infty} t\lambda e^{-\lambda t} dt = \left[-te^{-\lambda t}\right]_0^{+\infty} + \int_0^{+\infty} e^{-\lambda t} dt = -\frac{1}{\lambda}\left[e^{-\lambda t}\right]_0^{+\infty} = \frac{1}{\lambda} ,$$

对非负的随机变量 $X \geqslant 0$, 可以利用引理 4.6 和分布函数 $F(x) = P(X \leqslant x) = 1 - e^{-\lambda x}$ 有

$$E[X] = \int_0^{+\infty} P(X > t)dt = \int_0^{+\infty} 1 - F(t)dt = \int_0^{+\infty} e^{-\lambda t}dt = \frac{1}{\lambda} .$$

对于方差, 首先计算

$$E(X^2) = \lambda \int_0^{+\infty} t^2 e^{-\lambda t} dt = \left[-t^2 e^{-\lambda t}\right]_0^{+\infty} + \int_0^{+\infty} 2te^{-\lambda t} dt = \frac{2}{\lambda}E(X) = \frac{2}{\lambda^2} ,$$

于是得到 $\mathrm{Var}(X) = E(X^2) - [E(X)]^2 = 1/\lambda^2$.

下面研究指数分布的一个重要性质: 指数分布的无记忆性.

定理 4.4 若随机变量 $X \sim e(\lambda)$, 则对任意 $s > 0, t > 0$, 有

$$P(X > s + t | X > t) = P(X > s).$$

证明 对任意 $x > 0$, 根据分布函数有 $P(X > x) = 1 - F(x) = e^{-\lambda x}$, 从而得到

$$P(X > s + t | X > t) = \frac{P(\{X > s + t\} \cap \{X > t\})}{P(X > t)} = \frac{P(X > s + t)}{P(X > t)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = P(X > s) ,$$

定理得证.

指数分布是唯一具有无记忆性的连续型随机变量, 例如一盏灯泡的寿命 X 服从指数分布, 若已经使用了 s 个小时, 则再使用 t 个小时的概率与已使用过 s 个小时无关, 将这个经历给“忘记”了.

引理 4.8 若随机变量 $X_1, X_2, \cdots, X_n$ 是相互独立的、且分别服从参数为 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 的指数分布, 则有

$$X = \min\{X_1, X_2, \cdots, X_n\} \sim e(\lambda_1 + \lambda_2 + \cdots + \lambda_n) \ .$$

解 这里随机变量的相互独立性可以理解为随机变量取不同值的随机事件相互独立. 计算随机变量 X 的分布函数

$$\begin{aligned} F_X(x) &= P(X \leqslant x) = 1 - P(\min(X_1, X_2, \cdots, X_n) > x) \\ &= 1 - \prod_{i=1}^n P(X_i > x) = 1 - \prod_{i=1}^n \exp(-\lambda_i x) = 1 - \exp\left(-x \sum_{i=1}^n \lambda_i\right) \ , \end{aligned}$$

由此完成证明.

4.4.3 正态分布

正太分布是概率统计中最重要的一种分布, 最早由法国数学家棣莫弗 (De Moivre, 1667-1754) 在 1730s 提出, 用于近似抛硬币试验中随机事件的概率, 即中心极限定理的雏形. 德国数学家高斯 (Gauss, 1777-1855) 在 1800s 首次将正太分布应用于预测天文学中星体的位置, 由此才展示出正太分布的应用价值, 后来发现很多随机现象可以通过正太分布来描述, 正太分布因此被称为高斯分布.

正太分布的重要性主要体现在以下三个方面:

- 现实生活中很多随机现象可用正太分布进行描述, 例如人的身高或体重, 某地区的降雨量等;
- 很多分布可以通过正太分布来进行近似计算 (如后面第 X 章的中心极限定理);
- 数理统计中常用的统计分布都是由正太分布导出的, 如 χ^2 分布、t 分布和 F 分布.

定义 4.7 给定任何实数 u 和 $\sigma > 0$, 若随机变量 X 的概率密度为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad x \in (-\infty, +\infty) \ ,$$

称随机变量 X 服从 **参数为 (μ, σ^2) 的正态分布** (Normal distribution), 又称为 **高斯分布** (Gaussian distribution), 记为 $X \sim \mathcal{N}(\mu, \sigma^2)$.

特别地, 当 $\mu = 0$ 和 $\sigma = 1$ 时的正太分布 $\mathcal{N}(0, 1)$ 被称为 **标准正态分布**, 此时密度函数为

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \qquad x \in (-\infty, +\infty) \ .$$

对任意 $x \in (-\infty, +\infty)$ 有 $f(x) \geqslant 0$, 利用极坐标变换 $(x = r\cos\theta, y = r\sin\theta)$ 有

$$\left(\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx\right)^2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-\frac{x^2+y^2}{2}} dx dy$$