

第四章 随机变量的数字特征

数学期望 (expectation)

数学期望

方差

协方差、相关系数

其它数字特征

设连续型随机变量 X 的概率密度为 $f(x)$, 若积分 $\int_{-\infty}^{+\infty} x f(x) dx$ 绝对收敛 (即 $\int_{-\infty}^{+\infty} |x| f(x) dx < +\infty$), 则称积分 $\int_{-\infty}^{+\infty} x f(x) dx$ 的值为随机变量 X 的数学期望, 记为 $E(X)$

$$E(X) = \int_{-\infty}^{+\infty} x f(x) dx$$

设 Z 是随机变量 X, Y 的函数, $Z = h(X, Y)$ (h 是连续函数),

若二元连续型随机变量 (X, Y) 的概率密度为 $f(x, y)$, 则有 $E(Z) = E(h(X, Y)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x, y) f(x, y) dx dy$

其中积分绝对收敛

$$\text{特别地, } \begin{cases} E(X) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x f(x, y) dx dy \\ E(Y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f(x, y) dx dy \end{cases}$$

数学期望的特性

$$\begin{cases} E(C_0 + \sum_{i=1}^n C_i X_i) = C_0 + \sum_{i=1}^n C_i E(X_i) \\ E(\sum_{i=1}^n X_i) = \sum_{i=1}^n E(X_i), \text{ 其中 } X_i, i=1, \dots, n \text{ 相互独立} \end{cases}$$

方差

设 X 是一个随机变量, 若 $E((X - E(X))^2)$ 存在, 则称其为 X 的方差, 记为 $D(X)$ 或 $\text{Var}(X)$

$$D(X) = \text{Var}(X) = E((X - E(X))^2) = E(X^2) - E^2(X)$$

将 $\sqrt{D(X)}$ 记为 $\sigma(X)$, 称为 X 的标准差或均方差

对于离散型随机变量 X , 其分布律为 $P(X = x_i) = p_i, i = 1, 2, \dots$

$$D(X) = \sum_{i=1}^{\infty} [x_i - E(X)]^2 p_i$$

对于连续型随机变量 X , 其概率密度为 $f(x)$

$$D(X) = \int_{-\infty}^{+\infty} (x - E(X))^2 f(x) dx$$

[EX] 设随机变量 X 具有 $0-1$ 分布, 其分布律为: $P(X=0) = 1-p, P(X=1) = p$, 求 $D(X)$

$$E(X) = p$$

$$E(X^2) = p$$

$$\text{则 } D(X) = E(X^2) - E^2(X) = p - p^2$$

[EX] 设 $X \sim \pi(\lambda)$, 求 $D(X)$

$$P(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}, k=0, 1, \dots, \lambda > 0$$

$$E(X) = \lambda, E(X^2) = \lambda^2 + \lambda$$

$$D(X) = E(X^2) - E^2(X) = \lambda$$

$$\begin{aligned} E(X) &= \sum_{k=1}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} \cdot k \\ &= e^{-\lambda} \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = e^{-\lambda} \cdot \lambda \cdot e^{\lambda} \\ &= \lambda \end{aligned}$$

方差的性质

$$\begin{cases} D(C_0 + \sum_{i=1}^n C_i X_i) = \sum_{i=1}^n C_i^2 D(X_i) \\ D(X) = 0 \Leftrightarrow P(X=C)=1 \text{ 且 } C=E(X) \end{cases}$$

$$\text{标准化变量 } X^* = \frac{X - \mu}{\sigma}$$

$$\begin{cases} E(X^*) = 0 \\ D(X^*) = 1 \end{cases}$$

表1-几种常见分布的均值与方差

分布	分布律或密度函数	数学期望	方差
0-1分布	$P(X=k) = p^k (1-p)^{1-k}, k=0,1$	p	$p(1-p)$
二项分布 $B(n, p)$	$P(X=k) = C_n^k p^k (1-p)^{n-k}, k=0,1,\dots,n$	np	$np(1-p)$
泊松分布 $\pi(\lambda)$	$P(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}, k=0,1,\dots$	λ	λ
均匀分布 $U(a, b)$	$f(x) = \begin{cases} 1/(b-a), & a < x < b \\ 0, & \text{其它} \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
指数分布 $Exp(\lambda)$	$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x > 0 \\ 0, & \text{其它} \end{cases}$	$1/\lambda$	$1/\lambda^2$
正态分布 $N(\mu, \sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < +\infty$	μ	σ^2

协方差与相关系数

$E\{[X-E(X)][Y-E(Y)]\}$ 称为随机变量 X 与 Y 的协方差. 记为 $Cov(X, Y)$, 即

$$Cov(X, Y) = E\{[X-E(X)][Y-E(Y)]\}$$

称 $\rho_{XY} = \frac{Cov(X, Y)}{\sqrt{D(X)D(Y)}}$ 为随机变量 X 与 Y 的相关系数

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

$$D\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n D(X_i) + 2 \sum_{1 \leq i < j \leq n} Cov(X_i, X_j)$$

$$\Rightarrow D(X+Y) = D(X) + D(Y) + 2Cov(X, Y)$$

协方差的性质

$$\begin{cases} Cov(X, Y) = Cov(Y, X) \\ Cov(X, X) = D(X) \\ Cov(aX, bY) = abCov(X, Y), a, b \in \mathbb{R} \\ Cov(X_1 + X_2, Y) = Cov(X_1, Y) + Cov(X_2, Y) \end{cases}$$

相关系数的性质

$$\begin{cases} |\rho_{XY}| \leq 1 \\ |\rho_{XY}| = 1 \Leftrightarrow \text{存在常数 } a, b, \text{ s.t. } P(Y = a + bX) = 1 \end{cases}$$

特别的, $\rho_{XY} = 1$ 时, $b > 0$; $\rho_{XY} = -1$ 时, $b < 0$

当 $\rho_{XY} = 0$, 称 X 与 Y 不相关或零相关

$$\Leftrightarrow \begin{cases} Cov(X, Y) = 0 \\ E(XY) = E(X)E(Y) \\ D(X+Y) = D(X) + D(Y) \end{cases}$$

二元正态变量

1. 二元随机变量 (X, Y) 服从二元正态分布 $\Leftrightarrow$
 X, Y 的任意线性组合 $aX + bY$ 服从一元正态分布
其中 a, b 不全为零的常数,
且 $aX + bY \sim N(E(aX + bY), D(aX + bY))$

2. 若 (X, Y) 服从二元正态分布, 设 Z_1, Z_2 是 X, Y 的线性函数, 则
 (Z_1, Z_2) 也服从二元正态分布;
这一性质称为正态变量的线性变换不变性