Ch 6.4 条件期望

协方差

X和Y协方差$\mathrm{Cov}(X,Y)=E\big[\big(X-E(X)\big)\big(Y-E(Y)\big)\big]=E(XY)-E(X)E(Y)$

- $\mathrm{Cov}(X,c)=0$, $\mathrm{Cov}(X,X)=\mathrm{Var}(X)$, $\mathrm{Cov}(X,Y)=\mathrm{Cov}(Y,X)$
- $\mathrm{Var}(X\pm Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)\pm 2\mathrm{Cov}(X,Y)$
- $\mathrm{Cov}(aX,bY)=ab\mathrm{Cov}(X,Y)$
- $\mathrm{Cov}(X+a,Y+b)=\mathrm{Cov}(X,Y)$
- $\mathrm{Cov}(X_1+X_2,Y)=\mathrm{Cov}(X_1,Y)+\mathrm{Cov}(X_2,Y)$
- $\big(\mathrm{Cov}(X,Y)\big)^2\le \mathrm{Var}(X)\mathrm{Var}(Y)$

若随机向量$(X,Y)\sim \mathrm{N}(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho)$, 则$\mathrm{Cov}(X,Y)=\rho\sigma_x\sigma_y$

相关系数

随机变量X和Y的**相关系数**：$\rho_{XY} = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$

- 根据协方差性质可知$|\rho_{XY}| \le 1$
- $|\rho_{XY}| = 1$的充要条件为X与Y几乎处处有线性关系$Y = aX + b$
- 若X与Y相互独立, 则X与Y不相关 ($\rho_{XY} = 0$), 但反之不成立

随机向量 $(X, Y) \sim \mathrm{N}(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$, 则$X$与$Y$相关系数$\rho_{XY} = \rho$,
X与Y独立的充要条件是X与Y不相关

条件期望

前一章介绍了条件分布, 基于条件分布可以考虑条件期望, 分离散和连续性随机变量两种情况

对连续随机变量，在$Y = y$条件下X的条件密度函数$f_{X|Y}(x|y)$, 称

$$E(X|y) = E(X|Y = y) = \int_{-\infty}^{+\infty} x f_{X|Y}(x|y)dx$$

为在$Y = y$条件下X的**条件期望**

对 离 散 型 随 机 变 量 ， 在 $Y = y$ 条 件 下 X 的 条 件 分 布 列 为 $P(X = x_i|Y = j)$, 称

$$E(X|y) = E(X|Y = y) = \sum_i x_i P(X = x_i|Y = j)$$

为在$Y = y$条件下X的**条件期望**

性质

- 对任意常数a,b有$E(aX_1+bX_2|Y)=aE(X_1|Y)+bE(X_2|Y)$
- 对离散型随机变量(X,Y)和函数$g(X)$有

$$E(g(X)|Y)=\sum_i g(x_i)P(X=x_i|Y=y)$$

- 对连续型随机变量(X,Y)和函数$g(X)$有

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

设随机向量 $(X,Y)\sim \mathrm{N}(\mu_x,\mu_y,\sigma_x^2,\sigma_y^2,\rho)$，则在$Y=y$的条件下随机变量$X$服从正太分布$N(\mu_x-\rho\sigma_x(y-\mu_y)/\sigma_y,\left(1-\rho^2\right)\sigma_x^2)$,

$$E(X|y)=\mu_x-\rho\sigma_x\left(y-\mu_y\right)/\sigma_y$$

另一种方法计算期望

对二维离散随机变量(X,Y)有

$$E(X)=E_Y\big(E(X|Y)\big)=\sum_{y_j}E\left(X\middle|y_j\right)P\left(Y=y_j\right)$$

对二维连续随机变量(X,Y)有

$$E(X)=E_Y\big(E(X|Y)\big)=\int_{-\infty}^{+\infty}E(X|y)f_Y(y)dy$$

全期望公式

设$A_1, A_2, \cdots, A_n$是样本空间Ω一个分割，$A_iA_j = \emptyset$和$\Omega = \cup_{i=1\}}^{n} A_i$.
对任意随机变量X有

$$E(X) = E(X|A_1)P(A_1) + E(X|A_2)P(A_2) + \cdots + E(X|A_n)P(A_n)$$

特别地，随机事件A与其对立事件$\bar{A}$构成样本空间Ω的一个划分，
对任意随机变量X有

$$E(X) = E(X|A)P(A) + E\left(X\middle|\bar{A}\right)P\left(\bar{A}\right)$$

例题

一矿工被困在有三个门的矿井里，第一个门通一坑道，沿此坑道走3小时可使他到达安全地点；第二个门可使他走5小时后又回到原处；第三个门可使他走7小时后也回到原地. 如设此矿工在任何时刻都等可能地选定其中一个门，试问他到达安全地点平均要用多长时间?

例题

设 (X,Y)的联合概率密度为

$$f(x,y)=\begin{cases}\exp(-y) & 0<x<y<+\infty \\ 0 & \text{其它}\end{cases}$$

求条件期望$E(X|y)$.

例题

随机变量X和Y的期望、方差、相关系数分别为$\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho$，其中$\sigma_x > 0, \sigma_y > 0, \rho \in [-1, +1]$，求解最优的线性预测$Y = aX + b$使得$E\left((Y - aX - b)^2\right)$最小化

随机向量的数学期望与协方差阵

设随机向量$X=(X_1,X_2,\cdots,X_n)^{\top}$, 称

$$E(X)=\big(E(X_1),E(X_2),\cdots,E(X_n)\big)^{\top}$$

为**随机向量X的期望**, 以及称

$$\mathrm{Cov}(X)=\Sigma=\begin{pmatrix}\mathrm{Cov}(X_1,X_1) & \cdots & \mathrm{Cov}(X_1,X_n)\\ \mathrm{Cov}(X_2,X_1) & \cdots & \mathrm{Cov}(X_2,X_n)\\ \vdots & & \vdots\\ \mathrm{Cov}(X_n,X_1) & \cdots & \mathrm{Cov}(X_n,X_n)\end{pmatrix}$$

为**随机变量X的协方差矩阵**.

性质

随机向量$X=(X_1,X_2,\cdots,X_n)$的协方差矩阵是对称半正定的

设多维正态分布$X=(X_1,X_2,\ldots,X_n)^\top\sim N(\boldsymbol{\mu},\boldsymbol{\Sigma})$, 则有

$$\boldsymbol{\mu}=\big(E(X_1),E(X_2),\ldots,E(X_n)\big)^\top$$

$$\boldsymbol{\Sigma}=\left[\operatorname{Cov}(X_i,X_j)\right]_{n\times n}$$

Ch 7 集中不等式

Markov不等式

Markov不等式：对任意随机变量$X \geq 0$和$\epsilon > 0$, 有

$$P(X \geq \epsilon) \leq \frac{E(X)}{\epsilon}$$

推论：对任意随机变量X和$\epsilon \geq 0$, 及单调递增的非负函数$g(x)$, 有

$$P(X \geq \epsilon) \leq \frac{E(g(X))}{g(\epsilon)}$$

Chebyshev不等式

Chebyshev不等式： 设随机变量X的均值为μ, 则有

$$P(|X-\mu|>\epsilon)\leq\frac{\mathrm{Var}(X)}{\epsilon^2}$$

例题

设随机变量$X \sim N(-1,2)$和$Y \sim N(1,8)$，且X和Y的相关系数为-1，利用Chebyshev不等式估计

$$P(|X+Y| \geq 6) \leq ????$$

练习： 随机变量X和Y满足$E(X)=2$，$E(Y)=2$，$\mathrm{Var}(X)=1$，$\mathrm{Var}(Y)=4, \rho_{XY}=-1/2$．利用Chebyshev不等式估计$\mathrm{P}(|X-Y| \geq 6)$的上界.

单边Chebyshev不等式

单边Chebyshev不等式[Cantelli不等式]：随机变量X的均值$\mu > 0$，方差σ^2，则对任意$\epsilon > 0$有

$$P(X - \mu \geq \epsilon) \leq \frac{\sigma^2}{\sigma^2 + \epsilon^2}$$

$$P(X - \mu \leq -\epsilon) \leq \frac{\sigma^2}{\sigma^2 + \epsilon^2}$$

Chebyshev不等式推论

推论： 设独立同分布的随机变量$X_1, X_2, \ldots, X_n$满足$E(X_i) = \mu$和$\mathrm{Var}(X_i) \leq \sigma^2$, 对任意$\epsilon > 0$有

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2}$$

Young不等式

Young不等式： 给定常数$a > 0, b > 0$, 对满足$1/p + 1/q = 1$的实数$p > 0, q > 0$有

$$ab \le \frac{1}{p}a^p + \frac{1}{q}b^q.$$

Hölder不等式

Hölder不等式： 对任意随机变量X和Y以及实数$p > 0$和$q > 0$满足$1/p + 1/q = 1$, 有

$$E(|XY|) \leq \left(E(|X|^p)\right)^{\frac{1}{p}}\left(E(|Y|^q)\right)^{\frac{1}{q}}.$$

特别地, 当$p = q = 2$时Hölder不等式成为Cauchy-Schwartz不等式