第四章　随机变量的数字特征

知识点

1、数学期望

$$E(X)=\sum_{i=1}^{\infty}x_k p_k \qquad E(X)=\int_{-\infty}^{\infty}xf(x)dx$$

性质

$$E(aX+bY)=aEX+bEY$$

若X,Y独立，则 E(XY)=E(X)E(Y)

2、随机变量函数的数学期望

设 Y=g(X), g(x) 是连续函数，

$$E(Y)=\sum_{k=1}^{\infty}g(x_k)p_k \qquad E(Y)=\int_{-\infty}^{\infty}g(x)f(x)dx$$

若(X,Y) 是二维随机变量，$g(x,y)$是二元连续函数，

$$Z = g(X, Y)$$

(1)若(X,Y) 的分布律为 $P\{X = x_i, Y = y_j\} = p_{ij}$，

则

$$\boxed{E(Z) = \sum_{i,j=1}^{\infty} g(x_i, y_j) p_{ij}}$$

(2). 若(X,Y)的概率密度为 $f(x,y)$,

则

$$\boxed{E(Z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f(x,y) dxdy}$$

若二维连续型随机变量（X，Y）的联合密度为 $f(x,y)$，则

$$E[X]=\int_{-\infty}^{+\infty} x f_X(x)dx=\int_{-\infty}^{+\infty} x\left\{\int_{-\infty}^{+\infty} f(x,y)dy\right\}dx$$

$$=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x f(x,y)dxdy$$

$$E[Y]=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y f(x,y)dxdy$$

3、**方差**

$$DX = E(X - EX)^2$$

计算公式：

$$DX = E(X - EX)^2 = \sum_{i=1}^{\infty}(x_i - EX)^2 \cdot p_i\text{，}\quad \text{离散型。}$$

$$DX = \int_{-\infty}^{\infty}(x - EX)^2 f(x)dx \quad \text{连续型。}$$

性质

$$\boxed{DX = EX^2 - (EX)^2}$$

$$D(aX + bY) = a^2DX + b^2DY + 2abE(X - EX)(Y - EY)\text{，}$$

若 X，Y 独立，则 $D(aX + bY) = a^2DX + b^2DY$

特别，当 X、Y 相互独立时，$D(X \pm Y) = DX + DY$。

二、要熟记两点分布、二项分布、泊松分布、均匀分布、指数分布和正态分布的数学期望与方差

1. 两点分布

X	0	1
p_k	$1-p$	p

$EX=p$，$DX=EX^2-(EX)^2=pq$。

2. 二项分布　$X\sim B(n,p)$　$\boxed{EX=np}$　$\boxed{DX=npq}$

3. 泊松分布　$X\sim\pi(\lambda)$　$\boxed{EX=\lambda=DX}$

4. 均匀分布　$X\sim U(a,b)$　$\boxed{EX=\dfrac{a+b}{2}}$　$\boxed{DX=\dfrac{(b-a)^2}{12}}$

5. 正态分布　$X\sim N(\mu,\sigma^2)$　$\boxed{EX=\mu,\quad DX=\sigma^2}$

6. 指数分布

$$f(x)=\begin{cases}\lambda e^{-\lambda x} & x>0\\ 0 & x\le 0\end{cases}$$

$$\boxed{E(X)=\frac{1}{\lambda},\ D(X)=\frac{1}{\lambda^2}}$$

三、 要会用契比雪夫不等式作概率估计

$$EX = \mu, \quad DX = \sigma^2$$

$$P\{|X - \mu| \geq \varepsilon\} \leq \sigma^2 / \varepsilon^2$$

$$P\{|X - \mu| < \varepsilon\} \geq 1 - \sigma^2 / \varepsilon^2$$

随机变量的标准化：

$$\boxed{Y = (X - EX) / \sqrt{DX},}$$

称Y是随机变量X的标准化了的随机变量。

四、 引进了协方差、相关系数的概念，要掌握它们的性质与计算

1、协方差及相关系数的定义

协方差定义： COV(X,Y)= E(X-EX)(Y-EY)

特别 COV（X，X）=DX

计算公式: COV（X，Y）=E(XY)- E(X)E(Y)

D(aX+bY)= $a^2 DX + b^2 DY + 2abCOV(X,Y)$

特别 $D(X \pm Y) = DX + DY \pm 2COV(X,Y)$

相关系数 $\rho_{XY} = \frac{COV(X,Y)}{\sqrt{DX}.\sqrt{DY}}$

若 $\rho_{XY}=0$，称 X,Y 不相关.

$$X,Y\text{不相关} \Leftrightarrow \rho_{XY}=0 \Leftrightarrow COV(X,Y)=0$$
$$\Leftrightarrow E(XY)=EXEY \Leftrightarrow D(X\pm Y)=DX+DY$$

X，Y独立与X,Y不相关的关系：

定理：若X，Y独立，则X,Y不相关。

但是，X，Y不相关，不一定有X，Y相互独立。

2、**协方差的性质**

1) $COV(X,Y)=COV(Y,X)$;

2) $COV(aX,\ bY)=abCOV(X,Y)$;

3) $COV(X+Y,Z)=COV(X,Z)+COV(Y,Z)$;

3、相关系数的性质

1) $|\rho_{XY}| \le 1$;

2) $|\rho_{XY}| = 1 \Leftrightarrow \exists a, b,$ 使 $P\{Y = bX + a\} = 1$

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

说　明

相关系数是表征随机变量 X 与 Y 之间线性关系紧密程度的量.

当 $|\rho_{X,Y}| = 1$ 时，X 与 Y 之间以概率1存在着线性关系；

当 $|\rho_{X,Y}|$ 越接近于0时，X 与 Y 之间的线性关系越弱；

当 $|\rho_{X,Y}| = 0$ 时，X 与 Y 之间不存在线性关系（不相关）.

X与Y之间没有线性关系并不表示它们之间没有关系。

五、 要掌握二维正态随机变量的不相关与独立的等价性

设 (X,Y) 服从二维正态分布 $N(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$

则 X，Y独立 $\Longleftrightarrow \rho_{XY}=\rho=0 \Longleftrightarrow$ X，Y不相关。

n维正态分布的性质

1) **n** 维随机变量 $(X_1, \cdots, X_n)$ 服从 **n** 维正态分布的充分必要条件 $X_1, \cdots, X_n$ 的任意线性组合 $l_1 X_1 + \cdots + l_n X_n$ 服从一维正态分布。

2) 若 $(X_1, \cdots, X_n)$ 服从 **n** 维正态分布， $Y_1, \cdots, Y_n$ 是 $X_j (j = 1, \cdots, n)$ 的线性函数，则 $(Y_1, \cdots, Y_n)$ 也服从正态分布。

3) 若 $(X_1, \cdots, X_n)$ 服从 **n** 维正态分布，则 $X_1, \cdots, X_n$ 相互独立与两两不相关等价。

4) 相互独立的一维 正态随机变量的线性组合服从正态分布

5) n维 正态随机变量的边缘分布是一维正态分布；

反之，若 $X_1,\cdots,X_n$ 服从正态分布，且相互独立，则 $(X_1,\cdots,X_n)$ 服从n维正态分布。

注: 1) 若 $(X, Y) \sim N\left(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\right)$

则 $aX + bY \sim N(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2 + 2ab\rho\sigma_1\sigma_2)$

2) 若 X，Y 相互独立，且 $X \sim N\left(\mu_1, \sigma_1^2\right)$

$Y \sim N\left(\mu_2, \sigma_2^2\right)$

则 $aX + bY \sim N(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)$