1 二维随机变量及其分布

1.1 二维随机变量的联合分布和边际分布

1.1.1 二维随机变量的联合分布函数与性质

1. 联合分布函数:

$$F(x,y) = P(X \le x, Y \le y)$$

2. 性质: 1)F(x,y) 分别关于 x，y 单调不减:

$$x_1 < x_2 \Rightarrow F(x_1,y) \le F(x_2,y)$$

$$y_1 < y_2 \Rightarrow F(x,y_1) \le F(x,y_2)$$

2)F(x,y) 满足非负有界性:

$$0 \le F(x,y) \le 1; F(+\infty,+\infty) = 1$$

对任意 x，y:$F(x,-\infty) = F(-\infty,y) = F(-\infty,-\infty) = 0$
显然 $F(x,+\infty) = F_X(x)$
3)F(x,y) 分别关于 x，y 右连续:$F(x+0,y) = F(x,y), F(x,y+0) = F(x,y)$
4) 若 $x_1 < x_2, y_1 < y_2$, 则

$$P(x_1 < X < x_2, y_1 < Y < y_2) = F(x_2,y_2) - F(x_1,y_2) - F(x_2,y_1) + F(x_1,y_1)$$

1.1.2 边际分布函数

F(x,y) 是 (X,Y) 的联合分布函数，在二维变量里 X 的分布函数为:

$$F_X(x) = P(X \le x) = P(X \le x, Y \le +\infty) = F(x,+\infty); x \in R$$

称为边际分布函数,$F_Y(y)$ 同理

一般从联合分布函数可以求出边际分布函数，但是由两个边际分布函数很难求出联合分布函数、

1.2 二维离散随机变量

当 X，Y 都是离散型随机变量时
1. 联合分布列:$p_{ij} = P(X = x_i, Y = y_j)$
2. 联合分布列的性质:1) 非负性；2) 归一性
note: 由联合分布列可以求出联合分布函数:$F(x,y) = \sum_{x_i \le x} \sum_{y_i \le y} p_{ij}$

1.2.1 离散型随机变量的边际分布

1.X 的边际分布 $p_{i\cdot}$:$P(X = x_i) = P(X = x_i, Y < +\infty)$ $p_{i\cdot} = \sum_j p_{ij}$,Y 同理
Y 的边际分布 $p_{\cdot j}$ 同理

1.2.2 二维离散随机变量的独立性

1. 定义:(X,Y) 是二维随机变量，对于 X，Y 所有的取值 $x_1, X_2, \cdots y_1, y_2, \cdots$ 若有

$$P(X=x_i, Y=y_j) = P(X=x_i)P(Y=y_j)$$

则称 X，Y 相互独立

1.2.3 二维离散型随机变量的条件分布列

1. 定义:(X,Y) 是二维随机变量，对于 X，Y 所有的取值 $x_1, X_2, \cdots y_1, y_2, \cdots$ 称

$$p_{i|j} \triangleq P(X=x_i|Y=y_j) = \frac{P(X=x_i, Y=y_j)}{P(Y=y_j)}$$

为已知 $Y=y_i$ 条件下的 X 的分布列

对于 Y 同理可以定义

1.3 二维连续型随机变量及其分布

1.3.1 联合分布密度

1. 定义:F(x,y) 分布函数，存在非负函数 f(x,y) 使得所有的实数 x，y 有:

$$F(x,y) = P(X \le x, Y \le y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,dvdu$$

2.f(x,y) 性质: 1)$f(x,y) \ge 0$

2)$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,dxdy = 1$

3)F(x,y) 是二元连续函数

4) 在 f(x,y) 的连续点处有

$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y)$$

note:X,Y 一维连续，不一定有 (X,Y) 二维连续；但 (X,Y) 二维连续，一定有 X，Y 一维连续

1.3.2 二维连续型随机变量的边际密度

1. 定义:(X,Y) 是二维连续型随机变量，联合密度函数 f(x,y)，X 的边际密度函数:

$$f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy, x \in \mathbb{R}$$

Y 的边际密度函数同理定义

1.3.3 常用二维连续型随机变量的分布

1. 二维均匀分布设 D 为平面有界闭区间，其面积 S_D，若密度函数:

$$f(x,y) = \begin{cases} \frac{1}{S_D} & ,(x,y) \in D \\ 0 & ,(x,y) \notin D \end{cases}$$

称为二维变量 (X,Y) 服从 D 上的二维均匀分布

若 G 是 D 的子区域，则

$$P((x,y)\in G)=\iint_G f(x,y)\,d\sigma=\frac{1}{S_D}\iint_G d\sigma=\frac{S_G}{S_D}$$

2. 二维正态分布

记忆:$(X,Y)\ N(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)$，则 $X\ N(\mu_1,\sigma_1^2),Y\ N(\mu_2,\sigma_2^2)$

1.3.4 二维连续型随机变量的独立性

1. 定义:$f(x,y)=f_X(x)f_Y(y)$ 或 $F(x,y)=F_X(x)F_Y(y)$
2. 性质:$(X,Y)\ N(\mu_1,\mu_2,\sigma_1,\sigma_2,\rho)$, 则 XY 相互独立的充要条件 $\rho=0$

1.3.5 二维连续型随机变量的条件密度

由于二维连续型随机变量在某一条线上的概率是 0，所以无法用和离散型随机变量一样的公式处理，选择使用分布函数代替

1. 条件分布函数:1) 极限求 p78；2)$F_{X|Y}(x|y)=\int_{-\infty}^{x}\frac{f(u,y)}{f_Y(y)}\,du$
2. 条件密度函数:$f_{X|Y}(x|y)=\frac{f(x,y)}{f_Y(y)}(f_Y(y)>0)$