2.4 随机变量函数的分布

设 $\Omega \xrightarrow{X} \mathbb{R} \xrightarrow{g} \mathbb{R}$，则 $\Omega \xrightarrow{Y=g(X)} \mathbb{R}$

当 g 满足某些性质(可测)时，$g(X)$ 也为随机变量.

问题：已知 X 分布，如何求 Y 分布.

(根据定义求).

例1.（离散型） 设 X 分布列为

X	-2	-1	0	1	2
P	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{1}{10}$

求 $Y=X^2$ 的分布列.

解. Y 取值为 $4, 1, 0$.

$$P(Y=4) = P(X=2) + P(X=-2) = \frac{1}{5}$$

$$P(Y=1) = P(X=1) + P(X=-1) = \frac{2}{5}$$

$$P(Y=0) = P(X=0) = \frac{2}{5}$$

$\therefore$ Y 分布列为

Y	4	1	0
P	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$

例. 设 X 概率密度为 $f(x)$，$g(x) = I_{[0,\infty)}(x)$. 求 $Y=g(X)$ 分布.

解 Y 取值为 $0, 1$.

$$P(Y=0) = P(X<0) = \int_{-\infty}^{0} f(x)\,dx.$$

$$P(Y=1) = P(X\geq 0) = \int_{0}^{\infty} f(x)\,dx.$$ $\square$

(X 连续但 Y 离散).

例. 设 $X \sim N(\mu, \sigma^2)$，求 $Y = aX + b$ 的概率密度.

解. $F_Y(y) = P(Y \leq y) = P(aX + b \leq y)$

$$= P\left(X \leq \frac{y-b}{a}\right) = \int_{-\infty}^{\frac{y-b}{a}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$\therefore$ $f_Y(y) = F_Y'(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\frac{y-b}{a}-\mu)^2}{2\sigma^2}} \frac{d}{dy}\left(\frac{y-b}{a}\right)$

$$= \frac{1}{\sqrt{2\pi}\, a\sigma} e^{-\frac{(y-b-a\mu)^2}{2a^2\sigma^2}}$$

即 $Y \sim N(b + a\mu, a^2\sigma^2)$. □

(正态分布的线性变换仍为正态分布)

例. 设 $X \sim N(0,1)$. 求 $Y = X^2$ 的概率密度.

解. $Y \geq 0$. $\forall y > 0$,

$F_Y(y) = P(Y \leq y) = P(X^2 \leq y)$

$$= P(-\sqrt{y} \leq X \leq \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

法一：直接求导

$f_Y(y) = F_Y'(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \frac{d}{dy}\sqrt{y} - \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \frac{d}{dy}(-\sqrt{y})$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} \cdot \frac{1}{\sqrt{y}}$$

法二：由对称性 $F_Y(y) = 2\int_0^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$

再求导.

$$\therefore \quad f_Y(y) = \begin{cases} 0, & y \leq 0 \\ \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}}, & y > 0. \end{cases}$$

定理 2.6 设 X 概率密度为 $f_X(x)$. $y=g(x)$ 严格单调. g^{-1} 有连续导数. 则 $Y=g(X)$ 概率密度为

$$f_Y(y)=\begin{cases} f_X(g^{-1}(y))\,|g^{-1}(y)'|, & y\in(\alpha,\beta), \\ 0, & \text{其他}. \end{cases}$$

其中, (α,β) 为 $g(x)$ 的值域.

证明. 不妨设 $g\uparrow$. $\forall\ y\in(\alpha,\beta)$,

$$F_Y(y)=P(Y\le y)=P(g(X)\le y)$$
$$=P(X\le g^{-1}(y))=F_X(g^{-1}(y)).$$

$$\therefore f_Y(y)=\begin{cases} f_X(g^{-1}(y))\,g^{-1}(y)', & y\in(\alpha,\beta) \\ 0, & \text{其他} \end{cases}$$ ▨

推论 2.1 设 X 概率密度为 $f_X(x)$. $\{I_i\}$ 为 $\mathbb{R}$ 的划分. 对 $\forall i$, $g(x)$ 在 I_i 上满足严格单调且反函数连续可导. 记 $g_i(x)=g(x)I_{x\in I_i}$.

则
$$f_Y(y)=\begin{cases} \sum\limits_{i:\,y\in g(I_i)} f_X(g_i^{-1}(y))\,|g_i^{-1}(y)'|, & y\in(\alpha,\beta). \\ 0, & \text{其他}. \end{cases}$$

证明略.

总结

① 随机变量, 分布函数.

② 离散型: 二项分布、伯努利分布、泊松分布、超几何分布、几何分布

③ 连续型: 均匀分布、指数分布、正态分布.

④ 求随机变量函数的分布.

3. 多维随机变量及其分布

3.1 多维随机变量.

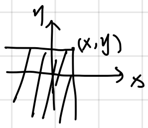
3.1.1 多维随机变量.

定义 3.1　设 $X_1, \cdots, X_n$ 是定义在 $(\Omega, \mathcal{F}, P)$ 上的 n 个随机变量，则称 $(X_1, \cdots, X_n)$ 为 n 维随机变量（或随机向量）.

例　掷两次骰子，X_1 = 第一次向上点数，X_2 = 第二次向上点数，则 (X_1, X_2) 为 2 维随机向量.

下面只讨论二维情形.　　r. v. { random vector / random variable.

定义 3.2　设 (X, Y) 为二维 r. v. 称



$$F(x, y) = P(X \le x, Y \le y), \quad x, y \in \mathbb{R},$$

为 (X, Y) 的联合分布函数，称 $F_X(x)$（或 $F_Y(y)$）为 (X, Y) 关于 X（或 Y）的边缘分布函数.

定理 3.1　设 (X, Y) 的联合分布函数为 $F(x, y)$，则

(1)　$0 \le F(x, y) \le 1, \quad \forall x, y.$

(2)　$F(x, y)$ 分别关于 x 和 y 单调非降.

(3)　$F(x, y)$ 分别关于 x 和 y 右连续.

(4)　$\forall x, y,$

$$F(-\infty, y) = F(x, -\infty) = F(-\infty, -\infty) = 0,$$

$$F(+\infty, +\infty) = 1.$$

(5)　$\forall x_1 < x_2, \quad y_1 < y_2,$

$$F(x_1, y_1) + F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) \ge 0.$$

(6) $\forall x, y$,

$$F(x,+\infty)=F_X(x), \qquad F(+\infty, y)=F_Y(y).$$

(7) 若 F 满足 (1)—(5)，则 F 为某二维 r.v. (X,Y) 的联合分布函数.

证明略.

例. 设

$$F(x,y)=\begin{cases} 0, & x<0 \text{ 或 } y<0, \\ \frac{1}{2}xy+\frac{1}{2}y, & 0\le y\le x<1, \\ \frac{1}{2}xy+\frac{1}{2}x, & 0\le x\le y<1, \\ x, & 0\le x<1,\ y\ge 1, \\ y, & 0\le y<1,\ x\ge 1, \\ 1, & x\ge 1,\ y\ge 1. \end{cases}$$

则

$$F_X(x)=\begin{cases} 0, & x<0, \\ x, & 0\le x<1, \\ 1 & x\ge 1. \end{cases}$$

3.1.2 二维离散型 r.v.

(X,Y) 只取有限个或可列个值.

联合分布列：$P(X=x_i, Y=y_j)=p_{ij}$

关于 X 或 Y 的边缘分布列：

$$P(X=x_i)=p_{i\cdot}, \qquad P(Y=y_j)=p_{\cdot j}$$

$X \backslash Y$	y_1	y_2	$\cdots$	y_j	$\cdots$	
x_1	p_{11}	p_{12}	$\cdots$	p_{1j}	$\cdots$	$p_{1\cdot}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2j}	$\cdots$	$p_{2\cdot}$
$\vdots$		$\vdots$				$\vdots$
x_i	p_{i1}	p_{i2}	$\cdots$	p_{ij}	$\cdots$	$p_{i\cdot}$
$\vdots$		$\cdots$				$\vdots$
	$p_{\cdot 1}$	$p_{\cdot 2}$	$\cdots$	$p_{\cdot j}$	$\cdots$	

性质 1) $p_{ij} \geq 0$

2) $\sum_{i,j} p_{ij} = 1$.

3) $\sum_j p_{ij} = p_{i\cdot}$, $\sum_i p_{ij} = p_{\cdot j}$.

例. 设

X	-1	0	1
P	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Y	0	1
P	$\frac{1}{2}$	$\frac{1}{2}$

且 $P(XY=0)=1$. 求 (X,Y) 联合分布列.

解:

X \ Y	0	1	
-1	$\frac{1}{4}$	0	$\frac{1}{4}$
0	0	$\frac{1}{2}$	$\frac{1}{2}$
1	$\frac{1}{4}$	0	$\frac{1}{4}$
	$\frac{1}{2}$	$\frac{1}{2}$	

例 袋中 a 个红球, b 个白球. 采用 1) 有放回方式 2) 无放回方式 取两次球, 定义

$$X = \begin{cases} 1 & \text{第一次取红} \\ 0 & \cdots \text{白} \end{cases} \qquad Y = \begin{cases} 1 & \text{第二次取红} \\ 0 & \cdots \text{白} \end{cases}$$

求 (X,Y) 联合分布列及边缘分布列.

解: 有放回

X \ Y	1	0	
1	$\left(\frac{a}{a+b}\right)^2$	$\frac{ab}{(a+b)^2}$	$\frac{a}{a+b}$
0	$\frac{ab}{(a+b)^2}$	$\left(\frac{b}{a+b}\right)^2$	$\frac{b}{a+b}$
	$\frac{a}{a+b}$	$\frac{b}{a+b}$	

无放回

X \ Y	1	0	
1	$\frac{a}{a+b}\,\frac{a-1}{a+b-1}$	$\frac{a}{a+b}\,\frac{b}{a+b-1}$	$\frac{a}{a+b}$
0	$\frac{b}{a+b}\,\frac{a}{a+b-1}$	$\frac{b}{a+b}\,\frac{b-1}{a+b-1}$	$\frac{b}{a+b-1}$
	$\frac{a}{a+b}$	$\frac{b}{a+b}$	