§2.3 随机变量函数的分布

问题：已知随机变量 X 的概率特性 —— 分布函数 或概率密度（分布律）

$$Y = g(X)$$

求 随机变量Y 的概率特性

方法：将与 Y 有关的事件转化成 X 的事件

一、离散型随机变量函数的分布

设随机变量 X 的分布律为

$$P\{X = x_k\} = p_k, \quad k = 1,2,\cdots$$

由已知函数 $g(x)$ 可求出随机变量 Y 的所有可能取值，则 Y 的分布律为

$$P\{Y = y_i\} = \sum_{k: g(x_k) = y_i} p_k, \quad i = 1,2,\cdots$$

例1 已知 X 的分布律为：

X	-1	0	1	2
p_k	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$

求 $Y_1 = 2X - 1$ 与 $Y_2 = X^2$ 的分布律

解

Y_1	-3	-1	1	3
P	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$

Y^2	1	0	1	4
p_i	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{2}$

Y^2	0	1	4
P	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{2}$

例2 已知 X 的分布律为：

$$P\{X=k\frac{\pi}{2}\}=pq^k, \quad k=0,1,2,\cdots$$

其中 $p+q=1,\ 0<p<1,$

求 $Y=\mathrm{Sin}\ X$ 的分布律。

解 $$P\{Y=0\}=P\left(\bigcup_{m=0}^{\infty}\{X=2m\cdot\frac{\pi}{2}\}\right)$$

$$=\sum_{m=0}^{\infty}pq^{2m}=\frac{p}{1-q^2}$$

$$P\{Y=1\} = P\left(\bigcup_{m=0}^{\infty}\{X=2m\pi+\frac{\pi}{2}\}\right)$$

$$= P\left(\bigcup_{m=0}^{\infty}\{X=(4m+1)\frac{\pi}{2}\}\right)$$

$$= \sum_{m=0}^{\infty} pq^{4m+1} = \frac{pq}{1-q^4}$$

$$P\{Y=-1\} = P\left(\bigcup_{m=0}^{\infty}\{X=2m\pi+\frac{3\pi}{2}\}\right)$$

$$= P\left(\bigcup_{m=0}^{\infty}\{X=(4m+3)\frac{\pi}{2}\}\right)$$

$$= \sum_{m=0}^{\infty} pq^{4m+3} = \frac{pq^3}{1-q^4}$$

故 Y 的分布律为

Y	-1	0	1
P	$\dfrac{pq^3}{1-q^4}$	$\dfrac{p}{1-q^2}$	$\dfrac{pq}{1-q^4}$

二、连续型随机变量函数的分布

已知随机变量 X 的概率密度 $f(x)$ (或分布函数) 求 $Y = g(X)$ 的概率密度或分布函数

方法：从分布函数出发,先求分布函数，再通过分布函数求概率密度.

1、$y=g(x)$是一个严格单调函数

例3 已知 X 概率密度为 $f_X(x), Y = aX + b, a, b$ 为常数，且 $a \neq 0$, 求 $f_Y(y)$

解 $$F_Y(y) = P\{Y \le y\}$$

$$= P\{aX + b \le y\}$$

当$a > 0$ 时，

$$F_Y(y) = P\{X \le \frac{1}{a}(y-b)\}$$

$$= \int_{-\infty}^{\frac{y-b}{a}} f_X(x)dx$$

$$\Longrightarrow f_Y(y) = \frac{1}{a} f_X\left(\frac{1}{a}(y-b)\right)$$

当$a<0$时，

$$F_Y(y)=P\{X\geq\frac{1}{a}(y-b)\}$$

$$=\int_{\frac{y-b}{a}}^{+\infty}f_X(x)dx$$

$$f_Y(y)=-\frac{1}{a}f_X\left(\frac{1}{a}(y-b)\right)$$

故

$$\boxed{f_Y(y)=\frac{1}{|a|}f_X\left(\frac{1}{a}(y-b)\right)}$$

例如，设 $X \sim N(\mu, \sigma^2)$，$Y = aX + b$，则

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{1}{a}(y-b)\right)$$

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

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

特别地，若 $X \sim N(\mu, \sigma^2)$，

则 $Y = \dfrac{X-\mu}{\sigma} \sim N(0,1)$

定理　设X为一连续型随机变量，　其概率密度为$f_X(x)$,若$y=g(x)$为一严格单调的函数，　其反函数$x=h(y)$有连续导数，则：$Y=g(X)$也是一个连续型随机变量，其概率密　度为：

$$f_Y(y)=\begin{cases} f_X[h(y)]\,|h'(y)|, & \alpha<y<\beta \\ 0, & \text{其它} \end{cases}$$

其中，$\alpha=\min\limits_{a\le x\le b} g(x),\ \beta=\max\limits_{a\le x\le b} g(x),$

证　设$y=g(x)$为严格单调增函数，值 域为(α,β),

其中$\alpha=g(-\infty),\beta=g(+\infty)$,则它的反函数　$x=h(y)$

在(α,β)内也是严格单调增函数。下求$F_Y(y)$：

当$y\le\alpha$时，$F_Y(y)=P\{Y\le y\}=0$；

当$y\ge\beta$时，$F_Y(y)=P\{Y\le y\}=1$；

当$\alpha<y<\beta$时，$F_Y(y)=P\{Y\le y\}=P\{g(X)\le y\}$

$$=P\{X\le h(y)\}=\int_{-\infty}^{h(y)}f_X(x)dx$$

于是Y的概率密度为：($h'(x)>0$)

$$f_Y(y)=\begin{cases} f_X[h(y)]h'(y), & \alpha<y<\beta \\ 0, & \text{其它} \end{cases}$$

同理可得：当 $y=g(x)$为严格单调减函数时，有：

$h'(x)\le 0$,且Y的概率密度为：

$$f_Y(y)=\begin{cases} f_X[h(y)][-h'(y)], & \alpha<y<\beta \\ 0, & 其它 \end{cases}$$

$\therefore$ 综上所述， $Y=g(x)$的概率密度为：

$$f_Y(y)=\begin{cases} f_X[h(y)]\left|h'(y)\right|, & \alpha<y<\beta \\ 0, & 其它 \end{cases}$$

例4 $X \sim E(2)$, $Y = -3X + 2$, 求 $f_Y(y)$

解 $f_X(x) = \begin{cases} 2e^{-2x}, & x > 0 \\ 0, & x \le 0 \end{cases}$

$$f_Y(y) = \frac{1}{|-3|} f_X\left(\frac{1}{-3}(y-2)\right)$$

$$= \begin{cases} \dfrac{1}{3} \cdot 2e^{-2\cdot\left(-\frac{y-2}{3}\right)}, & -\dfrac{y-2}{3} > 0 \\ 0, & \text{其他} \end{cases}$$

$$= \begin{cases} \dfrac{2}{3} e^{-\frac{2(2-y)}{3}}, & y < 2 \\ 0, & \text{其他} \end{cases}$$

2、$y=g(x)$不是一个严格单调函数

例5 已知 $X \sim N(0,1)$, $Y = X^2$, 求 $f_Y(y)$

解　从分布函数出发 $f_X(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad -\infty < x < +\infty$

$$F_Y(y) = P\{Y \le y\}$$

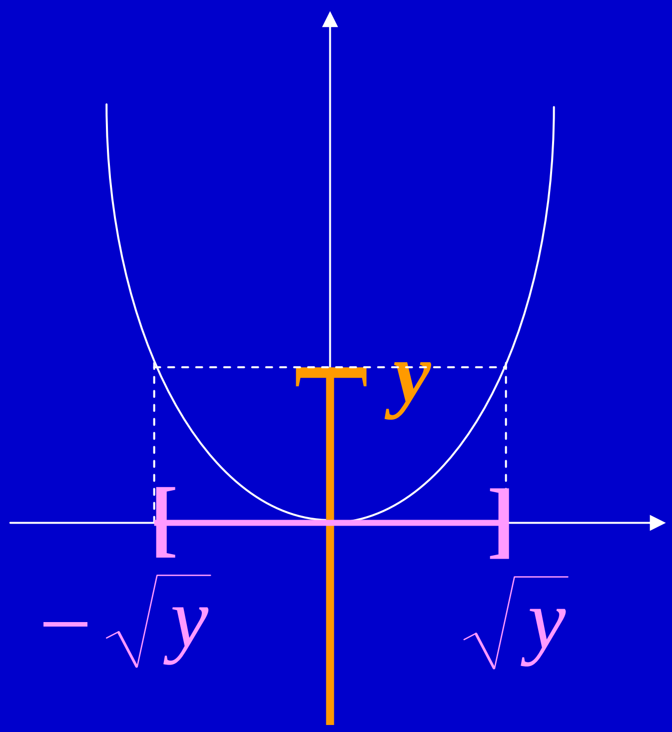
当$y \le 0$ 时，$F_Y(y) = 0$

当$y > 0$ 时，

$$F_Y(y) = P\{X^2 \le y\}$$

$$= P\{-\sqrt{y} \le X \le \sqrt{y}\}$$

$$= \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)dx$$



$$= \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

$$= \frac{2}{\sqrt{2\pi}} \int_0^{\sqrt{y}} e^{-\frac{x^2}{2}} dx$$

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

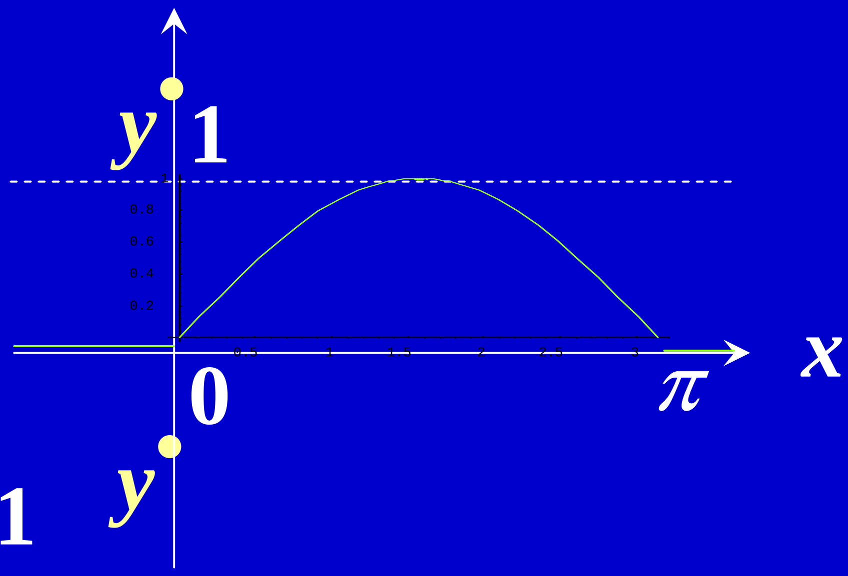
例6 设 X 的概率密度函数为

$$f_X(x)=\begin{cases}\dfrac{2x}{\pi^2}, & 0<x<\pi\\ 0, & \text{其他}\end{cases}$$

求$Y=\sin X$ 的概率密度函数

解 由图可知, Y 的取值范围为(0,1)

$y=\sin x(0<x<\pi)$



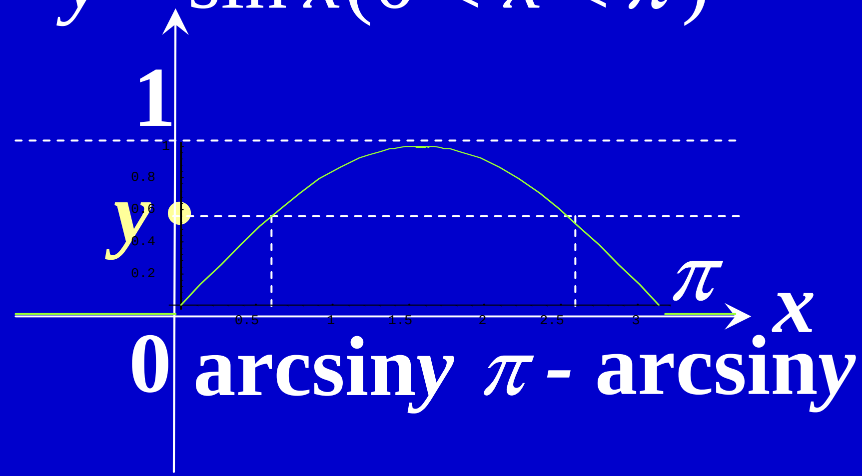
故当 $y\le 0$时,

$$F_Y(y)=0$$

当 $y\ge 1$时, $F_Y(y)=1$

当$0 < y < 1$时

$$y = \sin x (0 < x < \pi)$$



$$F_Y(y) = P\{Y \le y\} = P\{\sin X \le y\}$$

$$= \left[\int_0^{\arcsin y} + \int_{\pi - \arcsin y}^{\pi} \right] f_X(x) dx$$

$$= \left[\int_0^{\arcsin y} + \int_{\pi - \arcsin y}^{\pi} \right] \frac{2x}{\pi^2} dx$$

$$= \frac{2 \arcsin y}{\pi}$$

$$\therefore F_Y(y)=\begin{cases}0, & y\le 0\\ \dfrac{2\arcsin y}{\pi}, & 0<y<1\\ 1, & y\ge 1\end{cases}$$

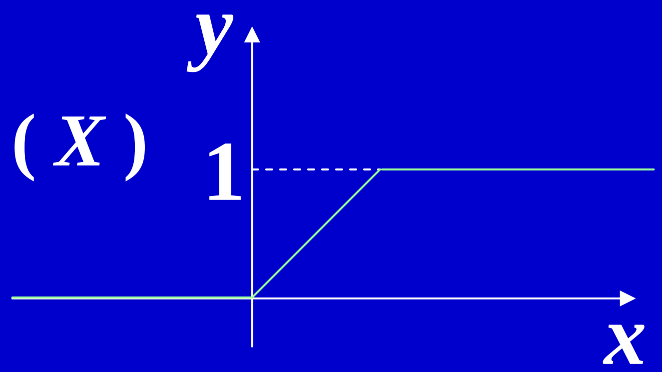
$$\text{故}\quad f_Y(y)=\begin{cases}\dfrac{2}{\pi\sqrt{1-y^2}}, & 0<y<1\\ 0, & \text{其他}\end{cases}$$

注意：连续型随机变量的函数的分布函数不一定是连续函数

例如：$X \sim U(0,2)$ $\quad f_X(x)=\begin{cases}\frac{1}{2}, & 0<x<2\\ 0, & \text{其他}\end{cases}$

$$g(x)=\begin{cases}0, & x\le 0\\ x, & 0<x\le 1\\ 1, & x>1\end{cases}$$

令 $Y=g(X)$



$$F_Y(y)=\begin{cases}0, & y<0,\\ \frac{y}{2}, & 0\le y<1\\ 1, & y\ge 1\end{cases}$$

$F_Y(y)$不是连续函数