第2.4节 随机变量的函数的分布

一、离散型随机变量的函数的分布

二、连续型随机变量的函数的分布

三、小结

一、离散型随机变量的函数的分布

设 X为离散型随机变量，其概率分布已知。

$Y = f(X)$为X的函数.

问题

如何根据随机变量X 的分布

求得随机变量$Y = f(X)$的分布?

例1 设 X 的分布律为

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

求 $Y = X^2$ 的分布律 .

解 Y 的可能值为 $(-1)^2,\ 0^2,\ 1^2,\ 2^2;$

即 $0,\ 1,\ 4.$

$$P\{Y=0\} = P\{X^2=0\} = P\{X=0\} = \frac{1}{4},$$

$$P\{Y=1\}=P\{X^2=1\}=P\{(X=-1)\cup(X=1)\}$$

$$=P\{X=-1\}+P\{X=1\}=\frac{1}{4}+\frac{1}{4}=\frac{1}{2},$$

$$P\{Y=4\}=P\{X^2=4\}=P\{X=2\}=\frac{1}{4},$$

故 Y 的分布律为

Y	0	1	4
p	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

由此归纳出离散型随机变量函数的分布的求法.

离散型随机变量函数概率分布的计算

如果 X 是离散型随机变量，其函数 $Y = f(X)$ 也是离散型随机变量.若 X 的分布律为

X	x_1	x_2	$\cdots$	x_k	$\cdots$
p_k	p_1	p_2	$\cdots$	p_k	$\cdots$

则$Y = f(X)$的分布律为

$Y = f(X)$	$f(x_1)$	$f(x_2)$	$\cdots$	$f(x_k)$	$\cdots$
p_k	p_1	p_2	$\cdots$	p_k	$\cdots$

若 $f(x_k)$中有值相同的,应将相应的 p_k 合并.

例2 设

X	-1	1	2
p_k	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$

求 $Y = X^2 - 5$ 的分布律.

解 Y 的分布律为

Y	-4	-1
p	$\frac{1}{2}$	$\frac{1}{2}$

二、连续型随机变量的函数的分布

例3 设随机变量 X 的概率密度为

$$f_X(x)=\begin{cases}\dfrac{x}{8}, & 0<x<4,\\ 0, & \text{其它}.\end{cases}$$

求随机变量 $Y=2X+8$ 的概率密度.

解 第一步 先求Y=2X+8 的分布函数 $F_Y(y)$.

$$F_Y(y)=P\{Y\le y\}=P\{2X+8\le y\}$$

$$= P\{X \le \frac{y-8}{2}\} = \int_{-\infty}^{\frac{y-8}{2}} f_X(x)\,\mathrm{d}\,x$$

第二步　由分布函数求概率密度.

$$f_Y(y) = F'_y(y)$$

$$= [\int_{-\infty}^{\frac{y-8}{2}} f_X(x)\,\mathrm{d}\,x]'$$

$$= f_X(\frac{y-8}{2})(\frac{y-8}{2})',$$

所以 $$f_Y(y)=\begin{cases}\dfrac{1}{8}\left(\dfrac{y-8}{2}\right)\cdot\dfrac{1}{2}, & 0<\dfrac{y-8}{2}<4,\\ 0, & \text{其它}.\end{cases}$$

$$=\begin{cases}\dfrac{y-8}{32}, & 8<y<16,\\ 0, & \text{其它}.\end{cases}$$

例4 设随机变量 X 的概率密度为

$$f_X(x)=\begin{cases}0, & x<0,\\ x^3e^{-x^2}, & x\ge 0.\end{cases}$$

求随机变量 $Y=X^2$ 和 $Y=2X+3$ 的概率密度.

解 先求随机变量 $Y=X^2$ 分布函数,

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

$$(y>0\text{时})=P\{-\sqrt{y}\le X\le \sqrt{y}\}$$

$$=F_X(\sqrt{y})-F_X(-\sqrt{y})$$

$$= \int_{-\infty}^{\sqrt{y}} f_X(x)\,\mathrm{d}\,x - \int_{-\infty}^{-\sqrt{y}} f_X(x)\,\mathrm{d}\,x.$$

再由分布函数求概率密度.

$$f_Y(y) = F_Y'(y) \;= f_X(\sqrt{y})(\sqrt{y})' - f_X(-\sqrt{y})(-\sqrt{y})'$$

$$= \frac{1}{2\sqrt{y}} \cdot (\sqrt{y})^3 \cdot e^{-(\sqrt{y})^2} + 0 \cdot \frac{1}{2\sqrt{y}}$$

$$= \begin{cases} \dfrac{ye^{-y}}{2}, y > 0, \\ 0, \qquad y \le 0. \end{cases}$$

当 Y=2X+3 时,有

$$y = 2x + 3 \Rightarrow x = \frac{y-3}{2},$$

$$f_Y(y) = F'_y(y) = [\int_{-\infty}^{\frac{y-3}{2}} f_X(x)\,\mathrm{d}\,x]'$$

$$= \begin{cases} (\frac{y-3}{2})^3 e^{-(\frac{y-3}{2})^2} (\frac{y-3}{2})', & y \geq 3, \\ 0, & y < 3. \end{cases}$$

$$
= \begin{cases} \dfrac{1}{2}\left(\dfrac{y-3}{2}\right)^3 e^{-\left(\frac{y-3}{2}\right)^2}, & y \geq 3, \\ 0, & y < 3. \end{cases}
$$

由上述例题可归纳 出计算连续型随机变量 的函数的概率密度的方法 .

定理 设$g(x)$是严格单调函数。随机变量X的密度函数为$f_X(x)$,试证随机变量$Y=g(X)$的密度函数为

$$f_Y(y)=f_X\left[h(y)\right]\cdot\left|\left[h(y)\right]'\right|$$

这里$x=h(y)$为$y=g(x)$的反函数。

证明：若$g(x)$为单调函数，则$g'(x)>0$,因其反函数

$x=h(y)$存在且亦为单调增函数即$\left[h(y)\right]'>0$,则

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

$$P\{X\le h(y)\}=F_X(h(y))$$

进而 $$f_Y(y)=f_X(h(y))\left[h(y)\right]'=f_X(h(y))\cdot\left|[h(y)]'\right|$$

若$g(x)$为单调减函数，则$g(x)^{'}<0$,此时$\left[h(y)\right]^{'}<0$,则

$$F_Y(y)=P\{Y\leq y\}=P\{f(X)\leq y\}=$$

$$P\{X\geq h(y)\}=1-F_X(h(y))$$

$$f_Y(y)=-f_X(h(y))\left[h(y)\right]^{'}=f_X(h(y))\cdot\left|\left[h(y)\right]^{'}\right|$$

因而证得

$$f_Y(y)=f_X\left[h(y)\right]\cdot\left|\left[h(y)\right]^{'}\right|$$

例5 设随机变量 $X \sim N(\mu, \sigma^2)$，试证明 X 的线性函数 $Y = aX + b\ (a \neq 0)$ 也服从正态分布.

证明 X 的概率密度为

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty.$$

设 $y = g(x) = ax + b,$

得 $x = h(y) = \dfrac{y-b}{a},$ 知 $[h(y)]' = \dfrac{1}{a} \neq 0.$

由公式 $f_Y(y)=\begin{cases} f_X[h(y)]\cdot\left|[h(y)]'\right|, \alpha<y<\beta, \\ 0, \qquad\qquad \text{其它}. \end{cases}$

得 $Y=aX+b$ 的概率密度为

$$f_Y(y)=\frac{1}{|a|}f_X\left(\frac{y-b}{a}\right), \qquad -\infty<y<\infty.$$

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

得 $Y=aX+b$

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

$$=\frac{1}{|a|\sigma\sqrt{2\pi}}e^{-\frac{[y-(b+a\mu)]^2}{2(a\sigma)^2}}, \quad -\infty<y<\infty.$$

请同学们思考

设 $f(x)$ 是连续函数，若 X 是离散型随机变量，则 $Y = f(X)$ 也是离散型随机变量吗？若 X 是连续型的又怎样？

答 若 X 是离散型随机变量，它的取值是有限个或可列无限多个，因此 Y 的取值也是有限个或可列无限多个，因此 Y 是离散型随机变量. 若 X 是连续型随机变量，那末 Y 不一定是连续型随机变量.

例如 设 X 在 (0,2) 上服从均匀分布,概率密度为

$$f(x)=\begin{cases}\dfrac{1}{2}, & 0<x<2\\ 0, & \text{其他}.\end{cases}$$

又设连续函数 $y=g(x)=\begin{cases}x, & 0\le x\le 1\\ 1, & 1<x\le 2\end{cases}$,

则 $\boldsymbol{Y=f(X)}$ 的分布函数 $\boldsymbol{F_Y(y)}$ 可以计算出来:

由于 Y 的取值为 [0,1]，所以

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

当 $y > 1$ 时, $F_Y(y) = P\{Y \le y\} = 1;$

当 $0 \le y \le 1$ 时, $F_Y(y) = P\{Y \le y\} = P\{f(X) \le y\}$

$$= \int_{-\infty}^{y} f(x)\,\mathrm{d}\,x = \int_{0}^{y} \frac{1}{2}\,\mathrm{d}\,x = \frac{y}{2}.$$

故 Y 的分布函数为 $F_Y(Y)=\begin{cases}0, & y\le 0,\\ \dfrac{y}{2}, & 0<y\le 1,\\ 1, & y>1.\end{cases}$

因为 $F_Y(y)$ 在 $y=1$ 处间断,故 $Y=f(X)$ 不是连续型随机变量 ,又因为 $F_Y(y)$ 不是阶梯函数,故 $Y=f(X)$ 也不是离散型随机变量 .

三、小结

1. 离散型随机变量函数的分布

如果 X 是离散型随机变量,其函数 $Y = f(X)$ 也是离散型随机变量.若 X 的分布律为

X	x_1	x_2	$\cdots$	x_k	$\cdots$
p_k	p_1	p_2	$\cdots$	p_k	$\cdots$

则 $Y = f(X)$ 的分布律为

$Y = f(X)$	$f(x_1)$	$f(x_2)$	$\cdots$	$f(x_k)$	$\cdots$
p_k	p_1	p_2	$\cdots$	p_k	$\cdots$

若 $f(x_k)$ 中有值相同的,应将相应的 p_k 合并.

2. 连续型随机变量的函数的分布

方法1 $F_Y(y) = P\{Y \le y\} = P\{g(X) \le y\}$

$$= \int_{g(x)\le y} f_X(x)dx \qquad (-\infty < x < \infty)$$

$F_Y(y)$关于y求导得到Y的密度函数.

方法2

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

注意条件.