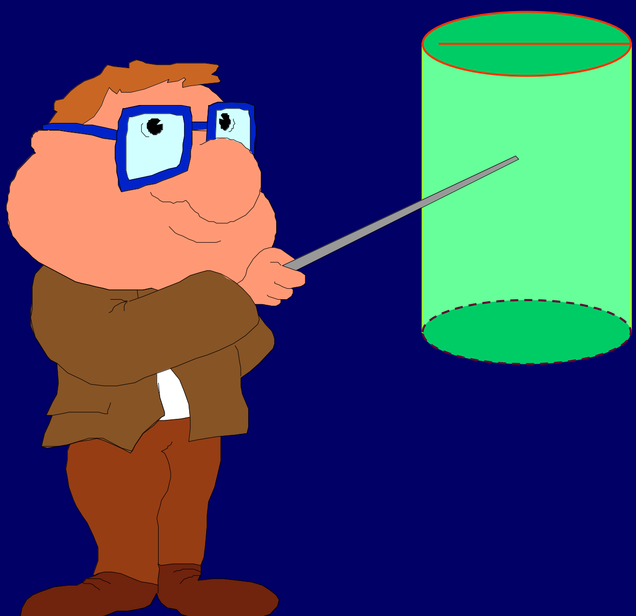
第五节 随机变量的函数的分布

- 问题的提出
- 离散型随机变量的函数的分布
- 连续型随机变量的函数的分布
- 小结 布置作业

一、问题的提出

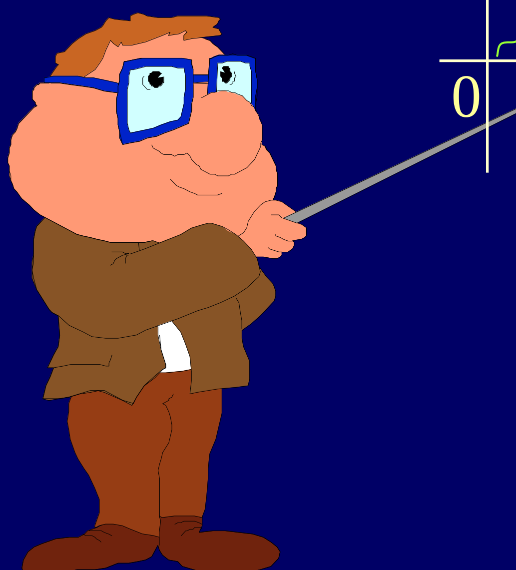
在实际中，人们常常对随机变量的函数更感兴趣.

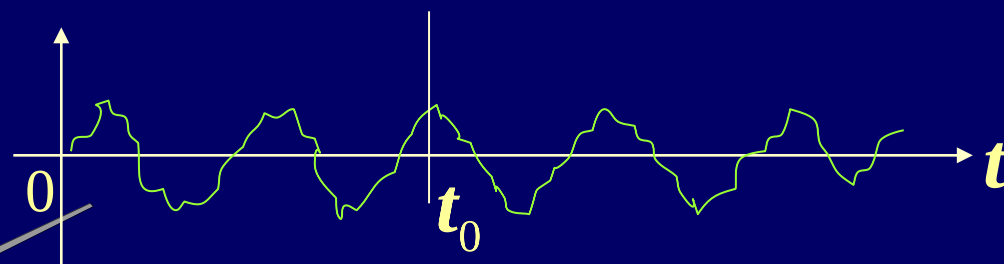
比如，已知圆轴截面直径 d 的分布，



求截面面积 $A=\dfrac{\pi d^2}{4}$ 的分布.

在比如， 已知 $t=t_0$ 时刻噪声电压 V 的分布，





求功率 $W=V^2/R$（R 为电阻）的分布等.

设随机变量 X 的分布已知，$Y=g(X)$ (设 g 是连续函数），如何由 X 的分布求出 Y 的分布？

这个问题无论在实践中还是在理论上都是重要的.

下面进行讨论.

二、离散型随机变量函数的分布

例1　设$X \sim \begin{Bmatrix} 1 & 2 & 5 \\ 0.2 & 0.5 & 0.3 \end{Bmatrix}$

求 $Y = 2X + 3$ 的概率函数.

解：　当 X 取值 1，2，5 时，

Y 取对应值 5，7，13，

而且X取某值与Y取其对应值是两个同时发生的事件，两者具有相同的概率.

故　$Y \sim \begin{Bmatrix} 5 & 7 & 13 \\ 0.2 & 0.5 & 0.3 \end{Bmatrix}$

一般地，若X是离散型 *r.v*，X 的分布律为

$$X \sim \begin{Bmatrix} x_1 & x_2 \cdots x_n \\ p_1 & p_2 \cdots p_n \end{Bmatrix}$$

则 $Y=g(X)$ $\sim \begin{Bmatrix} g(x_1) & g(x_2) \cdots g(x_n) \\ p_1 & p_2 \quad \cdots \quad p_n \end{Bmatrix}$

如果$g(x_k)$ 中有一些是相同的，把它们作适当并项即可.

如： $X \sim \begin{Bmatrix} -1 & 0 & 1 \\ 0.3 & 0.6 & 0.1 \end{Bmatrix}$

则 $Y=X^2$ 的分布律为：

$$Y \sim \begin{Bmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{0.6} & \mathbf{0.4} \end{Bmatrix}$$

三、连续型随机变量函数的分布

例2 设 $X \sim f_X(x)=\begin{cases} x/8, & 0<x<4 \\ 0, & 其它 \end{cases}$

求 $Y=2X+8$ 的概率密度.

解 设Y的分布函数为 $F_Y(y)$，

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

$$=P\{X\le \frac{y-8}{2}\}=F_X(\frac{y-8}{2})$$

于是Y的密度函数

$$f_Y(y)=\frac{dF_Y(y)}{dy}=f_X(\frac{y-8}{2})\cdot\frac{1}{2}$$

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

$$f_Y(y)=\frac{dF_Y(y)}{dy}=f_X(\frac{y-8}{2})\cdot\frac{1}{2}$$

$Y=2X+8$

注意到 $0<x<4$ 时， $f_X(x)\neq 0$

即 $8<y<16$ 时， $f_X(\frac{y-8}{2})\neq 0$

此时 $f_X(\frac{y-8}{2})=\frac{y-8}{16}$

故

$$f_Y(y)=\begin{cases} \frac{y-8}{32}, & 8<y<16 \\ 0, & 其它 \end{cases}$$

例3 设 X 具有概率密度 $f_X(x)$, 求 $Y=X^2$ 的概率密度.

解　设 Y 和 X 的分布函数分别为 $F_Y(y)$ 和 $F_X(x)$,

注意到 $Y=X^2 \geq 0$，故当 $y \leq 0$ 时，$F_Y(y)=0$.

当 $y>0$ 时，$F_Y(y)=P(Y \leq y)=P(X^2 \leq y)$

$$= P(-\sqrt{y} \leq X \leq \sqrt{y})$$

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

$$F_Y(y)=P(Y \leq y)$$

求导可得

$$f_Y(y)=\frac{dF_Y(y)}{dy}=\begin{cases}\dfrac{1}{2\sqrt{y}}\left[f_X(\sqrt{y})+f_X(-\sqrt{y})\right], & y>0\\ 0, & y\le 0\end{cases}$$

若 $f_X(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},\ -\infty<x<+\infty$

则 $Y=X^2$ 的概率密度为:

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

从上述两例中可以看到，在求$P(Y\le y)$ 的过程中，关键的一步是设法从$\{g(X)\le y\}$中解出X，从而得到与$\{g(X)\le y\}$等价的X 的不等式 .

例如，用 $\{X \le \frac{y-8}{2}\}$代替 $\{2X+8\le y\}$

用 $\{-\sqrt{y} \le X \le \sqrt{y}\}$代替$\{X^2\le y\}$

这样做是为了利用已知的 X的分布，从而求出相应的概率.

这是求r.v的函数的分布的一种常用方法.

例4 设随机变量X的概率密度为

$$f(x)=\begin{cases}\dfrac{2x}{\pi^2} & 0<x<\pi\\ 0 & \text{其它}\end{cases}$$

求 $Y=\sin X$ 的概率密度.

$-1\le Y\le 1$

解 注意到，当 $0<x<\pi$ 时 $0<y\le 1$

故

当 $y\le 0$ 时，$F_Y(y)=0,$

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

当 $y\ge 1$时，$F_Y(y)=1$

例4 设随机变量 X 的概率密度为

$$f(x)=\begin{cases}\dfrac{2x}{\pi^2} & 0<x<\pi \\ 0 & \text{其它}\end{cases}$$

求 $Y=\sin X$ 的概率密度.

$0<X<\dfrac{\pi}{2}$ 或 $\dfrac{\pi}{2}<X<\pi$

解 当 $0<y<1$ 时,

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

$$=P(0<X\le \arcsin y)+P(\pi-\arcsin y\le X<\pi)$$

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

$$= \left(\frac{\arcsin y}{\pi}\right)^2 + 1 - \left(\frac{\pi - \arcsin y}{\pi}\right)^2$$

而 $\boldsymbol{f_Y(y) = \dfrac{dF_Y(y)}{dy}}$

求导得:

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

例5 已知随机变量X的分布函数$F(x)$是严格单调的连续函数，证明$Y=F(X)$服从[0,1]上的均匀分布.

$0 \le Y \le 1$

证明 设 Y 的分布函数是 $G(y)$,

由于 $0 \le y \le 1$

$G(y) = P(Y \le y)$

于是 对 $y < 0$, $G(y) = 0$; 对 $y > 1$, $G(y) = 1$;

又由于X的分布函数F是严格递增的连续函数，其反函数 F^{-1} 存在且严格递增.

对$0 \le y \le 1$, $G(y)=P(Y\le y)=P(F(X)\le y)$

$=P(X\le F^{-1}(y))=F(F^{-1}(y))=y$

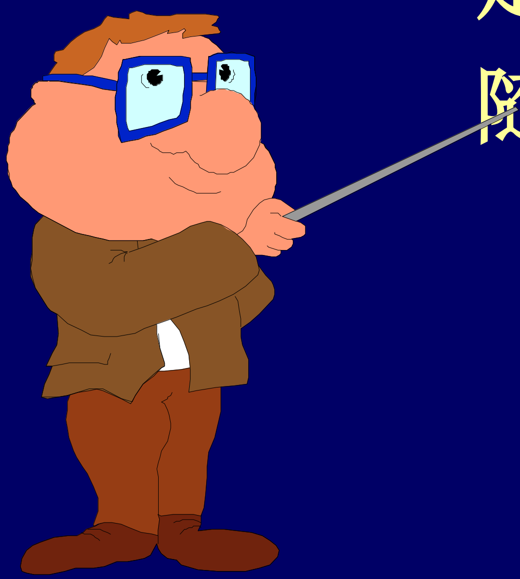
即Y的分布函数是

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

求导得Y的密度函数

$$g(y)=\begin{cases}1, & 0\le y\le 1\\ 0, & \text{其它}\end{cases}$$

可见, Y在[0,1]上服从的均匀分布.



下面给出一个定理，在满足定理条件时可直接用它求出随机变量函数的概率密度 .

定理 设 X是一个取值于区间$[a,b]$，具有概率密度$f(x)$的连续型 r.v,又设$y=g(x)$处处可导，且对于任意x，恒有 $g'(x)>0$或恒有 $g'(x)<0$，则$Y=g(X)$是一个连续型r.v，它的概率密度为

$$f_Y(y)=\begin{cases} f[h(y)]\left|\dfrac{dh(y)}{dy}\right|, & \alpha<y<\beta \\ 0, & \text{其它} \end{cases}$$

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

$x=h(y)$ 是 $y=g(x)$ 的反函数 .

此定理的证明与前面的解题思路类似

例7 设随机变量$X \sim N(\mu,\sigma^2)$服从正态分布，证明$Y = aX + b$ 也服从正态分布.

解 随机变量 X 的概率密度为

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

$y = g(x) = ax + b$ 解得$x = h(y) = \dfrac{y-b}{a}$

$h'(y) = \dfrac{1}{a}$ 所以$Y = aX + b$的概率密度为

$$f_y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right), \quad -\infty < y < +\infty$$

即

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

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

所以$Y=aX+b\sim N\left(a\mu+b,(a\sigma)^2\right)$

四、小结

这一节我们介绍了随机变量函数的分布.

对于连续型随机变量，在求 $Y=g(X)$ 的分布时，关键的一步是把事件 $\{g(X)\le y\}$ 转化为 X 在一定范围内取值的形式，从而可以利用 X 的分布来求 $P\{g(X)\le y\}$.

练习题

一、 设随机变量 X 的分布律为:

X	-2	-1	0	1	3
P	1/5	1/6	1/5	1/15	11/30

求 $Y=X^2$ 的分布律 ?

二、设 $X\sim N$（0，1）

（1）求 $Y=e^X$ 的概率密度

（2）求 $Y=2X^2+1$ 的概率密度。?

（3）求 $Y=|X|$的概率密度。

三、设随机变量 X 在（0，1）上服从均匀分布

（1）求 $Y=e^X$ 的分布密度

（2）求 $Y=-2\ln X$ 的概率密度。

一、　设随机变量 X 的分布律为：

X	-2	-1	0	1	3
P	1/5	1/6	1/5	1/15	11/30

求 $Y=X^2$ 的分布律

解：$Y=X^2$ 的所有可能取值为：$Y=0,1,4,9$

$$P\{Y=0\}=P\{X^2=0\}=P\{X=0\}=1/5$$

$$P\{Y=1\}=P\{X^2=1\}=P\{X=1 \text{或} X=-1\}$$

$$=P\{X=1\}+P\{X=-1\}=\frac{7}{30}$$

$$P\{Y=4\}=P\{X^2=4\}=P\{X=2 \text{或} X=-2\}$$

$$=P\{X=2\}+P\{X=-2\}=1/5$$

$$P\{Y=4\}=P\{X^2=4\}=P\{X=2\text{或}X=-2\}$$

$$=P\{X=2\}+P\{X=-2\}=P\{X=-2\}=1/5$$

$$P\{Y=9\}=P\{X^2=9\}=P\{X=3\text{或}X=-3\}$$

$$=P\{X=3\}+P\{X=-3\}=P\{X=3\}=11/30$$

设随机变量 Y 的分布律为:

Y	0	1	4	9
P	1/5	7/30	1/5	11/30

五、布置作业

《概率统计》标准化作业（二）三、3；四、