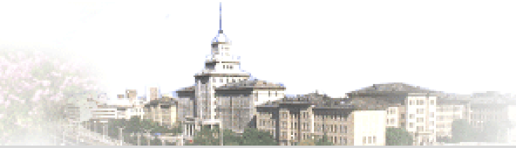
哈爾濱工業大學

第3章　随机变量及其分布

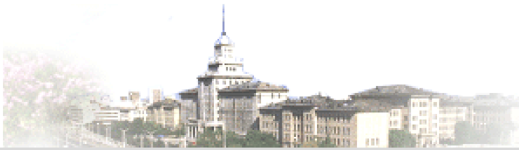
第17讲 随机变量函数的分布



问题 若已知分子运动的速率X的分布，如何求分子运动的动能 $Y=\frac{1}{2}mX^2$（m为分子的质量）分布？

对随机变量的函数$Y=g(X)$,
已知随机变量X的分布，
如何求随机变量Y的分布？

离散型随机变量函数的分布



例1 已知X的分布列为

X	-1	0	1	2
P	0.1	0.2	0.3	0.4

求$Y = X^2$ 的分布列.

解

P	0.1	0.2	0.3	0.4
X	-1	0	1	2
Y	1	0	1	4

$Y=X^2$	0	1	4
P	0.2	0.4	0.4

连续型随机变量函数的分布

例2 设$X\sim U(0,3)$, 求$Y=2X+3$ 的概率密度.

解 $f_X(x)=\begin{cases}1/3, & 0<x<3,\\ 0, & \text{其他}.\end{cases} \Rightarrow P(0<X<3)=1.$ 从而$P(3<Y<9)=1.$

当$y\le 3$时，$F_Y(y)=P(Y\le y)=0 \Rightarrow f_Y(y)=F_Y'(y)=0;$

当$y\ge 9$时，$F_Y(y)=P(Y\le y)=1 \Rightarrow f_Y(y)=F_Y'(y)=0;$

当$3<y<9$时，Y的分布函数为

$$F_Y(y)=P(Y\le y)=P(2X+3\le y)=P(X\le (y-3)/2)$$

$$=F_X((y-3)/2)$$

$$f_Y(y)=F_Y'(y)=f_X\left(\frac{y-3}{2}\right)\cdot\frac{1}{2}=\frac{1}{6}.$$

$$\boxed{f_Y(y)=\begin{cases}\dfrac{1}{6}, & 3<y<9,\\ 0, & \text{其它}.\end{cases}}$$

连续型随机变量函数概率密度的两种求法

分布函数法

已知X的概率密度$f_X(x)$,分布函数$F_X(x)$,$Y=g(X)$,

求Y概率密度$f_Y(y)$,分两步:

(1)先求Y的分布函数$F_Y(y)$

$$F_Y(y)=P(Y\le y)=P\{g(X)\le y\}\xrightarrow{\text{解出}X}\text{表示成}X\text{的分布函数;}$$

(2)求导数:$f_Y(y)=F_Y'(y)$.

例3 设随机变量X的概率密度为

$$f_X(x)=\begin{cases}3x^2/2, -1<x<1,\\ 0, \qquad 其他.\end{cases}$$ 求$Y=X^2$的概率密度.

解 由已知$P(-1<X<1)=1$，得$P(0<Y<1)=1$.

当$y\le 0$时，$F_Y(y)=P(Y\le y)=0 \Rightarrow f_Y(y)=F_Y'(y)=0;$

当$y\ge 1$时，$F_Y(y)=P(Y\le y)=1 \Rightarrow f_Y(y)=F_Y'(y)=0;$

当$0<y<1$时，Y的分布函数为（解法1）

$$F_Y(y)=P(Y\le y)=P(X^2\le y)=P(-\sqrt{y}\le X\le\sqrt{y})=F_X(\sqrt{y})-F_X(-\sqrt{y}),$$

$$f_Y(y)=F_Y'(y)=f_X(\sqrt{y})\frac{1}{2\sqrt{y}}+f_X(-\sqrt{y})\frac{1}{2\sqrt{y}}=\frac{3}{2}\sqrt{y}.$$

当$0 < y < 1$时，Y的分布函数为（解法2）

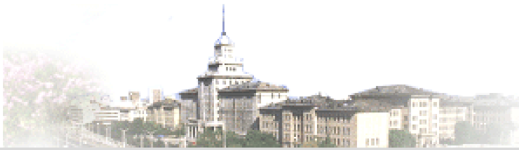
$$F_Y(y) = P(Y \le y) = P(X^2 \le y) = P(-\sqrt{y} \le X \le \sqrt{y})$$

$$= \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)dx = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{3}{2}x^2 dx = \left(\frac{1}{2}x^3\right)_{-\sqrt{y}}^{\sqrt{y}} = y^{3/2}.$$

$$f_X(x) = \begin{cases} 3x^2/2, -1 < x < 1, \\ 0, \quad \text{其他}. \end{cases}$$

$$f_Y(y) = F_Y'(y) = y^{3/2} = \frac{3}{2}\sqrt{y}.$$

$$f_Y(y) = \begin{cases} \frac{3}{2}\sqrt{y}, \ 0 < y < 1, \\ 0, \quad \text{其他}. \end{cases}$$

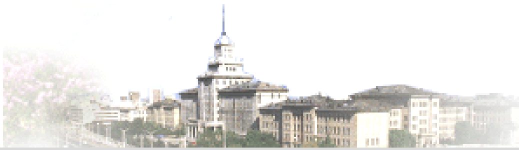


连续型随机变量函数概率密度的两种求法

- **公式法**
- **定理1** 设X的概率密度$f_X(x)$, $y=g(x)$为(a,b)上严格单调可微函数$(-\infty \le a < b \le +\infty)$,则$Y=g(X)$的概率密度为

$$f_Y(y)=\begin{cases} f_X(h(y))\,|h'(y)|, & A<y<B, \\ 0, & 其他. \end{cases}$$

其中$h(y)$为$g(x)$的反函数且$A=\min\{g(a),g(b)\}$,
$B=\max\{g(a),g(b)\}$.



例4设$X \sim N(\mu,\sigma^2)$,求$Y = aX + b$(a,b是常数,$a \neq 0$)概率密度.

解 $y = ax + b,(a \neq 0)$在$(-\infty, +\infty)$严格单调可微,

反函数为$x = h(y) = (y-b)/a, h'(y) = 1/a.$

$$\boxed{f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}}$$

$$f_Y(y) = f_X(h(y))|h'(y)|$$

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

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

即，$Y \sim N(a\mu + b, a^2\sigma^2)$.

结论 若$X \sim N(\mu,\sigma^2), a \neq 0$,则

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

例如，$X \sim N(1,2)$,则$Y = 3X - 4 \sim N(-1,18)$.

$$a\mu + b = 3 \cdot 1 + (-4) = -1$$

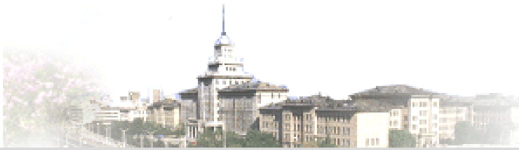
$$a^2\sigma^2 = 3^2 \cdot 2 = 18$$

连续型随机变量函数概率密度的两种求法

- 公式法
- 定理2　设X的概率密度$f_X(x)$, $y=g(x)$在不相交的区间$I_1, I_2, \cdots$上逐段严格单调，其反函数分别为$h_1(y), h_2(y), \cdots$，且$h_1'(y), h_2'(y), \cdots$均连续，则$Y=g(X)$的概率密度为

$$f_Y(y)=\sum_{i=1}^{n} f_X(h_i(y))\,|h_i'(y)|$$

练习



1\. 设X的密度$f_X(x)=\begin{cases} x/8, & 0<x<4 \\ 0, & 其它 \end{cases}$，求$Y=2X+8$的密度$f_Y(y)$.

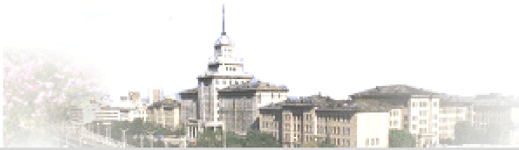
解 设Y的分布函数为 $F_Y(y)$，则 $P(0<X<4)=1 \Rightarrow P(8<Y<16)=1$

当$y\le 8$, $F_Y(y)=0$; 当$y\ge 16$, $F_Y(y)=1$;

当$8<y<16$时，

$$F_Y(y)=P(Y\le y)=P(2X+8\le y)=P(X\le \frac{y-8}{2})=F_X\left(\frac{y-8}{2}\right)$$

$$f_Y(y)=F_Y'(y)=f_X(\frac{y-8}{2})\cdot\frac{1}{2}=\begin{cases} \dfrac{y-8}{32}, & 8<y<16, \\ 0, & 其他. \end{cases}$$



2. 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.

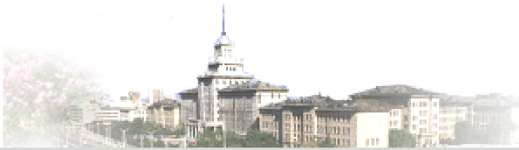
解 由于

$$P(0<X<\pi)=1 \Rightarrow P(0<Y\le 1)=1$$

所以

$$\text{当 } y\le 0,\ F_Y(y)=0, \Rightarrow\ f_Y(y)=0;$$

$$\text{当 } y>1,\ F_Y(y)=1, \Rightarrow\ f_Y(y)=0;$$



当$0 < y \le 1$时，$F_Y(y) = P(Y \le y) = P(\sin X \le y)$

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

$$= F_X(\arcsin y) - F_X(0) + F_X(\pi) - F_X(\pi - \arcsin y)$$

$$\Rightarrow f_Y(y) = f_X(\arcsin y)\frac{1}{\sqrt{1-y^2}} + f_X(\pi - \arcsin y)\frac{1}{\sqrt{1-y^2}}$$

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

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

方法2 当$0<y\le 1$时，$F_Y(y)=P(Y\le y)=P(\sin X\le y)$

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

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

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

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

3. 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.

证明　设Y的分布函数是$G(y)$,

由于$0 \le y \le 1$，故

当$y < 0$时，$G(y) = 0$, 当$y > 1$时，$G(y) = 1$,

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.

当$0 \le y \le 1$时，

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

$$= F(F^{-1}(y)) = y$$

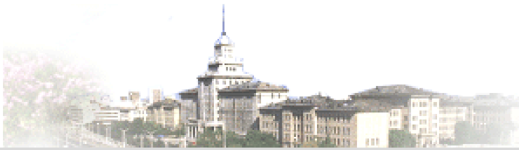
即Y的分布函数是 $G(y) = \begin{cases} 0, & y < 0 \\ y, & 0 \le y \le 1 \\ 1, & y > 1 \end{cases}$

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 服从[0，1]上的均匀分布.

本例的结论在计算机模拟中有重要的应用



例如，想得到具有密度函数为

$$f(x)=\begin{cases}\lambda e^{-\lambda x} & x>0\\ 0 & x\le 0\end{cases},\lambda>0$$

参数为 λ 的指数分布

的随机数应如何做呢？

由于 当x≥0时，$F(x)=1-e^{-\lambda x}$ 是严格单调的连续函数．

根据前面的结论， $Y=F(X)$服从[0,1]上的均匀分布.

于是得到产生指数分布的随机数的方法

均匀随机数 u_i

↓

给指数分布参数λ

↓

令 $x_i = -\dfrac{1}{\lambda}\ln(1-u_i)$

↓

指数随机数 → x_i

第三章内容总框图

随机变量及其分布

连续型随机变量及其概率密度

随机变量概念的引入

离散型随机变量及其概率分布列

概率密度与分布函数的关系

随机变量的分布函数

分布列与分布函数的关系

几种重要的连续型分布

随机变量函数的分布

几种重要的 离散型分布

二项分布的正态近似

均匀分布

指数分布

正态分布

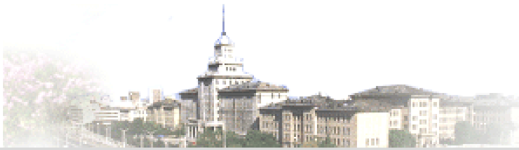
二项分布的泊松近似

二项分布

泊松分布

几何分布

超几何分布

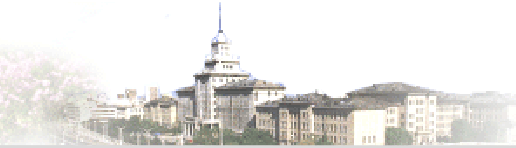


练习题

设随机变量的概率密度为 $f(x)=\begin{cases}\frac{1}{4}x^2 & 0<x<3\\ 0 & 其他\end{cases}$，令随机变量 $Y=\begin{cases}2 & x\le 1\\ x & 1<x<2\\ 1 & x\ge 2\end{cases}$

（1） 求 Y 的分布函数

（2） 求概率 $P(X\le Y)$



谢　谢！