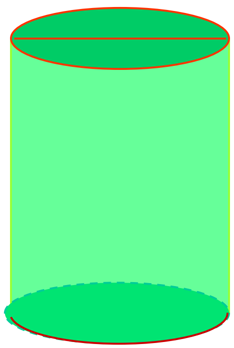
第二章 随机变量及其分布

§2.5 随机变量的函数的分布

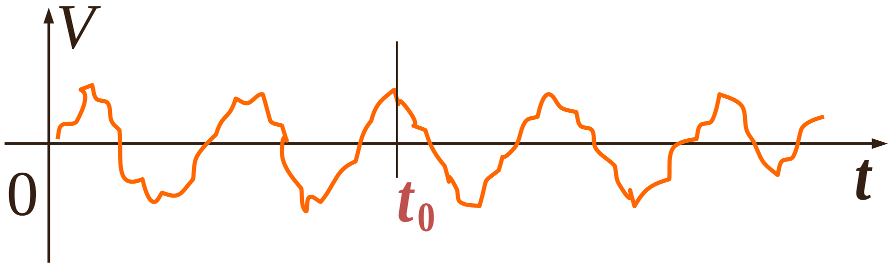
问题的提出：



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.

再如, 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.

在实际中，人们常常对随机变量 X 的函数$Y=g(X)$ (设g是连续函数)所表示的随机变量Y感兴趣

在分析和解决实际问题时，常常会遇到一些随机变量，它们的分布难于直接得到，但其与一些已知随机变量之间具有函数关系. 本节主要解决如何由随机变量X的概率分布求出随机变量$Y=g(X)$的概率分布.

对于随机变量X的函数的分布的讨论分两部分

一、离散型随机变量函数的分布律

二、连续型随机变量函数的概率密度

一、离散型随机变量函数的分布律

例1 设随机变量X的分布律为

X	-1	0	1	2	3
p_k	1/5	1/10	1/10	3/10	3/10

求 (1) $Y=X-1$ (2) $Y=-2X$ (3) $Y=X^2$ 分布律。

解:

X	-1	0	1	2	3
$X-1$	-2	-1	0	1	2
$-2X$	2	0	-2	-4	-6
X^2	1	0	1	4	9
p_k	1/5	1/10	1/10	3/10	3/10

一、离散型随机变量函数的分布律

例1 设随机变量X的分布律为

X	-1	0	1	2	3
p_k	$1/5$	$1/10$	$1/10$	$3/10$	$3/10$

求 (1) $Y=X-1$ (2) $Y=-2X$ (3) $Y=X^2$ 分布律。

解：

X	-1	0	1	2	3
$X-1$	-2	-1	0	1	2
$-2X$	2	0	-2	-4	-6
X^2		0	1	4	9
p_k		$1/10$	$3/10$	$3/10$	$3/10$

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

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

注意 若 $g(x_k)$ 中有值相同的，应将相应的 p_k 合并.

二、连续型随机变量函数的概率密度

对于连续型随机变量，需要由随机变量X的概率密度 $f_X(x)$ 去求随机变量 $Y=g(X)$ 的概率密度. 解决这类问题的一般方法是: ------分布函数法

第一步：求出Y的分布函数的表达式，

$$F_Y(y) = P\{Y \le y\} = P\{g(X) \le y\} = P\{X \in C_y\}$$

其中$C_y = \{x \mid g(x) \le y\}$

第二步：利用连续型随机变量分布函数与概率密度的关系，求导数即可得到所求概率密度.

例2 设随机变量X具有概率密度 $f_X(x)=\begin{cases}\frac{x}{8}, & 0<x<4\\ 0, & 其它\end{cases}$

求 Y=2X+8 的概率密度.

解 先求Y=2X+8的分布函数 $F_Y(y)$

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

于是，得 Y=2X+8的 概率密度为

$$f_Y(y)=f_X\left(\frac{y-8}{2}\right)\left(\frac{y-8}{2}\right)'=\begin{cases}\frac{1}{8}\left(\frac{y-8}{2}\right)\cdot\frac{1}{2}, & 0<\frac{y-8}{2}<4\\ 0, & 其它\end{cases}$$

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

例2 设随机变量X具有概率密度 $f_X(x)=\begin{cases}\dfrac{x}{8}, & 0<x<4\\ 0, & \text{其它}\end{cases}$

求 $Y=2X+8$ 的概率密度.

解 先求$Y=2X+8$的分布函数 $F_Y(y)$

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

于是，得 $Y=2X+8$的 概率密度为

$$f_Y(y)=f_X\left(\frac{y-8}{2}\right)\left(\frac{y-8}{2}\right)'=\begin{cases}\dfrac{1}{8}\left(\dfrac{y-8}{2}\right)\cdot\dfrac{1}{2}, & 0<\dfrac{y-8}{2}<4\end{cases}$$

特别 若 $X\sim N(\mu,\sigma^2)$，则 $aX+b\sim N(a\mu+b,a^2\sigma^2)$

$$a=\frac{1}{\sigma},b=-\frac{\mu}{\sigma}\Rightarrow\frac{X-\mu}{\sigma}\sim N(0,1)$$

例3 设随机变量X具有概率密度 $f_X(x)$，$-\infty < x < \infty$
求Y=X^2的概率密度.

解 先求Y的分布函数 $F_Y(y)$. 因为 $Y = X^2 \geq 0$

故当y≤0时， $F_Y(y) = P\{Y \leq y\} = 0$

当 y>0 时，$F_Y(y) = P\{X^2 \leq y\} = P\{-\sqrt{y} \leq X \leq \sqrt{y}\}$

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

于是，得Y的概率密度为

$$f_Y(y) = \begin{cases} \dfrac{1}{2\sqrt{y}}[f_X(\sqrt{y}) + f_X(-\sqrt{y})] \ , & y > 0 \\ 0 \ , & y \leq 0 \end{cases}$$

例如，设$X \sim N(0, 1)$，其概率密度为

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

则$Y=X^2$的概率密度为

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

注：(1) 此时称 Y 服从自由度为 1 的 χ^2分布;

(2) 若Y=g(X)中的g(.)是严格单调函数时，可由下面定理求出Y的概率密度.

定理 设随机变量X具有概率密度$f_X(x)$，又设函数$g(x)$处处可导且有$g'(x)>0$(或恒有$g'(x)<0$)，则$Y=g(X)$是连续型随机变量,其概率密度为

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

其中 $\alpha=\min\left(g(-\infty),g(\infty)\right),\beta=\max\left(g(-\infty),g(\infty)\right)$

$h(y)$是$g(x)$的反函数.

注: (1)若$g(x)$不是单调函数不能用此定理.

(2)若$f(x)$在有限区间$[a, b]$以外等于零，则只需假设在$[a,b]$上恒有$g'(x)>0$(或恒有$g'(x)<0$)，此时

$$\alpha=\min\{g(a),g(b)\},\beta=\max\{g(a),g(b)\}$$

证 我们只证 $g'(x) > 0$ 的情况.

此时 $g(x)$ 在 $(-\infty, \infty)$ 严格单调增加，其反函数 $h(y)$ 存在，且在 (α, β) 严格单调增加，可导. 分别记 X, Y 的分布函数为 $F_X(x), F_Y(y)$. 现在先求 Y 的分布函数 $F_Y(y)$.

因为 $Y = g(X)$ 在 (α, β) 内取值，

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

当 $y \ge \beta$ 时，$F_Y(y) = P\{Y \le y\} = 1$.

当 $\alpha < y < \beta$ 时，

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

$$= P\{X \le h(y)\} = F[h(y)].$$

将 $F_Y(Y)$ 关于 y 求导数，即得 Y 的概率密度

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

对于 $g'(x) < 0$ 的情况可以同样地证明，此时有

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

合并以上两式,定理得证.

补充定理：

若$\boldsymbol{g(x)}$在不相叠的区间 $I_1, I_2, \cdots$ 上逐段严格单调且处处可导，其反函数分别为 $h_1(y), h_2(y), \cdots$ 且均为连续函数，那么$\boldsymbol{Y = g(X)}$是连续型随机变量，其概率密度为

$$f_Y(y) = f_X(h_1(y))\left|h_1^{'}(y)\right| + f_X(h_2(y))\left|h_2^{'}(y)\right| + \cdots$$

例4 设电压 $V=A\sin\Theta$, 其中A是一个已知的正常数, 相角Θ是一个随机变量, 在区间$(-\pi/2, \pi/2)$服从均匀分布, 试求电压 V的概率密度.

解: Θ的概率密度为 $f(\theta)=\begin{cases}\dfrac{1}{\pi}, & -\dfrac{\pi}{2}<\theta<\dfrac{\pi}{2}\\ 0, & \text{其它}\end{cases}$

由于 $v=g(\theta)=A\sin\theta$, $g'(\theta)=A\cos\theta>0$, $\left(-\dfrac{\pi}{2},\dfrac{\pi}{2}\right)$

其反函数 $\theta=h(v)=\arcsin\dfrac{v}{A}, h'(v)=\dfrac{1}{\sqrt{A^2-v^2}}$

由定理得 $V=A\sin\Theta$的概率密度为

$$\psi(v)=\begin{cases}\dfrac{1}{\pi}\cdot\dfrac{1}{\sqrt{A^2-v^2}}, & -A<v<A\\ 0, & \text{其它}\end{cases}$$

例4 设电压 $V=A\sin\Theta$, 其中A是一个已知的正常数, 相角Θ是一个随机变量, 在区间$(-\pi/2, \pi/2)$服从均匀分布, 试求电压 V的概率密度.

解: Θ的概率密度为 $f(\theta)=\begin{cases}\dfrac{1}{\pi}, & -\dfrac{\pi}{2}<\theta<\dfrac{\pi}{2}\\ 0, & 其它\end{cases}$

由于 $v=g(\theta)=A\sin\theta$, $g'(\theta)=A\cos\theta>0$, $\left(-\dfrac{\pi}{2},\dfrac{\pi}{2}\right)$

其反函数 $\theta=h(v)=\arcsin\dfrac{v}{A}, h'(v)=\dfrac{1}{\sqrt{A^2-v^2}}$

注意

若$\Theta\sim U(0,\pi)$, 此时$v=g(\theta)=A\sin\theta$在$(0,\pi)$上不是单调函数.

练习题（正态分布）

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

解: X的概率函数为

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

现在$y=g(x)=ax+b$, 由这一式子解得

$$x = h(y) = \frac{y-b}{a}，\text{ 且有 } h'(y) = \frac{1}{a}$$

由定理得Y=aX+b的概率密度为

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

即

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

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

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

注: (1) 正态随机变量的线性函数仍然服从正态分布.

(2) 若 $a = \frac{1}{\sigma}, b = -\frac{\mu}{\sigma}$

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

2.某厂装配车间准备实行计件超产奖，为此需对生产定额做出规定.据以往纪录，工人每月装配产品数服从正态分布$N(4000,3600)$.假定车间主任希望10%的工人获得超产奖，求工人每月需完成多少件产品才能获奖？

解：设需完成件n件产品才能获奖,

记工人每月装配产品数为$X \sim N(4000,3600)$.

则依题意，有 $P\{X>n\}=10\%$,

即 $90\%=P\{X\le n\}=P\{\frac{X-4000}{60}\le\frac{n-4000}{60}\}=\Phi(\frac{n-4000}{60})$

查表得 $\frac{n-4000}{60}=1.28$, 解得 $n=4000+76.8\approx 4077$.

故每月需完成4077件以上才能获奖.

小结

1. 离散型随机变量的函数的分布

若 $Y = g(X)$ 且 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 $Y = g(X)$ 的分布律为

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

注意 若 $g(x_k)$ 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 $F_Y(y) = P\{Y \le y\} = P\{g(X) \le y\}$

$$= \int_{g(x)\le y} f_X(x)\,\mathrm{d}x, \quad (-\infty < y < +\infty),$$

再对 $F_Y(y)$ 求导得到 Y 的密度函数 $f_Y(y)$.

方法2

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

注意条件.

作业

- **第二章习题33，35**