§5　随机变量的函数的分布

- 离散型
- 连续型
- 定理及其应用

返回主目录

随机变量的函数

设 X 是一随机变量，Y 是 X 的函数，$Y=g(X)$ ，则 Y 也是一个随机变量．当 X 取值 x 时，Y 取值 $y=g(x)$

本节的任务就是：

已知随机变量 X 的分布，并且已知 $Y=g(X)$，要求随机变量 Y 的分布．

一、离散型随机变量的函数

设 X 是离散型随机变量，其分布律为

$$P\{X = x_n\} = p_n \qquad (n=1,\ 2,\ \cdots)$$

或

X	x_1	x_2	$\cdots,$	x_n	$\cdots$
P	p_1	p_2	$\cdots,$	p_n	$\cdots$

Y是X的函数：$Y=g(X)$，则Y也是离散型随机变量，它的取值为

$$y_1,\quad y_2,\quad \cdots,\quad y_n,\quad \cdots$$

其中 $y_n = g(x_n) \qquad (n=1,\ 2,\ \cdots)$

返回主目录

第 一 种 情 形

如果 $y_1, \quad y_2, \quad \cdots, \quad y_n, \quad \cdots$

两两不相同，则由

$$P\{Y = y_n\} = P\{X = x_n\} \qquad (n = 1, \quad 2, \quad \cdots)$$

可知随机变量Y的分布律为

$$P\{Y = y_n\} = p_n \qquad (n = 1, \quad 2, \quad \cdots)$$

或

Y	y_1	y_2	$\cdots,$	y_n	$\cdots$
P	p_1	p_2	$\cdots,$	p_n	$\cdots$

第 二 种 情 形

如果 y_1， y_2， $\cdots$， y_n， $\cdots$ 有相同的项，则把这些相同的项合并（看作是一项），并把相应的概率相加，即可得随机变量$Y = g(X)$的分布律.

例 1

设离散型随机变量 X 的分布律为

X	-3	-1	0	2	6	9
P	$\frac{1}{252}$	$\frac{5}{252}$	$\frac{15}{252}$	$\frac{35}{252}$	$\frac{70}{252}$	$\frac{126}{252}$

随机变量 $Y=2X-3$，试求 Y 的分布律.

解：

随机变量 $Y=2X-3$ 的取值为

$$-9,\quad -5,\quad -3,\quad 1,\quad 9,\quad 15,$$

例 1（续）

这些取值两两互不相同．　由此得随机变量

$$Y = 2X - 3$$

的分布律为

Y	-9	-5	-3	1	9	15
P	$\frac{1}{252}$	$\frac{5}{252}$	$\frac{15}{252}$	$\frac{35}{252}$	$\frac{70}{252}$	$\frac{126}{252}$

例 2

设随机变量 X 具有以下的分布律，试求

$$Y=(X-1)^2$$

的分布律.

X	-1	0	1	2
p_k	0.2	0.3	0.1	0.4

解: Y 有可能取的值为 0，1，4.

且 $Y=0$ 对应于 $(X-1)^2=0$，解得 $X=1$，

所以，$P\{Y=0\}=P\{X=1\}=0.1$，

例 2（续）

$Y=(X-1)^2$

X	-1	0	1	2
p_k	0.2	0.3	0.1	0.4

同理,

$P\{Y=1\}=P\{X=0\}+P\{X=2\}=0.3+0.4=0.7$,

$P\{Y=4\}=P\{X=-1\}=0.2$,

所以，$Y=(X-1)^2$ 的分布律为：

Y	0	1	4
p_k	0.1	0.7	0.2

返回主目录

例 3

设离散型随机变量 X 的分布律为

X	1	2	$\cdots$	n	$\cdots$
P	$\frac{1}{2}$	$\frac{1}{2^2}$	$\cdots$	$\frac{1}{2^n}$	$\cdots$

$$Y = g(X) = \begin{cases} -1 & \text{若}X\text{为奇数} \\ 1 & \text{若}X\text{为偶数} \end{cases}$$

试求随机变量 Y 的分布律.

解:

例 3（续）

$$P\{Y=-1\}=\sum_{n\text{为奇数}}P\{X=n\}=\sum_{k=0}^{\infty}P\{X=2k+1\}$$

$$=\sum_{k=0}^{\infty}\frac{1}{2^{2k+1}}=\frac{2}{3}$$

$$P\{Y=1\}=\sum_{n\text{为偶数}}P\{X=n\}=\sum_{k=1}^{\infty}P\{X=2k\}$$

$$=\sum_{k=1}^{\infty}\frac{1}{2^{2k}}=\frac{1}{3}$$

所以，随机变量Y的分布律为

Y	-1	1
P	$\frac{2}{3}$	$\frac{1}{3}$

二. 连续型随机变量函数的分布

设 X 是一连续型随机变量，其密度函数为 $f_X(x)$，再设 $Y=g(X)$ 是 X 的函数，我们假定 Y 也是连续型随机变量.我们要求的是 $Y=g(X)$ 的密度函数 $f_Y(y)$.

解题思路

(1). 先求 $Y=g(X)$ 的分布函数

$$F_Y(y)=P\{Y\leq y\}=P\{g(X)\leq y\}=\int_{g(x)\leq y} f_X(x)dx$$

(2). 利用 $Y=g(X)$ 的分布函数与密度函数之间的关系求 $Y=g(X)$ 的密度函数 $f_Y(y)=F_Y'(y)$

例 4

设随机变量 X 具有概率密度：

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

试求 $Y=2X+8$ 的概率密度.

解： (1) 先求 $Y=2X+8$ 的分布函数 $F_Y(y)$：

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

$$=P\{2X+8\le y\}=P\{X\le\frac{y-8}{2}\}$$

返回主目录

例 4（续）

$$F_Y(y)=\int_{-\infty}^{\frac{y-8}{2}} f_X(x)dx.$$

(2) 利用 $F_Y'(y)=f_Y(y)$ 可以求得:

$$f_Y(y)=f_X\left(\frac{y-8}{2}\right)\times\left(\frac{y-8}{2}\right)'$$

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

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

返回主目录

例 4（续）

整理得 $Y=2X+8$ 的概率密度为：

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

本例用到变限的定积分的求导公式

如果 $F(x)=\int_{\psi(x)}^{\varphi(x)} f(t)dt,$

则 $F'(x)=f[\varphi(x)]\varphi'(x)-f[\psi(x)]\psi'(x).$

例 5

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：(1) 先求 $Y = X^2$ 的分布函数 $F_Y(y)$:

1^0　由于 $Y = X^2 \geq 0$, 故当 $y \leq 0$ 时 $F_Y(y) = 0$.

2^0　当 $y > 0$ 时,

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

$$= P\{-\sqrt{y} \leq X \leq \sqrt{y}\} = \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)dx.$$

返回主目录

例 5（续）

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

(2)利用 $F_Y'(y)=f_Y(y)$及变限定积分求导公式得：

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

例如,设 $X \sim N(0,1)$,其概率密度为:

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

则 $Y=X^2$ 的概率密度为:

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

此时称 Y 服从自由度为1的 χ^2分布。

返回主目录

例 6

设随机变量 X 的密度函数为 $f_X(x)$，$Y=|X|$，试求随机变量Y的密度函数 $f_Y(y)$.

解：

设随机变量 X 的分布函数为$F_X(y)$，随机变量Y的分布函数为$F_Y(y)$

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

(1). 若 $y<0$，则

$$F_Y(y)=P\{Y\le y\}=P\{|X|\le y\}=P(\Phi)=0$$

返回主目录

例 6（续）

(2). 若 $y \ge 0$，则

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

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

综上所述，得随机变量 Y 的分布函数为

$$F_Y(y) = \begin{cases} F_X(y) - F_X(-y) & y \ge 0 \\ 0 & y < 0 \end{cases}$$

对上式求导，可得 $Y = |X|$ 的密度函数为

$$f_Y(y) = \begin{cases} f_X(y) + f_X(-y) & y \ge 0 \\ 0 & y < 0 \end{cases}$$

返回主目录

定理

设随机变量 X 具有概率密度 $f_X(x)$，$-\infty < x < \infty$，

又设函数 $g(x)$ 处处可导，且有 $g'(x) > 0$ (或恒有 $g'(x) < 0$).

则 $Y = g(X)$ 是一个连续型随机变量 Y，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 $h(y)$ 是 $g(x)$ 的反函数，$\alpha = \min\{g(-\infty), g(\infty)\}$,

即 $x = g^{-1}(y) = h(y)$　　$\beta = \max\{g(-\infty), g(\infty)\}$.

定理（续）

若 $f(x)$ 在有限区间 $[a,b]$ 以外等于零，则只须假设在 $[a,b]$ 上恒有 $g'(x)>0$（或恒有 $g'(x)<0$），此时仍有：

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

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

证明：

设随机变量$Y=g(X)$的分布函数为$F_Y(y)$，

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

由题设，不妨假设$g'(x)>0$，则$g(x)$是严格增加的函数.

因此，$$F_Y(y)=P\left\{X\le g^{-1}(y)\right\}=P\{X\le h(y)\}$$

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

定 理 的 证 明

由题设，当随机变量 X 在区间 $(-\infty, +\infty)$ 上变化时，随机变量 Y 在区间 (α, β) 上变化.

其中，

$$\alpha = \min\{g(-\infty), g(+\infty)\}, \quad \beta = \max\{g(-\infty), g(+\infty)\}$$

因此，当 $y \in (\alpha, \beta)$ 时，

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

所以，$f(y) = F_Y'(y) = \dfrac{d}{dy}\left(\int_{-\infty}^{h(y)} f_X(x)dx\right)$

返回主目录

定 理 的 证 明

$$= f_X(h(y)) \cdot h'(y) = f_X(h(y)) \cdot |h'(y)|$$

若 $g'(x) < 0$，则 $g(x)$ 是严格减少的函数.

因此，　当 $y \in (\alpha,\ \beta)$ 时，

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

$$= P\{X \ge g^{-1}(y)\} = P\{X \ge h(y)\} = \int_{h(y)}^{+\infty} f_X(x)dx$$

所以，$f(y) = F_Y'(y) = \dfrac{d}{dy}\left(\int_{h(y)}^{+\infty} f_X(x)dx\right)$

定 理 的 证 明

$$= -f_X\left(h(y)\right)\cdot h'(y) \quad = f_X\left(h(y)\right)\cdot\left|h'(y)\right|$$

综上所述，得 $Y=g(X)$ 的密度函数为

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

返回主目录

补充定理：

若g(x)在不相叠的区间 $I_1, I_2, \cdots$ 上逐段严格单调，其反函数分别为 $h_1(y), h_2(y), \cdots$ 均为连续函数，那么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 7

设随机变量 $X \sim N(\mu,\ \sigma^2)$, $Y = e^X$，试求随机变量 Y 的密度函数 $f_Y(y)$.

解：

由题设，知 X 的密度函数为

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad (-\infty < x < +\infty)$$

因为函数 $y = e^x$ 是严格增加的，它的反函数为 $x = \ln y$.

例 7（续）

并且当随机变量 X 在区间 $(-\infty, \ +\infty)$ 上变化时，$Y=e^X$ 在区间 $(0, \ +\infty)$ 上变化．所以，当 $y\in(0, \ +\infty)$ 时，

$$f_Y(y)=f_X(\ln y)\cdot\left|(\ln y)'\right|=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(\ln y-\mu)^2}{2\sigma^2}\right\}\cdot\frac{1}{y}$$

由此得随机变量 $Y=e^X$ 的密度函数为

$$f_Y(y)=\begin{cases}\dfrac{1}{\sqrt{2\pi}y\sigma}\exp\left\{-\dfrac{(\ln y-\mu)^2}{2\sigma^2}\right\} & y>0\\ 0 & y\le 0\end{cases}$$

例 8

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 X的概率密度为:

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

$y = g(x) = ax + b, g'(x) = a,$满足定理的条件,

$y = g(x)$的反函数为: $x = h(y) = \dfrac{y-b}{a}$,且$h'(y) = \dfrac{1}{a}$.

例 8(续)

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

由定理的结论得： $x=h(y)=\dfrac{y-b}{a}$,且$h'(y)=\dfrac{1}{a}$.

$$f_Y(y)=f_X[h(y)]|h'(y)|=\frac{1}{|a|}f_X(\frac{y-b}{a})$$

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

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

返回主目录

例 9

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$\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$上服从

均匀分布，试求电压V的概率密度.

解:

$$v=g(\theta)=A\sin\theta,$$ 在（$-\frac{\pi}{2},\frac{\pi}{2}$）上恒有

$$g'(x)=A\cos\theta>0,$$ 且有反函数 $\theta=h(v)=\arcsin\frac{v}{A}$,

以及 $$h'(v)=\frac{1}{\sqrt{A^2-v^2}},$$

返回主目录

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

利用定理的结论: $f_Y(y)=\begin{cases}f_X[h(y)]\,|h'(y)|, & \alpha<y<\beta,\\ 0, & 其它.\end{cases}$

$$\boxed{h'(v)=\frac{1}{\sqrt{A^2-v^2}},}$$

得$V=A\sin\Theta$的概率密度为:

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

返回主目录

第二章　小　结

1 引进了随机变量的概念，要求会用随机变量表示随机事件。
2 给出了分布函数的定义及性质，要会利用分布函数示事件的概率。
3 给出了离散型随机变量及其分布率的定义、性质，要会求离散型随机变量的分布率及分布函数，掌握常用的离散型随机变量分布：两点分布、二项分布、泊松分布。
4 给出了连续型随机变量及概率密度的定义、性质，要掌握概率密度与分布函数之间关系及其运算，掌握常用的连续型随机变量分布：均匀分布、指数分布和正态分布。
5 会求随机变量的简单函数的分布。