概率论与数理统计

第十二讲

随机变量函数的分布

第12讲　随机变量函数的分布

有时我们关心的随机变量不是直接观测得到的随机变量，而是它的函数.

例 （电功率）设某供电线路上电流值 X 为一随机变量, 其分布函数为$F_X(x)$. 若线路上有一电阻 R, 试求 R 上的电功率 Y 的分布函数 $F_Y(y)$.

- 分析 显然有$Y=RX^2$, 且 Y 取值为非负的, 由分布函数定义, 对于$y \geq 0$, $F_Y(y) = P\{Y \leq y\}$

$$= P\{RX^2 \leq y\} = P\{-\sqrt{\frac{y}{R}} \leq X \leq \sqrt{\frac{y}{R}}\}$$

$$= F_X(\sqrt{\frac{y}{R}}) - F_X(-\sqrt{\frac{y}{R}} - 0)$$

一般地，需要研究：设 X 为一随机变量，分布已知. $Y=g(X)$，其中 g 为一确定的实函数，要求 Y 的分布.

仍讨论离散型与连续型两种情况：

离散型：Y 的分布律

连续型：Y 的密度函数

1. 离散型情形

设随机变量 X 的分布律如下：

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

$Y=g(X)$, 则只须

1. 逐个算出 Y 的取值 $g(x_1), g(x_2), \ldots,$ 每个 $g(x_k)$ 对应的概率为 p_k;
2. 合并所有相同的 $g(x_k)$，并将对应的 p_k 相加.

例　设 X 的分布律为：

X	0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$	1
p_k	0.1	0.2	0.4	0.2	0.1

设 $Y=\sin^2(\pi X)$, 求 Y 的分布律.

计算

Y	0	$\frac{1}{2}$	1	$\frac{1}{2}$	0
p_k	0.1	0.2	0.4	0.2	0.1

合并

Y	0	$\frac{1}{2}$	1
p_k	0.2	0.4	0.4

有时X是连续型的, Y仍有可能是离散的.

例　(儿童智商)设儿童智商$X \sim N(100,\ 100)$, 将儿童按智商分为3类, 类标号 Y 规定如下:

$$Y=\begin{cases}1, & X>110\\ 0, & 90<X\le 110\\ -1, & X\le 90\end{cases}$$

求 Y 的分布律.

答:

Y	-1	0	1
p_k	0.16	0.68	0.16

2. 连续型情形

设 X 为连续型随机变量, 具有密度函数 $f_X(x)$, 又设 $Y=g(X)$亦为连续型随机变量, 求其密度函数 $f_Y(y)$.

步骤：

1) 从分布函数着手

$$F_Y(y)=P\{Y\le y\}=P\{g(X)\le y\}=\int_{\{x:\,g(x)\le y\}} f_X(x)\mathrm{d}x$$

其中积分区域 $\{x: g(x)\le y\}$ 表示满足 $g(x)\le y$ 的 x 的点集.

进一步化简的关键：寻求上述集合的 x 的显式表达.

2) 对分布函数求导, 得到概率密度函数.

例（电功率续）前述问题中，设X有密度函数$f_X(x)$.
又为简单计算, 令$R=1$, 即求$Y=X^2$的密度函数.

已经求得 $F_Y(y) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$

求导得 $f_Y(y) = \dfrac{1}{2\sqrt{y}}(f_X(\sqrt{y})+f_X(-\sqrt{y}))$

或者直接从对积分限求导法则得到, 即对下式求导

$$F_Y(y) = \int_{\{x:\, x^2 \le y\}} f_X(x)\mathrm{d}x = \int_{-\sqrt{y}}^{\sqrt{y}} f_X(x)\mathrm{d}x$$

例　按上例，若$X \sim N(0,1)$，则

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

今后知道Y服从自由度为1的χ^2分布.

定理 设$Y=g(X)$是X的严格单调函数，有反函数$X=h(Y)$，且$h(Y)$可微，则有

$$f_Y(y)=\begin{cases}f(h(y))\,|h'(y)|, & y\text{在}Y\text{的取值范围内}\\ 0, & \text{其他}\end{cases}$$

证明提示:不妨设$g(x)$严格增, 则$h(y)$严格增, 此时

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

求导, 并注意到 $0<h'(y)=|h'(y)|$.

- 问题 若$g(x)$严格减, 上述证明要做什么修正?

例 (线性函数) 设 X 的密度函数为 $f_X(x)$,

$Y=aX+b,\ a\neq 0$, 求 $f_Y(y)$.

答: $f_Y(y)=\dfrac{1}{|a|}f_X(\dfrac{y-b}{a})$

特别, 若 $f_X(x)=\dfrac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$, $a=\sigma, b=\mu$,

则 $f_Y(y)=\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

即知: 1. 任何正态分布都可表成标准正态分布的线性函数;

2. 正态分布的线性函数仍为正态分布.

- 注 在形如 $\dfrac{1}{|a|}f_X(\dfrac{y-b}{a})$ 的密度函数中, b称为位置参数, a称为尺度参数.

例（股票价格）考虑时间u后股票价格S_u，已知 $S_u=S_0e^{X_u}$，而$X_u\sim N(u\mu,u\sigma^2)$，$S_0$为常数，求$S_u$的密度函数$f_S(s)$.

解: $h(S)=\ln\dfrac{S}{S_0}$，$h'(S)=\dfrac{1}{S}$，故

$$f_S(s)=\frac{1}{\sqrt{2\pi u\sigma^2}\,s}e^{-\frac{(\ln\frac{s}{S_0}-u\mu)^2}{2u\sigma^2}}$$

- 注 $\dfrac{S_u}{S_0}$ 称服从参数为$(u\mu,u\sigma^2)$的对数正态分布，记为$LN(u\mu,u\sigma^2)$.

定理的推广

X, Y 同前, 若连续函数 $g(x)$ 在 $(-\infty, b_1), (b_1, b_2), \cdots, (b_k, +\infty)$ 分段严格单调, 且$g(x)$在上述小区间内均有可微的反函数 $h_1(y), h_2(y), \cdots, h_k(y)$, 则

$$f_Y(y) = \begin{cases} \sum\limits_{i=1}^{k} f(h_i(y))\,|h_i'(y)|, & y\text{在}Y\text{的取值范围内} \\ 0, & \text{其他} \end{cases}$$

例（电功率 问题, 续）利用上述公式重做此例 (设$R=1$), 则$g(x)=x^2$在 $(-\infty, 0), (0, +\infty)$ 分别严格单调,

$$h_1(y) = -\sqrt{y}, h_2(y) = \sqrt{y}, \text{ 故} |h_1'(y)| = |h_2'(y)| = \frac{1}{2\sqrt{y}},$$

$$\text{故 } f_Y(y) = \frac{1}{2\sqrt{y}}(f_X(\sqrt{y}) + f_X(-\sqrt{y})).$$