# 随机变量的函数及其分布

$X$为概率空间$(\Omega,\mathcal{F},P)$上的随机变量，$g(x)$为(分段)连续函数或(分段)单调函数，则$Y=g(X)$也是$(\Omega,\mathcal{F},P)$上的随机变量，称它为随机变量$X$的函数。

## 问题

已知$X$的分布和函数$g(x)$，求它的函数$Y=g(X)$的分布

# 离散型随机变量的函数的分布

若$X$为离散型随机变量, 则$Y = g(X)$也是离散型随机变量

## 问题

已知$X$的分布率和函数$g(x)$，求它的函数$Y = g(X)$的分布率

# 离散型随机变量的函数的分布

例1 设$X$的分布律为

| $X$ | $-1$ | $0$ | $1$ | $2$ |
|---|---|---|---|---|
| $P$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ |

求$Y=X^2$的分布律

解:

| $X$ | -1 | 0 | 1 | 2 |
|---|---|---|---|---|
| $Y=X^2$ | 1 | 0 | 1 | 4 |
| $P$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ | $\frac{1}{4}$ |

$\Rightarrow$

| $Y$ | $0$ | $1$ | $4$ |
|---|---|---|---|
| $P$ | $\frac{1}{4}$ | $\frac{1}{2}$ | $\frac{1}{4}$ |

# 离散型随机变量的函数的分布

| $X$ | $x_1$ | $x_2$ | $\cdots$ | $x_k$ | $\cdots$ |
|---|---|---|---|---|---|
| $P$ | $p_1$ | $p_2$ | $\cdots$ | $p_k$ | $\cdots$ |

$\Longrightarrow$

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

$$q_k = P(Y = y_k) = \sum_{\{i;\, y_k = g(x_i)\}} p_i$$

$\Longrightarrow$

| $Y=g(X)$ | $y_1$ | $y_2$ | $\cdots$ | $y_k$ | $\cdots$ |
|---|---|---|---|---|---|
| $P$ | $q_1$ | $q_2$ | $\cdots$ | $q_k$ | $\cdots$ |

# 连续型随机变量的函数的分布

## 问题

已知连续型随机变量$X$的概率密度$f_X(x)$

求它的函数$Y = g(X)$的概率密度$f_Y(x)$

# 连续型随机变量的函数的分布

设随机变量$X$的分布函数$F\ (x)$是连续的，且除有限多个点

$(C_1 < C_2 < \cdots < C_n)$ 外，导数$F'(x)$存在且连续，

则$X$是连续型随机变量，它具有如下的概率密度

$$f(x) = \begin{cases} F'(x), & x \notin \{c_1, \ldots c_n\} \\ 0, & x \in \{c_1, \ldots c_n\} \end{cases}$$

# 连续型随机变量的函数的分布

例2 设随机变量$X$的概率密度函数为 $f(x)=\begin{cases}\frac{x}{8}, & 0<x<4\\ 0, & 其它\end{cases}$

求随机变量$Y$=2$X$+8的概率密度.

解：第一步 先求$Y$=2$X$+8 的分布函数$F_Y(y)$

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

$$F_Y(y) = F_X\left(\frac{y-8}{2}\right)$$

第二步　由分布函数求概率密度.

$$f_Y(y) = F'_Y(y) = \left[F_X\left(\frac{y-8}{2}\right)\right]' = \frac{1}{2} f_X\left(\frac{y-8}{2}\right)$$

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

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

# 分布函数微分法

- 化简$Y=g(X)$的分布函数 $F_Y(y)=P(g(X)\le y)$, ($如$: $F_X\left(\frac{y-8}{2}\right)$)

  直到看出除有限多个点外导数$F'_Y(y)$存在且连续为止.

- 求导数$F'_Y(y)$, 并令 $f_Y(y)=\begin{cases} F_Y{}'(y), & 如果F_Y{}'(y)存在 \\ 0 & 如果F_Y{}'(y)不存在 \end{cases}$

- 结论: $Y=g(X)$的概率密度为 $f_Y(y)$.

# 分 布 函 数 微 分 法

例3：设随机变量$X$有（分段连续的）概率密度$f_X(x)$, 而$Y=aX+b$,

这里$a, b$是常数且$a\neq 0$. 求$Y$的概率密度$f_Y(y)$.

解：$F_Y(y) = P(aX+b\leq y)$

$$= \begin{cases} P\left(X\leq \dfrac{y-b}{a}\right) = F_X\left(\dfrac{y-b}{a}\right), & a>0 \\ P\left(X\geq \dfrac{y-b}{a}\right) = 1-F_X\left(\dfrac{y-b}{a}\right), & a<0 \end{cases}$$

# 分布函数微分法

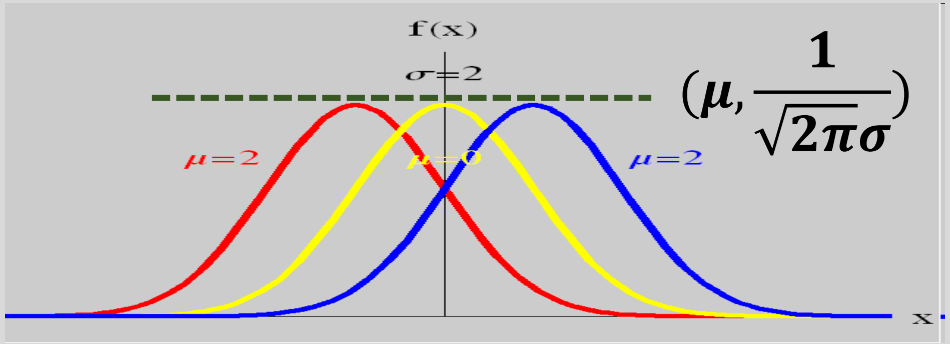
$$f_Y(y)=\begin{cases}\left[F_X\left(\dfrac{y-b}{a}\right)\right]'=\dfrac{1}{a}f_X\left(\dfrac{y-b}{a}\right), & a>0\\[2ex] \left[-F_X\left(\dfrac{y-b}{a}\right)\right]'=-\dfrac{1}{a}f_X\left(\dfrac{y-b}{a}\right), & a<0\end{cases}$$

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

# 正态随机变量的线性函数仍然服从正态分布

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

若$X \sim N(\mu, \sigma^2)$, 则 $\quad Y = \dfrac{X-\mu}{\sigma} \sim N(0,1)$




$$F(x) = \Phi(\frac{x-\mu}{\sigma}) \quad P(X \le x) = P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = \Phi(\frac{x-\mu}{\sigma})$$

# 分 布 函 数 微 分 法

例4： 设随机变量$X$的概率密度为$f_X(x)$, 求随机变量$Y$=$X^2$的概率密度.

解：注意到随机变量$Y = X^2 \geq 0$, 故

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

$$f_Y(y) = F'_Y(y) = [F_X(\sqrt{y})]' - [F_X(-\sqrt{y})]' = f_X(\sqrt{y})\frac{1}{2\sqrt{y}} + f_X(-\sqrt{y})\frac{1}{2\sqrt{y}}$$

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

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

若$X\sim N(0,1)$, 则$Y=X^2$的概率密度为

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

# 分 布 函 数 微 分 法

例5、设$F(x)$是任何一个连续的严格单调增的分布函数，$F^{-1}(x)$为其反函数

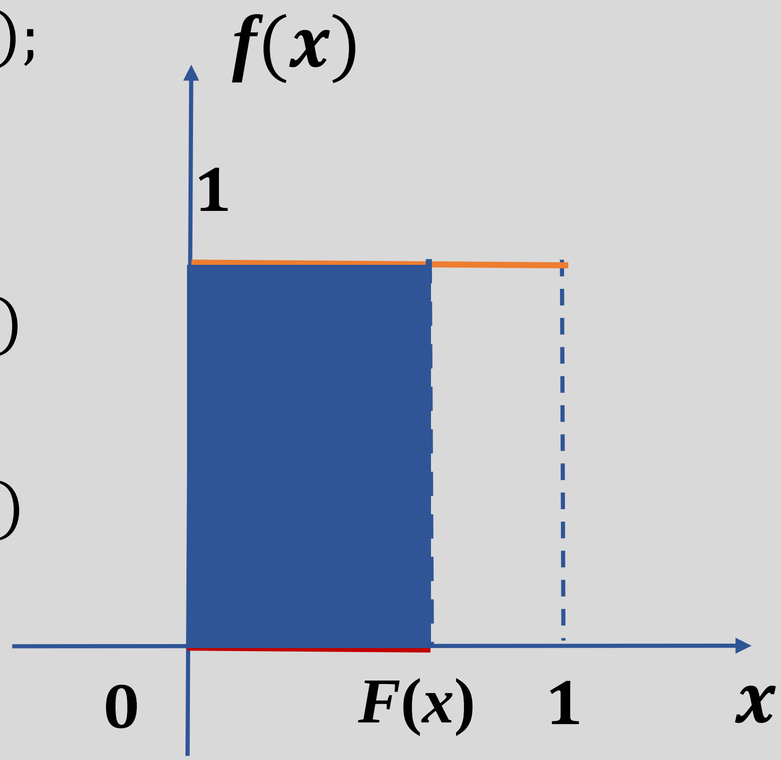
1) 若$X\sim U[0,1]$，则$F^{-1}(X)$的分布函数恰为$F(x)$；

2) 若$Y\sim F(x)$，则$F(Y)\sim U[0,1]$.

分析1)：令 $Y=F^{-1}(X)$，要证 $F_Y(x)=F(x)$

$$F_Y(x)=P(Y\le x)=P(F^{-1}(X)\le x)$$

$$=P\big(X\le F(x)\big)=F(x)$$



# 积 分 转 化 法

设随机变量$X$的概率密度为$f_X(x)$, $g(x)$是(分段)连续或

(分段)单调函数, $Y=g(X)$, 如果对任何有界连续函数$h(x)$, 成立

$$\int_{-\infty}^{+\infty} h[g(x)]f_X(x)dx=\int_{\alpha}^{\beta} h(y)p(y)dy$$

(其中$-\infty \le \alpha < \beta \le +\infty$), 则$Y=g(X)$的概率密度为

$$f_Y(y)=\begin{cases} P(y), & \alpha < y < \beta \\ 0, & \text{其它} \end{cases}$$

# 积 分 转 化 法

例5： 设随机变量$X \sim E(2)$. 证明: $Y = 1 - e^{-2X} \sim U(0,1)$.

证： 由已知条件可得 $X \sim f_X(x) = \begin{cases} 2e^{-2x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$,

$g(x) = 1 - e^{-2x}$，任意给定有界连续函数$h(x)$,

$$\int_{-\infty}^{+\infty} h[g(x)] f_X(x) dx = \int_{0}^{+\infty} h\left(1 - e^{-2x}\right) 2e^{-2x} dx$$

$$\int_{-\infty}^{+\infty} h[g(x)]f_X(x)dx = \int_{\alpha}^{\beta} h(y)p(y)dy \Longrightarrow f_Y(y) = \begin{cases} P(y), & \alpha < y < \beta \\ 0, & \text{其它} \end{cases}$$

$$\int_{-\infty}^{+\infty} h[g(x)]f_X(x)dx = \int_{0}^{+\infty} h\left(1-e^{-2x}\right)2e^{-2x}dx$$

$$y = 1 - e^{-2x}$$

$$e^{-2x} = 1 - y$$

$$dx = \frac{1}{2(1-y)}\mathrm{dy}$$

$$= \int_0^1 h(\mathrm{y})2(1-y)\frac{1}{2(1-y)}dy$$

$$= \int_{\boxed{0}\to\alpha}^{\boxed{1}\to\beta} h(\mathrm{y})\underbrace{\boxed{1}}_{p(y)}dy \Longrightarrow f_Y(y) = \begin{cases} 1, & 0 < y < 1 \\ 0, & \text{其它} \end{cases}$$

$$\Longrightarrow Y = 1 - e^{-2X} \sim U(0,1).$$

# 思　考

设$X$为连续型随机变量，$Y = g(X)$也一定是连续型吗？

设若$X \sim N(0,1)$，$g(x) = \begin{cases} a, & x \geq 0 \\ b, & x < 0 \end{cases}$ $a \neq b$,

则$Y = g(X)$的可能取值为$a, b$,

$$P(Y = a) = P(X \geq 0) = \frac{1}{2} \qquad P(Y = b) = P(X < 0) = \frac{1}{2}$$

即：$Y = g(X)$ 服从两点分布.

# 思　考

是否存在既非离散型也非连续型的随机变量？

设$X \sim U[0,2]$, 又设 $g(x)=\begin{cases} x, & 0 \leq x \leq 1 \\ 1, & 1 \leq x \leq 2 \end{cases}$,

试求$Y=g(X)$的概率分布.

# 思　　考

设$X \sim U[0,2]$，又设 $g(x)=\begin{cases} x, & 0 \le x \le 1 \\ 1, & 1 \le x \le 2 \end{cases}$，试求$Y=g(X)$的概率分布.

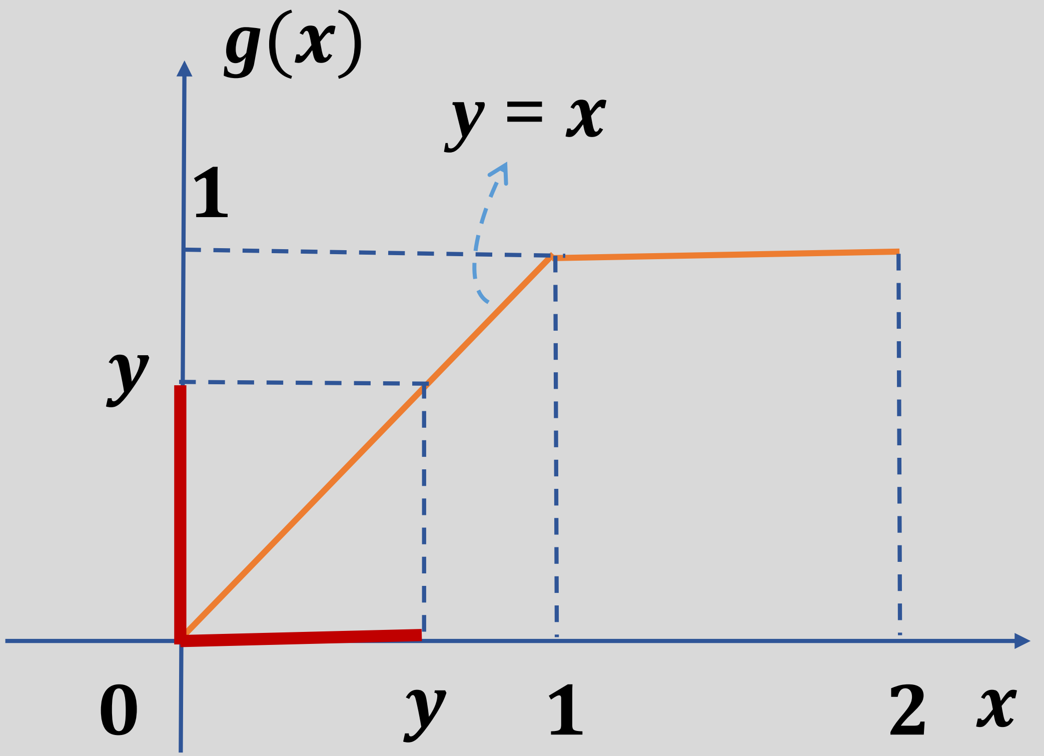
解：$X \sim f(x)=\begin{cases} \frac{1}{2}, & 0 \le x \le 2 \\ 0, & else \end{cases}$，

$Y=g(X)$的可能取值范围为$[0,1]$

当 $0 \le y < 1$ 时，

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

$= P(0 \le X \le y) = \int_0^y \frac{1}{2} dx = \frac{y}{2}$

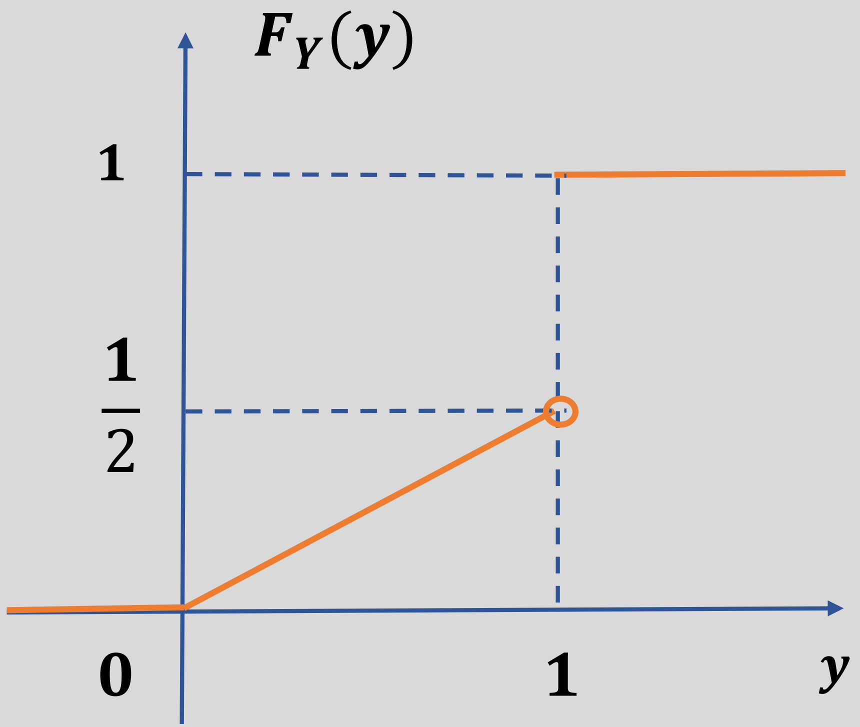


# 思　考

设$X \sim U[0,2]$, 又设 $g(x)=\begin{cases} x, & 0 \le x \le 1 \\ 1, & 1 \le x \le 2 \end{cases}$，试求$Y=g(X)$的概率分布.

故 $Y$ 的分布函数为

$$F_Y(y)=\begin{cases} 0, & y<0 \\ \dfrac{y}{2}, & 0 \le y < 1 \\ 1, & y \ge 1 \end{cases}$$




即随机变量$Y$既不是离散型也不是连续型

# 小　结

设 $g(x)$为(分段)连续函数或(分段)单调函数

$X$为离散型随机变量 $\longrightarrow$ $Y=g(X)$一定是离散型.

$X$为连续型随机变量 $\not\longrightarrow$ $Y=g(X)$一定是连续型.

求$Y=g(X)$的概率分布

- 分布函数微分法
- 积分转化法
- 分布律（随机变量函数为离散型时）