§2.5 随机变量的函数

- 函数$f:\mathbb{R}\to\mathbb{R}$, $x\mapsto y=f(x)$. (定理5.1. 要求是Borel函数). 随机变量X 的函数指: 一个新的随机变量

$$Y=f(X):\omega\mapsto f(X(\omega)).$$

- 目标: 求Y 的分布列或密度函数.
- 例. 假设X 的分布列为$P(X=x_i)=p_i, \forall i$. 则Y 也是离散型, 将其可能取值记为$y_j, \forall j$.

$$P(Y=y_j)=\sum_{i:f(x_i)=y_j} p_i.$$

例5.1. $X \sim N(\mu, \sigma^2)$, 求$Y = \frac{X-\mu}{\sigma}$ 的分布.

- 分布函数法. 第一步, 将$Y \leqslant y$ 改写为$X \leqslant \mu + \sigma y$. 从而,

$$F_Y(y) = P(Y \leqslant y) = P(X \leqslant \mu + \sigma y).$$

- 第二步, 代入X 的密度, 做变量替换:

$$F_Y(y) = \int_{-\infty}^{\mu+\sigma y} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx$$
$$= \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{t^2}{2}\right\} dt.$$

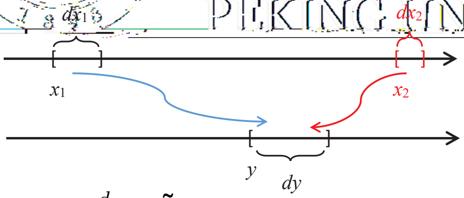
- 第三步, 求导, 得到$p_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$, 即$Y \sim N(0, 1)$.
- 一般地, $Z := a + bX \sim N(a + b\mu, b^2\sigma^2)$, 其中$b \neq 0$.

- 一对一情形: 定理5.2. 假设f 连续可导, $f'(x) > 0$, $\forall x$, 则

$$p_X(x)dx = p_Y(y)dy \to p_Y(y) = p_X(x)\frac{1}{f'(x)} = p_X(g(y))g'(y).$$

- 对一情形*: f 为分段的一对一情形的函数. 例如, f 是 项式函数, 都有

$$p_Y(y) = \sum_{x_i:f(x_i)=y} p_X(x_i)\frac{1}{|f'(x_i)|}.$$



- 若$X \stackrel{d}{=} \tilde{X}$, 则$f(X) \stackrel{d}{=} f(\tilde{X})$.

例5.2. 假设$X \sim N(0,1)$, $Y = X^2$. 求Y 的密度函数.

- 方法一、分布函数法: 对任意$y > 0$,

$$F_Y(y) = P(-\sqrt{y} \leqslant X \leqslant \sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

$$\Rightarrow p_Y(y) = p_X(\sqrt{y})\frac{d\sqrt{y}}{dy} - p_X(-\sqrt{y})\frac{d(-\sqrt{y})}{dy} = \frac{1}{\sqrt{2\pi y}}e^{-\frac{y}{2}}.$$

- 方法二、用 对 公式:

  第一步, 确定每个$y > 0$ 的原像点: $x_1 = \sqrt{y}$, $x_2 = -\sqrt{y}$.

- 第二步, 求出每个x_i 的贡献: $p_X(x_i)\left|\frac{dx_i}{dy}\right| = \frac{1}{2\sqrt{2\pi y}}e^{-\frac{y}{2}}$.
- 第三步, 对i 求和:

$$p_Y(y) = \sum_{i=1}^{2} p_X(x_i)\left|\frac{dx_i}{dy}\right| = \frac{1}{\sqrt{2\pi y}}e^{-\frac{y}{2}}, \quad 其中y > 0.$$

- 注: $X^2 \sim \Gamma\left(\frac{1}{2}, \frac{1}{2}\right)$.

反例. 假设$X \sim N(0,1)$, $Y = f(X)$, 其中,

$$f(x) = \begin{cases} x^2, & 若\ |x| > 1; \\ 1, & 若\ |x| \leqslant 1. \end{cases}$$

- Y 不是连续型:

$$P(Y=1) = P(|X| \leqslant 1) > 0.$$

- f 在$(-1,1)$ 上恒有$f'(x) = 0$.

例5.4. 已知$Y=\ln X\sim N(\mu,\sigma^2)$. 求$X$ 的分布.

- 称X 服从对数正态分布.
- $X=e^Y>0$. 因此, $\forall x>0$, 有

$$G_X(x)=P(X>x)=P(Y>\ln x)$$
$$=\int_{\ln x}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(y-\mu)^2}{2\sigma^2}}dy$$
$$=\int_{x}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(\ln u-\mu)^2}{2\sigma^2}}\frac{1}{u}du.$$

- $p_X(x)=-G_X'(x)=\frac{1}{\sqrt{2\pi}\sigma x}e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}$, $x>0$.

- 定义5.1. 设F 是分布函数(满足单调、规范、右连续). 令

$$F^{-1}(p) := \inf\{x : F(x) \geqslant p\}, \quad \forall p \in (0,1).$$

称F^{-1} 为F 的广义反函数.

- 定理5.3. 假设F 是分布函数.

若$U \sim U(0,1)$, 则$X = F^{-1}(U)$ 满足$F_X = F$.

- 证明: $F^{-1}(p) \leqslant x$ 当且仅当$p \leqslant F(x)$. 于是,

$$P(X \leqslant x) = P(F^{-1}(U) \leqslant x) = P(U \leqslant F(x)) = F(x).$$

- 例5.3. 假设F_X 连续, 则

$$Y = F_X(X) \sim U(0,1).$$

- F_X 不连续时的反例, $X \sim B(1,p)$.