《初等概率论》第 7 讲

邓 婉 璐

清华大学
统计学研究中心

October 19, 2018

一、条件分布和条件密度

A. 离散型的情况

设 (X, Y) 是离散型随机向量，有概率分布

$$p_{ij} = \mathbb{P}(X = x_i, Y = y_j) > 0, \quad i, j = 1, 2, \ldots,$$

则 X, Y 分别有边缘分布

$$p_i = \mathbb{P}(X = x_i), \quad q_j = \mathbb{P}(Y = y_j) \quad i, j = 1, 2, \ldots,$$

定义 1.1 (条件概率分布)

对每个固定的 j, 称

$$\mathbb{P}(X = x_i | Y = y_j) = \frac{\mathbb{P}(X = x_i, Y = y_j)}{\mathbb{P}(Y = y_j)} = \frac{p_{ij}}{q_j}, \quad i = 1, 2, \ldots$$

为给定条件 $Y = y_j$ 下，X 的条件分布列.

一、条件分布和条件密度

定理 1.1

X, Y 独立的**充分必要条件**是对任何 $i, j \geq 1$,

$$\mathbb{P}(X = x_i | Y = y_j) = p_{i\cdot}$$

一、条件分布和条件密度

例 1.1

甲向一个目标射击，用 S_n 表示第 n 次击中目标时的射击次数. 如果甲每次击中目标的概率是 $p=1-q$, 则 $(X,Y)=(S_1,S_2)$ 的联合分布为:

$$\mathbb{P}(X=i,\ Y=j)=p^2q^{j-2},\quad j>i\geq 1.$$

X,Y 的边缘分布分别为:

$$\mathbb{P}(X=i)=\sum_{j=i+1}^{\infty}\mathbb{P}(X=i,\ Y=j)=\sum_{j=i+1}^{\infty}p^2q^{j-2}=pq^{i-1};$$

$$\begin{aligned}\mathbb{P}(Y=j)&=\sum_{i=1}^{j-1}\mathbb{P}(X=i,\ Y=j)=\sum_{i=1}^{j-1}p^2q^{j-2}\\&=(j-1)p^2q^{j-2},\quad j=2,3,\ldots.\end{aligned}$$

一、条件分布和条件密度

于是对确定的 $j(j \geq 2)$，得 X 的条件分布

$$\mathbb{P}(X = i | Y = j) = \frac{p^2 q^{j-2}}{(j-1)p^2 q^{j-2}} = \frac{1}{j-1}, \quad 1 \leq i < j;$$

对确定的 i，得 Y 的条件分布

$$\mathbb{P}(Y = j | X = i) = \frac{p^2 q^{j-2}}{(j-1)pq^{i-1}} = pq^{j-i-1}, \quad j > i.$$

上式表明，已知 $S_2 = j$ 时，S_1 在 $\{1, 2, ..., j-1\}$ 中的取值是等可能的.

此外，若令 $X_1 = S_1$, $X_2 = S_2 - S_1, ...$, 则

$$\begin{aligned}\mathbb{P}(X_1 = i, X_2 = j) &= \mathbb{P}(X = i, Y = i + j) = p^2 q^{i+j-2} \\ &= pq^{i-1} \cdot pq^{j-1} = \mathbb{P}(X_1 = i)\mathbb{P}(X_2 = j).\end{aligned}$$

可见随机变量 $X_1, X_2, \ldots$ 是独立同分布的. ♣ 如果随机变量 $X_1, X_2, \ldots$ 是相互独立并且有相同的分布函数，则称 $X_1, X_2, \ldots$ 独立同分布.

一、条件分布和条件密度

例 1.2

设某超级市场周日的顾客数 N 是随机变量，单个顾客的消费（单位：元）与 N 独立，且服从 $\mathcal{P}(\mu)$. 用 S 表示该日的全天营业额，求条件概率 $\mathbb{P}(S=k|N)$.

解. X_j 表示第 j 个顾客的消费额，则 $X_1, X_2, \ldots$ 独立同分布. 用 $S_m = X_1 + \ldots + X_m$ 表示前 m 个顾客的消费额，则 $S_m \sim \mathcal{P}(m\mu)$. 于是

$$\mathbb{P}(S=k|N=m) = \mathbb{P}(S_m=k) = \frac{(m\mu)^k}{k!}\mathrm{e}^{-m\mu}, \quad m=0,1,\ldots.$$

故

$$\mathbb{P}(S=k|N) = \frac{(N\mu)^k}{k!}\mathrm{e}^{-N\mu}.$$

一、条件分布和条件密度

B. 连续型的情况

背景：北京夏季的高温闷热天气会造成北京电网的负荷过高，用 Y 表示夏季未来某天的最高气温，用 X 表示同一天北京电网的最大负荷. 可以认为 (X, Y) 是连续型随机变量，有联合密度 $f(x, y)$. 如果已有对 Y 的预测值 y，在已知 $Y = y$ 的条件下研究 X 的概率分布是有实际意义的工作. 用

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

表示已知 $Y = y$ 的条件下，X 的分布函数，称为条件分布函数. 注意：条件分布函数 $\mathbb{P}(X \le x | Y = y)$ 就是最高气温为 y 的那天北京电网最大负荷的概率分布函数，是有明确的意义的.

♣ 如何计算条件分布 $\mathbb{P}(X \le x | Y = y)$. X, Y 分别有边缘密度

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx.$$

一、条件分布和条件密度

对于充分小的 $\varepsilon > 0$，可以理解

$$\mathbb{P}(X \le x | y - \varepsilon < Y \le y) \approx \mathbb{P}(X \le x | Y = y).$$

另一方面，如果 $f_Y(y)$ 在 y 连续，$f_Y(y) > 0$，并且 $\partial F(x, y)/\partial y$ 存在，就有

$$\lim_{\varepsilon \to 0} \mathbb{P}(X \le x | y - \varepsilon < Y \le y) = \lim_{\varepsilon \to 0} \frac{\mathbb{P}(X \le x, y - \varepsilon < Y \le y)}{\mathbb{P}(y - \varepsilon < Y \le y)}$$

$$= \lim_{\varepsilon \to 0} \frac{F(x, y) - F(x, y - \varepsilon)}{F_Y(y) - F_Y(y - \varepsilon)}$$

$$= \frac{\frac{\partial}{\partial y} \int_{-\infty}^{y} \left(\int_{-\infty}^{x} f(s, t)\, ds \right) dt}{f_Y(y)}$$

$$= \frac{\int_{-\infty}^{x} f(s, y)\, ds}{f_Y(y)}.$$

一、条件分布和条件密度

定义 1.2

设随机向量 (X, Y) 有联合密度 $f(x, y)$，Y 有边缘密度 $f_Y(y)$. 若在 y(确定的 y) 处 $f_Y(y) > 0$, 则称

$$\mathbb{P}(X \le x | Y = y) = \frac{\int_{-\infty}^{x} f(s, y)\, ds}{f_Y(y)}, \quad x \in R$$

为给定条件 $Y = y$ 下，X 的**条件分布函数**，简称**条件分布**，记作 $F_{X|Y}(x|y)$. 称

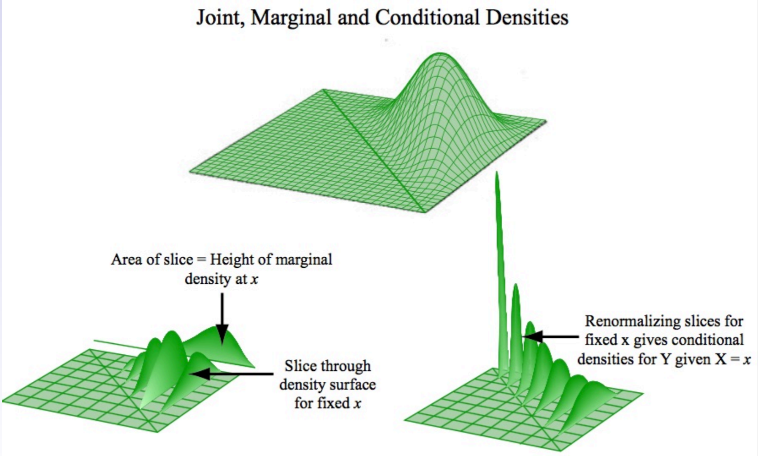
$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = \frac{f(x, y)}{\int_{-\infty}^{\infty} f(x, y)\, dx}, \quad x \in R$$

为给定条件 $Y = y$ 下，X 的**条件概率密度**，简称**条件密度**.

♣ $f(x, y) = f_Y(y) f_{X|Y}(x|y)$.

一、条件分布和条件密度

对任意固定的 x，$f_{Y|X}(y|x)$ 是一个概率密度函数.



一、条件分布和条件密度

♣ 条件密度与条件分布的关系：

对使得 $f_Y(y) > 0$ 的 y,

1. $F_{X|Y}(x|y) = \mathbb{P}(X \le x | Y = y) = \int_{-\infty}^{x} f_{X|Y}(s|y)ds$, $x \in R$;
2. 如果 $F_{X|Y}(x|y)$ 关于 x 连续，且除去至多可列个点外有连续的导数，则

$$f_{X|Y}(x|y) = \begin{cases} \dfrac{\partial F_{X|Y}(x|y)}{\partial x}, & \text{当偏导数存在,} \\ 0, & \text{其它} \end{cases}$$

是给定条件 $Y = y$ 下，X 的条件密度.

一、条件分布和条件密度

定理 1.2

X, Y 独立的**充分必要条件**是对 $y \in \{y|f_Y(y) > 0\}$,

$$f_{X|Y}(x|y) = f_X(x), \quad x \in R.$$

一、条件分布和条件密度

例 1.3

设 $(X, Y) \sim \mathcal{N}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$，已知 $X = x$ 时，求 Y 的条件密度.

解. $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$. 故

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{1}{\sigma_2\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{(y-\mu_x)^2}{2(1-\rho^2)\sigma_2^2}\right),$$

其中 $\mu_x = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)$.

此说明

$$Y\big|_{X=x} \sim \mathcal{N}\left(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1),\ (1-\rho^2)\sigma_2^2\right).$$

同理，

$$X\big|_{Y=y} \sim \mathcal{N}\left(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2),\ (1-\rho^2)\sigma_1^2\right).$$

当 X 和 Y 独立时，$\rho = 0$，于是 $f_{Y|X}(y|x) = f_Y(y)$.

一、条件分布和条件密度

例 1.4

设计算机使用的环境指标 $Y \sim \Gamma(\alpha, \beta)$，概率密度是

$$f_Y(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} \mathrm{e}^{-\beta y}, \quad y > 0.$$

已知 $Y = y$ 时，软件的使用寿命 $X \sim \mathcal{E}(y)$. 求 X 的分布.

解. X 的条件密度 $f_{X|Y}(x|y) = y\mathrm{e}^{-xy}$, $x > 0$. 于是 (X, Y) 的联合密度为

$$f(x, y) = f_{X|Y}(x|y) f_Y(y) = y\mathrm{e}^{-xy} \cdot \frac{\beta^\alpha}{\Gamma(\alpha)} y^{\alpha-1} \mathrm{e}^{-\beta y}, \quad x, y > 0.$$

一、条件分布和条件密度

最后对 $x>0$, 有

$$
\begin{aligned}
f_X(x) &= \int_0^\infty f(x,y)\,dy = \int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} y^\alpha \mathrm{e}^{-y(x+\beta)}\,dy \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{1}{(x+\beta)^{\alpha+1}} \int_0^\infty t^\alpha \mathrm{e}^{-t}\,dt \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+1)}{(x+\beta)^{\alpha+1}} \\
&\overset{\Gamma(\alpha+1)=\alpha\Gamma(\alpha)}{=\!=\!=\!=\!=\!=} \frac{\alpha\beta^\alpha}{(x+\beta)^{\alpha+1}}.
\end{aligned}
$$

于是 X 有概率密度

$$
f_X(x) = \frac{\alpha\beta^\alpha}{(x+\beta)^{\alpha+1}}, \quad x>0.
$$

一、条件分布和条件密度

♣ 随机向量 $(\mathbf{X},\mathbf{Y})$ 的条件分布

定义 1.3

设 $\mathbf{X}=(X_1,...,X_n)$ 和 $\mathbf{Y}=(Y_1,...,Y_m)$ 是随机向量，$f(\mathbf{x},\mathbf{y})$ 是 $(\mathbf{X},\mathbf{Y})$ 的联合密度，此时 $\mathbf{Y}$ 有边缘密度

$$f_{\mathbf{Y}}(\mathbf{y})=\int_{R^n} f(\mathbf{x},\mathbf{y})\,d\mathbf{x}.$$

若在 $\mathbf{y}$(确定的 $\mathbf{y}$) 处 $f_{\mathbf{Y}}(\mathbf{y})>0$，则称

$$\mathbb{P}(\mathbf{X}\le\mathbf{x}|\mathbf{Y}=\mathbf{y})=\frac{\int_{\mathbf{s}\le\mathbf{x}} f(\mathbf{s},\mathbf{y})\,d\mathbf{s}}{f_{\mathbf{Y}}(\mathbf{y})}$$

为给定条件 $\mathbf{Y}=\mathbf{y}$ 下，$\mathbf{X}$ 的条件分布函数，简称条件分布，记作 $F_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y})$. 称

$$f_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y})=\frac{f(\mathbf{x},\mathbf{y})}{f_{\mathbf{Y}}(\mathbf{y})}$$

为给定条件 $\mathbf{Y}=\mathbf{y}$ 下，$\mathbf{X}$ 的条件概率密度，简称条件密度.

二、次序统计量

定义 2.1

设 $X_1, X_2, ..., X_n$ 是概率空间 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的随机变量，对 $\omega \in \Omega$，将 $X_1(\omega), X_2(\omega), ..., X_n(\omega)$ 从小到大排列得到

$$X_{(1)}(\omega) \leq X_{(2)}(\omega) \leq \cdots \leq X_{(n)}(\omega).$$

称 $X_{(1)}, X_{(2)}, ..., X_{(n)}$ 为 $X_1, X_2, ..., X_n$ 的次序统计量 (order statistics).

♣ 次序统计量的分布密度

以下设随机变量 $X_1, ..., X_n$ 独立同分布，有公共的分布函数 $F(x)$ 和密度函数 $f(x)$.

二、次序统计量

例 2.1

$(X_{(1)}, X_{(2)}, ..., X_{(n)})$ 的联合分布 $F_n(x_1, x_2, ..., x_n)$ 是连续函数.

证明. 只须证明每个 $X_{(k)}$ 的分布函数连续. 利用 $\mathbb{P}(X_1 = x) = 0$, 得

$$
\begin{aligned}
&\mathbb{P}(X_{(k)} = x) \\
&= \mathbb{P}(\{X_j\}\text{中有 } k-1 \text{ 个} < x, \text{有一个} = x, \text{有 } n-k \text{ 个} > x) \\
&= \frac{n!}{(k-1)!1!(n-k)!}[\mathbb{P}(X_1 < x)]^{k-1}\mathbb{P}(X_1 = x)[1 - F_1(x)]^{n-k} \\
&= 0.
\end{aligned}
$$

二、次序统计量

例 2.2

对于 $-\infty \le a < x_1 < x_2 < \cdots < x_k < b \le \infty$，有

$$\int_{a<x_1<\cdots<x_k<b} f(x_1)\cdots f(x_k)\,dx_1\cdots dx_k = \frac{(F(b)-F(a))^k}{k!}.$$

证明. （归纳法）$k=1$ 时结论成立. 假设结论对 $k-1$ 成立，对于 k，用 Fubini 定理得到

$$\begin{aligned}
&\int_{a<x_1<\cdots<x_k<b} f(x_1)\cdots f(x_k)\,dx_1\cdots dx_k\\
&=\int_a^b \Big(\int_{a<x_1<\cdots<x_k} f(x_1)\cdots f(x_{k-1})\,dx_1\cdots dx_{k-1}\Big) f(x_k)\,dx_k\\
&=\int_a^b \frac{(F(x_k)-F(a))^{k-1}}{(k-1)!} f(x_k)\,dx_k\\
&=\frac{(F(b)-F(a))^k}{k!}.
\end{aligned}$$

二、次序统计量

例 2.3

$(X_{(1)}, X_{(2)}, \ldots, X_{(n)})$ 有联合密度

$$g(x_1, \ldots, x_n) = \begin{cases} n! \prod_{i=1}^{n} f(x_i), & \text{if } x_1 < \cdots < x_n, \\ 0, & \text{others} \end{cases}$$

例 2.4

$X_{(k)}$ 有密度

$$g_k(x_k) = n \binom{n-1}{k-1} [F(x_k)]^{k-1} [1 - F(x_k)]^{n-k} f(x_k).$$

例 2.5

对 $k_1 < k_2$, $(X_{(k_1)}, X_{(k_2)})$ 有联合密度

$$
\begin{aligned}
g(x_{k_1}, x_{k_2}) = {} & \frac{n!}{(k_1-1)!(k_2-k_1-1)!(n-k_2)!} \\
& \times [F(x_{k_1})]^{k_1-1}[F(x_{k_2})-F(x_{k_1})]^{k_2-k_1-1} \\
& \times [1-F(x_{k_2})]^{n-k_2} f(x_{k_1}) f(x_{k_2}), \quad x_{k_1} < x_{k_2}.
\end{aligned}
$$

例 2.6

某家庭原来有 4 只灯泡用于室内照明，新装修后有 24 只灯泡用于室内照明. 装修入住后主人总认为灯泡更容易坏了，试解释其中的原因.

二、次序统计量

解. 设所有灯泡的使用寿命相互独立，且服从 $\mathcal{E}(\lambda)$. 用 X_i 表示第 i 只灯泡的使用寿命，则装修前等待第一只灯泡烧坏的时间长度 X 为

$$X = \min\{X_1, X_2, X_3, X_4\};$$

装修后等待第一只灯泡烧坏的时间长度 Y 为

$$Y = \min\{X_1, X_2, ..., X_{24}\}.$$

利用例 3.4 的结论分别得到 X 和 Y 的密度函数

$$f_X(t) = 4\lambda e^{-4\lambda t}, \quad f_Y(t) = 24\lambda e^{-24\lambda t}, \quad t > 0,$$

即 $X \sim \mathcal{E}(4\lambda)$, $Y \sim \mathcal{E}(24\lambda)$. 于是有 $\mathbb{P}(X > t) = e^{-4\lambda t}$ 和 $\mathbb{P}(Y > t) = e^{-24\lambda t}$. 当 $\lambda = 1/1500$(小时) 时 (比如 Philips 牌长寿灯泡),

$$\mathbb{P}(X > 400) = 0.3442, \quad \mathbb{P}(X > 200) = 0.5866;$$
$$\mathbb{P}(Y > 400) = 0.0017, \quad \mathbb{P}(Y > 200) = 0.0408.$$

从中可以看到 Y 要比 X 随机地小很多.

三、随机变量的 p 分位数

设 X 是 $(\Omega, \mathscr{F}, \mathbb{P})$ 上的 r.v.，$F(x)$ 是其分布函数.

定义：随机变量的 p 分位数

对 $p \in (0,1)$，定义

$$F^{-1}(p) = \inf\{x | F(x) \geq p\}.$$

称 $F^{-1}(p)$ 为 F 的或者 X 的 p **分位数**，通常用 ξ_p 表示. 特别地，称 $\xi_{1/2}$ 为 F 或 X 的中位数.

♣ 分位数的性质

1. $F^{-1}(p)$ 单调非降。

证明. 对任意的 $p_1 < p_2$，

$\Longrightarrow$ 如果 $F(t) \geq p_2$，则 $F(t) \geq p_2 > p_1$，

$\Longrightarrow \{t : F(t) \geq p_1\} \supset \{t : F(t) \geq p_2\}$

$\Longrightarrow F^{-1}(p_1) = \inf\{t : F(t) \geq p_1\} \leq \inf\{t : F(t) \geq p_2\} = F^{-1}(p_2)$.

三、随机变量的 p 分位数

② $F^{-1}(F(x)) \le x, \quad \forall x \in R.$

证明. $F^{-1}(F(x)) = \inf\{t : F(t) \ge F(x)\} \le x$ as $x \in \{t : F(t) \ge F(x)\}$.

③ $F(F^{-1}(p)) \ge p, \quad \forall p \in (0,1).$

证明. 首先集合 $\{t : F(t) \ge p\}$ 一定具有如下形式：

$$\{t : F(t) \ge p\} = (a, \infty), \quad \text{或} \quad [a, \infty). \tag{1}$$

(事实上，一定不是这种 (a, ∞) 形式.)
为此，假定 $r \in \{t : F(t) \ge p\}$ (蕴含了 $F(r) \ge p$), 则，对任意的 $r' > r$, 有 $F(r') \ge F(r) \ge p$, 即 $r' \in \{t : F(t) \ge p\}$. 由 (1), 得 $F^{-1}(p) = \inf\{t : F(t) \ge p\} = a$. 由于 $a + n^{-1} \in \{t : F(t) \ge p\}$, 有 $F(a + n^{-1}) \ge p$. 令 $n \to \infty$, 并利用 F 的右连续性，得 $F(F^{-1}(p)) = F(a) = \lim_{n\to\infty} F(a + n^{-1}) \ge p$.

注：上述证明中，$a = F^{-1}(p) \in \{t : F(t) \geq p\}$. 因此有

$$\{t : F(t) \geq p\} = [a, \infty) = [F^{-1}(p), \infty).$$

故

$$F^{-1}(p) = \inf\{t : F(t) \geq p\} = \min\{t : F(t) \geq p\}.$$

4. $F^{-1}(p) \leq t \iff p \leq F(t)$.

证明. $(\Longleftarrow)$ 如果 $F(t) \geq p \Longrightarrow t \in \{t : F(t) \geq p\} \Longrightarrow t \geq \inf\{t : F(t) \geq p\} = F^{-1}(p)$.

$(\Longrightarrow)$ 如果 $F^{-1}(p) \leq t$, 由于 F 是单调非降的，利用 (3) 有 $F(t) \geq F(F^{-1}(p)) \geq p$.

三、随机变量的 p 分位数

5. $F^{-1}(p)$ 是左连续的.

证明. 令 $p_n \nearrow p$.

$\Longrightarrow F^{-1}(p_n) \le F^{-1}(p)$, 不妨设 $F^{-1}(p_n) \nearrow b$

$\Longrightarrow F^{-1}(p_n) \le b \le F^{-1}(p), \forall n.$

$\Longrightarrow p_n \le F(b)$ (利用 (4)).

$\Longrightarrow F(b) \ge \lim\limits_{n\to\infty} p_n = p.$

$\Longrightarrow b \ge F^{-1}(F(b)) \ge F^{-1}(p)$ (利用 (1),(2)).

$\Longrightarrow \lim\limits_{n\to\infty} F^{-1}(p_n) = b = F^{-1}(p).$

6. 如果 F 连续的, 则 $F(F^{-1}(p)) = p, \quad \forall p \in (0,1).$

证明. 由 (3), $F(F^{-1}(p)) \ge p$. 现在证明等号成立. (反证法) 假设不成立, 令 $a = F^{-1}(p)$, 有 $F(a) > p$. 利用 F 的连续性和单调性, 则存在 $\delta > 0$ 使得 $F(a-\delta) \ge p$, 因此 $a-\delta \ge F^{-1}(F(a-\delta)) \ge F^{-1}(p) = a$(利用 (1),(2)) 此蕴含了 $\delta \le 0$, 矛盾!

三、随机变量的 p 分位数

定理 3.1（产生服从分布函数 $F(x)$ 的随机变量）

设 $X \sim \mathcal{U}(0,1)$, $F(x)$ 是连续分布函数，则 $Y = F^{-1}(X) \sim F$.

证明. 显然

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

因此 $Y \sim F$.

例 3.1

如何生成服从下列分布函数的随机变量，(1)$\mathcal{E}(\lambda)$, (2)$\mathcal{N}(0,1)$, (3)Cauchy 分布，(4)$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I(x_i \le x)$, 其中 $x_1, \ldots, x_n$ 是给定的观测值.

解. 设 $U, V \sim \mathcal{U}(0,1)$.

(1). $X = -\lambda^{-1}\log(1-U)$.

三、随机变量的 p 分位数

(2). 令 $\theta = 2\pi U$, $R = \sqrt{-2\log V}$. 则 $X = R\cos\theta$ 和 $Y = R\sin\theta$ 独立的 $\mathcal{N}(0,1)$.

(3). $Z = X/Y$, 其中 X, Y 是独立的且服从 $\mathcal{N}(0,1)$.

(4). 注意到 $F_n(x)$ 是阶梯函数. 令 $x_{(1)} \leq \ldots \leq x_{(n)}$ 是 $x_1, \ldots, x_n$ 次序观测值. 因为

$$\begin{aligned} F^{-1}(x) &= x_{(1)}, \qquad 0 < x \leq \frac{1}{n} \\ &= x_{(2)}, \qquad \frac{1}{n} < x \leq \frac{2}{n} \\ &= \ldots\ldots \\ &= x_{(n)}, \qquad \frac{n-1}{n} < x \leq 1. \end{aligned}$$

所以，$F^{-1}(U)$ 的值域为 $\{x_{(1)}, \ldots, x_{(n)}\}$ 且

$\mathbb{P}(F^{-1}(U) = x_{(k)}) = \mathbb{P}((k-1)/n < U \leq k/n) = 1/n$, $k = 1, \ldots, n$.

即只须等概率地从观测值中抽样.

三、随机变量的 p 分位数

其它定义：随机变量的 p 分位数

对 $p \in (0,1)$，如果

$$\mathbb{P}(X \le x) \ge p, \quad \mathbb{P}(X \ge x) \ge 1-p,$$

称 x 为 X 的 p-分位数.

♣ 上述定义方式所确定的 p-分位数不一定唯一.

其它定义：随机变量的 p 分位数

对 $p \in (0,1)$，

$$F^{-1}(p) = \sup\{x | F(x) < p\}.$$

小结

重点知识点

- 基于随机变量的条件分布
- 基于连续型随机变（向）量的条件密度
- 基于分布函数得到的 p-分位数

技巧

- 密度与分布间的转换 (e.g. 理解条件密度的定义).
- 殊途同归的选择，同一问题的不同解法，可能曲线救国计算更简便 (e.g. Cauchy 分布 PDF 的计算)
- 生成服从某特定分布的随机变量.

思考

条件概率的形式？相应的乘法法则、全概率公式、贝叶斯法则？

作业

- 教材：第二章 36, 37; 第三章 29; 第四章 5, 8, 10, 14
- 设 $X \sim \mathcal{U}(0,1)$，求 $h(x)$ 使得 $Y = h(X)$ 服从两点分布 $B(1,p)$.

$$Thank\quad you!$$