概率论

胡太忠
thu@ustc.edu.cn

安徽 合肥

2023 年 9 月

第 4 章　随机向量

(12 课时)

- 联合分布与边际分布
- 条件分布
- 独立性
- 随机变量的函数

4.1.1 随机向量的定义

设概率空间为 $(\Omega, \mathscr{A}, \mathrm{P})$. $(X_1, \ldots, X_n)$ 为一个随机向量当且仅当对每个 i, X_i 为随机变量.

- **定理 4.1.1** $(X_1, \ldots, X_n)$ 为一个随机向量当且仅当对任意 $x_i \in \Re$,

$$\{X_1 \le x_1, \ldots, X_n \le x_n\} \in \mathscr{A}.$$

※ $(X_1, \ldots, X_n)$ 为一个随机向量当且仅当对任意 $B \in \mathscr{B}(\Re^n)$,

$$\{w : (X_1(w), \ldots, X_n(w)) \in B\} \in \mathscr{A}.$$

4.1.1 随机向量的定义

设概率空间为 $(\Omega, \mathscr{A}, \mathrm{P})$. $(X_1, \ldots, X_n)$ 为一个随机向量当且仅当对每个 i, X_i 为随机变量.

- **定理 4.1.1** $(X_1, \ldots, X_n)$ 为一个随机向量当且仅当对任意 $x_i \in \Re$,

$$\{X_1 \le x_1, \ldots, X_n \le x_n\} \in \mathscr{A}.$$

※ $(X_1, \ldots, X_n)$ 为一个随机向量当且仅当对任意 $B \in \mathscr{B}(\Re^n)$,

$$\{w : (X_1(w), \ldots, X_n(w)) \in B\} \in \mathscr{A}.$$

Borel-Cantelli 引理的应用

- 【例 4.2.a】 设随机变量序列 $\{X_n\}$ 不必独立，满足

$$\mathrm{P}(X_n = 0) = \frac{1}{n^2} = 1 - \mathrm{P}(X_n = 1), \quad \forall\, n \geq 1,$$

则 $\lim X_n = 1$, a.s..

- 【例 4.2b】 设随机变量序列 $\{X_n\}$ 相互独立，满足

$$\mathrm{P}(X_n = 0) = \frac{1}{n} = 1 - \mathrm{P}(X_n = 1), \quad \forall\, n \geq 1,$$

则 $\mathrm{P}(X_n = 0, \mathrm{i.o.}) = 1 = \mathrm{P}(X_n = 1, \mathrm{i.o.})$. 于是, X_n 的极限不存在.

Borel-Cantelli 引理是概率极限理论的基础

Borel-Cantelli 引理的应用

- 【例 4.2.a】 设随机变量序列 $\{X_n\}$ 不必独立，满足
$$\mathrm{P}\left(X_n=0\right)=\frac{1}{n^2}=1-\mathrm{P}\left(X_n=1\right), \quad \forall\, n \geq 1,$$
则 $\lim X_n=1$, a.s..

- 【例 4.2b】 设随机变量序列 $\{X_n\}$ 相互独立，满足
$$\mathrm{P}\left(X_n=0\right)=\frac{1}{n}=1-\mathrm{P}\left(X_n=1\right), \quad \forall\, n \geq 1,$$
则 $\mathrm{P}\left(X_n=0, \mathrm{i.o.}\right)=1=\mathrm{P}\left(X_n=1, \mathrm{i.o.}\right)$. 于是, X_n 的极限不存在.

Borel-Cantelli 引理是概率极限理论的基础

【例 4.2.11】 设 X, Y iid $\sim U(0,1)$, $Z = \{X+Y\}$, 则 $Z \sim U(0,1)$, 且 X, Y, Z 两两独立, 但不相互独立.

解: 先证 $Z \sim U(0,1)$. 对任意 $x \in (0,1)$,

$$
\begin{aligned}
F_Z(x) &= \mathrm{P}\left(X+Y \le x\right) + \mathrm{P}\left(1 \le X+Y \le 1+x\right) \\
&= \int_0^x (x-u)du + \int_0^1 \mathrm{P}\left(1-u \le Y \le (1+x-u)\wedge 1\right)du \\
&= \frac{1}{2}x^2 + \int_0^1 [(1+x-u)\wedge 1 - 1 + u]du \\
&= \frac{1}{2}x^2 + \int_0^x udu + \int_x^1 xdu \\
&= \frac{1}{2}x^2 + \left(x - \frac{1}{2}x^2\right) = x.
\end{aligned}
$$

再证 X, Y, Z 两两独立. 对任意 $0 < x \le y \le 1$,

$$
\begin{aligned}
& \mathrm{P}(X \le x, Z \le y) \\
= \; & \mathrm{P}(X \le x, X+Y \le y) + \mathrm{P}(X \le x, 1 \le X+Y \le 1+y) \\
= \; & \int_0^x (y-u)du + \int_0^x \mathrm{P}(1-u \le Y \le (1+y-u) \wedge 1)du \\
= \; & xy - \frac{1}{2}x^2 + \int_0^x [(1+y-u) \wedge 1 - 1 + u]du \\
= \; & xy - \frac{1}{2}x^2 + \int_0^x udu \\
= \; & xy;
\end{aligned}
$$

对任意 $0 < y \le x \le 1$,

$$\begin{aligned} &\mathrm{P}\left(X \le x, Z \le y\right) \\ &= \mathrm{P}\left(X \le y, Z \le y\right) + \mathrm{P}\left(y < X \le x, 1 \le X + Y \le 1 + y\right) \\ &= y^2 + \int_y^x \mathrm{P}\left(1 - u \le Y \le 1 + y - u\right) du \\ &= y^2 + \int_y^x y\, du = y^2 + (xy - y^2) = xy, \end{aligned}$$

即 X, Z 相互独立. 类似, Y, Z 相互独立. 但 X, Y, Z 不相互独立, 因为

$$\mathrm{P}\left(X \le \frac{1}{4}, Y \le \frac{1}{4}, Z > \frac{1}{2}\right) = 0 \ne \frac{1}{4} \cdot \frac{1}{4} \cdot \frac{1}{2}.$$

4.4.5 随机变量的随机加权平均

考虑如下特殊的随机变量的随机加权平均：

$$Z = J X_1 + (1 - J) X_2,$$

其中 X_1, X_2, J 是定义于 $(\Omega, \mathscr{A}, \mathrm{P})$ 的随机变量，$J \sim \mathrm{Bernoulli}(p)$.

- 设 $X_1 \sim F_1$，$X_2 \sim F_2$，且 X_1, X_2, J 相互独立，则

$$F_Z(x) = pF_1(x) + (1-p)F_2(x), \quad x \in \Re.$$

- 设 $X_1 = \min\{Y_1, Y_2\}$，$X_2 = \max\{Y_1, Y_2\}$，其中 $Y_1 \sim G_1$，$Y_2 \sim G_2$ 且 Y_1, Y_2, J 相互独立，则

$$F_Z(x) = pG_1(x) + pG_2(x) + (1-2p)G_1(x)G_2(x), \quad x \in \Re.$$

- 设 $X_2 = \min\{Y_1, Y_2\}$，$X_1 = \max\{Y_1, Y_2\}$，其中 $Y_1 \sim G_1$，$Y_2 \sim G_2$ 且 Y_1, Y_2, J 相互独立，求 $F_Z(x)$.

4.4.7 记录值

设 $\{X_n, n \geq 1\}$ iid, 共同分布函数为 F, 视 X_n 为时刻 n 的观测值.

- X_1 是第一个纪录值.
- 当 $n \geq 2$ 时, 如果 X_n 超过它前面的所有变量的值, 即

$$X_n > \max\{X_1, \ldots, X_{n-1}\},$$

则称于时刻 n 创造了一个记录, 记录值为 X_n.

再定义

$$I_j = \begin{cases} 1, & X_j \text{ 是一个记录值}, \\ 0, & \text{否则}, \end{cases}$$

则 $Z_n = \sum_{i=1}^n I_j$ 表示到时刻 n 前 n 个观测值中记录值出现的次数.

设 $\{X_n, n \geq 1\}$ iid, 共同分布函数为 F, 且 F 为连续函数.

问题：研究序列 $\{I_j, j \geq 1\}$ 性质，特别是独立性.

- **命题** 对任意 $j \geq 1$,

$$\mathrm{P}(I_j = 1) = \frac{1}{j}, \quad \mathrm{P}(I_j = 0) = 1 - \frac{1}{j}.$$

证明: 利用对称性, 有

$$\mathrm{P}(X_\ell = \max\{X_1, \ldots, X_j\}) = \frac{1}{j}, \quad \ell = 1, \ldots, j.$$

设 $\{X_n, n \geq 1\}$ iid, 共同分布函数为 F, 且 F 为连续函数.

问题：研究序列 $\{I_j, j \geq 1\}$ 性质，特别是独立性.

- 命题 对任意 $j \geq 1$,

$$\mathrm{P}(I_j = 1) = \frac{1}{j}, \quad \mathrm{P}(I_j = 0) = 1 - \frac{1}{j}.$$

证明: 利用对称性, 有

$$\mathrm{P}(X_\ell = \max\{X_1, \ldots, X_j\}) = \frac{1}{j}, \quad \ell = 1, \ldots, j.$$

设 $\{X_n, n \geq 1\}$ iid $\sim F$, 且 F 连续. 定义

$$R_n = \sum_{k=1}^{n} I(X_k \geq X_n),$$

R_n 表示 X_n 在 $X_1, X_2, \ldots, X_n$ 中处于第几大位置.

- $R_n \in \{1, 2, \ldots, n\}$.
- $(R_1, R_2, \ldots, R_n)$ 唯一确定 $X_1, X_2, \ldots, X_n$ 的排序. 例如, 当 $n = 3$ 时,

$$\begin{aligned}
\{R_1 = 1, R_2 = 1, R_3 = 1\} &= \{X_1 < X_2 < X_3\}, \\
\{R_1 = 1, R_2 = 1, R_3 = 2\} &= \{X_1 < X_3 < X_2\}, \\
\{R_1 = 1, R_2 = 1, R_3 = 3\} &= \{X_2 > X_1 > X_3\}, \\
\{R_1 = 1, R_2 = 2, R_3 = 1\} &= \{X_3 > X_1 > X_2\}, \\
\{R_1 = 1, R_2 = 2, R_3 = 3\} &= \{X_1 > X_2 > X_3\}, \\
\{R_1 = 1, R_2 = 2, R_3 = 2\} &= \{X_1 > X_3 > X_2\}.
\end{aligned}$$

- 定理 4.4.5 (Renyi 定理) 设 $\{X_n, n \geq 1\}$ iid $\sim F$, 且 F 连续, 则

  (1) $\{R_n, n \geq 1\}$ 为独立随机变量序列, 且

  $$\mathrm{P}(R_n = k) = \frac{1}{n}, \quad k = 1, \ldots, n.$$

  (2) $\{I_n, n \geq 1\}$ 是独立的 Bernoulli 随机变量序列, 且

  $$\mathrm{P}(I_n = 1) = \frac{1}{n}, \quad \mathrm{P}(I_n = 0) = 1 - \frac{1}{n}.$$

分析: $\{I_n, n \geq 1\}$ 相互独立 $\Longleftrightarrow$ 事件序列 $\{A_n, n \geq 1\}$ 相互独立, 其中

$$A_n = \{I_n = 1\} = \{R_n = 1\}.$$

- 定理 4.4.5 (Renyi 定理) 设 $\{X_n, n \geq 1\}$ iid $\sim F$, 且 F 连续, 则

  (1) $\{R_n, n \geq 1\}$ 为独立随机变量序列, 且

$$\mathrm{P}(R_n = k) = \frac{1}{n}, \quad k = 1, \ldots, n.$$

  (2) $\{I_n, n \geq 1\}$ 是独立的 Bernoulli 随机变量序列, 且

$$\mathrm{P}(I_n = 1) = \frac{1}{n}, \quad \mathrm{P}(I_n = 0) = 1 - \frac{1}{n}.$$

分析: $\{I_n, n \geq 1\}$ 相互独立 $\Longleftrightarrow$ 事件序列 $\{A_n, n \geq 1\}$ 相互独立, 其中

$$A_n = \{I_n = 1\} = \{R_n = 1\}.$$

证明: 首先, $\mathrm{P}(R_n = k) = 1/n$ 显然, 其中 $k = 1, \ldots, n$. 以下仅证明对任意 $r_j \in \{1, \ldots, j\}$, $j = 1, \ldots, n$, 有

$$\mathrm{P}\left(R_1 = r_1, R_2 = r_2, \ldots, R_n = r_n\right) = \frac{1}{n!} = \prod_{k=1}^{n} \mathrm{P}\left(R_k = r_k\right). \tag{1}$$

对取定的 $(r_1, \ldots, r_n)$, 必存在 $X_1, \ldots, X_n$ 的一种排序使得

$$\left\{R_1 = r_1, R_2 = r_2, \ldots, R_n = r_n\right\} = \left\{X_1 = X_{i_1:n}, \ldots, X_n = X_{i_n:n}\right\},$$

其中 $(i_1, \ldots, i_n)$ 为 $(1, 2, \ldots, n)$ 的某个置换. 注意到 F 连续, 所以

$$\mathrm{P}\left(X_1 = X_{i_1:n}, \ldots, X_n = X_{i_n:n}\right) = \frac{1}{n!}. \quad \blacksquare$$

作业

第 4 章第一次作业

§4.1:　2，3，5，8

§4.2:　6，7，11，12

第 4 章第二次作业

§4.3:　2

§4.4:　1，2，3，6，8，11，13