

高等概率论

菠萝包萝卜

2025 年 9 月 18 日

参考书: 《Probability: Theory and Examples》-Rick Durrett, 《概率论》-应坚刚, 何萍

1 初等概率论

1.1 历史

从略

1.2 计数

n 个中取 r 个

排列数: 无放回下是 $\frac{n!}{(n-r)!}$, 有放回下 n^r

组合数: $C_n^r = \binom{n}{r} = \frac{n!}{r!(n-r)!}$

1.3 古典概率论

Definition. 随机试验的样本空间 Ω , 其中 Ω 样本的个数有限且等可能出现, 那么对事件 $A \subset \Omega$ 定义

$$P(A) = \frac{|A|}{|\Omega|}$$

对于 $\Omega = \{\omega_1, \dots, \omega_n\}$, 有

$$p(\omega_i) = \frac{1}{n}$$

值得一提的是, 上述 A 是由若干基本事件 ω_i 组成的, $|A|$ 就是它所含 ω_i 的个数, 如果有 l 个, 那么

$$P(A) = \frac{l}{n}$$

样本空间 Ω 是全集; 事件的集合又叫集族 (簇) \mathcal{F} , 它是以 Ω 某些子集为元素的集合; 概率 P 是一个实值集函数, 它的定义域是 \mathcal{F} 值域是 \mathbb{R} 。

容易看出, 在古典的情况下, 我们的 P 总是满足

1) 非负: $\forall A \subset \Omega, P(A) \geq 0$

2) 正则: $P(\Omega) = 1$

3) 有限可加: $A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$

3') 可列可加: A_i 互不相容则 $P(\cup_i A_i) = \sum_i P(A_i)$

独立随机试验中的事件 $A \subset \Omega_1, B \subset \Omega_2, A \times B \subset \Omega = \Omega_1 \times \Omega_2$, 我们有

$$P(A \times B) = P_1(A)P_2(B)$$

Theorem 1.1. 容斥原理: 如果 $A_1 \cdots A_n \in \mathcal{F}$, 我们有

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k-1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} P(A_{i_1} \cap \cdots \cap A_{i_k})$$

Proof. 利用数学归纳法, $n = 2$ 的时候显然

假设 $n - 1$ 的时候成立, $P(\bigcup_{i=1}^{n-1} A_i \cup A_n) = P(\bigcup_{i=1}^{n-1} A_i) + P(A_n) - P(\bigcup_{i=1}^{n-1} A_i \cap A_n)$ 再利用集合性质即可。□

Problem 1.1. 匹配问题: n 个人取自己的帽子, 恰有 k 个人拿到自己的帽子的概率。

Solution. 我们先取定 k 个人, 有 C_n^k 种取法, 这 k 个人全部拿对, 则概率为 $\frac{1}{n \cdot (n-1) \cdots (n-k+1)}$, 其余的 $n - k$ 个人全拿错, 也就是全部没匹配上 $1 - \sum_{i=1}^{n-k} (-1)^{i-1} \frac{1}{i!} = \sum_{i=0}^{n-k} (-1)^i \frac{1}{i!}$, 组合起来概率就是

$$\frac{n!}{k!(n-k)!} \cdot \frac{(n-k)!}{n!} \sum_{i=0}^{n-k} (-1)^i \frac{1}{i!} = \frac{1}{k!} \sum_{i=0}^{n-k} (-1)^i \frac{1}{i!}$$

□

1.4 几何概型

Definition. 在一个有界区域 Ω 上随机地取一个点, 落在子区域 D 上点概率是

$$P(\text{落在 } D \text{ 中}) = \frac{m(D)}{m(\Omega)}$$

Problem 1.2. 约会问题: 从略

2 概率空间和随机变量的分布

2.1 概率空间

Definition 2.1. 三元体 (Ω, \mathcal{F}, P) 称为概率空间如果:

1. Ω 是所有结果的集合
2. \mathcal{F} 是 Ω 子集的非空集合
3. P 是定义在 \mathcal{F} 上的实值集函数, 值域是 $[0, 1]$
且 \mathcal{F} 是一个 σ -代数, 所谓 σ -代数是指
 - a. $\Omega \in \mathcal{F}$ 即全集在 \mathcal{F} 里面
 - b. 若 $A \in \mathcal{F} \Rightarrow \bar{A} \in \mathcal{F}$ 即对补封闭
 - c. 一系列 $A_i \in \mathcal{F}$ 有

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$$

即对可列并封闭

Definition 2.2. 可测空间 (measurable space): (Ω, \mathcal{F}) 的 \mathcal{F} 是 σ -代数, 且能够在 \mathcal{F} 上定义测度。称定义在 \mathcal{F} 上的实值集函数 μ 为测度, 如果 μ 满足

1. $\mu(A) \geq 0, \forall A \in \mathcal{F}$
2. 一系列两两不交的 $A_i \in \mathcal{F}$ 有

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

3. 如果 $\mu(\Omega) = 1$ 称 μ 为概率测度

Theorem 2.1. 设 μ 是 (Ω, \mathcal{F}) 上的一个测度, 则 μ 满足

1. 单调性: $A \subset B \Rightarrow \mu(A) \leq \mu(B)$
2. 次可加: $A \subset \bigcup_{i=1}^{\infty} A_i, \mu(A) \leq \mu\left(\bigcup_{i=1}^{\infty} A_i\right)$
3. 连续性: A_i 递增收敛于 A 则 $\mu(A_i) \rightarrow \mu(A)$, 递减相同

Proof. 1. $B = A \cup (B - A) \Rightarrow \mu(B) = \mu(A) + \mu(B - A) \geq \mu(A)$

2. 进入首次分解 $A'_n = A_n \cap A, B_1 = A'_1$; 当 $n > 1$ 令 $B_n = A'_n - \bigcup_{i=1}^{n-1} A'_i$ 从而它们互不相容, 容易证明一件事 $\bigcup_{i=1}^n B_i = \bigcup_{i=1}^n A'_i$ 对任意 n 成立, 以及 $\bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A'_i$
由于

$$A \subset \bigcup_{i=1}^{\infty} A_i \Rightarrow A = A \cap \left(\bigcup_{i=1}^{\infty} A_i\right) = \bigcup_{i=1}^{\infty} (A_i \cap A) = \bigcup_{i=1}^{\infty} A'_i$$

通过性质 1 以及 $B_i \subset A_i$ 我们就有

$$\mu(A) = \mu\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \mu(B_i) \leq \sum_{i=1}^{\infty} \mu(A_i)$$

3. 设 $B_n = A_n - A_{n-1}$ 于是他们两两不交, 有 $\bigcup_{i=1}^{\infty} B_i = A, \bigcup_{i=1}^n B_i = A_n$, 从而

$$\mu(A) = \sum_{m=1}^{\infty} \mu(B_m) = \lim_{n \rightarrow \infty} \sum_{i=1}^n \mu(B_i) = \lim_{n \rightarrow \infty} \mu(A_n)$$

□

Example 2.1. 如伯努利试验 $(\Omega = \{A, \bar{A}\}, \mathcal{F} = \{\emptyset, A, \bar{A}, \Omega\}, P)$, 其中

$$P(\emptyset) = 0, P(\Omega) = 1, P(A) = p, P(\bar{A}) = 1 - p$$

Claim 2.2. $\alpha \in A$ 若 \mathcal{F}_α 是 σ -代数, 则 $\bigcap_{\alpha \in A} \mathcal{F}_\alpha$ 仍然是 σ -代数

Proof. 1. 首先对于这些 \mathcal{F}_α 都有 $\Omega \in \mathcal{F}_\alpha$ 从而 $\Omega \in \bigcap_{\alpha \in A} \mathcal{F}_\alpha$

2. 设 $A \in \bigcap_{\alpha \in A} \mathcal{F}_\alpha$, 于是 $A \in \mathcal{F}_\alpha, \forall \alpha \in A$, 从而 $\bar{A} \in \mathcal{F}_\alpha, \forall \alpha \in A$, 因此 $\bar{A} \in \bigcap_{\alpha \in A} \mathcal{F}_\alpha$

3. 设 $A_1, \dots, A_n, \dots \in \bigcap_{\alpha \in A} \mathcal{F}_\alpha$, 于是它们同时属于 $\mathcal{F}_\alpha, \forall \alpha \in A$, 所以

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}_\alpha, \forall \alpha \in A$$

故

$$\bigcup_{i=1}^{\infty} A_i \in \bigcap_{\alpha \in A} \mathcal{F}_\alpha$$

□

Claim 2.3. 我们称 $\sigma(\mathcal{A})$ 是由 \mathcal{A} 生成的 σ -域, 如果 \mathcal{A} 是一个 Ω 的集类, $\sigma(\mathcal{A})$ 是所有 $A \subset \mathcal{F}_\alpha$ 的 \mathcal{F}_α 的交

2.2 随机变量以及其分布

Definition 2.3. 称 X 是一个随机变量, 如果 X 是一个 $\Omega \rightarrow \mathbb{R}$ 的实值函数, 如果它满足 $\forall A \in \mathcal{B}$, 满足

$$X^{-1}(A) = \{\omega : X(\omega) \in A\} \in \mathcal{F}$$

换句话说, 对 Borel 集 A , 使得 $X(\omega)$ 落在 A 上的这些 ω 形成的集合是在事件域 \mathcal{F} 中的, 我们就称 X 是 \mathcal{F} -可测的.

这里有一些要素, 就是在概率论中我们讨论的 \mathcal{F} 一般是不变的, 而在随机分析中为了方便测度变换, 我们的 \mathcal{F} 不一定是变化的.

Example 2.2. 示性函数是可测的 $I_A(\omega) = \begin{cases} 1, \omega \in A \\ 0, \omega \notin A \end{cases}$, 其中 $A \in \mathcal{F}$.

Definition 2.4. 设 X 是一个随机变量, 引入一个概率测度 $\mu(\cdot)$ 使得

$$\mu(A) = P(X \in A) = P(\{\omega : X(\omega) \in A\}) = P(X^{-1}(A))$$

其中 $A \in \mathcal{B}$, 称这样的测度为 X 的概率分布.

Definition 2.5. 称 $F(x)$ 是一个分布函数, 如果 P 是随机变量 X 的一个概率分布, 且

$$F(x) = P(X \leq x)$$

Theorem 2.4. 分布函数 F 的性质

- (i) 单调不减;
- (ii) $F(+\infty) = 1, F(-\infty) = 0$;
- (iii) 右连续 $\lim_{y \rightarrow x^+} F(y) = F(x)$
- (iv) $F(x-) = \lim_{y \rightarrow x^-} F(y) \Rightarrow F(x-) = P(X < x)$
- (v) $P(X = x) = F(x) - F(x-)$

Proof. (i) 如果 $x \leq y$ 有 $\{X \leq x\} \subset \{X \leq y\}$, 从而由概率测度的性质

$$F(x) \leq F(y)$$

(ii) 因为 $\{X \leq x\} \xrightarrow{x \rightarrow \infty} \Omega; \{X \leq x\} \xrightarrow{x \rightarrow -\infty} \emptyset$ 从而

$$F(+\infty) = 1, F(-\infty) = 0$$

(iii) 由于 $y \searrow x \Rightarrow \{X \leq y\} \searrow \{X \leq x\}$, 考虑 $y_n = x + \frac{1}{n}$ 它是从右边趋向 x 的, 即 $\bigcap_{n=1}^{\infty} \{X \leq y_n\} = \{X \leq x\}$ 从而

$$F(x) = P(X \leq x) = P\left(\bigcap_{n=1}^{\infty} \{X \leq y_n\}\right) = \lim_{n \rightarrow \infty} P(X \leq y_n) = \lim_{n \rightarrow \infty} F(y_n)$$

(iv) 取 $y_n = x - \frac{1}{n}$ 同上即可

(v) 由 $\{X = x\} = \{X \leq x\} \setminus \{X < x\}$ 立得. □

Theorem 2.5. 如果一个函数满足刚才的前三条性质, 则它是某个随机变量的分布函数.

Proof. 令 $\Omega = (0, 1), \mathcal{F} = \mathcal{B}, P$ 是 Lebesgue 测度, 并设

$$X(\omega) = \sup\{y : F(y) < \omega\}$$

于是如果有

$$\{\omega : X(\omega) \leq x\} = \{\omega : \omega \leq F(x)\}$$

那么结论就可以得证, 下面说明这件事

\Leftarrow 如果 $\omega \leq F(x)$ 于是 $x \notin \{y : F(y) < \omega\}$ 从而 $X(\omega) \leq x$

\Rightarrow 如果 $X(\omega) \leq x$ & $\omega > F(x)$ 于是由于 F 右连续, 存在一个 $\epsilon > 0, \omega > F(x + \epsilon)$

$\Rightarrow x + \epsilon \leq X(\omega) \Rightarrow x < x + \epsilon \leq X(\omega)$ 矛盾! □

Remark 2.1. 称刚才的 $X(\omega) = F^{-1}$ 为 F 广义逆, 并且有 $P(\{\omega : F^{-1}(\omega) \leq x\}) = F(x)$, 从而我们作一个随机变量 $\xi \sim U(0, 1)$, $F^{-1}(\xi)$ 就是以 F 为分布函数.

Definition 2.6. 称随机变量 X, Y 同分布, 如果

$$P(X \leq x) = P(Y \leq x)$$

对 $\forall x$ 成立.

Definition 2.7. 如果

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(y)dy$$

我们称 X 有密度函数 $f \Leftrightarrow f \geq 0, \int_{\mathbb{R}} f(x)dx = 1$

Remark 2.2. 在有密度的情况下 $P(X = x) = \lim_{\epsilon \rightarrow 0} \int_{x-\epsilon}^{x+\epsilon} f(y)dy = 0$

Definition 2.8. 称 X 是一个随机变量, 如果 X 是一个 $\Omega \rightarrow S$ 的可测映射, 它把 (Ω, \mathcal{F}) 映到了 (S, \mathcal{S}) , 如果它满足 $\forall A \in \mathcal{S}$, 满足

$$X^{-1}(A) = \{\omega : X(\omega) \in A\} \in \mathcal{F}$$

如果 $(S, \mathcal{S}) = (\mathbb{R}, \mathcal{B}^d)$, 称其为 d 维的随机向量。

Theorem 2.6. 如果 $\{\omega : X(\omega) \in A\} \in \mathcal{F}$ 对任意的 $A \in \mathcal{A}$, 其中 $\mathcal{S} = \sigma(\mathcal{A})$, 则 X 是 \mathcal{S} 可测的。

Proof. 注意到 $\{X \in B\} = \{\omega : X(\omega) \in B\}$, 我们有

$$\begin{aligned} \{X \in \cup_i B_i\} &= \cup_i \{X \in B_i\} \\ \{X \in B^c\} &= \{X \in B\}^c \end{aligned}$$

于是

$$\mathcal{B} = \{B : \{X \in B\} \in \mathcal{F}\}$$

是一个 σ -代数, 因为 $\mathcal{A} \subset \mathcal{B}$, 并且 \mathcal{A} 生成了 \mathcal{S} , 从而 $\mathcal{S} \subset \mathcal{B}$ □

由上我们得到了如果 \mathcal{S} 是 σ -代数, 则 $\{\{X \in B\} : B \in \mathcal{S}\}$ 是 σ -代数, 它是使得 X 是可测映射的最小的 Ω 上的 σ -代数, 我们称其为由 X 生成的 σ -域, 记作 $\sigma(X) = \{\{X \in B\} : B \in \mathcal{S}\}$

Definition 2.9. 如果 $X : (\Omega, \mathcal{F}) \rightarrow (S, \mathcal{S})$, $(S, \mathcal{S}) \rightarrow (T, \mathcal{T})$ 都是可测映射, 那么 $f \circ X$ 是 $(\Omega, \mathcal{F}) \rightarrow (T, \mathcal{T})$ 的可测映射

Proof. 令 $B \in \mathcal{T}$ 有

$$\{\omega \in \Omega : f(X(\omega)) \in B\} = \{\omega : X(\omega) \in f^{-1}(B)\} \in \mathcal{F} \Rightarrow f^{-1}(B) = \{S : f(S) \in B\} \in \mathcal{S}$$

□

Theorem 2.7. 如果 X_1, \dots, X_n 是随机变量, 并且, $f : (\mathbb{R}^n, \mathcal{R}) \rightarrow (\mathbb{R}, \mathcal{R})$ 是可测的, 那么 $f(X_1, \dots, X_n)$ 是一个随机变量

Proof. 由于

$$\{(X_1, \dots, X_n) \in A_1 \times \dots \times A_n\} = \bigcap_i \{X_i \in A_i\} \in \mathcal{F} \Rightarrow A_i$$

□