Probability and Statistics Tutorial 2

Siyi Wang

Southern University of Science and Technology 11951002@mail.sustech.edu.cn

September 21, 2020

Outline

- Review
- 2 Homework
- Supplement Exercises
- 4 Further Reading

Review

- 1. Definition of Probability Measure P in (Ω, \mathcal{F}, P)
 - $P(A) \ge 0$, for any $A \in \mathcal{F}$.
 - $P(\Omega) = 1$.
 - $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$, where $A_i \cap A_j = \emptyset$, for any $i \neq j$. (Countable Additivity)
- 2. Properties of Probability Measure P
 - $P(A) \in [0,1]$, for any $A \in \mathcal{F}$.
 - $P(\overline{A}) = 1 P(A)$, for any $A \in \mathcal{F}$. $(P(\emptyset) = 0)$
 - $P(\bigcup_{i=1}^{N} A_i) = \sum_{i=1}^{N} P(A_i)$, where $A_i \cap A_j = \emptyset$, for any $i \neq j$. (Finite Additivity)
 - If $A \subset B$, then $P(A) \leq P(B)$.
 - $P(A \cup B) = P(A) + P(B) P(A \cap B)$.



Review

- 2. Properties of Probability Measure P
 - $P(A_1 \cup A_2 \cup ... \cup A_n) = \sum_{i=1}^n P(A_i) \sum_{1 \leq i < j \leq n} P(A_i A_j) + ... + (-1)^{n-1} P(A_1 A_2 ... A_n)$. (Proof Method: (1) Induction (2) Indicator Function and Expectation.)
 - P(A B) = P(A) P(AB).
 - $P(A \cup B) \le P(A) + P(B)$.
- 3. Counting Methods
 - $\Omega = \{\omega_1, ..., \omega_N\}$ and $P(\omega_i) = 1/N$, for i = 1, 2, ..., N. Then, we have $P(A) = \frac{\text{number of outcomes that A contains}}{\text{number of all outcomes}}$, for $A \subset \Omega$.
 - Addition Principle. $N = m_1 + ... + m_s$.
 - Multiplication Principle. $N = m_1 \cdot m_2 \cdot ... \cdot m_p$.



P20, 4

4. 证明:

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leqslant \sum_{i=1}^{n} P(A_i)$$

Solution

Method 1. By the Supplement Exercises last time, let $B_1 = A_1$ and

$$B_k = A_k - \bigcup_{i=1}^{k-1} A_i \text{ for } k = 2, ..., n.$$

Then, we have
$$\bigcup_{i=1}^{n} B_i = \bigcup_{i=1}^{n} A_i$$
, where $B_i \cap B_j = \emptyset$, for $i \neq j$.

Also,
$$P(B_k) \leq P(A_k)$$
, for $k = 2, ..., n$.

Hence,
$$P(\bigcup_{i=1}^{n} A_i) = P(\bigcup_{i=1}^{n} B_i) = \sum_{i=1}^{n} P(B_i) \le \sum_{i=1}^{n} P(A_i)$$
.

Method 2. Since
$$P(A \cup B) \leq P(A) + P(B)$$
, then we have

$$P(\bigcup_{i=1}^{n} A_i) \le P(\bigcup_{i=1}^{n-1} A_i) + P(A_n) \le ... \le \sum_{i=1}^{n} P(A_i).$$



P20, 7

7. 证明邦费罗尼 (Bonferroni) 不等式:

$$P(A \cap B) \geqslant P(A) + P(B) - 1$$

Solution

Since
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
 and $0 \le P(A \cup B) \le 1$, then $P(A) + P(B) \ge P(A \cap B) \ge P(A) + P(B) - 1$.

P21, 28

山光明州平たタン・

28. 扑克游戏中, 5 个玩家从 52 张纸牌中每人分得 5 张. 共有多少种分法?

Solution

$$\binom{52}{5,5,5,5,5,27}. \ (\textit{or} \ C_{52}^5 \ C_{47}^5 \ C_{42}^5 \ C_{37}^5 \ C_{32}^5, \ \textit{or} \ \binom{52}{5} \binom{47}{5} \binom{42}{5} \binom{37}{5} \binom{27}{5}, \ \textit{or} \ \frac{52!}{5!5!5!5!5!27!})$$

P21, 29

29. 扑克玩家分到 3 个黑桃和 2 个红心. 他甩掉其中 2 个红心后再抽 2 张扑克. 他再抽到 2 个黑桃的概率是多少?

Solution

$$\frac{\binom{10}{2}}{\binom{47}{2}} = \frac{10*9}{47*46}.$$

12 / 49

Exercise 1

 $P(AC) = \frac{1}{9}$, 求A,B,C至少有一个发生的概率. \checkmark

Solution

Since $ABC \subset AB$ and P(AB) = 0, then P(ABC) = 0.

Hence, $P(A \cup B \cup C) =$

$$P(A) + P(B) + P(C) - P(AB) - P(BC) - P(AC) + P(ABC) = \frac{3}{4} - \frac{1}{8} = \frac{5}{8}$$

Exercise 2

2. 已知 A,B 两个事件满足条件 $P(AB) = P(\overline{AB})$,且 P(A) = p . 求 P(B).

Solution

We have
$$P(\overline{A}B) = P(\overline{A}) - P(\overline{A} \cap \overline{B}) = (1 - p) - P(\overline{A} \cap \overline{B})$$
.
Then, $P(B) = P(\overline{A}B) + P(AB) = (1 - p) - P(\overline{A} \cap \overline{B}) + P(AB) = 1 - p$.

Exercise 3

- 1. 从 n 双尺码不同的鞋子中任取 2r (2r≤n)只, 求下列事件的概率: ↔
 - 1) 所取 2r 只鞋子中没有两只成对;←
 - 2) 所取 2r 只鞋子中只有两只成对;←
 - 3) 所取 2r 只鞋子恰好配成 r 对.↩

Solution

(1)
$$\frac{C_n^{2r}2^{2r}}{C_n^{2r}}$$
.

$$(2)^{\frac{C_n^{2r-1}C_{2r-1}^12^{2r-2}}{C_{2n}^{2r}}} \text{ or } \frac{C_n^1C_{n-1}^{2r-2}2^{2r-2}}{C_{2n}^{2r}}$$

$$(3)\frac{C_n^r}{C_{2n}^{2r}}$$

Exercise 4

2. (匹配问题) 将 4 把能打开 4 间不同房门的钥匙随机发给 4 个人,试求至少有一人能 打开门的概率.↩

Solution

Method 1. Number of the outcomes that no one choose the right key=3*3=9. Then, $P=1-\frac{9}{24}=\frac{5}{8}$.

Method 2. Let A_i be the event that i-th key is given to the right people.

Then, we have
$$P(A_i) = \frac{A_3^3}{A_4^4} = \frac{1}{4}$$
, $P(A_i A_j) = \frac{A_2^2}{A_4^4} = \frac{1}{12}$,

$$P(A_i A_j A_k) = \frac{1}{A_4^4} = \frac{1}{24}$$
 and $P(A_i A_j A_k A_l) = \frac{1}{A_4^4} = \frac{1}{24}$.

Then, we have
$$P(A_1 \cup A_2 \cup A_3 \cup A_4) = \sum_{i} P(A_i) - \sum_{i < j} P(A_i A_j) + \cdots$$

$$\sum_{i < j < k} P(A_i A_j A_k) - P(A_1 A_2 A_3 A_4) = 1 - \frac{1}{2} + \frac{1}{6} - \frac{1}{24} = \frac{5}{8}.$$

Method 3. Consider the k keys case. Let M_k =number of the outcomes that no one choose the right key. We have $M_1=0$ and $M_2=1$. Also,

$$M_k = (k-1)M_{k-1} + (k-1)M_{k-2}$$
. Hence, $M_3 = 2$ and $M_4 = 9$. Hence,

$$P = 1 - \frac{9}{24} = \frac{5}{8}.$$

Exercise 4'

How about n keys rather than 4 keys?

Solution

Method 1. Let A_i be the event that i-th key is given to the right people.

Then, we have
$$P(A_i) = \frac{A_{n-1}^{n-1}}{A_n^n} = \frac{1}{n}$$
, $P(A_i A_j) = \frac{A_{n-2}^{n-2}}{A_n^n} = \frac{1}{n(n-1)}$, ..., $P(A_{i_1} A_{i_2} ... A_{i_k}) = \frac{1}{n(n-1)...(n-k+1)}$, ... and $P(A_i A_j A_k A_l) = \frac{1}{n!}$. Then, we have
$$P(A_1 \cup ... \cup A_n) = \sum_i P(A_i) - \sum_{i < j} P(A_i A_j) + ... + (-1)^{n-1} P(A_1 ... A_n) = \frac{C_n^1}{n} - \frac{C_n^2}{n(n-1)} + \frac{C_n^3}{n(n-1)(n-2)} + ... + (-1)^{n-1} \frac{C_n^n}{n!} = 1 - \frac{1}{2!} + \frac{1}{3!} - ... + (-1)^{n-1} \frac{1}{n!}$$
.

Remark. When $n \to \infty$, then $P(A_1 \cup ... \cup A_n) \to 1 - \frac{1}{e}$.

Solution

Method 2. Consider the k keys case. Let M_k =number of the outcomes that no one choose the right key. We have $M_1 = 0$ and $M_2 = 1$. Also,

$$M_k = (k-1)M_{k-1} + (k-1)M_{k-2}.$$

Let
$$F_k = \frac{M_k}{k!}$$
 Then, we have $F_k = \frac{k-1}{k} F_{k-1} + \frac{1}{k} F_{k-2}$.

Henceforth,
$$F_k - F_{k-1} = -\frac{1}{k}(F_{k-1} - F_{k-2})$$
.

Henceforth,
$$F_k - F_{k-1} = -\frac{1}{k}(F_{k-1} - F_{k-2})$$
.
Then, $F_k - F_{k-1} = (-1)^{k-2} \frac{1}{k(k-1)...3} (F_2 - F_1) = (-1)^k \frac{1}{k!}$.

Then,
$$F_k = \frac{1}{2!} - \frac{1}{3!} + ... + (-1)^k \frac{1}{k!}$$

Hence,
$$P(\text{no key matches}) = F_n = \frac{1}{2!} - \frac{1}{3!} + ... + (-1)^n \frac{1}{n!}$$
.

Therefore.

$$P(\text{at least one key matches}) = 1 - F_n = 1 - \frac{1}{2!} + \frac{1}{3!} - ... + (-1)^{n-1} \frac{1}{n!}.$$

Exercise 5

6

对任意事件 A, B, C 证明:

- $P(AB) + P(AC) P(BC) \le P(A)$;
- $P(AB) + P(AC) + P(BC) \ge P(A) + P(B) + P(C) 1$

Solution

Answer:

$$(1) \qquad (A \cap B) \cup (A \cap C) = A \cap (B \cup C)$$

$$P((A \cap B) \cup (A \cap C)) = P((A \cap B)) + P((A \cap C)) - P(ABC)$$

$$P(A) \ge P(A \cap (B \cup C)), P(BC) \ge P(ABC)$$

$$P(A) + P(BC) \ge P(A \cap (B \cup C)) + P(ABC)$$

结合上面所有式子可得。

(2) $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(BC) - P(AC) + P(ABC)$ 再结合所有概率在 [0,1] 之间可得。

Exercise 6

- 15. 设 A,B 是两事件,且 P(A) = 0.6, P(B) = 0.8,问:
- (1) 在什么条件下 P(AB) 取到最大值,最大值是多少?
- (2) 在什么条件下 P(AB) 取得最小值,最小值是多少?

Solution

解 (1) 因为 $P(AB) \le P(A) = 0.6$, $P(AB) \le P(B) = 0.8$, 所以当P(AB) = P(A) 时, P(AB) 的最大值是 0.6.

(2) 因为 $P(AB) = P(A) + P(B) - P(A \cup B) \ge P(A) + P(B) - 1 = 0.4$,所以有 $P(AB) \ge 0.4$.而当 $P(A \cup B) = 1$ 时,有P(AB)达到最小值0.4.

Exercise 7

23. 证明: $|P(AB) - P(A)P(B)| \leq \frac{1}{4}$.

Solution

证 不妨设 $P(A) \ge P(B)$,则

$$P(AB) - P(A)P(B) \le P(B) - P(B)P(B) = P(B)[1 - P(B)] \le \frac{1}{4}.$$

另一方面,还有

$$P(A)P(B) - P(AB) = P(A)[P(AB) + P(\overline{AB})] - P(AB)$$

$$= P(A)P(\overline{AB}) + P(AB)[P(A) - 1]$$

$$\leq P(A)P(\overline{AB}) \leq P(A)P(\overline{A}) = P(A)[1 - P(A)] \leq \frac{1}{A}.$$

综合上述两方面,可得

$$|P(AB) - P(A)P(B)| \leq \frac{1}{4}.$$

Exercise 8

17. 把 n 个"0"与 n 个"1"随机地排列,求没有两个"1"连在一起的概率.

Solution

解 考虑n个"1"的放法:2n个位置上"1"占有n个位置,所以共有 $\binom{2n}{n}$ 种

放法,这是分母. 而"没有两个 1 连在---起",相当于在 n 个"0"之间及两头(共n+1 个位置)去放"1",这共有 $\binom{n+1}{n}$ 种放法,于是所求概率为

$$p_n = \frac{\binom{n+1}{n}}{\binom{2n}{n}} = \frac{n+1}{\binom{2n}{n}}.$$

具体可算得 $p_3 = 0.2$, $p_5 = 0.0238$, $p_7 = 0.00233$. 随着n的增加,此种事件发生的概率愈来愈小,最后趋于零.

Exercise 9

0

13. 把 10 本书任意地放在书架上,求其中指定的四本书放在一起的概率.

Solution

解 10本书任意地放在书架上所有可能的放法数为10!,这是分母. 若把指定的四本书看作一本"厚"书,则与其他的6本书一起随意放,有7! 种可能放法,这是第一步. 第二步再考虑将这指定的四本书作全排列,共有4! 种可能放法. 故总共有7!×4! 种可能放法,这是分子. 于是所求概率为

$$\frac{7! \ 4!}{10!} = \frac{1}{30}$$

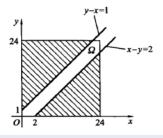
Exercise 10

on 213/ 30

24. 甲乙两艘轮船驶向一个不能同时停泊两艘轮船的码头,它们在一昼夜内到达的时间是等可能的. 如果甲船的停泊时间是一小时,乙船的停泊时间是两小时,求它们中任何一艘都不需要等候码头空出的概率是多少?

Solution

$$P(A) = \frac{S_A}{S_B} = \frac{\frac{1}{2}(23^2 + 22^2)}{24^2} = 0.879.$$



Exercise 11

3.-

- 22. 将n个完全相同的球(这时也称球是不可辨的)随机地放人N个盒子中,试求:
 - (1) 某个指定的盒子中恰好有 k 个球的概率;
 - (2) 恰好有 m 个空盒的概率;
 - (3) 某指定的 m 个盒子中恰好有 j 个球的概率.

Solution

解 先求样本点总数. 我们用N+1 根火柴棒排成一行,火柴棒之间的N个间隔恰好形成N个盒子,并依次称它们为第1个盒子,第2个盒子,…,第N个盒子,n个球用"0"表示,考虑到两端必须是火柴棒方能形成N个盒子,所以n个(不可辨)球放入N个(可辨)盒子中,就相当于把N-1 根火柴棒(N+1 根火柴棒中去掉两端的两根)和n个"0"随机地排成一行. 譬如N=4,n=3 时,"10010111"表示第1个盒子中有2个球、第2个盒子中有1个球、第3、4个盒子中无球. 这样一来,n个球放入N个盒子所有的样本点总数相当于:M0~1+n个位置任选n个位置放"0"、其他位置放火柴棒. 故样本点总数为(N+n-1).

(1) 记 A 为事件"指定的某个盒子中恰有 k 个球",不失一般性,可认为第 1 个盒子中有 k 个球,则余下 n-k 个球放人另外 N-1 个盒子中. 类似于样本点总数的计算,此种样本点共有 $\binom{N-1+n-k-1}{n-k}$,考虑到球不可辨故

4 0 5 4 40 5 4 2 5 4 2 5 2

Solution

$$P(A) = \frac{\binom{N+n-k-2}{n-k}}{\binom{N+n-1}{n}}, \qquad 0 \leq k \leq n.$$

(2) 记 B_m 为事件"恰有 m 个空盒". 它的发生可分两步描述:

第一步,从N个盒子任取m个盒子,共有 $\binom{N}{m}$ 种取法.

第二步,将n个球放人余下的N-m个盒中,且这N-m个盒子中都要有球.这当然要求 $n \ge N-m$ (或 $m \ge N-n$),否则第二步发生的概率为零.为了使第二步能发生,我们设想先把n个球排成一行,随机抽取球与球之间的n-1个间隔中的N-m-1个间隔放火柴棒即可,这有 $\binom{n-1}{N-m-1}$ 种可能.

综合上述两步,所求概率为

Solution

$$P(B_m) = \frac{\binom{N}{m} \binom{n-1}{N-m-1}}{\binom{N+n-1}{n}}, \qquad N-n \leq m \leq N-1.$$

(3) 若事件 C 表示"指定的 m 个盒子中恰有 j 个球",这意味着另外 N - m 个盒子中放 <math>n-j 个球. 由类似于样本点总数的计算知:j 个球放入 m 个盒子中共有 $\binom{m+j-1}{m-1}$ 种放法,而另外 n-j 个球放入 余下的 N-m 个盒子中有 $\binom{N-m+n-j-1}{n-j}$ 种放法. 于是所求概率为 $P(C) = \frac{\binom{m+j-1}{m-1}\binom{N-m+n-j-1}{n-j}}{\binom{N-m+n-j-1}{n-j}}, \qquad 1 \leq m \leq N, \quad 0 \leq j \leq n.$

Exercise 12

9

从 n 个不同元素中每次取出一个,放回后再取出下一个,如此连续取 r 次 所得的组合称为重复组合。

Solution

Answer:

我们可以把 n 个元素看作是 n 个盒子,第一个盒子放第一个元素,第二个盒子放第二个元素,依次推理下去,盒子里不放球相当于没有抽到对应的元素。如果 把 "o" 和" | " 看作一个位置,相当于从总位置 (n+r-1) 个选出 r 个位置给小 o,即 (n+r-1)

Exercise 13

31. (巴拿赫问题)某数学家有两盒火柴,每盒都有 n 根. 每次使用时,他任取一盒并从中抽出一根. 问他发现一盒空而另一盒还有 $r(0 \le r \le n)$ 根的概率是多少?

Solution

解 由对称性知,只要计算事件 E = "发现 A 盒空而 B 盒还有 r 根"的概率即可,所求概率是此概率的 2 倍.

先计算样本空间中的样本点个数. 因为每次都是等可能地取 A 盒或 B 盒, 共取了 2n-r+1 次, 故样本空间中共有 2^{2n-r+1} 个样本点.

事件 E 发生可分两段考察,前 2n-r 次中 A 盒恰好取到 n 次,且次序不论,最后一次(第 2n-r+1 次) 必定取到 A 盒,这样才能发现 A 盒已空,此种样本点共有 $\binom{2n-r}{n}$ 个,因此 $P(E)=\binom{2n-r}{n}/2^{2n-r+1}$. 所求概率为 $p=2P(E)=\binom{2n-r}{n}/2^{2n-r}$.

Exercise 14

$$= P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC).$$

25. 甲掷硬币 n+1 次, 乙掷 n 次. 求甲掷出的正面数比乙掷出的正面数多的概率.

Solution

解记

$$X_1 =$$
 甲掷出的正面数, $X_0 =$ 甲掷出的反面数 $= n + 1 - X_1$,

$$Y_1 =$$
乙掷出的正面数, $Y_0 =$ 乙掷出的反面数 = $n - Y_1$.

又记

$$E = \{X_1 > Y_1\}, \qquad F = \{X_0 > Y_0\},$$

由于正反面的地位是对称的,因此 P(E) = P(F). 又因为

$$F = \{X_0 > Y_0\} = \{n + 1 - X_1 > n - Y_1\}$$
$$= \{X_1 - 1 < Y_1\} = \{X_1 \le Y_1\} = \overline{E}.$$

所以由 $P(E) = P(F) = P(\overline{E})$, 得 P(E) = 0.5.

Further Reading

1. Benferroni Inequality and Inclusion-Exclusion Identity

For sets $A_1, A_2, \ldots A_n$, we create a new set of nested intersections as follows. Let

$$P_1 = \sum_{i=1}^n P(A_i)$$

$$P_2 = \sum_{1 \le i < j \le n}^n P(A_i \cap A_j)$$

$$P_3 = \sum_{1 \le i < j < k \le n}^n P(A_i \cap A_j \cap A_k)$$

÷

$$P_n = P(A_1 \cap A_2 \cap \cdots \cap A_n).$$

Than the inclusion explanion identity corn that

Further Reading

Then the inclusion-exclusion identity says that

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P_1 - P_2 + P_3 - P_4 + \cdots \pm P_n.$$

Moreover, the P_i are ordered in that $P_i \ge P_j$ if $i \le j$, and we have the sequence of upper and lower bounds

$$\begin{array}{rcl} P_1 \geq P(\cup_{i=1}^n A_i) & \geq & P_1 - P_2 \\ P_1 - P_2 + P_3 \geq P(\cup_{i=1}^n A_i) & \geq & P_1 - P_2 + P_3 - P_4 \\ & \vdots \end{array}$$

See Exercises 1.42 and 1.43 for details.

These bounds become increasingly tighter as the number of terms increases, and they provide a refinement of the original Bonferroni bounds. Applications of these bounds include approximating probabilities of runs (Karlin and Ost 1988) and multiple comparisons procedures (Naiman and Wynn 1992).

Further Reading

b. We illustrate the proof that the P_i are increasing by showing that $P_2 \geq P_3$. The other arguments are similar. Write

$$P_{2} = \sum_{1 \leq i < j \leq n} P(A_{i} \cap A_{j}) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} P(A_{i} \cap A_{j})$$

$$= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left[\sum_{k=1}^{n} P(A_{i} \cap A_{j} \cap A_{k}) + P(A_{i} \cap A_{j} \cap (\cup_{k} A_{k})^{c}) \right]$$

Now to get to P_3 we drop terms from this last expression. That is

$$\begin{split} &\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left[\sum_{k=1}^{n} P(A_i \cap A_j \cap A_k) + P(A_i \cap A_j \cap (\cup_k A_k)^c) \right] \\ &\geq \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left[\sum_{k=1}^{n} P(A_i \cap A_j \cap A_k) \right] \\ &\geq \sum_{i=1}^{n-2} \sum_{j=i+1}^{n-1} \sum_{k=j+1}^{n} P(A_i \cap A_j \cap A_k) &= \sum_{1 \leq i < j < k \leq n} P(A_i \cap A_j \cap A_k) &= P_3. \end{split}$$

The sequence of bounds is improving because the bounds $P_1, P_1 - P_2 + P_3, P_1 - P_2 + P_3 - P_4 + P_5, \ldots$, are getting smaller since $P_i \geq P_j$ if $i \leq j$ and therefore the terms $-P_{2k} + P_{2k+1} \leq 0$. The lower bounds $P_1 - P_2, P_1 - P_2 + P_3 - P_4, P_1 - P_2 + P_3 - P_4 + P_5 - P_6, \ldots$, are getting bigger since $P_i \geq P_j$ if $i \leq j$ and therefore the terms $P_{2k+1} - P_{2k} \geq 0$.

Thank you!