

第一章 随机事件及其概率（下）

Chapter 1 Random Events and Probability II

张昕

西南民族大学 经济学院

三月 2022 年春季学期

目录

- 1 条件概率
- 2 乘法公式
- 3 全概率公式
- 4 贝叶斯公式
- 5 独立

你在做选择题，不确定正确答案是哪一个，于是

| 答案 | A | B | C | D |
|----|-------|-------|-------|-------|
| 概率 | $1/4$ | $1/4$ | $1/4$ | $1/4$ |

现在你看了书后面的答案，答案说 A 是正确选项，于是

| 答案 | A | B | C | D |
|----|---|---|---|---|
| 概率 | 1 | 0 | 0 | 0 |

- ① 拥有更多的信息，就更接近真相；
- ② 拥有更多的信息，不符合条件的选项被排除，概率更新为 0；
- ③ 拥有更多的信息，概率的样本空间就发生了变化。

例

考察老师家中两个小朋友的性别，样本空间 $S = \{bb, bg, gb, gg\}$ ，其中 bg 代表大的孩子为男孩、小的是女孩，其他类似。求以下概率

- ① 家中至少一个女孩；
- ② 今天你看到老师带着男孩逛街，家中至少有一个女孩；
- ③ 家中有一男一女两个孩子，和上述事件有什么关系？

① $A = \{\text{家中至少一个女孩}\}, P(A) = \frac{3}{4};$

② $B = \{\text{家中至少一个男孩}\}$ ，当 B 发生时，样本空间变成 $S_B = \{bb, bg, gb\}$ ，此时 $P(A) = \frac{2}{3}$ 。我们将由于 B 发生后形成的新样本空间中 A 发生的概率记作 $P(A|B) = \frac{2}{3}$ 。

③ $AB = \{\text{家中一男一女两个孩子}\}, P(AB) = \frac{2}{4}.$

$$P(A|B) = \frac{2/4}{3/4} = \frac{P(AB)}{P(B)}.$$

条件概率

定义 (条件概率 (conditional Probability))

If $P(B) > 0$, then

$$P(A|B) = \frac{P(AB)}{P(B)}$$

定理

条件概率 $P(\cdot|B)$ 也是概率, 也满足概率三公理。

- ① 非负性: $0 \leq P(A|B) \leq 1$
- ② 完备性: $P(S|B) = 1$
- ③ 可列可加性: 对于互斥的事件 A_1, A_2, \dots (即 $A_i A_j = \emptyset$ 当 $i \neq j$), 则

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

条件概率

- $P(A|B)$ 称为“已知事件 B 发生的条件下, 事件 A 发生的**条件概率**”;
- 条件概率也是概率, 所以 $P(\cdot)$ 的性质对 $P(\cdot|B)$ 同样适用;
- $P(B) > 0$ 在形式上保证了分母不为 0, 可以理解为没必要以不会发生的 0 概率事件作为讨论的前提。

条件概率

例

我的钥匙找不到了，但是我有八成的把握我的钥匙在我外衣的衣兜里，我又确信有一半的概率在左衣兜，有另一半的概率在右衣兜。现在我去摸摸了我外衣的左衣兜却没发现钥匙，问钥匙在右衣兜的条件概率。

$$\begin{aligned} P(R|L^c) &= \frac{P(RL^c)}{P(L^c)} = \frac{P(R)}{1 - P(L)} \\ &= \frac{0.4}{1 - 0.4} \\ &= 2/3 \end{aligned}$$

“公式法”和“压缩样本空间法”的比较

例

投掷一枚硬币两次，问以下条件两枚硬币都朝上的概率。

- ① 第一次正面朝上
- ② 至少一次正面朝上

- ① 样本空间 $S=\{(H,H),(H,T),(T,H),(T,T)\}$, 事件“两枚硬币都朝上”记作 $A=\{(H,H)\}$, 事件“第一枚朝上”记作 $B=\{(H,H),(H,T)\}$ 事件“至少一枚朝上”记作 $C=\{(H,H),(H,T),(T,H)\}$ 。

条件概率

例 (橘 P20, 例 2)

3 红 2 白共 5 个球, 不放回抽 2 个, 已知第一次抽到红球, 问第二次抽到白球。

A 表示 “第一次取到红球”, B 表示 “第二次取到红球”, 那么 AB 就表示 “第一次取到红球且第二次取到红球”。

条件概率

例 (橘 P20, 例 2)

3 红 2 白共 5 个球，不放回抽 2 个，已知第一次抽到红球，问第二次抽到白球。

A 表示 “第一次取到红球”，B 表示 “第二次取到红球”，那么 AB 就表示 “第一次取到红球且第二次取到红球”。

$$\therefore P(AB) = \frac{3 \times 2}{5 \times 4} = \frac{3}{10}$$

$$P(A) = \frac{3 \times 4}{5 \times 4} = \frac{6}{10}$$

$$\therefore P(B|A) = \frac{P(AB)}{P(A)} = \frac{3}{6} = \frac{1}{2}$$

条件概率

例 (橘 P20, 例 2)

3 红 2 白共 5 个球，不放回抽 2 个，已知第一次抽到红球，问第二次抽到白球的概率。

条件概率会将不符合条件的“错误答案”排除，在新的样本空间重新计算概率，这种思路被称为“**压缩样本空间**”。

条件概率

例 (橘 P20, 例 2)

3 红 2 白共 5 个球，不放回抽 2 个，已知第一次抽到红球，问第二次抽到白球的概率。

条件概率会将不符合条件的“错误答案”排除，在新的样本空间重新计算概率，这种思路被称为“**压缩样本空间**”。

- 先计算，第一次抽到红球有几种可能。即 3×4 种。
- 再计算其中符合条件的有几种，即第一次抽到红球后第二次抽到白球，即 3×2 种。
- 根据古典概型

$$P(B|A) = \frac{\text{第一次抽到红球后第二次抽到白球}}{\text{第一次抽到红球后第二次无所谓}} = \frac{n(AB)}{n(A)} = \frac{3 \times 2}{3 \times 4} = \frac{1}{2}$$

条件概率

例 (橘 P20, 例 1)

3 黑 7 白共 10 个球，不放回地抽两个球，记 A 为第一次摸到黑球，B 为第二次摸到黑球，问

- ① 已知第一次摸到黑球，第二次又摸到黑球；
- ② 已知第二次摸到黑球，第一次也摸到黑球；
- ③ 两者相等吗？为什么？

目录

1 条件概率

2 乘法公式

3 全概率公式

4 贝叶斯公式

5 独立

条件概率与乘法公式

例

你要选课了，选我的课得 A 的概率是 $1/2$ ，选别的老师的课则得 A 的概率是 $2/3$ 。你犹豫不决，决定投个硬币来决定选谁的课。问你选了我的课并且得了 A 的概率。

- A: “成绩得 A”， B: “选我的课”， AB: “选我的课且得 A”。
- $P(AB) = P(B)P(A|B) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

条件概率与乘法公式

例

你要选课了，选我的课得 A 的概率是 $1/2$ ，选别的老师的课则得 A 的概率是 $2/3$ 。你犹豫不决，决定投个硬币来决定选谁的课。问你选了我的课并且得了 A 的概率。

- A: “成绩得 A”，B: “选我的课”，AB: “选我的课且得 A”。
- $P(AB) = P(B)P(A|B) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

例

8 红 4 白共 12 个球，无放回地取出两个球。问两个球都是红球的概率。
 A_i 表示第 i 次取到红球。

条件概率与乘法公式

例

你要选课了，选我的课得 A 的概率是 $1/2$ ，选别的老师的课则得 A 的概率是 $2/3$ 。你犹豫不决，决定投个硬币来决定选谁的课。问你选了我的课并且得了 A 的概率。

- A: “成绩得 A”, B: “选我的课”, AB: “选我的课且得 A”。
- $P(AB) = P(B)P(A|B) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

例

8 红 4 白共 12 个球，无放回地取出两个球。问两个球都是红球的概率。
 A_i 表示第 i 次取到红球。

- $P(A_1 A_2) = P(A_1)P(A_2|A_1) = \frac{2}{3} \times \frac{7}{11} = \frac{14}{33}$

回答了 “不放回地抽 2 个球” 和 “一次性抽 2 个球” 是一回事的原因。

乘法公式

定理 (乘法公式 (The multiplication rule))

$$P(AB) = P(B)P(A|B) = P(A)P(B|A)$$

$$P(A_1 A_2 A_3 \cdots A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 A_2) \cdots P(A_n|A_1 \cdots A_{n-1})$$

$$P(A_1) \frac{P(A_1 A_2)}{P(A_1)} \frac{P(A_1 A_2 A_3)}{P(A_1 A_2)} \cdots \frac{P(A_1 \cdots A_n)}{P(A_1 \cdots A_{n-1})} = P(A_1 A_2 A_3 \cdots A_n)$$

乘法公式

例 (橘 P21, 例 4; 同理蓝 P17, 例 4)

100 个灯泡, 其中 10 个是次品。不放回取 3 个, 问第三次才抽到正品。
 A_i 代表第 i 次抽到正品。

乘法公式

例 (橘 P21, 例 4; 同理蓝 P17, 例 4)

100 个灯泡, 其中 10 个是次品。不放回取 3 个, 问第三次才抽到正品。
 A_i 代表第 i 次抽到正品。

$$\begin{aligned} P(\bar{A}_1 \bar{A}_2 A_3) &= P(\bar{A}_1) \times \frac{P(\bar{A}_1 \bar{A}_2)}{P(\bar{A}_1)} \times \frac{P(\bar{A}_1 \bar{A}_2 A_3)}{P(\bar{A}_1 \bar{A}_2)} \\ &= P(\bar{A}_1) \times P(\bar{A}_2 | \bar{A}_1) \times P(A_3 | \bar{A}_1 \bar{A}_2) \\ &= \frac{10}{100} \times \frac{9}{99} \times \frac{90}{98} \\ &\approx 0.0083 \end{aligned}$$

乘法公式

例

把 52 张牌随机分成 4 堆，每堆 13 张，问每堆都恰好有一张 A 的概率。

$A_1 = \{\spadesuit A \text{ 放在任意一堆}\}$

$A_2 = \{\spadesuit A \text{ 和 } \heartsuit A \text{ 在不同堆}\}$

$A_3 = \{\spadesuit A, \heartsuit A \text{ 和 } \diamondsuit A \text{ 在不同堆}\}$

$A_4 = \{\spadesuit A, \heartsuit A, \diamondsuit A \text{ 和 } \clubsuit A \text{ 在不同堆}\}$

$$\begin{aligned} P(A_1 A_2 A_3 A_4) &= P(A_1) \frac{P(A_1 A_2)}{P(A_1)} \frac{P(A_1 A_2 A_3)}{P(A_1 A_2)} \frac{P(A_1 A_2 A_3 A_4)}{P(A_1 A_2 A_3)} \\ &= P(A_1) P(A_2 | A_1) P(A_3 | A_1 A_2) P(A_4 | A_1 A_2 A_3) \\ &= 1 \times \frac{52 - 13}{51} \times \frac{52 - 26}{50} \times \frac{13}{49} \\ &\approx 0.105 \end{aligned}$$

乘法公式

例

在 10 对夫妇中随机选出 5 人，问 5 人中一对夫妇都没有的概率。

乘法公式

例

在 10 对夫妇中随机选出 5 人，问 5 人中一对夫妇都没有的概率。

$$\frac{20}{20} \times \frac{18}{19} \times \frac{16}{18} \times \frac{14}{17} \times \frac{12}{16}$$

乘法公式

例

在 10 对夫妇中随机选出 5 人，问 5 人中一对夫妇都没有的概率。

$$\frac{20}{20} \times \frac{18}{19} \times \frac{16}{18} \times \frac{14}{17} \times \frac{12}{16}$$

例

现将 15 名新生平均分到 3 个班，其中 3 名优秀生，问每班 1 名优秀生的概率。

乘法公式

例

在 10 对夫妇中随机选出 5 人，问 5 人中一对夫妇都没有的概率。

$$\frac{20}{20} \times \frac{18}{19} \times \frac{16}{18} \times \frac{14}{17} \times \frac{12}{16}$$

例

现将 15 名新生平均分到 3 个班，其中 3 名优秀生，问每班 1 名优秀生的概率。

$$1 \times \frac{10}{14} \times \frac{5}{13} = \frac{25}{91}$$

乘法公式

例 (罐子问题; 蓝 P17, 例 3)

袋子中有 r 红 t 白个球, 每次取一个观察颜色并放回, 再放入 a 个与那只球同色的球, 若连续取 4 次。求第一、二次为红球, 且三、四次是白球的概率

A_i 代表第 i 次取到红球。

$$\begin{aligned} P(A_1 A_2 \bar{A}_3 \bar{A}_4) &= P(A_1) \times \frac{P(A_1 A_2)}{P(A_1)} \times \frac{P(A_1 A_2 \bar{A}_3)}{P(A_1 A_2)} \frac{P(A_1 A_2 \bar{A}_3 \bar{A}_4)}{P(A_1 A_2 \bar{A}_3)} \\ &= P(A_1) \times P(A_2 | A_1) \times P(\bar{A}_3 | A_1 A_2) \times P(\bar{A}_4 | A_1 A_2 \bar{A}_3) \\ &= \frac{r}{r+t} \times \frac{r+a}{r+t+a} \times \frac{t}{r+t+2a} \times \frac{t+a}{r+t+3a} \end{aligned}$$

例 (罐子问题; 蓝 P17, 例 3)

袋子中有 r 红 t 白个球，每次取一个观察颜色并放回，再放入 a 个与那只球同色的球，若连续取 4 次。求第一、二次为红球，且三、四次是白球的概率

- 当 $a = -1$, 这种情况为**不放回抽样**, 因为抽一次就少一个。我们已经知道, 此时概率不依赖抽出球的次序。
- 当 $a = 0$, 这种情况为**放回抽样**, 因为抽一次就少一个。我们已经知道, 此时每次抽球的概率不会改变。
- 当 $a > 0$, 这种情况为**传染病模型**, 增加同色球, 就相当于传染了一个新的病人。

目录

- 1 条件概率
- 2 乘法公式
- 3 全概率公式
- 4 贝叶斯公式
- 5 独立

全概率公式

定理 (乘法公式与全概率公式)

$$\begin{aligned} \because E &= E \cap S = E \cap (F \cup F^c) = EF \cup EF^c \\ \therefore P(E) &= P(EF) + P(EF^c) \\ &= P(E|F)P(F) + P(E|F^c)P(F^c) \end{aligned}$$

- 可以理解为事件 E 在 F 发生的条件下和 F 不发生的条件下的加权平均概率。
- $P(E|F) \leq P(E) \leq P(E|F^c)$ 或 $P(E|F^c) \leq P(E) \leq P(E|F)$ 必有一个成立。

全概率公式

例 (1.5.1)

- a 保险公司把客户分成两类，一类为容易出事故者，另一类为安全者。第一类一年内出事故的概率为 0.4，而另一类为 0.2。假如全部人口中容易出事故者占 30%。现在有一新客户来购买保险，问这个人在购买保险后一年内出事故的概率为多少。
- b 假如一个投保人在购买保险后的一年内出了事故，问他是容易出事故者的概率。

a A 表示 “容易出事故者”， A_1 表示 “1 年内出事故”。

$$P(A_1) = P(A_1|A)P(A) + (A_1|A^c)P(A^c) = 0.4 \times 0.3 + 0.2 \times 0.7 = 0.26$$

b

$$P(A|A_1) = \frac{P(AA_1)}{P(A_1)} = \frac{P(A)P(A_1|A)}{P(A_1)} = \frac{0.3 \times 0.4}{0.26} = \frac{6}{13}$$

全概率公式

例 (1.5.2)

你做选择题，你的策略是，如果会就直接答，如果不会就猜一个。现在你知道正确答案的概率是 p ，猜中的概率是 $1/m$ 。我判卷的时候发现你把这道题做对了，我就想知道你真正会做这道题的概率。

全概率公式

例 (1.5.2)

你做选择题，你的策略是，如果会就直接答，如果不会就猜一个。现在你知道正确答案的概率是 p ，猜中的概率是 $1/m$ 。我判卷的时候发现你把这道题做对了，我就想知道你真正会做这道题的概率。

C 表示 “这道题做对 (Correct)”， K 表示 “你会这道题 (Know)”。

$$\begin{aligned} P(K|C) &= \frac{P(KC)}{P(C)} \\ &= \frac{P(K)P(C|K)}{P(K)P(C|K) + P(K^c)P(C|K^c)} \\ &= \frac{p}{p + (1/m)(1 - p)} \\ &= \frac{mp}{1 + (m - 1)p} \end{aligned}$$

全概率公式

例 (1.5.3)

一种血液化验的新方法，有 95% 的把握能够把患有某癌症的病人筛查出来，但是该化验同样对健康人有 1% 的假阳性概率。假如目前人群中患有这种癌症的病人仅占 0.5%。现在某病人化验结果为阳性，问该病人确实罹患该癌症的概率。

D 表示 “患病 (Disease)”，E 表示 “检验阳性 (Positive)”。

全概率公式

例 (1.5.4)

本医生认为，如果我认为某病人罹患某病的可能性达到 80%，我会建议他手术，如果我不确定，我会建议他做进一步检查，但是该检查痛苦且昂贵。

现在老李来找我看病，我对他患该病的把握仅有 60%，于是我让老李去做检查，结果检查结果是阳性。正当我要做手术时，我发现老李隐瞒了病史，老李患有糖尿病。

这让我很不开心，虽然这并不影响我 60% 的判断，但却影响了检查的结果，因为这个又贵又痛苦的检查虽然不会给健康人给出阳性报告，但是对患有糖尿病且又患有该病的患者有 30% 的概率给出阳性报告。

请问，现在我该怎么办，是直接做手术，还是让老李重新做检查呢？

$$P(D|E) = \frac{P(DE)}{P(E)} = \frac{P(D)P(E|D)}{P(D)P(E|D) + P(D^c)P(E|D^c)} = \frac{(.6)1}{(.6)1 + (.4)(.3)} \approx .833$$

全概率公式

例 (1.5.5)

柯南本来有六成把握某人是犯罪嫌疑人。现在得知犯案的人是左撇子，而这个人恰好是个左撇子。已知，人群中左撇子的占比是 20%。问新的证据下，柯南认为这个人犯罪的把握有几成？

G 表示嫌疑人有罪 (Guilty), C 表示嫌疑人特征 (characteristic)。

全概率公式

例 (1.5.5)

柯南本来有六成把握某人是犯罪嫌疑人。现在得知犯案的人是左撇子，而这个人恰好是个左撇子。已知，人群中左撇子的占比是 20%。问新的证据下，柯南认为这个人犯罪的把握有几成？

G 表示嫌疑人有罪 (Guilty), C 表示嫌疑人特征 (characteristic)。

$$P(G|C) = \frac{P(GC)}{P(C)} = \frac{P(G)P(C|G)}{P(G)P(C|G) + P(G^c)P(C|G^c)} = \frac{(.6)1}{(.6)1 + (.4)(.2)} \approx .882$$

全概率公式

例 (1.5.6)

箱子 1 里有 n 个红球，箱子 2 里有 n 个蓝球。现在，从箱子 1 里随机拿走一个球，再从箱子 2 里拿一个球放到箱子 1 里。直到两个箱子的球都被取走，即箱子 1 取了 $2n$ 次，箱子 2 拿取了 n 次。问箱子 1 里取走的最后一个球是红球的概率。

F 表示某特定的红球的最后被取走， N_i 表示该特定的球不是第 i 次被取走的球。

$$\begin{aligned} P(F) &= P(N_1 N_2 \cdots N_n F) = \\ P(N_1) P(N_2 | N_1) \cdots P(N_n | N_1 \cdots N_{n-1}) P(F | N_1 \cdots N_n) &= \\ (1 - 1/n) \cdots (1 - 1/n) (1/n) \end{aligned}$$

R_j 表示 j 号红球最后被取走。

$$\begin{aligned} P(R_j) &= \left(1 - \frac{1}{n}\right)^n \frac{1}{n} \\ P(R) &= P\left(\sum_{j=1}^n R_j\right) = \sum_{j=1}^n P(R_j) = \left(1 - \frac{1}{n}\right)^n \approx e^{-1} \end{aligned}$$

Odd ratio

定义 (Odd ratio)

$$\text{Odd ratio} = \frac{P(A)}{P(A^c)}$$

当我们找到新的证据 B, 在 B 成立的条件下 A 发生与不发生的概率为:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \quad P(A^c|B) = \frac{P(B|A^c)P(A^c)}{P(B)}$$

则 A 新的 odd ration 为

$$\frac{P(A|B)}{P(A^c|B)} = \frac{P(A)}{P(A^c)} \frac{P(B|A)}{P(B|A^c)}$$

Odd ratio

例 (1.5.7)

硬币 A 有 $1/4$ 的概率正面朝上，硬币 B 有 $3/4$ 的概率正面朝上。现在箱子里有 A 和 B 硬币各一枚，我随机摸了一枚，抛了一次，结果正面朝上，问这枚硬币是 A 硬币的概率。

A 表示抛掷了 A 硬币，B 表示抛掷了 B 硬币，显然 $B = A^c$ 。

$$P(A|H) = \frac{P(A)P(H|A)}{P(H)} = 1/4$$

$$P(B|H) = \frac{P(B)P(H|B)}{P(H)} = 3/4$$

$$\text{Oddratio} = 1/3$$

Odd ratio

例 (1.5.5)

柯南本来有六成把握某人是犯罪嫌疑人。现在得知犯案的人是左撇子，而这个人恰好是个左撇子。已知，人群中左撇子的占比是 20%。新的证据对柯南破案是否有帮助。

G 表示嫌疑人有罪 (Guilty), C 表示嫌疑人特征 (characteristic)。

Odd ratio

例 (1.5.5)

柯南本来有六成把握某人是犯罪嫌疑人。现在得知犯案的人是左撇子，而这个人恰好是个左撇子。已知，人群中左撇子的占比是 20%。新的证据对柯南破案是否有帮助。

G 表示嫌疑人有罪 (Guilty), C 表示嫌疑人特征 (characteristic)。

$$\frac{P(G)}{P(G^c)} = \frac{0.6}{0.4} = \frac{3}{2}$$

$$\frac{P(G|C)}{P(G^c|C)} = \frac{P(G)}{P(G^c)} \frac{P(C|G)}{P(C|G^c)} = \frac{3}{2} \times \frac{1}{0.2} = \frac{15}{2}$$

目录

- 1 条件概率
- 2 乘法公式
- 3 全概率公式
- 4 贝叶斯公式
- 5 独立

全概率公式和贝叶斯公式

定理 (贝叶斯公式 (Bayes's formula))

$$P(F_j|E) = \frac{P(EF_j)}{P(E)} = \frac{P(E|F_j)P(F_j)}{\sum_{i=1}^n P(E|F_i)P(F_i)}$$

该公式的含义是我们可以根据试验的结果 (E) 来修订我们对试验的猜测 ($P(F_j)$)。

贝叶斯公式

例 (1.5.10)

我有三种电池，第一种连续使用超过 100 小时的概率达到了 0.7，另两种分别是 0.4 和 0.3。已知箱子里三种电池的占比是 2:3:5。

- ① 我随机拿了一个电池，问使用时长超过 100 小时的概率。
- ② 现在我随便拿的电池的使用时长已经超过 100 小时了，那么他是第 i 种电池的概率。

A 代表使用时长超过 100 小时， F_i 表示挑出第 i 种手电。

贝叶斯公式

例 (1.5.10)

我有三种电池，第一种连续使用超过 100 小时的概率达到了 0.7，另两种分别是 0.4 和 0.3。已知箱子里三种电池的占比是 2:3:5。

- ① 我随机拿了一个电池，问使用时长超过 100 小时的概率。
- ② 现在我随便拿的电池的使用时长已经超过 100 小时了，那么他是第 i 种电池的概率。

A 代表使用时长超过 100 小时， F_i 表示挑出第 i 种手电。

$$\begin{aligned} P(A) &= P(A|F_1)P(F_1) + P(A|F_2)P(F_2) + P(A|F_3)P(F_3) \\ &= .7 \times .2 + .4 \times .3 + .3 \times .5 = .41 \end{aligned}$$

贝叶斯公式

例 (1.5.10)

我有三种电池，第一种连续使用超过 100 小时的概率达到了 0.7，另两种分别是 0.4 和 0.3。已知箱子里三种电池的占比是 2:3:5。

- ① 我随机拿了一个电池，问使用时长超过 100 小时的概率。
- ② 现在我随便拿的电池的使用时长已经超过 100 小时了，那么他是第 i 种电池的概率。

A 代表使用时长超过 100 小时， F_i 表示挑出第 i 种手电。

贝叶斯公式

例 (1.5.10)

我有三种电池，第一种连续使用超过 100 小时的概率达到了 0.7，另两种分别是 0.4 和 0.3。已知箱子里三种电池的占比是 2:3:5。

- ① 我随机拿了一个电池，问使用时长超过 100 小时的概率。
- ② 现在我随便拿的电池的使用时长已经超过 100 小时了，那么他是第 i 种电池的概率。

A 代表使用时长超过 100 小时， F_i 表示挑出第 i 种手电。

$$\begin{aligned} P(F_i|A) &= P(AF_i)/P(A) = P(A|F_i)P(F_i)/0.41 \\ P(F_1|A) &= .7 \times .2/0.41 = 14/41 \\ P(F_2|A) &= .4 \times .3/0.41 = 12/41 \\ P(F_3|A) &= .3 \times .5/0.41 = 15/41 \end{aligned}$$

贝叶斯公式

例 (1.5.8)

一架飞机失踪了。现在推测该飞机会坠落在三个区域中的一个。如果对每个区域进行搜索，在该区域搜索到飞机的概率为 $1 - \beta_i$ 。搜索队已经完成对区域 1 的搜索，并未发现飞机。我们就可以根据这个信息去修订我们原来对三个区域搜索到飞机的概率。试问如何修订呢？直觉告诉我们，在第一区域没搜索到飞机，那飞机坠毁在第一区域的概率会下降，坠毁在其他区域的概率会增加。

R_i 表示飞机坠毁在区域 i (Region i)， E 表示搜索区域 1 没有发现飞机。

$$P(R_1|E) = \frac{P(ER_1)}{P(E)} = \frac{P(E|R_1)P(R_1)}{\sum_{i=1}^3 P(E|R_i)P(R_i)} = \frac{\beta_1}{\beta_1 + 2}$$

贝叶斯公式

例 (1.5.8)

一架飞机失踪了。现在推测该飞机会坠落在三个区域中的一个。如果对每个区域进行搜索，在该区域搜索到飞机的概率为 $1 - \beta_i$ 。搜索队已经完成对区域 1 的搜索，并未发现飞机。我们就可以根据这个信息去修订我们原来对三个区域搜索到飞机的概率。试问如何修订呢？直觉告诉我们，在第一区域没搜索到飞机，那飞机坠毁在第一区域的概率会下降，坠毁在其他区域的概率会增加。

$$P(R_j|E) = \frac{P(E|R_j)P(R_j)}{P(E)} = \frac{1}{\beta_1 + 2} \quad j = 2, 3$$

贝叶斯公式

例 (例 1.5.9)

我有三张形状相同的牌，第一张两面全是红色，第二张两面全是黑色，第三张一面红一面黑。现在随机拿出一张，向上的一面为红色，问你应该赌另一面是黑色呢还是红色呢？

贝叶斯公式

例 (例 1.5.9)

我有三张形状相同的牌，第一张两面全是红色，第二张两面全是黑色，第三张一面红一面黑。现在随机拿出一张，向上的一面为红色，问你应该赌另一面是黑色呢还是红色呢？

RR, RB, BB 分别代表抽出的牌是“两面红”，“一面红一面黑”，“两面黑”三个事件。

$$\begin{aligned} P(RB|R) &= \frac{P(RB \cap R)}{P(R)} \\ &= \frac{P(R|RB)P(RB)}{P(R|RR)P(RR) + P(R|RB)P(RB) + P(R|BB)P(BB)} \\ &= \frac{\frac{1}{2} \times \frac{1}{3}}{1 \times \frac{1}{3} + \frac{1}{2} \times \frac{1}{3} + 0 \times \frac{1}{3}} = \frac{1}{3} \end{aligned}$$

目录

- 1 条件概率
- 2 乘法公式
- 3 全概率公式
- 4 贝叶斯公式
- 5 独立

独立

定义 (独立 (Independent))

$$P(A|B) = P(A)$$
$$P(AB) = P(A|B)P(B) = P(A)P(B)$$

例 (例 1.6.1)

投掷两个硬币，第一个正面为上的概率是 $1/2$ ，第二个正面的概率是 $1/2$ ，两者是独立的，则两个都是正面的概率是 $1/4$ 。

- 互斥和独立是两个不同的概念。
 - 如果 $P(A) > 0, P(B) > 0$ 时，互斥和独立不能同时发生。
 - 若 $A = \emptyset$ 或 $B = \emptyset$ ，A 与 B 才可能既独立又互斥。

独立

例 (教材 P26, 例 1)

一副 52 张的扑克牌中任取一张, A 表示 “抽到 K”, B 表示 “抽到黑色”。问两者独立吗?

$$P(A) = \frac{4}{52} = \frac{1}{13}$$

$$P(B) = \frac{26}{52} = \frac{1}{2}$$

$$P(AB) = \frac{2}{52} = \frac{1}{26}$$

独立

例 (1.6.2)

投掷两个骰子, E_1 表示 “骰子之和是 6”, F 表示 “第一个骰子数字是 4”, 问两者独立吗? 现在 E_2 为 “两个骰子数字和是 7”, 问它和 F 独立吗?

$$P(E_1 F) = P(\{4, 2\}) = 1/36$$

$$P(E_1)P(F) = 5/36 \times 1/6 = 5/216$$

$$P(E_2 F) = P(\{4, 2\}) = 1/36$$

$$P(E_2)P(F) = 6/36 \times 1/6 = 1/36$$

例 (1.6.2)

$$\begin{aligned} P(E_1 F) &= P(\{4, 2\}) = 1/36 \\ P(E_1)P(F) &= 5/36 \times 1/6 = 5/216 \\ P(E_2 F) &= P(\{4, 2\}) = 1/36 \\ P(E_2)P(F) &= 6/36 \times 1/6 = 1/36 \end{aligned}$$

◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡ ▶ ↺ 🔍 ↻

独立

定理

- 如果 A 和 B 独立, 那么
 - A 和 B^c 独立;
 - A^c 和 B 独立;
 - A^c 和 B^c 独立

多个事件独立

例 (1.6.3)

如果 A 和 B 独立, A 也和 C 独立, 那么 A 和 BC 独立吗? 我们再来投掷两个骰子, A 表示 “骰子之和是 7”, B 表示 “第一个骰子为 4”, C 表示 “第二个骰子为 3”。已知 A 和 B 独立, 同理 A 和 C 也独立。但是很明显 A 和 BC 不独立, 因为 $P(A|BC) = 1$ 。

多个事件独立

例 (1.6.3)

如果 A 和 B 独立, A 也和 C 独立, 那么 A 和 BC 独立吗? 我们再来投掷两个骰子, A 表示 “骰子之和是 7”, B 表示 “第一个骰子为 4”, C 表示 “第二个骰子为 3”。已知 A 和 B 独立, 同理 A 和 C 也独立。但是很明显 A 和 BC 不独立, 因为 $P(A|BC) = 1$ 。

定义 (三个事件独立)

A, B and C are independent. \Leftrightarrow

$$P(AB) = P(A)P(B)$$

$$P(AC) = P(A)P(C)$$

$$P(BC) = P(B)P(C)$$

$$P(ABC) = P(A)P(B)P(C)$$

多个事件独立

定义 (多个事件相互独立)

A_1, A_2, \dots, A_n 中任意 k 个事件满足下式, 则称为以上事件**相互独立**。

$$P(A_{i_1}, A_{i_2}, \dots, A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k})$$

问上式共有多少项?

定义 (多个事件两两独立)

A_1, A_2, \dots, A_n 中任意两个事件相互独立, 称为以上事件**两两独立**。

- 若 A_1, A_2, \dots, A_n 相互独立, 则其中任意 k 个事件也相互独立。
- 若 A_1, A_2, \dots, A_n 相互独立, 则其中任意 k 个换成对立事件也成立。

多个事件独立

例 (1.6.4)

如果 A , B 和 C 相互独立, A 和 $B \cup C$ 独立吗?

多个事件独立

例 (1.6.4)

如果 A , B 和 C 相互独立, A 和 $B \cup C$ 独立吗?

$$\begin{aligned} P(A(B \cup C)) &= P(AB \cup AC) = P(AB) + P(AC) - P(ABC) \\ &= P(A)P(B) + P(A)P(C) - P(A)P(BC) \\ &= P(A)[P(B) + P(C) - P(BC)] \\ &= P(A)P(B \cup C) \end{aligned}$$

定理

如果 A , B 和 C 相互独立, A 、 B 、 C 的组合事件仍然独立。

多次独立重复试验

定义 (重复试验和伯努利试验)

一个大的试验由多个相互独立的子试验构成，且每个子试验都有相同的子样本空间及相同的概率结果。那么这种试验就被称为**重复试验 (Trials)**。

假如重复试验的每个子试验的只有发生或不发生两种互斥的结果，那么这样的重复试验称为**伯努利试验 (Bernoulli Trials)**，重复 n 次则称为 n **重伯努利试验**。

伯努利试验

例 (1.6.5)

一个重复 n 次的伯努利试验，成功的概率为 p ，不成功的概率为 $1-p$ ，问：

- ① 至少成功一次的概率
- ② 任意成功 k 次的概率
- ③ 全部成功的概率
- ④ 如果将伯努利试验重复无限次，那么全部成功的概率。

伯努利试验

例 (1.6.5)

一个重复 n 次的伯努利试验，成功的概率为 p ，不成功的概率为 $1-p$ ，问：

- ① 至少成功一次的概率
- ② 任意成功 k 次的概率
- ③ 全部成功的概率
- ④ 如果将伯努利试验重复无限次，那么全部成功的概率。

① $1 - (1 - p)^n$;

② $\binom{n}{k} p^k (1 - p)^{n-k}$;

③ $\binom{n}{n} p^n (1 - p)^0 = p^n$;

④ 如果 $p=1$ ，则为 1; 如果 $p<1$ ，则为 0。

n 重伯努利试验也称为伯努利概型，服从二项分布。

多次独立重复试验

例 (1.6.6)

进行独立重复试验，每次试验投掷两个骰子，并记录数字和。问“和为 5” 出现在“和为 7” 之前的概率。

E_n 表示，前 $n-1$ 次均没有出现 5 和 7，第 n 次才出现 5。
 $P(\{\text{和为 5}\}) = \frac{4}{36}$, $P(\{\text{和为 7}\}) = \frac{6}{36}$

$$P(E_n) = \left(1 - \frac{10}{36}\right)^{n-1} \times \frac{4}{36}$$
$$P\left(\bigcup_{n=1}^{\infty} E_n\right) = \frac{1}{9} \sum_{n=1}^{\infty} \left(\frac{13}{18}\right)^{n-1} = \frac{1}{9} \times \frac{1}{1 - 13/18} = \frac{2}{5}$$

多次独立重复试验

例 (1.6.6)

进行独立重复试验，每次试验投掷两个骰子，并记录数字和。问“和为5”出现在“和为7”之前的概率。

E 表示“和为5出现在和为7之前；A 表示“第一次投掷和为5”；
B 表示“第一次投掷和为7；C 表示“第一次投掷和既不是5也不是7”。

$$P(E) = P(A)P(E|A) + P(B)P(E|B) + P(C)P(E|C)$$

其中， $P(E|A)=1$ ， $P(E|B)=0$ 。而 $P(E|C)=P(E)$ ，相当于重新开始，又因为独立性，前面对后面的投掷没有影响。

$$P(E) = \frac{1}{9} + \frac{13}{18}P(E)$$

所以， $P(E)=2/5$

多次独立重复试验

例 (1.6.6)

进行独立重复试验，每次试验投掷两个骰子，并记录数字和。问“和为 5” 出现在“和为 7” 之前的概率。

- 和为 5 个概率是 $\frac{4}{36}$;
- 和为 7 个概率是 $\frac{6}{36}$;
- 所以“和为 5” 出现在“和为 7” 之前的比例应该是 $4/6$;
- 所以“和为 5” 出现在“和为 7” 之前的概率是 $\frac{4}{10}$;
- 若 A 和 B 是一次独立重复试验中两个互斥事件，那么事件 A 出现在事件 B 之前的概率是

$$\frac{P(A)}{P(A) + P(B)}$$

多次独立重复试验

例 (1.6.7: 氪金游戏)

我在开手办盲盒，每次开出手办 i 的概率为 p_i ，且 $\sum_{i=1}^n p_i = 1$ 。假定每次开盲盒都是独立。现在我已经开出 k 个手办了。 A_i 表示“至少有一个手办 i 了”，若 $i \neq j$ 。问：

- ① $P(A_i)$
- ② $P(A_i \cup A_j)$
- ③ $P(A_i | A_j)$

多次独立重复试验

例 (1.6.7: 氮金游戏)

我在开手办盲盒，每次开出手办 i 的概率为 p_i ，且 $\sum_{i=1}^n p_i = 1$ 。假定每次开盲盒都是独立。现在我已经开出 k 个手办了。 A_i 表示“至少有一个手办 i 了”，若 $i \neq j$ 。问：

- 1 $P(A_i)$
- 2 $P(A_i \cup A_j)$
- 3 $P(A_i | A_j)$

- 1 $P(A_i) = 1 - P(A_i^C) = 1 - (1 - p_i)^k$
- 2 $P(A_i \cup A_j) = 1 - P((A_i \cup A_j)^c) = 1 - (1 - p_i - p_j)^k$
- 3 $P(A_i A_j) = P(A_i) + P(A_j) - P(A_i \cup A_j) =$
 $1 - (1 - p_i)^k + 1 - (1 - p_j)^k - [1 - (1 - p_i - p_j)^k]$
 $P(A_i | A_j) = \frac{A_i A_j}{A_j}$

多次独立重复试验

例 (1.6.8: 点数问题 (Problem of points))

多次独立重复试验, 每次成功的概率为 p , 失败概率是 $1-p$ 。在第 m 次失败前已经成功 n 次的概率。假如每次成功 A 得一分, 失败则 B 的一分。如果 A 先得到 n 分 A 就获胜, 如果 B 先得到 m 分 B 就获胜。问 A 获胜的概率有多大?

多次独立重复试验

例

游戏有 r 个玩家，每个玩家 i 手里有 n_i 单位的赌资。赌局的每个阶段随机选两个人对局，胜者从败者处拿走 1 单位赌资。如果某玩家输光所有赌资就出局，直到仅剩一人时游戏结束，他就是最终获胜者。假设每场对局都是独立的，且玩家实力势均力敌，即每场对局两个玩家均有等概率击败对方。问玩家 i 获胜的概率。

例 (1.7.1)

保险公司把客户分成两类，一类为容易出事故者，另一类为安全者。第一类任意一年内出事故的概率为 0.4，而另一类为 0.2。假如全部人口中容易出事故者占 30%。现在某客户在第一年已经出了一次事故，问他在第二年再出事故的概率。

A 表示 “该客户是易出事故的人”， A_i 表示 “他在第 i 年出了一次事故”。

$$P(A_2|A_1) = P(A|A_1)P(A_2|AA_1) + P(A^c|A_1)P(A_2|A^cA_1)$$

$$P(A|A_1) = \frac{P(A_1A)}{P(A_1)} = \frac{P(A_1|A)P(A)}{P(A_1)} = \frac{0.4 \times 0.3}{0.26} = \frac{6}{13}$$

$$P(A^c|A_1) = 1 - P(A|A_1) = \frac{7}{13}$$

因为 $P(A_2|AA_1) = 0.4$, $P(A_2|A^cA_1) = 0.2$,

$$P(A_2|A_1) = 0.4 \times \frac{6}{13} + 0.2 \times \frac{7}{13} \approx 0.29$$

条件独立

定义 (条件独立)

如果在 B 已经发生的条件下, A_1 发生的概率不因 A_2 是否发生而改变, 即在给定 B 发生的条件下 A_1 和 A_2 独立, 我们就称给定 B 发生条件下 A_1 和 A_2 条件独立 (conditionally independent)。

$$P(A_1|A_2B) = P(A_1|B) \text{ or, equivalently } P(A_1A_2|B) = P(A_1|B)P(A_2|B)$$

先验概率与后验概率

- $P(B_j)$ 称为 B_i 的先验概率 (prior probability), 是我们对一个事件发生概率的初始判断。
- 事件 A 发生后带来了新的信息, 修正了我们对事件 B 发生概率的判断, 被称为后验概率 (posterior probability)。

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{P(A)} = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^n P(A|B_i)P(B_i)}$$

$$P(B_j|A_1 A_2) = \frac{P(A_1 A_2|B_j)P(B_j)}{\sum_{i=1}^n P(A_1 A_2|B_i)P(B_i)}$$

先验概率与后验概率

例 (狼来了)

伊索寓言中有一个 “狼来了” 的故事。

A 表示 “小朋友说谎”，B 表示 “小朋友可信”。假设村民们对小孩最初的判断为 $P(B)=0.8$ 。

$$P(B|A) = \frac{P(BA)}{P(A)} = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$$

假设 $P(A|B) = 0.1$ ，表示 “可信小朋友说谎概率为 0.1”， $P(A|B^c) = 0.5$ 。

$$P(B|A) = \frac{0.8 \times 0.1}{0.8 \times 0.1 + 0.2 \times 0.5} = 0.444$$

再用一次贝叶斯公式

$$P(B|AA) = \frac{0.444 \times 0.1}{0.444 \times 0.1 + 0.556 \times 0.5} = 0.138$$

谢谢!