概率统计第三讲：独立性

史灵生　清华数学系

1、本讲提要

1. 独立性
   - 两事件的独立性
   - 多事件的独立性
   - 应用
2. 相关性
   - 相关系数
   - 相关性质

2、独立性定义

- 如果两个事件其中任何一个事件发生与否都不改变另外一个事件发生的概率，具体地说，

$$P(A) = P(A|B) = P(A|\bar{B}), \quad P(B) = P(B|A) = P(B|\bar{A}),$$

- 那么我们认为它们在概率意义上与后者无关，这就是事件独立的本意。当然为使涉及的条件概率都有意义，我们需要假设$0 < P(A) < 1, 0 < P(B) < 1$。
- 经分析可发现,上面6个等式中任一个都蕴含其他5个.比如,
  - 由条件概率定义，$P(A|B) = P(A|\bar{B})$等价于 $P(AB)P(\bar{B}) = P(A\bar{B})P(B)$，
  - 而后者又等价于

$$\begin{aligned} P(AB) &= P(AB)P(\bar{B}) + P(AB)P(B) \\ &= P(A\bar{B})P(B) + P(AB)P(B) = P(A)P(B) \end{aligned}$$

  - 这个等式中A, B地位一样，它等价于$P(B|A) = P(B|\bar{A})$。
- 其他等式的等价性请作为练习自己证明。

3、独立性定义

定义

称事件A与B相互独立（或简称独立），如果

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

这里我们将最初的直观想法推广到零概率事件和满概率事件。总结以上分析，得到

性质

1. 零概率事件/满概率事件与任何事件独立。
2. 若A, B独立，则由它们各自产生的任一对事件也独立，即任何$C \in \mathcal{F}_A := \{\emptyset, \Omega, A, \bar{A}\}$与任何$D \in \mathcal{F}_B := \{\emptyset, \Omega, B, \bar{B}\}$独立。这里$\mathcal{F}_A$是包含$A$的最小事件域。

4、独立性

例

若A, B独立，则$A, \bar{B}$独立。

证明：

- $P(A) = P(AB) + P(A\bar{B})$，
- $$\begin{aligned} P(A\bar{B}) &= P(A) - P(AB) \\ &= P(A) - P(A)P(B) \\ &= P(A)[1 - P(B)] \\ &= P(A)P(\bar{B})。 \end{aligned}$$

5、双胞胎问题

双胞胎问题续

新搬来的邻居家有一对双胞胎小孩，其中有男孩与这对双胞胎是龙凤胎相互独立吗？

解

- 样本空间：{(男,男),(男,女),(女,男),(女,女)}，古典概型。
- 事件A表示这是一对龙凤胎，事件B表示至少有一个男孩，$A=\{(男,女),(女,男)\}$，$B=\{(男,男),(男,女),(女,男)\}$。
- 于是，$P(A|B)=2/3\neq 1/2=P(A)$，
- 或，$P(AB)=P(A)=1/2\neq 1/2\times 3/4=P(A)P(B)$，
- 不独立！

6、Monty Hall问题

Monty Hall问题续

在一个猜奖游戏中，主持人事先在A、B、C三个箱子中的一个箱子里藏了奖品。则奖品在A里与主持人在B、C两个箱子中选一个（比如B）打开示意其中无奖品相互独立吗？

解：

- 记A表示事件“奖品在A箱中”，类似定义事件B, C。
- 记B^*表示事件“主持人打开B箱示意其中无奖”。
- 则$P(A|B^*) = 1/3 = P(A)$，所以相互独立！

注：因为主持人事先知道奖品藏在哪个箱子里，所以，如果奖品在观众猜的箱子里，那么他可以打开剩下两个箱子中的任何一个示意其中无奖。

7、独立性

例

从一副牌中随机抽取一张，则此牌是“A”与此牌为“♠”相互独立吗？

解：

1. $P(A) = 4/54$，$P(\spadesuit) = 13/54$，$P(A\spadesuit) = 1/54$。

   $P(A\spadesuit) = 1/54 \neq 52/(54 \times 54) = P(A)P(\spadesuit)$。

   不独立！

2. 若不含两王牌，则

   $P(A\spadesuit) = 1/52 = 4/52 \times 13/52 = P(A)P(\spadesuit)$，

   独立！

注：　可见，随机事件的性质严重依赖于其生活的样本空间！

8、独立性

例

两射手彼此独立地向同一目标射击，设甲射中目标的概率为0.9，乙射中目标的概率为0.8，求目标被击中的概率为几？

解：

记 A为“甲射中目标”，B为“乙射中目标”，则$P(AB)=P(A)P(B)$。

1. $P(A\cup B)=P(A)+P(B)-P(AB)=P(A)+P(B)-P(A)P(B)$
$$=0.9+0.8-0.9\times 0.8=0.98\text{，或}$$
2. $P(A\cup B)=1-P(\overline{A\cup B})=1-P(\bar{A}\bar{B})=1-P(\bar{A})P(\bar{B})$
$$=1-(1-0.9)(1-0.8)=0.98\text{。}$$

9、多个事件的独立性

- 现在考虑多个事件的独立性。我们先来看三个事件的情形。
- 按直观想法，三个事件A, B, C相互独立是指其中任意一个事件的概率都不因由其他两个事件产生的任何事件的发生而改变。即A与任何$D \in \mathcal{F}_{B,C}$独立，其中

$$\begin{aligned}\mathcal{F}_{B,C} := \{&\emptyset, \Omega, B, C, \bar{B}, \bar{C}, BC, B\bar{C}, \bar{B}C, \bar{B}\bar{C},\\ &BC \cup \bar{B}\bar{C}, \bar{B}C \cup B\bar{C}, B \cup C, B \cup \bar{C}, \bar{B} \cup C, \bar{B} \cup \bar{C}\}\end{aligned}$$

  是含有B, C的最小事件域。
- 而且该结论对轮换A, B, C后依然成立。跟两个事件的独立性一样，这里有些独立性可以蕴含其他的独立性。三个事件的独立蕴含其中任意两个事件独立，并且一个事件和另两个事件的乘积事件独立。可以证明这两种说法是等价的。

10、独立性的定义

定义

- 称事件A, B, C相互独立（或简称独立），如果

$$
\begin{aligned}
P(AB) &= P(A)P(B), \\
P(BC) &= P(B)P(C), \\
P(CA) &= P(C)P(A), \\
P(ABC) &= P(A)P(B)P(C).
\end{aligned}
$$

- 称n个事件$A_1, A_2, \ldots, A_n$相互独立（或简称独立），如果对$1, 2, \ldots, n$中的任意一组指标$i_1, i_2, \ldots, i_k$，有
$$P(A_{i_1}A_{i_2}\cdots A_{i_k}) = P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k}).$$
- 称一组事件$\{A_\alpha\}$相互独立,如果其中任意有限个事件独立.
- 称事件组$\mathcal{A}_1, \ldots, \mathcal{A}_m$相互独立，如果任意$A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2, \ldots, A_m \in \mathcal{A}_m$相互独立。

11、独立性质

注：

- 多个事件两两独立并不能保证它们是相互独立的。
- 比如掷一颗正四面体的骰子，A表示得到1点或2点，B表示得到2点或3点，C表示得到3点或1点。则这三个事件两两独立，但是$ABC = \emptyset$，从而它们不相互独立。
- 因为$P(A) = P(B) = P(C) = 1/2$，
- $P(AB) = 1/4 = P(A)P(B)$，...
- 但是，$P(ABC) = 0 \neq 1/8 = P(A)P(B)P(C)$！

12、独立性质

定理

将相互独立的事件$A_1, \ldots, A_n$任意分组，则由各组事件分别产生的事件域$\mathcal{F}_1, \ldots, \mathcal{F}_m$相互独立。

推论

设事件$A_1, \ldots, A_n$相互独立。则$\overline{A_1}, \ldots, \overline{A_n}$也是相互独立的，并且

$$P(A_1 \cup \cdots \cup A_n) = 1 - [1 - P(A_1)] \cdots [1 - P(A_n)].$$

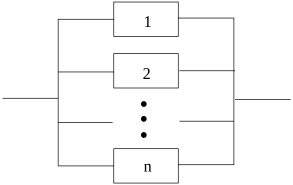
13、定理的证明

我们只对3个事件的情形证明，其方法可以适用于任意多个事件的情形。设A, B, C独立。我们证明A与任何$D \in \mathcal{F}_{B,C}$独立。

1. A, B, C独立$\Rightarrow A, BC$独立。这由定义即可得到。
2. 我们证明：A, B, C独立$\Rightarrow A, B, \bar{C}$独立。根据定义，由$A, B, C$独立知它们两两独立，知$A, B, \bar{C}$两两独立，而
$$\begin{aligned} P(AB\bar{C}) &= P(AB) - P(ABC) = P(A)P(B)[1 - P(C)] \\ &= P(A)P(B)P(\bar{C}). \end{aligned}$$
3. 因此$A, B, \bar{C}$相互独立。根据前两步，我们知道A与BC，$B\bar{C}$，$\bar{B}C$，$\bar{B}\bar{C}$中的每一个都是独立的，而$\mathcal{F}_{B,C}$中的非空集合都可以表示为这4个集合中若干个集合的并。下面的第三步使我们最终证明A与任何$D \in \mathcal{F}_{B,C}$独立
4. 我们证明：A, E独立，A, F独立，$EF = \emptyset \Rightarrow A, E \cup F$独立.
$$\begin{aligned} P[A(E \cup F)] &= P(AE \cup AF) = P(AE) + P(AF) \\ &= P(A)P(E) + P(A)P(F) = P(A)P(E \cup F)。 \end{aligned}$$

14、系统的可靠性

1. 串联系统：–□–□–...–□–
2. 并联系统：



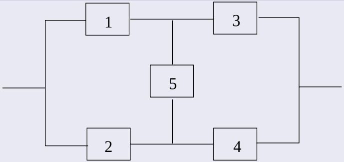
记 A_i= "第i个元件正常工作"，$p_i = P(A_i)$，则

$$p_{串} = P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} p_i,$$

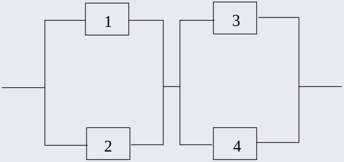
$$p_{并} = P\left(\bigcup_{i=1}^{n} A_i\right) = 1 - \prod_{i=1}^{n}(1 - p_i).$$

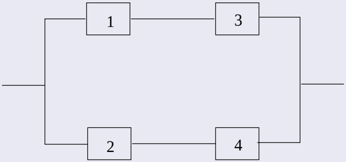
15、系统的可靠性

例 求下面系统的可靠性。



解：





$p_5[1-(1-p_1)(1-p_2)][1-(1-p_3)(1-p_4)]+(1-p_5)[1-(1-p_1p_3)(1-p_2p_4)]$。

16、小概率原则

例

在一个随机试验中事件A发生的概率为$\varepsilon > 0$。无论$\varepsilon > 0$多小，只要我们不断的独立重复做该试验，A迟早会发生的概率为1，或者不严格地说，A迟早会出现。

证明：

设A_k表示在第k次试验时A发生。则$A_1, \ldots, A_n, \ldots$相互独立，且$P(A_k) = \varepsilon$。

$$
\begin{aligned}
P\left(\bigcup_{k=1}^{\infty} A_k\right) &= \lim_{n\to\infty} P\left(\bigcup_{k=1}^{n} A_k\right) = \lim_{n\to\infty}\left\{1 - \prod_{k=1}^{n}[1 - P(A_k)]\right\} \\
&= 1 - \lim_{n\to\infty}(1-\varepsilon)^n = 1.
\end{aligned}
$$

17、小概率原则

例

在一个随机试验中事件A发生的概率为$\varepsilon > 0$。无论$\varepsilon > 0$多小，只要我们不断的独立重复做该试验，A迟早会发生的概率为1，或者不严格地说，A迟早会出现。

注：

- 小概率原则的一个有趣例子是：如果孙悟空不知疲倦乱敲键盘，那么它几乎注定会写出《西游记》。实际上，还可以证明：这猴子可以无穷多次写出这同一部名著。
- 因此，只要你不懈努力，机遇总会光顾你的。当然，学习可以使你以更大的概率更早地（在你有限的人生旅途中）见到成功的那一天。

18、抽样模型

例

口袋中有a个黑球与b个白球，每次随机地摸出一个球。问第一次摸出的是黑球与第二次摸出的是黑球相互独立吗？

解: 记$A_i =$“第i次取出的是黑球”，$i = 1, 2$。

1. 有放回抽样：$P(A_1A_2) = a^2/(a+b)^2 = P(A_1)P(A_2)$，故$A_1$与$A_2$相互独立。
2. 不放回抽样：

$$P(A_1) = a/(a+b)$$
$$P(A_1A_2) = P(A_1)P(A_2|A_1) = \tfrac{a}{a+b}\tfrac{a-1}{a+b-1}$$
$$\begin{aligned} P(A_2) &= P(A_1A_2) + P(\bar{A}_1A_2) \\ &= P(A_1)P(A_2|A_1) + P(\bar{A}_1)P(A_2|\bar{A}_1) \\ &= \tfrac{a}{a+b}\cdot\tfrac{a-1}{a+b-1} + \tfrac{b}{a+b}\tfrac{a}{a+b-1} = a/(a+b) \end{aligned}$$

$P(A_1A_2) \neq P(A_1)P(A_2)$，故$A_1$与$A_2$相关。

19、相关系数

注记：从例子可以看出，不同的抽样方式对独立性是有影响的，但当样本空间很大时，这种影响是可忽略的：

$$
\begin{aligned}
D &= |P(A_1A_2) - P(A_1)P(A_2)| = \left|\frac{a}{a+b}\left(\frac{a-1}{a+b-1} - \frac{a}{a+b}\right)\right| \\
&= \frac{ab}{(a+b)^2(a+b-1)} \le \frac{1}{4(a+b-1)} \to 0 \quad (\text{当} a+b \to \infty).
\end{aligned}
$$

定义

设$0 < P(A), P(B) < 1$。称

$$
r(A,B) = \frac{P(AB) - P(A)P(B)}{\sqrt{P(A)[1-P(A)]P(B)[1-P(B)]}}
$$

为事件A和B的相关系数。

20、相关性质

$$r(A,B)=\frac{P(AB)-P(A)P(B)}{\sqrt{P(A)[1-P(A)]P(B)[1-P(B)]}}$$

定理

1. $r(A,B)=0 \Leftrightarrow A$与$B$相互独立。
2. $-1\le r(A,B)\le 1$；
   $r(A,B)=1 \Leftrightarrow P(A)=P(AB)=P(B)$；
   $r(A,B)=-1 \Leftrightarrow P(A)=P(A\bar{B})=P(\bar{B})$。
3. $r(A,B)>0 \Leftrightarrow P(A|B)>P(A) \Leftrightarrow P(B|A)>P(B)$；
   $r(A,B)<0 \Leftrightarrow P(A|B)<P(A) \Leftrightarrow P(B|A)<P(B)$。

21、相关性质

（2）$-1 \leq r(A,B) = \dfrac{P(AB) - P(A)P(B)}{\sqrt{P(A)[1-P(A)]P(B)[1-P(B)]}} \leq 1$。

证明：（2）

- $P(AB) - P(A)P(B) \leq P(A)[1-P(B)]$
- $P(AB) - P(A)P(B) \leq P(B)[1-P(A)]$
- $P(AB) - P(A)P(B) \geq -P(A)P(B)$
- $P(AB) - P(A)P(B) \geq -[1-P(A)][1-P(B)]$
- 当$P(AB) - P(A)P(B) \geq 0$时，由前两个不等式得到（当$P(AB) - P(A)P(B) \leq 0$时，由后两个不等式得到）$[P(AB) - P(A)P(B)]^2 \leq P(A)[1-P(A)]P(B)[1-P(B)]$。