第四次习题课

王瑞

March 24, 2022

All generalizations are false, including this one. – Mark Twain

1 事件的独立性

需要掌握的内容包括事件独立的定义.

定义 1.1 两个事件 A 和 B 被称为**独立**的如果:

$$P(A \cap B) = P(A)P(B) \tag{1}$$

例 1.1 如果 A 和 B 是独立的事件, 那么 A^c 和 B, A 和 B^c, A^c 和 B^c 也是独立的.

例 1.2 A 和 B 是独立的事件等价于以下条件 *(假设 A 和 B 发生概率都大于 0):*

$$P(A|B) = P(A) \tag{2}$$

$$P(B|A) = P(B) \tag{3}$$

$$P(A|B) = P(A|B^c) \tag{4}$$

$$P(B|A) = P(B|A^c) \tag{5}$$

定义 1.2 两个事件 A 和 B 被称为**正相关**的如果:

$$P(A \cap B) > P(A)P(B) \tag{6}$$

例 1.3 试说明 *(6)* 和下列条件之间是等价的 *(假设 A 和 B 发生的概率均大于 0):*

$$P(A|B) > P(A) \tag{7}$$

$$P(B|A) > P(B) \tag{8}$$

$$P(A|B) > P(A|B^c) \tag{9}$$

$$P(B|A) > P(B|A^c) \tag{10}$$

定义 1.3 两个事件 A 和 B 被称为**条件在事件 C 上是正相关**的如果:

$$P(A \cap B|C) > P(A|C)P(B|C) \tag{11}$$

例 1.4 试说明 *(11)* 和下列条件之间是等价的 *(*假设 A 和 B 发生的概率均大于 *0,* 且 $P(A|B,C)$ 的含义是 $\frac{P(A\cap B|C)}{P(B|C)}$*):*

$$P(A|B,C) > P(A|C) \tag{12}$$
$$P(B|A,C) > P(B|C) \tag{13}$$
$$P(A|B,C) > P(A|B^c,C) \tag{14}$$
$$P(B|A,C) > P(B|A^c,C) \tag{15}$$

例 1.5 ***(*辛普森悖论*):***

$$\left\{\begin{array}{c} P(A|B,C) > P(A|B^c,C) \\ P(A|B,C^c) > P(A|B^c,C^c) \end{array}\right\} \nRightarrow P(A|B) > P(A|B^c) \tag{16}$$

为了看清楚上式背后的逻辑，我们考虑著名的伯克利大学录取的数据。为了简化起见，我们考虑只有两个系，分别是物理系和英语系：

物理系	接收	拒绝	条件概率
女生	60	40	$P(A|F,Ph) = 0.6$
男生	50	50	$P(A|M,Ph) = 0.5$

英语系	接收	拒绝	条件概率
女生	250	750	$P(A|F,En) = 0.25$
男生	20	80	$P(A|M,En) = 0.2$

合计	接收	拒绝	条件概率
女生	310	790	$P(A|F) = 0.28$
男生	70	130	$P(A|M) = 0.35$

可以看出，在英语系和物理系中，女生的录取率都比男生高，但是合计的数据中，女生的录取率却比男生低！(解释：大多数女生都申请了英语系，但是英语系的录取率更低！)

例 1.6 请说明在式 *(16)* 中，如果事件 B 和事件 C 是独立的，那么式 *(16)* 中的推出是成立的。

2 随机变量及其分布

需要掌握的内容包括离散型随机变量的概率分布，连续型随机变量的密度函数. 分布函数的定义，事件发生概率的计算.

例 2.1 在我们经常看到的新闻里，时常会以“万分之一”或者“百年一遇”的字眼来描述某事件发生的概率有多么低，从而吸引读者的眼球 *(*也许是因为大家总是对罕见的事情感兴趣*)*。然而如果我们仔细审视这些事件发生的概率，或许发现这些事件发生的概率并没有想象中那么低. 许多年前，新闻报道了美国纽约州的一个小学某年新入学的学生里有 *5* 个双胞胎，这则新闻被当时的媒体称为“几乎不可能发生”的事件。让我们来仔细地考虑下列的事件发生的概率：

1. 为了简化我们的计算，假设一个儿童是双胞胎的概率为 1/90，且不同儿童是否是双胞胎是独立的。我们假设小学一共有 *60* 个儿童入学，试计算这个小学有 *5* 个及以上双胞胎入学的概率。你认为这个概率是足够小以至于有新闻报道价值的吗？
2. 即使 *1.* 中的概率是很小的，而且有新闻报道价值，考虑“有 *5* 个及以上双胞胎入学”的事件可能在纽约州的任何一个县的任何一个学校发生。纽约州一共有 *62* 个县，而且假设每个县有 *5* 个类似的小学。请计算“纽约州至少有一个学校有 *5* 个及以上双胞胎入学的概率”，这个事件你觉得是“几乎不可能发生的事件”吗？
3. 如果你认为 *2.* 中计算的概率仍然是很小的，进一步考虑“有 *5* 个及以上双胞胎入学”的事件可能会在美国的 *50* 个州的任何一个，过去 *10* 年中的任何一年中发生至少一次的概率。

例 2.2 在许多生物医学和社会科学的研究中, 很多时候我们感兴趣的研究对象是“到某事件发生的时间”的分布。例如在生物医学的研究中，我们感兴趣的是服用某种药的病人从服药到死亡时间的分布；在社会科学的研究中，我们可能感兴趣具有某特征的一对夫妻从结婚到离婚的时间的分布. 这种研究关心的对象是非负随机变量, 它们的分布称为寿命分布. 若 X 是非负随机变量, 其分布函数是 $F(x)$, 密度函数是 $f(x)$, 为了简化起见，我们这里认为 X 服从指数分布，即 $f(x) = e^{-x}(x \geq 0), f(x) = 0(x < 0)$. 学者们为了研究寿命分布，还引入生存函数 $S(x) = P\{X > x\}$ 以及风险函数 $\lambda(x) = \frac{f(x)}{1-F(x)}$. 试以给出的指数分布验证下列的结论正确*:*

$$S'(x) = -f(x) \tag{17}$$

$$\log S(x) = -\int_0^x \lambda(t)dt \tag{18}$$

$$f(x) = \lambda(x)\exp[-\int_0^x \lambda(t)dt] \tag{19}$$

事实上，上式对任意的非负连续随机变量都是成立的！

3 随机变量的数字特征

此节需要掌握的内容包括连续型和离散型随机变量期望的计算和方差的计算，以及随机变量函数的数字特征的计算.

例 3.1 假设 A 代表一个我们感兴趣的事件，$P(A) = p$。定义随机变量 I_A 如下：

$$I_A = \begin{cases} 1 & (\text{如果 } A \text{ 发生}) \\ 0 & (\text{如果 } A \text{ 不发生}) \end{cases} \tag{20}$$

请计算 I_A 的期望和方差.

例 3.2 求服从下列概率分布的随机变量 X 的数学期望:

1.
$$f(x) = \begin{cases} \frac{1}{2}x^{-\frac{x}{2}}, & x > 0 \\ 0 & x \leq 0 \end{cases} \tag{21}$$

2.
$$f(x) = \frac{1}{4}e^{-\frac{|x|}{2}} \tag{22}$$

3.
$$f(x) = \begin{cases} 8x^{-3}, & x > 2 \\ 0, & x \leq 2 \end{cases} \tag{23}$$

例 3.3 *Betteley* 给出了一个有趣的期望的运算法则: 令 X 和 Y 分别代表两个随机变量, 定义 $X \wedge Y = min(X, Y)$, 以及 $X \vee Y = max(X, Y)$。类似于 $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, 我们有
$$E(X \vee Y) = EX + EY - E(X \wedge Y) \tag{24}$$

References

[1] Casella, George, and Roger L. Berger. Statistical inference(2nd ed). Cengage Learning, 2002.

[2] 李贤平. " 基础概率论 (第三版)." (2010).

[3] Shao, Jun. Mathematical statistics. Springer Science & Business Media, 2003.

[4] Michael, Perlman. Probability and mathematical statistics, `https://sites.stat.washington.edu/people/mdperlma/STAT%20512%20MDP%20Notes.pdf`