

# Homework 7

**Problem 1.** Find an example to verify the claim that ‘(pairwise) independence does not imply mutual independence’. Pls give a detailed proof.

*Solution.* One possible example can be like this,  $Pr(A) = \frac{1}{2}, Pr(B) = \frac{1}{2}, Pr(C) = \frac{1}{2}$  and  $Pr(A \cap B) = \frac{1}{4}, Pr(A \cap C) = \frac{1}{4}, Pr(B \cap C) = \frac{1}{4}, Pr(A \cap B \cap C) = \frac{1}{7}$ . We say  $A, B$  and  $C$  are pairwise independent since  $Pr(A \cap B) = Pr(A)Pr(B), Pr(A \cap C) = Pr(A)Pr(C)$  and  $Pr(B \cap C) = Pr(B)Pr(C)$ . However,  $A, B, C$  are not mutually independent, since  $Pr(A \cap B \cap C) \neq Pr(A)Pr(B)Pr(C)$ .  $\square$

**Problem 2.** Show that, if  $E_1, E_2, \dots, E_n$  are mutually independent, then so are  $\overline{E_1}, \overline{E_2}, \dots, \overline{E_n}$ .

*Solution.* Since  $E_1, E_2, \dots, E_n$  are mutually independent, for any subset  $I \subseteq [1, n]$ , we have  $Pr(\bigcap_{i \in I} E_i) = \prod_{i \in I} Pr(E_i)$ . And we have,  $Pr(\bigcap_{i \in I} \overline{E_i}) = 1 - Pr(\bigcup_{i \in I} E_i)$ . By inclusion-exclusion principle, we further have  $Pr(\bigcap_{i \in I} \overline{E_i}) = 1 - \sum_{i \in I} Pr(E_i) + \sum_{i \in I, j \in I, i < j} Pr(E_i \cap E_j) + \dots = \prod_{i \in I} (1 - Pr(E_i)) = \prod_{i \in I} Pr(\overline{E_i})$ . Thus,  $\overline{E_1}, \overline{E_2}, \dots, \overline{E_n}$  are mutually independent.  $\square$

**Problem 3.** The problem on the 37<sup>st</sup> page of slide on ‘Probability: a quick review’ (i.e., the more complicated example). (What is  $Pr(U|W)$ ?)

*Solution.*

$$\begin{aligned} Pr(U|W) &= \frac{Pr(U \cap W)}{Pr(W)} \\ &= \frac{Pr(R) \cdot Pr(U \cap W|R) + Pr(\neg R) \cdot Pr(U \cap W|\neg R)}{Pr(R) \cdot Pr(W|R) + Pr(\neg R) \cdot Pr(W|\neg R)} \end{aligned}$$

We have  $Pr(R) = 0.8$ , thus  $Pr(\neg R) = 0.2$  and  $Pr(U \cap W|R) = Pr(U|R) \cdot Pr(W|R), Pr(U \cap W|\neg R) = Pr(U|\neg R) \cdot Pr(W|\neg R)$ , then we can figure out that  $Pr(U|W) = 0.8$ .  $\square$

**Problem 4.** Suppose  $X$  and  $Y$  are two independent random variables, show that

$$E(X \cdot Y) = E[X] \cdot E[Y]$$

*Solution.* By definition we have

$$\begin{aligned}
 E[X \cdot Y] &= \sum (x \cdot y) \cdot Pr(X \cdot Y = x \cdot y) \\
 &= \sum_x \sum_y (x \cdot y) \cdot Pr(X = x \cap Y = y) \\
 &= \sum_x \sum_y (x \cdot y) \cdot Pr(X = x) \cdot Pr(Y = y) \\
 &= \sum_x x \cdot Pr(X = x) \sum_y y \cdot Pr(Y = y) \\
 &= E[X] \cdot E[Y]
 \end{aligned}$$

□

**Problem 5.** A monkey types on a 26 -letter keyboard that has lowercase letters only. Each letter is chosen independently and uniformly at random from the alphabet. If the monkey types 1,000,000 letters. what is the expected number of times the sequence “proof” appears?

*Solution.*

$$E[X] = \frac{1}{26} \times (1000000 - 4) = 0.08$$

□

**Problem 6** (Open question). Find a real-life case which is related to the so-called Simpson’s Paradox. Try to explain (the reason and the possible effects) and solve the ‘paradox’ in your case.

*Solution.* Assume we have carried out a survey at two high schools, the result is shown as below. We can see School A is better if we look at the data from

School	Student gender	Num	Average Score
A	Male	275	71.5%
	Female	25	62.0%
	Total	300	70.7%
B	Male	150	73.5%
	Female	150	64.5%
	Total	300	69.0%

the perspective of gender, but if we look at the total average score, then it seems school B is better. The reason is that in this survey we only choose 25 female students but 275 male students in school A. To solve this problem, the number of female students should be the same as that of male students that is, both are 150. □