

P8104 Probability - Homework 1

Yongyan Liu

2025-09-15

Problem 1

Let the sample space be defined as $C = \{1, 2, 3, 4, 5, 6\}$. Define the events: $A = \{2, 4, 6\}$, $B = \{1, 2, 3\}$.

(a) Compute the following set operations: $A \cup B$, $A \cap B$, A^c , and $(A \cup B)^c$.

- $A \cup B = \{1, 2, 3, 4, 6\}$
- $A \cap B = \{2\}$
- $A^c = \{1, 3, 5\}$
- $(A \cup B)^c = \{5\}$

(b) Assume each outcome in C is equally likely. Compute $P(A)$, $P(B)$, $P(A \cap B)$, and $P(A \cup B)$. Then, verify $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

- $P(A) = 3/6 = 1/2$
- $P(B) = 3/6 = 1/2$
- $P(A \cap B) = 1/6$
- $P(A \cup B) = 5/6$
- Verified $P(A) + P(B) - P(A \cap B) = 1/2 + 1/2 - 1/6 = 5/6 = P(A \cup B)$

(c) Define a set function Q that maps a subset $A \subseteq C$ to the number of even elements in A . That is, $Q(A) = \text{number of even numbers in } A$. Compute $Q(A)$, $Q(B)$, and $Q(A \cup B)$.

- $Q(A) = 3$
 - $Q(B) = 1$
 - $Q(A \cup B) = 3$
-

Problem 2

(a) An experiment consists of tossing a fair coin three times. Define the events A : the first toss is heads; and B : at least two heads occur in the three tosses.

- List the full sample space of the experiment
 - {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}
- Identify the outcomes that make up events A and B
 - $A = \{HHH, HHT, HTH, HTT\}$
 - $B = \{HHH, HHT, HTH, THH\}$
- Compute $P(B)$, $P(A \cap B)$, and $P(B | A)$
 - $P(B) = 4/8 = 1/2$
 - $P(A \cap B) = 3/8$
 - $P(B | A) = P(A \cap B)/P(A) = 3/4$

(b) A hand of 5 cards is drawn at random (without replacement) from a standard 52-card deck. Compute the probability that it contains all spades, *given that* it contains at least 4 spades.

Let event $A = \{\text{all five cards are spades}\}$ and $B = \{\text{at least 4 out of the 5 cards are spades}\}$

$$P(A) = \frac{\binom{13}{5}}{\binom{52}{5}}$$

$$P(B) = \frac{\binom{13}{5} + \binom{13}{4}\binom{39}{1}}{\binom{52}{5}}$$

Since $A \subseteq B$,

$$P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} = \frac{\binom{13}{5}}{\binom{13}{5} + \binom{13}{4}\binom{39}{1}} \approx 4.41\%$$

Problem 3

Suppose two events A and B satisfy $P(A) = 0.4$, $P(B) = 0.5$, and $P(A \cap B) = 0.2$.

(a) Are A and B independent? Justify using at least two equivalent definitions.

- Yes. They are independent. See the two verification below.
 - $P(A) * P(B) = 0.4 * 0.5 = 0.2 = P(A \cap B)$
 - $P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{0.2}{0.5} = 0.4 = P(A)$

(b) Are A and B mutually exclusive? Why or why not?

- No. A and B are not mutually exclusive, since $P(A \cap B) \neq 0$.

(c) Suppose C is independent of A , and $P(C) = 0.6$. What is $P(A \cap C)$?

- $P(A \cap C) = P(A) * P(C) = 0.4 * 0.6 = 0.24$

Problem 4

A rare disease affects 1 in 1,000 people. A diagnostic test for the disease has 98% sensitivity (true positive rate) and 97% specificity (true negative rate). That is,

- $P(\text{Test is positive} \mid \text{disease}) = 0.98$
- $P(\text{Test is negative} \mid \text{no disease}) = 0.97$

Let D be the event that a person has the disease, and T^+ be the event that the test result is positive.

(a) Compute the probability that a randomly selected individual tests positive.

Since $P(D) = 1/1000 = 0.1\%$, $P(D^c) = 1 - P(D) = 99.9\%$ and $P(T^+ \mid D^c) = 1 - P(T^- \mid D^c) = 0.03$, We have

$$P(T^+) = P(T^+ \mid D) * P(D) + P(T^+ \mid D^c) * P(D^c) = 0.98 * 0.1\% + 0.03 * 99.9\% = 3.095\%$$

(b) Compute the probability that a person actually has the disease given a positive test result, i.e., $P(D \mid T^+)$.

$$P(D \mid T^+) = \frac{P(D \cap T^+)}{P(T^+)} = \frac{P(T^+ \mid D) * P(D)}{P(T^+)} = \frac{0.98 * 0.1\%}{3.095\%} \approx 3.17\%$$

(c) Interpret the result from (b). What does it suggest about the usefulness of the test?

The test has a very high false positive rate. When the test is negative, we can almost be certain that the patient have no such a disease. But when the test is positive, the chance that the patient has the disease is also very slim.

The high false positive rate is mainly due to the rarity of the disease itself. The test is still useful to help filter out most of the healthy cases. For the cases which are tested positive, further diagnosis is required to confirm the disease.

Problem 5

Let E_1 , E_2 , and E_3 be three events with $P(E_1) = 0.3$, $P(E_2) = 0.4$, $P(E_3) = 0.5$.

(a) Use Boole's inequality to give an upper bound for $P(E_1 \cup E_2 \cup E_3)$.

- $P(E_1 \cup E_2 \cup E_3) \leq P(E_1) + P(E_2) + P(E_3) = 0.3 + 0.4 + 0.5 = 1.2$
- This is not useful though, because $P(E_1 \cup E_2 \cup E_3) \leq 1$ by definition.

(b) Suppose that $P(E_1 \cap E_2) = 0.12$, $P(E_1 \cap E_3) = 0.10$, and $P(E_2 \cap E_3) = 0.08$. Use the following second-order Bonferroni inequality

$$P\left(\bigcup_{i=1}^n A_i\right) \geq \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j)$$

to give a lower bound for the same union $P(E_1 \cup E_2 \cup E_3)$.

$$\begin{aligned} P(E_1 \cup E_2 \cup E_3) &\geq P(E_1) + P(E_2) + P(E_3) - P(E_1 \cap E_2) - P(E_1 \cap E_3) - P(E_2 \cap E_3) \\ &= 0.3 + 0.4 + 0.5 - 0.12 - 0.10 - 0.08 \\ &= 0.9 \end{aligned}$$

(c) Briefly interpret your bounds from (a)–(b). What extra information would tighten them further?

Based on (a) and (b), $0.9 \leq P(E_1 \cup E_2 \cup E_3) \leq 1$. We can tighten the bounds further if we know the value of $P(E_1 \cap E_2 \cap E_3)$.