

Please upload your work as a **single PDF file** to ELMS (under the "Assignments" tab)

- Submissions on paper or by email will not be accepted.
- Please do not submit your solutions as multiple separate files (pictures of individual pages). Such submissions are difficult to grade and will not be accepted.
- Justification of solutions is required.
- Each problem is worth 10 points.

**Problem 1.** A rare disease affects 1% of the population. A diagnostic test returns a positive result in 92% of infected individuals and in 4% of healthy individuals.

- (a) If a randomly selected person tests positive, what is the probability that they actually have the disease?  
(b) Suppose a person takes the test twice independently and both results are positive. What is now the probability that the person has the disease?

We write  $D$  for disease,  $\bar{D}$  for no disease;  $P, N$  for positive / negative

Given,  $P(D) = 0.01$ ;  $P(P|D) = 0.92$ ;  $P(P|\bar{D}) = 0.04$

(a) Compute using LOTP

$$P(P) = P(P|D)P(D) + P(P|\bar{D})P(\bar{D}) \\ = 0.92 \cdot 0.01 + 0.04 \cdot 0.99 = 0.049$$

Bayes formula:

$$P(D|P) = \frac{P(P|D)P(D)}{P(P)} = \frac{0.92 \cdot 0.01}{0.049} = \frac{0.0092}{0.049} = \underline{0.189}$$

(b) LOTP:

$$P(P, P_2) = P(P, P_2|D)P(D) + P(P, P_2|\bar{D})P(\bar{D}) = (0.92)^2 \cdot 0.01 + (0.04)^2 \cdot 0.99 \approx 0.01$$

Bayes formula:

$$P(D|P, P_2) = \frac{P(P, P_2|D)P(D)}{P(P, P_2)} = \frac{(0.92)^2 \cdot 0.01}{0.01} \approx \underline{0.84}$$

**Problem 2.** A company has two departments, Research (R) and Marketing (M). Overall, 60% of employees work in R and 40% in M. Among employees in R, 70% work remotely at least two days per week, while among employees in M, 30% do so.

(a) What is the probability that a randomly chosen employee works remotely at least two days per week?

(b) Given that an employee works remotely at least two days per week, what is the probability that they are in Research?

(c) Consider the events  $E_1$  "works in Research" and  $E_2$  "works remotely at least two days per week". Is it true that  $P(E_1)P(E_2) = P(E_1E_2)$ ?

(d) A pair of employees are chosen independently at random. What is the probability that they work in different departments?

We write  $r$  for remote,  $\bar{r}$  for not remote

$$(a) \text{ LOTA: } P(r) = P(r|R)P(R) + P(r|M)P(M) = 0.7 \cdot 0.6 + 0.3 \cdot 0.4 = 0.54$$

$$(b) \text{ Bayes: } P(R|r) = \frac{P(r|R)P(R)}{P(r)} = \frac{0.7 \cdot 0.6}{0.54} \approx 0.778$$

$$(c) \quad P(E_1E_2) = P(E_2|E_1)P(E_1) = P(r|R)P(R) = 0.7 \cdot 0.6 = 0.42$$

$$P(E_1)P(E_2) = P(R)P(r) = 0.6 \cdot 0.54 \neq 0.42 \quad \text{Ans: No}$$

$$(d) \quad P(\text{both in the same dept}) = P(R_2|R_1)P(R_1) + P(M_2|M_1)P(M_1)$$

$$= 0.6^2 + 0.4^2 = 0.36 + 0.16 = 0.52$$

$$P(\text{different}) = 1 - 0.52 = 0.48$$

**Problem 3.** (A variant of the Monty Hall problem)

There are 5 closed boxes; exactly one contains a prize. You choose one box (it stays closed). The host, who knows where the prize is, opens two empty boxes among the remaining four and then offers you the option to switch to one of the two unopened boxes other than your original choice (you choose randomly between those two if you switch).

- (a) What is the probability of winning if you keep your original choice?
- (b) What is the probability of winning if you switch?
- (c) Which strategy is better?

(a)  $\frac{1}{5}$  b/c the initial probability is  $\frac{1}{5}$  and nothing changes

(b) Let  $B_i = \{\text{prize in box } i\}$ ,  $i = 1, 2, \dots, 5$

Without loss of generality (the game is symmetric)  
assume that you choose Box 1.

LOTP

$$\begin{aligned} P(\text{win} | \text{switching}) &= P(W|B_1)P(B_1) + \dots + P(W|B_5)P(B_5) \\ &= (0 + 4 \cdot \frac{1}{2}) \cdot \frac{1}{5} = \frac{2}{5} \end{aligned}$$

(c) Switching is better

**Problem 4.** Two players alternate rolling a fair die until the first 6 appears. The player who rolls the first 6 wins. Player A rolls first.

Is  $P(\text{A wins}) >, =, \text{ or } < 1/2$ ?

- (1) Let  $p = P(\text{A wins})$ . We compute  $p$  by conditioning on what happens in the first two rolls (A's roll followed by B's roll).
- (2) What is the probability that A wins immediately on the first roll?
- (3) What is the probability that A does *not* win on the first roll and B wins on the second roll?
- (4) What is the probability that neither A nor B rolls a 6 in the first two rolls?
- (5) Use the hint: if neither player rolls a 6 in the first two rolls, the situation is exactly the same as at the beginning of the game. Explain why, in this case, the probability that A eventually wins is again  $p$ .
- (6) Combine the previous steps to write an equation for  $p$  of the form
$$p = (\text{probability A wins immediately}) + (\text{probability no 6 in first two rolls}) \cdot p.$$
- (7) Find  $p$  and argue whether A has an advantage.
- (8) Simulate the game by computer to confirm your answer. What is the smallest number of trials you used to match the answer to within  $10^{-2}$ ?  $10^{-3}$ ? You can use any software package. Include the code in your paper.

$$(2) \quad 1/6$$

$$(3) \quad 5/6 \cdot 1/6 = 5/36$$

$$(4) \quad \frac{25}{36}$$

(5) The game does not have memory, so if it goes into the 3<sup>rd</sup> round, it resets to the initial state.

$$(6) \quad p = \frac{1}{6} + \frac{25}{36} p$$

$$\therefore p = \frac{6}{11} > 1/2$$

so A has an advantage

**Problem 5.** An urn contains 4 red balls and 6 blue balls. Three balls are drawn without replacement.

- (a) What is the probability that exactly two of the three balls are red?  
 (b) Given that at least one of the three balls is red, what is the probability that all three are red?  
 (c) Suppose instead the balls are drawn with replacement. How do your answers to (a) and (b) change?

$$(a) \quad P(2 \text{ red}) = \frac{\binom{4}{2} \binom{6}{1}}{\binom{10}{3}} = \frac{36 \cdot 6}{10 \cdot 9 \cdot 8} = \frac{3}{10}$$

In the notation of L8, this is the hypergeometric distribution  $\text{Hgeom}(4, 6, 3)$

$$(b) \quad P(\geq 1 \text{ red}) = 1 - P(\text{all blue}) = 1 - \frac{\binom{6}{3}}{\binom{10}{3}} = 1 - \frac{6 \cdot 5 \cdot 4}{10 \cdot 9 \cdot 8} = 1 - \frac{5}{30} = \frac{25}{30}$$

$$P(3 \text{ red}) = \frac{\binom{4}{3}}{\binom{10}{3}} = \frac{4 \cdot 6}{10 \cdot 9 \cdot 8} = \frac{1}{30}$$

$$P(3 \text{ red} | \geq 1 \text{ red}) = \frac{P(\{3 \text{ red}\} \cap \{\geq 1 \text{ red}\})}{P(\geq 1 \text{ red})} = \frac{1/30}{25/30} = \frac{1}{25}$$

(c) recompute (a) with replacement

$$P(\text{red}) = 0.4; \quad P(\text{blue}) = 0.6$$

$$P(2 \text{ red}) = \binom{3}{2} 0.4^2 \cdot 0.6 = 0.288$$

recompute (b)

$$P(3 \text{ red} | \geq 1 \text{ red}) = \frac{0.4^3}{1 - 0.6^3} \approx 0.0816$$

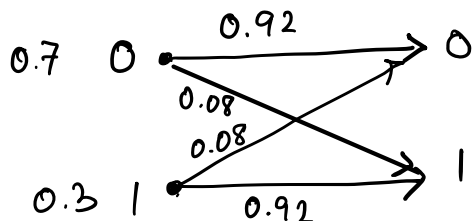
Note that 2 red is easier w/out replacement while  $\{3 \text{ red} | \geq 1 \text{ red}\}$  has better odds with replacement

**Problem 6.** A message bit  $X \in \{0, 1\}$  is transmitted over a noisy channel. The sender chooses  $X = 1$  with probability 0.3 and  $X = 0$  with probability 0.7. The channel flips a transmitted bit independently with probability 0.08.

(a) If the receiver observes  $Y = 1$ , what is the posterior (conditional) probability that  $X = 1$ ? In other words, find  $P(X = 1 | Y = 1)$

(b) Suppose now that each bit is transmitted three times independently (i.e.,  $X$  is sent as  $XXX$ ), and the receiver decides by majority rule. What is the probability that the receiver decodes incorrectly?

(c) Compare the error probability in part (b) with the error probability when a single transmission is used.



$$\begin{aligned}
 \text{(a)} \quad P(X=1 | Y=1) &= \frac{P(Y=1 | X=1) P(X=1)}{P(Y=1 | X=1) P(X=1) + P(Y=1 | X=0) P(X=0)} \\
 &= \frac{0.92 \cdot 0.3}{0.92 \cdot 0.3 + 0.08 \cdot 0.7} \approx 0.831
 \end{aligned}$$

(b) Since the action of error does not depend on the input

$$P(\text{error} | X=0) = P(\text{error} | X=1)$$

$$\begin{aligned}
 \therefore P(\text{error}) &= P(\text{error} | X=0) P(X=0) + P(\text{error} | X=1) P(X=1) \\
 &= P(\text{error} | X=0) = P(\geq 2 \text{ bits received incorrectly}) \\
 &= \binom{3}{2} 0.08^2 \cdot 0.92 + 0.08^3 \approx 0.0182
 \end{aligned}$$

(c) Let us compute  $P(\text{error})$  in a single transmission:

$$\begin{aligned}
 P(E) &= P(E | X=0) P(X=0) + P(E | X=1) P(X=1) \\
 &= P(E | X=0) = 0.08
 \end{aligned}$$

$\therefore$  Repeating helps