

Name: Solution
Section:

each problem graded out of 10
final grade awarded out of 20 as sum of best 2 scores

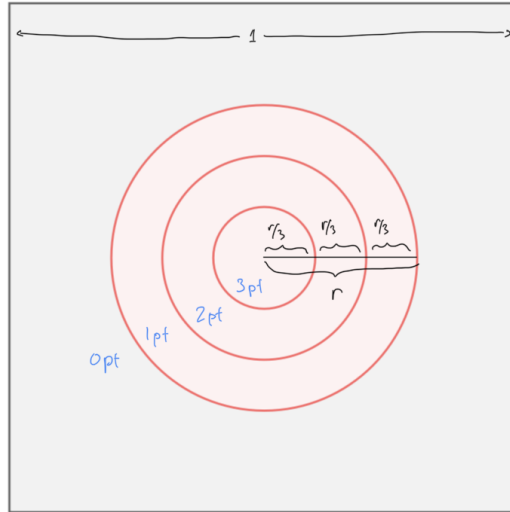
Instructions: write your name and section above. Then read each problem carefully and provide answers in the spaces below each prompt, using the reverse side of the page if necessary. Your answers should display your reasoning clearly. If you use results from class in your answer, you can simply write, “by a theorem from class...”, and need not identify the result by name. You do not need to justify small steps in this way. You also do not need to calculate decimals or perform any burdensome calculations, but you should simplify your answers as much as possible. Please feel free to ask any clarifying questions about the problems. You have 50 minutes to complete the exam. Good luck!

1. Consider a dart board comprising concentric circles with radii $r, \frac{2}{3}r, \frac{1}{3}r$, as depicted below. Suppose this board is mounted on a unit square, and that for an inexperienced player, the probability of hitting each region is simply the area of that region. Assume that a throw will land within the unit square. The scoring rules are: 0 points are awarded if the board is missed, 1 point is awarded for hitting the outermost circle, 2 points are awarded for hitting the middle circle, and 3 points are awarded for hitting the inner circle. Find an expression for the probability of scoring i points on a single throw and show (or argue) that it is a valid probability measure.

score out of 10

2pt for associating areas to probs
1pt for writing an expression in terms of i
2pt for correct probabilities

2pt for approach to check axioms
1pt per axiom verified



$$P(0 \text{ pt}) = 1 - \pi r^2 \quad p_0$$

$$P(1 \text{ pt}) = \pi r^2 - \pi \left(\frac{2}{3}r\right)^2 \quad p_1$$

$$P(2 \text{ pt}) = \pi \left(\frac{2}{3}r\right)^2 - \pi \left(\frac{1}{3}r\right)^2 \quad p_2$$

$$P(3 \text{ pt}) = \pi \left(\frac{1}{3}r\right)^2 \quad p_3$$

$$P(i \text{ pts}) = \begin{cases} 1 - \pi r^2, & i=0 \\ \left[\left(\frac{(3-i+1)r}{3} \right)^2 - \left(\frac{(3-i)r}{3} \right)^2 \right], & i=1,2,3 \end{cases}$$

since the sample space is finite & each outcome has probability p_i where $0 \leq p_i \leq 1$ and $\sum_i p_i = 1$, this is a valid probability measure

2. Imagine you work a summer job at a theme park. At the entrance there are 5 ticket booths, and on a particular morning, there are 200 people waiting at the entrance. Everyone waiting must choose one line and buy a ticket individually. You're working at the first ticket booth, and it takes you exactly 2 minutes to process each transaction. After you finish selling the morning tickets, you'll take a break. Assume that customers select their lines at random and no new customers arrive. Find an expression for the probability that you'll be on break in an hour or less. (*Hint:* start by finding the probability that you'll be on break in $2j$ minutes.)

outcomes are allocations of customers to lines, so, e.g.,

$(L_1, L_1, L_5, L_2, \dots, L_4)$ drawn w/ rep from $\{L_1, L_2, L_3, L_4, L_5\}$
 $\uparrow \quad \uparrow \quad \uparrow \quad \dots \quad \uparrow$
 person person person ... person
 $1 \quad 2 \quad 3 \quad \dots \quad 200$

since identities of customers are not relevant, combinations are unordered.

$$\text{so: } |S| = \binom{5 + 200 - 1}{200} = \binom{204}{200}$$

$$\text{let } E_j = \{\text{on break in } 2j \text{ minutes}\} \\ = \{j \text{ customers in line 1}\}$$

form outcomes in E_j by allocating j customers $\rightarrow L_1$
 $200-j$ customers $\rightarrow L_2, \dots, L_5$

$$\text{so } |E_j| = \binom{4 + (200-j) - 1}{200-j} \\ = \binom{203-j}{200-j}$$

then $E = \{\text{on break in an hour}\} = \bigcup_{j=1}^{30} E_j$ and since E_j are disjoint,

$$P(E) = P\left(\bigcup_{j=1}^{30} E_j\right) = \sum_{j=1}^{30} P(E_j) = \sum_{j=1}^{30} \frac{\binom{203-j}{200-j}}{\binom{204}{200}}$$

score out of 10

2pt for correct set-up: allocating lines to customers

2pt for counting the sample space correctly

2pt for counting the ways to put 30 (or j) customers in one line

2pt for a probability consistent with prior calculations

2pt for writing event of interest as a disjoint union and expressing probability as a sum

3. Suppose you've developed a machine learning model to classify activity patterns as "at work" or "not at work" based on user cell phone data. In testing prior to deployment, you assess performance on timepoints sampled from 1000 users whose information was not used in the development of the model; of those, 800 users were working during the sampled timepoints. Denote the event that a user is at work by W and the event that a user is predicted as at work by V . Suppose you observe that for a randomly selected user from this test group:

$$P(V|W) = 0.9$$

$$P(V^C|W^C) = 0.3$$

- How many of your model predictions were correct?
- What is the probability that the prediction was correct for a randomly selected user in the test group?
- If your model predicts that a user in the test group is not at work, what is the probability that the prediction is correct?

$$\begin{aligned} \text{i. } P(V|W) = 0.9 &\Rightarrow 0.9 \cdot 800 = 720 \text{ correct} \\ P(V^C|W^C) = 0.3 &\Rightarrow 0.3 \cdot 200 = 60 \text{ correct} \end{aligned} \quad \left. \vphantom{\begin{aligned} P(V|W) = 0.9 \\ P(V^C|W^C) = 0.3 \end{aligned}} \right\} 780 \text{ correct}$$

$$\text{ii. } P(\text{correct}) = \frac{780}{1000} = 0.78$$

$$\begin{aligned} \text{OR: } P(\text{correct}) &= P(\text{correct}|W)P(W) + P(\text{correct}|W^C)P(W^C) \\ &= P(V|W)P(W) + P(V^C|W^C)P(W^C) \\ &= 0.9 \cdot 0.8 + 0.3 \cdot 0.2 \\ &= 0.78 \end{aligned}$$

$$\text{iii. } P(W^C|V^C) = \frac{P(V^C|W^C)P(W^C)}{P(V^C|W^C)P(W^C) + P(V^C|W)P(W)}$$

$$= \frac{(0.3)(0.2)}{(0.3)(0.2) + (0.1)(0.8)}$$

score out of 10

$$= \frac{0.06}{0.06 + 0.08}$$

2pt for number of correct predictions, with 1pt partial credit for a reasonable or partial calculation

$$= \frac{6}{14}$$

$$= \frac{3}{7}$$

2pt for probability in (ii), with full credit awarded if consistent with (i) but error in (i)

$$(\approx 0.43)$$

2pt for identifying $P(W^C|V^C)$ in (iii)
2pt for using bayes' rule
2pt for correct calculations