Midterm 2 - DSC 40A, Spring 2023

Instructions

- This is a 50-minute exam consisting of 6 questions worth a total of 32 points.
- You may use any number of handwritten note sheets, and no other resources.
- No calculators.
- Please write neatly and stay within the provided boxes.
- You may fill out the **front page only** until you are instructed to start.
- Leave all answers **unsimplified** in terms of permutations, combinations, factorials, exponents, etc.
- You **do not** need to show your work or provide justification, unless a problem specifically asks you to.

Statement of Academic Integrity

By submitting your exam, you are attesting to the following statement of academic integrity.

I will act with honesty and integrity during this exam.

Name:	Solutions
PID:	A12345678
Seat you are in:	

Lecture Section: ◯ A00 (10-10:50AM) ◯ B00 (11-11:50AM)

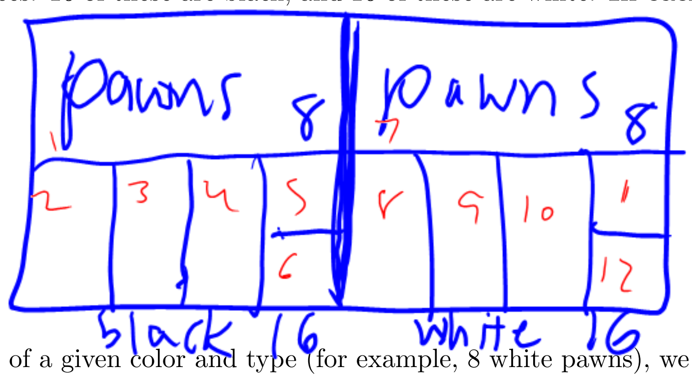
Version - A

♘ Chess Pieces ♕

A set of chess pieces has 32 pieces. 16 of these are black, and 16 of these are white. **In each color**, the 16 pieces are

types

- 8 pawns,
- 2 bishops,
- 2 knights,
- 2 rooks,
- 1 queen, and
- 1 king.



When there are multiple pieces of a given color and type (for example, 8 white pawns), we will assume they are **indistinguishable** from one another.

1. (3 points) Consider an experiment where each of n people selects one piece from their own set of 32 chess pieces, uniformly at random. The **result of** the experiment is a description of the **colors and types** of the pieces each person selected. For example, if $n = 3$, one possible result is:

- Person 1 selected a white knight.
- Person 2 selected a black queen.
- Person 3 selected a black pawn.

How many results are possible for this experiment with n people?

- ○ ~~2^n~~
- ○ ~~6^n~~
- ● 12^n
- ○ ~~16^n~~
- ○ ~~32^n~~
- ○ None of the above.

12 different (color, type) pairs

2 colors

6 types

each piece is equally likely

with replacement, can be duplicates

sequences

2. (5 points) Suppose you randomly select 2 pieces from a set of 32 chess pieces, **without replacement**.

a) (3 points) You glance at the pieces just long enough to see that both pieces are white. What is the probability that you have 2 pawns?

Solution: $\frac{8}{16} * \frac{7}{15} = \frac{C(8,2)}{C(16,2)} = \frac{\frac{8}{32} * \frac{7}{31}}{\frac{16}{32} * \frac{15}{31}} = \frac{7}{30}$ P(2 pawns | 2 white)

b) (2 points) True or False: Having two pawns is **independent** of having two white pieces.

◯ True
● False

if ind., P(2 pawns | 2 white) = P(2 pawns)

$P(2 \text{ pawns} | 2 \text{ white}) = \frac{8}{16} \cdot \frac{7}{15}$

direct interpretation

what is this?

$\neq \frac{16}{32} \cdot \frac{15}{31}$

formula for cond. prob.

$$P(2 \text{ pawns} | 2 \text{ white}) = \frac{P(2 \text{ pawns} \overset{\text{and}}{\cap} 2 \text{ white})}{P(2 \text{ white})}$$

$$= \frac{P(2 \text{ white pawns})}{P(2 \text{ white})} = \frac{\frac{8}{32} \cdot \frac{7}{31}}{\frac{16}{32} \cdot \frac{15}{31}}$$

using mult rule

P (1st white and 2nd white)

rook pawn queen pawn

1 2 3 4 ... 14 15 16

3. (9 points) In this problem, a **lineup** is a way of arranging items in a straight line.

a) (3 points) A chess player lines up all **16 white pieces** from the set of chess pieces. How many different-looking lineups can be created? Remember, some pieces look the same.

similar to rearranging letters MISSISSIPPI

Solution: $\boxed{\frac{16!}{8!2!2!2!}} = \frac{P(16,8)}{2^3} = C(16,8)*C(8,2)*C(6,2)*C(4,2)*C(2,1)*C(1,1)$

many equivalent expressions

b) (3 points) A chess player lines up all **16 pawns** from the set of chess pieces. How many lineups have white pawns on both ends?

8 black, 8 white

$C(16,8)$ H/T

Solution: $C(14,6) = C(14,8) = \frac{14!}{8!6!}$

c) (3 points) A chess player lines up all **16 pawns** from the set of chess pieces. Assuming that each different-looking lineup is equally likely, what is the probability that the lineup has two of the same-colored pawns on both ends (both black or both white)?

- ○ $\frac{1}{4}$
- ○ $\frac{1}{2}$ close
- ○ $\frac{7}{30}$
- ● $\boxed{\frac{7}{15}}$
- ○ None of the above.

$$P(\text{both ends same}) = P(\text{both ends white}) + P(\text{both ends black})$$

$$\frac{C(14,6)}{C(16,8)} + \frac{C(14,6)}{C(16,8)} = 2 * \frac{C(14,6)}{C(16,8)}$$

W ___ ___ ___ ___ ___ ___ W

middle 14 spots

P(last matches first)

b ___ ___ ___ ___ ___ ___ ___

1st 16th

$$P(\text{black first}) * \underbrace{P(\text{black last} \mid \text{black first})}_{7/15} + P(\text{white first}) * P(\text{white last} \mid \text{white first})$$

$$\frac{1}{2} \qquad\qquad \frac{1}{2} \cdot \frac{7}{15}$$

$$\cancel{2} \cdot \frac{1}{\cancel{2}} \cdot \frac{7}{15} = \frac{7}{15}$$

90% Avi wins against beginner

4. (6 points) Suppose that there are three possible experience levels in chess (beginner, intermediate, advanced). Only **10%** of beginner players beat Avi at a game of chess, while **50%** of intermediate players and **80%** of advanced players do.

Avi signs up to participate in a certain chess tournament called the Avocado Cup. Aside from Avi, **50%** of the players in the tournament are beginners, **40%** are intermediate, and **10%** are advanced.

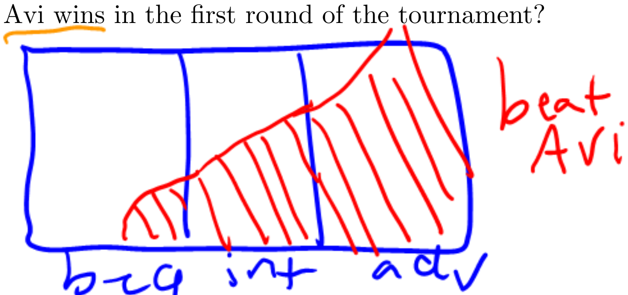
The tournament proceeds in rounds. In the first round of the tournament, players are **randomly paired up** for a game of chess.

Avi equally likely to have any opponent

a) (3 points) What is the probability that Avi wins in the first round of the tournament?

- ○ 33%
- ○ 50%
- ● 67%
- ○ 83%
- ○ None of the above.

$1 - P(\text{beat Avi})$



b) (3 points) It turns out that, sadly, Avi loses to his opponent in the first round. What is the probability that Avi's opponent is an advanced player? Choose the **closest answer** among the choices listed.

- ○ 15%
- ● 25%
- ○ 35%
- ○ 45%

Bayes: $P(\text{adv} \mid \text{beat Avi}) = \dfrac{P(\text{adv}) \cdot P(\text{beat Avi opponent} \mid \text{adv})}{P(\text{beat Avi})}$

$P(\text{beg}) = 0.5 \qquad P(\text{int}) = 0.4 \qquad P(\text{adv}) = 0.1$

$P(\text{beat Avi} \mid \text{beg}) = 0.1 \quad + \quad P(\text{beat Avi} \mid \text{int}) = 0.5 \quad + \quad P(\text{beat Avi} \mid \text{adv}) = 0.8$

$P(\text{beat Avi}) = P(\text{beat Avi} \cap \text{beg}) + P(\text{beat Avi} \cap \text{int}) + P(\text{beat Avi} \cap \text{adv})$

$= P(\text{beg}) \cdot P(\text{beat Avi} \mid \text{beg}) + \ldots$

$P(\text{beat Avi}) = 0.05 + 0.20 + 0.08 = \boxed{0.33}$

b) $\dfrac{0.08}{0.33}$

5. (3 points) You have a large historical dataset of all competitors in past years of the Avocado Cup chess tournament. Each year, hundreds of chess players compete in the tournament, and one person is crowned the winner. For each competitor in each year of the competition's history, you have information on their

- experience level (beginner, intermediate, advanced),
- birth month (January through December), and
- whether they won the tournament that year (yes or no).

Naive Bayes

Assume that birthdays of competitors are evenly distributed throughout the months.

You want to predict who will win this year's Avocado Cup. To do so, you use this historical data to train a Naive Bayes classifier and classify each competitor as a winner or non-winner, given their experience level and birth month. Which of the following reasons best explains **why your classifier is ineffective** in identifying the winner?

- ○ Because it uses a variable (birth month) that likely has nothing to do with a person's chances of winning the tournament.
- ○ Because it uses a variable (experience level) that likely has a strong connection with a person's chances of winning the tournament.
- ● Because it uses a dataset where there are many more non-winners than winners.
- ○ Because it uses a categorical response variable.

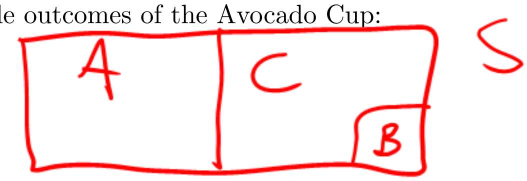
$$P(\text{win} \mid \text{exp.} + \text{birth month}) \propto P(\text{win}) * P(\text{exp} \mid \text{win}) * P(\text{birth month} \mid \text{win}) \approx \frac{1}{12}$$

$$P(\text{don't win} \mid \text{exp} + \text{birth month}) \propto P(\text{don't win}) * P(\text{exp} \mid \text{don't win}) * \frac{1}{12} \approx P(\text{birth month} \mid \text{don't win})$$

6. (6 points) The Avocado Cup is organized into rounds. In each round, players who win advance to the next round, and players who lose are eliminated. Rounds continue on like this until there is a single tournament winner.

Define the following events in the sample space of possible outcomes of the Avocado Cup:

- A = Avi loses in the first round.
- B = Avi wins the tournament.
- C = Avi wins in the first round.



a) (3 points) Which of the following statements is true? **Select all that apply.**

- ☐ ~~A and B are independent.~~
- ■ A and B are conditionally independent given C.
- ☐ ~~A, B, and C form a partition of the sample space.~~
- ☐ None of the above.

mutually exclusive

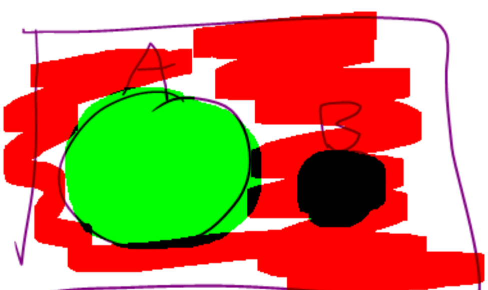
$B \cap C \neq \emptyset$

b) (3 points) The events A and B are mutually exclusive, or disjoint. More generally, for **any** two disjoint events A and B, show how to express $P(\overline{A}|(A \cup B))$ in terms of $P(A)$ and $P(B)$ **only**. For this problem only, **show your work and justify each step.**

Solution:

$$\underbrace{P((A \cap B)|C)}_{0} = \underbrace{P(A|C)}_{0} \cdot \underbrace{P(B|C)}_{\text{nonzero}}$$

$$P(\overline{A}|(A \cup B)) = \frac{P(\overline{A} \cap (A \cup B))}{P(A \cup B)} = \frac{P(B)}{P(A) + P(B)}$$



other solutions: — Bayes Thm
— complement