

5 pages. 6 problems. 100 points. No calculators. Show all work.

Problem 1 (5 points each). Suppose you have the following sets:

$$A = \{\text{green}, 3, \text{hat}, \text{green}\}$$

$$B = \{3, \text{green}, \text{hat}\}$$

$$C = \{A, B\}$$

(a) How many elements does A have? List them without repetition.

3: 3, hat, green

(b) How many elements does C have? List them without repetition.

2: A, B

(c) Is A a subset of B ? Why or why not?

Yes, all of its elements are in B .

(d) Which (if any) of the three sets have the empty set as a subset?

All of them. \emptyset is ~~the~~ a subset of any set.

Problem 2 (5 points each). You have a standard deck of cards.

(a) You shuffle the deck. What is the probability that the third card from the top is a queen?

$$\frac{4}{52}, \text{ or } \frac{1}{13}$$

(b) You and your friend each randomly take a card and compare them. What is the probability that you have a higher card than your friend?

$$\frac{48}{51} \cdot \frac{1}{2} = \frac{24}{51}$$

↑
the chances
that we don't
have the same card

↑
the chances
that mine is higher,
if they are different.

Problem 3 (5 points each). Consider the following games:

A: You flip a coin. If it's heads, you get 3 points. If it's tails, you get 5 points.

B: You roll two dice. If they're the same, you lose 6 points. Otherwise, you gain 6 points.

(a) Compute the expected value for game A.

$$(50\%)(3) + (50\%)(5) = \frac{3+5}{2} = 4$$

(b) Compute the expected value for game B.

$$\frac{1}{6}(-6) + \frac{5}{6}(6) = -1 + 5 = 4$$

(c) Compute the variance each game.

$$A: \frac{1}{2}(3-4)^2 + \frac{1}{2}(5-4)^2 = \frac{1}{2} + \frac{1}{2} = 1$$

$$B: \frac{1}{6}(-6-4)^2 + \frac{5}{6}(6-4)^2 = \frac{100}{6} + \frac{5 \cdot 4}{6} = \frac{100+20}{6} = \frac{120}{6} = 20$$

~~$\frac{1}{6}(100+20) = \frac{120}{6} = 17.33...$~~

(d) If you're feeling lucky, which game should you play to get more points? Why?

B: it has a higher chance of producing really high winnings (A is more likely to be close to the 4 points/turn).

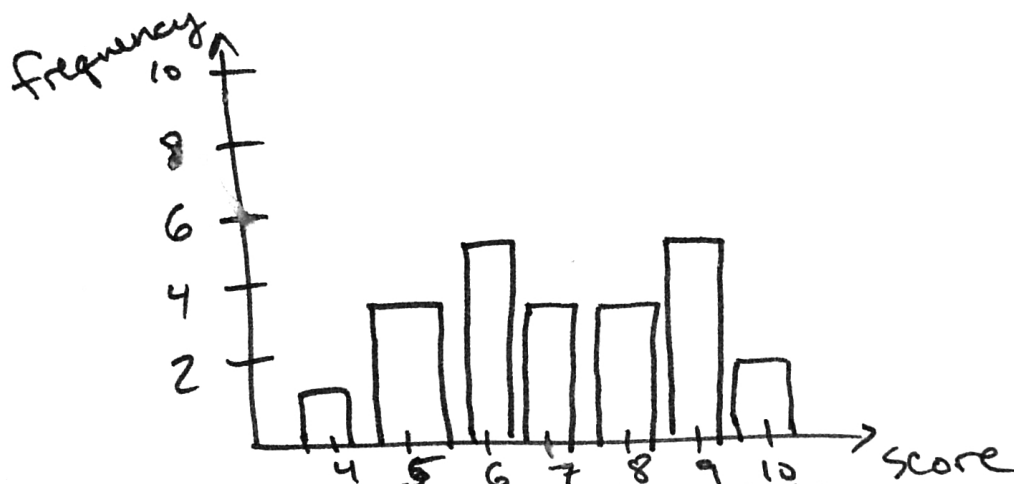
Problem 4 (5 points each). Your students get the following grades on a quiz (out of 10 points):

7, 6, 9, 5, 9, 10, 8, 9, 6, 8, 6, 5, 8, 9, 6, 7, 10, 5, 7, 9, 4, 6

(a) What are the maximum, minimum, and median scores?

max: 10
 min: 4
 median: 7

(b) Use a bar graph to display the scores.



(c) Use a pie chart to show how many students get a 9 or better.

$\frac{7}{22}$



The number of students who got a 9 or better is 7.

Problem 5 (20 points total: 5, 10, 5). Columbia University has 1,500 incoming students each year, and a 4-year graduation rate of 90%. Suppose you have developed a test that predicts whether a student will leave early. Suppose your test has a 1% false positive rate, and an 20% false negative rate (here, a "positive" means the student will **not** graduate).

(a) How many of the incoming students will graduate? How many won't?

$$90\% \times 1500 = 1350 \text{ will graduate.}$$

$$10\% \times 1500 = 150 \text{ will not graduate}$$

(b) If you test all the students, how many will you decide are not going to graduate?

From the 1350, 1% are incorrectly tested: 13-14 students.

From the 150, 80% are correctly tested: 120 students.

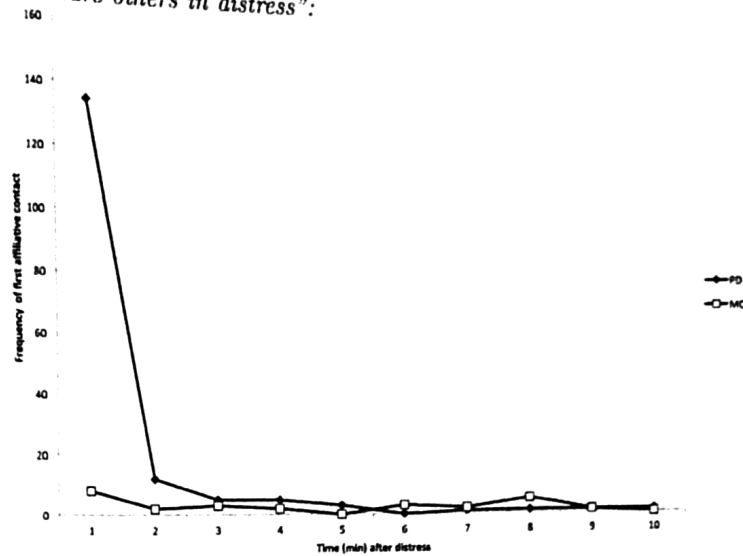
So ≈ 133.5 students are diagnosed as "not likely to graduate".

(c) If you decide a specific student will not graduate, what are the chances that the student actually won't graduate?

$$\frac{120}{133.5} \leftarrow \begin{array}{l} \text{actually not graduating.} \\ \text{total diagnosed} \end{array}$$

since we essentially picked a random person out of the group diagnosed as non-graduating.

Problem 6 (5 points each). The following picture is from a recent scientific publication "Asian elephants (*Elephas maximus*) reassure others in distress":



In the study, the scientists waited for one elephant to get upset and saw whether other elephants tried to comfort the unhappy elephant. The number of interactions is graphed as the "PD" (post-distress) line. For comparison, they also provide an "MC" (matched control) graph of observed elephant interactions when nothing bad was happening.

(a) Is this an experimental or observational study? Why do you think this choice was made?

Observational. Scaring elephants is not nice.

(b) What kind of graph is used to display the data? Is it appropriate for this experiment? Why?

Line graph. Yes, it's useful to see the reactions over time. A bar graph would have worked too.

(c) The image seems pretty convincing by itself, but the article is 17 pages long. What supporting information would you expect to see in the article?

Previous studies, information about the elephants in question, comparison to other animals. Data that show that this behavior counts as "reassuring", etc. More explanation of what "distress" and "matched control" mean.