

Name: _____

SID: _____

Section number and time: _____

-
- You are allowed a calculator that cannot communicate with other devices (no phones).
 - **Explain all your reasoning and show your work.**
 - The test begins on the next page. You may use this page for scratch work.
 - **WRITE YOUR NAME ON EACH PAGE**
-

Name: _____

1. Below is the distribution table for the average number of hours per day spent by students at a certain university on their laptops, doing things unrelated to academics. The right endpoint is included, but not the left. Assume we can continuously measure time.

Number of hours	0-2	2-4	4-8	8-12	12-20
Height (% per hour)	6	14	8		1.25

(a) Find the missing value (2 points)

(b) Find the median of the distribution. (2 points)

(c) True or False: The median for this data is greater than the average of this data. (2 points)

-
2. Indicate whether the following statements are true or false, and give a brief explanation: (2 points each)

(a) If in a large statistics course, the scores for the final followed the normal curve closely. The average was 70 points out of 100, and three-fourths of the class scored between 60 and 80 points, we can conclude that the sd of the scores was less than 10 points. **T** **F**

(b) If you add 10 to each entry on a list, that leaves the sd of the list unchanged. **T** **F**

(c) If the sd of a list is 0, it must be true that all the numbers on the list are 0. **T** **F**

(d) It is always possible to design a double-blind controlled experiment, provided the budget is sufficiently large. **T**
F

Name: _____

3. Suppose A and B are independent events such that $P(A) = 0.2$ and $P(B) = 0.6$.

(a) Is it possible for A and B to be mutually exclusive? (2 points)

(b) What is the probability of **neither** A **nor** B occurring? (2 points)

4. I roll a fair **ten**-sided die **three** times.

(a) What is the probability that all the rolls are different? (2 points)

(b) What is the probability that not all the rolls are the same? (2 points)

5. Two cards are dealt off the top of a well-shuffled standard deck (52 cards, 4 suits: hearts, clubs, spades, diamonds, 13 cards in each suit labeled 2-10, Jack, Queen, King, Ace). Compute probabilities of the following events:

(a) The second card is a queen. (2 points)

(b) The cards are a king and a queen. (2 points)

Name: _____

6. Can we model the following situations using binomial random variables? If yes, indicate what n and p should be. If not, indicate why. (2 points each)

(a) The number of aces in a 6-card hand dealt off a well-shuffled standard deck of cards.

(b) If a die is rolled three times and a coin is tossed three times, the sum of the number of times the coin lands heads and the die lands with an odd number of spots.

7. Let X be a discrete random variable such that X takes the values 1 and -1 with probabilities 0.4 each and the value 2 with probability 0.2. Find the expected value of X . (2 points)

8. Consider the following box of tickets:

1	1	1	-3
---	---	---	----

(a) What is the standard deviation of the tickets in the box? (2 points)

(b) If we draw 300 times from this box, what will be the standard error of the sum of these 300 draws? (2 points)

(c) What is the approximate probability that the sum of 300 draws will fall between -30 and 60? (2 points)

(d) How would you use `pnorm()` to compute this probability? (2 points)

Name: _____

9. Consider a binomial random variable X with parameters $n = 100$, and $p = 0.05$.

(a) In R, how would you create a vector that consisted of ninety-five 0's, and five 1's? (2 points)

(b) How would you use R to compute the probability that $X \leq 2$? (2 points)

10. Consider the following data frame that is a subset of the one that you have used in a homework assignment, giving total earnings and salaries of some top athletes, in millions of dollars.

Name	Total	Salary	Sport
Cristiano Ronaldo	108.0	61.0	Soccer
Neymar	90.0	73.0	Soccer
LeBron James	85.5	33.5	Basketball
Roger Federer	77.2	12.2	Tennis
Stephen Curry	76.9	34.9	Basketball
Kevin Durant	57.3	25.3	Basketball

(a) Assume that `dplyr` is loaded in R. Write the code that you would use to create a new column called "Endorsements" that gives the difference between the player's total earnings (called Total) and his salary. (2 points)

(b) Write code to compute the average salary, grouped by sport. (2 points)

(c) What would you expect the following code to return?
`arrange(filter(forbes18, sport=="Soccer"), total)` (2 points)

Name: _____
