

STAT 151 Midterm Exam: Section D01, Version 1

Name: **Solution**

Instructions: This is a closed book exam. You may use the formula sheet and the z -score table provided and a non-programmable calculator. The exam has eight pages and has multiple choice and short answer questions. In each multiple choice question, choose the answer you think is closest to being correct. There are no deductions for incorrect guesses. Mark your choices clearly in the answer section below by writing the letter corresponding to your chosen answer. Make sure your answers are correctly located and clearly marked. Each correct answer is worth 1 mark. In the short answer section, use the space provided and/or the back of other pages. Clearly mark your answers for visibility and legibility. Mark worth is denoted after each specific question.

Hand in ALL pages. Print your name at the top of this page and your student ID number at the top of the next page.

1. _____
 2. _____
 3. _____
 4. _____
 5. _____

11. _____
 12. _____
 13. _____
 14. _____
 15. _____

6. _____
 7. _____
 8. _____
 9. _____
 10. _____

16. _____
 17. _____
 18. _____
 19. _____
 20. _____

Student ID: _____

1. Which of the following statements is TRUE?

- A) This is a false statement.
- B) Standard deviations always come from samples.
- C) Proportions always come from populations.
- D) A median can be a parameter or a statistic.**

2. Let the probability of enjoying *Transformers 2: Revenge of the Fallen*, or *T*, be 0.25. Also, the probability of enjoying *The Dark Knight*, or *DK*, is 0.95. If the probability of enjoying both movies is 0.2375, one can conclude that

- A) *T* and *DK* are independent events.**
- B) $P(T | DK)$ is greater than 0.5.
- C) $P(DK \text{ or } T)$ is less than 0.5.
- D) Optimus Prime will run into bad boys on an island that might be Pearl Harbor.

3. While studying, a friend throws a number at you. After you pick it up, you notice its value of -0.15. If this were a correlation value, what could you DEFINITELY say?

- A) The value represents a strong, negative linear association between two variables.
- B) There is no strong relationship possible between the two variables.
- C) The value represents a weak, negative linear association between two variables.**
- D) There is no relationship possible between you and your metaphorically violent friend.

Use the following information for questions 4-5:

The World Cups of Soccer won by **SOME** of the nations of the world were:

Athlete	World Cup total
Brazil	5
England	1
France	1
Germany	3
Italy	4

4. The *interquartile range* for this data is

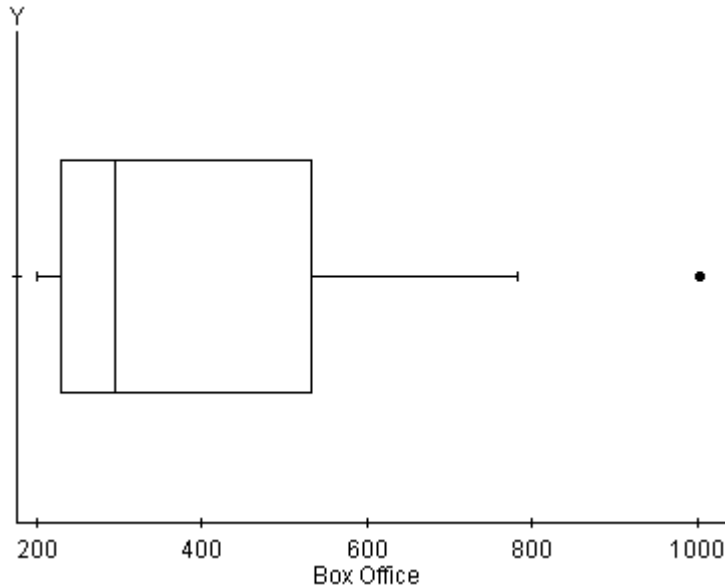
- A) 3.500** B) 3.200 C) 1.789 D) 3.000

5. By including the totals of Argentina and Uruguay (which each won two),

- A) the mean will remain the same.
- B) the median will increase.
- C) the range will remain the same.**
- D) the mode will decrease.

Use the following information for questions 6-7:

Boxplot of films making over \$US 200 million in 2008



6. In the above boxplot, the third quartile and the shape, respectively, are

- A) 525 and right-skewed
- B) 525 and left-skewed
- C) 790 and left-skewed
- D) 790 and right-skewed

7. From the above boxplot, which one of the following statements is CORRECT?

- A) Less than 50% of the films made between \$US 200 and 400 million.
- B) None of the values for the 5-number summary are exactly the same.**
- C) All values in the dataset are within the lower and upper inner fences.
- D) Jokers, archaeologists, and pandas make better money than teenage vampires.

8. Which of the following statements is TRUE?

- A) Random assignment is not possible with every kind of study.**
- B) An observational study never allows for population inferences.
- C) Causal inferences are allowed only when samples are randomly selected.
- D) Random sampling is only possible with experiments.

9. Bleeding Gums Murphy is a jazz musician/singer. His performances have been 3.5 hours long on average, with a standard deviation of thirty minutes. If the time of his performances follows a normal distribution, how many shows will be longer than 3.8 hours?

- A) 0.6000
- B) 0.2743**
- C) 0.4960
- D) 0.7257

10. Which one of the following statements is TRUE?

- A) Every dataset has a mode.
- B) 95% of observations lie within 3σ of the mean on a bell-shaped curve.
- C) The median and standard deviation are resistant to outliers.
- D) 50% of observations lie within the first and third quartiles.**

Use the following information for question 11:

The following 2-way table is a survey from students at a foreign university in the United Kingdom. They were asked if they knew what a fight club was and if they knew how to spell ‘seven’.

	Spell ‘se7en’	Spell ‘seven’	Punched surveyor
Knows of fight clubs	12	16	
Not know of fight clubs	40	130	
Punched Surveyor			2

11. If one person is selected at random, what is the approximate probability that they spell ‘seven’, given that they do not know of fight clubs?

- A) 0.235 **B) 0.765** C) 0.650 D) 0.308

12. If you roll a single die 6 times, what is the chance that not all rolls are odd numbers?

- A) $\left(\frac{1}{2}\right)^6$ B) $1 - \left(\frac{1}{6}\right)^6$ C) $\left(\frac{5}{6}\right)^1$ **D) $1 - \left(\frac{1}{2}\right)^6$**

Use the following information about movies from Pixar Animation Studios to answer questions 13-14:

x = the film’s rating on IMDB.com (out of 10, 10 being the highest)
 y = money made by the film in the United States (in millions of \$US),
 $n = 10$, $\bar{x} = 8.080$, $\bar{y} = 240.554$,
 $s_x = 0.444$, $s_y = 48.901$, $r = 0.407$

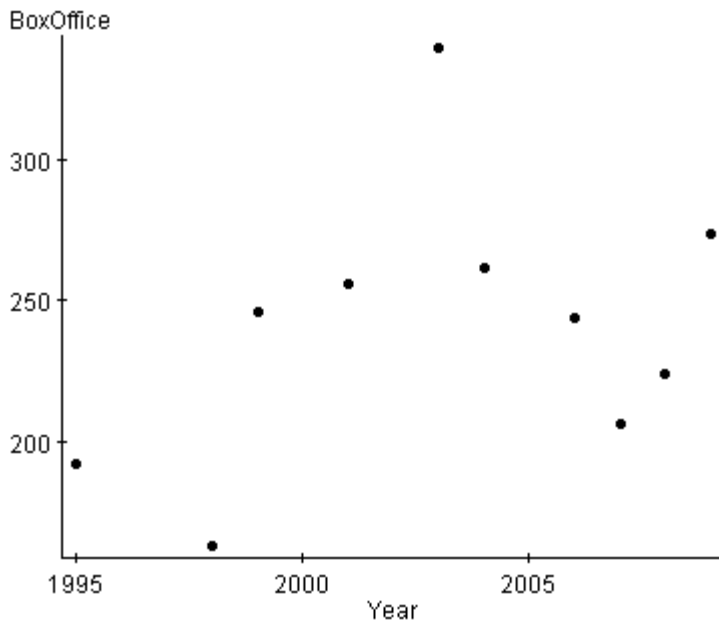
13. Respectively, the intercept and slope are approximately

- A) -121.639 and 44.826** B) -190.486 and 53.347
- C) 240.524 and 0.00370 D) 142.649 and 12.117

14. If Pixar Animation Studios creates a new film with a rating of 9.0 on IMDB.com, what amount of money does the regression line predict the film will make?

- A) 251.702 **B) 281.794** C) 289.637 D) 240.557

Use the following graph of movies from Pixar Animation Studios that relates the year the film came out with the money the film made in the U.S. for question 15:



15. Using the graph above, which statement is TRUE?

- A) It is appropriate to predict how much money *Toy Story 3* will make in 2010.
- B) The residual corresponding to $x = 2003$ will be large.**
- C) The year a film comes out causes an increase in the money the film will make.
- D) The coefficient of determination will be close to one.

16. Suppose the proportion of people who prefer the hairstyle Tom Hanks has in *Angels & Demons* over his hairstyle in *The Da Vinci Code* is 0.963. If a random sample consists of 125 people, what can be said about the sampling distribution of the sample proportion?

- A) It may not be normal with a standard error of 0.0356.
- B) It is approximately normal with a standard error of 0.0169.
- C) It is approximately normal with a standard error of 0.00029.
- D) It may not be normal with a standard error of 0.0169.**

17. Which of the following statements about probability distributions is TRUE?

- A) For a discrete distribution, $P(X < a) = 1 - P(X > a)$.
- B) For a continuous distribution, $P(X \leq a) = P(X > a)$.
- C) For a discrete distribution, $P(X = a) = 0$.
- D) For a continuous distribution, $P(X \leq a) = 1 - P(X > a)$.**

18. Suppose you toss five coins at once. Let Y be the overall number of tails from the five tosses. Then, the probability of having either all heads or all tails is _____ whereas the probability of finding no more than three heads can be shown by _____. Choose the correct pairing below to fill in the blanks.

- A) $\frac{1}{32}; P(Y > 2)$
- B) $\frac{1}{16}; P(Y \geq 2)$**
- C) $\frac{1}{16}; P(Y > 3)$
- D) $\frac{1}{32}; P(Y \geq 3)$

19. Regarding the general properties of the sampling distribution of \bar{x} , it is TRUE that

- A) as the sample size increases, the standard deviation of \bar{x} also increases.
- B) no matter the population distribution, the sampling distribution of \bar{x} is approximately normal if the sample size is greater than 20.
- C) the spread of the sampling distribution of \bar{x} is larger than the spread of the population distribution.
- D) the sampling distribution of \bar{x} is normal for any sample size, if the population distribution is normal.**

20. If you draw a single card from a full deck of 52 cards, which of the following probability statements is CORRECT?

- A) $P(\text{Face}) = P(\text{Ace})$
- B) $P(\text{Heart} \cup \text{Diamond}) = P(\text{Club} \cap \text{Spade})$
- C) $P(\text{Black} \cap \text{Red}) = 1$
- D) $P(\text{Black}) = P(\text{Heart} \cup \text{Spade})$**

Short Answer Questions

1. Say you have a fair coin and you associate obtaining a head with the value 0 whereas obtaining a tail has a value of 2. Let X be the sum of the values obtained from tossing the coin three times. Calculate the following probabilities. (4 points)

A) $P(X \leq 5)$ (2 points)

B) $P(0 < X \leq 4)$ (2 points)

Using the probability distribution below,

$$\begin{aligned} P(X \leq 5) &= P(X = 0) + P(X = 2) + P(X = 4) \\ &= 1/8 + 3/8 + 3/8 \\ &= 7/8 \end{aligned}$$

$$\begin{aligned} P(0 < X \leq 4) &= P(X = 2) + P(X = 4) \\ &= 3/8 + 3/8 \\ &= 6/8 = 3/4 \end{aligned}$$

x	$P(X = x)$
0	1/8
2	3/8
4	3/8
6	1/8

2. Suppose you have a continuous random variable X following a uniform distribution where the endpoints are 10 and 18. Calculate the following probabilities. (4 points)

A) $P(X \geq 15.5)$ (2 points)

$$\begin{aligned} &\text{width} \times \text{height} \\ (18 - 15.5) &\times (1/8) \\ (2.5) &\times (1/8) \\ &0.3125 \end{aligned}$$

B) $P(13 \leq X < 20)$ (2 points)

$\rightarrow P(13 \leq X < 18)$ since range ends at 18.

$$\begin{aligned} (18 - 13) &\times (1/8) \\ (5) &\times (1/8) \\ &0.625 \end{aligned}$$

3. Don sells airtime for commercials on TV. Suppose the lengths of the airtime he sells follow a normal distribution with a mean of 30 seconds and a standard deviation of 2.5 s.

A) If you randomly select a commercial with airtime sold by Don, what is the probability that the commercial is more than 27 seconds long? (3 points)

$$P(X > 27) \rightarrow P\left(\frac{X - \mu}{\sigma} > \frac{27 - 30}{2.5}\right) = P(Z > -1.20)$$

$$= 1 - P(Z < -1.20) = 1 - 0.1151 = 0.8849$$

B) What is the maximum time to be in Don's shortest 18% of commercials? (3 points)

The shortest 18% means looking for an area of 0.18 inside the z-score table. The corresponding z value is -0.92.

$$x = \mu + (z)(\sigma) = 30 + (-0.92)(2.5) = 27.7$$

4. Three friends are exceptionally good at playing poker. Jay and Bob each have a probability of 0.91 for winning a single game of poker while Brody wins a game of poker with a probability of 0.96. Since they *never play each other*, assume that each of them winning a game is independent of the others.

A) What is the probability that Jay wins a single game or Brody loses one? (3 points)

Let 'J' be "Jay wins", 'Bo' be "Bob wins", and 'Br' be "Brody wins". Then,

$$P(J \cup Br^c) = P(J) + P(Br^c) - P(J \cap Br^c) = 0.91 + (1 - 0.96) - (0.91)(1 - 0.96)$$

$$= 0.9136$$

B) What is the probability that Jay wins 2 games and Bob loses 2 games? (3 points)

$$P(J \cap J \cap Bo^c \cap Bo^c) = P(J) \times P(J) \times P(Bo^c) \times P(Bo^c) =$$

$$= (0.91) \times (0.91) \times (1 - 0.91) \times (1 - 0.91)$$

$$= 0.006708$$