STAT 151 Midterm Exam: Section W01, Version A

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. _____
6. _____
7. _____
8. _____
9. _____
10. _____
11. _____
12. _____
13. _____
14. _____
15. _____
16. _____
17. _____
18. _____
19. _____
20. _____

Student ID: _____

1. Which of the following statements is TRUE?

A) Not this one.
B) A group of ten Canadian provinces is a sample.
C) A group of ten American states is a sample.
D) A median can be a parameter, but not a statistic.

2. Let the probability of a person aging backwards, or *BB*, be 0.002. Also, the probability of a person robbing a casino, or *RR*, is 0.04. If the probability of a person aging backwards and robbing a casino is 0.0012, one can conclude that

A) *BB* and *RR* are mutually exclusive events.
B) $P(BB \mid RR)$ is greater than 0.05.
C) $P(RR \text{ or } BB)$ is less than 0.05.
D) this person is also a good fighter yet hard to understand.

3. You're at a party and you get a number. For a brief second, you HAVE to see it as 0.5557828. If this was a correlation value, you could DEFINITIVELY say that

A) $r^2 = 0.309$ and the figure shows a moderate, positive association between 2 variables.
B) $r^2 = 0.746$ and the figure shows a strong, positive association between 2 variables.
C) $r^2 = -0.309$ and the figure shows a moderate, negative association between 2 variables.
D) $r^2 = 0.309$ and the figure shows a strong, positive association between 2 variables.

Use the following information for questions 4-5:

The goals scored by ***SOME*** Edmonton Oilers by January 28 were:

Athlete	**Goal total**
Gagner, Sam	4
Hemsky, Ales	14
Horcoff, Sean	11
Moreau, Ethan	10
Penner, Dustin	11

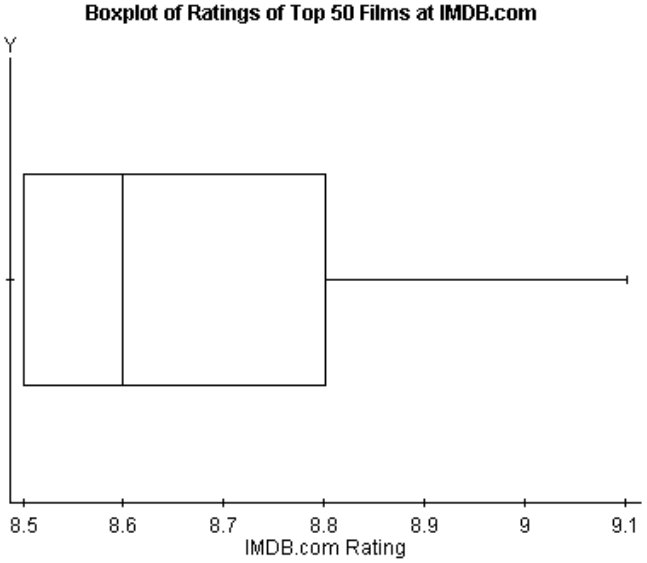
4. The *standard deviation* for this data is

A) 5.500 **B)** 3.674 C) 13.500 D) 1.500

5. By including the goal totals of 12 and 7 for defensemen Souray and Visnovsky,

A) the range will not change.
B) the median will decrease.
C) the mean will remain the same.
D) the mode will change.

Use the following information for questions 6-7:



6. In the above boxplot, the lower inner fence and the shape, respectively, are

A) 8.50 and right-skewed B) 8.05 and left-skewed
C) 8.50 and left-skewed **D)** 8.05 and right-skewed

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

A) More than 50% of the films have a rating less than 9.0.
B) The standard deviation is approximately 0.3.
C) A value of 9.2 is an outlier.
D) *The Usual Suspects* has a rating of 8.7, so you should have watched it by now.

8. Which of the following statements is TRUE?

A) Population inferences are possible when random assignment occurs.
B) An observational study allows for both kinds of inference.
C) Causal inferences are possible when random sampling occurs.
D) A particular experiment may have population inferences but not causal inferences.

9. Consuming chocolate is a well-known activity on certain days of the year (one can never tell which day). Unfortunately, large consumption of this food (say, a minimum of 1.1 kg) brings about the condition known as "choco-pox". If chocolate consumption for any individual follows a normal distribution with a mean of 0.65 kg and a standard deviation of 0.2 kg, what is the probability of a randomly selected individual consuming enough chocolate to get "choco-pox"?

A) 0.9878 B) 0.0838 C) 0.2578 **D)** 0.0122

10. Which one of the following statements is CORRECT regarding the Empirical Rule?

A) The Empirical Rule only works for standard normal distributions.
B) 95.44% of the data is contained within three standard deviations of the data.
C) The Empirical Rule has very little force in galaxies far, far away.
D) A z-score interval of (-1.00, 1.00) has an area of 68.26%.

Use the following information for questions 11-12:
The following 2-way table is a survey from a place with…um…people who know…uh…stuff and can write exams, maybe?

	Knows stuff	Does not know stuff
Can write exams	70	10
Can't write exams	7	3

11. If one person is selected at random, what is the approximate probability that they do not know stuff, given that they can't write exams?

A) 0.125 **B)** 0.300 C) 0.144 D) 0.033

12. If one person is selected at random, what is the approximate probability that they know stuff or can write exams?

A) 0.967 B) 0.922 C) 0.910 D) 0.778

Use the following info to answer questions 13-14:

x = the year a *Star Wars* film came out in theatres,
y = its rating on IMDB.com (out of 10, 10 being the highest)
$n = 6$, $\bar{x} = 1991.000$, $\bar{y} = 7.833$,
$s_x = 12.345$, $s_y = 1.021$, $b = -0.0654$

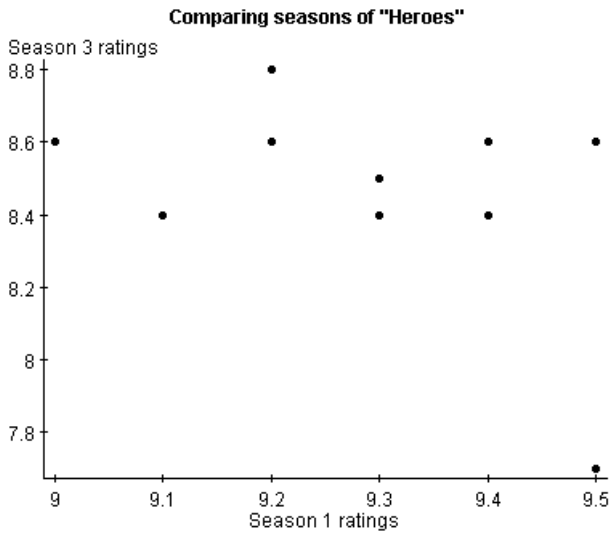
13. Respectively, the intercept and correlation value are approximately

A) 18.602 and -0.0054 **B)** 138.044 and -0.791
C) 138.044 and 0.791 D) 138.044 and -0.0654

14. *Star Wars: The Clone Wars* came out in 2008, made $US 35.2 million, and has an IMDB.com rating of 5.3. What is the residual for this new observation?

A) 6.721 **B)** -1.421 C) -2.441 D) 1.421

Use the following graph of viewer ratings for select episodes of the show *Heroes* from seasons 1 and 3 for question 15:



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

A) the corresponding correlation value will be negative.
B) the outlier in the bottom right-hand corner has no effect on the correlation.
C) predicting the rating in Season 3 for a rating of 8.8 in Season 1 is appropriate.
D) simple linear regression will prove, without a doubt, that the show has gotten worse.

Use the following information to answer questions 16-17:

In the movie *Slumdog Millionaire*, a contestant on the Indian version of "Who Wants to be a Millionaire?" attempts to answer 15 multiple choice questions correctly. Suppose the **proportion** of contestants able to answer the first ten questions correctly is 0.1990.

16. From a random sample of 100 people, what is the probability of no more than 20 people answering the first ten questions correctly?

A) 0.5120 B) 0.0250 C) 0.2000 D) 0.4880

17. If a random sample consists of 24 people, what can be said about the sampling distribution of the sample proportion?

A) It may not be normal with a standard error of 0.00664.
B) It is approximately normal with a standard error of 0.0815.
C) It is approximately normal with a standard error of 0.0962.
D) It may not be normal with a standard error of 0.0815.

18. Which of the following probability statements is CORRECT?

A) $P(Z < 0.55) = 0.2912$
B) $P(Z > -0.55) = 0.7088$
C) $P(Z < 0.55) = 0.5500$
D) $P(Z < -0.55) = 0.3085$

19. From a deck of 52 playing cards, suppose you draw a single card four times, each time without replacement. Let Y be the number of kings out of the four cards. Then, Y is a _____ random variable and the probability of finding no more than two kings can be shown by _____. Choose the correct pairing below to fill in the blanks.

A) continuous; $P(Y \leq 2)$
B) discrete; $P(Y \leq 2)$
C) discrete; $P(Y < 2)$
D) discrete; $P(Y > 2)$

20. During reading week, a student has the opportunity to relax after great exams like THIS one (R), go on vacation (V), and study for exams AFTER Reading Week ($Yuck$). Let's assume all 3 events are independent. Which of the following probability statements is DEFINITELY true?

A) $P(R \cup V \cup Yuck) = P(R) \times P(V) \times P(Yuck)$
B) $P(V \cup Yuck) = P(V) + P(Yuck) - P(R \cap Yuck)$
C) $P(R \cap V \cap Yuck) = P(R) + P(V) + P(Yuck)$
D) $P(V \cup Yuck) = P(V) + P(Yuck) - P(V \cap Yuck)$

Short Answer Questions

1. Suppose you have a continuous random variable X following a uniform distribution where the endpoints are -3 and 2. Calculate the following probabilities. (6 points)

A) $P(X < 1.5)$ (2 points)

width	×	height
(1.5 – (-3))	×	(1/5)
(4.5)	×	(1/5)
0.9		

Version B: $P(X < 1.5) \rightarrow P(X < 1) = 1$

B) $P(X \geq -0.25)$ (2 points)

width	×	height
(2 – (-0.25))	×	(1/5)
(2.25)	×	(1/5)
0.45		

Version B: $P(X \geq -0.25) = 0.25$

C) $P(-1.5 < X \leq 3)$ (2 points)

→ $P(-1.5 < X \leq 2)$ since range ends at 2.

width	×	height
(2 – (-1.5))	×	(1/5)
(3.5)	×	(1/5)
0.7		

For all questions, the height of any rectangle will always be 1/5.

Version B: $P(-1.5 < X \leq 2) = P(-1.5 < X \leq 1) = 0.5$

2. Say you have two fair coins and you associate obtaining a head with the value 1 whereas obtaining a tail has a value of -1. Let X be the sum of the values obtained from tossing each of the two coins once. Calculate the following probabilities. (4 points)

A) $P(X \geq -1)$ (2 points)

$P(X \geq -1) = P(X = 0) + P(X = 2) = ½ + ¼ = ¾$

x	$P(X = x)$
-2	¼
0	½
2	¼

Version B: $P(X \geq -1) = P(X = -1) + P(X = 2) = ½ + ¼ = ¾$

B) $P(-1.5 \leq X < 2)$ (2 points)

$P(-1.5 \leq X < 2) = P(X = 0) = ½$

Version B: $P(-1.5 \leq X < 2) = P(X = -1) = ½$

Version B:

x	$P(X = x)$
-4	¼
-1	½
2	¼

3. Jimmy is a baker. He bakes gingerbread people. Suppose the heights of his gingerbread people follow a normal distribution with a mean of 10 cm and a standard deviation of 2.5 cm.

A) If you randomly select one of Jimmy's gingerbread people, what is the probability that they are between 6.5 cm and 13.5 cm tall? (6 points)

$P(6.5 \le X \le 13.5) \rightarrow P\left(\frac{6.5-10}{2.5} \le \frac{X-\mu}{\sigma} \le \frac{13.5-10}{2.5}\right) = P(-1.40 \le Z \le 1.40)$
$= P(Z \le 1.40) - P(Z \le -1.40) = 0.9192 - 0.0808 = 0.8384$

Version B: $P(6.25 \le X \le 13.75) \rightarrow P(-1.50 \le Z \le 1.50) = 0.8664$

B) What's the minimum height to be among Jimmy's tallest 15% of gingerbread people? (4 points)

The tallest 15% means looking for an area of 0.85 inside the z-score table. The corresponding z value is 1.04.
$x = \mu + (z)(\sigma) = 10 + (1.04)(2.5) = 12.6$ *Version B*: $z = 0.84 \rightarrow x = 12.1$

4. Dr. Gregory House has four patients under his care. Two of them (H_1 and H_2) each have a probability of survival of 0.15 while the other two (H_3 and H_4) each have a probability of survival of 0.10. Assume that each of them surviving is independent of the others.

A) What is the probability of H_1 and H_3 surviving while H_2 and H_4 do not? (3 points)

Let 'H_1' be "H_1 surviving", 'H_2' be "H_2 surviving", and so on for H_3 and H_4. Then,

$P(H_1 \cap H_2^C \cap H_3 \cap H_4^C) = P(H_1)P(H_2^C)P(H_3)P(H_4^C) = (0.15)(1 - 0.15)(0.10)(1 - 0.10)$
$= (0.15)(0.85)(0.10)(0.90) = 0.011475$

Version B: $P(H_1 \cap H_2 \cap H_3^C \cap H_4^C) = (0.20)(0.20)(0.85)(0.85) = 0.0289$

B) What is the probability that, out of H_1, H_3, and H_4, only ONE survives? (7 points)

Using the same notation as in A),

$P(\text{only one survives}) = P(H_1 \cap H_3^C \cap H_4^C) + P(H_1^C \cap H_3 \cap H_4^C) + P(H_1^C \cap H_3^C \cap H_4)$
$= P(H_1)P(H_3^C)P(H_4^C) + P(H_1^C)P(H_3)P(H_4^C) + P(H_1^C)P(H_3^C)P(H_4)$
$= (0.15)(0.90)(0.90) + (0.85)(0.10)(0.90) + (0.85)(0.90)(0.10)$
$= 0.1215 + 0.0765 + 0.0765 = 0.2745$

Version B: $P(\text{only one survives}) =$
$(0.20)(0.80)(0.85) + (0.80)(0.20)(0.85) + (0.80)(0.80)(0.15) = 0.136 + 0.136 + 0.096 = 0.368$