Midterm 1 - STAT 301
Spring 2015
Type A

Name:

UIN:

Signature:

1. Do not open this test until told to do so.
2. Turn in your exam with your answers circled when you are done with the exam. You should not take the exam with you.
3. This is a closed book examination. You can only bring the normal z-tables to this exam. A cheat sheet will be provided.
4. You have 60 minutes to work on this exam. There are 15 multiple choice questions. If you cannot do one question move on to the next.
5. You may use a calculator in the exam (but not a phone).
6. No partial credit is given.
7. If you are unsure of what a question is asking for, do not hesitate to ask the instructor or course assistant for clarification.
8. I use sd to denote standard deviation.
9. Good Luck!!!

(1) Adults at an airport were randomly selected and asked four questions. What type of variable do we associate to each answer:

1	2	3	4
No. of children	Preferred destination	Flight number	Your Weight.

Answer:

	1	2	3	4
A	Numerical continuous	Categorical	Numerical discrete	Numerical continuous
B	Numerical continuous	Categorical	Numerical discrete	Numerical discrete
C	Numerical discrete	Categorical	Categorical	Numerical continuous
D	Numerical discrete	Numerical discrete	Numerical discrete	Numerical continuous
E	Numerical discrete	Categorical	Categorical	Binary

(2) You are given the data set: $-10, -9, 0, 3, 4, 5, 6$. The lowest value in this data set, -10, is changed to a lower (smaller) number (a number less than -10). How do the summary statistics change?

(A) The IQR stays the same, the mean will get lower (smaller), but the standard deviation is likely to get bigger.

(B) The median stays the same, but IQR, mean and standard deviation will get larger.

(C) The IQR stays the same, but the mean and standard deviation are likely to get lower.

(D) The median stays the same, but the 1st quartile will get lower/smaller.

(E) The median stays the same, but the 3rd quartile will get lower/smaller.

(5) The mean age of undergraduates at A&M is 20 years old with standard deviation one year. These students will be followed over 30 years. In 30 years time, what will be the mean age and standard deviation of this same set of students (converted to months)?

(A) mean $= 20 + 30 = 50$, sd $= 12 \times 1$.

(B) mean $= 20 + 30 = 50$, sd $= 1 + 1 \times 30 = 31$.

(C) mean $= 12 \times 20 + 12 \times 30 = 600$, sd $= 12 \times 1 + 12 \times 30 = 372$

(D) mean $= 12 \times 20 + 12 \times 30 = 600$, sd $= 1 + 1 \times 30 = 31$

(E) mean $= 12 \times 20 + 12 \times 30 = 600$, sd $= 12 \times 1 = 12$.

(4-5) We consider the following four data sets: (i) Distribution of numbers in a lottery (there are 49 balls), (ii) Weight of new born calves (in pounds) (iii) Salary of Walmart staff (in dollars) (iv) Size of rats (in inches).

The plots are given below.

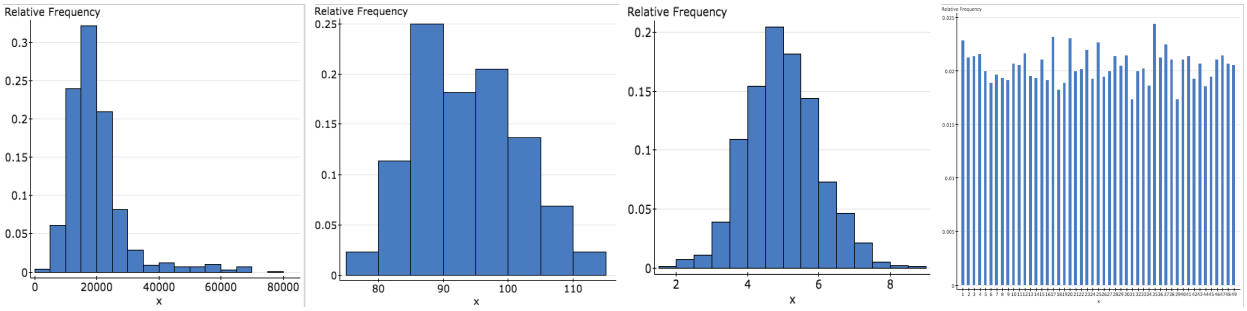


Figure 1: Plots labelled (1), (2), (3) and (4) (left to right).

(4) Match the plot to the data.

	1	2	3	4
A	Salary	Calves	Rats	Lottery
B	Calves	Lottery	Salary	Rats
C	Rats	Lottery	Calves	Salary
D	Salary	Rats	Calves	Lottery
E	Lottery	Rats	Calves	Salary

(5) Match the mean and standard deviation to each data set.

	Lottery	Calves	Salary	Rats
A	mean = 49, sd =40	mean = 110, sd = 50	mean = 10K, sd = 10K	mean = 8, sd = 6
B	mean = 10K, sd = 10K	mean = 110, sd = 50	mean = 49, sd 40	mean = 8, sd = 6
C	mean = 20K, sd = 10K	mean = 5, sd = 1	mean = 25, sd = 14	mean = 93, sd =8
D	mean = 25, sd = 14	mean = 93, sd =8	mean = 20K, sd = 10K	mean = 5, sd = 1
E	mean = 10K, sd = 10K	mean = 8, sd = 6	mean = 49, sd 40	mean = 110, sd = 50

(6) OkCupid, an online dating website, has spend a decade collecting and analyzing data from people who use the website. The team randomly selected 5000 female users and compares the average attractiveness scores they received (based on people who viewed them) with the number of messages they were sent in a month. This is an example of

(A) Experimental Study. (B) Observational Study.

(C) Selective sampling. (D) Stratified sampling. (E) [A] and [D].

(7) Suppose the sample standard deviation of a data set is zero, what does this tell you about the data?

(A) The data must only take zero values ie. 0, 0, 0, ..., 0.

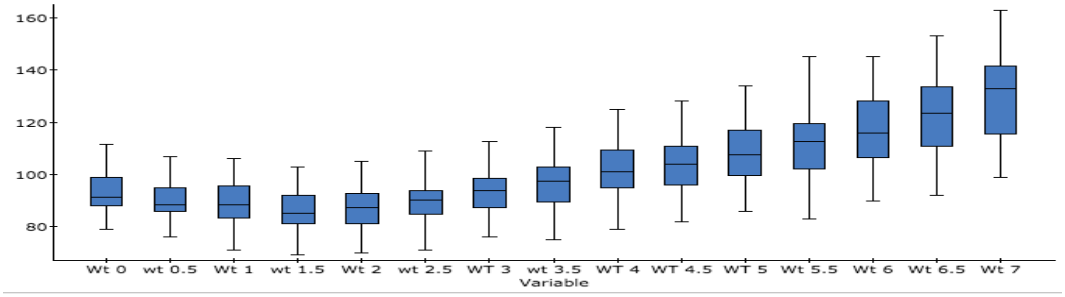
(B) The sample mean and the sample standard deviation have to be the same.

(C) The data is biased.

(D) Without seeing the data we cannot say anything about it.

(E) The data and the sample mean are the same.

(8) The weight of 45 calves are tracked from birth to 7 weeks (weights are taken every half a week). The data is plotted below. What does this plot suggest?



(A) The weights appear to be normally distributed.

(B) There appears to be small drop in the weights early on (week 0-1.5), then the weights starts to increase.

(C) There seems to an increase in the spread of weights over time.

(D) [A] and [B] (E) [B] and [C].

(9) The length of parrots beaks is known to be normally distributed with mean $\mu = 5$inches and standard deviation 2inches. What is the chance that beak of a randomly selected parrot will be between $4 - 8$ inches?

(A) 0.5 (B) 0.62 (C) 2 (D) 0.93 (E) 0.977

(10) The Daily Eagle recently had the headline '20% of the people who died last year in the Bryan/College Station area have smoked marijuana (cannabis) in the past 10 years'. What conclusions can be drawn from this headline?

(A) Marijuana is bad for you, but only if smoked. (B) Marijuana is bad for you.

(C) The data suggests that marijuana is bad for you when smoked.

(D) Nothing. (E) Marijuana is good for you.

(11) Female heights are known to be normally distributed with mean 64.5 inches and standard deviation 2.5 inches. Male heights are known to be normally distributed with mean 70 inches and standard deviation 4 inches. Peter is 72 inches tall.

Using equivalent percentiles, calculate how tall Peter would be if he were female.

(A) 65.75 inches (B) 72 inches (C) 65 inches (D) 65.19 inches (E) 64 inches.

(12) Educationists claim that students who do well in one subject do well in other subjects too. To check these claims 1000 students who attained either A, B or C, in Mathematics and English were randomly sampled. The data is collected in the Table below. What proportion of those students who scored an A in English also scored an A in Mathematics?

		Maths			
		A	B	C	
English	A	200	100	100	400
	B	120	100	50	270
	C	100	180	50	330
		420	400	200	1000

(A) 50% (B) 20% (C) 47.6%

(D) 40% (C) 42%.

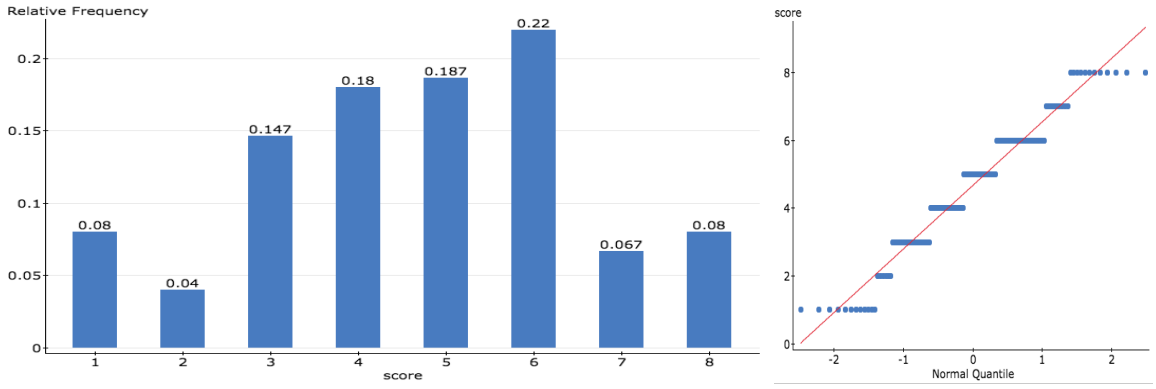


Figure 2:

(13-15)

(13) Students were asked to rate a course from 1-8 (only giving integer scores). The mean score is 4.68 and standard deviation is 1.88. Assuming the scores follow a normal distribution calculate the probability of a student rating the class **6 points or over**?

(A) 70.2% (B) 29.7% (C) 24.1% (D) 75.9% (E) 30.2%

(14) The scores for the survey is plotted in the relatively frequency histogram and QQplot in Figure 2. Use the information in Figure 2 to calculate the exact probability that a student rates the course **6 points or over**?

(A) 36.7% (B) 22% (C) 63.3% (D) 30.2% (E) 29.7%

(15) Explain why there is a discrepency between the two probabilities calculated in (13) and (14).

(A) The sample size in the calculation of (14) is not large enough. If a larger sample size was larger the data would look more normal.

(B) By the central limit theorem they would be close if a large enough sample size were collected.

(C) (13) uses an approximation of the true distribution, whereas (14) uses the true distribution.

(D) [A] and [B]

(E) [B] and [C].