Solution

STP 231 Statistics for Life Sciences

Your Name Print name here_____

Test #1 Summer 2015

Honor Statement:

By signing below you confirm that you have neither given nor received any unauthorized assistance on this exam. This includes any use of a graphing calculator beyond those uses specifically authorized by the Mathematics Department and your instructor. Furthermore, you agree not to discuss this exam with anyone until the exam testing period is over. In addition, your calculator's memory and menus may be checked at any time and cleared by any testing center proctor or Mathematics Department instructor.

_____ _____

Signature **Date**

Instructions:

- The exam is worth a total of 105 points; please make sure your exam has all pages (9) before you begin. Each multiple choice problem is worth 5 points each.
- Show all work in detail or your answer will not receive any credit. **If you think it, then write it.** Include appropriate units on all questions that apply. Write neatly and box all answers.
- Please ask your instructor you need scratch paper. **Do not use your own.**
- **No calculators or computers that do symbolic algebra, like the Casio FX-2, TI-89, or TI- 92, may be used.**
- The formulas are on the last page of the exam. You may take this page off to help you during the exam
- **Part 1 of the exam is for free response. Show your work or explain the process for each problem to receive credit.**
- Part 2 is the multiple choice part of the exam. A table is given on page four for you to write in the letter for the correct answer. You may take the pages apart to help you.
- **Any student who accesses a phone or any internet-capable device during an exam for any reason automatically receives a score of zero on the exam. All such devices must be turned off and put away and made inaccessible during the exam.**

Brain Dump

For Part I – Free Response. Show all work (as described on page 1).

1. Construct the following tables, graphs, and statistics as requested [20 points]
The weight in pounds of each of the 22 members of a freshman 4A football team are listed below:

i	1	2	3	4	5	6	7	8	9	10	11
y_i	144	152	142	151	160	152	131	164	141	153	140

i	12	13	14	15	16	17	18	19	20	21	22
y_i	144	175	156	147	133	172	159	135	159	148	129

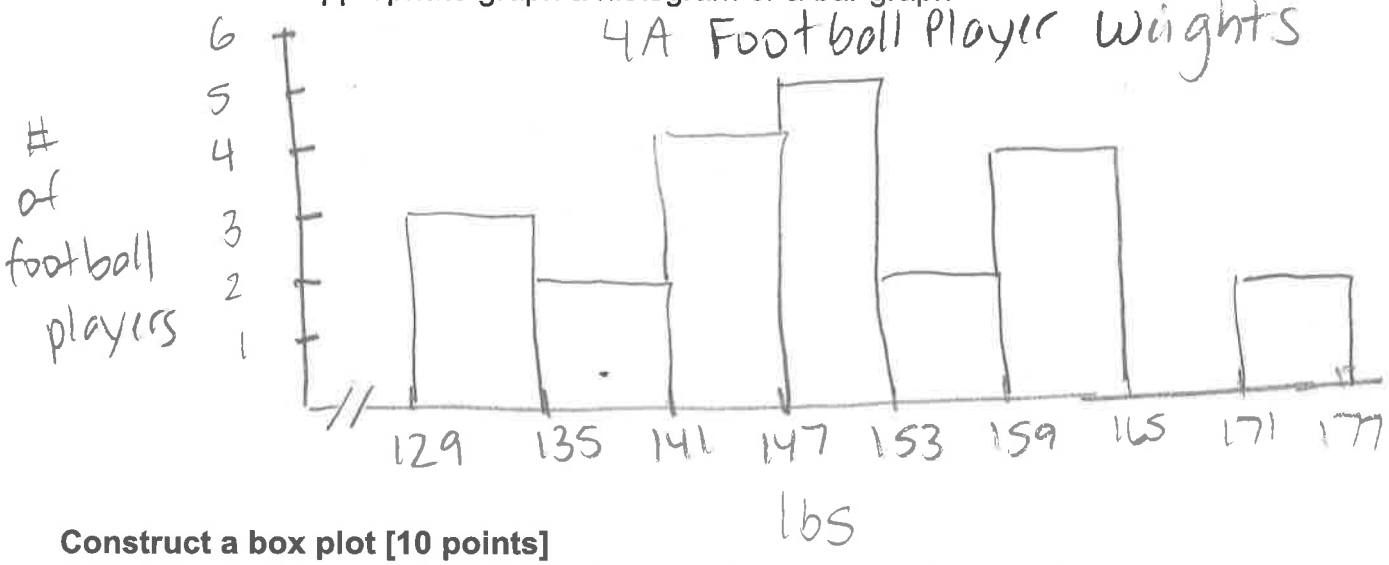
A. Construct a frequency distribution table use 8 classes

Freshman Football Players Weight Data

Y: Weight in pounds	Talley	F Number of Football Players
[129 - 135)	III	3
[135 - 141)	II	2
[141 - 147)	IIII	4
[147 - 153)	IIII	5
[153 - 159)	II	2
[159 - 165)	IIII	4
[165 - 171)		0
[171 - 177)	II	2

$$\text{width} = \frac{175 - 129}{8}$$
$$= 5.75 \approx 6$$

B. Construct the appropriate graph a histogram or a bar graph



Construct a box plot [10 points]

2. The normal monthly precipitation (in inches) for August is listed for 20 different US cities.

i	1	2	3	4	5	6	7	8	9	10
y_i	0.4	1	1.5	1.6	2	2.2	2.4	2.7	3.4	3.4

i	11	12	13	14	15	16	17	18	19	20
y_i	3.5	3.6	3.6	3.7	3.7	3.9	4.1	4.1	4.1	7.0

A. Given: If Q_1=2.1, Q_2=3.45, and Q_3=3.80. What was the interquartile range for this data?

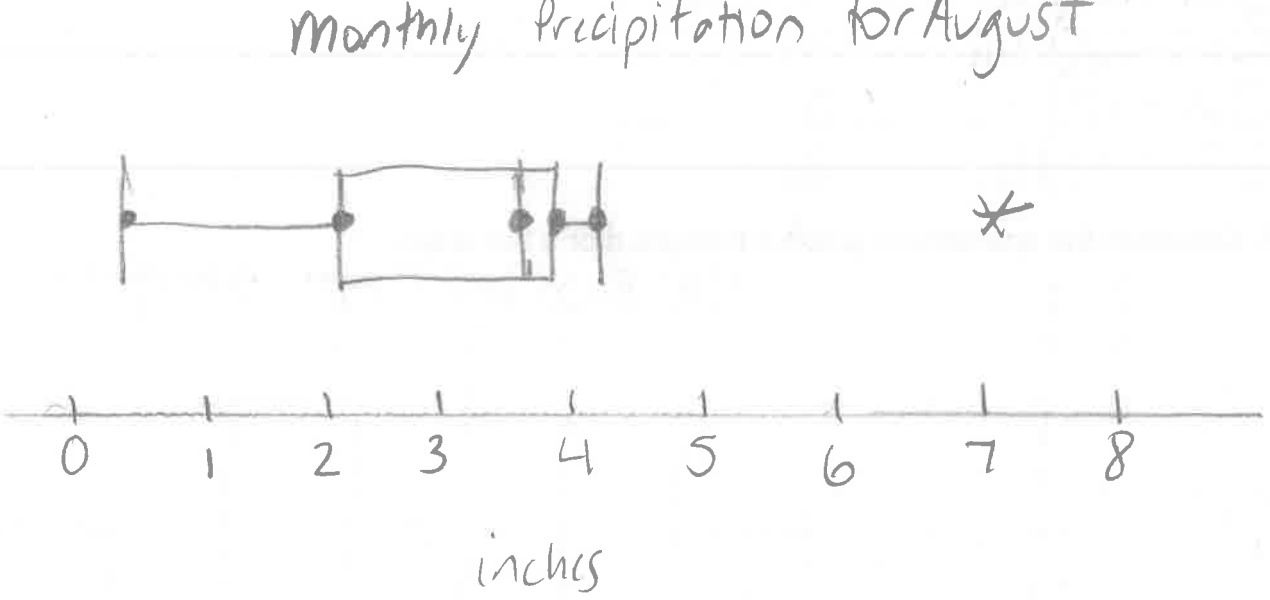
$$IQR = Q_3 - Q_1 = 3.80 - 2.1 = 1.7$$

B. What are the upper and lower limits? Are there any outliers in this data set?

$$LL = Q_1 - 1.5 \cdot IQR = 2.1 - 1.5 \cdot 1.7 = 2.1 - 2.55 = -0.45$$

$$UL = Q_3 + 1.5 \cdot IQR = 3.80 + 1.5 \cdot 1.7 = 6.35$$

C. Draw the appropriate box plot (regular or modified) based upon the data



Part 2 Multiple Choice: Circle your answer choice on the exam **AND** fill in the answer sheet shown below with the letter of the answer that you believe is the correct answer.

Problem Number	Letter of Answer	Problem Number	Letter of Answer	Problem Number	Letter of Answer	Problem Number	Letter of Answer	Problem Number	Letter of Answer
3.		6		9		12		15	
4		7		10		13		16	
5		8		11		14		17	

Use the following data for problems 3 and 5

3. Serum cotinine level is a predictor of risk of lung cancer among smokers. The level of serum cotinine in ng/ml for ~~27~~ 26 smokers is shown below. Determine the five number summery for the data.

i	1	2	3	4	5	6	7	8	9	10
y_i	11.12	12.58	13.73	14.42	18.22	19.28	20.16	23.67	25.00	25.39

1 2 3 4 5 6 7

i	11	12	13	14	15	16	17	18	19	20
y_i	29.41	30.71	32.54	32.56	34.21	36.73	37.73	39.48	48.58	51.21

i	21	22	23	24	25	26
y_i	56.74	58.69	72.37	104.54	114.49	145.43

median is at the $\frac{26+1}{2}$ = 13.5 position

$X_{13} = 32.54$ $\quad$ Q_1 is at $\frac{13+1}{2} = 7^{th}$

$X_{14} = 32.56$ $\quad$ $Q_1 = 20.16$

$Q_2 = 32.55$ $\quad$ $Q_3 = 51.21$

11.12, 20.16, 32.55, 51.21, 145.43

A 11.12, 20.16, 32.55, 51.21, 145.43
B 11.12, 20.16, 32.55, 56.74, 145.43
C 11.12, 19.28, 32.54, 51.21, 145.43
D 11.12, 19.28, 32.56, 56.74, 145.43
E None of these

4. What is the range of this data set?

$R = 145.43 - 11.12 = 134.31$

A. 31.05 ng/ml B. 36.58 ng/ml C. 134.31 ng/ml
D. 101.91 ng/ml E. None of these

5. If the standard deviation is 34.03 ng/ml and the mean is 42.85 ng/ml, what percentage of the observations are within 1 standard deviation of the mean?

$42.85 \pm 34.03 = 8.82$
$42.85 + 34.03 = 76.88$

$\frac{23}{26} \times 100\% =$
$88.46\% =$

A. 73.4 B. 94.4 C. 68.6
D. 88.5 E. None of these

Use the following information to answer questions 6-8
At All-Breed Animal Rescue Shelter, researchers studied the breed preference of dog adopters. The researchers recorded what the breed of the dog was and the length of time that the dogs were at the shelter before they were adopted.

6. What type of study were the researchers conducting?

A. Observational Study B. Experimental Design C. Neither

7. What is the independent (input) variable in this study?

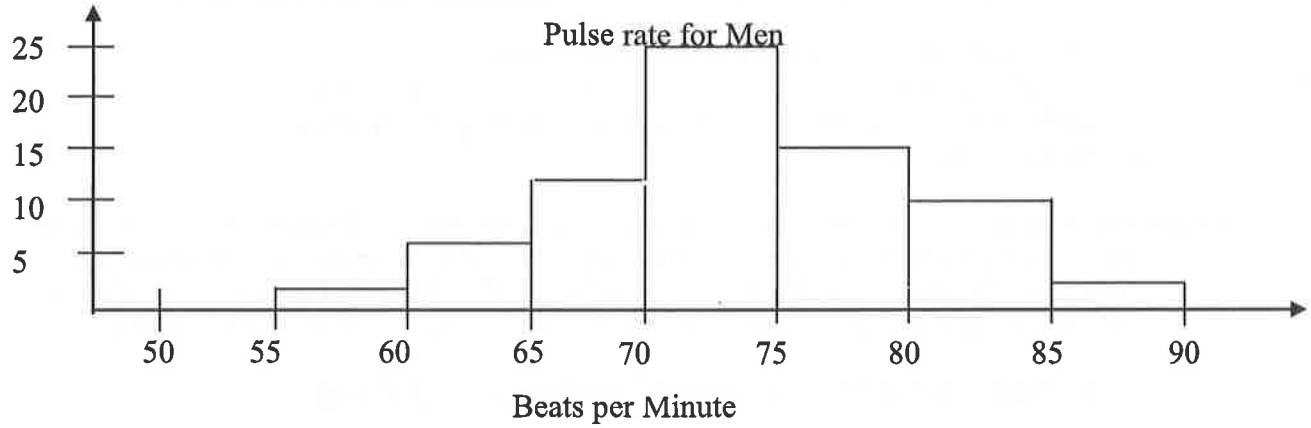
A. The length of time B. Breed Preference
C. The researchers D. There is not an observational unit in this study

8. "Length of time" is what type of variable?

A. Numerical and continuous
B. Categorical and ordinal
C. Categorical
D. None of these

Use the following information for problems 9 and 10

9. A random sample of men aged 25-50 was taken to determine the pulse rate per minute. What is the shape of the histogram?

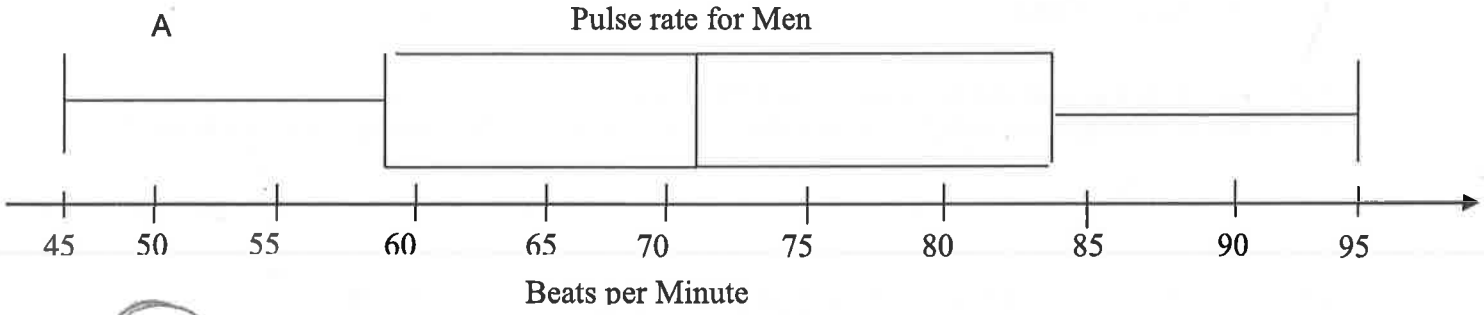


A. Right Skewed B. Left Skewed C. Uniform

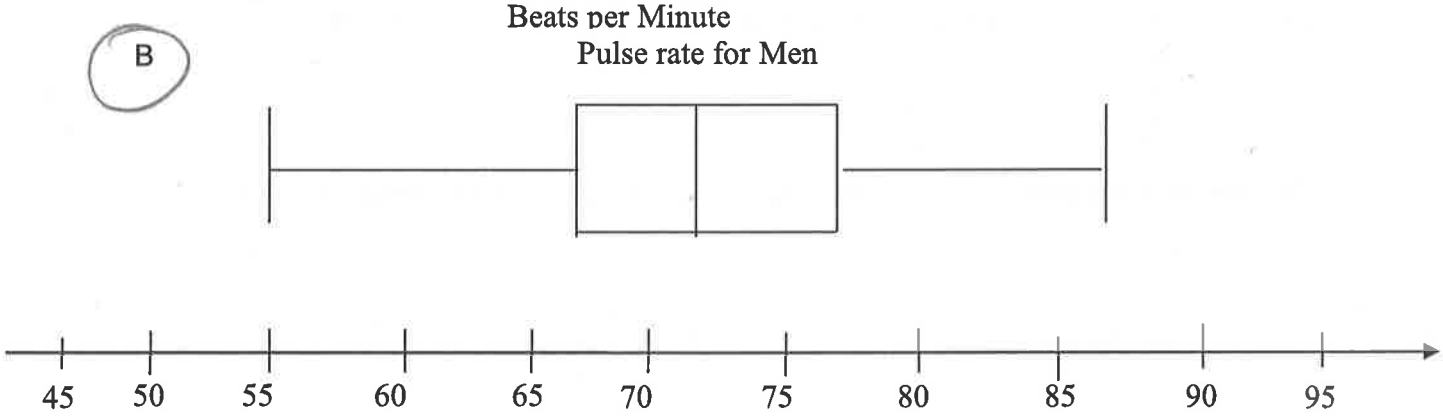
D. Bell Shaped E. None of these.

10. What would the box plot of this data look like?

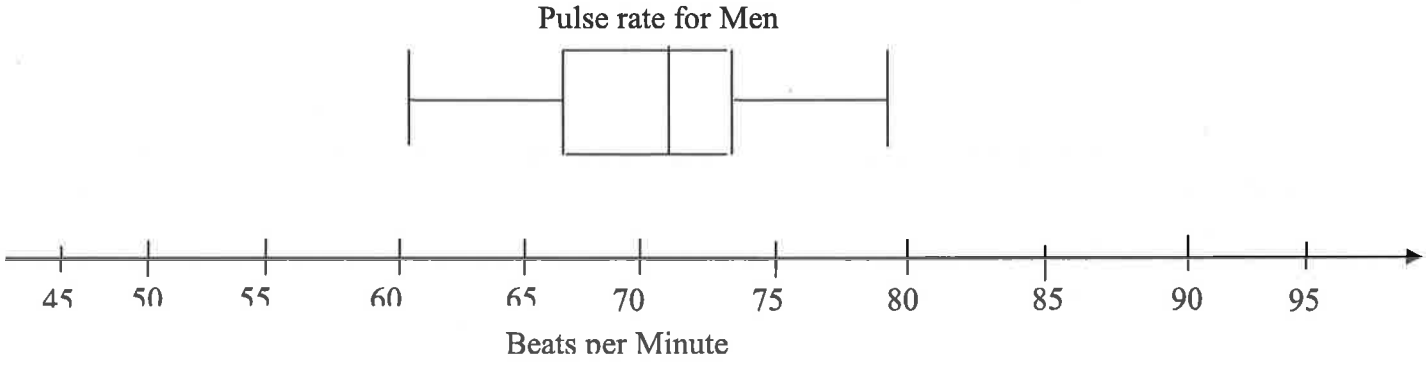
A



B



C



11. Suppose you look at two boxplots comparing the weights of male cats vs. female cats, and you find that the box for the males is much wider than the box for the females. What does this mean about the data sets?

A. Male cats weigh more than female cats overall.
B. Weights of male cats are more skewed than for female cats.
C. Male cats have more variability in their weights than female cats.
D. None of the above.

12. A Consider a population consisting of 192 individuals with unique IDs: 001, 002,…,192. Use a fragment of a random numbers table below to make the selection of 5 individuals needed in the study. List the ID's of the individuals selected for your sample. Start at the very beginning, go right, **make sure** to consider appropriate number of digits.

25319188315411009073456234610315690024911123075440129 87

Selected numbers are: 191, 154, 110 090 103

A. 253, 191. 883, 154, 110
B. 253, 191, 154, 110, 090
C. 154, 110, 090, 103, 156
D. 191, 154, 110, 090, 103
E. None of these

Determine if the following statements are true or false

13. In simple random sampling, each possible sample is equally likely to be the one obtained.

A. True
B. False

14. If an outlier is observed it should be immediately removed from the data.

A. True
B. False

15. Data for a variable Y can be transformed if it is not bell-shaped using the following transformations: $\sqrt{Y}, \log(Y), \frac{1}{Y}, or\ Y^2$.

A. True
B. False

16. Samples being taken from different depths in the soil to determine the amount of organic and inorganic material at each depth demonstrate an example of stratified sampling.

A. True
B. False

17. We do not expect any error when we estimate a population parameter from a sample statistic?

A. True
B. False

Formulas

Sample Statistics

$$\bar{y} = \frac{\Sigma y}{n}$$

Range = Max – min

$$s = \sqrt{\frac{\Sigma (y - \bar{y})^2}{n-1}}$$

$$\bar{y} = \frac{\Sigma (f_i \cdot y_i)}{n}$$

$$s = \sqrt{\frac{\Sigma (y_i - \bar{y})^2 \cdot f_i}{n-1}}$$

$$\text{Class width} = \frac{\text{Maximum Observation - Minimum Observation}}{\text{Number of Classes}}$$

$$\textit{Coefficient of Variation} = \frac{s}{\bar{y}} \times 100\% \qquad IQR = Q_3 - Q_1$$

$$\textit{Lower fence} = Q_1 - 1.5 \times IQR \qquad \textit{Upper fence} = Q_3 + 1.5 \times IQR$$