Yerevan State University

YSU Applied Statistics with R
Self-Assessment Quiz
Fall 2018

Descriptive Statistics

Michael Poghosyan

Last Revision Date: October 2, 2018

Instructions

Click on the *"Begin Quiz"* button in green to start the quiz. At the end click on the *"End Quiz"* button in green, and the system will calculate the number of correct answers. Also at the end you can hit the *"Correct"* button next to the quiz score and go back to the beginning of the quiz to see the correct answers (but try to avoid this and just try to take the quiz again and again to find your mistakes, if any). Also try not to spend more than 2-3 minutes for each question.

And do not use **R** to answer to the questions.

Quiz on Descriptive Statistics

Begin Quiz

1. We have the following dataset:

Name	Age (y)	Height (cm)	Blood Press	Salary (K AMD mntly)
Stud1	20	160	130	0
Stud2	21	164	128	200
Stud3	20	155	120	0
Stud4	30	165	122	300
Stud5	22	158	131	0
Inst	50	175	220	150

How many variables we have?

A 0 B 1 C 2 D 3 E 4 F 5 G 6

2. For the dataset above, how many observations we have?

A 0 B 1 C 2 D 3 E 4 F 5 G 6

3. For the dataset above, how many categorical variables we have?

A 0 B 1 C 2 D 3 E 4 F 5 G 6

4. For the dataset above, how many ordinal variables we have?

A 0 B 1 C 2 D 3 E 4 F 5 G 6

5. What is the product of the medians of

$$x: \ -1, 3, 4, 7, 2, 0, -1, 2, 1$$

and

$$y: \ 0, 2, -4, -1, 2, 6,$$

i.e., what is

$$median(x) * median(y) \ ?$$

A 0;

B 1;

C 3;

D 2;

E -1

6. Calculate the value of

$$x_{(2)} \cdot x_{(6)} + x_4,$$

if

$$x: \ 0, 3, 2, 1, 1, 1, 0, -1, -1, 1$$

A 0;

B 1;

C 3;

D 2;

E -1

7. What will be the output of **R** when running:

```
x <- c(-1,2,-3,2,1)
range(x)
```

Please do not use **R** to answer to this question!

A -3;

B 5;

C -3 -1 1 2 2;

D -3 2;

E -1

8. What will give in the output **R**:

```
x <- c(-1, 0, 1)
var(x)
```

A 1;

B -1;

C 2;

D 2/3;

E 3/2

9. Which of the following is not a measure of location for a dataset?

A Mean;

B Range;

C Mode;

D 2nd Quartile;

10. Which of the following is not a measure of spread for a dataset?

A IQR;

B Range;

C Standard Deviation;

D 2nd Quartile;

E Variance

11. Is there a difference between the 2nd quartile and 50% quantile?

A Yes

B No;

12. Our dataset has 9 elements. How many elements are less than or equal to the Median of that dataset?

A 4

B 5

C 6

13. Our dataset has 10 elements. How many elements are less than or equal to the third quartile of that dataset?

A 5

B 6

C 7

D 8

14. Is it always true that the number of data points less than or equal to the Median is equal to the number of data points larger than or equal to the Median?

A Yes

B No;

15. Let q_α be the $100 \cdot \alpha\%$ quantile of the dataset

$$x: \ 0, 2, -1, -1, 3, 3, 1, 4, 3, -1, 4, 3, 1, 0$$

Calculate

$$|q_{0.5} - median(x)| + q_{0.3}$$

A -1

B 0

C 0.5

D 1

E 1.5

16. How many outliers have the following dataset:

$$-10, -10, -10, 1, 1, 1, 100, 100 \text{ ?}$$

A 0

B 1

C 2

D 3

17. Calculate the MAD (Mean Absolute Deviations from the Mean) of the dataset:

$$x: \ 101, 104, 103, 99, 99, 100.$$

A 1

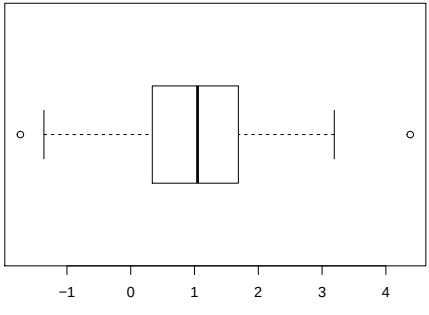
B 120

C 3

D 5/3

E 4.5

18. Here is a boxplot of some dataset. Which statement is not true?



A Dataset is symmetric

B Dataset has outliers

C 50% of observations are larger than 1.8

D The third quartile is ≈ 1.8

E IQR ≈ 1.3

19. Which of the followings is not giving the shape of the PDF of a r.v. behind the data (all bin widths are the same for the histograms):

A KDE

B Density Histogram

C Frequency Histogram

D Boxplot

E Stem-and-Leaf Plot

F Relative Frequency Histogram

20. The following Stem-and-Leaf plot is obtained in **R** (no rounding was made):

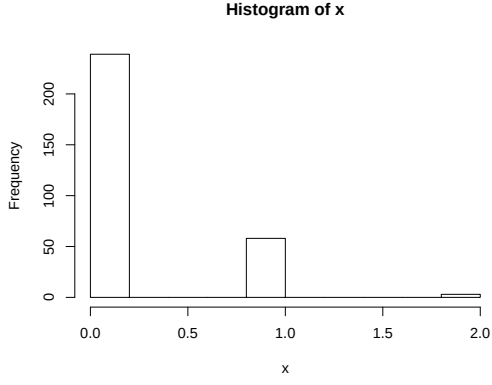
```
The decimal point is 1 digit(s) to the left of the |

 -1 | 10
 -0 |
  0 |
  1 | 015
  2 | 11122
  3 | 2
  4 | 01
```

Find the mode of the dataset.

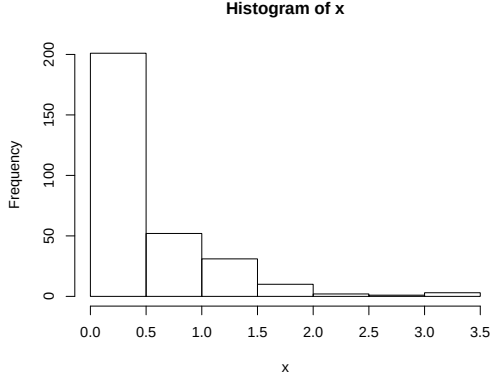
A 2	B 0.2	C 4
D 40	E 2.1	F 21
G 0.21	H 2.111	

21. The following histogram is obtained using **R** to generate 300 random numbers from some distribution. Which distribution was used?



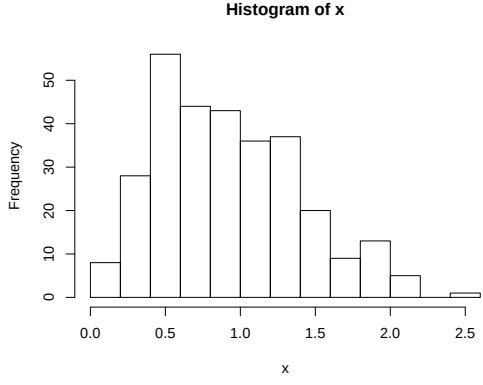
A Normal
B Exponential
C Poisson
D Uniform
E Bernoulli
F None of the above

22. The following histogram is obtained using **R** to generate 300 random numbers from some distribution. Which distribution was used?



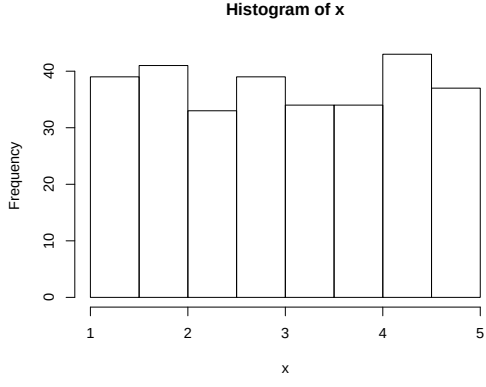
A Normal
B Exponential
C Poisson
D Uniform
E Bernoulli
F None of the above

23. The following histogram is obtained using **R** to generate 300 random numbers from some distribution. Which distribution was used?



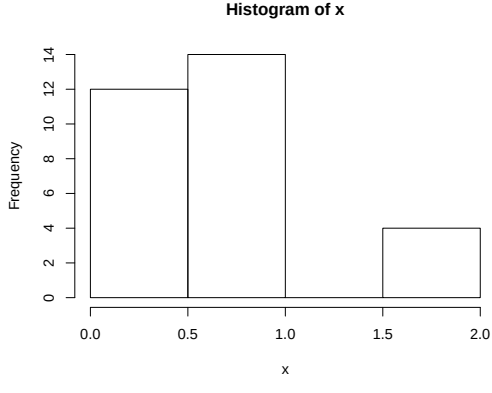
A Normal
B Exponential
C Poisson
D Uniform
E Bernoulli
F None of the above

24. The following histogram is obtained using **R** to generate 300 random numbers from some distribution. Which distribution was used?



A Normal
B Exponential
C Poisson
D Uniform
E Bernoulli
F None of the above

25. The following histogram is obtained using **R** to generate random numbers from the $Binom(4, 0.2)$ distribution. Find the sample size.



A 10 B 20 C 25
D 24 E 32 F 30

26. For the sample in the previous problem, calculate the mean.

A 0.5

B 0.72

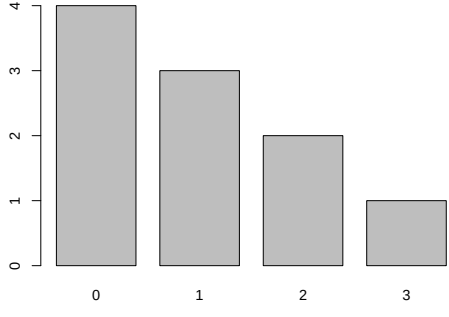
C 11/15

D 0.75

E 1

F 0.6(3)

27. What is representing the following graph?



A Boxplot
B Stem-and-Leaf plot
C Frequencies Histogram
D Bar Plot
E Pie Chart

28. For the dataset from the previous question, calculate the median.

A 2

B 1

C 3

D 1.5

E 0.5

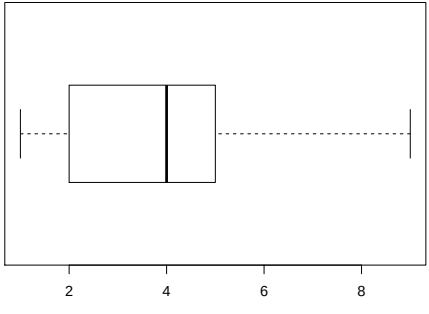
29. Is the Median or the Mean shown in a Boxplot?

A Median

B Mean

C Both

30. The following is the boxplot for some dataset. What is the median of the dataset?



A 1
B 2
C 3
D 4
E 5

31. For the dataset above, what is the average of the first and third quartiles?

A 2.5

B 3

C 3.5

D 4

E 4.5

32. For the dataset above, what is the IQR?

A 2.5

B 3

C 3.5

D 4

E 4.5

33. For the dataset above, what is the range?

A 3

B 4

C 6

D 8

E 9

34. For the dataset above, assume that the number 7 was among the observations. Is it an outlier for that dataset?

A Yes

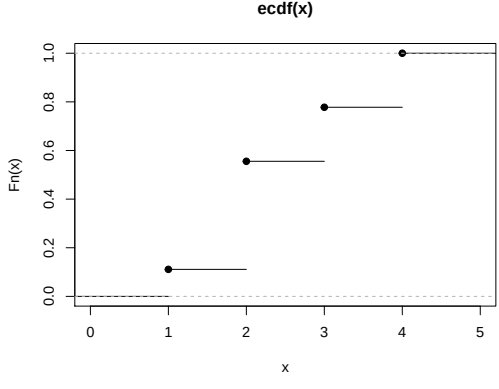
B No

35. For the dataset above, can 0 be a data point?

A Yes

B No

36. The following is the Empirical CDF for some dataset. What is the median of the dataset?



A 1
B 1.5
C 2
D 2.5
E 3

37. Is it possible to calculate the number of datapoints from the previous question ECDF graph?

A Yes

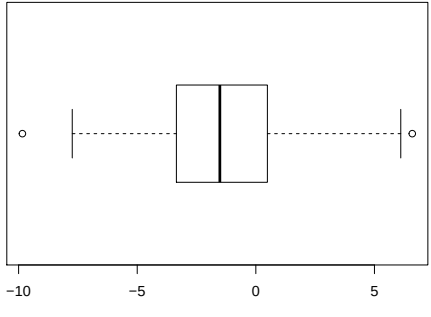
B No

38. Is it possible to calculate the frequency of datapoints from the previous question ECDF graph?

A Yes

B No

39. The following is the BoxPlot for some dataset of size 50. The quartiles are: $Q_1 = -3.348907$, $Q_2 = -1.516847$, $Q_3 = 0.486812$. Approximately how many datapoints we have in [-3.348907, 0.486812]?



A 10 B 15 C 20
D 25 E 30

40. For the dataset from the previous question, approximately how many datapoints we have in [-3.348907, -1.516847]?

A 8

B 12

C 18

D 22

E 26

41. From which distribution I have generated the dataset giving the boxplot from the previous question?

A Bernoulli

B Binomial

C Normal

D Exponential

E Poisson

42. What will output the following **R** command:

```
x <- c(-1, -2, 3, -4, 5)
y <- sort(x)
```

A
```
-1 -2  3 -4  5
-4 -2 -1  3  5
```
B
```
[1] -1 -2  3 -4  5
[1] -4 -2 -1  3  5
```
C Nothing

D Other

43. The following is the output of running the

```
x <- rbinom(150, 10, 0.6)
x
```

command in **R**:

```
  [1] 6 2 6 8 10 7 7 2 7 6 8 7 5 7 6 2 4 6 6 9 6 6 3 6 6 8 4
 [28] 7 6 6 6 7 5 6 5 5 9 7 5 7 4 5 5 3 7 6 4 7 7 5 7 7 6 6
 [55] 9 9 7 6 4 5 4 6 6 5 4 6 7 5 9 7 6 9 5 8 7 5 8 6 5 5 2
 [82] 3 6 6 6 3 5 6 4 6 5 8 5 7 8 7 5 3 6 4 5 5 6 3 7 5 4 5
[109] 4 6 6 4 5 5 9 6 6 8 5 8 6 7 8 6 8 5 6 7 4 4 4 7 6 6 4
[136] 6 7 4 7 9 7 9 6 8 8 7 4 8 4 6
```

What is the value of the 86-th element of x?

A 3

B 6

C 7

D 5

E 8

44. Which of the following sets can be the set of all deviations for some dataset?

A 2, -2, 1, -2

B 2, 4, 0, 2

C 3, 2, -1, -4

45. For the previous example, if the set can be the set of deviations, calculate the sample variance (by using $n-1$ in the denominator):

A 4.(3)

B 7

C 10

D -2

E 0

46. Is it possible to recover the mean of a dataset, if we will have only deviations of that dataset?

A Yes

B No

47. What is the value of

$$\int_{-\infty}^{+\infty} e^{-x^2/2}\,dx\ ?$$

A 1

B -1

C $+\infty$

D $\sqrt{2\pi}$

E $2\sqrt{\pi}$

F 2π

48. Was this Quiz easy?

A Yes

B No

End Quiz

Quiz Result

Score: Correct