Topics: Descriptive Statistics and Probability
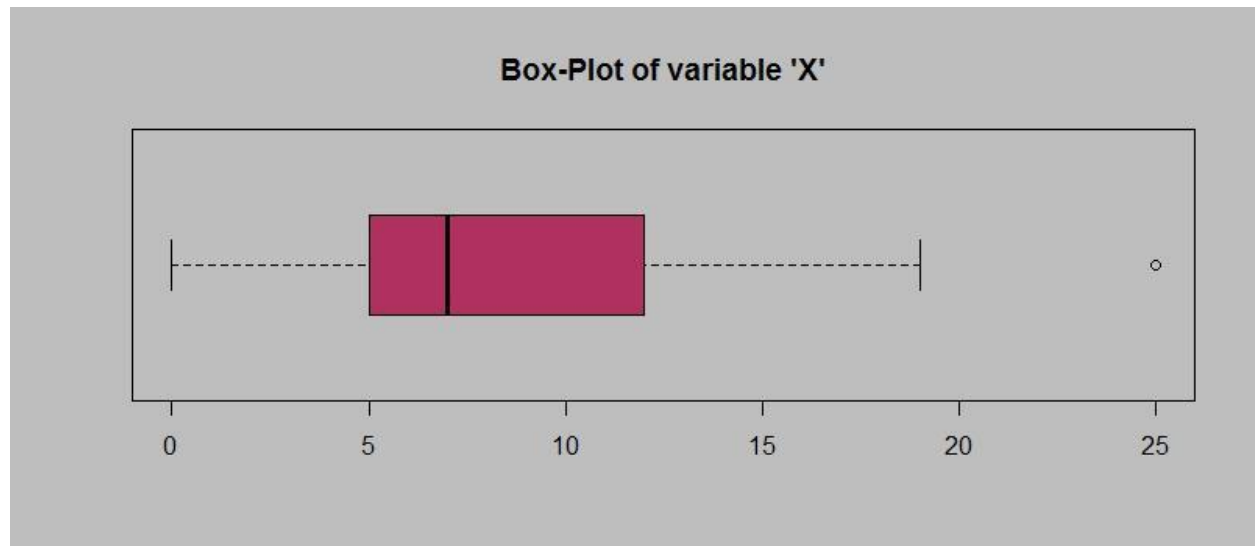
1. Look at the data given below. Plot the data, find the outliers and find out μ, σ, σ^2

Name of company	Measure X
Allied Signal	24.23%
Bankers Trust	25.53%
General Mills	25.41%
ITT Industries	24.14%
J.P.Morgan & Co.	29.62%
Lehman Brothers	28.25%
Marriott	25.81%
MCI	24.39%
Merrill Lynch	40.26%
Microsoft	32.95%
Morgan Stanley	91.36%
Sun Microsystems	25.99%
Travelers	39.42%
US Airways	26.71%
Warner-Lambert	35.00%

2.

Questions referred to from *Aczel A., Sounderpandian J., Complete Business Statistics (7ed.)*

Box-Plot of variable 'X'

Answer the following three questions based on the box-plot above.

(i)   What is inter-quartile range of this dataset? (please approximate the numbers) In one line, explain what this value implies.

Ans:- Frist Quartile at 25 = Q1=5

Third Quartile at 75 = Q3 = 12

IQR= Q3-Q1 =12-5 =7

A larger IQR indicates a greater spread of the data within the middle 50%, while a smaller IQR indicates that the data is more tightly clustered around the median.

(ii)   What can we say about the skewness of this dataset?

Ans:- Right skewness, mode < median < mean

(iii)   If it was found that the data point with the value 25 is actually 2.5, how would the new box-plot be affected?

Ans:-   If the data point is in 2.5, then it is not situated in the outlier it is in the boxplot and the boxplot range is in 0 to 20, ANd there is no outliers will be there.

3.

Histogram of variable 'Y'
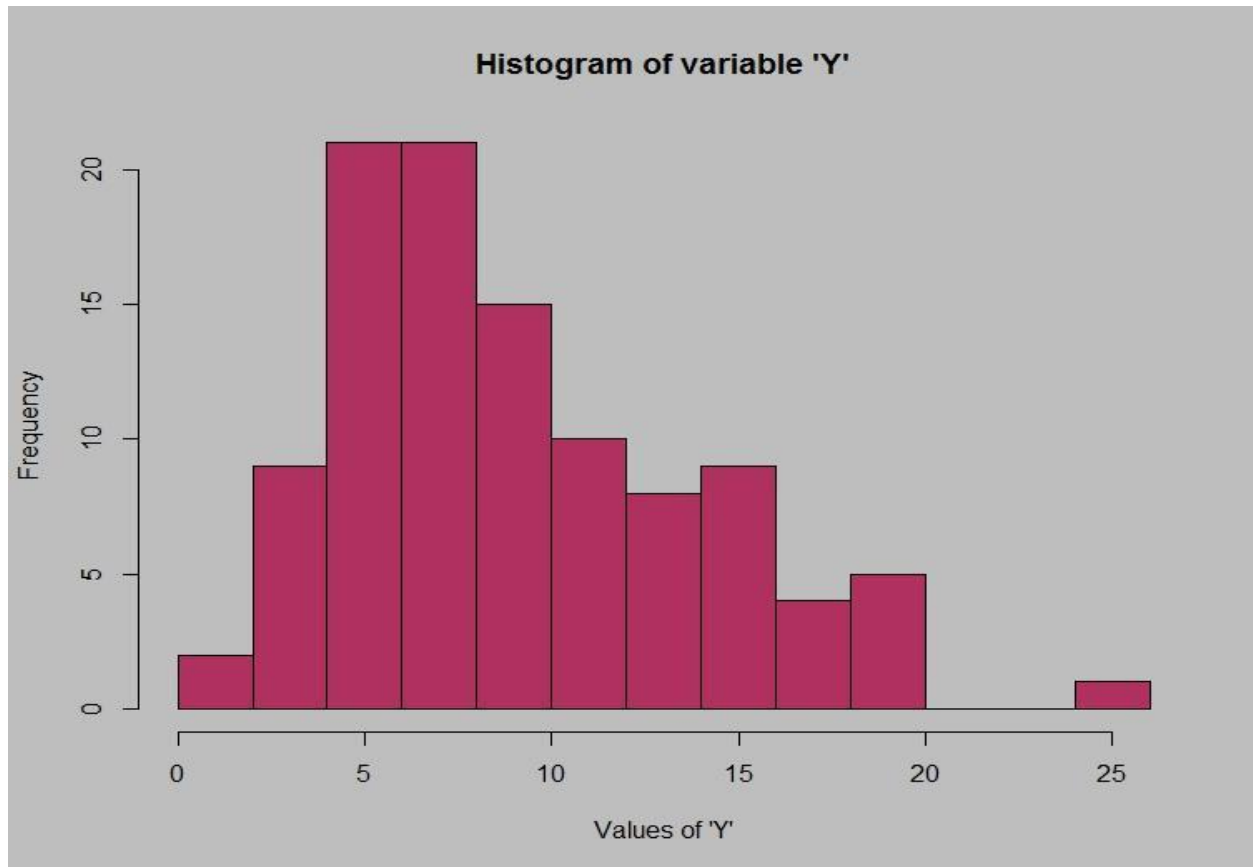
Answer the following three questions based on the histogram above.

(i)   Where would the mode of this dataset lie?

       Ans:-   Mode of the dataset lie in range of 5 to 7

(ii)   Comment on the skewness of the dataset.

       Ans:-   The skewness is positively skewed or Right skewed mode < median < mean

(iii)   Suppose that the above histogram and the box-plot in question 2 are plotted for the same dataset. Explain how these graphs complement each other in providing information about any dataset.

       Ans:-   We can find Median in Boxplot and Mode in histogram.

By using Histogram we can find the data is continuous or discrete and provides the frequency distribution.

By using Boxplot we can find that the majority of the data lies in the certain region or certain area and it provides data in quantile distribution. And it also provides whisker length which provides data which are not in the majority area.

4. AT&T was running commercials in 1990 aimed at luring back customers who had switched to one of the other long-distance phone service providers. One such commercial shows a businessman trying to reach Phoenix and mistakenly getting Fiji, where a half-naked native on a beach responds incomprehensibly in Polynesian. When asked about this advertisement, AT&T admitted that the portrayed incident did not actually take place but added that this was an enactment of something that "could happen." Suppose that one in 200 long-distance telephone calls is misdirected. What is the probability that at least one in five attempted telephone calls reaches the wrong number? (Assume independence of attempts.)

5. Returns on a certain business venture, to the nearest $1,000, are known to follow the following probability distribution

x	P(x)
-2,000	0.1
-1,000	0.1
0	0.2
1000	0.2
2000	0.3
3000	0.1

(i)   What is the most likely monetary outcome of the business venture?
        Ans:-   Most likely monetary outcome of the business venture is x=2000
(ii)   Is the venture likely to be successful? Explain
        Ans:-   P(x>0) = 0.6, implies that there is a 60% chance that the venture returns would be in profit and greater than expected results. P(x=0), implies that the venture with no profits and loss is 0.2 . And P(X<0), implies that the venture with losses is 0.2 . So, by analysing the probability of profits and loss and neutral we decide that it has chances to be likely to be successful.
(iii)   What is the long-term average earning of business ventures of this kind? Explain
        Ans:-   The long-term average earning of business ventures of this kind is x*P(x)= (-2000*0.1)+(-1000*0.1)+(0*0.2)+(1000*0.2)+(2000*0.3)+(3000*0.1)=**800**
(iv)   What is the good measure of the risk involved in a venture of this kind? Compute this measure
        Ans:-   The good measure of the risk involved in a venture of this kind =0.2
        Because, the loss probability is 0.2