**Topics: Descriptive Statistics and Probability**

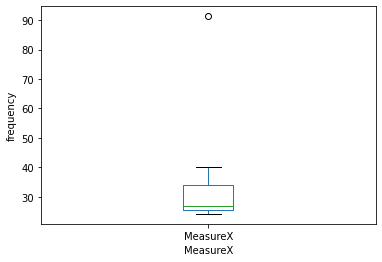
1. Look at the data given below. Plot the data, find the outliers and find out

|  |  |
| --- | --- |
| **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% |

Mean value is: 33.27133333333333

Standard deviation value is: 16.945400921222028

Variance for the given set is: 287.1466123809524



One outlier existing above the upperfence and the outlier is

| **Nameofcompany** | **MeasureX** |
| --- | --- |
| **10** | Morgan Stanley | 91.36 |



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

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

Ans:) IQR = upperQuartile - lowerQuartile = (12 - 5) = 7

The above IQR tells us that half of the values in the dataset below median are closer or nearer to median and there is wide spread of data on right side with outliers

1. What can we say about the skewness of this dataset?

Ans:) The dataset can be considered as Right Skewed.I.e half of the values above the median are spreadout very far with outilers.

1. 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) There would be no outlier existing , and the median value be slightly increased



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans: The mode of the dataset lies in an interval of [4-6]. 20 datavalues lies in this interval

1. Comment on the skewness of the dataset.

Ans: Dataset is right Skewed.

1. 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) Through the above histogram we can conclude that the data is right skewed as data is continuous until 20 but has another value at 25 with a gap in b/w 20 and 25 which implies may be outliers are existing, which can be confirmed with the box plot that outliers are existing above the upper fence

1. 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.)

Ans:

****Given :****one in 200 long-distance telephone calls is misdirected.

****To find :****probability that at least one in five attempted telephone calls reaches the wrong number

****Solution:****

one in 200 long-distance telephone calls is misdirected

=>  probability of call misdirecting  p = 1/200

     Probability of call not Misdirecting = 1 - 1/200 = 199/200

Number of Calls = 5

P(x) = ⁿCₓpˣqⁿ⁻ˣ

n = 5

p = 1/200

q = 199/200

at least one in five attempted telephone calls reaches the wrong number

= 1  -  none of the call reaches the wrong number

= 1  - P(0)

= 1   -  ⁵C₀(1/200)⁰(199/200)⁵⁻⁰

= 1  -  (199/200)⁵

= 0.02475

****probability that at least one in five attempted telephone calls reaches the wrong number = 0.02475****

1. 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 |

1. What is the most likely monetary outcome of the business venture?

Ans: The most likely monetary outcome of business venture is, 2000.I.e the most likely probability p(x) = 0.3 occurring at x = 2000.

1. Is the venture likely to be successful? Explain

Ans: The venture is likely to be successful as avg probability of getting profits is more compared to avg probability of getting loses or being neural(1100>300)

E(x being successful) = 1000\*0.2+2000\*0.3+3000\*0.1 =1100

E(x being not successful) = -(2000\*0.1+1000\*0.1+0\*0.2) =300

1. What is the long-term average earning of business ventures of this kind? Explain

Ans: Long term avg earnings of business venture is given by its expected value  
 E(x) = -(2000\*0.1)+(-1000\*0.1)+0\*0.2+1000\*0.2+2000\*0.3+3000\*0.1 = 800

Avg earnings is 800 in long run

1. What is the good measure of the risk involved in a venture of this kind? Compute this measure

Ans: To find the risk involved in a venture of this kind we can calculate the probability of getting losses

The Returns in first two observations are in negative. If those probabilities are summed

up, then the total probability of losses will be 0.1+0.1 = 0.2

Therefore, the probability of risk involved in the venture is 0.2

In long Run the venture being in losses is

E(x being not successful) = -(2000\*0.1+1000\*0.1+0\*0.2) =300