**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% |

Soln: Solved in (AssignmentSet1\_Q1.ipynb)



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.

Soln: Lets say Q1=5 and Q3=12(approximately)

Inter-quartile range(IQR)=Q3-Q1=12-5=7

This values implies that the most of the data points are above 7.

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

Soln: From the above plot we can see that most of the data is distributed over the right

side, hence it is left skewed.

(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?

Soln: There will be no outlier.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Soln: The mode of this dataset lie between 4 to 8.

1. Comment on the skewness of the dataset.

Soln: Most of the data is concentrated on the left side, hence it is right skewed histogram.

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.

Soln: Histogram is used to find the mean, median, mode, skewness and kurtosis and the box

Plot is used to find the IQR, quartiles and outliers.

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.)

Soln: One in 200 long-distance telephone calls is misdirected

Probability of call misdirected = 1/200=0.005

Probability of calls not misdirected=1-0.005=0.995

Number of calls = 5

At least one in five attempted calls reaches wrong number

= 1 – At least one in five attempted calls not reaches wrong number

= 1- (0.995)^5

= 1-0.975

=0.025

=2.5%

The probability of at least one in five attempted calls reaches wrong number is 2.5%.

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?

Soln: The most likely monetary outcome of the business venture is 0.3.

1. Is the venture likely to be successful? Explain

Soln: Yes, the venture is likely to be successful.

P(x=1000)+P(x=2000)+P(x=3000)

= 0.2+0.3+0.1

= 0.6

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

Soln: (0.1)(-2000)+(0.1)(-1000)+(0)(0.2)+(1000)(0.2)+(2000)(0.3)+(3000)(0.1)

= 800

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

Soln: The good measure of the risk involved in a venture of this kind is the standard

deviation .