**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% |

Ans) Box plot is shown in attached python file.

Outlier is – Morgon Stanley

Mean – 0.3320

Variance – 0.0287

Standard Deviation – 0.1685



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) Inter-Quartile range = Upper Quartile range – Lower Quartile range = 12-5 = 7

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

Ans) Looks like it is right-skewed or positively skewed.

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) The outlier would have been removed.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans) Most probably we can assume that the mode can lie in range of 3 to 12.

1. Comment on the skewness of the dataset.

Ans) The dataset given is right skewed or positively 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) We can say that both the dataset is right skewed or positively skewed and also both have a outlier too.

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) Lets say that P(a) = 1/200 as one phone call is mis-connected

Then right connected calls P(b) = 199/200

Probability for 5 connected calls = [1-(199/200)^5] \* 100

= 2.48

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) Most likely will be one with high probability which is $2000.

1. Is the venture likely to be successful? Explain

Ans) Looking forward to the probabilities we can assume that it can be a good venture as positive probability are more.

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

Ans) Long-term average earning = (-2000\*0.1) + (-1000\*0.1) + (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

Ans) Risk can be computed by variance and standard deviation.

Here, Variance = 3500000 Standard Deviation = 1870.82

As standard deviation’s value is too large from average earning we can compute that the venture is highly risky.