**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: >> Morgan Stanley- 91.36% is an outlier.

>> Mean= 33.27

>> Standard deviation= 16.37

>> Variance= 268.003



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.
2. What can we say about the skewness of this dataset?
3. 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: 1). IQR = Q3- Q1= 12-5= 7.

2). We can say that the data is positively skewed.

3). There would be no outliers in the data and the data would have followed normal distribution.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?
2. Comment on the skewness of the dataset.
3. 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: 1). The mode would be 9 approximately.

2). The dataset is positively skewed.

3). We can see the outliers through both graphs but boxplot gives the specific range meaning anything outside that range is an outlier. Also, through boxplot we can find the exact mean value while through histogram we cannot tell the exact mean value. Moreover, through both graphs we can say about the skewness of the dataset. Furthermore, we can also know the 25th and 75th percentile through boxplot.

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: n= 5, p= 0.005

The probability that at least one in five attempted telephone calls reaches the wrong number is 0.0245.

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?
2. Is the venture likely to be successful? Explain
3. What is the long-term average earning of business ventures of this kind? Explain
4. What is the good measure of the risk involved in a venture of this kind? Compute this measure

Ans: 1). The most likely monetary outcome of the business venture is the one with the highest probability that is 2000 with a probability of 0.3.

2). Yes, the venture is likely to be successful as the probability of profit is 0.2+0.3+0.1= 0.6 which is

greater than the probability of loss 0.1+0.1= 0.2.

3). The long-term average earning of business ventures of this kind is sum(x\*p(x)) = 800.

4). The measure of risk can be determined through standard deviation. Therefore, the risk involved is around 1572.33.