**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**: There is one Outlier in the above data , i.e the data of Morgan Stanley with data point of 91.36%

The Mean of the data is 33.27

The Std of the data is 287.14

The Variance of the data is 16.945



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:** The IQR of the above data set is 12 -5 = 7(approximately). This values implies that the most of the data points occurs in between this range

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

**Ans:**  By observing the above box plot , we can say that the skewness of the data right skewed , as the median is in left portion of the IQR

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:**  Then there would be no outliers in the data set.

The IQR might not get affected but Median value of the set will be changed



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Ans :**  The mode of Data set would lie in between 4-8

1. Comment on the skewness of the dataset.

**Ans:**  By observing the histogram we can say that the data 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:**  They both are right-skewed and both have outliers the median can be easily visualized in box plot where as in histogram mode is more visible.

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:  If** 1 in 200 long-distance telephone calls are getting misdirected.

probability of call misdirecting   = 1/200

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

**The** probability for at least one in five attempted telephone calls reaches the wrong number

Number of Calls = 5

n = 5

p = 1/200

q = 199/200

P(x) = at least one in five attempted telephone calls reaches the wrong number

P(x) = (nCx) (p^x) (q^n-x)

P(1) = (5C1) (1/200)^1 (199/200)^5-1

P(1) = 0.0245037

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:** For X = 2000 since probability is 0.3 (more than others)

1. Is the venture likely to be successful? Explain

**Ans:** P(X>0) = 0.2+0.2+0.3+0.1 = 0.8

There is a good chance 80% profit

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

**Ans:** The long-term average is Expected value = Sum (X \* P(X)) = 800$ which means on an average the returns will be + 800$

1. 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 depends on the Variability in the distribution. Therefore a good measure to evaluate the risk for a venture of this kind would be variance or standard deviation of the variable X.

SD = 1870.829

Variance =3500000

The large value of standard deviation of $1870 is considered along with the average returns of $800 indicates that this venture is highly risky.