**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:**



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=Q3-Q1=12-5=7, only 2 values are before the IQR falling. IQR is close to median & whisker is higher towards upper side.

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

**Ans:** As whisker is larger on upper side, it is Right 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:** There would not be any outlier & whisker on left side will be almost equal to right side, plot will look like No Skewed.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Ans:** Most of the data set lie between 4-8.

1. Comment on the skewness of the dataset.

**Ans:** Data set is Left 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 differentiate Mode in box plot but cannot do that in Histogram.

Both are Right Skewed & outliers at 25.

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 each call is misdirected with the same probability, then that probability is 1/200. Which means that the probability of all five calls directed properly is (199/200)^5, i.e., the probability of at least one misdirect is ***1-(199/200)^5*** which is about ***0.02475(2.475%)***

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 the business venture  is $2000 as it has maximum probability = 0.3.

1. Is the venture likely to be successful? Explain

**Ans:** Yes, the venture is likely to be successful because the monetary outcome is a positive number that is $ 800(= (0.1)(−2,000) + (0.1)(−1,000) + (0.2)(0) + (0.2)(1,000) + (0.3)(1,000) + (0,1)(3,000)= 800) which is near to $1000 as positive indicator.

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

**Ans:** Long term business earning will remain as expected value= ∑E(X)P(X)  = $800

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

**Ans:** Good measure of the risk can be measured through Standard Deviation.

Population Std Dev post data normalization of given data set=0.341565

Population Std Dev post data normalization of (x\*(Px))=0.335927