**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 :

Mean (µ) : 33.27133%

SD (σ) : 16.9454%

Var (σ²) : 02.8714%

Outlier : 91.36% (Morgan Stanley)



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 = (12-5) = 7 i.e. 50% of the whole data is present inside this range.

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

Ans : The data is 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 : May be the boxplot is lacking its symmetry due to the outlier 25. So replacing 25 by 2.5 , we may get changed mean and median value over there. May be mean approximately equal to the median value in this case.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans : The Mode lies in between 5 to 7

1. Comment on the skewness of the dataset.

Ans : Data is 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: In Histogram we are able to get the modal value bot not exactly the median value. But incase of the Boxplot we can obtain the median value. Looking either of the plot we can comment on the distribution of the data. In histogram its little bit difficult to comment outlier, whereas in boxplot we can see the outlier in dot or oval shape.

Looking at the histogram we can’t comment the Interquartile range (major 50% of the data) but incase of boxplot we can define the Inter quartile Range.

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: We are reaching one wrong number in 200 calls. So, prob. of reaching a wrong no is 1/200.

Hence in cycle of 40 attempts of 5 calls we are reaching one wrong number.

So, we conclude that probability of reaching one wrong number in 5 calls is 1/40 = 0.025.

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: $2,000

1. Is the venture likely to be successful? Explain

Ans: Probability of Getting profit is (0.2+0.3+0.1) = 0.6. So, we can say the venture is likely to be successful.

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

Ans: As we know E(X) = Σ ( Pi Xi )

E(X) = 800 which is positive.

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