**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 |

One outlier is existing at 91.36

Mean = 33.271333

Standard Deviation = 16.945401

Variance = 287.1466123809



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.

A: Inter Quartile Range = 12 – 5 = 7, IQR value depicts the spread of the middle values and also tells how the other values are far from the mean.

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

A: The data in the above diagram represents, it to be right skewed which implies the data distribution 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?

A: Though the value 25 is replaced by 2.5, the data distribution is right skewed which implies the data distribution is positively skewed.

The new box plot would have a greater Inter Quartile Range as well as

outlier will be excluded as the outlier was present at value 25 perviously.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

A: Mode of the dataset lies between 4 to 8.

1. Comment on the skewness of the dataset.

A: The data in the above graph represents rightly skewed which implies positive

skewness.

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.

A: Histogram has contains data continuously until 20 but has another value 25 with

the gap in between 20 & 25 which implies has an outlier being present. Boxplot

directly shows a indication of outlier 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.)
2. 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?

A: The most likely monetary outcome of the business venture is 2000 as the highest

probability lies at this value.

1. Is the venture likely to be successful? Explain

A: Yes, the venture is likely to be successful as values at three positions with their

probabilities depict higher rate as compared to other values.

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

A: E = (-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

A: The returns in first two observations are in negative. If those probabilities are summed up, then the total probability will be 0.2.

Therefore, the probability of risk involved in the venture is 0.2.

E(venture not be successful ) = (2000\*0.1) + (1000\*0.1) = 300