**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- 0.3327133, Sigma-0.169454, Var-2.871466

Morgan Stanley is an outlier of 91.36 using barplot.



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, it means that the data distributed around from 5 to 12 with median around 7.

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

Ans- The mean is greater than the median as per the boxplot so it is a right skewed dataset.

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- if it was found that the data point is actually 2.5 instead of 25, the outlier in the boxplot will be removed. Whether the median shifts or not depend on the size of the data. And it will reduces right skewness of the data.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans- The mode lie in two numbers (4 and 10) consider the highest rectangle and the point meet on x- axis.

1. Comment on the skewness of the dataset.

Ans- The skewness of the histogram moves towards right side so it is a 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- The box plot showed that the an outlier at 25 at Y value. Both the plots indicates the +ve skewness of the dataset.

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- let us consider the probability of I call misdirected out of 200 as an event A

Probability of occurring of event A= 1/200.

P(A)=1/200

Probability of having at least one successful call will be= 1-P(A)= 1-1/200=0.967

As every event is independent of other event the probability will be 1-(0.967)^5=0.02475=2%chance.

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- x= 2000 with the highest probability of 0.3 by using expected value formula x.p(x)

1. Is the venture likely to be successful? Explain

Ans- success of the venture can be defined as many ways, but based on the data we can look at and positive returns can make the ventures successful and the probability distributions gives us an idea about the long term chances of the earning given values of returns so there would be 60% probability that the venture would be successful. = 0.2+0.3+0.1=0.6\*100=60%

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

Ans- The expected value is nothing but the weighted mean or average= x.p(x)=800

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

Ans- Standard deviation (sigma) is always use to find the risk,

lesser the value of SD, lower the risk of the venture.

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