**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=

Outlier is Morgan Stanley 91.36%



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= inter-quartile range : 3q-1q

Interquartile range (IQR) is a measure of the spread of a dataset, defined as the difference between the 75th and 25th percentiles, also known as the first and third quartiles. The first quartile (Q1) is also called the lower quartile and the third quartile (Q3) is also called the upper quartile. In other words, the IQR is the range of the middle 50% of the data.

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

Ans= 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=

• The median (middle value) of the data set would change, as the value of 2.5 would likely be closer to the middle of the data set than the value of 25.

• The range of the data set would change, as the value of 2.5 is much smaller than the value of 25. This would likely result in a smaller range and a smaller interquartile range (the range of values between the first and third quartiles).

• The outliers would change, as the value of 25 would no longer be considered an outlier, but the value of 2.5 would be considered an outlier.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans= Mode will be at 6

1. Comment on the skewness of the dataset.

Ans= It 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.
2. 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= The answer would be 1 - (199/200)^5

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= venture is $2000

1. Is the venture likely to be successful? Explain

Ans=

So, in this case, the expected return would be:

(-2,000 \* 0.1) + (-1,000 \* 0.1) + (0 \* 0.2) + (1,000 \* 0.2) + (2,000 \* 0.3) + (3,000 \* 0.1) = $500

As the expected return is positive, it is likely that the venture will be successful.

However, it is worth noting that expected value is a measure of central tendency and it doesn't account for the variability of the possible outcomes. As a better picture of the performance of the venture, one should look at the standard deviation or variance of the distribution as well.

So, it's possible that the venture is likely to be successful but there's also a chance for large losses.

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

Ans=

1.long-term average earning of business ventures of this kind is $1000.

2.It is important to note that the expected value is not the same as the most likely outcome, but rather a measure of the central tendency of the distribution. It represents the average outcome if many similar business ventures were taken and their returns were added together and divided by the number of ventures.

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

Ans=

To calculate the standard deviation, we need to know the mean (expected value) of the distribution.

Mean = Σ (x \* P(x))

= (-2,0000.1) + (-1,0000.1) + (00.2) + (10000.2) + (20000.3) + (30000.1)

= -200 + -100 + 0 + 200 + 600 + 300

= 800

Now we can calculate the standard deviation using the following formula:

σ = √Σ (x - Mean)^2 \* P(x)

= √[(-2,000-800)^20.1 + (-1,000-800)^20.1 + (0-800)^20.2 + (1000-800)^20.2 + (2000-800)^20.3 + (3000-800)^20.1]

= √[(1,200^20.1) + (200^20.1) + (800^20.2) + (200^20.2) + (1,200^20.3) + (2,200^20.1)

= √(14400000.1 + 400000.1 + 6400000.2 + 400000.2 + 14400000.3 + 48400000.1)

= √(144000 + 4000 + 128000 + 8000 + 432000 + 484000)

= √(744000)

= 864

So the standard deviation of the given distribution is 864. So the risk involved in this venture is 864.