**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% |

Solution:

|  |  |
| --- | --- |
| Min | 24.14% |
| Q1 | 25.47% |
| Median | 26.71% |
| Q3 | 33.97% |
| Max | 91.36% |

|  |  |
| --- | --- |
| Mean | 33.27% |
| Standard Deviation | 16.94% |
| Variance | 287.14% |

|  |  |  |
| --- | --- | --- |
| IQR | Q3-Q1 | 8.50% |
| UL | Q3+1.5IQR | 46.72% |
| LL | Q1-1.5IQR | 12.72% |

Lower limit is 12.72% and the Upper limit is 46.72%. Any data falls outside this interval are considered as outliers. From the data, we see that 91.36% is found to be an outlier.



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: Interquartile range = Q3 – Q1 = 12 – 5 = 7 (approximately)

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

Ans: It 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: From the box plot we see that the value 25 is the outlier. If the value 25 is actually 2.5, it is within the lower limit. Therefore we don’t consider it as an outlier and hence Q1 will be 2.5.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans: Mode: Class Interval from 4-8 with highest frequency 20. Mode is 20

1. Comment on the skewness of the dataset.

Ans: It is positively skewed distribution.

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:Both the graphs show that the distribution of the data is positively skewed and the outlier for both the graphs is 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.)

Solution:

P(Wrong Number) = 1/200

Therefore, P(Not a wrong number) = 1-1/200 = 199/200

Now,

P(at least one in five attempted telephone calls reaches the wrong number)

= 1 – P(Not all5 attempted calls reaches wrong number)

= 1 – (199/200)5

= 1 – 0.975

= 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: The Most likely monetary of business venture is $2000 as it has the highest probability 0.3.

1. Is the venture likely to be successful? Explain

Ans: Yes, the venture is likely to be successful since the probability of non-negative return is greater than 0.5. That is 0.2+0.3+0.1 = 0.6.

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

|  |  |  |
| --- | --- | --- |
| x | P(x) |  |
| -2000 | 0.1 | -200 |
| -1000 | 0.1 | -100 |
| 0 | 0.2 | 0 |
| 1000 | 0.2 | 200 |
| 2000 | 0.3 | 600 |
| 3000 | 0.1 | 300 |
|  | ∑ xP(x) | 800 |

Ans:

Hence, the long-term average earning of business ventures of this kind is $800.

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

Ans: The good measure of the risk involved in a venture of this kind is standard deviation.

|  |  |  |  |
| --- | --- | --- | --- |
| x | P(x) | x P(x) | x2P(x) |
| -2000 | 0.1 | -200 | 400000 |
| -1000 | 0.1 | -100 | 100000 |
| 0 | 0.2 | 0 | 0 |
| 1000 | 0.2 | 200 | 200000 |
| 2000 | 0.3 | 600 | 1200000 |
| 3000 | 0.1 | 300 | 900000 |
|  |  | 800 | 2800000 |

Mean = ∑ xP(x) = 800.

Variance = ∑x2P(x) = 2800000

Standard deviation = = 1673.32