**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% |

Answer: I’ve provided the code in the folder.



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.
2. What can we say about the skewness of this dataset?
3. If it was found that the data point with the value 25 is actually 2.5, how would the new box-plot be affected?

Answer:

1. IQR = Q3 - Q1 = 20 - 5 = 15

Mean of the IQR will be 15/2 = 7.5

1. The box plot has positive skewness because the median is nearer to start of the plot.
2. The median and IQR will change.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?
2. Comment on the skewness of the dataset.
3. 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.

Answer:

1. The mode would be around 15.5 .
2. Skewed right. Because the tail of the dataset is on the right side.
3. By using them together, you can get a more complete picture of the data and identify important patterns and trends.
4. 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.)

Answer: We can use binomial distribution,

P(at least one call misdirected) = 1 - P(no calls misdirected)

= 1 - (199/200)^5

= 1 - 0.995025

= 0.004975

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?
2. Is the venture likely to be successful? Explain
3. What is the long-term average earning of business ventures of this kind? Explain
4. What is the good measure of the risk involved in a venture of this kind? Compute this measure

Answer:

1. The most likely monetary outcome of the business venture is 2000. This is because the probability of making a profit of 2000 is 0.3, which is higher than the probability of any other outcome.
2. Yes, the venture is likely to be successful. This is because the probability of making a profit is 0.8, which is much higher than the probability of making a loss of 0.2.
3. The long-term average earnings of this business is around 500.

The measure is called variance. I’ve calculated the variance using the code and received the output as follows,

Outcome 2.916667e+06

Probability 5.555556e-03

dtype: float64