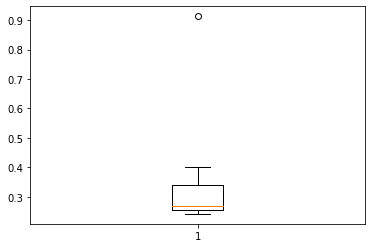
**Topics: Descriptive Statistics and Probability**

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:



Mean = 0.3327133333333333

Sd = 0.16945400921222029

Variance = 0.028714661238095233



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: Q1 = 5, Q3 = 12

IQR = Q3 – Q1

= 12 – 5

IQR = 7

The value of inter-quartile range (IQR) implies that how far is how far apart the low quartile and upper quartile is in the boxplot. 50% of the data lies between Q1 and Q3.

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

Ans: The skewness of the data is positive hence it is a right skewed data.

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: 25 is an outlier value. If it is 2.5 then there is no outlier value. The mean and the median has to be calculated to see if there is a shift in the data.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans: The data has two modes. The modes lie between 2 and 8. But they cannot be two modes in a data.

1. Comment on the skewness of the dataset.

Ans: The skewness of the data is positive hence it is right skewed data.

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 boxplot and the histogram both have outliers as 25 and both are positively skewed.

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: Probability of 1 misdirected call out of 200 calls = 1/200

Probability of at least 1 successful call = 1- 1/200

= 199/200

= 0.995

As every event is independent of the other event the probability will be

= 1 - (0.995) ^ 5

= 0.0247

= 2% chance.

5. 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: Maximum probability = 0.3 for P(x=2000). Hence, $2000 has the highest probability occurrence.

1. Is the venture likely to be successful? Explain

Ans: P (x>0) = P(x=1000) + P(x=2000) + P(x=3000) = 0.2 + 0.3 + 0.1 = 0.6

Hence, the probability of the venture being successful is 60% and the probability of risk is 20%.

(iii) What is the long-term average earning of business ventures of this kind? Explain

Ans: Average = x\*P(x) = (-2000\*0.1) + (-1000\*0.1) + (0\*0.2) +(1000\*0.2) + (2000\*0.3) + (3000\*0.1) = 800. In the long-term the average earning of the business venture is $800.

(iv) What is the good measure of the risk involved in a venture of this kind? Compute this measure

Ans: P (risk of loss) = P(x<0) = P(x=-1000) + P(x=-2000) = 0.1 + 0.1 = 0.2

So, the risked involved is 20% in this venture.