**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 : In Coding File.



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 = Lower Quartile(25% of data)

Q3 = Upper Quartile(75% of data)

Inter Quartile Range(IQR) = Q3 -Q1

= 12 – 5

= 7

This shows that 50% of the data lies between IQR.

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

Ans:

The Skewness of the Dataset is Positive.

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:

We can see there is an outlier on 25, so if value is become 2.5 then there will be no outlier. This change in value will affect the mean in a way that it reduces, and the median also reduces.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans:

The mode can lie between 4 to 8 because majority of the entry in this range.

To pin point the actual Mode we will have analyze the data.

1. Comment on the skewness of the dataset.

Ans:

The skewness of the Dataset is Positive and it is Right skewness

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:

* There is an outlier of the value 25 and both the plot has positive skewness.
* Also the histogram gives the frequency of the distribution of values of Y and the box plot provides the IQR and the Whiskers

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:

Given: One in 200 long-distance telephone calls is misdirected.

To find : probability that at least one in five attempted telephone calls reaches the

Wrong number

Solution:

One in 200 long-distance telephone calls is misdirected

Probability of call misdirecting, p = 1/200

Probability of call not Misdirecting = 1-1/200 =199/200

Number of calls = 5

Probability of having at least one successful call will be

1-p(X) = 1-1/200 = 199/200 = 0.995

Probability of at least one in five attempted calls reaches wrong number

= 5\*(1/200) = 0.025

= 2.5%

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?

The value of x with highest probability = 2000

1. Is the venture likely to be successful? Explain

Probability of the venture to fail = 0.1+0.1 = 0.2

Probability of the venture to neither fail nor be successful = 0.2

Probability of the venture to be successful = 1 – (0.2+0.2) = 0.6 = 60%

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

Expected mean = sum(x\*p(x))

= [(-2000\*0.1)+(-1000\*0.1)+(0)+(1000\*0.2)+(2000\*0.3)+(3000\*0.1)]

= 800

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

Risk involved in a venture = E(X2)-[E(X)]2 {E(X2)=X2\*P(X})

Var = 2800000 – 800^2 = 2160000

Standard Deviation = sqrt(Var) = 1470

Since the deviation from the expected value is more high the risk involved is also very high.