**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 = 0.332713333333333

Sd = 0.16945400921222

Var = 0.0287146612380952





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.

**Solution :**

IQR is the range between upper quartile (Q3) and lower quartile (Q1)

IQR= Q3-Q1= 12-5 = 7

50% of the data lies between IQR

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

**Solution :**

Skewness = 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?

**Solution :**

If the value of 25 was actually 2.5, there won't be an outlier. The mean and median must then be determined in order to determine whether the data have changed.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Solution :**

Since the majority of entries fall within this range, the mode can be anywhere between 3 and 10. We must analyse the data in order to identify the actual Mode.

1. Comment on the skewness of the dataset.

**Solution :**

Skewness = Positive

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.

**Solution :**

The plot has positive skewness and the value 25 is an outlier.

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 :**

X = Probability of 1 call Misdirected out of 200

Probability of Occurring of X = 1/200

P(X) = 1/200

Probability of having at least one successful call will be

1-P(X) = 1-1/200

= 199/200

= 0.967

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

1 - (0.967)^5

0.02475 = 2% chance.

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?

**Solution :**

$2000 as it has the highest probability of occurrence

1. Is the venture likely to be successful? Explain

**Solution :**

if Success == positive returns as a measure

Then there is a 60% probability that the venture would be successful (0.3+0.2+0.1=0.6=>0.6\*100=>60%).

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

**Solution :**

(-2000\*0.1)+(-1000\*0.1)+(0\*0.2)+(1000\*0.2)+(2000 \*0.3)+(3000\*0.1 )= 800

The long-term average earning for these type of ventures would be around $800

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

**Solution :**

A good measure to evaluate the risk would be variance and standard deviation of the variable x

Var = 3500000

Sd = 1870.83

When the average returns of $800 are taken into account, the high standard deviation value of $1870 shows that this investment is very risky.