**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:** Mean = 33.27133333333333

Variance = 287.1466123809524

Standard Deviation = 16.945400921222028



**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:** Inter Quartile Range lies between upper quartile(q3) and lower quartile(q1) [q3=12,q1=5]

So IQR=q3-q1=12-5=7.

So 50% of data lies in IQR.

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

**Ans**: The Data 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**: If data point with the value 25 is actually 2.5, then No Outliers will present. The value actually lies in between 2 and 5. That means before q1. So mean and median need to calculated to see if there is shift in the data or not.



**Answer the following three questions based on the histogram above.**

1. **Where would the mode of this dataset lie?**

**Ans**: The mode of data lies in between 3 and 10. Because when we see above histogram then we notices that the peak is range between 3-10 , so it means most of the data lies in that range only.

1. **Comment on the skewness of the dataset.**

**Ans**: Positively Skewed. Because on left side the graph is peaked and it slowly skewed to right side. As it skewed to right side. So it is called Right skewed or positively skewed.

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**: These both plots have outlier value 25 and both plots has positive right skewness.

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**: Event= 1 call misdirected happened out of 200

The probability of happening an event=1/200

Probability happening one successful call=1-1/200=199/200=0.967

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

(1-0.967)^5=0.0245= 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?**

**Ans**:2000 has more chance to occur.

1. **Is the venture likely to be successful? Explain**

**Ans**: Yes successful. Because here success rate is 60%. Here 60% is profit.20% is loss.20% no profit and no loss. [PROFIT (1000,2000,3000) =0.2+0.3+0.1=0.6(60%)]. [LOSS (-2000,-1000) =0.1+0.1=0.2(20%)]. [NO PROFIT AND LOSS (0) = 0.1(10%)]

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

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

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

**Ans**: We can calculate risk by variance and standard deviation.

Variance =3500000

Standard deviation = 1870.8286

This venture is at high risk because standard deviation is 1870.8286 and with the average return is 800.