**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: After Plotting the data we find that it has only one outlier**

**Mean ()=33.9236**

**Variance ()=302.3628**

**Standard deviation ()=17.3886**

**The outlier is 91.36**

* **Refer A2 Set-1 Descriptive Statistics and Probability.ipynb 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: Inter- Quartile Range=Q3-Q1=12-5**

**IQR=7**

**And Median=7**

**Hence, Inter-Quartile Range is same as Median value.**

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

**ANS: It is Right Skewed and it is not follows the Normal distribution.**

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: The boxplot is completely changes. The IQR and mean also changes.**



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**ANS: The mode of this dataset would lie between 4 to 8.**

1. Comment on the skewness of the dataset.

**ANS: The dataset is Right 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: The Boxplot gives Inter-Quartile Range of the dataset and the Histogram gives the**

**Range of the data values. The Histogram and Boxplot tells us that there is only one**

**outlier Present in the data. Also, the boxplot clearly shows the outliers and the**

**histogram Provides the mode of the dataset.**

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: The probability that the call is misdirected=1/200**

**The probability that the call is not misdirected=1- P(E)=1-1/200**

**=199/200**

**Probability that at least one in 5 attempted call reaches the wrong number is**

**=1-probability that no attempted call reaches the wrong number**

**=1-(199/200)5**

**=0.025**

**Therefore, the probability that at least one in 5 attempted call reaches the wrong number**

**Is 0.025.**

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: The most likely monetary outcome of the business venture=2000**

**And it has the maximum probability i.e., “0.3”.**

1. **I**s the venture likely to be successful? Explain

**ANS: The venture is successful if X is positive. Hence, if X is 1000,2000 or 3000**

**Then the probability is 0.2+0.3+0.1=0.6**

**Then 0.6>0.5 Hence, venture likely to be successful.**

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

**ANS: The long-term average earning of business ventures=E(X)**

**E(X)=∑X.P(X)=800**

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

**ANS: The Risk involved in a venture is Var(X)= E(X2)-{E(X)}2**

**=2800000-8002 =2160000**

**SD=√Var=1470**

**As Variability is high Hence, Risk is high.**

**Submitted By**

**K. KRISHNAVENI**

**(krishnavenikolipaka1000@gmail.com)**