**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: N=15

Mean=µ=33.27%

Variance=σ^2=268.0035

S.D=σ=16.37

Median=26.71%

|  |  |
| --- | --- |
| Q1 | 25.47 |
| Q3 | 33.975 |
| IQR | 8.505 |
| L Bound | 12.7125 |
| U Bound | 46.7325 |

Outlier: 91.36%





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.

Sol:

Q1=5, Q3=12 (Approx. Number taken)

IQR=Q3-Q1=12-5=7

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

Ans: Boxplot is skewed right i.e. Dataset is having Positive skewness

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: New Boxplot will be skewed left and the dataset will having negative skewness.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans: 4 to 8

1. Comment on the skewness of the dataset.

**Ans: Histogram veers to the left hence it is right skewed and its tail is having positive skew to the right**

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: **Histogram is used to find out mode, skewness, kurtosis and boxplot is used to find out IQR and outliers.**

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: 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

P(x) = ⁿCₓpˣqⁿ⁻ˣ

n = 5

p = 1/200

q = 199/200

at least one in five attempted telephone calls reaches the wrong number

= 1  -  none of the call reaches the wrong number

= 1  - P(0)

= 1   -  ⁵C₀(1/200)⁰(199/200)⁵⁻⁰

= 1  -  (199/200)⁵

= 0.02475

**probability that at least one in five attempted telephone calls reaches the wrong number = 0.02475**

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:**0.3**

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

**ANS: yes the venture likely to be successful p(x=1000)+p(x=2000)+p(x=3000)=0.2+0.3+0.1=0.6 Probability of getting positive returns is more.**

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

Ans: long term average (μ) is simply multiply each value of the random variable by its probability and add the products.

|  |  |  |  |
| --- | --- | --- | --- |
| x | P(x) | x.p(x) | (x- μ)^2.p(x) |
| -2,000 | 0.1 | (-2000)(0.1)=-200 | 784000 |
| -1,000 | 0.1 | (-1000)(0.1)=-100 | 324000 |
| 0 | 0.2 | (0)(0.2)=0 | 128000 |
| 1000 | 0.2 | (1000)(0.2)=200 | 8000 |
| 2000 | 0.3 | (2000)(0.3)=600 | 432000 |
| 3000 | 0.1 | (3000)(0.1)=300 | 484000 |
|  |  | Total | 2160000 |

So Long Term average μ =(-200)+(-100)+0+200+600+300=800

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

Ans: Standard deviation is the good measure of Risk Involved.

Standard Deviation=1469.69