**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% |

|  |  |
| --- | --- |
| Mean | 33.271333 |
| Variance | 287.146612 |
| Standard  Deviation | 16.945401 |

The Outlier is **Morgan Stanley : 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.
2. What can we say about the skewness of this dataset?
3. If it was found that the data point with the value 25 is actually 2.5, how would the new box-plot be affected?
4. **Ans :**  The Inter-quartile range of the dataset lies between Q1-Q3 is 5 – 12

It implies that most of the data lies between Q2-Q3

Range = Max-Min=12-5=7

1. **Ans :**  The data is Positively Skewed (Right Skewed), And the whisk on the right side is larger when compared to the Left.
2. **Ans :** If it was found that the data point with the value 25 is actually 2.5,the out lairs will be at lower extreme side.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Ans** : Mode of the dataset lies Between 4-8

1. Comment on the skewness of the dataset.

**Ans :** The data is Positively Skewed(Right Skewed),the tail of the distribution is on the right side.

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 and boxplot are very similar in that they both help to visualize.

Although histograms are better in determining the underlying distribution of the

data, box plots allow you to compare multiple data sets better than histograms as

they are less detailed and take up less space .It is easy to find the outliers in the

boxplot than hist.

And the would skewed right .The outliers are present at upper extreme.

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)= **nCx\*p^x\*q^(n-x)**

n = 5

p= 1/200

q= 199/200

atleast 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⁻

= 1  -  (199/200)⁵

= 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 :** As the probability (0.3) is more for 2000$ as compared to others, Therefore,

the most likely monetary outcome of the business venture = 2000$

1. Is the venture likely to be successful? Explain

**Ans :** Yes, there are 60% chances of getting a positive return and 20% chances of negative returns or debts.

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

**Ans :** Long term returns = ((-2000\*1)+ (-1000\*1)+ (1000\*2)+ (2000\*3)+

(3000\*1) / 6) =8000/6 =1333

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

**Ans :** Good measure is, Positive returns (profits) probability tends to be more than negative returns (loss). i.e. 60% probability of profits and 20% probability of loss