**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 = 33.271%= 0.3327**

**Standard deviation= 16.37%= 0.1637**

**Variance= 268.00% = 0.0268**

**Outliers= Morgan Stanley is an outlier of 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, Q2=7, Q3=12

Inter Quartile Range = Q3-Q1= 12-7=3

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

**SOL=** Right skewed or Positive 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?

**SOL=**  The boxplot would shrink as all the values becomes smaller.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**SOL**= From 4 to 8

1. Comment on the skewness of the dataset.

**SOL=** 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.

**SOL=** They both complement each other by indicating the skewness of the data

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.)

**SOL=** 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

= P(1)

=  ⁵C1(1/200)1(199/200)⁵⁻1

= 5\* (199)4/(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?

**SOL=**Business venture 1000 with P(x)=0.3 is the most likely monetary outcome

1. Is the venture likely to be successful? Explain

**SOL=** Since P(-2000)+P(-1000)+P(0)+P(1000)+P(2000)+P(3000)

**=** 0.1+0.1+0.2+0.2+0.3+0.1

**=** 1

**=** yes the business venture is going to be sucessful

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

**SOL=** Long term average earning of the business = mean(X)

**= (**-2000-1000-0+1000+2000+3000)/100

**=** 500

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

**SOL=** Measure of risk involved in this venture = var(X)

**=** 2916666.66

**=** Since the variance is very large the risk involved will be very high.