**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 : 0.3327

Variance : 0.0268

Standard Deviation : 0.163708



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.

**Answer: IQR = q3-q1 = 12-5 = 7**

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

**Answer: From the above graph its seems to be like positive or right skew .**

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?

**Answer: then upper outlier and +ve skewness will remove. Data will look like normal distributed.**



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Answer: Mode of value Y id lies at 5-8 of the data.**

1. Comment on the skewness of the dataset.

**Answer: Data is positive or right skew.**

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.

**Answer: Histogram give the frequency of value of Y for some range and box plot show the range of upper extreme lower extreme and outlier data from histogram we can’t say that from which point is outlier start but boxplot give exact value and give outlie point perfectly.**

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.)

**Answer: The probability of the event E is**

**P(E)= 1/200**

**Probability of having at least one successful call will be**

**1-P(X)= 1-1/200= 199/200= 0.967**

**As every event is independent of other event the probability will be**

**1- (0.967) ^5**

**0.02475 = 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?

**Answer: 2000**

1. Is the venture likely to be successful? Explain

**Answer: Success is positive returns as a measure**

**Then there is a 60% probability that the venture would be successful (0.3+0.2+0.1=0.6=>0.6\*100=>60%).**

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

**Answer: profit=(0.2\*1000+0.3\*2000,0.1\*3000)=1100**

**Loss-(0.1\*2000+0.1\*1000)=300**

**success=1100-300=800**

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

**Answer:**

**We can use Standard deviation or Variance to measure risk**

**Variance = $3500000**

**Standard Deviation: = $1870.829**

**Compared with standard deviation and average returns it is risky.**