**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% |



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 : Q – Q1 = 12-5=7

I looks like median value

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

Ans : mean < median, Left skew data

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 : There is no outliers on the given data and curves follows the normal distribution



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 : Right skew data, majority data left 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: I both diagram skewness can be identifies, both have outliers. In Boxplot median can be easily visualized and in Histogram mode can be easily visualized.

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.)

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**

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 : 2000 due to more probability 0.3

1. Is the venture likely to be successful? Explain

Ans : 800, Average value is 800

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

Ans : Aveareg returns = 800

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

Ans : Standard deviation is high, hence risk is high

|  |  |  |  |
| --- | --- | --- | --- |
| x | P(x) | E(X)= X . P(X) | E(X²) = X² . P(X) |
| -2,000 | 0.1 | -200 | 400,000 |
| -1,000 | 0.1 | -100 | 100,000 |
| 0 | 0.2 | 0 | - |
| 1000 | 0.2 | 200 | 200,000 |
| 2000 | 0.3 | 600 | 1,200,000 |
| 3000 | 0.1 | 300 | 900,000 |
|  |  | **800** | **2,800,000** |
| Var (X) = E(X²) - { E(X) }² | 2800000-(800\*800) | 2,160,000 |  |
| SD = √Var |  | 1,470 |  |