**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: Outliers: Morgan Stanley 91.36%

Mean = 33.271333

Standard Deviation = 16.945401

Variance = 287.1466123809524



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: Inter-Quartile Range = 7, Approximately First Quantile = 5, Third Quantile = 12, Median = 7, **IQR** = Third Quantile - First Quantile = 12 - 5 = 7, Median Value.

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

Ans: It is Right-Skewed median is towards the left side, so it is not normally distributed.

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: If the data point with the value 25 is actually 2.5 there would be no outliers on the given dataset because of the outlier the data had positive skewness it will reduce and the data will be in normal distribution.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans: The mode of the dataset approximately lie in between 4 to 8

1. Comment on the skewness of the dataset.

Ans: The data is Right-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.

Ans: Both boxplot and histogram are right-skewed and have outliers. Median is easily visible in boxplot and mode is visible in histogram

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:

200 long-distance telephone calls is misdirected.

To find the probability that atleast one in five attempted telephone calls reaches wrong number

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 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 |

|  |  |  |  |
| --- | --- | --- | --- |
| X | P(X) | E(X)=X.P(X) | E(X^2)=X^2.P(X) |
| -2000 | 0.1 | -200 | 400000 |
| -1000 | 0.1 | -100 | 100000 |
| 0 | 0.2 | 0 | 0 |
| 1000 | 0.2 | 200 | 200000 |
| 2000 | 0.3 | 600 | 1200000 |
| 3000 | 0.1 | 300 | 900000 |
|  |  | 800 | 2800000 |

1. What is the most likely monetary outcome of the business venture?

Ans: The most likely monetary outcome of the business venture is 2000 as it has maximum probability is 0.3

1. Is the venture likely to be successful? Explain

Ans: Yes, the probability that the venture will make more than 0 or a profit

if X is 1000, 2000, 3000

The probability is 0.2+0.3+0.1=0.6

0.6 > 0.5 this states that 60% chances for this venture to be making a profit.

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

Ans: The long-term average is Expected value E(X) = X.P(X) = 800, which means on an average the returns will be +800

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

Ans: Risk involved in venture depends on the Variability in the distribution.

Higher Variance means more chances of risk Var(X) = E(X^2)-(E(X))^2 = 2800000 - 800^2 = 2160000