**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% |

Solution: -

1. Mean, µ = 24.23+25.53+25.41+24.14+29.62+28.25+25.81+24.39+40.26+32.95+91.36+25.99+39.42+26.71+35.00/15

|  |
| --- |
| Mean, µ =33.2713 |

1. Variance, σ^2 = ∑(X-Xbar)^2/n

=(24.23-33.27)^2+(25.53-33.27)^2+(25.41-33.27)^2+(24.14-33.27)^2+(29.62-33.27)^2+(28.25-33.27)^2+(25.81-33.27)^2+(24.39-33.27)^2+(40.26-33.27)^2+(32.95-33.27)^2+(91.36-33.27)^2+(25.99-33.27)^2+(39.42-33.27)^2(26.41-33.27)^2+(35.00-33.27)^2 / 15

|  |
| --- |
| Variance, σ^2 = 268.0035 |

1. Standard deviation, σ =

σ

|  |
| --- |
| Variance, σ = 16.3708 |

1. 

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?

Solution: -

2. Min value = 0
3. Max value = 25

Q1: - 1st Quantile = 5

Q2: - 2nd Quantile (or) median = 7

Q3: - 3rd Quantile = 13

Inter quantile range IQR = 13-5 = 8

1. Right skewed

I take that the positively skewed, because left quantile is the further from the median.

1. a. Above plots shows two points 20 and 25 and point 25 is the outliers as they are not included in the box of other normal observation.

b. If box plot was found that the data point with the value 25 is actually 2.5, then it will lie on point ‘0’ and point ‘5’ so the box plot distribution nothing will be change. It always positively (or) right skewed.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?
2. Comment on the skewness of the dataset.
3. 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.

Solution: -

1. Positive skewness, skewed to the right, 4 to 8 is the mode of the dataset.
2. Positive skewness, skewed to the right.
3. We can’t say diff mode in box plot but we can do that in hist.
4. 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.)

Solution: -

One in 200 long-distance telephone calls is misdirected.

Probability of calls misdirecting P = 1/200.

= 1-1/200 = 199/200

Number of calls = 5

P(X) = nCxP^xQ^n-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

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?
2. Is the venture likely to be successful? Explain
3. What is the long-term average earning of business ventures of this kind? Explain
4. What is the good measure of the risk involved in a venture of this kind? Compute this measure

Solution: -

1. The most likely monetary outcome of the business venture: X = 2000 with the highest probability of 0.3
2. The venture is likely to be successful, because (X = 1000) + (X = 2000) + P(X = 3000)

= 0.1+0.20.3

= 0.6

1. IS Standard deviation

Sd

(1) 0.08164966

1. The long-term average earning business kind is probability of 0.3 because, X = 2000 and highest probability in the list is 0.3 is equal to X = 2000 and P(X) = 0.3
2. The good measure of the risk involved in a venture is probability of 0.1 because,

X = -1000 and that -1000 probability is also 0.1 not in ± values like -2000 has 0.1 and 2000 has 0.3, but it was X = -1000 and P(X) = 0.1