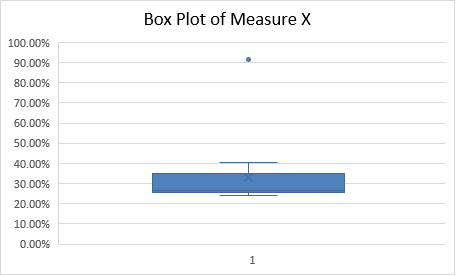
**Topics: Descriptive Statistics and Probability**

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: -



Outlier of this following box plot is : Morgan Stanley - 91.36%

Measure X describe: -

Mean :- 33.27

Standard Deviation: - 16.945

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.

**Answer: -** Approx. (1st Quartile range) Q1 = 5

(3rd Quartile Range) Q3= 12

(2nd Quartile Range) Q2 = 7

Inter-Quartile Range) IQR= Q3-Q1= 17-5=7

Second Quartile Range is Median Value

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

**Answer: -**Right skewed median is towards the left side it is not normal distribution

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: -** If it was found that the data point with the value 25 is actually 2.5, then in given

Dataset there would be no outliers and positive skewness will reduce and data become

normally distributed.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Answer: -** Most of the data lies in between approximately 4 to 10, So mode of the lies according to

the given plot is 5 to 10.

1. Comment on the skewness of the dataset.

**Answer: -** Right Skewed or Positive skewness (Mean>median>Mode)

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: -** If above histogram and the box-plot in question 2 are plotted for the same dataset. Then

both are right skewed and both have outliers the median can be easily visualized in box plot

where as in histogram mode is more visible.

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: -**  **If** 1 in 200 long-distance telephone calls are getting misdirected.

probability of call misdirecting   = 1/200

Probability of call not Misdirecting = 1-1/200 = 199/200

**The** probability for at least one in five attempted telephone calls reaches the wrong number

Number of Calls = 5

n = 5

p = 1/200

q = 199/200

P(x) = at least one in five attempted telephone calls reaches the wrong number

P(x) = ⁿCₓ pˣ qⁿ⁻ˣ

P(x) = (nCx) (p^x) (q^n-x) # nCr = n! / r! \* (n - r)!

P(1) = (5C1) (1/200)^1 (199/200)^5-1

P(1) = 0.0245037

Therefore, the probability that at least one in five attempted telephone calls reaches the wrong number is approximately 0.0248, or about 2.48%

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: -**  From the table we can see that the probability of the x=2000 is the highest with the

probability of 0.3.

This means 30% chance that the monetary outcomes of the business venture will be

$2000.

1. Is the venture likely to be successful? Explain

**Answer: -**  Yes, the probability that the venture will make more than 0 or a profit

p(x>1000) = 0.2

p(x>2000) = 0.3

p(x=3000) = 0.1

Total = 0.6

This means it have good percentage of 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

**Answer: -**

For long term average earning of business ventures of this kind. We need to calculate expected value

The formula for calculating the expected value is :

E(X)= ∑[(X\*P(X)]

Where, X= possible return and P(X)= possibility return occurring

E(X)= (-2,000 \* 0.1) + (-1,000 \* 0.1) + (0 \* 0.2) + (1,000 \* 0.2) + (2,000 \* 0.3) + (3,000 \* 0.1)

= 200-100+0+200+600+300

= 800

Which means on average return will be 800$, but this one amount is not accurate might be it will differ from actual amount

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

**Answer:** A good measure to evaluate the risk involved in a venture would be the variance and standard deviation of the variable

Var = 3500000

Sd = 1870.83

The large value of standard deviation of $1870 is considered along with the average returns of $800 indicates that this venture is highly risky