**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:-

Morgan Stanley = 91.36% this is outlier

mu = 33.27% , vairence= 287.1466 , SD= 16.9454



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?

Ans:-

1. Approximately (First Quantile Range) Q1= 5 (Third Quantile Range) Q3 = 12,

Median (Second Quartile Range) = 7 (Inter Quartile Range) IQR = Q3-Q1 = 12-5=**7** Second Quartile Range is the Median Value.

1. Highly Skeweed towards positive or Right Tail Skeweed and outlier present.
2. In that case 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 normal distributed.



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.

Ans:-

1. The mode of this data set lie in between 5 to 10 and approximately between 4 to 8.
2. Right-Skeweed.Mean>Median>Mode
3. They 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.
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.)

Ans:-

Given one in 200 long-distance telephone calls is misdirected.

To find : probability that at least one in five attempted telephone calls reaches the wrong number.

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) = "Cxpxqn-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(O)

=1 - 5Co(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

Ans:-

1. Most likely monetary outcome of the business venture is $ 2000 as it has maximum Probability 0.3.
2. Venture is successful if X is +ve,

Hence if X is 1000, 2000 or 3000

Probability is 0.2+0.3+0.1 0.6 as 0.6 > 0.5

Hence venture likely to be **successful.**

1. Long-term average earning of business ventures = E(X)

E(X) = Sum(X.P(X)) = $800

1. Risk involved in a venture

Var (X) = E(X^2) (E(X))^2

= 2800000 - 800²

=2160000 (Quite High)

SD = √Var = $ 1470

**As Variability is Quite high hence Risk is high.**