**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:

Population mean of the data () = **33.27%.**

Population Variance of the data () = **2.87%**

Population Standard deviation of the data () = **16.94%.**



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.

Inter Quartile Range = Q3-Q1 = 12-5 = 7

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

In the above data,More number of observations presented above the median.

So, **data is left skewed and will have a negative value.**

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?

In the above data, the observation “25” is lying out side the value of upper limit, so it is a outlier . As box plot is used to visualize the presence of outliers in data, the revised value of “25” as “2.5” will slightly **move the median value upwards.**



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie? 5 to 10
2. Comment on the skewness of the dataset. Left Skeewed
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.

From the above Boxplot and Histogram we can say that

Median will be 7 (approximately)

25 is Considered as Outlier

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 one in the 200 calls reaches wrong number, probability of call reaching wrong number is 1/200.

And probability of call not reaching wrong number is 1- 1/200 = 199/200.

S

probability of atleast one of five calls reach wrong number =

1 - probability of none of the calls reach wrong number.

P(≥1) = 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?

Most likely outcome or Expected value of this business venture is,

*EV*=∑*P*(*Xi*​)×*Xi*

*So,E(x)*= **800**

1. Is the venture likely to be successful? Explain

yes

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

Answer:

Long term average earnings or mean of the returns is () = **500.**

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

Answer:

Good measure for finding risk involved in this venture is, standard deviation().

Because it tell us about volatility of the data.

Standard deviation () of the data is  **1870.82**