**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:**

Mean(𝜇)= 33.27133333333333

Standard Deviation(𝜎)= 16.945400921222028

Variance(𝜎2)= 287.1466123809524

Attached python file with name basic statistics level 2



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:**

(i)- Inter Quartile Range (IQR) = Q3-Q1 = 12-5 = 7(IQR), Mean>Median

(ii)-This is Positively Right Skewed Data. Where mean>Median

(iii)-There would be no Outliers in the given dataset. And the data is Normally 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:**

(i)-The mode of this dataset lies in between 4 to 8.

(ii)-Positively right Skewed Data.

(iii)- 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.

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:**

Probability of call being misdirected is p = 1/200 = 0.005

Number of attempted phone calls is n = 5

probability that at least one in five attempted telephone calls reaches the wrong number

= 0.005\*5 = 0.025

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 :**

(i) The most likely monetary outcome of the business venture: 𝑥 = 2,000 with probability of 0.3

(ii) The venture is likely to be successful, because P(𝑥 = 1,000) + P(𝑥 = 2,000) + 𝑃(𝑥 = 3,000) = 0.2+0.3+0.1 = 0.6

(iii) (0.1) (−2,000) + (0.1) (−1,000) + (0.2) (0) + (0.2) (1,000) + (0.3) (1,000) + (0.1) (3,000) = 800

(Iv) The good measure of the risk involved in a venture of this kind 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

Std = sqrt (2160000) = 1469.69