**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 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).** IQR = Q3 – Q1 = 12 – 5 = 7

The inter-quartile range of the data implies the amount of spread of the data, so the middle 50% data is spread over 7 units.

**Ans (ii).** The dataset is positively-skewed because the median (Q2) is near to the 1st quartile which is shorter.

**Ans (iii).** The minimum and maximum would remain unchanged however, Q1 and Q3 would likely get shifted which would cause a right shift to the box.



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 would be closer to the lower values in x axis to the left of median. So, Mode = 5

**Ans (ii).** The skewness is positively skewed since the longest bar is towards the lower values.

**Ans (iii).** The histogram will talk about the median, mean and mode of the same data as plotted in the box plot and will also provide information of the data beyond the box. The box plot will tell that under what range 50% of the dataset values are occurring.

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 that a single call is misdirected = 1/200

So, probability that not a single call is misdirected = 1 – 1/200 = 199/200

Now, probability that none of the 5 directed calls were misdirected = (199/200)^5

Thus, the probability that at least one of the 5 directed calls were misdirected = 1 – (199/200)^5 = **0.024**

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 will be the one having highest probability i.e., $2000.

Ans (ii). If we add the probabilities of the positive returns,

So, 0.2 + 0.2 + 0.3 + 0.1 = 0.8

0.8 is greater than sum of probabilities of negative returns. So, venture is likely to be successful.

Ans (iii). CHECK PYTHON NOTEBOOK

Ans (iv). CHECK PYTHON NOTEBOOK