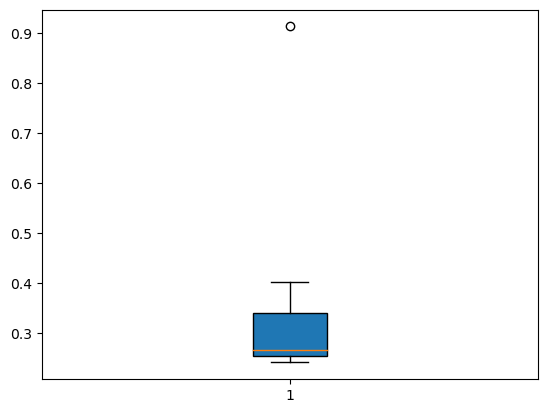
**Topics: Descriptive Statistics and Probability**

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:



**Outliers = 91.36%**

= **0.332713**

= **0.169454** ^2= **0.02871**



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.

**Ans: The IQR for the data is about 7.5**

**The IQR is nothing but the Quartile Deviation it gives information about spread of the middle data.**

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

**Ans: The data has right or positively skewed.**

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?

**Ans: The new boxplot will not have any outlier i.e. data has within box or it has approximate normal.**



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Ans: This is bimodal and have the mode about 5 -7**

1. Comment on the skewness of the dataset.

**Ans: The distribution has longest right tail hence it has skewness at right side, It is positive skewed**

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.

**Ans: Histograms give a good sense of the distribution of a variable. Box plots attempt to do the same thing however, don't give as good of a picture of the distribution of this variable.**

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:** **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. It is a binomial distribution problem.**

**We have to find prob of at least 1 in 5 five telephone calls reaches wrong ,**

**P(x) = ⁿCₓ pˣ qⁿ⁻ˣ P(x)**

**= 0.024751**

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?

**Ans: The most likely monetary outcome of the business venture is 2000$, as the probability for 2000$ is the most compared to other.**

1. Is the venture likely to be successful? Explain

**Ans: Yes, the probability of venture make 0 or more profit is 0.8 i.e., 80% chances that the venture will be successful.**

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

**Ans: The long-term average earning of business ventures is the expected value = sum(x\*p(x)) = 800$**

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

**Ans: The good measure of the risk involved in a venture is depends on the variability of the distribution, the more the variance the more the risk. Here, Var (X) = E(X^2) –(E(X)) ^2 = 2800000 – 800^2 = 2160000. The standard deviation is 1469.6938 $**