**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.

ANS:First quantile range (Q1)=5,Third quantile range(Q3)=12

**Inter quantile range IQR=Q1-Q3=7**

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

ANS: It is not normal distribution Has median at the left side, **it is Right Skewed** **(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:Then there will be no outliers in the given dataset because of the outlier 25 data**

**is positively skewed there for data becomes normally distributed data**

\



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**ANS: Mode of the data can be lies between 5 to 10 or approximately in between 4 to 8**

1. Comment on the skewness of the dataset.

**Ans: mean>median>mode, data distribution is positively 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: Both of them are rightly skewed,and have same outlier median can be easily visualized.**

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 out of 200 long distance calls are getting misdirected probability of call misdirecting = 1/200**

**Probability of call not misdirecting = 1-1/200=199/200**

**p(x) = (nCx) (p^x) (q^n-x)**

**Number of calls(n)=5**

**P=1/200**

**q=199/200**

**(5C1) (1/200)^1 (199/200)^5-1 P(1)**

**=0.0245037**

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: most monetary outcome of the business venture is 2000$ because the probability of the 2000$ is 0.3 maximum as compare to others.**

1. Is the venture likely to be successful? Explain

**Ans: Yes, The probability of the venture will make more than 0 or a profit.**

**0.2+0.2+0.3+0.1=0.8**

**This state show that there is good 80% chances for this venture to make profit.**

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

**Ans: The long-term average is Expected value = Sum (X \* P(X)) = 800$ which means on average the returns will be + 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 of this kind depends on the Variability in the distribution. Higher Variance means more chances o risk Var (X) = E(X^2) –(E(X))^2 = 2800000 –800^2 = 2160000**