**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% |

plt.boxplot(df[‘Measure\_X’],vert = False)



df.describe()

0.3327133

df.var()

0.02871466

df.std()

0.169454

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.

IQR= Q3-Q1 =12-5=7\*1.5=10.5

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

It skewed towards right side and it is positive skewness

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?

if the value is 2.5 instead of 25 then the median will change it will affect the boxplot .

3.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans) the mode would be lie between 4 to 12 because most of the data fall between those value.

1. Comment on the skewness of the dataset.

Ans) it is +ve skewness and skewed towards right side.

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 values of y is 25 and both skewed towards right side

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)

Let us consider for 1 call p(A)=1/200

1-p(A)=1-1/200 = 199/200 =0.995

Similarly now assume for 5 call =1-(0.995)\*5

=0.2475 = 2% chance

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?

The most likely monetary outcome of the business venture is 2000 because it has the highest probability p(X)

1. Is the venture likely to be successful? Explain

It may successful in many way 0.1\*0.2\*0.3=0.6=60% it is successful

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

(-2000\*0.1)+(-1000\*0.1)+(0\*0.2)+(1000\*0.2)+(2000\*0.3)+(3000\*0.1)=800

Long term earning Is $800

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

df[“x”].std()

1870.829

df[“x”].var()

3500000