**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: ---🡪7.5(approx.…)

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

Ans: ----🡪 Right 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: ------🡪There will be no outliers and mean of the data will be reduced and data will become more normal



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans: ----🡪Approx (5 to 7)

1. Comment on the skewness of the dataset.

Ans………>Right skewness

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-------🡪By looking at both (Histogram ang boxplot) we can say that data is right skewed and 25 is the Outlier and also majority of data lies between 5\_10

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 misdirecting P=1/200

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

Number of calls=5

P(x)=nCxPXQn-x

N=5

P=1/200

Q=199/200

At least one in five attempted telephone call reaches the Wrong number

=1-none of the call reaches the wrong number

=1-P(0)

=1-5c0(1/200)0(199/200)5-0

=1-(199/200)5

=0.02475

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: ---🡪2000rs as it has highest probability

1. Is the venture likely to be successful? Explain

Ans: ----🡪Venture is successful if X is +ve

Hence if X is 1000,2000, or 3000

Probability is 0.2+0.3+0.1=0.6

As 0.6> 0.5 Hence venture likely to be successful

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

Expected value

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

800.0

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

((-2000\*\*2) \*0.1) +((1000\*\*2) \*0.1) +((0\*\*2) \*0.2) +((1000\*\*2) \*0.2) +((2000\*\*2) \*0.3) +((3000\*\*2) \*0.1)

Variance=116000.0

Standard deviation: --1077

Variability is Quite high hence Risk is high