**Topics: Descriptive Statistics and Probability**

* 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: Code is attached to the File Assignment-2(1)*



Answer the following three questions based on the box-plot above.

* What is inter-quartile range of this dataset? (please approximate the numbers) In one line, explain what this value implies.

*Ans: Approximately Q1 = 5, Q3 = 12, Median = 7 , IQR = q3-q1 = 7.*

* What can we say about the skewness of this dataset?

*Ans : The data is left Skewed, and Normal Distribution.*

* 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 present in the data set due to positive skewness*



Answer the following three questions based on the histogram above.

* Where would the mode of this dataset lie?

*Ans: Approximately around 4-8*

* Comment on the skewness of the dataset.

*Ans: Right skewed data, as mean>median*

* 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: We can see outliers are present in the data set, mean > median ( RIGHT-SKEWED) ,*

*mode lies in b/w 4-8*

* 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: call misdirected = 1/200*

*call not misdirected = 1-1/200 = 199/200*

*no.of attempts (n) = 5*

*no.of successful misdirected attempts (q) = 1/200*

*P(x) = ⁿCₓ pˣ qⁿ⁻ˣ*

*ⁿCₓ= n!/x!\*(n-x)!*

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

*P(1) = 0.0245037*

* 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 |

* What is the most likely monetary outcome of the business venture?

*Ans: Outcome would be around $2000 as it has the highest P(x)=0.3*

* Is the venture likely to be successful? Explain

*Ans: yes, The venture P(x) > 0 are 0.2+0.2+0.3+0.1 = 0.8 .*

*it has 80% probability to be success*

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

*Ans: Expected value = Sum(x\*p(x)) = $800*

*so Avg returns are $800*

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

*Ans: higher variance means higher risk*

*Var(x) = p(x^2)-(p(x))^2 = 2800000 - 800^2 = 2160000*