**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: x.mean()=33.27133333333333

x.var()=287.1466123809524

x.std()=16.945400921222028

outlier in boxplot : Morgan Stanley 91.36%

2)



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 the numbers Q1 = 5 Q3 = 12,

Median = 7

IQR = Q3 – Q1 = 12 – 5 = 7

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

Ans: Right/positive-Skewed median is towards the left side it is not 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: In that case there would be no Outliers on the given dataset because of the outlier the data had Right/positive skewness it will reduce and the data will normal distributed.



Answer the following three questions based on the histogram above.

* Where would the mode of this dataset lie?

Ans: The mode of this data set lie in between 5 to 10 and approximately between 4 to 8 .

* Comment on the skewness of the dataset.

Ans: Right-Skewed. Mean>Median>Mode

* 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: They both are right/positive -skewed and both have outliers median can be easily visualized in box plot as well as in histogram.

3) 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: Given – 1 in 200 long-distance telephone calls is misdirected then probability of call misdirecting is 1/200 so 1-1/200 ie 199/200 is Probability of not Misdirecting calls.

Then Probability (P(x)) of 5 attempted telephone calls reaches the wrong number.

P(x) = ⁿCₓ pˣ qⁿ⁻ˣ P(x) = (nCx) (p^x) (q^n-x)

# nCr = (5C1) (1/200)^1 (199/200)^5-1 P(1) = 0.0245037

5)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 |

Ans:

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

Ans: he most likely monetary outcome of the business venture is 2000$ which has maximum probability is 0.3 as compared to others.

* Is the venture likely to be successful? Explain

Ans: Yes, the probability that the venture will make more than 0 or a profit p(x>0)+p(x>1000)+p(x>2000)+p(x=3000)

0.2+0.2+0.3+0.1 = 0.8 t

his states that there is a 80% chances of making a profit.

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

which means on an average the returns will be + 800$

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

Ans: This depends on the Variability in the distribution. Higher Variance means more chances of risk. Var (X) = E(X^2) –(E(X))^2

=2800000 – 800^2 = 2160000