**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- Mean= 0.33271

Variance= 0.028715

Std = 0.169454

Outliers = Above 46.74% & below 12.72%

So here 91.36% or Morgan Stanley is the outliers.

|  |
| --- |
|  |
|  |



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.
2. What can we say about the skewness of this dataset?
3. 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- \* The inter- quartile range of this dataset is 7.5 approx. It varies 5 to 12.5 range and has median

Value is 7 approx.

\* The skewness of this dataset is positively skewed.

\* If it was found that the data point with the value 25 is actually 2.5 the outlier from the

boxplot will be removed .

it will reduce the right skewness of the data.





Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?
2. Comment on the skewness of the dataset.
3. 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- \* The mode of this dataset lies within 4 to 10

\* The dataset has positive skewness.

\* Both the boxplot and histogram show the median value is approximately 7 , it is positively skewed and 25 is the outlier of the dataset.

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 misdirected =1/200

Probability of call not misdirected = 1-(1/200) = 199/200

No of calls = 5

P(x) = nCx (P^x) (q^n-x)

At least one in five attempted calls reaches the wrong number= 1 - P(0)

= 1- 5C0(1/200)^0 (199/200)^5-0

= 1 – (199/200)^5

= 0.0248

= 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?
2. Is the venture likely to be successful? Explain
3. What is the long-term average earning of business ventures of this kind? Explain
4. What is the good measure of the risk involved in a venture of this kind? Compute this measure

ANS- 1) The most likely monetary outcome of the business venture is x=2000 with the highest

Probability of 0.3 .

2) If the venture can maintain for long term business then eventually it will be successful since

the probability of non-negative return is higher than 0.50 and the expected value for return is

a positive number . P(1000+2000+3000)=0.2+0.3+0.1=0.6 = 60%

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

4) The good measures of the risk involved in a venture of this kind is standard deviation.

Std = 1870.829

The large value of std of 1870 is considered along with the average returns of 800 indicates that this venture is highly risky. .