**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% |

Mean – 33.27%

Variance – 0.028715

Standard Deviation - 0.169454

Q 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.

Inter quartile range is the range between upper quartile Q3 and the lower quartile Q1

IQR= Q3-Q1

= 12-5= 7

50% of the data lies between IQR

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

The above distribution we can said that the data is right skewed or positive 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?

1) If it was found that the data point is 2.5 instead of 25 then the outlier from the boxplot

will be removed

2) The right skewness of the data will also be get reduced.

2.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

1) This is the bimodal distribution It has TWO modes

2) The mode will be found between 4 to 10 where most of the data is lies

But we can not say that always the mode can be lie on the high peak its just the assumption

For actual value we need to do calculations

1. Comment on the skewness of the dataset.

It is right skewed +ve skewed data

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.

From the above histogram and bar plot we can say that the outlier is 25 also in both the

plots , data is positive skewed.

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.)

Let consider 1 call misdirected out of 200 call Is event A

Probability of occurring event A is = 1/200

P(A)= 1/200

Probability of having 1 successful call will be

1-P(A) = 1-1/200 = 0.995

As every event Is independent of other event the probability will be

1-(0.995)^5 = 2.4 % 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 is $2000 because it is having highest probability of occurrence.

1. Is the venture likely to be successful? Explain

By provided data we can say that the venture is having positive monetary outcome as here

The probability of the monetary outcome is given by (x) which indicates that

0.2+0.3+0.1 = 0.6=> 0.6\*100 =60%

so the venture will have the probability of 60% of being successful .

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

To calculate the expected/average business venture of the same kind

We can say that the expected value or earning is average of

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

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

Risk can be define by possible variability in expected returns . Therefore to calculate the good measure for the risk of venture like this would be Standard deviation and variance of variable X

Here the standard deviation is 1870.82 and the variance is 3500000.0

So for the standard deviation of the 1870.82 with average return of the 800$ the venture will be risky.