**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

x.mean()=33.27133333333333

# Vairance

x.var()=287.1466123809524

# Standard Deviation

x.std()=16.945400921222028



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.

**Answer= IQR = 12-5 =7, this represents the range between upper quartile(Q3)&lower quartile(Q1) which contains 50% of the data points**

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

**Ans= Right skewed,positively 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= 2.5 will be not considered an outlier .The boxplot will start from 0 and send at 20 in representation.**



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**ANS=Mode can lie between 3&10 because majority of the entry in this range.To pin point the actual mode we will have analyze the data.**

1. Comment on the skewness of the dataset.

**ANS=positively skewed.**

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=There is an outliers of the value 25& both plots are positively skewed.median in boxplot and mode in Histogram**

**Histogram provides the frequency distribution so we can see how many time each data point is occurring however, boxplot provides the quintile distribution i.e.50%data lies between** %5

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= **X=probability of 1 call misdirected out of 200**

**Probability of occouring of X=1/200**

**P(X)=1/200**

**Probability of having at least one successful call will be**

**1-P(X)=1-1/200=199/200=0.967**

**As every event is independent of other event the probability will be**

**1-(0.967)^5**

**0.02475=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?

**ANS**= **Max,p=0.3 $2000 because it has high probability occurrence.**

1. Is the venture likely to be successful? Explain

**ANS=(P(x>0)=0.6,implies there is a 60%chance that the venture would yield profits or greater than expected returns .P is only 0.2 so the venture is likely to be successful.**

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

**ANS= Weighted average= x\*P(x)**

**=(-2000\*0.1)+(-1000\*0.1)+(0\*0.2)+(1000\*0.2)+(2000\*03)+(3000\*0.1)**

**=800**

**The long term average earning of this kind if business ventures is 800**

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

**ANS=A good measures to evaluate the risk would be variance and standard deviation of variable X**

**Var=3500000**

**Sd = 1870.83**

**The large value of standard deviation is 1870.83 is consider with Avarge earning 800,**

**It indicates that this venture is highly risk**