**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% |



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.

ANS: From the above boxplot we can identify that Q1=5, Q3=12

We know that ( Inter Quartile Range=> IQR=Q3-Q1 )

IQR=>12-3=7, here median of above data is 7

Therefore both median and IQR are same in above dataset.

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

ANS: Above boxplot indicates right skewed, As most of the data is left biased so,the median is on left most side of the box plot.And it is not going to follow Normal distrubution.The skewness value will be positive(+).

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: If the value 25 is replaced with 2.5 there will be no outliers and the positive skewness will be reduced then the data will be Normally distributed



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

ANS: The mode of this dataset lie between 4 and 8 of ‘y’ values.

1. Comment on the skewness of the dataset.

ANS: The data is right skewed.And most of the data is left biased.

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: Explicitly the boxplot shows median, histogram shows the mode, And also historgram graphs the exact outliers of the dataset.Coming to skewness both the graphs are right skewed and left biased data,both indicated the distribution is Abnormal.

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: Let us consider the probavility of 1 call misdirected out of 200 as event A.

Probability of occuring of event A=1/200

P(A)=1/200

Probability of having at least one successful call will be 1-P(A)=1-1/200=0.967

As every event is independent of other event the probability will be 1-(0.767)^5

0.0247=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 |

|  |  |  |  |
| --- | --- | --- | --- |
| X | P(X) | E(X)= X . P(X) | E(X²) = X² . P(X) |
| -2000 | 0.1 | -200 | 400000 |
| -1000 | 0.1 | -100 | 100000 |
| 0 | 0.2 | 0 | 0 |
| 1000 | 0.2 | 200 | 200000 |
| 2000 | 0.3 | 600 | 1200000 |
| 3000 | 0.1 | 300 | 900000 |
|  |  | 800 | 2800000 |

1. What is the most likely monetary outcome of the business venture?

ANS: Most likely monetary outcome of the business venture is $2000 as it has maximum probability is 0.3

1. Is the venture likely to be successful? Explain

ANS: Venture is successful if X is + ve, Hence if X is 1000 , 2000 or 3000 Probability is  0.2+0.3+0.1 = 0.6 as 0.6 > 0.5 Hence venture likely to be successful.

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

ANS: The long term average earning of bussiness ventures is expected value=

E(X)= ∑ x.P(x).

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

ANS: Risk involved in a venture Var (X) = E(X²)  - { E(X) }²=   2800000 -   800²= 2160000  ( Quite High)

SD = √Var  ≈ $14700

As **Variability is Quite high**  hence **Risk is high.**