**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.

ANSWER: IQR = Q3-Q1

Here Q3 = 12(approx.) Q1 =5

= 12-5 = 7

50% of data lies between Q1 and Q3.

So,Q2 quartile present at point 7. It is the median value.

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

ANSWER: it is positively skewed. Most data points lies between Q2 and q4.so,it is skewed towards right. Median>mean.median is towards left.

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?

ANSWER: 25 is an outlier. if data point with the value 25 is actually 2.5,then it will be falling under first quartile Q1.therefore no outliers present in this data.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

ANSWER: mode of the data set lies in middle between 5- 10 ,approximately between 4-8.where data is peaked. distribution skewed towards right. So ,it is positively skewed.

1. Comment on the skewness of the dataset.

ANSWER: it is positively skewed .since distribution is skewed towards right

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.

ANSWER : it is possible that these plots(histogram ,boxplot ) are of same data set .both plots shows that data is right skewed .both are positively skewed. data point 25 is far from the remaining data points in histogram which is related to outlier of boxplot which is again 25.median can be easily visualized using boxplot.whereas mode can be easily visualized using histogram.

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.)

ANSWER: if 1 in 200 calls is misdirected,the probability will be = 1/200=199/200

Probability of cars not misdirecting = 1-1/200

the probability that at least one in five attempted telephone calls reaches

the wrong number=5

n=5 , p=1/200=0.005 , q = 199/200=0.995

p(x)= at least one in five attempted telephone calls reaches the wrong number

= 0.0002475093624999536

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?

ANSWER : the most likely monetary outcome of the business venture will be 2000 beacause it has maximum probability compared to remaining.

1. Is the venture likely to be successful? Explain

ANSWER : p(x) =0.2+0.2+0.3+0.1 = 0.8

Since the probability is not zero the venture likely to be successful by 80% profit.

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

ANSWER : sum(x\*p(x))= [-2000\*0.1 + -1000\*0.1 + 0\*0.2 +1000\*0.2 + 2000\*0.3+3000\*0.1 = 800$

800$ is the long-term average earning of business ventures of this kind.

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

ANSWER : it depends on variance of the distribution.more variance means more risk involved.

Var(x) = E(x\*\*2) – (P(x))\*\*2 = 2800000-800\*\*2 = 2160000

Here 2160000 is variance so, it is high .so, the risk involved is high.