**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: Data above Morgan Stanley 91.36 % is outlier

Mean= = 33.27

Std dev = 16.94

Var = = 287.15



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: Q1 = 5, Q3=12, IQR = Q3-Q1 = 7

IQR is difference between upper and lower data, which implies that data is spread more in that range or IQR is measure of spread of data.

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

Ans: skewness is on the right side. Median is on left so its not normally distributed.

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: The box plot might not have an outlier and IQR value will be changed. Median will be the middle and curve will be almost normally distributed



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans: The Mode would be between 4 to 8

(ii) Comment on the skewness of the dataset.

Ans: The skewness is on the 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.

Ans: Both are right skewed. And in histogram modes can be visualized . and in boxplot outliers and medians are visualized well

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 calls get misdirected = 1/200

probability of calls not get misdirected =1- (1/200) = 199/200

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: will be 0.3 which is 2000 , as the probability is higher than others

1. Is the venture likely to be successful? Explain
2. Ans: yes, the probability of profit will be , p(x>0)+p(x>1000)+p(x>2000)+p(x=3000) = 0.2+0.2+0.3+0.1 = 0.8. 80% chances that the venture will make profit.
3. What is the long-term average earning of business ventures of this kind? Explain

|  |  |  |
| --- | --- | --- |
| x | P(x) | Sum(x\*P(x)) |
| -2,000 | 0.1 | -200 |
| -1,000 | 0.1 | -100 |
| 0 | 0.2 | 0 |
| 1000 | 0.2 | 200 |
| 2000 | 0.3 | 600 |
| 3000 | 0.1 | 300 |
|  |  | 800 |

Ans: Expected value = Sum (X \* P(X)) = 800$

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

Ans: Good measure is depend on the variance of the venture , Higher Variance means more chances of risk Var (X) = E(X^2) –(E(X))^2 = 2800000 – 800^2 = 2160000