**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: From Boxplot-------> Margan Stanely = 91.36% is outlier

Mean = 33.27

Std = 16.94

Var = 287.14



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.
2. What can we say about the skewness of this dataset?
3. If it was found that the data point with the value 25 is actually 2.5, how would the new

Ans. (I) Inter quartile range (IQR) = 12-5 =7

The median is at 7.

50% of the datapoints lie between 5 to 12.

(ii) This data is right skewed i.e outliers lies on the right tail.

(iii) If the data point with value of 25 is actually 2.5, then the distribution will have no

outliers.box- plot be affected?



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?
2. Comment on the skewness of the dataset.
3. 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 (i) Mode lies at 5.5

(ii) The data is right skewed with large outliers on the right tail. Most of the data is

concentrated on the left side of the mean

(iii)Both graphs give an idea about distribution of the data, whether data is normally

distributed or not, skewness of data. Both graphs indicate about the outliers.

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.)

Sol. Let E be the event that a call is misdirected.

P(E) = 1/200

Then, the probability that the call is not misdirected = 1-1/200 = 199/200

No. of calls = 5

P(X) = n Cx \* px \* qn-x

n = 5

p = 1/200

q = 199/200

At least 1 in 5 calls reaches wrong number = 1 – none of the calls reaches wrong number.

= 1 – P(0)

= 1 – 5C0 (1/200)0 (199/200)5-0

= 1 – (199/200)5

= 0.02475

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?
2. Is the venture likely to be successful? Explain
3. What is the long-term average earning of business ventures of this kind? Explain
4. What is the good measure of the risk involved in a venture of this kind? Compute this measure

Ans.

(i) Most likely monetary outcome of the business venture is $2000 because it has the highest probability of 0.3

(ii) Yes, the venture is likely to be successful coz P(X) > 0.6 (0.2+0.3+0.1)

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

(iii) The long-term average earning of business ventures of this kind is the

expected value = Sum [ E(X) \* P(X) ] = $ 800

(iv) Good measure of the risk involved in this venture is the Standard deviation