**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.
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 box-plot be affected?

**Ans =**

**1.**  The IQR range of the plot is approximately **10** (15-5). This value implies that 50% of the data set point lies b/w 5 to 15.

**2.** The plot appears to be **Positive skewed** as the median (10) is closer to the lower quartile (5) than the upper quartile (15).

**3.** If the plot points with the values of 25 the new box plot affects by Shorter in upper whisker, decrease in range, less positive skewed.

=====================================================================================

1. 

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 = 1.** The mode of the plot would lie around the value of 10 as it as the highest frequency.

**2.** The plot appears to be **Positive skewed** as the tail of graph extends to right.

**3.** The **Histogram** provides information about the **frequency distribution** of the plot.

The boxplot provides information about the **Median, quartiles and 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.)

**Ans =**  P(M) = Probability of a single call being misdirected = 1/200.

P(C) = Probability of a single call being correctly directed = 1-P(M) = 199/200.

N = Total no of attempts (5).

P(All 5 calls are correct) = P(C)5 = (199/200)5

P(at least 1 call Is misdirected) = 1-P (all 5 calls are correct)

P(at least 1 call Is misdirected) = 1-P(199/200)5

P(at least 1 call Is misdirected) = 1-(199/200)5 = 0.0249.

So the probability that at least one in five attempted telephone calls reached the wrong is approximately 0.0249 or 2.49%.

==============================================================================

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** =

1. The highest probability if for 2000

2. Yes, because the total earnings of the venture is positive in value i.e. 800 and highest probability of earnings is 2000.

3. Income = (x\*P(X))

|  |  |  |
| --- | --- | --- |
| x | P(x) | (x\*P(X)) |
| -2000 | 0.1 | -200 |
| -1000 | 0.1 | -100 |
| 0 | 0.2 | 0 |
| 1000 | 0.2 | 200 |
| 2000 | 0.3 | 600 |
| 3000 | 0.1 | 300 |

**Total 800**

4.

|  |  |  |
| --- | --- | --- |
| x | P(x) | x\*P(X) |
| -2000 | 0.1 | -2 |
| -1000 | 0.1 | -1 |
| 0 | 0.2 | 0 |
| 1000 | 0.2 | 2 |
| 2000 | 0.3 | 6 |
| 3000 | 0.1 | 3 |

**Var 7.222d**

**Std 2.68**