**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:**

IQR = Q3 - Q1 = 12 - 5

IQR = 7; this means that majority of the data points are located between Q1 & Q3.

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

**Ans:**

Dataset is positively skewed here, or Right skewed data.

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:**

1. new box-plot would not contain any outlier data point
2. and mean of the dataset would have shifted towards approx 5.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Ans:** approximately around 5 to 12.

1. Comment on the skewness of the dataset.

**Ans:** dataset is positively skewed / right skewed

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:**

1) histogram shows information about skewness, kurtosis, more confidently, and it also gives info about one outlier data point, and the frequency of the values.

2) whereas for box-plot it only shows outlier data point, and where the majority of the data points lie

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(single calls misdirected) = 1/200

P(call not being misdirected) = 1- 1/200 = 199/200

As attempts are independent therefore,

P(all of 5 call aren’t misdirected) = (199/200)^5

P(atleast 1 call is misdirected) = 1- (199/200)^5

= 0.02475

The probability that atleast one in 5 attempted telephone call reaches the wrong number is 0.02475 i.e. 2.475%

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: 2000**

1. Is the venture likely to be successful? Explain

Ans:

probability of positive returns = 0.2+0.3+0.1 = 0.6

probability of negative returns = 0.1+0.1 = 0.2

As the probability of positive returns > probability of negative return, we say that venture is likely to succeed.

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

Ans:

long term average earning means expected value = x \* p(x)

E = (-2000 \* 0.1) + (-1000 \* 0.1) + (0 \* 0.2) + (1000 \* 0.2) + (2000 \* 0.3) + (3000 \* 0.1)

E = -200 - 100 + 0 + 200 + 600 + 300

E = 800

The long-term average earning of business ventures of this kind is **$800**

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

**Ans:**

Standard Deviation (SD) = √[Σ((x - E)^2 \* P(x))

SD = √[((-2000 - 800)^2 \* 0.1) + ((-1000 - 800)^2 \* 0.1) + ((0 - 800)^2 \* 0.2) + ((1000 - 800)^2 \* 0.2) + ((2000 - 800)^2 \* 0.3) + ((3000 - 800)^2 \* 0.1)]

SD = √[784000+324000+128000+8000+432000+484000]

SD = √2158000

SD ≈ 1469.69

(measure of risk) of this venture is approximately **$1469.69**