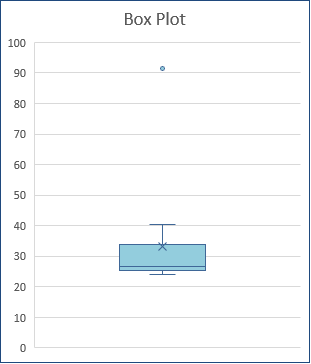
**Topics: Descriptive Statistics and Probability**

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% |

|  |  |
| --- | --- |
| Mean | 25.41 |
| Std | 16.37081 |
| var | 287.1466 |

From the data we can see Morgan Stanley has measure x = 91.36% which is an outlier.



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.

IQR = Q3 - Q1= 12 - 5 = 7.

The interquartile range (IQR) tells us the range where the bulk of the values lie.

i.e., 50% data

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

(Q2-Q1) < (Q3-Q2)

Median tends towards lower quartile and mean is higher than median. So, the dataset has positive skewness.

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?

Mean will decrease, but median remain same. So, difference between both decrease which result in decrease in skewness of distribution. Width of (Q2-Q1) increase and (Q3-Q2) decreases.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Mode lies in range of 4 to 8.

1. Comment on the skewness of the dataset.

We can observe median < mean as more value of lies in between 4 to 10. So, dataset has positive skewness.

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.

From boxplot we identified the outlier and median. In histogram we get mode or frequency distribution and skewness of the dataset. Skewness can be identified in both plots; one plot provide evidence to other plot.

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.)

from scipy.stats import binom

X= 1- binom(5, 1/200).cdf(1)

X

Ouput:

0.0002475093624999536

Given

n = 5

Probability of occurring 1 misdirected call = p(x) = 1/200

Probability of having at least 1 successful call = 1-p(x) = 1-1/200 = 0.967

P(x>=1) = 1-p(x<1)

= 1- (0.967) ^5

=0.0002475 or 0.024%

The probability that at least one in five attempted telephone calls reaches the wrong number is 0.024%.

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?

In probability distribution highest probability is 0.3.

So, most likely monetary outcome of the business venture is $2000.

1. Is the venture likely to be successful? Explain

Venture has both positive and negative returns.

Positive returns = 0.2+0.3+0.1= 0.6

= 60%

Negative returns = 0.1+ 0.1 = 0.2

= 20%

Since positive returns are more than negative returns, the venture is likely to be successful.

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

Long term Earning =

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

= $800

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

Standard deviation can be good measure of risk involved in a venture as it gives range of possible deviation of returns in the venture. The more the standard deviation more is the risk involved.

Standard deviation = $2213.94

Returns =$ 8002213.94