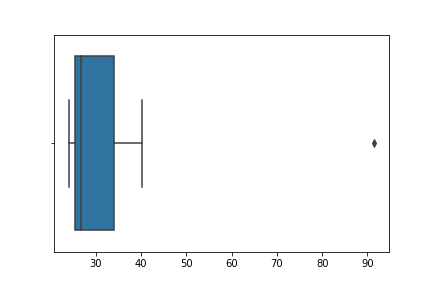
**Topics: Descriptive Statistics and Probability**

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:** using sns.boxplot(x) we get this figure 🡪



From the above Figure we can clearly see that Morgan Stanley 91.36 is the outlier.

And using x.mean() we get = 33.2713333333333

x.var() we get = 287.146612380952

x.std() we get = 16.9454009212220



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(inter-quartile range) = 12-5 = 7, this value implies that the range

which contains 50% of the data points.

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

Ans Right skewed or Positively skewed

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 2.5 would not be considered as an outlier because it would fall inside

the boxplot range i.e 0 to 19.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Ans It can be seen that the mode would lie above 4 and below 8.

1. Comment on the skewness of the dataset.

Ans Right skewed or Positively 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 Median is the main analyzing factor in boxplot whereas Mode in

histogram

Histogram provides the frequency distribution so we can see how many times each data point is occurring whereas boxplot provides the quantile distribution or rane i.e. 50% data lies between 5 and 12.

Boxplot provides whisker length to identify outliers but from histogram it is very difficult to say this, not sure that 25 is outlier.

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 let us consider the probability of 1 call misdirected out of 200 as event A.

Probability of occurring of event A= 1/200

P(A)= 1/200

Probability of having at least one successful call will be

1. P(A)= 1-1/200= 199/200= 0.967

As every event is independent of other event the probability will be

1 - (0.967) ^ 5

0.02475 = 2% chance.

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 Maximum P(x) is 0.3 for x= 2000 so most likely outcome is 2000.

1. Is the venture likely to be successful? Explain

Ans P(x>0) = P(1000) + P(2000) + P(3000)

= 0.2 + 0.3 + 0.1

= 0.6, Which means 60% chances of being successful.

P(x<0) = P(-1000) + P(-2000)

= 0.1 + 0.1

= 0.2, Which means 20% chances of loss.

So chances of venture being successful is more.

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

Ans x\*P(x) = (-200) + (-100) + 0 + 200 + 600 + 300

= 800

the long-term average earning of business ventures is 800

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

Ans P(x<0) = P(-1000) + P(-2000)

= 0.1 + 0.1

= 0.2, Which means 20% chances of risk.