**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: Mean is 33.27, standard deviation is 16.94 and variance is 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.

Answer: The inter quartile range is 7

IQR=Q3-Q1

12-5=7

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

Answer: Right 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?

Answer: In that case there would be no Outliers on the given dataset because of the outlier the data had positive skewness it will reduce and the data will normal distributed



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

Answer: The mode of this dataset lie 4-8

1. Comment on the skewness of the dataset.

Answer: 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.

Answer: They both are right skewed and both have outliers the median can be easily

visualized in box plot where as in histogram mode is more visible.

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.)

Answer: The probability of reaching one in five attempted calls reaches the wrong number is 0.024

probability of calls getting misdirected =1/200=.005

probability of calls not getting misdirected =199/200=.995

since one in 5 attempts the call reaches the wrong number

5\*(1/200)\*(199/200)

=0.0248

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?

Answer: The most likely monetary outcome of the business venture is 2000$ As for 2000$ the probability is 0.3 which is maximum as compared to others.

1. Is the venture likely to be successful? Explain

Answer: Yes the venture likely to be successfull.0.79 this states that there is a 79% chances for this venture to be making a profit.

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

Answer: long term average earning of business venture is 800

(-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

Answer: The measure of the risk depends on the variability of the distribution ,higher the variance more chances of risk , since the variance is higher the risk is higher

(-2000\*-200)+(-1000\*-100)+(0\*0)+(1000\*200)+(6000\*200)+(3000\*300)

= 2800000

= 800\*800= 640000

= 2800000-640000

= 2160000

np.sqrt(2160000)