**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=33.27%

Standard Deviation = 16.945

Variance 287.1466



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?

Answer:

1. In the above diagram Q1 =5 and Q3 = 12

IQR = Q3-Q1 = 12-5 = 7

The median lies between 5 and 12 which is in 7 variability measurement apart

1. The whisker is longer on the upper end hence it is positively skewed.
2. (iii) The upper whicker which is at 19 will be eliminated if we change the value of the outlier to 2.5



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.

Answer:

1. The mode will lie 5 and 6. Hence mode is 6
2. As the tail is increasing towards right so it is right skewed or positive skewness.
3. Both plots have a common outlier with value 25. Both are positive skewed. The mode is 6 and median is at 7. More than 50 % of the data lies between 5 and 12. Hence both the plats are drawn from the same dataset.
4. 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:

Probability of the call misdirecting is p which is 1/200

Hence Probability of the call not misdirecting is 1- 1/200

n= 5, q= 199/200 and p = 1/200

P(x) = nCx\*p^x\*q^n-x

As per the problem statement we have to find out at least one in five attempted telephone calls reaching the wrong number is

1-None of the call reaches the wrong number

1-(199/200)^5

0.02475

Hence at least one in five attempted telephone calls reaching the wrong number is 0.02475

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

Answer:

1. The most likely monetary outcome of the business venture is 2000 as it has the highest probability of 0.3
2. The probability of positive distribution is 0.6 where as for negative and zero is 0.4 hence the venture likely to be successful.
3. If we multiply x and p(x) and sum it the value will be 800 which is the average outcome of the business venture.
4. Total population Mean = (-2000-1000+0+1000+2000+3000)/6 =500

Variance = sum (x-mean)^2/6

=(-2000-500)^2+(-1000-500)^2+(0-500)^2+(1000-500)^2+(2000-500)^2+(3000-500)^2/6

=(-2500)^2+(-1500)^2+(-500)^2+(500)^2+(1500)^2+(2500)^2/6

=6250000+2250000+250000+250000+2250000+6250000/6

=17500000/6

=2910000.666

Standard Deviation = Square root of variance = sqrt 2910000.66 = 1707.

It’s a good measure of this business venture.