**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.

**Solution:** *IQR = Q3 – Q1*

*Q3 = approx. 12*

*Q1 = 5*

*IQR = 12-5 = approx. 7*

*Explanation : IQR signifies, the middle 50% of the data is between 12 and 5.*

*Min value of variable ‘X’ : 0*

*Max value of variable ‘X’ = approx. 19*

*Mid point / Median = approx. 7*

*Outlier is at 25*

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

**Solution:** Positive/right 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?

**Solution**: If ‘25’ was 2.5, then it would come under 1st Quartile , and Q1 may shift to a bit left.

Doubt in the above iii) point, can I get some clarifications on that.



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Solution:** *there are two modes here, approximately Y bar of (4-6) – 1 mode, has the highest peak*

*2nd mode is Y bar of (6-8) has the same peak.*

*Correct me, If I am wrong.*

1. Comment on the skewness of the dataset.

**Solution:** *Positive/ 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.

**Solution:**

**From Box plot: below points can be concluded**

* *Positive/right Skewed*
* *Median at around 7*
* *Min & Max values : 0 & approx. 19*
* *Outlier is at 25*
* *50% of data lies between 0 & 7, rest 50% lies between 7 & 19*

**From Histogram:**

* *Positive/right skewed*
* *Outliers at around 24-26*
* *Min & max value : 0 & 20*

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.)

**Solution:**

*one in 200 long-distance telephone calls is misdirected*

*=>  probability of call misdirecting  p = 1/200*

*Probability of call not Misdirecting = 1 - 1/200 = 199/200*

*Number of Calls = 5*

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

*n = 5*

*p = 1/200*

*q = 199/200*

*at least one in five attempted telephone calls reaches the wrong number*

*= 1  -  none of the call reaches the wrong number*

*= 1  - P(0)*

*= 1   -  5C0\*(1/200)^0(199/200)^5-0*

*= 1  -  (199/200)^5*

*= 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?

**Solution:** 2000, it has maximum probability = 0.3

1. Is the venture likely to be successful? Explain

**Solution** : as per the Graph, Yes. We can see the probability distribution for the profitable business is comparatively more than the loss business.

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

**Solution:** Need some idea on this

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

**Solution :** Probability of profitable business = (0.2 + 0.3 + 0.1) = 0.6

Probability of non-profitable business = 1-0.6 = 0.4(out of which 0.2 is for disinterested business)

So the only loss we can assume for P(x) = 0.2

So we can conclude on seeing this probability, there is a good risk involved.