**Topics: Descriptive Statistics and Probability**

1. Look at the data given below. Plot the data, find the outliers and find out the outliers and find

|  |  |
| --- | --- |
| **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.

**ANS**:

**Approximately (First Quantile Range) Q1 = 5**

**(Third Quantile Range) Q3 = 12**

**Median (Second Quartile Range) = 7**

**(Inter-Quartile Range) IQR = Q3 – Q1 = 12 – 5 = 7 the IQR describes the middle 50% of data values when ordered from lowest to highest.**

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

**Ans:**

**From the above boxplot we can say that it is slightly positively skewed because the length of the right side tail is longer when compared to the left side tail.**

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:**

**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?

**ANS: The mode of this data set lie in between 5 to 10.**

1. Comment on the skewness of the dataset.

**ANS:  Right-Skewed. Mean>Median>Mode**

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:**

**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.)

**ANS:**

**IF 1 in 200 long-distance telephone calls are getting misdirected.**

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

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

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

**Number of Calls = 5**

**n = 5**

**p = 1/200**

**q = 199/200**

**P(x) = at least one in five attempted telephone calls reaches the wrong number**

**P(x) = ⁿCₓ pˣ qⁿ⁻ˣ**

**P(x) = (nCx) (p^x) (q^n-x) # nCr = n! / r! \* (n - r)!**

**P(1) = (5C1) (1/200)^1 (199/200)^5-1**

**P(1) = 0.0245037**

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 |

### Solution:

* E(X) = ∑X . P(X)
* E(X²) = ∑X² . P(X)
* Var (X) = E(X²)  - { E(X) }²
* SD = √Var

|  |  |  |  |
| --- | --- | --- | --- |
| X | P(X) | E(X)= X . P(X) | E(X2)=X2.P(X) |
| -2000 | 0.1 | -200 | 400000 |
| -1000 | 0.1 | -100 | 100000 |
| 0 | 0.2 | 0 | 0 |
| 1000 | 0.2 | 200 | 200000 |
| 2000 | 0.3 | 600 | 1200000 |
| 3000 | 0.1 | 300 | 900000 |
|  |  | 8000 | 280000 |

1. What is the most likely monetary outcome of the business venture?

**ANS:**

**most likely monetary outcome of the business venture is $ 2000 as it has maximum Probability  0.3**

1. Is the venture likely to be successful? Explain

**ANS:**

**Venture is successful if X is + ve**

**Hence if X is 1000 , 2000 or 3000**

**Probability is  0.2 + 0.3 + 0.1 = 0.6**

**as 0.6 > 0.5 Hence venture likely to be successful**

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

**ANS:**

**long-term average earning of business ventures  = E(X)**

**E(X) = ∑ X.P(X)  = $ 800**

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

**ANS:**

**Risk involved in a venture**

**Var (X) = E(X²)  - { E(X) }²**

**=   2800000 -   800²**

**= 2160000  ( Quite High)**

**SD = √Var  ≈ $ 1470**

**As Variability is Quite high  hence Risk is high.**