**Topics: Descriptive Statistics and Probability**

* 1) Look at the data given below. Plot the data, find the outliers and find out mean, standard deviation and variance.

|  |  |
| --- | --- |
| **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:- Outliers**: Company Measure X (%)

Morgan Stanley **91.36**

**Mean: 33.27133333333333**

**Standard Deviation: 16.945400921222028**

**Variance: 287.1466123809524**

2)



Answer the following three questions based on the box-plot above.

* 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****, (Inter-Quartile Range)*

***IQR =Q3 – Q1 = 12 – 5 = 7***

***THIS SHOWS THAT 50% OF THE DATA LIES BETWEEN IQR.***

* What can we say about the skewness of this dataset?

***ANS:-* THE SKEWNESS OF THE DATA IS POSITIVE. IT IS RIGHT**

**SKEWED**.

* 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:-* IF THE VALUE IS 2.5 THEN THERE WILL BE NO OUTLIER ON**

**THE NEW BOXPLOT.**

* 3)



Answer the following three questions based on the histogram above.

* Where would the mode of this dataset lie?

**ANS:-** **THE MODE OF THIS DATASET LIE IN BETWEEN**

**4 TO 8 APPROXIMATELY** .

* Comment on the skewness of the dataset.

**ANS:-** **AS WE OBSERVE FROM THE ABOVE HISOGRAM , IT IS**

**RIGHT SKEWED . IT IS HAVING POSITIVE SKEWNESS**.

* 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:-** **BOTH OF THEM HAVE THE SAME OUTLIER DATAPOINT i.e.**

**25 AND THEY ARE RIGHT SKEWED. THROUGH BOXPLOT WE**

**CAN FIND OUT THE MEDIAN AND THROUGH HISTOGRAM**

**WE CAN FIND THE MODE.**

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.)

**ANS:- GIVEN :-** one in 200 long-distance telephone calls is misdirected.

**TO FIND**:- probability that at least one in five attempted telephone calls reaches the wrong number.

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

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

Number of calls **= 5**

Probability of having atleast one successful call will be

**1-P(X) = 1-1/200 = 199/200 = 0.995**

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

**(0.995)^5 = 0.97524875312**

**=> 1 - 0.97524875312 = 0.02475124688**

**approximately = 0.025**

5) 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 |

* a)What is the most likely monetary outcome of the business

venture?

**ANS:-** 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.

* b)Is the venture likely to be successful? Explain

**ANS**:- It is successful because

p(x=1000 )+p(x=2000)+p(x=3000) = 0.2+0.3+0.1 = **0.6** = **60%**

**more than 50%.**

* c)What is the long-term average earning of business ventures of this kind? Explain

**ANS**:- The long-term average earning of business ventures of this

kind are average of probability distribution.

**(-2000)(0.1)+(-1000)(0.1)+0(0.2)+(1000)(0.2)+(2000)(0.3)+(3000)**

**(0.1) = $800**

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

**ANS**:- **Var (X) = 2160000**

**Standard Deviation**: 1469.6938456699068

**Approximately = 1470**

The good measure of the risk involved in a venture of this kind depends on the Variability in the distribution. **Higher Variance means more chances of risk .**