**Topics: Descriptive Statistics and Probability**

* 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% |

**The outlier of the above data is Morgan Stanley 91.36%**

**Mean of the above data set is 33.27133**

**Standard deviation is 16.945401**

**Variance is 287.146612**



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.

**First Quartile Q1= 5 Third Quartile Q3=12**

**Inner Quartile Range(IQR)=Q3-Q1=12-5=7**

**The Inner Quartile range is equal to the second quartile (median)**

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

**Since the median is towards the first quartile this implies that it is a right skewed(positive skew)**

* If it was found that the data point with the value 25 is actually 2.5, how would the new box-plot be affected?

**Then their will be no outliers on the given dataset since it is positive skewness then the given data set reduces to normal distribution**



Answer the following three questions based on the histogram above.

* Where would the mode of this dataset lie?

**The mode of the given dataset lie between the range (4,8)**

* Comment on the skewness of the dataset.

**It is Right skewed (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.

**In both the plot i.e in histogram and in boxplot we have right skewness and also we the outliers whereas in histogram it is easy to find the mode of the given data set**

* 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.)

**If 1 in 200 long distance telephone calla are getting misdirected**

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

**probability of call not misdirecting=1-1/200=199/200**

**The probability that at least one in five attempted telephone calls reaches the wrong number=P(x)=(5C1)\*(1/200)\*(199/200)^(5-1)=0.0245037**

* 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 |

* What is the most likely monetary outcome of the business venture?

**The most likely monetary outcome of the business venture is $2000 whose probability value is 0.3**

* Is the venture likely to be successful? Explain

**Yes, the probability that the venture will make more than 0 are in profit**

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

**Their is 60% chance for this venture to make the profit**

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

**The long term average earnings of the business ventures is given by value**

**=Sum(x\*P(x))=(-2000\*0.1)+(-1000\*0.1)+(0\*0.2)+(1000\*0.2)+(2000\*0.3)+(3000\*0.1)=800**

**that means means on average the returns will be 800$+**

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

**The good measure of the risk involved in a venture of this kind depends on the variability in the distribution.**

**Higher the variance means more the chances of risk**

**Mean=0.1+0.1+0.2+0.2+0.3+0.1=1**

**Standard deviation =(0.1-1)^2+(0.1-1)^2+(0.2-1)^2+(0.2-1)^2+(0.3-1)^2(0.1-1)^2=4.2**

**variance=sqrt(4.2)=2.04939**

**Since the variance is low their is less chances of getting risk**