**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% |

Ans:The outliers in the boxplot is Morgan stanley=91.36%

Mean=33.271333

Standard deviation=16.945401

Measure\_x.var()

Variance=287.146613809524



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?

(1)Ans:Q is quartile

HereQ1=5Q3=12,Median is Q2=7(Inter quartile range)IQR=Q3-Q1=12-5=7 Q2 is the median value

(2)ans: positive skewed toward left side

(3)Ans: if that happens than there would be no outliers on the given data set only because of the dataset the outliers having the positive skeweness



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. (1)Ans: The mode of these data set lies in between 5to 10 and between4to 8

(2)Ans: positive\_skewed

(3)Ans:both are positively skewed and bothe have outliers the median can be easily seen in boxplot but where as in histogram mode is more visible

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: One in 200 long distance telephone calls is misdirected

* Probability of call misdirecting p=1/200

\* Probability of calls which is not misdirected=1-1/200=

Number of call=5

P(X)=nCxpXqn-X

N=5

P=1/200

q=199/200

At least one in five attempts telephone calls reaches the wrong number

=1-no calls 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?
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

(1)Ans: The probability (0.3) is high for 2000$ as compared to others so therefore, mostlikely monetary outcome of the business venture =2000$

(2)Ans: Long term average =sum[p(xi)\*XI]=(-2000\*0.1)+(-1000\*0.1)+(0)+(1000\*0.2)+(2000\*0.3)+(3000\*0.1)=800$ as the long term average give positive numbers business venture seems to be successful

(3)Ans:Long term average=sum[p(xi)\*Xi]=(-2000\*0.1)+(-1000\*0.1)+(0)+(1000\*0.2)+(2000\*0.3)+(3000\*0.1)=800$ Average return will be 800$ (4)ANS: variability is high hence there is high risk

Var(x)=E(X^2)-(E(X))^2=2800000-800^2=2160000