**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.

**Answer:** Approximately Q1 = 5(first quartile) and (Third Quantile) Q3 = 12

**We know that** (Inter-Quartile Range) IQR = Q3 – Q1 = 12 – 5 = 7,

**In the above bar plot** we also see that median value is approximately 7 so the IQR

Value shows the **median value that is second quartile value.**

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

**Answer:** This is right skewed dataset(Positively skewed), because median is shift towards left

Side.

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?

**Answer:** If this datapoint change into 2.5 then the dataset would be no outliers, then it will

Change from right skewed to the normal distribution(approximately symmetric type

Dataset)



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Answer:**  The mode of this dataset lie approximately between 4 to 8

1. Comment on the skewness of the dataset.

**Answer:** Positively skewed dataset in which 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.

**Answer:**  We clearly see that both plot shows right skewed dataset and also both of them

Contains outliers, histogram easily shows **MODE**  of the dataset where as box plot

Shows **MEDIAN**  of the dataset easily.

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.)

**Answer: In python:- from scipy.stats import binom**

**binom.pmf(1,5,1/200)= 0.02450373751562501**

**We know that in binomial distribution- P(X)=** **P(x) = ⁿCₓ pˣ qⁿ⁻ˣ**

So here in this problem n=5,x=1, and probability of call misdirected(p) = 1/200 ,Probability of call not Misdirected(q) = 1-(1/200) = 199/200

So after putting the values in the above binomial distribution we get 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) | E(X)=X\*p(x) | E(X^2)=X^2 \* P(X) |
| -2,000 | 0.1 | -200 | 400000 |
| -1,000 | 0.1 | -100 | 100000 |
| 0 | 0.2 | 0 | 0 |
| 1000 | 0.2 | 200 | 200000 |
| 2000 | 0.3 | 600 | 1200000 |
| 3000 | 0.1 | 300 | 900000 |
|  | Total | 800 | 2800000 |

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

**Answer:** The most likely monetary outcome of the business venture is 2000$, because the prob

of this venture is maximum as compared to the others.

1. Is the venture likely to be successful? Explain

**Answer:** Yes, because here at least the probability of x=0 is 0.2+0.2+0.3+0.1=0.8. It means 80% profit it returns.

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

**Answer:** The long-term average earning of business ventures of this kind is 800 which means

On an average returns on a certain business venture is 800

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

**Answer:** For finding the good measure of the risk involved in a venture we have to find the variability of the risk. So we know that **Var(x)=E(x2) – (E(x))2**

**So here var(x)=** 2800000 – 8002 =2160000