**Topics: Descriptive Statistics and Probability**

* Look at the data given below. Plot the data, find the outliers and find out μ,σ,σ^2

|  |  |
| --- | --- |
| **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 following is the outlier in the boxplot:

Morgan Stanley 91.36%

Measure\_X.describe()

Mean = 33.289333

Standard deviation = 16.935149

measure\_x.var()

Variance = 286.7992638095238



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: IQR= Q3-Q1 = 12-5 = 7

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

Ans: Median is skewed towards right. Hence not normally distributed.

* 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: There will be no outliers in the data. The distribution could reduce skewedness towards right. Also might get normally distributed.



Answer the following three questions based on the histogram above.

* Where would the mode of this dataset lie?

Ans: Mode of the data should lie between 5 and 10.

* Comment on the skewness of the dataset.

Ans: Right skewed. with outliers present on right side.

* 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 provide information about presence of outlier. also that data is rightly skewed and not normally distributed.

* 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: Number of calls n=200

Prob of call been misdirected p=1/200

Prob of call not been misdirected q=1- 1/200= 199/200

Probablility that atleast 1 in 5 calls been misdirected= P(x) = ⁿCₓ pˣ qⁿ⁻ˣ

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

P(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?

Ans: Most likely outcome is profit of 2000$.

* Is the venture likely to be successful? Explain
* Ans. Probablility of making 0 or more than that= P(0)+P(1000)+P(2000)+P(3000)=0.2+0.2+0.3+0.1=0.8

80% chances of venture making a profit

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

Ans: The long-term average is Expected value = Sum (X \* P(X)) = 800$ which means on an average the returns will be + 800$

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

Ans: Variance is a good measure to risk involve in a venture. Higher vairance means more risk.

E(X) =Sum X.\*P(X) | E(X^2) =X^2\*P(X)

-200             | 400000

-100                 | 100000

0             | 0

200       | 200000

600         | 1200000

300         | 900000

Total: 800         | 2800000

Var (X) = E(X^2) –(E(X))^2

= 2800000 – 800^2

Ans: = 2160000