**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% |

Ans)



Outliner for the given above data is Morgan Stanley which is 91.36.

Mean for the above data = 33.28

Standard deviation for the above data = 16.94

Variance for the above data = 287.14



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) First quartile range is approximately = 5

Third quartile rage is approximately = 12

Inter quartile range = Third quartile range - First quartile range

= 12-5

= 7

Inter quartile range is approximately = 7

It is the Median value.

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

Ans) Right skewed median is towards the left side, it is not normal distribution.

* 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) In that case there would be no outliners on the given dataset and the data will be normally distributed.



Answer the following three questions based on the histogram above.

* Where would the mode of this dataset lie?

Ans) Repeated value is 20, hence the mode would lie in between from 4 to 8 approximately.

* Comment on the skewness of the dataset.

Ans) The above given dataset is right skewed because mean is greater than the median and mode is greater than the median.

* 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) Above both histogram and box plot are right skewed and both have the outliners. But in boxplot median is visible and in the histogram mode is 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) This problem can be modeled using the binomial distribution, where each attempted telephone call is a Bernoulli trial with a probability of success (i.e. being misdirected) of 1/200. We want to find the probability of at least one success in five trials.

Using the complement rule, we can find the probability of no successes in five trials and subtract it from 1 to get the desired probability:

P(at least one success in five trials) = 1 - P(no successes in five trials)

P(no successes in five trials) = (199/200)^5

P(at least one success in five trials) = 1 - (199/200)^5

P(at least one success in five trials) ≈ 0.024

Therefore, the probability that at least one in five attempted telephone calls reaches the wrong number is approximately 0.024, or about 2.4%.

* 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) The most likely monetary outcome of the buisness venture is 2000, as for 2000 P(x) is 0.3 which is maximum.

* Is the venture likely to be successful? Explain

Ans) Have to calculate the sum of the P(x) where the buisness venture made profit

= P(x>0) + P(x>1000) + P(x>2000) + P(x>3000)

= 0.2 + 0.2 + 0.3 + 0.1

= 0.8

Yes, venture likely to be successful as the sum of probability is approx 80%.

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

Ans)

|  |  |
| --- | --- |
| E(X^2) =X^2\*P(X) | E(X) =Sum X.\*P(X) |
| 400,000 | -200 |
| 100,000 | -100 |
| 0 | 0 |
| 200,000 | 200 |
| 1,200,000 | 600 |
| 900,000 | 300 |

long-term average earning of business ventures  = E(X)

E(X) = ∑ X.P(X)  = $ 800

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

Ans) Risk involved in a venture of this kind

Var(X) = E(X2) – {E(X)}2

= 2,800,000 – 800 2

= 2,800,000 – 640,000

= 2,160,000

SD = √Var

= √2,160,000

= $1470

As variability is Quite high hence Risk is high