**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% |

Outlier=91.36



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.

* 5 to 13 or 7, it show us the data is positively skew

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

* positively skew

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?

* The IQR value may lower and the skewness may be lesser



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

* Between 5 to 6

1. Comment on the skewness of the dataset.

* positively skew

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.

* Both are good at showing skewness and the the boxplot some times better to find the outliers

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.)

Probability of call misdirecting  p = 1/200

     Probability of call not Misdirecting = 1 - 1/200 = 199/200

Number of Calls = 5

P(x) = ⁿCₓpˣqⁿ⁻ˣ n=5 p = 1/200 q = 199/200

at least one in five attempted telephone calls reaches the wrong number

= 1  -  none of the call reaches the wrong number

= 1  - P(0)

= 1   -  ⁵C₀(1/200)⁰(199/200)⁵⁻⁰

= 1  -  (199/200)⁵

= 0.02475

**probability that at least one in five attempted telephone calls reaches the wrong number = 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?

* 200

1. Is the venture likely to be successful? Explain

* The probability values are good in case of positive values

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

* 1000 to 2000 because the probability holds at 50 per

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

=> Probability of success of this venture is p>0

0.2**+**0.2**+**0.3**+**0.1 = 0.7999999999999999

#computing long term average income sum(X)\*p(x)

(**-**2000**\***0.1)**+**(**-**1000**\***0.1)**+**(0**\***0.2)**+**(1000**\***0.2)**+**(2000**\***0.3)**+**(3000**\***0.1) = 800.0

#computing the risk factor by variance.

(**-**2000**\*-**200)**+**(**-**1000**\*-**100)**+**(0**\***0)**+**(1000**\***200)**+**(2000**\***600)**+**(3000**\***300) = 2800000

#square of long term average income

800**\***800 = 640000

#var(x) = 2800000**-**640000

2160000

#standard deviation is

np**.**sqrt(2160000) = 1469.6938456699068