**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% |

* There is one outlier at 91.36 which is “Morgan Stanley”
* Mean = 33.27
* Standard deviation = 16.945
* Variance = 287.14



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. **Ans**: (12-5)=7, IQR implies that middle half of data set falls in this range
2. What can we say about the skewness of this dataset? **Ans**: Positive skewed
3. f it was found that the data point with the value 25 is actually 2.5, how would the new box-plot be affected? **Ans**: The left whisker and right whisker would be of same length



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie? **Ans**: Mode lies in (4-8)interval on x-axis as more number of values of Y fall in that interval
2. Comment on the skewness of the dataset. **Ans**: Right skewed
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. **Ans**: We can find mode using histogram and median, outliers using boxplots
4. 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.)

**Binomial distribution**

*Ans:  The probability of failure(q) is just 1 minus the probability of success(p): P(F) = 1 – p.*

One wrong number in 200 calls

Probability of one call misdirecting p = 1/200 = 0.005

Probability of one call not misdirecting q= 1-0.005 = 0.995

Number of calls = 5

Using,

P(x) = ⁿCₓpˣqⁿ⁻ˣ

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

= 1- None of the call reaches wrong number

= 1- 5C0 \* 0.005^0 \* 0.995^5

= 1- 0.995^5

=0.024

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?

Ans: The most likely monetary outcome of the business venture is 0.3

1. Is the venture likely to be successful? Explain\

Ans: Yes the venture is likely to be successful

p(x=1000)+p(x=2000)+p(x=3000)=0.2+0.3+0.1=0.6

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

Ans: (0.1)(-2000)+(0.1)(-1000)+(0.2)(0)+(0.2)(1000)+(0.3)(2000)+(0.1)(3000)=800

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

Good Measure of risk is standard deviation