**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:**

**IQR = Q3-Q1**

**IQR = 12-5 = 7**

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

**Answer: Data is right skewed as there are outliers and data is not normally distributed.**

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: in that case the data would be normally distributed data.**



Answer the following three questions based on the histogram above.

1. Where would the mode of this dataset lie?

**Answer: as data is right skewed the mode of data lies between 5-10**

1. Comment on the skewness of the dataset.

**Answer: the data is right skewed where 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: they are both right skewed and there is outlier at 25. Where mean > Median> Mode**

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**: **(Problem of Combinational Probability)**

**1 in 200 long-distance telephone calls are getting misdirected.  
Probability of call misdirecting = 1/200.**

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

**Probability for at least one in five attempted telephone calls reaches the wrong number Number of Calls (n)= 5**

**p = 1/200**

**q = 199/200**

**P(x) = at least one in five attempted telephone calls reaches the wrong number**

**P(x) = ⁿCₓpˣqⁿ⁻ˣ**

**=  -  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**

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?

**Answer: The most likely monetary outcome of business venture P(X) = 0.3**

1. Is the venture likely to be successful? Explain

**Answer: Yes, the probability that the venture likely to be successful =**

**P(x>0)= P(X=1000,2000,3000) = 0.2+0.2+0.3+0.1 = 0.8**

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 ∑Xi\*P(xi) =(-2000\*0.1)+(-1000\*01)+(1000\*0.2)+(2000\*0.3)+(3000\*0.1) = 800**

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

**Answer: The good measure of the risk involved in a venture of this kind depends on variability in the expected returns**

**Higher the variance means more the chances of risk**

**Varience = E(X2)-(E(X))2**

**= 2800000-800**

**=2160000**

|  |  |  |  |
| --- | --- | --- | --- |
| **X** | **P(x)** | **E(X)=Xi\*p(xi)** | **(E(X))2** |
| **-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** |
|  |  | **800** | **2800000** |