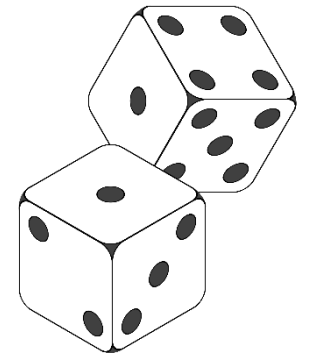


STATISTICS AND DATA ANALYSIS

TA Session: Probability Practice
October 6, 2014

Practice 1

A fare dice is rolled and a fair coin is tossed. Find the probability that the dice shows an odd number and the coin shows a head.



Practice 2

Suppose A and B are independent events, B and C are mutually exclusive, and A and C are independent events. Moreover, we have $\Pr(A)=0.4$, $\Pr(B)=0.9$ and $\Pr(C)=0.1$. Find the following probabilities:

- a) $\Pr(A \cap B)$
- b) $\Pr(B \cap C)$
- c) $\Pr(A \cap B \cap C')$
- d) $\Pr((A \cap C) \cup B)$
- e) $\Pr((A \cap C') \cup B)$

Practice 3

You decide you're only going to buy a lottery ticket if your expected winning is larger than the ticket price. Suppose a ticket costs \$10:

With probability 0.01, you win \$1000.

With probability 0.05, you win \$100.

With probability 0.1, you win \$10.

Should you buy a ticket for this lottery? Why?

Practice 4

Define X = number of **heads** after **3 flips** of an **unfair coin** with the following distribution:

$$\Pr(X = \text{Head}) = 0.3 \text{ and } \Pr(X = \text{Tail}) = 0.7.$$

- a) List all the possible outcomes of X .
- b) What are the probabilities of all the outcomes of X ? (You may use R as a calculator.)

Practice 4

Define X = number of **heads** after **3 flips** of an **unfair coin** with the following distribution:

$$\Pr(X = \text{Head}) = 0.3 \text{ and } \Pr(X = \text{Tail}) = 0.7.$$

- c) Using the R code mentioned in the video, find the expected value of X .
- d) Using the R code mentioned in the video, find the variance and the standard deviation of X .

Practice 5

Consider the wholesale data set:

- a) Extract sales data collected from channel 1 and region 1, with 4 columns: Channel, Region, Fresh and Milk.

Hint:

Use “which” and “data.frame” function.

Practice 5

- b) For sales data collected from channel 1 and region 1, calculate the means, medians, and sample variances for milk sales.

Hint:

Use “mean”, “median” and “var” function.

You can also use “summary” function to see what happened.

Practice 5

- c) For sales data collected from channel 1 and region 1, draw a histogram for milk sales data with the default number of classes and class intervals.

Hint:

Use “hist” function.

- d) For each of the six channel-region combination, calculate the sample correlation coefficient between fresh food sales and milk sales.

Practice 5

- e) Draw scatter plots for the channel-region combinations with the highest and lowest correlation coefficients.

Hint:

Use “plot” function.