Lab 2 Assignment

Dr. Purna Gamage

1. This question refers to Experiment 1 in Lab 2:

```
library(gtools)
```

   a. Practice Experminet 1: example 1 and example 2.(You don't have to submit this part)

   b. What is the difference if between "%in%" and "=="?

   c. Try to find the *probabilities* of part (a) with a different set of codings without using %in%. (Hint:Only change the codes with %in%)

Eaxample1: (i). Probability of getting the first card a Queen card?

(ii). Conditional probability of the second card being a Queen given that the first card was a Queen?

Example2: Probability of a Natural 21 in Blackjack?

2. (This is fictional) It is believed that for COVID-19 a highly sensitive test has a low false-negative rate. A highly specific test has a low false-positive rate. In a certain testing center, they have found out that for a certain COVID-19 testing, the test is positive 88% of the time when tested on a patient with the disease (high sensitivity). The test is negative 90% of the time when tested on a healthy patient (high specificity). Probability of having COVID-19 is 0.25 for the patients come to that testing center.

   a. Specify each probability using conditional probability notations.

   b. Set the seed. Let's try this example for 1000000 simulations.

Let's asssume that the patients arrive at this testing center are testing only for the Corona Virus. Simulate 1000000 patients coming to this testing center using sample() and specifying the probabilities of the disease. (Name the variable as "disease"). Use the values 0 and 1 where "disease=0" is the set of patients who does not have the virus and "disease=1" is the set of patients who have the virus.

   c. Simulate 1000000 tests for these patients. First, generate an empty vector for "test". Then fill in the values 0 and 1 where "disease=0" is the set of patients who does not have the virus and "disease=1" is the set of patients who have the virus. (Hint: use sample() function for creating each subsets)

Answer the following questions by first calculating by hand and then using the simulations. Compare your results.

   d. What is the probability that a test is positive?

   e. Calculate the probability that a patient have the disease if the test is positive.

   f. What is the probability that an individual has the disease if the test is negative?

   g. If a patient's test is positive, how much does that increase their risk of having the disease?

(Hint: First calculate the probability of having the disease given a positive test, then divide by the probability of having the disease.)