STAT 41600/41700

Project # 2

Task #1: A very powerful probabilistic tool used in all areas of science is what is called *simulation*. This is where a computer is used to simulate an experiment many times to approximate the probabilities of certain events. The fundamental idea is based on the *long-run frequency interpretation* of probability. That is, the probability $\Pr(E)$ of an event, E to occur is just the proportion of times the event E occurs in 'infinitely' many repetitions of the experiment or, in other words,

$$P(E) = \lim_{n\to\infty} \frac{\text{\# of times E occurs in n trials}}{n}$$

Of course, we can't perform any experiment infinitely many times, but we can use a computer to do it a large number of times, say N. Then we should have, for a large N

$$P(E) \approx \frac{\text{\# of times E occurs in the N trials}}{N}$$

Your **task** now is to simulate, using R, different events, and to figure out their probabilities by using the above relative frequency approach and determine which game is more favorable to "Winning". That is, use R to simulate 3,000 plays of the following games:

1) **Game 1**: roll a die one time, **WIN** if you have an Ace (a '6').
2) **Game 2**: roll a die four times, **WIN** if you have one or more Aces.
3) **Game 3**: roll a pair of dice 24 times, **WIN** if you have one or more Double Aces.

Note, you don't have to perform all 3,000 plays, at once. You can perform n=100, plays at first, then n=200, n=400, n=500, n=800 and finally n=1000 additional plays and then record the results and complete the tables below.

- For example, for n=100 plays of Game 1:

```
> xx<-sample(seq(1,6, by=1), 100, replace=T)
> xx
[1] 4 6 5 1 1 1 2 5 2 1 1 1 4 1 3 2 1 2 6 5 5 1 1 6 4 5 5 2 5 1 2
4 1 3 5 2 2 1 6 3 4 3 5 6 3 5 1 1 6 6 1 3 1 4 6 3 4 2 4 6 2 3 6 4
5 6 2 2 6 5 2 1 1 5 6 3 3 6 1 4 1 5 2 5 5 2 3 4 5 6 6 4 2 3 1 4 2
1 6 1

> table(xx==6)
FALSE  TRUE
   83    17
```

Here 'TRUE' means a **'WIN'** .

- This would be equivalent to using the 'Binomial' distribution (to be discuss in Chapter 4) to simulate the game

```
> xx<-rbinom(n=100, 1, 1/6)
xx
0 0 0 0 0 1 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0
0 0 0 0 0 0 0 0 1 0 1 0 0 0 1 0 0 0 0 1 0 1 0 0 0 1 1 0 0 1 0 1 1
0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 1
0

> table(xx==1)
xx

FALSE  TRUE
  80     20
```

- Similarly, for n=100 plays of Game 2,

```
> xx<-rbinom(n=100, 4, 1/6)
> xx
0 1 1 1 0 1 0 0 2 0 0 2 2 0 0 0 2 0 1 0 0 1 0 1 2 1 1 0 0 0 0 2 2
0 1 1 1 0 0 2 0 1 0 0 1 1 1 0 1 1 1 1 0 1 1 1 1 0 2 0 0 1 1 1 0 1
1 0 2 1 0 0 1 1 0 1 3 2 0 1 0 0 0 1 1 1 0 1 1 0 1 1 0 0 2 1 1 1 1
0

> table(xx>0)

FALSE  TRUE
   42    58
```

- And, for n=100 plays of Game 3

```
> xx<-rbinom(n=100, 24, 1/36)
> xx
0 0 1 3 1 0 2 0 2 1 0 1 0 0 1 0 1 1 0 2 0 0 0 1 1 0 0 1 1 1 1 0 1
0 0 1 0 1 0 0 0 1 2 1 1 0 1 0 0 0 0 0 0 0 1 0 2 1 0 1 0 1 2 0 1 1
0 2 1 0 0 2 2 0 0 1 3 0 0 0 0 0 1 1 1 1 0 0 0 0 0 1 1 1 0 0 1 0 1
2

> table(xx>0)

FALSE  TRUE
   51    49
```

STAT 416/417

Project 1

Name: ___________________________

Complete the following tables with the results of your simulations and turn this page in with a copy of your R code.

GAME 1					$P(Win) =$
#of Plays n=	# Losses	# Wins	Cumulative # of plays	Cumulative # of Wins	Relative Frequency of Wins
100	80	20	100	20	0.2
200	162	38	300	58	01933
400	326	74	700	132	0.188
500			1200		
800			2000		
1000			3000		0.1677

GAME 2					$P(Win) =$
#of Plays n=	# Losses	# Wins	Cumulative # of plays	Cumulative # of Wins	Relative Frequency of Wins
100	42	58	100	58	0.58
200			300		
400			700		
500			1200		
800			2000		
1000			3000		

GAME 3					$P(Win) =$
#of Plays n=	# Losses	# Wins	Cumulative # of plays	Cumulative # of Wins	Relative Frequency of Wins
100	51	49	100	49	0.49
200			300		
400			700		
500			1200		
800			2000		
1000			3000		