ST 314 Data Analysis 01

Ivan Timothy Halim

January 13, 2019

Part 1

Random Variable 1

We should use the Binomial Distribution because the probability of outcome (p) and the number of trials (N) is known. In this case:

p = 0.02, N = 10

x	0	1	2	3	4	5	6	7	8	9	10
p(x)	0.817	0.167	0.015	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Random Variable 2

In this case, we don't know the number of trials (N) and we don't know the probability of outcome (p), but we do know the average number of claims, which is 0.8 claims in 2 years. Therefore we should use the Poisson Distribution:

$\lambda = 0.8$

x	0	1	2	3	4	5	6	7	8	9	10
p(x)	0.449	0.359	0.144	0.038	0.008	0.001	0.000	0.000	0.000	0.000	0.000

Part 2

a. $300

b. 0.083

c. $E(x) = \sum x \,.\, p(X = x)$

$= 0(0.083) + 300(0.208) + 350(0.042) + 400(0.083) + 450(0.042) + 500(0.125) + 550(0.042) + 600(0.125) + 700(0.042) + 800(0.083) + 900(0.083) + 5000(0.042)$

$= \$670$

d. We can solve this by using conditional probability

$p(\text{All zeros}) = p(X = 0)^3 = (0.083)^3 = 5.72e\text{-}4$

Or we can use the binomial distribution where X = 3, N = 3 and p = 0.083

```
dbinom(3, 3, 0.083)
```

```
## [1] 0.000571787
```

e. We can solve this problem using the cumulative distribution function

$p(X >= 1) = F(X = 3) - F(X = 0)$

```
pbinom(3, 3, 0.083) - pbinom(0, 3, 0.083)
```

```
## [1] 0.2289048
```

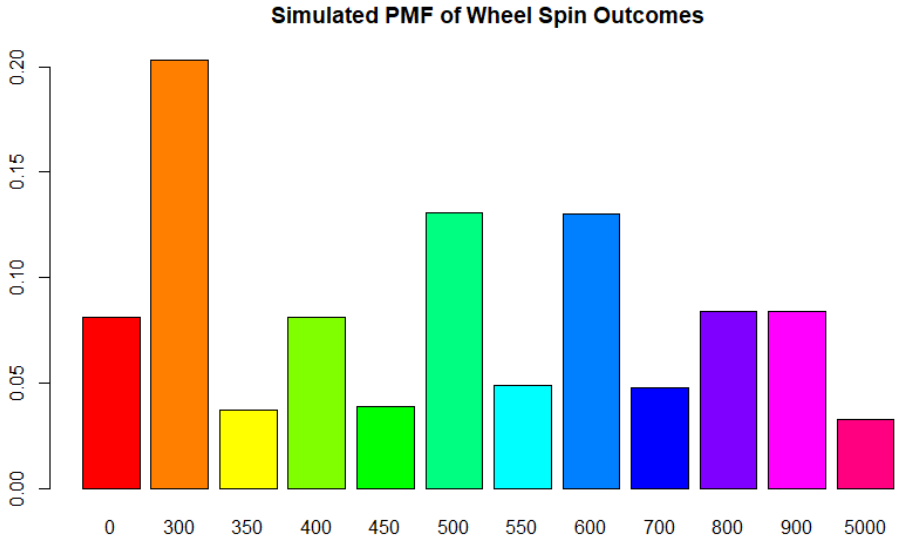
Part 3

a. $300

b. $664.4 It is very close since the calculated average is $670

c.

x	0	300	350	400	450	500	550	600	700	800	900	5000
p(x)	0.073	0.233	0.048	0.093	0.034	0.11	0.036	0.139	0.037	0.084	0.071	0.042



d. It is very close with a margin of error of about 5%

e. The highest bar on the plot is $300 so yes it is the same

f. As the number of spins approaches infinity, the simulated values will approach the theoretical values