Q1b: $\boxed{}\ \boxed{}\ \boxed{}\ \boxed{W}$

m1 t1 m2 t2

$$\frac{\binom{3}{2}\binom{17}{1}}{\binom{20}{3}} \quad \frac{1}{17}$$

Week 3 Lab

Question 1

Total # of balls: 17 + 3 = 20

An urn contains 17 balls marked LOSE and 3 balls marked WIN. You and an opponent take turns selecting at random a single ball from the urn without replacement.

The person who selects the third WIN ball wins the game. It does not matter who selected the first two WIN balls. Use R to calculate the probabilities asked in the following.

Note: The `choose(...)` *function, which takes two arguments* `n` *and* `k`*, will be useful here.*

(a)

(a) $\boxed{W}\ \boxed{W}\ \boxed{W}$

If you draw first, find the probability that you win the game on your second draw.

(b)

4th draw

If you draw first, find the probability that your opponent wins the game on his second draw.

(c)

k = 1, me win at 2nd draw

If you draw first, the probability that you win can be found from:

$$P(\text{You win if you draw first}) = \sum_{k=1}^{9} \frac{\binom{3}{2}\binom{17}{2k-2}}{\binom{20}{2k}} \times \frac{1}{20-2k}.$$

W W W

1 1

$$\underbrace{\frac{\binom{3}{2}\binom{17}{0}}{\binom{20}{2}}}_{\text{2 draws}} \cdot \underbrace{\frac{1}{18}}_{\text{my draw}}$$

Why is this?

Note: You could win on your second, third, fourth, ..., or tenth draw, not on your first.

My answer:

Now calculate this value:

(d)

If you draw second, the probability that you win can be found from:

$$P(\text{You win if you draw second}) = \sum_{k=1}^{9} \frac{\binom{3}{2}\binom{17}{2k-1}}{\binom{20}{2k+1}} \times \frac{1}{19-2k}$$

Why is this? k = 1, win on 2nd draw

My answer: k = 2, win on 3rd draw

Now calculate this value:

To understand the answer in (a) & (b). The first fraction describes how the first 2 draws take place:
- of the 3 WIN balls, 2 are selected
- of the 17 LOSE balls, 0 are selected
- 1/18 describes the chances of the 3rd draw being a win (of the 18 balls left)

(e)

Based on your results in (c) and (d), would you prefer to draw first or second?

My answer: draw 2nd, p is higher

Question 2

An urn contains n balls numbered from 1 to n. A random sample of n balls is selected from the urn, one at a time. A match occurs if ball numbered i is selected on the i^{th} draw.

If the draws are done **with replacement**, it can be shown that:

n: # of balls

$$P(\text{at least one match}) = 1 - \left(1 - \frac{1}{n}\right)^n.$$

If the draws are done **without replacement**, it can be shown that:

$$P(\text{at least one match}) = 1 - \left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^n}{n!}\right) = 1 - \sum_{k=0}^{n} \frac{(-1)^k}{k!}.$$

sum()

(a)

For each value of n given in the table at the bottom of this document, use R to find the probability of having at least one match, and write down the results in the appropriate entries of the table. The `factorial(...)` function will be useful here.

Note: To get the "without replacement" probability, you will need to define `k` *as the sequence of values* `k` *is summed over (that is,* `k <- 0:n`*), and use the* `sum(...)` *function — See Week 2 lab.*

Note: To calculate the theoretical probability for $n = \infty$, use the truncated value n $= 10^4$.

With replacement: 0.4605...

Without replacement: 0.6321...

(b)

R is widely used as a tool to conduct simulations of random experiments and processes. The idea is that simulated outcomes should closely match real-world outcomes, and by observing simulated outcomes, researchers can gain insight on the real-world processes. To perform a simulation, you outline a model of the process, and allow the computer to generate random trials of this model and record their results.

Theoretical probabilities can be compared with the results of a simulation to check for their correctness. In part (a), we calculated the theoretical probabilities of each ball-drawing process. We can also use R to simulate these processes, using the relevant relative frequencies to simulate the probability of at least one match.

Run the following code to create the function `match.f` in R:

```
match.f = function(n, simsize, rep){
  freq=0
  for(i in 1:simsize){
    sam=sample(1:n, size=n,replace=rep)
    # (sum(sam==1:n)>=1) checks whether or not there is at least 1 match in sam
    freq=freq + (sum(sam==1:n)>=1)
  }
```

```
  freq/simsize
}
```

This function will perform simulations of the random experiment. We just need to give it the value of each of the parameters: `n` (the number of balls), `simsize` (the size of the simulation), and `rep` (whether the draws are done with replacement).

Now we simulate the drawing process 1000 times (`simsize=1000`) for each given `n` and `rep` (`rep=TRUE` indicates the "with replacement" procedure is used).

The code chunk below displays the code for simulating 1000 scenarios for which there is 1 ball and replacement is done after a draw. Execute the following and write down the result in the appropriate entry of the table below. Then change the parameter values to complete the rest of the table (keep the number of simulations at 1000).

Note: To obtain simulation results for $n = \infty$, use the truncated value `n = 10^4`.

```
match.f(1, 1000, TRUE)
```

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

Note: The results obtained by simulations, being randomly generated, have no true answer. They should be approximately equal to the theoretical probabilities.

P(at least one match)

	with replacement	with replacement	without replacement	without replacement
n	Theoretical Prob.	Simulated Result	Theoretical Prob.	Simulated Result
1	1	1	1	1
3	0.704		0.667	0.685
10	0.651		0.632	0.641
15	0.645		0.632	0.627
100	0.634			0.61
∞	0.632		0.632	0.626