Homework #2

1. (10 points) The blood type among Americans is approximately distributed as: 37% type A, 13% type B, 44% type O, and 6% type AB. Suppose the blood types are distributed the same in both male and female populations. And assume that the blood types are independent of marriage.
(A) An individual with type B blood can safely receive transfusions only from persons with type B or type O blood. What is the probability of a husband has type B or type O blood? If a woman has type A blood, what is the probability that her husband is NOT an acceptable donor for her?
(B) What is the probability that, in a randomly chosen couple, the husband has type A blood and the wife has type AB blood?

2. (10 points) To overcome the difficulty in getting truthful answers on sensitive issues, "randomized response" is proposed. As an example, we ask the students whether they have gone to a high COVID-risk event. However, they are asked to toss a coin first in private (which has a 50% chance landing on heads and 50% chance landing on tails). The coin toss result is known only to the students themselves. If a "head" is tossed, they are to answer "Yes" regardless of whether they went or not. If a "tail" is tossed, they are to answer truthfully. This way, the surveyor has no way of telling whether a particular individual went to the event or not.
Suppose that in fact 20% of the students have gone to this event. What is the probability of a student answering "No" (that is, he/she did not go AND tossed a "tail")? What is the probability that a student did not go and answered "yes" (because a "head" toss)? What is the conditional probability of a student did not go given that he/she answered "yes"?

3. (10 points each) Exercises 11, 12, 17 in section 7.6.

4. (5 points) Use R to do this problem.
(a) Compute P(980 < Z < 1032.6) for the random variable Z following a Chi-squared distribution with degrees of freedom 1000.
(b) Use <u>a normal approximation</u> to calculate this probability.
(c) Include in the submission those R commands calculating probabilities in (a) and (b).

5. (10 points)
20 employees of a company drive to work. During the time period of 11am to 11:10am, each employee has a 6% chance of driving in front their office building, and they do this independent of each other. Other cars driving by during the same time period follows an independent Poisson distribution with mean 1. Let X and Y denote, respectively, the number of employee cars and other cars during this time period. W=X+Y denote the total number of cars driving by.

(a) Find the mean and variance of W.
(b) Find the probability that exactly W=1 car drove by.
(c) Use <u>the normal approximation</u> to calculate probability in part (b) instead.
(d) Submit R codes calculating the probabilities in parts (b) and (c).

R commands for computing probabilities

```
> ##### Example program to produce a table of CDF
> ##Set x values where to calculate the probability: 0,1,2,...,20
> probTable <- data.frame(x=(0:20))
> ## Calculate CDF using exact Binomial, then Poisson and Norm
al approximation.
> probTable$binprob = pbinom(probTable$x, size=1000, p=0.01)
> probTable$posprob = ppois(probTable$x, lambda=10)
> probTable$normprob = pnorm((probTable$x+0.5-10)/sqrt(9.9))
> ## Display results, round to 5 digits for better display
> round(probTable, digits=5)
    x binprob posprob normprob
1   0 0.00004 0.00005  0.00127
2   1 0.00048 0.00050  0.00345
3   2 0.00268 0.00277  0.00857
4   3 0.01007 0.01034  0.01942
5   4 0.02869 0.02925  0.04023
6   5 0.06614 0.06709  0.07633
7   6 0.12888 0.13014  0.13299
8   7 0.21886 0.22022  0.21344
9   8 0.33169 0.33282  0.31678
10  9 0.45730 0.45793  0.43687
11 10 0.58304 0.58304  0.56313
12 11 0.69735 0.69678  0.68322
13 12 0.79251 0.79156  0.78656
14 13 0.86556 0.86446  0.86701
15 14 0.91759 0.91654  0.92367
16 15 0.95213 0.95126  0.95977
17 16 0.97361 0.97296  0.98058
18 17 0.98617 0.98572  0.99143
19 18 0.99310 0.99281  0.99655
20 19 0.99671 0.99655  0.99873
21 20 0.99850 0.99841  0.9995
```