Name ____

Spring 2021 **STAT 587-3** **Exam I**

(26 points)

Instructions:

- You are allowed to use any resource except aid from another individual.
- Aid from another individual, will automatically earn you a 0.
- All problems are worth 1 point except the last problem that is worth 4 points.

1. Suppose the following table describes the joint probability that it will snow today and tomorrow.

	Tomorrow	
Today	Snow	No snow
Snow	0.2	0.1
No snow	0.3	

For example, the probability that it will snow today and not snow tomorrow is 0.1.

(a) Determine the probability that it will not snow today and it will not snow tomorrow.
Answer: Since the probabilities must sum to 1 we have $1 - 0.2 - 0.1 - 0.3 = 0.4$.

(b) Determine the marginal probability that it will snow tomorrow.
Answer: We need to sum the column for Snow Tomorrow, so $0.2 + 0.3 = 0.5$.

(c) Determine the conditional probability that it will snow tomorrow given that it snowed today. Answer: The conditional probability is the joint probability of snow today and tomorrow divided by the marginal probability that it will snow today, so $0.2/(0.2+0.1) = 0.67$.

(d) Are the events of snow today and snow tomorrow independent? Answer: No, because the joint probability of snow today and snow tomorrow is not equal to the product of the marginal probabilities of snow today and snow tomorrow, i.e. $0.2 \neq 0.5 * 0.3 = 0.15$.

2. Let $X \sim Unif(10, 20)$.

   Answer:

```
a = 10
b = 20
```

   (a) What is the image of X?

   Answer: $[10, 20]$

   (b) Determine $E[X]$.

   Answer:

```
(ex = (a+b)/2)

## [1] 15
```

   (c) Determine $Var[X]$.

   Answer:

```
(vx = (b-a)^2/12)

## [1] 8.333333
```

   (d) Determine $E[-2X + 3]$.

   Answer:

```
-2*ex+3

## [1] -27
```

   (e) Determine $Var[-2X + 3]$.

   Answer:

```
(-2)^2 * vx

## [1] 33.33333
```

3. For the following experiments, determine the best distribution to use to model the outcome.

   (a) Whether or not the Mars Perseverance Rover has a successful landing.
   Answer: Bernoulli/binomial

   (b) The price of Bitcoin at the end of 2021.
   Answer: normal

   (c) The number of traffic accidents that occur on I-35 in Iowa in February of 2021.
   Answer: Poisson

   (d) A manufacturing line produces semiconductor chips. Out of the next 1,000 chips, the number that fail quality control testing is recorded.
   Answer: binomial

   (e) The logarithm of salaries of 2021 graduates of any Iowa State University Engineering program.
   Answer: normal

4. Let $Y \sim Bin(60, 0.3)$.

   Answer:

```
n = 60
p = 0.3
```

   (a) Determine the image of Y.

   Answer:

```
0:n

##  [1]  0  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 2
## [26] 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 4
## [51] 50 51 52 53 54 55 56 57 58 59 60
```

   (b) Determine $E[Y]$.

   Answer:

```
n*p

## [1] 18
```

   (c) Determine $SD[Y]$.

   Answer:

```
sqrt(n*p*(1-p))

## [1] 3.549648
```

   (d) Determine $P(Y \leq 30)$.

   Answer:

```
pbinom(30, size = n, prob = p)

## [1] 0.999636
```

   (e) Determine $P(Y \geq 20)$.

   Answer: $P(Y \geq 20) = 1 - P(Y < 20) = 1 - P(Y \leq 19)$

```
1-pbinom(19, size = n, prob = p)

## [1] 0.3308409
```

5. A hydraulic fracturing oil well consumes, on average, 0.000015 km^3 of water per year with a standard deviation of 0.001 km^3. We would like to estimate how much water will be used by the 300,000 oil wells in 2021 (assuming pumping will continue as usual).

   Amounts were loosely based off `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4758395/`. The abstract for this article may help you answer the last question.

   Answer:

```
m = 0.000015
s = 0.001
n = 300000
```

   (a) Determine the expected amount of water that will be consumed next year by the 300,000 wells in km^3..

   Answer:

```
m*n

## [1] 4.5
```

   (b) Determine the standard deviation of the amount of water that will be consumed next year by the 300,000 wells in km^3.

   Answer:

```
sqrt(n*s^2)

## [1] 0.5477226
```

   (c) Determine the approximate probability that the 300,000 wells will use more than 5 km^3 of water.

   Answer: By the Central Limit Theorem, we have

$$S \dot{\sim} N(4.5, 0.2465^2)$$

```
1-pnorm(5, mean = n*m, sd = sqrt(n*s^2))

## [1] 0.1806552
```

   (d) Name reasons why you may not want to use a normal distribution to model the amount of water used by each well. (4 points)

   Answer: Here are some reasons you may not want to use a normal distribution:

   - water usage is strictly positive
   - standard deviation is much larger than the mean
   - different types of wells use very different amounts of water