Name ____

Spring 2019 **STAT 587C** **Exam I**

(100 points)

Instructions:

- Full credit will be given only if you show your work.
- The questions are not necessarily ordered from easiest to hardest.
- You are allowed to use any resource except aid from another individual.
- Aid from another individual, will automatically earn you a 0.

1. The State of California is out-sourcing its bridge crack detection program to Amazon Mechanical Turk. The process works like this 1) a citizen takes a picture of a bridge, 2) using a mobile app, they upload the picture to Amazon, and 3) a user in the Mechanical Turk program indicates whether or not the picture shows a crack in the bridge. Based on prior experience, the probability of the user indicating a crack when there really is a crack is 0.98, the probability of the user indicating no crack when there really is no crack is 0.90, and only 2 out of every 1000 pictures uploaded actually have cracks in bridges. Calculate the probability the bridge actually has a crack when the Mechnical Turk program user indicates it has a crack. (20 points)

   Answer: Let

   - C indicate the bridge has a crack
   - Y indicate the Mechanical Turk user indicates that "Yes, there is a crack."

   We are given the following probabilities

$$\begin{aligned} P(Y|C) &= 0.98 \\ P(Y^C|C^C) &= 0.90 \\ P(C) &= 2/1000 = 0.002. \end{aligned}$$

   We are interested in

$$\begin{aligned} P(C|Y) &= \frac{P(Y|C)P(C)}{P(Y|C)P(C)+P(Y|C^C)P(C^C)} \\ &= \frac{P(Y|C)P(C)}{P(Y|C)P(C)+[1-P(Y^C|C^C)][1-P(C)]} \\ &= \frac{0.98*0.002}{0.98*0.002+(1-0.90)*(1-0.002)} \end{aligned}$$

```
(0.98*0.002)/(0.98*0.002 + (1-0.90)*(1-0.002))

## [1] 0.01926101
```

2. Let X be a random variable with the following probability mass function:

x	1	2	3	4	5
P(X=x)	0.1	0.1	0.2	0.3	0.3

(a) Show that this is a valid probability mass function. (4 points)

Answer:

```
p <- c(0.1, 0.1, 0.2, 0.3, 0.3)
x <- 1:5
all(p>=0)

## [1] TRUE

sum(p)==1

## [1] TRUE
```

Thus this is a valid probability mass function.

(b) Determine the expected value of X. (4 points)

Answer:

```
(mu <- sum(p*x))

## [1] 3.6
```

(c) Determine the variance of X. (4 points)

Answer:

```
(vr <- sum(p*(x-mu)^2))

## [1] 1.64
```

(d) Let $Y = 2 \times X - 5$. Determine the expected value of Y. (4 points)

Answer: $E[Y] = 2 \times E[X] - 5 = 2 * 3.6 - 5 = 2.2$

(e) Let $Y = 2 \times X - 5$. Determine the variance of Y. (4 points)

Answer: $Var[Y] = 2^2 Var[X] = 4 \times 1.64 = 6.56$

3. A distributed censor network has 650 censors each with an 11% probability of failure. Assume censor failures are independent.

   Answer: Let Y be the number of censors that have failed. Then $Y \sim Bin(n, p)$ with

```
n <- 650
p <- 0.11
```

   (a) Calculate the expected number of censor failures. (5 points)

   Answer:

```
n*p

## [1] 71.5
```

   (b) Calculate the standard deviation of the number of censor failures. (5 points)

   Answer:

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

## [1] 7.977155
```

   (c) Calculate the probability of exactly 72 failures. (5 points)

   Answer: $P(Y = 72)$

```
dbinom(72,n,p)

## [1] 0.0497022
```

   (d) Calculate the probability that there are fewer than 100 failures. (5 points)

   Answer: $P(Y < 100) = P(Y \leq 99)$

```
pbinom(99,n,p)

## [1] 0.9995998
```

4. Electrical resistors indicate their expected resistance in ohms and tolerance around this resistance using a series of bands. The standard deviation of the resistance is the tolerance times the expected resistance. For example, if the bands indicate an expected resistance of 200 ohms and a tolerance of 10%, then the standard deviation is $200\times 0.1 = 20$ ohms. Variability in manufacturing means the actual resistance of a partiular resistor is normally distributed with a mean that is the expected resistance and standard deviation that can be calculated as described above.

   The following questions concern a particular resistor with bands that indicate an expected resistance of 350 ohms and a tolerance of 5%.

   (a) Calculate the standard deviation for resistors with these bands. (5 points)

   Answer: Let X be the actual resistance of the resistor and assume $X \sim N(\mu, \sigma^2)$ with

```
mu <- 350
tolerance <- 0.05
(sigma <- 350*tolerance)

## [1] 17.5
```

   Thus the standard deviation for resistors with these bands is 17.5 ohms.

   (b) Calculate the probability the actual resistance is greater than 400. (5 points)

   Answer:

```
1-pnorm(400, mean = mu, sd = sigma)

## [1] 0.002137367
```

   (c) Calculate the probability the actual resistance is within 20 ohms of the expected resistance. (5 points)

   Answer:

```
pnorm(mu + 20, mean = mu, sd = sigma) - pnorm(mu - 20, mean = mu, sd = sigma)

## [1] 0.7469021
```

   (d) Determine the tolerance needed so that the probability the actual resistance is within 10 ohms of the expected value is at least 90%. (5 points)

   Answer:

$$0.90 = P(|X-\mu| < 10) \implies 0.05 = P(X-\mu < -10) = P\left(\frac{X-\mu}{\sigma} < \frac{-10}{\sigma}\right) = P(Z < -10/\sigma)$$

```
qnorm(0.05)

## [1] -1.644854
```

   So $\sigma = -10/-1.6448536 = 6.0795683$ and thus the tolerance must be smaller than $\sigma/\mu = 1.7370195\%$.

5. The time it takes for a printer to complete a print job can be modeled as an exponential random variable. The distribution for an exponential random variable has a single parameter λ which is both the mean and standard deviation for the random variable. Suppose you sent a job to the single lab printer and there are 64 jobs in line before yours. Assume the mean time to completion for a job is 30 seconds and that completion time for each job is independent of all other jobs and is exponentially distributed.

   (a) Determine the approximate distribution for the time until your job starts based on the Central Limit Theorem. (5 points)

   Answer: Let X_i be the time for the ith job. Then $S = \sum_{i=1}^{64} X_i$ is the time until your job starts and, by the CLT, $S \dot{\sim} N(64 \times 30, 64 \times 30^2)$. (5 points)

```
n <- 64
(mu <- n*30)

## [1] 1920

(vr <- n*30^2)

## [1] 57600

(sigma <- sqrt(vr))

## [1] 240
```

   In minutes, this is $S \dot{\sim} N(32, 16)$.

   (b) Determine the probability it will take exactly 15 minutes for your job to start. (5 points)

   Answer: Since the probability for any continuous random variable being equal to a particular value is 0, this probability is 0.

   (c) Calculate the approximate probability that it will take more than 10 minutes for your job to start. (5 points)

   Answer:

```
1-pnorm(10*60, mean = mu, sd = sigma)

## [1] 1
```

   (d) The electricity costs while the printer is running is \$0.20/hour. Determine the expected costs to print the 64 jobs. (5 points)

   Answer: Expected cost in dollars is the cost times the expected time. In order for the units to cancel, we convert the expected time in seconds to expected time in hours.

```
0.20*mu/60^2

## [1] 0.1066667
```