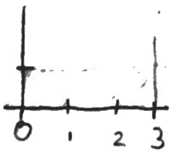
Math 183 HW7, Name: Chandler Blaid Burgess, **PID:** A98029477

1. Suppose that if you arrive at a bus stop at 8:00, the amount of time that you will have to wait for the next bus (in minutes) is uniformly distributed on $[0, 3]$.
   a) Find the probability that you will have to wait at least 1 minute for the bus.
   b) Find the probability that you will have to wait between 30 seconds and 2 minutes.
   c) Find the expected value and standard deviation of the amount of time that you have to wait for the next bus.

$\frac{1}{b-a} = \frac{1}{3}$



a) $P(X \geq 1)$?

$P(X \geq 1) = 1 - P(X < 1)$

$= 1 - f(0 \leq X < 1)$

$= 1 - \int_0^1 f(x)\,dx = 1 - \int_0^1 \frac{1}{3}\,dx = \frac{1}{3}x\Big|_0^1$

$\boxed{P(X \geq 1) = \frac{1}{3}}$

b) $P(0.5 < X < 2) = f(0.5 < X < 2)$

$P(0.5 < X < 2) = \int_{1/2}^{2} \frac{1}{3}\,dx = \frac{1}{3}x\Big|_{1/2}^{2} = \frac{2}{3} - \frac{1}{6} = \frac{3}{6}$

$\boxed{P(0.5 < X < 2) = 1/2}$

c) $E(X) = \frac{b-a}{2}$

$\boxed{E(X) = \frac{3}{2} = 1.5}$

$SD(X) = \sigma = \frac{b-a}{\sqrt{12}}$

$\sigma = \frac{3-0}{\sqrt{12}}$

$\boxed{\sigma = \frac{3}{\sqrt{12}} \approx 0.866}$

2. Suppose that the amount of time (in months) that a light bulb lasts before it burns out has an exponential distribution with parameter $\lambda = 1/4$.

a) What is the probability that it lasts at least three months?

b) If it has already lasted two months, what is the probability that it will last at least three more months?

c) On average, how many months will the light bulb last?

$$f(x) = \begin{cases} (1/4)e^{-(1/4)x}, & x \geq 0 \\ 0, & - \end{cases}$$

a) $P(X<3) = \int_0^3 \frac{1}{4e^{x/4}} dx = -e^{-(x/4)}\Big|_0^3 = -e^{-3/4} + e^0$

$P(X<3) = 1 - \frac{1}{e^{3/4}}$

$P(X \geq 3) = 1 - f(X<3)$

$= 1 - \left(1 - \frac{1}{e^{0.75}}\right) = \frac{1}{e^{0.75}}$

$\boxed{P(X \geq 3) \approx 0.529}$

b) $P(X \geq s+t \mid X \geq s) = P(X \geq t)$

$\boxed{\approx 0.528}$

c) $\boxed{E(X) = \frac{1}{\lambda} = \frac{1}{1/4} = 4 \text{ months}}$

3. Accidents occur at a busy intersection at the rate of three per year. What is the probability that it will be at least one year before the next accident at the intersection? Compute the answer using the following two methods:

a) Let X be the number of accidents in the next year. Find the distribution of X and calculate $P(X = 0)$.

b) Let T be the amount of time until the next accident. Find the distribution of T and calculate $P(T > 1)$.

a) $P(x) = \frac{\lambda^x \cdot e^{-\lambda}}{x!}$

$P(X=0) = \frac{3^0 \cdot e^{-3}}{3!}$

$\boxed{P(X=0) = \frac{1}{6e^3} \approx 0.008}$

b) $f(x) = \begin{cases} 3e^{-3x}, & x \geq 0 \\ 0 & - \end{cases}$

$P(T \leq 1) = \int_0^1 3e^{-3x} dx = -e^{-3x}\Big|_0^1$

$= -e^{-3} + 1$

$P(T > 1) = 1 - P(T \leq 1)$

$= \cancel{1} + e^{-3} \cancel{-1}$

$\boxed{P(T>1) = e^{-3} \approx 0.0498}$

3.1 Triathlon times, Part I In triathlons, it is common for racers to be placed into age and gender groups. Friends Leo and Mary both completed the Hermosa Beach Triathlon, where Leo competed in the *Men, Ages 30 - 34* group while Mary competed in the *Women, Ages 25 - 29* group. Leo completed the race in 1:22:28 (4948 seconds), while Mary completed the race in 1:31:53 (5513 seconds). Obviously Leo finished faster, but they are curious about how they did within their respective groups. Can you help them? Here is some information on the performance of their groups:

- The finishing times of the *Men, Ages 30 - 34* group has a mean of 4313 seconds with a standard deviation of 583 seconds.
- The finishing times of the *Women, Ages 25 - 29* group has a mean of 5261 seconds with a standard deviation of 807 seconds.
- The distributions of finishing times for both groups are approximately Normal.

Remember: a better performance corresponds to a faster finish.

(a) Write down the short-hand for these two normal distributions.
(b) What are the Z-scores for Leo's and Mary's finishing times? What do these Z-scores tell you?
(c) Did Leo or Mary rank better in their respective groups? Explain your reasoning.
(d) What percent of the triathletes did Leo finish faster than in his group?
(e) What percent of the triathletes did Mary finish faster than in her group?

a) $x_1 = 4948 \quad \mu_1 = 4313 \quad \sigma_1 = 583 \; : \; M = N(4313, 583)$

$x_2 = 5513 \quad \mu_2 = 5261 \quad \sigma_2 = 807 \; : \; W = N(5261, 807)$

b) $z_{x_1} = \dfrac{x_1 - \mu_1}{\sigma_1}$

$z_{x_1} = \dfrac{4948 - 4313}{583}$

$\boxed{z_{x_1} = \dfrac{635}{583} \approx 1.089}$

$z_{x_2} = \dfrac{x_2 - \mu_2}{\sigma_2}$

$z_{x_2} = \dfrac{5513 - 5261}{807}$

$\boxed{z_{x_2} = \dfrac{252}{807} \approx 0.312}$

The z-scores tell us that Leo had a higher time than Mary did w/ respect to their groups.

c) Mary ranked better than Leo w/in her group, because the z-scores tell us that Mary had a lower time than Leo w/in her group, thus she is ranked higher because she outperformed more people.

d) $P(z \leq 1.089) \approx 0.8621$

$P(z > 1.089) = 1 - P(z \leq 1.089) = 1 - 0.8621$

$\boxed{P(z > 1.089) \approx 0.1379 \Rightarrow 13.79\%}$

e) $P(z \leq 0.312) \approx 0.6217$

$P(z > 0.312) = 1 - P(z \leq 0.312)$

$= 1 - 0.6217$

$\boxed{P(z > 0.312 \approx 0.3783 \Rightarrow 37.83\%}$

3.8 CAPM. The Capital Asset Pricing Model (CAPM) is a financial model that assumes returns on a portfolio are normally distributed. Suppose a portfolio has an average annual return of 14.7% (i.e. an average gain of 14.7%) with a standard deviation of 33%. A return of 0% means the value of the portfolio doesn't change, a negative return means that the portfolio loses money, and a positive return means that the portfolio gains money.

(a) What percent of years does this portfolio lose money, i.e. have a return less than 0%?

(b) What is the cutoff for the highest 15% of annual returns with this portfolio?

a) $P(X<0)$ $\quad \mu = 14.7 \quad \sigma = 33 \quad z_0 = \frac{0-14.7}{33} \approx -0.45$

$\Downarrow$

$P(z < -0.45) = 0.3264$

$$\boxed{P(X<0) = P(z<-0.45) = 0.3264 \Rightarrow 32.64\%}$$

b) $P(X>y) = 0.15$

$P\left(z > \frac{y-\mu}{\sigma}\right) = 0.15$

$P\left(z < \frac{y-\mu}{\sigma}\right) = 1 - P\left(z > \frac{y-\mu}{\sigma}\right)$

$= 1 - 0.15 = 0.85$

$P\left(z < \frac{y-\mu}{\sigma}\right) = 0.85$

$\frac{y-\mu}{\sigma} = z^{-1}(0.85)$

$y-\mu = \sigma(z^{-1}(0.85))$

$y = \mu + \sigma(z^{-1}(0.85))$

$y = 14.7 + 33(1.04)$

$$\boxed{y = 49.02}$$

3.12 Speeding on the I-5, Part I. The distribution of passenger vehicle speeds traveling on the Interstate 5 Freeway (I-5) in California is nearly normal with a mean of 72.6 miles/hour and a standard deviation of 4.78 miles/hour.[4]

(a) What percent of passenger vehicles travel slower than 80 miles/hour?

(b) What percent of passenger vehicles travel between 60 and 80 miles/hour?

(c) How fast do the fastest 5% of passenger vehicles travel?

(d) The speed limit on this stretch of the I-5 is 70 miles/hour. Approximate what percentage of the passenger vehicles travel above the speed limit on this stretch of the I-5.

a) $P(X < 80\,\text{mph}) = P\left(Z < \frac{80-72.6}{4.78}\right)$

$\boxed{P(X<80) = P(Z<1.55) = 0.9394 \Rightarrow 93.94\%}$

b) $P(60 < X < 80) = P\left(\frac{60-72.6}{4.78} < Z < \frac{80-72.6}{4.78}\right)$

$P(60<X<80) = P(-2.64 < Z < 1.55)$

$P(60<X<80) = P(Z<1.55) - P(Z<-2.64)$

$= 0.9394 - 0.0041$

$\boxed{P(60<X<80) = 0.9353 \Rightarrow 93.53\%}$

c) $P(X > y) = 0.05$

$P(X<y) = 1 - P(X>y) = 0.95$

$P\left(Z < \frac{y-\mu}{\sigma}\right) = 0.95$

$\frac{y-\mu}{\sigma} = z^{-1}(0.95)$

$y = \mu + \sigma\left(z^{-1}(0.95)\right) = 72.6 + 4.78(1.65)$

$y = 80.487$ ∴ They must drive faster than 80.487 mph $\Rightarrow X > 80.487$

d) $P(X>70) = 1 - P(X<70)$ $\quad P(X<70) = P(Z<-0.54) = 0.2946$

$P(X>70) = 1 - 0.2946$

$\boxed{P(X>70) = 0.7054 \Rightarrow 70.54\%}$

3.14 Find the SD. Find the standard deviation of the distribution in the following situations.

(a) MENSA is an organization whose members have IQs in the top 2% of the population. IQs are normally distributed with mean 100, and the minimum IQ score required for admission to MENSA is 132.

(b) Cholesterol levels for women aged 20 to 34 follow an approximately normal distribution with mean 185 milligrams per deciliter (mg/dl). Women with cholesterol levels above 220 mg/dl are considered to have high cholesterol and about 18.5% of women fall into this category.

a) $z^{-1}(0.98) = \frac{132-100}{\sigma}$

$2.05 = \frac{32}{\sigma} \Rightarrow \boxed{\sigma = \frac{32}{2.05} = 15.61}$

b) $P(X>220) = 0.185$

$P(X<220) = 0.815$

$\frac{220-185}{\sigma} = z^{-1}(0.815)$

$\boxed{\sigma = \frac{35}{0.9} = 38.89}$

3.16 SAT scores. SAT scores (out of 2400) are distributed normally with a mean of 1500 and a standard deviation of 300. Suppose a school council awards a certificate of excellence to all students who score at least 1900 on the SAT, and suppose we pick one of the recognized students at random. What is the probability this student's score will be at least 2100? (The material covered in Section 2.2 would be useful for this question.)

$P(X>1900) = 1 - P(X<1900)$ $\quad$ $P(X<1900) = P(z<1.3\bar{3}) = 0.9082$

$P(X>1900) = 0.0918$

$$P(X>2100 \mid X>1900) = \frac{P(X>2100 \cap X>1900)}{P(X>1900)} = \frac{P(X>2100)}{P(X>1900)}$$

$$\boxed{P(X>2100 \mid X>1900) = \frac{0.0228}{0.0918} = 0.248 \approx 25\%}$$

$P(X>2100) = 1 - P(X<2100)$ $\quad$ $P(X<2100) = P(z<2) = 0.9772$

$P(X>2100) = 1 - 0.9772 = 0.0228$

R Question

The following problem comes from the website https://fivethirtyeight.com, which is a news site that believes in using rigorous data science and statistics when reporting the news (gasp!). Each week they have two interesting problems related to statistics and probability. This problem is from 1/13/2017:

"It's your 30th birthday (congrats, by the way), and your friends bought you a cake with 30 candles on it. You make a wish and try to blow them out. Every time you blow, you blow out a random number of candles between one and the number that remain, including one and that other number. How many times do you blow before all the candles are extinguished, on average?"

The answer, which is about 3.995 blows, can be found here. This uses probability theory to find the answer.

Your goal is to write a **simulation** of the candle-blowing experience. Then you'll run it 10,000 times, and each time record how many blows it took to extinguish all the candles. Average the blow counts together and you should get an answer around 3.995. **[code, average of 10,000 blow counts]**

Hints: You might want to use a while loop (Google this!), and the sample function can generate random integers if you use it correctly.

```
# Counter for the number of trials
n <- 0

# Empty vector to store the number of blows for each trial
toAvg <- vector(mode = "logical", 10000)

# Loop to perform 10,000 trials
while(n < 10000) {
  # Initial number of candles
  totalCandles <- 30
  blows <- 0

  # Loop until all of the candles are out
  while(totalCandles > 0) {
    # Provide a random number of candles blown out (between 1 and the remaining candles)
    candles <- sample(1:totalCandles, 1, replace = FALSE)

    # Subtract the candles just blown out from the total candles left
    totalCandles <- totalCandles - candles

    # accumulate the number of blows thus far
    blows = blows + 1
  }
  toAvg[n] <- blows
  n = n + 1
}

# Find the average number of blows across the 10,000 trials
s <- sum(toAvg)
avgBlows <- s / length(toAvg)

print(avgBlows)
```

```
[1]
3.9734
```