Homework 8

Zoe Farmer - 101446930

Kleiber

1	2	3	4	T

Chapter 5 #7 – The density function of X is given by

$$f(x) = \begin{cases} a + bx^2 & \to 0 \le x \le 1 \\ 0 & \to Otherwise \end{cases}$$

If $E(X) = \frac{3}{5}$, find a and b.

(a) We can rewrite $E(X)$ as the below and solve for a and b.

$$
\begin{aligned}
E(X) &= \int_0^1 x \left(a + bx^2\right) \, dx \\
&= \int_0^1 ax + bx^3 \, dx \\
&= \left. \frac{ax^2}{2} + \frac{bx^4}{4} \right|_0^1 \\
&= \frac{a}{2} + \frac{b}{4} = \frac{3}{5} \\
&= \frac{2a + b}{4} = \frac{3}{5} \\
&= 10a + 5b = 12 \\
10a &= 12 - 5b \Rightarrow a = \frac{12 - 5b}{10} \\
1 &= \int_0^1 \frac{12 - 5b}{10} + bx^2 \, dx \\
&\boxed{b = \frac{6}{5} \qquad a = \frac{3}{5}}
\end{aligned}
$$

Chapter 5 # 10 – Trains headed for destination A arrive at the train station at 15-minute intervals starting at 7 AM, whereas trains headed for destination B arrive at 15-minute intervals starting at 7:05 AM.

(a) If a certain passenger arrives at the station at a time uniformly distributed between 7 and 8 AM and then gets on the first train that arrives, what proportion of time does he or she go to destination A?

    i. We can use the definition of a uniform distribution and find that the probability that a passenger arrives at any given time is $1/60$. Train A arrives at $\{0700, 0715, 0730, 0745, 0800\}$

1

while train B arrives at $\{0705, 0720, 0735, 0750\}$. The probability that the passenger gets on any one train can be expressed as

$$A \rightarrow \{\frac{1}{60}, \frac{1}{60}, \frac{1}{60}, \frac{1}{60}, \frac{1}{60}\}$$
$$B \rightarrow \{\frac{1}{60}, \frac{1}{60}, \frac{1}{60}, \frac{1}{60}\}$$

Now to find the total probability of either destination, we can simply sum the sets and obtain the probability that he boards train A to be $\frac{1}{12}$ and the probability that he boards train B to be $\frac{1}{15}$. He will board train A more often.

(b) What if the passenger arrives at a time uniformly distributed between 7:10 and 8:10 AM?

    i. Now our sets adjust slightly.

$$A \rightarrow \{0700, 0715, 0730, 0745, 0800\}$$
$$B \rightarrow \{0705, 0720, 0735, 0750, 0805\}$$

Rewriting with probabilities

$$A \rightarrow \{\frac{1}{60}, \frac{1}{60}, \frac{1}{60}, \frac{1}{60}, \frac{1}{60}\}$$
$$B \rightarrow \{\frac{1}{60}, \frac{1}{60}, \frac{1}{60}, \frac{1}{60}, \frac{1}{60}\}$$

And we see that the probabilities are the same, $\frac{1}{12}$.

Chapter 5 # 11 – A point is chosen at random on a line segment of length L. Interpret this statement, and find the probability that the ratio of the shorter to the longer segment is less than $\frac{1}{4}$.

(a) If the probability that any point is chosen is equal, than we have a uniform distribution, and the probability that any given point being chosen is $\frac{1}{L}$. Let X the ratio of the shorter to longer segments. We're interested in $P(X < 0.25) = 1 - P(X \geq 0.25)$.

Let's first find the probability that any given point is chosen.

$$p(k; L) = \begin{cases} \frac{1}{L} & \rightarrow 0 \leq k \leq L \\ 0 & \rightarrow Otherwise \end{cases}$$

Now let's look at the probability that the ratio is equal to any given number.

$$f(k; L) = \begin{cases} p(k \cdot L; L) & \rightarrow k \geq 0 \\ 0 & \rightarrow Otherwise \end{cases}$$

Finally we can determine

$$1 - P(X \geq 0.25) = 1 - \int_{0.25}^{L} \frac{1}{L} dx = \boxed{1 - \frac{1}{L} \cdot (L - 0.25)}$$

Chapter 5 # 13 – You arrive at a bus stop at 10 o'clock, knowing that the bus will arrive at some time uniformly distributed between 10 and 10:30.

(a) What is the probability that you will have to wait longer than 10 minutes?

    i. If the arrival time is uniformly distributed, than we can assume that the mean time is 10:15, and that we can use the definition of the pdf of a uniform distribution, which states that a distribution is uniform if the pdf is equal to

$$f(k; A, B) = \begin{cases} \frac{1}{B-A} & \to A \le k \le B \\ 0 & \to Otherwise \end{cases} = \begin{cases} \frac{1}{30} & \to 0 \le k \le 30 \\ 0 & \to Otherwise \end{cases}$$

Using this allows us to determine the cdf

$$F(k; A, B) = \begin{cases} 0 & \to x < A \\ \frac{x-A}{B-A} & \to A \le k < B \\ 1 & \to x \ge B \end{cases} = \begin{cases} 0 & \to x < 0 \\ \frac{x}{30} & \to 0 \le k < 30 \\ 1 & \to x \ge 30 \end{cases}$$

By which we can determine $P(X > 10) = 1 - P(X \le 10) = 1 - \frac{1}{3} = \boxed{0.66667}$ where X is the number of minutes needed to wait.

(b) If, at 10:15, the bus has not yet arrived, what is the probability that you will have to wait at least an additional 10 minutes?

   i. We've now changed the time period, and thusly we need to adjust the probability distributions.

$$F(k; A, B) = \begin{cases} 0 & \to x < 0 \\ \frac{x}{15} & \to 0 \le k < 15 \\ 1 & \to x \ge 15 \end{cases}$$

By which we can new determine $P(X \ge 10) = \frac{2}{3} = \boxed{0.66667}$.

Chapter 5 # 16 – The annual rainfall (in inches) in a certain region is normally distributed with $\mu = 40$ and $\sigma = 4$. What is the probability that, starting with this year, it will take over 10 years before a year occurs having a rainfall of over 50 inches? What assumptions are you making?

(a) Recall that a normally distributed random variable has the form

$$f(x, \mu = 40, \sigma = 4) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{4\sqrt{2\pi}} e^{-\frac{(x-40)^2}{32}}$$

If we let X be how many inches of rain the region receives, we are then interested in $P(X > 50) = 1 - P(X \le 50)$, which we can write as such.

$$1 - \int_0^{50} \frac{1}{4\sqrt{2\pi}} e^{-\frac{(x-40)^2}{32}} \, dx = 0.00620967$$

This number is the probability that any given year has more than 50 inches of rain. If we now let Y be the random variable denoting how many years it is until 50 inches is received, then we can create a new Geometric distribution with the probability of success being our previously calculated probability.

$$f(k; p = 0.00620967) = (1-p)^{k-1} p = (1 - 0.00620967)^{k-1} 0.00620967$$

We now find $P(Y > 10) = 1 - P(Y \le 10)$ to be $\boxed{0.933775}$. This is with the assumption that each year's rainfall is independent of the previous years.

Chapter 5 # 18 – Suppose that X is a normal random variable with mean 5. If $P\{X > 9\} = 0.2$, approximately what is $Var(X)$?

(a) We can rewrite the given probability to that in terms of a cdf.

$$P(X > 9) = 1 - P(X \le 9) = 1 - \int_{-\infty}^{9} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-5)^2}{2\sigma^2}} \, dx = \qquad 0.2$$

$$\int_{-\infty}^{9} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-5)^2}{2\sigma^2}} \, dx = \qquad 0.8$$

$$\sigma = \quad \boxed{4.75273}$$

Chapter 5 # 23 – One thousand independent rolls of a fair die will be made. Compute an approximation to the probability that the number 6 will appear between 150 and 200 times inclusively. If the number 6 appears exactly 200 times, find the probability that the number 5 will appear less than 150 times.

(a) We can model this as a binomial distribution, where the probability of success is 1/6 and the number of "flips" is 1000.

Histogram of rbinom(1e+06, 1000, 1/6)
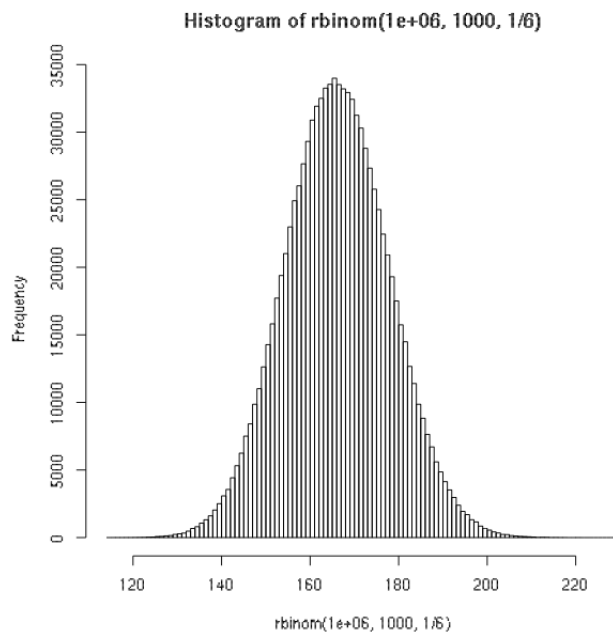
Which can be rewritten and solved for the appearance of the number 6 between 150 and 200 times.

$$\sum_{k=150}^{200} \binom{1000}{k} (1/6)^k 1 - (1/6)^{1000-k} \, dk \approx 0.9264531$$

If the number 6 appears *exactly* 200 times than we can rewrite the problem to examine 5 instead.

$$\sum_{k=0}^{149} \binom{800}{k} (1/5)^k 1 - (1/5)^{800-k} \, dk \approx 0.17698455$$

Chapter 5 # 26 – Two types of coins are produced at a factory: a fair coin and a biased one that comes up heads 55 percent of the time. We have one of these coins, but do not know whether it is a fair coin or a biased one. In order to ascertain which type of coin we have, we shall perform the following statistical test: We shall toss the coin 1000 times. If the coin lands on heads 525 or more times, then we shall conclude

that it is a biased coin, whereas if it lands on heads less than 525 times, then we shall conclude that it is a fair coin. If the coin is actually fair, what is the probability that we shall reach a false conclusion? What would it be if the coin were biased?

(a) We can model this as a binomial distribution:

Histogram of rbinom(n = 1e+06, size = 1000, prob = 0.5)
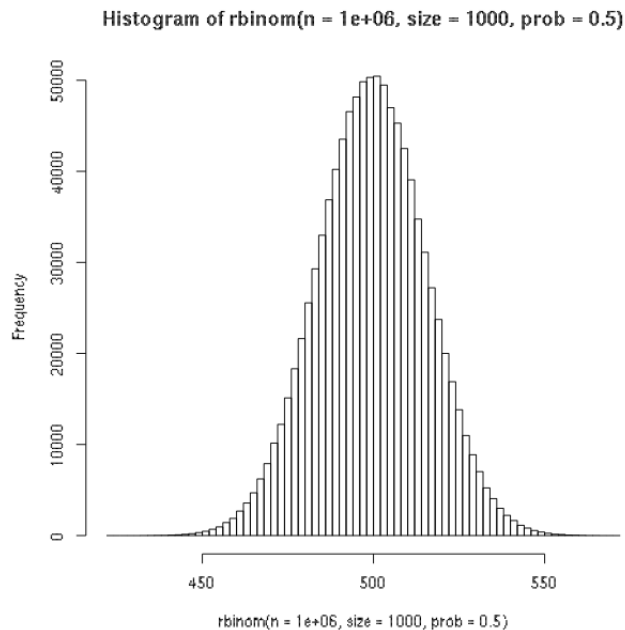
Whereupon we see that not only is the binomial distribution uniform, it also tells us the probability that a coin will yield 525 or more heads as

$$\sum_{k=525}^{\infty} \binom{1000}{k} 0.5^k 0.5^{1000-k} \, dk = 0.0606071$$

This number is the probability that any given coin will have 525 or more heads. This means that if we are given a fair coin, $\approx 5\%$ of the time we will incorrectly categorize the coin. On the flipside, if our coin is biased our probability shifts from 0.5 to 0.55, and our equation becomes

$$\sum_{k=525}^{\infty} \binom{1000}{k} 0.55^k 0.55^{1000-k} \, dk = 0.9473182$$

And we see that we catch biased coins 95% of the time.

Chapter 5 # 27 – In 10,000 independent tosses of a coin, the coin landed on heads 5800 times. Is it reasonable to assume that the coin is not fair? Explain.

(a) In this situation, if someone was bored enough to flip a coin 10,000 times which resulted in 58% heads, it would be a reasonable assumption to assume that the coin is not fair. Sure, there is a margin of error in this assumption due to the fact that the coin is supposed to be entirely random, and thereby you cannot be 100% sure that it is an unfair coin, however this is enough trial to see a consistent trend.

Chapter 5 # 7 – The standard deviation of X, denoted $SD(X)$, is given by

$$SD(X) = \sqrt{Var(X)}$$

Find $SD(aX + b)$ if X has variance σ^2.

(a) We can simply use the definition of standard deviation. The standard deviation of $aX + b$ is simply $a \cdot \sigma$ where σ is the standard deviation of X, defined as the square root of the variance.

Chapter 5 # 9 – Show that Z is a standard normal random variable, then, for $x > 0$,

(a) $P(Z > x) = P(Z < -x)$

    i. The definition we need here is that of the cdf of a continuous variable, which is defined as

$$F(x) = \int_{-\infty}^{x} f(y)\, dy$$

    Using this we can rewrite this as the below equation and solve.

$$1 - \int_{-\infty}^{x} f(y)\, dy = \quad 1 - \int_{-x}^{\infty} f(y)\, dy$$
$$\int_{-\infty}^{x} f(y)\, dy = \quad \int_{-x}^{\infty} f(y)\, dy$$

    Since we are working with a uniform distribution, then these two parts are equal.

(b) $P(|Z| > x) = 2P(Z > x)$

    i. Using our definitions from before it is trivial to see that since we have a uniform distribution, everything on the left side of the mean is a mirror of the right side of the mean. This means that if we look at the absolute value of Z, we are in fact looking at the two equal end parts, which is equal to twice one.

(c) $P(|Z| < x) = 2P(Z < x) - 1$

    i. Again, we can use our previous definitions to prove this. Similar to (b) we see that the absolute value indicates a doubled probability, however in this case we are concerned about the space contained, because it is much larger, therefore we need to pull out the middle part, which has probability 1.

Chapter 5 # 10 – Let $f(x)$ denote the probability density function of a normal random variable with mean μ and variance σ^2. Show that $\mu - \sigma$ and $\mu + \sigma$ are points of inflection of this function. That is, show that $f(x) = 0$ when $x = \mu - \sigma$ or $x = \mu + \sigma$.

(a) A normal distribution has form

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

    With first derivative

$$\frac{df(x, \mu, \sigma)}{dx} = -\frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}(x - \mu)}{\sqrt{2\pi}\sigma^3}$$

    And second derivative

$$\frac{d}{dx} = \frac{(x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma^5} - \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\sigma^3}$$

    Now if we simply plug in $\mu \pm \sigma$ we can see that the second derivative is equal to zero, which by definition makes these two points inflection points.