

Homework 03 - STAT416

Joseph Sepich (jps6444)

09/22/2020

Chapter 2 Problem 72

Part a

Here we have 5 independent and identically distributed Normal random variables with a mean of 100 and variance of 100. To find the probability that at least one month of sales exceeds 115 we can find the probability that no months exceed 115 and subtract that value from one. The probability that no months exceed 115 is the value of the CDF at 115 multiplied 5 times:

$$P(X_1, X_2, X_3, X_4, X_5 \leq 115) = \prod_{i=1}^5 P(X_i \leq 115)$$

We can use the standard normal CDF if we transform the distribution.

$$P(X_i \leq 115) = P\left(\frac{X_i - \mu}{\sigma} \leq \frac{115 - 100}{10}\right) = P(Z \leq 1.5) = \Phi(1.5) = 0.9332$$

$$\prod_{i=1}^5 P(X_i \leq 115) = (0.9332)^5 = 0.7077$$

This gives us our desired value of $1 - 0.7077 = 0.2923$. The probability that at least one month exceeds 115 in sales is **29.23%**.

Part b

The probability that the total sales exceeds 530 in the next five months requires us to define a new random variable Y . $Y = \sum_{i=1}^5 X_i$, which gives a mean of $\sum_{i=1}^5 \mu_{X_i} = 500$ and a variance of $\sum_{i=1}^5 \sigma_{X_i}^2 = 500$. We can then search for:

$$1 - P(Y \leq 530) = 1 - P\left(\frac{Y - 500}{\sqrt{500}} \leq \frac{530 - 500}{\sqrt{500}}\right) = 1 - \Phi(1.34) = 1 - 0.9099 = 0.0901$$

The probability that total sales exceeds 530, $P(Y > 530)$, is **9.01%**.

Chapter 2 Problem 78

Part a

Recall the Markov inequality:

$$P(X \geq a) \leq \frac{E[x]}{a}$$

We also know that the sum of independent Poisson R.V.s is a Poisson R.V. whose mean is the sum of the individual means:

$$Y = \sum_{i=1}^{10} X_i = \text{Poisson}(\sum_{i=1}^{10} \lambda_i)$$

Here we can conclude the bound is:

$$P(Y \geq 15) \leq \frac{\sum_{i=1}^{10} \lambda_i}{15} = \frac{10}{15}$$

Part b

We can use the central limit theorem to approximate $P(\sum_{i=1}^{10} X_i \geq 15)$. Using the central theorem we will approximate by assuming that the sample mean is normally distributed. We make sure to transform the value to the standard normal and then use the CDF to determine the probability.

$$P(\sum_{i=1}^{10} X_i \geq 15) = 1 - P\left(\frac{\sum_{i=1}^{10} X_i - 10}{\sqrt{10}} \leq \frac{15 - 10}{\sqrt{10}}\right) = 1 - \Phi(1.58) = 1 - 0.9429 = 0.0571$$

Using CLT to approximate the probability that the sum is greater than 15, we get that the probability is **5.71%**, which is less than our bound of 66.67%.

Chapter 3 Problem 3

Compute $E[X|Y = i]$ for $i = 1, 2, 3$. Recall the formula for the conditional expectation:

$$E[X|Y = i] = \sum_x x P(X|Y = i)$$

For $i = 1$:

$$E[X|Y = 1] = 1P(1|Y = 1) + 2P(2|Y = 1) + 3P(3|Y = 1) = \frac{1}{5} + \frac{6}{5} + \frac{3}{5} = 2$$

For $i = 2$:

$$E[X|Y = 2] = 1P(1|Y = 2) + 2P(2|Y = 2) + 3P(3|Y = 2) = \frac{2}{3} + 0 + \frac{3}{3} = \frac{5}{3}$$

For $i = 3$:

$$E[X|Y = 3] = 1P(1|Y = 3) + 2P(2|Y = 3) + 3P(3|Y = 3) = 0 + \frac{6}{5} + \frac{6}{5} = \frac{12}{5}$$

Chapter 3 Problem 7

$$E[X|Y = 2] = 1P(1|Y = 2) + 2P(2|Y = 2) = 1(\frac{1}{5} + 0) + 2(0 + \frac{4}{5}) = \frac{1}{5} + \frac{8}{5} = \frac{9}{5}$$

$$E[X|Y = 2, Z = 1] = 1P(1|Y = 2, Z = 1) + 2P(2|Y = 2, Z = 1) = 1 + 0 = 1$$

Chapter 3 Problem 11

Here we are finding $E[X|Y = y]$. We can use the equation for expectation of a continuous random variable:

$$E[X|Y = y] = \int x f_{X|Y}(X|Y = y) dx$$

First we need to find the conditional pdf $f_{X|Y}(X|Y = y)$. $f_{X|Y}(X|Y = y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$.

$$f_Y(y) = \int_{-y}^y \frac{e^{-y}(y^2 - x^2)}{8} dx = \frac{e^{-y}}{8} (xy^2 - \frac{1}{3}x^3)|_{-y}^y$$
$$f_Y(y) = \frac{e^{-y}}{8} ((y^3 - \frac{1}{3}y^3) - (-y^3 + \frac{1}{3}y^3)) = \frac{e^{-y}}{6} y^3$$

Plug this back in to get the conditional pdf:

$$f_{X|Y}(X|Y = y) = \frac{\frac{e^{-y}(y^2 - x^2)}{8}}{\frac{e^{-y}}{6} y^3} = \frac{3(y^2 - x^2)}{4y^3} = \frac{3}{4}y^{-1} - \frac{3}{4}x^2y^{-3}$$

Now we can find the conditional expectation:

$$\int_{-y}^y x f_{X|Y}(X|Y = y) dx = \int_{-y}^y (\frac{3}{4}xy^{-1} - \frac{3}{4}x^3y^{-3}) dx = \frac{3}{4}(\frac{x^2}{2y} - \frac{x^4}{4y^3})|_{-y}^y = \frac{3}{4}((\frac{y}{2} - \frac{y}{4}) - (\frac{y}{2} - \frac{y}{4})) = 0$$

Thus we confirm that the conditional expectation $E[X|Y = y] = 0$.

Chapter 3 Problem 15

Here we are finding $E[X^2|Y = y]$. We can use the equation for the second moment of a continuous random variable:

$$E[X^2|Y = y] = \int x^2 f_{X|Y}(X|Y = y) dx$$

First we need to find the conditional pdf $f_{X|Y}(X|Y = y)$. $f_{X|Y}(X|Y = y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$.

$$f_Y(y) = \int_0^y \frac{e^{-y}}{y} dx = \frac{e^{-y}}{y} x|_0^y = e^{-y}$$

Plug this back in to get the conditional pdf:

$$f_{X|Y}(X|Y=y) = \frac{\frac{e^{-y}}{y}}{e^{-y}} = \frac{1}{y}$$

Now we can find the conditional second moment:

$$\int_0^y x^2 f_{X|Y}(X|Y=y) dx = \int_0^y \frac{x^2}{y} dx = \frac{x^3}{3y} \Big|_0^y = \frac{y^2}{3}$$

Chapter 3 Problem 23

We are trying to find $E[N]$ where N is the number of flips until 2 out of three most recent coin flips are heads. We know that we start with 2 flips, then we can condition on the result of these two flips:

$$E[N] = 2 + E[N|H, H]P(H, H) + E[N|H, T]P(H, T) + E[N|T, H]P(T, H) + E[N|T, T]P(T, T)$$

We must define what the expectations are. First we have $E[N|H, H]$. The expected number of further flips then would be 0, because two heads have already been flipped. Next we have $E[N|T, T]$, which is $E[N]$, since it would essentially be like starting the coin flips over, since no progress towards flipping heads has been made.

Next let's look at the two more complicated values. $E[N|H, T]$ and $E[N|T, H]$ show progress toward our goal of two heads in the last three flips. For $E[N|H, T]$ we add 1 flip more. If that next flip is heads, then we met our goal, otherwise we reset. This gives us $E[N|H, T] = 1 + 0P(H) + E[N]P(T)$. For $E[N|T, H]$ we have the same 1 flip more, that either results in meeting the goal of 2 heads, or results in the previous expectation of a H, T . $E[N|T, H] = 1 + 0P(H) + E[N|H, T]P(T)$. Let's plug these values back into the main equation:

$$E[N] = 2 + 0p^2 + (1 + 0P(H) + E[N]P(T))p(1-p) + (1 + 0P(H) + (1 + 0P(H) + E[N]P(T))P(T))p(1-p) + E[N]p(1-p)$$

$$E[N] = 2 + (1 + E[N](1-p))p(1-p) + (1 + (1 + E[N](1-p))(1-p))p(1-p) + E[N](1-p)$$

$$E[N] = 2 + p(1-p) + E[N]p(1-p)^2 + p(1-p) + p(1-p)^2 + E[N]p(1-p)^3 + E[N](1-p)^2$$

$$E[N](1-p(1-p)^2 - (1-p)^2 - p(1-p)^3) = 2 + 2p(1-p) + p(1-p)^2$$

$$E[N] = \frac{2 + 2p(1-p) + p(1-p)^2}{1 - p(1-p)^2 - (1-p)^2 - p(1-p)^3}$$

Chapter 3 Problem 24

Part a

We are trying to find $E[N]$ when flipping a coin with probability p of landing on heads. We will flip until at least one head and one tail have been flipped. We start at two flips minimum:

$$E[N] = 1 + E[N|H]P(H) + E[N|T]P(T)$$

So let's look at our next flip for each. If we get heads first:

$$E[N|H] = 1 + E[N|H]P(H) + 0P(T)$$

$$E[N|H] = 1 + (1 + E[N|H]P(H))P(H)$$

So this expectation becomes an infinite geometric series:

$$E[N|H] = 1 + P(H) + P(H)^2 \dots = \frac{1}{1 - P(H)}$$

Now let's look at if we get tails first:

$$E[N|T] = 1 + 0P(H) + E[N|T]P(T)$$

This looks exactly like the previous expectation for $E[N|H]$, so mirroring that we get:

$$E[N|T] = \frac{1}{1 - P(T)}$$

Now let's plug these back in for $P(H) = p$ and $P(T) = 1 - p$:

$$\begin{aligned} E[N] &= 1 + \left(\frac{1}{1-p}\right)p + \frac{1}{p}(1-p) \\ E[N] &= 1 + \frac{p^2}{(1-p)p} + \frac{(1-p)^2}{p(1-p)} \\ E[N] &= 1 + \frac{p^2 + (1-p)^2}{p(1-p)} \end{aligned}$$

Part b

To get the expected number of heads we need to consider two cases. The first case is that a tail is flipped first. If this occurs, then we only expect there to be 1 head, since we only need at least one tail and one head to reach our goal; however if a head is flipped first then we will keep flipped until we do not have a head:

$$E[N] = 1P(T) + E[N|H]P(H)$$

We know that $E[N|H] = 1 + 0P(T) + E[N|H]P(H)$. We can see that this also follows the pattern of $E[N|H]$ in part a, so we know it forms a geometric series resulting in $E[N|H] = \frac{1}{1-p}$. Plugging this back in we get:

$$E[N] = 1 - p + p\left(\frac{1}{1-p}\right) = 1 - p + \frac{p}{1-p}$$

Part c

The expected number of flips that land on tails is the same as heads, but we swap the probabilities.

$$E[N] = p + (1-p)\frac{1}{p} = p + \frac{1-p}{p}$$

Part d

$$E[N] = 2 + E[N|H, H]P(H, H) + E[N|H, T]P(H, T) + E[N|T, H]P(T, H) + E[N|T, T]P(T, T)$$

$E[N|H, H] = 1 + 0P(T) + E[N|H]P(H)$, since we just need to know how many more heads we expect to get before our single tails. We will be able to replace the value of $E[N|H]$ with that found in part a. Both $E[N|H, T]$ and $E[N|T, H]$ need just one more heads, so we see $1 + 0P(H) + E[N|T]P(T)$. We will be able to replace the value of $E[N|T]$ with the value of that found in part a. If both of the first two flips are heads then we need to find two more heads:

$$E[N|T, T] = 1 + E[N|H]P(H) + E[N|T, T]P(T)$$

This will become another geometric series where $a = E[N|H]P(H)$ and $r = P(T)$, which we know the value from part a. Putting all this together we get:

$$E[N] = 2 + \frac{p^2}{1-p} + \frac{2p(1-p)}{p} + \frac{p(1-p)^2}{(1-p)p}$$

$$E[N] = 2 + \frac{p^3 + 3p(1-p)^2}{p(1-p)}$$

Chapter 3 Problem 26

$$E[N_A] = E[N_A|\text{win}]p_A + E[N_A|\text{lose}](1 - p_A)$$

We know that if you lose with A, then you essentially start anew with B $E[N_A|\text{lose}] = 1 + E[N_B]$. On the other hand if you win you could either win again or lose. If you win again, then you are done, but otherwise you start over:

$$E[N_A|\text{win}] = E[N_A|\text{win}, \text{lose}](1 - p_B) + 2p_B$$

$$E[N_A|\text{win}] = (2 + E[N_A])(1 - p_B) + 2p_B$$

$$E[N_A] = 1 + p_A + (p_A - p_A p_B)E[N_A] + (1 - p_A)E[N_B]$$

$$E[N_B] = 1 + p_B + (p_B - p_A p_B)E[N_B] + (1 - p_B)E[N_A]$$

$$E[N_A] - E[N_B] = p_A - p_B + (1 - p_A - p_B + p_A p_B)E[N_B] + (p_B - 1 + p_A - p_A p_B)E[N_A]$$

$$E[N_A] - E[N_B] = p_A - p_B + (p_A - p_A p_B - 1 + p_B)(E[N_A] - E[N_B])$$

$$(-p_A + p_A p_B + 2 - p_B)(E[N_A] - E[N_B]) = p_A - p_B$$

Here we know that the term on the left must be positive, since the probabilities are each less than 1. Since the probability of B is greater than A we know this difference $E[N_A] - E[N_B]$ will end up being negative, since the term on the right is $p_A - p_B$. A negative difference implies that the expected number of games played starting with B will be higher than the expected number of games starting with A, therefore we should start by playing player A first.

Problem A

Problem Constraints:

- $\mu = 5$
- $\sigma = 2$
- $n = 50$

Part a

Here we are approximating the probability that it will take more than 260 minutes to process all 50 items. The sample mean ΣX_i is a random variable that can be approximated via the Normal distribution given the Central Limit theorem.

$$P(\Sigma_{i=1}^{50} X_i > 260) = 1 - P\left(\frac{\Sigma_{i=1}^{50} X_i - 250}{\sqrt{100}} \leq \frac{260 - 250}{\sqrt{100}}\right) = 1 - \Phi(1) = 1 - 0.8413 = 0.1587$$

There is a **15.87%** probability that it will take more than 260 minutes to process all 50 items.

Part b

The probability that at least 30 items will be processed in the first 140 minutes is the same as saying the probability it will take us less or equal to 140 minutes to process 30.

$$P(\Sigma_{i=1}^{30} X_i \leq 140) = P\left(\frac{\Sigma_{i=1}^{30} X_i - 150}{\sqrt{60}} \leq \frac{140 - 150}{\sqrt{60}}\right) = \Phi(-1.29) = 0.0985$$

There is a **9.85%** probability that 30 or more items will be process in the first 140 minutes.