Mandatory 5

9. april 2019 12:52

Exercise 1

(i) A light bulb factory relies on two machines, machine A and machine B. Machine A produces 65% of the stock and machine B the remaining 35%. Some light bulbs have defects: after coming out of machine A, 8% of light bulbs have deffects, while after coming out of machine B, 5% of light bulbs have deffects. We pick a random light bulb from the stock. Given that it does not present any defect, what is the probability that it came out of machine A?

Events

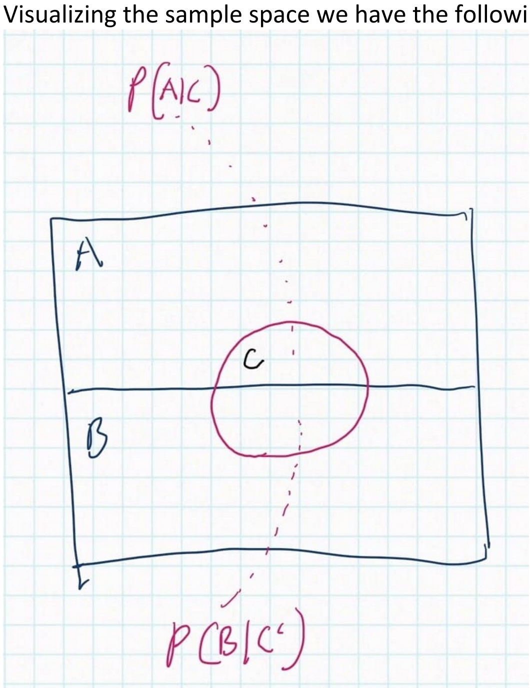
$A = Bulb\ is\ from\ machine\ A$
$B = Bulb\ is\ from\ machine\ B$
$C = defect\ bulb$

Probabilities

$P(A) = 0.65$
$P(B) = 0.35$

$P(C|A) = 0.08$
$P(C|B) = 0.05$
$P(C^c|A) = 0.92$
$P(C^c|B) = 0.95$

Visualizing the sample space we have the following Venn diagram



We need to calculate the probability that a random picked lightbulb is from Machine A and not defect.
Using conditional probability we can write the probability we need to solve
$P(from\ A \mid not\ defect) = P(A|C^c)$

Recall Bayes' rule

$$P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{P(A) * P(C|A)}{P(A_1)P(B|A_1) + \cdots + P(A_n)P(B|A_n)}$$

Which combines the chain rule and total probability theorem to calculate a conditional probability.

And so we can use Bayes' rule to calculate $\mathrm{P}(A|C^c)$

$P(C^c|A) = 1 - P(C|A) = 0.92$
$P(C^c|A) = 1 - P(C|B) = 0.95$
$A_1 = A$
$A_2 = B$

$$\mathrm{P}(A|C^c) = \frac{P(A) * P(C^c|A)}{P(A) * P(C^c|A) + P(B)P(C^c|B)}$$

$$\mathrm{P}(A|C^c) = \frac{0.65 * 0.92}{0.65 * 0.92 + 0.35 * 0.95} = 0.64267 \ldots$$

$\mathrm{P}(A|C^c) \approx 0.643$

And so the probability of randomly picking a working lightbulb produces by Machine A is approximated to be 64 percent

(ii) Suppose the lifetime of a lightbulb without defect follows an exponential distribution with mean 10000 hours, what is the probability that a lightbulb functions for more than 5000 hours?

Let X be the time elapsed until the event of interests measured in hours
We know the mean of X is
$E[X] = 10000$

Recall that

$$E[X] = \frac{1}{\lambda}$$

Where λ is a positive parameter . We can rewrite the equation in relation to λ

$$E[X] = \frac{1}{\lambda} \Rightarrow$$

$$\lambda = \frac{1}{\mathrm{E[X]}}$$

$$\lambda = \frac{1}{10000}$$

Using the formula
$P(X \geq a) = P(X > a) = e^{-\lambda a}$

We can calculate the probability that X exceeds 5000 hours
$Let\ a = 5000$

$P(X > 5000) = e^{-\frac{1}{10000} * 5000}$

$P(X > 5000) = e^{-\frac{1}{2}}$

$P(X > 5000) = 0.60653065971263342360379953499118 \approx 0.61$

And so the probability that a light functions more than 5000 hours is approximately 61 percent

Exercise 2

Exercise 2: Consider an infinite number of bags S_1, S_2, S_3, etc. S_1 contains 3 yellow marbles and 2 green ones. Each of the following bag contains 2 green and 2 yellow marbles. We draw a marble from S_1 and put it in S_2, then from we draw a marble from S_2 and put it in S_3, etc. For $n \geq 1$ we denote by E_n the event "the marble drawn from S_n is green".

(i) Calculate $P(E_1)$, $P(E_2 \mid E_1)$, $P(E_2 \mid E_1^c)$ and $P(E_2)$.
(ii) Express $P(E_{n+1})$ as a function of $P(E_n)$ for all $n \geq 1$.
(iii) *Bonus question:* Can you show that $P(E_n) \to 1/2$ when $n \to \infty$?

(1)
We need to show $P(E_1)$, $P(E_2|E_1)$, $P(E_2|E_1^c)$, $P(E_2)$

Let the following set describe what a bag contains
$S_1 = \{G, G, Y, Y, Y\}$
$S_{n>1} = \{G, G, Y, Y\}$

We define the events
$E_1 = Marble\ from\ S_1\ is\ green$
$E_2 = Marble\ from\ S_2\ is\ green$
$E_n = Marble\ from\ S_n\ is\ green$

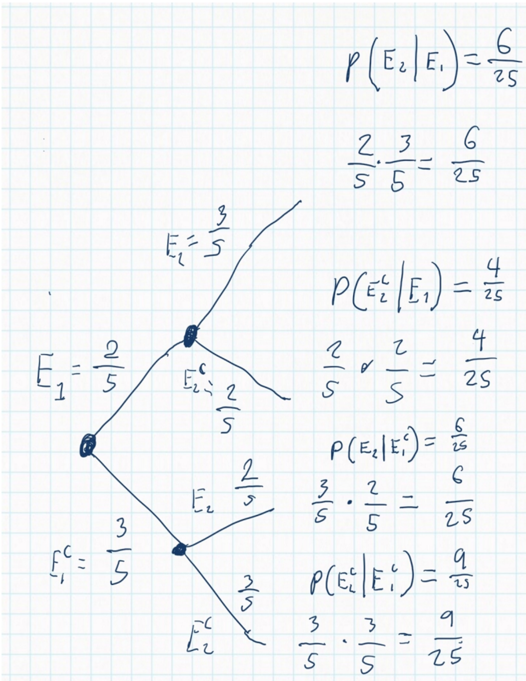
Assume that all balls are equally likely to be pulled from a bag we can define the following probabilities
$P(E_1) = \frac{2}{5}$
$P(E_1^c) = 1 - P(E_1) = \frac{3}{5}$

If we pick a ball from S1 and insert in S2, we know that if the previous picked ball was green then we know that S_{n+1} contains
$S_{n+1} = \{G, G, G, Y, Y\}$

In the other hand we know that if the previous ball was not green then we have that S_{n+1} contains
$S_{n+1} = \{G, G, Y, Y, Y\}$

$$P(E_{n+1}|E_n) = \frac{(number\ of\ G\ in\ S_{n+1}) + 1}{(number\ of\ Balls\ S_{n+1}) + 1} * P(E_n)$$

$$P(E_{n+1}|E_n^c) = \frac{(number\ of\ G\ in\ S_{n+1})}{(number\ of\ Balls\ in\ S_{n+1}) + 1}) * (1 - P(E_n))$$



$$P(E_2|E_1) = \frac{P(E_2 \cap E_1)}{P(E_1)} = \frac{\frac{3}{5} * \frac{2}{5}}{\frac{2}{5}} = \frac{\frac{6}{25}}{\frac{2}{5}} = \frac{30}{50} = \frac{3}{5}$$

$$P(E_2|E_1^c) = \frac{P(E_2 \cap E_1^c)}{P(E_1^c)} = \frac{\frac{2}{5} * \frac{3}{5}}{\frac{3}{5}} = \frac{\frac{6}{25}}{\frac{3}{5}} = \frac{30}{75} = \frac{2}{5}$$

Applying the total probability theorem we can calculate P(E2)
$P(E_2) = P(E_1) * P(E_2|E_1) + P(E_1^c) * P(E_2|E_1^c)$

$$P(E_2) = \frac{2}{5} * \frac{3}{5} + \frac{3}{5} * \frac{2}{5}$$

$$P(E_2) = \frac{6}{25} + \frac{6}{25}$$

$$P(E_2) = \frac{12}{25}$$

(2)
Express $P(E_{n+1})$ as a function of $P(E_n)$ for all $n \geq 1$

Recall that

$let\ g = number\ of\ green\ balls\ in\ S_{n+1}$
$let\ y = number\ of\ yellow\ balls\ S_{n+1}$

$$P(E_{n+1}|E_n) = \frac{g+1}{g+y+1}$$

$$P(E_{n+1}|E_n^c) = \frac{g}{g+y+1}$$

We can use these probabilities with the total probability theorem to define a general function for $P(E_{n+1})$ of $P(E_n)$ for all $n \geq 1$

$$P(E_{n+1}) = P(E_{n+1}|E_n) * P(E_n) + P(E_{n+1}|E_n^c) *1(-P(E_n))$$

$$P(E_{n+1}) = \frac{g+1}{g+y+1} * P(E_n) + \frac{g}{g+y+1} *1(-P(E_n))$$

Reducing the equation

$$P(E_{n+1}) = \frac{P(E_n) * (g+1)}{g+y+1} + \frac{g * (1 - P(E_n))}{g+y+1}$$

$$P(E_{n+1}) = \frac{P(E_n) * g + P(E_n)}{g+y+1} + \frac{g *1(-P(E_n))}{g+y+1}$$

$$P(E_{n+1}) = \frac{P(E_n) * g + P(E_n)}{g+y+1} + \frac{g - g * P(E_n)}{g+y+1}$$

$$P(E_{n+1}) = \frac{P(E_n)}{g+y+1} + \frac{g}{g+y+1}$$

$$P(E_{n+1}) = P(E_n) * \frac{1}{g+y+1} + \frac{g}{g+y+1}$$

Where the initial condition is

$$P(E_1) = \frac{2}{2+3}$$

Inserting the values we have our function

$$P(E_1) = \frac{2}{2+3} = \frac{2}{5}$$

$$P(E_{n+1}) = P(E_n) * \tfrac{1}{5} + \tfrac{2}{5}$$

Testing that we get the right results

$$P(E_2) = P(E_1) * \frac{1}{5} + \frac{2}{5}$$

$$P(E_2) = \frac{2}{5} * \frac{1}{5} + \frac{2}{5}$$

$$P(E_2) = \frac{2}{25} + \frac{2}{5}$$

$$P(E_2) = \frac{2}{25} + \frac{10}{25}$$

$$P(E_2) = \frac{12}{25}$$

Which corresponds the resulting $P(E_2)$ from the previous assignment

Likewise for $P(E_3)$

$$P(E_3) = P(E_2) * \frac{1}{5} + \frac{2}{5}$$
$$P(E_3) = \frac{12}{25} * \frac{1}{5} + \frac{2}{5} = \frac{62}{125}$$

We may generalize our function further.
Since we have

$$P(E_1) = \frac{2}{5}$$

$$P(E_2) = P(E_1) * \frac{1}{5} + \frac{2}{5}$$

$$P(E_3) = \left(P(E_1) * \frac{1}{5} + \frac{2}{5}\right) * \frac{1}{5} + \frac{2}{5}$$

$$P(E_4) = \left(\left(P(E_1) * \frac{1}{5} + \frac{2}{5}\right) * \frac{1}{5} + \frac{2}{5}\right) * \frac{1}{5} + \frac{2}{5}$$

$$P(E_n) = P(E_{n-1}) * \frac{1}{5} + \frac{2}{5} \quad : for\ all\ n \geq 2$$

(3)
Show that $P(E_n) \to \frac{1}{2} \mid n \to \infty$

Proof by induction
we can then show that when $n \to \infty$ the probability is $\frac{1}{2}$

We show that
$P(E_n) \leq \frac{1}{2}$
And also
$P(E_n) \leq P(E_{n+1})$

Proving the base case for n=1

$$P(E_1) \leq \frac{1}{2} \Rightarrow \frac{2}{5} \leq \frac{1}{2}$$
$$P(E_1) \leq P(E_2) \Rightarrow \frac{2}{5} \leq \frac{12}{25}$$

Assume that n=k is true
$P(E_k) \leq \frac{1}{2}$

And also
$P(E_k) \leq P(E_{k+1})$

Currently stuck right here... My induction skills are not that good.

I'm not sure how to correctly proof this by induction. Yet, the reason why we get $\frac{1}{2}$ probability for $P(E_{n\to\infty})$ is because the initial probability plays an insignificant role on the overall probability of getting a green ball when n grows towards ∞.

Exercise 3: Let X be an exponential random variable with parameter 1 and ϵ a Bernouilli random variable taking its values in $\{-1, +1\}$ with parameter $1/2$. Give the CDF of $Y = \epsilon X$. Is it a continuous random variable? If yes, use its CDF to compute its PDF.

ϵ is a discrete random variable and so the PMF of a Bernoulli random variable ϵ with parameter 1/2 is

$$p_\epsilon(k) = \begin{Bmatrix} \frac{1}{2} & if\ k = -1{,}1 \\ 0 & otherwise \end{Bmatrix}$$

The PDF X with parameter 1 is

$$f_X(x) = \begin{Bmatrix} e^{-x} & if\ x \geq 0 \\ 0 & otherwise \end{Bmatrix}$$

1)

Give CDF of $Y = \epsilon X$

We can use the total probability rule to calculate this. We know

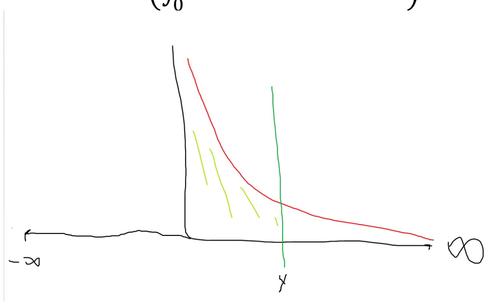
$$P(Y \leq y) = P(\epsilon X \leq y) = P(\epsilon X \leq y \cap \epsilon = 1) + P(\epsilon X \leq y \cap \epsilon = -1)$$

$$= P(\epsilon = 1) * P(\epsilon X \leq y | \epsilon = 1) + P(\epsilon = -1) * P(\epsilon X \leq y | \epsilon = -1)$$

$$= \frac{1}{2} * P(X \leq y) + \frac{1}{2} * P(-X \leq y)$$

We can then compute the CDFs of $P(X \leq y)$ and $P(-X \leq y)$ for the cases if $y \geq 0$ and $y < 0$

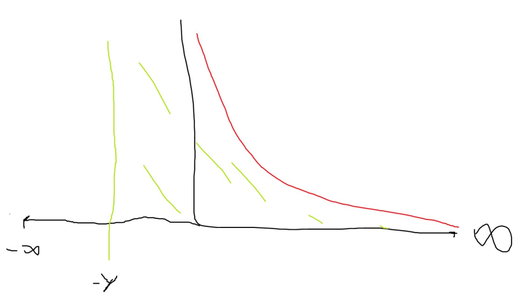
$$P(X \leq y) = \begin{Bmatrix} 0 & if\ y < 0 \\ \int_0^y f_X(x)dx & y \geq 0 \end{Bmatrix}$$



When y < 0 we get zero because we are outside the area of interest whereas when y >= 0 we calculate the area from 0 to y

On the other case we have

$$P(-X \leq y) = P(X \geq -y) = \begin{Bmatrix} \int_{-y}^{\infty} f_X(x)dx & if\ y < 0 \\ 1 & y \geq 0 \end{Bmatrix}$$



When y < 0 we calculate the area between -y to infinity else if y >= 0 we know that y is the area outside of interest added with all the area of interest and so we know that 0 + 1 = 1

Looking at the cases of y we get two cases

If $y < 0$

$$P(Y \leq y) = \frac{1}{2} * 0 + \frac{1}{2} * \int_{-y}^{\infty} f_x(x)dx$$

If $y \geq 0$

$$P(Y \leq y) = \frac{1}{2} * \int_{0}^{y} f_x(x)dx + \frac{1}{2} * 1$$

Calculating the integrals

$$\int_{-\infty}^{y} f_x(x)dx = \int_{-\infty}^{y} e^{-x}\,dx \Rightarrow -e^{-x}\Big|_0^y = -e^y - (-e^0) = -e^y - (-1) = -e^{-y} + 1$$

$$P(-X \leq y) = P(X \geq -y) = \int_{-y}^{\infty} e^{-x}\,dx \Rightarrow -e^{-x}\Big|_{-y}^{\infty} = -e^{\infty} - (-e^y) = 0 - (-e^y) = e^y$$

Inserting the value

If $y < 0$

$$P(Y \leq y) = \frac{1}{2} * 0 + \frac{1}{2} * (e^y) = \frac{1}{2} * (e^y) = \frac{e^y}{2}$$

If $y \geq 0$

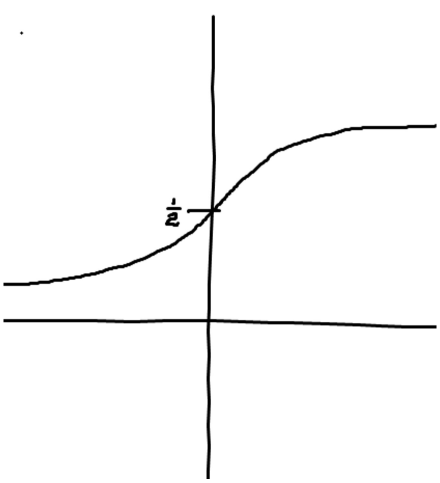
$$P(Y \leq y) = \frac{1}{2} * (-e^{-y} + 1) + \frac{1}{2} * 1 = \frac{-e^{-y} + 1}{2} + \frac{1}{2} = \frac{-e^{-y} + 1}{2} + \frac{1}{2} = \frac{-e^{-y} + 2}{2} = \frac{-e^{-y}}{2} + \frac{2}{2} = \frac{-e^{-y}}{2} + 1$$

And so the CDF is

$$F_Y(y) = \begin{Bmatrix} \frac{e^y}{2} & if\ y < 0 \\ \frac{-e^{-y}}{2} + 1 & y \geq 0 \end{Bmatrix}$$

2)

Yes Y is a continuos random variable because when you visualize the CDF we have



Since the graph is connected and contains no singularity we can differentiate the CDF and thus derive a PDF

$$f_Y(y) = \frac{dF_y}{dy}(y)$$

$$f(y) = \begin{cases} \dfrac{d\left(\frac{1}{2} * e^{y}\right)}{dy}(y) & if\ y < 0 \\ \dfrac{d\left(\frac{1}{2} * -e^{-y}\right)}{dy}(y) & y \geq 0 \end{cases}$$

$$f(y) = \begin{cases} \dfrac{1}{2} * e^{y} & if\ y < 0 \\ \dfrac{1}{2} * e^{-y} & y \geq 0 \end{cases}$$

Which is the pdf of Y