Problem 1. (2 children) Consider a family that has two children. We are interested in the children's genders. Our sample space is:

$$S = \{(G,G), (G,B), (B,G), (B,B)\}$$

where G represents a girl and B represents a boy. Assume that all four possible outcomes are equally likely.

1. What is the probability that both children are girls given that the first child is a girl?
2. Suppose we ask the father: "Do you have at least one daughter?" He responds "Yes!" Given this extra information, what is the probability that both children are girls? In other words, what is the probability that both children are girls given that we know at least one of them is a girl?

$\frac{3}{4}$

$(G,G),\ (G,B),\ \cancel{(B,G)}$ $\frac{3}{4}$

$\frac{1}{3}$

$\frac{1}{3}$ $\frac{1}{2}$

$P(A) = \frac{1}{4}$

$P(B) = \frac{2}{4}$

$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$ $P(A \cap B)$

$= \frac{\frac{1}{4}}{\frac{2}{4}} = \frac{1}{2}$

$\underbrace{(G,G)\ (G,B)\ (B,G)}$

$\frac{1}{3}$

Problem 2. (2 Children, one Lilia) A family has two children. We ask the father, "Do you have at least one daughter named Lilia?" He replies, "Yes!" What is the probability that both children are girls?

In other words, we want to find the probability that both children are girls, given that the family has at least one daughter named Lilia. Here, you can assume the following:

- If a child is a girl, her name will be Lilia with probability $\alpha \ll 1$, independently from other children's names.
- If the child is a boy, his name will not be Lilia.

$\alpha \cdot \alpha$

$$P(GG|L) = \frac{P(L|GG) \cdot \frac{1}{4}}{P(L)}$$

$$P(L|GG) = \alpha \cdot \alpha + \alpha(1-\alpha) + (1-\alpha)\alpha = 2\alpha - \alpha^2$$

L L

L L

L L

GG   GB   BG   BB

L   L   L   L

$\boxed{GB}$ $\frac{1}{4}$

BG

BB

GG

α

$$P(L \cap GB) = P(GB) \cdot P(L|GB)$$

$$= P(GG) \cdot L(L|GG)$$

$\frac{1}{4}$

$P(BB) \cdot P(L|BB)$

$\frac{1}{4} \cdot 0$

$0.8 \quad 0.65 \quad \alpha \neq 1$

$$\frac{\frac{1}{4}(2\alpha - \alpha^2)}{\frac{1}{4}\alpha + \frac{1}{4}\alpha + \frac{1}{4}(2\alpha - \alpha^2)} = \frac{2\alpha - \alpha^2}{2\alpha + 2\alpha - \alpha^2} = \frac{2\alpha - \alpha^2}{4\alpha - \alpha^2} = \frac{2-\alpha}{4-\alpha}$$

$$\frac{\frac{1}{2} \cdot \frac{1}{2}}{\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2}}$$

$GG = GB = h$

Problem 3. (elections) In a certain town, 30% of the people are conservatives and 70% socialists. At the last election, 65% of conservatives voted and 80% of socialists. A person from the town is selected at random, and states that she voted at the last election. What is the probability that she is a socialist?

Problem 4. (Spam emails) A spam filter tags emails as spam or not spam. Based on historical data:

- 80% of spam emails contain the word "lottery."
- 30% of non-spam emails contain the word "lottery."
- 40% of emails are spam.

What is the probability that an email is spam given it contains the word "lottery"?

Problem 5. (PMF and CDF for 2 coin flips) I toss a coin twice. Let X be the number of observed heads. Find and plot the PMF and CDF of X.

H H

$P(\ldots 0) = \frac{1}{4}$

$P(\{X = 1\}) = \frac{1}{2}$

$X = \{0, 1, 2\}$ $P(\ldots 2) = \frac{1}{4}$

$0 \to \frac{1}{4}$

$1 \to \frac{1}{2}$

$2 \to \frac{1}{4}$

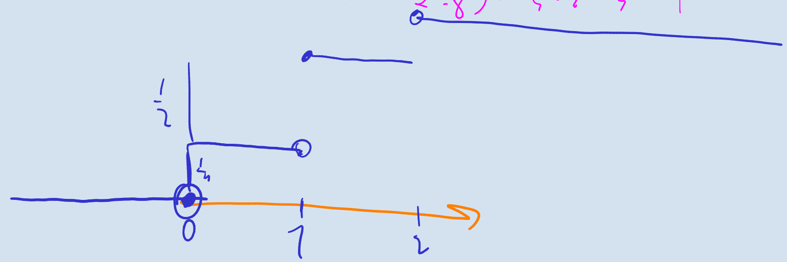
$F_X(x) = P(X < x)$

$F(0) = 0$

$F(0.2) = \frac{1}{4}$

$F(1.5) = \frac{1}{4} + \frac{1}{2}$

$F(2.8) = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$

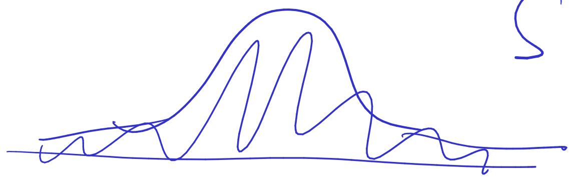


Problem 6. (PDF, CDF) Let X be a continuous random variable with the following PDF:

$$f_X(x) = \begin{cases} ce^{-x} & \text{if } x \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

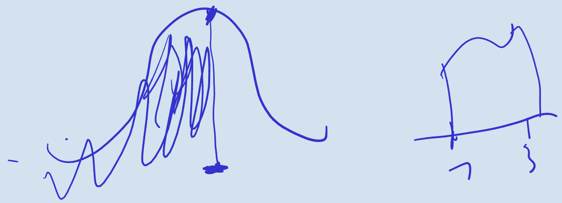
where c is a positive constant.

$\begin{cases} e^{-x} \\ 0, \end{cases}$



$\int$

1. Find c.
2. Find the CDF of X, denoted $F_X(x)$.
3. Find $P(1 < X < 3)$.



$1 = \int_{-\infty}^{\infty} f_X(x)\,dx = \int_0^{\infty} ce^{-x}\,dx =$

$= c\int_0^{\infty} e^{-x}\,dx = c\,-e^{-x}\Big|_0^{\infty} =$

$= -c(e^{-\infty} - e^{0}) = -c(0-1) = c = 1$

$\int_a^b f(x)\,dx$

$g(b) - g(a)$

$g(x)\Big|_a^b$

2 $F_X(x) = \int_0^x e^{-u}\,du$

$\int_1^3$

$\int e^{-x} = -e^{-x}$

$P(1 < X < 3) = F(3) - F(1)$

$Var(X) = E[(X - E(X))^2]$

$\int (x-c)^2 f(x)\,dx$

$Var(X) = EX^2 - (EX)^2$

Problem 7. (Expected value and Variance of the Uniform distribution) Let $X \sim \text{Uniform}(a, b)$.

1. Find $\mathbb{E}[X]$ (the expected value of X).
2. Find $\mathbb{V}\text{ar}(X)$ (the variance of X).

3 3 3 3 4 4 4 4

1 2 3 4 5 6

1 2 3 4 5 6

$\sum x P(x)$ $\quad 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + \dots = 3.5$

10 $\quad$ 30 $\quad$ 1000\$

100% $\quad$ 20% $\quad$ 998

$\frac{1}{2}$ $\quad$ $\frac{1}{8}$

$\mathbb{E}Y = \int_{-\infty}^{\infty} x^2 f(x)\,dx$

$\mathbb{E}[X^2]$ $\quad \sum_{i=1}^{6} i^2 \cdot \frac{1}{6}$ $\quad K$

log

$\text{Var} = K - (3.5)^2$

$\int x f(x)\,dx$ $\left\{ \frac{3}{x^4} \quad x \geq 1 \right.$ $\quad 0$

$\frac{1}{6} \cdot (1-3.5)^2 + \frac{1}{6} \cdot (2-3.5)^2$

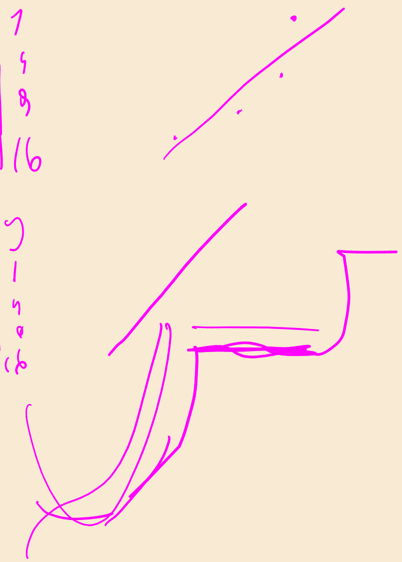
$x^2 y$

feature engineering

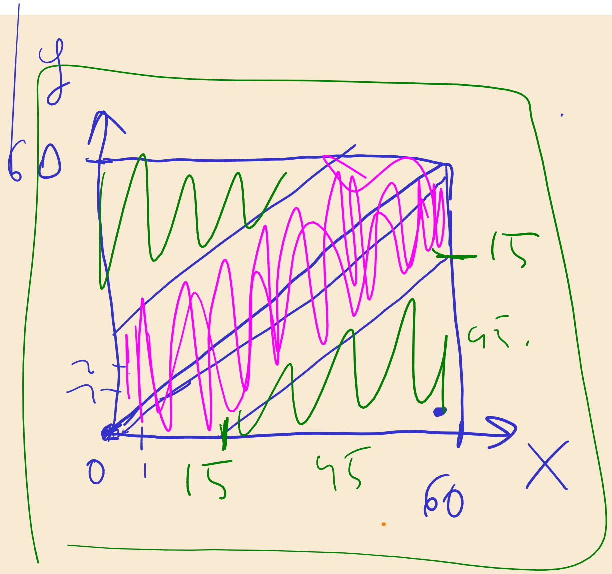
Problem 9. Anush and Nairi are shopping at the mall. They agree to split up for a time and then meet for lunch. They plan to meet in front of Kinopark between 12:00 and 13:00. The one who arrives first agrees to wait 15 minutes for the other to arrive. After 15 minutes, that person will leave and continue shopping. What is the probability that they will meet if each one of them arrives at any time between 12:00 and 13:00?

Hint: Try to represent the problem on the coordinate system, by letting x denote the time Anush arrives, and y, the time Nairi arrives.

x	0
1	1
2	4
3	9
4	16



$$\boxed{\begin{array}{l} |x-y| \le 15 \\ x, y \in [0, 60] \end{array}}$$



$x = y + 15$

$$\frac{\#\text{succ}}{\#\text{total}}$$

$y = x - 15$

$y = x + 15$

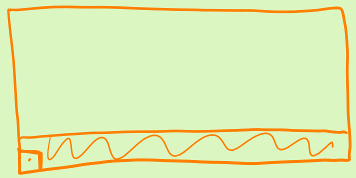
3600

🧪 Problem: Medical Test

A disease affects 1 in 1,000 people (i.e., 0.1%). There is a test for the disease:

- If a person **has** the disease, the test is **positive** 99% of the time (true positive rate = sensitivity).
- If a person **does not have** the disease, the test is **positive** 5% of the time (false positive rate = 5%).

Question: If a person tests positive, what is the probability they actually have the disease?

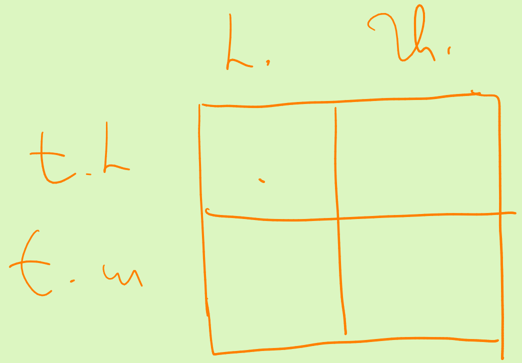


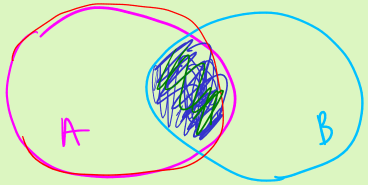
1000
50

0.99 + 50

$$\frac{1}{0.99+50} \approx 2\%$$

$$P(A|B) = \frac{P(A\cap B)}{P(B)}$$





$$P(A\cap B) = P(A|B)\cdot \cancel{P(B)} =$$

$$P(B|A) = \frac{P(B\cap A)}{P(A)} \Rightarrow = \frac{P(B|A)\cdot P(A)}{P(B)}$$

1.9

$$P(H|T^+) = \frac{P(T^+|H)\ P(H)}{P(T^+)}$$

$$P(H) = 0.1$$

$$P(H_{gild}) = 0.99$$

$$P(T^+) = P(T^+|H) + P(T^+|H_{gild})$$

0.99   0.05