Applied Stochastic Processes

Kipngeno Koech (bkoech - andrew ID)

September 13, 2024

Question 1: Probability Desntiy Function

1. Consider the PDF of a continuous random variable Y that models the duration (in hours) of a sudden electrical outage in a city, described by the function:

$$g(y) = \lambda^2 y e^{-\lambda y} \quad \text{for} \quad y > 0$$

(a) (2 points): Prove that g(y) is a valid probability density function by verifying the normalization condition:

normalization condition: $\int_0^\infty \lambda^2 y e^{-\lambda y} dy = 1$

$$\int_0^\infty \lambda^2 y e^{-\lambda y} dy = \lambda^2 \int_0^\infty y e^{-\lambda y} dy$$

let $u = y$, $dv = e^{-\lambda y} dy$, $du = dy$, $v = -\frac{1}{\lambda} e^{-\lambda y}$

$$\lambda^2 \left[-\frac{1}{\lambda} y e^{-\lambda y} \Big|_0^\infty + \frac{1}{\lambda} \int_0^\infty e^{-\lambda y} dy \right]$$

$$\lambda^2 \left[0 + \frac{1}{\lambda} \left(-\frac{1}{\lambda} e^{-\lambda y} \Big|_0^\infty \right) \right]$$

$$\lambda^2 \left[\frac{1}{\lambda^2} e^{-\lambda y} \Big|_0^\infty \right]$$

$$\lambda^2 \left[\frac{1}{\lambda^2} e^{-\infty} - \frac{1}{\lambda^2} e^0 \right]$$

$$\lambda^2 \left[0 - \frac{1}{\lambda^2} \right]$$

$$\lambda^2 \left[-\frac{1}{\lambda^2} \right]$$

$$\lambda^2 \left[-\frac{1}{\lambda^2} \right] = -1$$

(b) (4 points): Show whether g(y) is concave or convex over its domain by analyzing its second derivative. Discuss the practical implications of this concavity in predicting the likelihood of prolonged outages.

The second derivative of the function is given by:

$$\frac{d^2}{dy^2}(\lambda^2 ye^{-\lambda y}) = \frac{d}{dy}(\lambda^2 e^{-\lambda y} - \lambda^3 ye^{-\lambda y})$$

$$\frac{d}{dy}(\lambda^2 e^{-\lambda y} - \lambda^3 ye^{-\lambda y}) = \lambda^3 e^{-\lambda y} - \lambda^4 ye^{-\lambda y} - \lambda^3 e^{-\lambda y}$$

$$\lambda^3 e^{-\lambda y} - \lambda^4 ye^{-\lambda y} - \lambda^3 e^{-\lambda y} = -\lambda^4 ye^{-\lambda y}$$

The second derivative of the function is negative, which means that the function is concave over its domain. This means that the function is decreasing over its domain. This has practical implications in predicting the likelihood of prolonged outages. If the function is concave, it means that the likelihood of prolonged outages is decreasing over time. This means that the likelihood of prolonged outages is decreasing as time goes on. This can be useful in predicting the likelihood of prolonged outages and taking appropriate action to prevent them.

(c) (4 points): Find the mode of the distribution g(y) by solving for the critical points of the function and determining which point corresponds to the maximum or minimum value. Using λ = [0.5 2] relate the mode to real-life scenarios where outages are frequent and brief or infrequent but prolonged.

The mode of the distribution is given by the point where the first derivative of the function is equal to zero. The first derivative of the function is given by:

$$\frac{d}{dy}(\lambda^2 ye^{-\lambda y}) = \lambda^2 e^{-\lambda y} - \lambda^3 ye^{-\lambda y}$$

Setting the first derivative equal to zero:

$$\lambda^2 e^{-\lambda y} - \lambda^3 ye^{-\lambda y} = 0$$

$$\lambda^2 e^{-\lambda y}(1 - \lambda y) = 0$$

$$\lambda^2 e^{-\lambda y} = 0 \quad \text{or} \quad 1 - \lambda y = 0$$

$$\lambda^2 e^{-\lambda y} = 0$$

is not possible because the exponential function is always positive. Therefore, the mode of the distribution is given by:

$$1 - \lambda y = 0$$

$$1 = \lambda y$$

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

The mode of the distribution is given by $y = \frac{1}{\lambda}$. The mode of the distribution is the point where the likelihood of an outage is the highest. If the outages are frequent and brief, the mode of the distribution will be closer to zero. If the outages are infrequent but prolonged, the mode of the distribution will be further from zero.

Two manufacturing plants, Plant A and Plant B, produce goods with daily outputs X (in tons) and Y (in tons) respectively. The production efficiency between the two plants can be modeled by the joint probability density function:

$$f(x, y) = e^{-(x+y)} \text{ for } x > 0 \text{ and } y > 0.$$

(a) (2 points): Show that f(x, y) is a valid joint probability density function by verifying the normalization condition:

$$\iint_{x>0,y>0} e^{-(x+y)}\, dx\, dy = 1.$$

$$\int_0^\infty \int_0^\infty e^{-(x+y)}\, dx\, dy = \int_0^\infty e^{-y}\, dy \int_0^\infty e^{-x}\, dx$$

$$\int_0^\infty e^{-y}\, dy = -e^{-y}\Big|_0^\infty = -e^{-\infty} + e^0 = 1$$

$$\int_0^\infty e^{-x}\, dx = -e^{-x}\Big|_0^\infty = -e^{-\infty} + e^0 = 1$$

$$\int_0^\infty \int_0^\infty e^{-(x+y)}\, dx\, dy = 1 \times 1 = 1$$

Therefore, the joint probability density function is valid.

(b) (4 points): Find the probability that both plants produce more than 1 ton of goods in a day by calculating:

$$P(X > 1, Y > 1) = \iint_{x>1,y>1} e^{-(x+y)}\, dx\, dy.$$

$$\int_1^\infty \int_1^\infty e^{-(x+y)}\, dx\, dy = \int_1^\infty e^{-y}\, dy \int_1^\infty e^{-x}\, dx$$

$$\int_1^\infty e^{-y}\, dy = -e^{-y}\Big|_1^\infty = -e^{-\infty} + e^{-1} = e^{-1}$$

$$\int_1^\infty e^{-x}\, dx = -e^{-x}\Big|_1^\infty = -e^{-\infty} + e^{-1} = e^{-1}$$

$$\int_1^\infty \int_1^\infty e^{-(x+y)}\,dx\,dy = e^{-1} \times e^{-1} = e^{-2}$$

Therefore, the probability that both plants produce more than 1 ton of goods in a day is e^{-2}.

Discuss the practical interpretation of this result for plant managers who are aiming to maintain high production levels.

The probability that both plants produce more than 1 ton of goods in a day is e^{-2}. This means that the likelihood of both plants producing more than 1 ton of goods in a day is e^{-2}. This can be interpreted as the likelihood of both plants maintaining high production levels. If the probability is high, it means that the likelihood of both plants maintaining high production levels is high. If the probability is low, it means that the likelihood of both plants maintaining high production levels is low. Plant managers can use this information to make decisions about how to maintain high production levels. They can use this information to make decisions about how to improve production efficiency and increase output.

(c) (4 points): Calculate the probability that Plant A produces more than 2 tons and Plant B produces less than 2 tons:

$$P(X > 2, Y < 2) = \iint_{x>2, 0<y<2} e^{-(x+y)}\,dx\,dy.$$

$$\int_2^\infty \int_0^2 e^{-(x+y)}\,dx\,dy = \int_2^\infty e^{-y}\,dy \int_0^2 e^{-x}\,dx$$

$$\int_2^\infty e^{-y}\,dy = -e^{-y}\Big|_2^\infty = -e^{-\infty} + e^{-2} = e^{-2}$$

$$\int_0^2 e^{-x}\,dx = -e^{-x}\Big|_0^2 = -e^{-2} + e^0 = 1 - e^{-2}$$

$$\int_2^\infty \int_0^2 e^{-(x+y)}\,dx\,dy = e^{-2} \times (1 - e^{-2}) = e^{-2} - e^{-4}$$

Therefore, the probability that Plant A produces more than 2 tons and Plant B produces less than 2 tons is $e^{-2} - e^{-4}$.

Explain how this information might be useful for balancing production targets between the twoplants.

The probability that Plant A produces more than 2 tons and Plant B produces less than 2 tons is $e^{-2} - e^{-4}$. This information can be useful for balancing production targets between the two plants. Plant managers can use this information to set production targets for each plant. They can use this information to determine how much each plant should produce in order to meet overall production targets. They can use this information to balance production levels between the two plants and ensure that production targets are met. This information can be useful for optimizing production efficiency and increasing output.

Question 2: Eigenvalues and Eigenvectors

You are analyzing a system modeled by a 3x3 state transition matrix, which represents the interaction of three interconnected features in a neural network, given by:

$$B = \begin{pmatrix} 4 & 2 & 1 \\ 1 & 3 & 2 \\ 0 & 1 & 2 \end{pmatrix}$$

To get the eigenvectors, you use the formula $\det(B - \lambda I) = 0$ where I is the identity matrix.

1 Eigen Values

The eigenvalues of the matrix B are given by the roots of the equation $\det(B - \lambda I) = 0$. $\det(B - \lambda I) = 0$ is given by:

$$\det \begin{vmatrix} 4-\lambda & 2 & 1 \\ 1 & 3-\lambda & 2 \\ 0 & 1 & 2-\lambda \end{vmatrix} = 0$$

$$(4-\lambda)\underline{((3-\lambda)(2-\lambda)-2)} - 2(2-\lambda) + 1 = 0$$

$$(4-\lambda)(6-5\lambda+\lambda^2-2) - \underline{2(2-\lambda)} + 1 = 0$$

$$(4-\lambda)(4-5\lambda+\lambda^2) - \underline{4} + 2\lambda + \underline{1} = 0$$

$$\underline{(4-\lambda)(4-5\lambda+\lambda^2)} + 2\lambda - 3 = 0$$

$$4(4-5\lambda+\lambda^2) - \lambda(4-5\lambda+\lambda^2) + 2\lambda - 3 = 0$$

$$16 - 20\lambda + 4\lambda^2 - 4\lambda + 5\lambda^2 - \lambda^3 + 2\lambda - 3 = 0$$

$$-\lambda^3 + 9\lambda^2 - 18\lambda + 13 = 0$$

$$\lambda^3 - 9\lambda^2 + 18\lambda - 13 = 0$$

The roots of the equation $\lambda^3 - 9\lambda^2 + 18\lambda - 13 = 0$ are the eigenvalues of the matrix B:

$$\lambda_1 \approx 5.33005874, \quad \lambda_2 \approx 2.79836032, \quad \lambda_3 \approx 0.87158094$$

since all the three eigen values are positive, the system is unstable

2 Modification of the off diagonal elements of the matrix

let us say the off-diagonal elements of the matrix are modified to be 0.5 as follows:

$$\begin{pmatrix} 4 & 0.5 & 0.5 \\ 0.5 & 3 & 0.5 \\ 0.5 & 0.5 & 2 \end{pmatrix}$$

The eigenvalues of the matrix B are given by the roots of the equation $\det(B - \lambda I) = 0$. which is given by:

the matrix B is given by:

$$\begin{vmatrix} 4-\lambda & 0.5 & 0.5 \\ 0.5 & 3-\lambda & 0.5 \\ 0.5 & 0.5 & 2-\lambda \end{vmatrix} = 0$$

$$4-\lambda\underline{((3-\lambda)(2-\lambda)-0.5^2)}-\underline{0.5((0.5(2-\lambda)-0.5(0.5)))}+0.5\underline{((0.5)(0.5)-0.5(3-\lambda))} = 0$$

$$(4-\lambda)(6-5\lambda+\lambda^2-0.5^2)-0.5(0.75-0.5\lambda)+0.5(0.5\lambda-1.25) = 0$$

$$4(6-5\lambda+\lambda^2-0.25)-\lambda(6-5\lambda+\lambda^2-0.25)-0.375+0.25\lambda+0.25\lambda-1.25 = 0$$

$$24-20\lambda+4\lambda^2-1-6\lambda+5\lambda^2-\lambda^3+0.25\lambda-0.375+0.25\lambda-1.25 = 0$$

$$-\lambda^3+9\lambda^2-18\lambda+13 = 0$$

The roots of the equation $\lambda^3 - 9\lambda^2 + 18\lambda - 13 = 0$ are the eigenvalues of the matrix B:

$$\lambda_1 \approx 5.33005874, \quad \lambda_2 \approx 2.79836032, \quad \lambda_3 \approx 0.87158094$$

the eigen values still remain the same, hence the system is still unstable

Question 3: Markov Chains

The transition matrix for a Markov chain is given by:

$$Q = \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{pmatrix}$$

It is modelling the transition of a system between three states high performance, medium performance, low performance, for employees respectively.

initial state is given by: $\pi_0 = \begin{pmatrix} 0.5 & 0.3 & 0.2 \end{pmatrix}$

3 find the state of the system after 1 step

The state of the system after 1 step is given by:

$$\pi_1 = \pi_0 Q$$

$$\pi_1 = \begin{pmatrix} 0.5 & 0.3 & 0.2 \end{pmatrix} \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{pmatrix}$$

$$\pi_1 = \begin{pmatrix} 0.5(0.7) + 0.3(0.3) + 0.2(0.2) & 0.5(0.2) + 0.3(0.4) + 0.2(0.5) & 0.5(0.1) + 0.3(0.3) + 0.2(0.3) \end{pmatrix}$$

$$\pi_1 = \begin{pmatrix} 0.48 & 0.32 & 0.2 \end{pmatrix}$$

4 steady state distribution

The steady state distribution is given by:

$$\pi Q = \pi$$

$$\begin{pmatrix} \pi_a & \pi_b & \pi_c \end{pmatrix} \begin{pmatrix} 07 & 0.2 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{pmatrix} = \begin{pmatrix} \pi_a & \pi_b & \pi_c \end{pmatrix}$$

also:

$$\pi_a + \pi_b + \pi_c = 1 - \text{ since the sum of the probabilities must be 1}$$

The above matrix equation can be written as:

$$\pi_a(0.7) + \pi_b(0.3) + \pi_c(0.2) = \pi_a \ldots \text{equation 1}$$

$$\pi_a(0.2) + \pi_b(0.4) + \pi_c(0.5) = \pi_b \ldots \text{equation 2}$$

$$\pi_a(0.1) + \pi_b(0.3) + \pi_c(0.3) = \pi_c \ldots \text{equation 3}$$

$$\pi_a + \pi_b + \pi_c = 1 \ldots \text{equation 4}$$

with this system of equations, we can solve for the steady state distribution. The solution is given by:

step 1: we can rewrite the equations as:

$$\pi_a(0.7 - 1) + \pi_b(0.3) + \pi_c(0.2) = 0 \ldots \text{equation 1}$$

$$\pi_a(0.2) + \pi_b(0.4 - 1) + \pi_c(0.5) = 0 \ldots \text{equation 2}$$

$$\pi_a(0.1) + \pi_b(0.3) + \pi_c(0.3 - 1) = 0 \ldots \text{equation 3}$$

$$\pi_a + \pi_b + \pi_c = 1 \ldots \text{equation 4}$$

this is equivalent to:

$$-(0.3)\pi_a + (0.3)\pi_b + (0.2)\pi_c = 0 \ldots \text{equation 1}$$

$$0.2\pi_a + -(0.3)\pi_b + (0.5)\pi_c = 0 \ldots \text{equation 2}$$

$$0.1\pi_a + (0.3)\pi_b + -(0.7)\pi_c = 0 \ldots \text{equation 3}$$

multiply equation 1 by 10, equation 2 by 10 and equation 3 by 10:

$$-3\pi_a + 3\pi_b + 2\pi_c = 0 \ldots \text{equation 1}$$

$$2\pi_a - 3\pi_b + 5\pi_c = 0 \ldots \text{equation 2}$$

$$\pi_a + 3\pi_b - 7\pi_c = 0 \ldots \text{equation 3}$$

$$\pi_a + \pi_b + \pi_c = 1 \ldots \text{equation 4}$$

we can write

$$\pi_a = 1 - \pi_b - \pi_c$$

and substitute in the above equations to get:

$$-3(1 - \pi_b - \pi_c) + 3\pi_b + 2\pi_c = 0 \ldots \text{equation 1}$$

$$2(1 - \pi_b - \pi_c) - 3\pi_b + 5\pi_c = 0 \ldots \text{equation 2}$$

$$1 - \pi_b - \pi_c + 3\pi_b - 7\pi_c = 0 \ldots \text{equation 3}$$

let us expand equation 1:

$$-3 + 3\pi_b + 3\pi_c + 3\pi_b + 2\pi_c = 0$$

$$-3 + 6\pi_b + 5\pi_c = 0$$

$$6\pi_b + 5\pi_c = 3 \ldots \text{equation 5}$$

let us expand equation 2:

$$2 - 2\pi_b - 2\pi_c - 3\pi_b + 5\pi_c = 0$$

$$2 - 5\pi_b + 3\pi_c = 0$$

$$-5\pi_b + 3\pi_c = -2 \ldots \text{equation 6}$$

the above equation 5 and 6 can be written as:

$$\begin{pmatrix} 6 & 5 \\ -5 & 3 \end{pmatrix} \begin{pmatrix} \pi_b \\ \pi_c \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$$

its argumented matrix is:

$$\begin{pmatrix} 6 & 5 & | & 3 \\ -5 & 3 & | & -2 \end{pmatrix}$$

by row reduction, we get:

$$R_1 => \frac{1}{6}R_1 => \begin{pmatrix} 1 & \frac{5}{6} & | & \frac{1}{2} \\ -5 & 3 & | & -2 \end{pmatrix}$$

$$R_2 => R_2 + 5R_1 => \begin{pmatrix} 1 & \frac{5}{6} & | & \frac{1}{2} \\ 0 & \frac{43}{6} & | & \frac{1}{2} \end{pmatrix}$$

$$R_2 => \frac{6}{43}R_2 => \begin{pmatrix} 1 & \frac{5}{6} & | & \frac{1}{2} \\ 0 & 1 & | & \frac{3}{43} \end{pmatrix}$$

$$R_1 => R_1 - \frac{5}{6}R_2 => \begin{pmatrix} 1 & 0 & | & \frac{19}{43} \\ 0 & 1 & | & \frac{9}{43} \end{pmatrix}$$

this gives us:

$$\pi_b = \frac{19}{43}, \quad \pi_c = \frac{9}{43}$$

$$\Pi_a = 1 - \frac{19}{43} - \frac{9}{43} = \frac{15}{43}$$

therefore the steady state distribution is given by:

$$\pi = \begin{pmatrix} \frac{15}{43} & \frac{19}{43} & \frac{9}{43} \end{pmatrix}$$

A lot of employees are in the medium performance state, while the least number of employees are in the low performance state. The high performance state has the second highest number of employees. So the company might introduce Initiatives to push the medium performance employees to high performance state and the low performance employees to medium performance state.

Question 5: Matrix and Vector Manipulation Operations

Given a matrix M that represents a quadratic form xTMx, discuss under what conditions the function is concave or convex. Connect this to how concavity affects the optimization landscape in machine learning models, such as in support vector machines or neural network loss surfaces

solution:

A quadtratic form $x^T Mx$ is convex if and only if:

- M it is positive semidefinite. A positive semidefinite matrix is a symmetric matrix where all the eigenvalues are non-negative.

A quadratic form $x^T Mx$ is concave if and only if:

- M is negative semidefinite. A negative semidefinite matrix is a symmetric matrix where all the eigenvalues are non-positive.

In machine learning, the optimization landscape is the surface that is formed by the loss function. The loss function is a function that measures the difference between the predicted value and the actual value. The optimization landscape is the surface that is formed by the loss function. The optimization landscape can be convex or concave.
A convex optimization landscape is one where the loss function is convex. A concave optimization landscape is one where the loss function is concave. Convex optimization landscapes are easier to optimize because they have a single global minimum. Concave optimization landscapes are harder to optimize because they have multiple local minima.

Question 5: Conditional Probability

Question 6: Conditional Probability

5 storyline

An insurance company examines its pool of auto insurance customers and makes the following observations:

- All customers insure at least one car
- 70% of the customers insure more than one car
- 20% of the customers insure a sports car
- of those customers who insure more than one car, 15% insure a sports car

6 find the probability that a randomly chosen customer insures one car and it is not a sports car

for Conditional probability, the formula is:

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

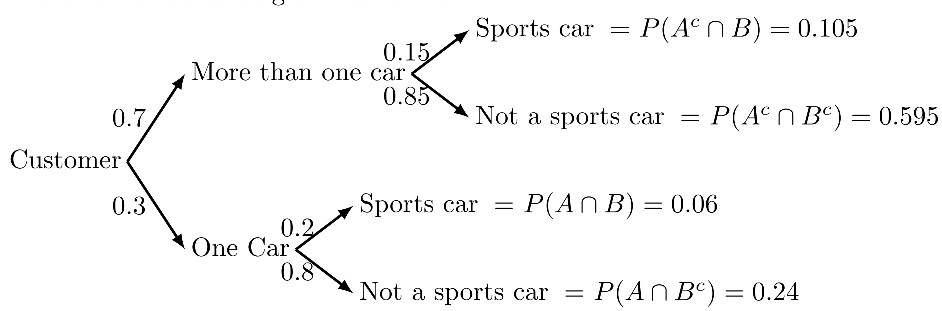
let A be the event that a customer insures one car and B be the event that the customer insure a sports car.

let us list out the probability of the events we have:

- $P(A) = 0.3 =$ probability that a customer insures one car
- $P(A^c) = 0.7 =$ probability that a customer insures more than one car
- $P(B) = 0.2 =$ probability that a customer insures a sports car

- $P(B^c = 0.8)$ = probability that a customer does not insure a sports car
- $P(A^c \cap B) = 0.105$ = probability that a customer insures more than one car and it is a sports car
- $P(A^c \cap B^c) = 0.595$ = probability that a customer insures more than one car and it is not a sports car

this is how the tree diagram looks like:



so, let us do the calculations:

What is the probability that a customer insures one car and it is a sports car? this are dependent events

- $P(B^c = 0.8)$ = probability that a customer does not insure a sports car
- $P(B^c|A) = \frac{P(B^c \cap A)}{P(A)}$ = probability that a customer insures one car and it is not a sports car
- $P(B^c \cap A) = P(A) - P(A \cap B) = 0.3 - (0.3 \times 0.2) = 0.24$ = probability that a customer insures one car and it is not a sports car
- $P(A \cap B) = P(A) \times P(B) = 0.3 \times 0.2 = 0.06$ = probability that a customer insures one car and it is a sports car

Therefore the probability that a randomly chosen customer insures one car and it is not a sports car is $\mathbf{0.24}$

Question 7: Bayes Theorem

7 questions

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

(a) (4 points) A medical test for a new strain of a viral respiratory disease has a 98% sensitivity and a 97% specificity. If 0.5% of the population has the disease, calculate the probability that a person has the disease given they tested positive.

- $P(D) = 0.005$ = probability that a person has the disease

- $P(D^c) = 0.995 =$ probability that a person does not have the disease
- $P(T|D) = 0.98 =$ probability that a person tests positive given they have the disease
- $P(T|D^c) = 0.03 =$ probability that a person tests positive given they do not have the disease
- $P(T) = P(T|D)P(D) + P(T|D^c)P(D^c) = (0.98 \times 0.005) + (0.03 \times 0.995) = 0.03475 =$ probability that a person tests positive
- $P(D/T) = \frac{P(T|D)P(D)}{P(T)} = \frac{0.98 \times 0.005}{0.03475} \approx 0.141$

(b) (4 points) A financial credit scoring model is 95 percent effective in identifying individuals who are likely to default on their loans when they actually will default. However, the model also yields a "false positive" result for 1 percent of individuals who are creditworthy. (That is, if a creditworthy individual is tested, then, with probability 0.01, the model will incorrectly classify them as likely to default.) If 0.5 percent of the population actually defaults on their loans, what is the probability that an individual will default given that the model predicts they are likely to default?

- $P(E) = 0.005 =$ probability that an individual will default on their loans
- $P(E|D) = 0.95 =$ probability that an individual will default on their loans given they actually will default
- $P(D) = P(E|D) \cdot P(D) + P(E|D^c) \cdot P(D^c) =$ probability that an individual will default on their loans
- $P(E|D^c) = 0.01 =$ probability that an individual will default on their loans given they are creditworthy
- $P(E) = (0.95)(0.005) + (0.01)(0.995)$
- $P(D) = 0.0147 =$ probability that an individual will default on their loans

Bayes Theorem $= P(D|E) = \frac{P(E|D)P(D)}{P(E)} = \frac{0.95 \times 0.005}{0.0147} \approx 0.323$

Question 8: Independence of Events, Inclusion-Exclusion Principle & Mutual Exclusivity

8 Positive probabilities & Independent events

1. show that if A & B are events with positive probabilities and are indepedent, then so are:

1. A^c and B

let us show that A^c and B are independent events:

$P(A \cap B) = P(A)P(B) =$ since A and B are independent events
$P(A^c \cap B) = P(B) - P(A \cap B) =$ by the definition of the complement of A
$P(A^c \cap B) = P(B) - P(A)P(B) =$ since A and B are independent events
$P(A^c \cap B) = P(B)(1 - P(A)) =$ factor out P(B)
$P(A^c \cap B) = P(B)P(A^c) =$ by the definition of the complement of A
Therefore, A^c and B are independent events

to show that they have positive probabilities, we can use the fact that the probability of the complement of an event is 1 minus the probability of the event:
$P(A^c) = 1 - P(A) =$ by the definition of the complement of A
since A has a positive probability, then $P(A) > 0$
therefore, $P(A^c) = 1 - P(A) > 0$
since B has a positive probability, then $P(B) > 0$
therefore, $P(A^c \cap B) = P(B)P(A^c) > 0$
(b) (2 points) A and Bc
to show that A and Bc are independent events, we can use the fact that the probability of the complement of an event is 1 minus the probability of the event:
$P(A \cap Bc) = P(A) - P(A \cap B) =$ by the definition of the complement of B
$P(A \cap Bc) = P(A) - P(A)P(B) =$ since A and B are independent events
$P(A \cap Bc) = P(A)(1 - P(B)) =$ factor out P(A)
$P(A \cap Bc) = P(A)P(Bc) =$ by the definition of the complement of B
Therefore, A and Bc are independent events

to show that they have positive probabilities, we can use the fact that the probability of the complement of an event is 1 minus the probability of the event:
$P(Bc) = 1 - P(B) =$ by the definition of the complement of B
since B has a positive probability, then $P(B) > 0$
therefore, $P(A \cap Bc) = P(A)P(Bc) > 0$
(c) (2 points) Ac and Bc

to show that Ac and Bc are independent events, we can use the fact that the probability of the complement of an event is 1 minus the probability of the event:
$P(Ac \cap Bc) = P(Ac) - P(Ac \cap B) =$ by the definition of the complement of B
$P(Ac \cap Bc) = P(Ac) - P(Ac \cap B) =$ by the definition of the complement of A
$P(Ac \cap Bc) = P(Ac) - P(Ac)P(B) =$ since A and B are independent events
$P(Ac \cap Bc) = P(Ac)(1 - P(B)) =$ factor out P(Ac)
$P(Ac \cap Bc) = P(Ac)P(Bc) =$ by the definition of the complement of B
Therefore, Ac and Bc are independent events

to show that they have positive probabilities, we can use the fact that the probability of the complement of an event is 1 minus the probability of the event:

$P(Bc) = 1 - P(B) =$ by the definition of the complement of B

since B has a positive probability, then $P(B) > 0$

therefore, $P(Ac \cap Bc) = P(Ac)P(Bc) > 0$

2. (3 points) In the ECE class of 2026 consisting of 25 students, 15 take Robotics, 10 take Introduction to Systems Software Engineering, and 5 take both. Calculate the number of students who take either Robotics or Intro. to Systems Software Engineering or both.

total population = 25
students who take Robotics = 15
students who take Introduction to Systems Software Engineering = 10
students who take both = 5 $P(R \cap E)$
we are trying to find the number of students who take either Robotics or Intro. to Systems Software Engineering or both.

$P(R \cup E) = P(R) + P(E) - P(R \cap E)$
$P(R \cup E) = 15 + 10 - 5 = 20$

therefore the number of students who take either Robotics or Intro. to Systems Software Engineering or both is 20

3. You are at Nyandungu amusement park with your friends and you want to play a game where you roll a fair six-sided die. There are prizes for rolling specific numbers:

- If you roll a 1 or a 2, you win a small prize (Event A).
- If you roll a 3, 4, or 5, you win a medium prize (Event B).
- If you roll a 6, you win a large prize (Event C).

(a) (2 points) Are Events A and B mutually exclusive? Why or why not?

Yes they are mutually exclusive because the events do not have any outcomes in common. If you roll a 1 or 2, you cannot roll a 3, 4, or 5. Therefore, the events are mutually exclusive.

(b) (2 points) Are Events A and C mutually exclusive? Why or why not? Yes they are mutually exclusive because the events do not have any outcomes in common. If you roll a 1 or 2, you cannot roll a 6. Therefore, the events are mutually exclusive.

(c) (2 points) What is the probability of winning either a small or a medium prize? Probability of winning a small prize = $P(A) = \frac{2}{6} = \frac{1}{3}$
Probability of winning a medium prize = $P(B) = \frac{3}{6} = \frac{1}{2}$
Probability of winning either a small or a medium prize = $P(A \cup B) = P(A) + P(B) = \frac{1}{3} + \frac{1}{2} = \frac{5}{6}$

answer = $\mathbf{\frac{5}{6}}$

Question 9: Combinatorics

9 questions

1. calculate the number of ways to arrange the letters in the word "ALGORITHM" such that the vowels are together.

permutations: $P(n) = \frac{n!}{(n-r)!}$

the number of letters in the word "ALGORITHM" is 8, with 3 vowels and 5 consonants.

the number of ways to arrange the vowels is $P(3) = \frac{8!}{(8-3)!} = 336$

2. (4 points) Every fall semester, elections are held at CMU-Africa to choose members of the student guild and other club representatives. After, the elections were held, the elected members conducted a survey on students problems and it was revealed that tuition funding and housing ranked first. The guild decided on decentralized task forces to address specific issues from a group of 5 females and 7 males, how many different committees consisting of 2 females and 3 males can be formed? What if 2 of the males are feuding and refuse to serve on the committee together?

Combinations to choose 2 females in a group of 5(here order doesn't matter): $C(n, r) = \frac{n!}{r!(n-r)!}$

Combinations

$C\binom{5}{2} = \frac{5!}{2!(5-2)!} = \frac{120}{12} = 10$

Combinations to choose 3 males in a group of 7(here order doesn't matter):

$C\binom{7}{3} = \frac{7!}{3!(7-3)!} = \frac{5040}{144} = 35$

Therefore the number of different committees without restrictions: $10 \times 35 = 350$

If 2 of the males refuse to work together, then the number of different committees is given by:

we need three males from the 7

let us say we are forced to choose the two of them, as in the C(2,2) = 1 we need an extra male from the remaining 5 males so that we fill up the committee slots for males

this means C(5,1) = 5, there are five different ways to form the committee since this two male can be paired up with any of the remaining gents

this means the number of invalid committees (ones with the two feuding male) is $1 \times 5 = 5$

so the total number of valid committees is 350 - 5 = 345

answer = **345**

3. A recent census summary reveals that there are about 1.75 million people in Kigali, of which approximately 55% are females. A survey conducted by a major exchange student in cosmetics reveals an alarming balding rate among men. The survey reports that the average hair count for people in Kigali is 150, 000 hairs. Men aged 35 years and above have about 35% of the average hair count. This age group constitutes 60% of the male population. Assume the hair counts for men aged 35 years and above follow a normal distribution with a mean of 52, 500 hairs and a standard deviation of 10, 000 hairs.

This is what we have so far:

- total population = 1.75 million
- female population = 0.55 of 1.75 million
- male population = 0.45 of 1.75 million
- male 35 years and above = 0.6 of 0.45 of 1.75 million

(a) (2 points) Calculate the expected number of men aged 35 years and above in Kigali.

$E(X) = \mu = 0.6 \times 0.45 \times 1.75 million = \mathbf{472,500}$

(2 points) Determine the range within which approximately 68% of the hair counts for men aged 35 years and above lie, using the given normal distribution

- $\mu = 52,500$
- $\sigma = 10,000$
- 68% of the hair counts lie within 1 standard deviation from the mean
- the range is given by: $\mu \pm \sigma$
- the range is given by: $52,500 \pm 10,000 = 42,500$ and $62,500$

3. (2 points) Considering the calculated range, use the pigeonhole principle to show that at least 2 men will have the same hair count within this range. Assume hair counts are discrete values.

the range is given by: $42,500$ and $62,500$
men in kigali each 35 and above = 472,500
the range is 20,000 + 1 = the range is inclusive
men in kigali that are likely to fall within this range = 0.86 of 472,500 = 321,300

since the range is 20,000 and the number of men is 321,300, then by the pigeonhole principle, at least 2 men will have the same hair count within this range.

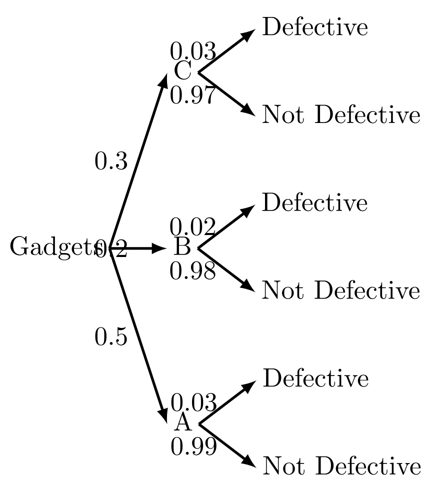
Question 10: LAW OF TOTAL PROBABILITY

Questions

1. A factory produces three types of gadgets. Type A constitutes 50%, Type B constitutes 30%, and Type C constitutes 20% of the total production. The defect rates for these gadgets are 1%, 2%, and 3% respectively. Calculate the probability that a randomly selected gadget is defective

this is what we have:

- $P(A) = 0.5$, $P(D) = 0.01$
- $P(B) = 0.3$, $P(D) = 0.02$
- $p(c) = 0.2$, $P(D) = 0.03$



a probability that a randomly selected gadget is defective is given by:

$$P(D) = P(D|A)P(A) + P(D|B)P(B) + P(D|C)P(C)$$

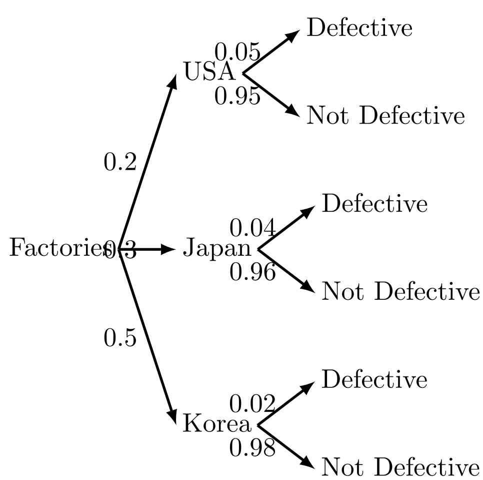
=

$$0.5 \times 0.01 + 0.3 \times 0.02 + 0.2 \times 0.03 = 0.017$$

2. (4 marks) Samsung has three factories—Factory Korea, Factory Japan, and Factory USA that produce the same type of electronic component. The factory in Korea produces 50% of the components, 30% are produced in Japan, and the factory in the USA is responsible for 20% of production. The probability of a component being defective is 2% for Factory Korea, 4% for Factory Japan, and 5% for Factory USA. Suppose a randomly selected component from the company's entire production is found to be defective. What is the probability that it was produced by Factory Japan?

This is what we have so far:

- $P(K) = 0.5$, $P(D) = 0.02$ - Korea
- $P(J) = 0.3$, $P(D) = 0.04$ - Japan
- $P(U) = 0.2$, $P(D) = 0.05$ - USA



the probability that a component was produced by Factory Japan given that it is defective is given by:

we first need to calculate the probability that a component is defective using law of total probability:

$$P(D) = P(D|K)P(K) + P(D|J)P(J) + P(D|U)P(U)$$

=

$$0.5 \times 0.02 + 0.3 \times 0.04 + 0.2 \times 0.05 = 0.032$$

$$P(J|D) = \frac{P(D|J)P(J)}{P(D)} = \frac{0.04 \times 0.3}{0.032} = \frac{0.012}{0.032} \approx 0.375$$