1. A hospital operates a screening test for a rare inherited disease among newborns.

0.1% (P(D) = 0.001) of the babies will be expected to have the disease.

The test is 99% sensitive (P(Positive | Disease) = 0.99).

It also has a 2% false positive rate (P(Positive | No Disease) = 0.02).

If an infant is found to be positive, how likely is the infant to actually have the disease?

import numpy as np

import matplotlib.pyplot as plt

# Given probabilities

P\_D = 0.001 # Prior: P(Disease)

P\_not\_D = 1 - P\_D

P\_Pos\_given\_D = 0.99 # Sensitivity

P\_Pos\_given\_not\_D = 0.02 # False Positive Rate

# Simulate repeated positive tests

n\_tests = 10

posterior\_history = [P\_D] # Start with prior

current\_posterior = P\_D

for i in range(n\_tests):

P\_Pos = P\_Pos\_given\_D \* current\_posterior + P\_Pos\_given\_not\_D \* (1 - current\_posterior)

current\_posterior = (P\_Pos\_given\_D \* current\_posterior) / P\_Pos

posterior\_history.append(current\_posterior)

# Plotting

plt.figure(figsize=(8, 5))

plt.plot(range(n\_tests + 1), posterior\_history, marker='o', linestyle='-', color='purple')

plt.xlabel('Number of Positive Tests')

plt.ylabel('P(Disease | Positive Tests)')

plt.title('Bayesian Update After Repeated Positive Tests')

plt.grid(True)

plt.ylim(0, 1)

plt.xticks(range(n\_tests + 1))

plt.show()

2. In a forensic database, 0.1% of people share a specific rare DNA marker. A forensic lab reports a match with a crime scene sample, and the test has a 99.9% accuracy rate (i.e., the chance of a false match is 0.1%).

If a person is identified by the DNA test, what is the probability they actually left the DNA at the crime scene?

import numpy as np

import matplotlib.pyplot as plt

# Given probabilities

P\_D = 0.001 # Prior: P(Disease)

P\_not\_D = 1 - P\_D

P\_Pos\_given\_D = 0.99 # Sensitivity

P\_Pos\_given\_not\_D = 0.02 # False Positive Rate

# Simulate repeated positive tests

n\_tests = 10

posterior\_history = [P\_D] # Start with prior

current\_posterior = P\_D

for i in range(n\_tests):

P\_Pos = P\_Pos\_given\_D \* current\_posterior + P\_Pos\_given\_not\_D \* (1 - current\_posterior)

current\_posterior = (P\_Pos\_given\_D \* current\_posterior) / P\_Pos

posterior\_history.append(current\_posterior)

# Plotting

plt.figure(figsize=(8, 5))

plt.plot(range(n\_tests + 1), posterior\_history, marker='o', linestyle='-', color='purple')

plt.xlabel('Number of Positive Tests')

plt.ylabel('P(Disease | Positive Tests)')

plt.title('Bayesian Update After Repeated Positive Tests')

plt.grid(True)

plt.ylim(0, 1)

plt.xticks(range(n\_tests + 1))

plt.show()

3. In a certain population, only 10% of people carry a recessive gene for red hair. Two carriers have a 25% chance of having a red-haired child. If one parent is known to carry the gene and their child has red hair, there's a 50% chance the other parent is also a carrier.

If a child has red hair and one parent is a confirmed carrier, what is the probability that the other parent is a carrier?

import numpy as np

import matplotlib.pyplot as plt

# Initial prior: Probability other parent is a carrier

prior = 0.1

P\_red\_given\_carrier = 0.25

P\_red\_given\_not\_carrier = 0.0

# Simulate belief update after multiple red-haired children

n\_children = 5

posterior\_history = [prior]

current\_posterior = prior

for i in range(n\_children):

# Total probability of red hair

P\_red = P\_red\_given\_carrier \* current\_posterior + P\_red\_given\_not\_carrier \* (1 - current\_posterior)

# Bayes' update

current\_posterior = (P\_red\_given\_carrier \* current\_posterior) / P\_red

posterior\_history.append(current\_posterior)

# Plotting the belief update

plt.figure(figsize=(8, 5))

plt.plot(range(n\_children + 1), posterior\_history, marker='o', linestyle='-', color='crimson')

plt.xlabel('Number of Red-Haired Children')

plt.ylabel('P(Other Parent is Carrier)')

plt.title('Belief Update: Probability Other Parent is a Carrier')

plt.ylim(0, 1.05)

plt.grid(True)

plt.xticks(range(n\_children + 1))

plt.show()