Clex 2: The rules of probability

Problem 1: Joint Distributions and the Fundamental Rules

Consider two discrete random variables, $X \in \{0, 1\}$ and $Y \in \{0, 1, 2\}$, with the joint probability mass function $P(X, Y)$ defined by the following table:

$P(X,Y)$	$Y = 0$	$Y = 1$	$Y = 2$
$X = 0$	0.10	0.15	0.05
$X = 1$	0.20	0.30	0.20

a) Use the **Sum Rule** (Marginalization) to calculate the marginal distributions $P(X)$ and $P(Y)$. Verify that $\sum_x P(X = x) = 1$.

Solution: To find $P(X)$, we sum across the rows:

- $P(X = 0) = 0.10 + 0.15 + 0.05 = 0.30$
- $P(X = 1) = 0.20 + 0.30 + 0.20 = 0.70$

Verification: $\sum_x P(X = x) = 0.30 + 0.70 = 1$.

To find $P(Y)$, we sum down the columns:

- $P(Y = 0) = 0.10 + 0.20 = 0.30$
- $P(Y = 1) = 0.15 + 0.30 = 0.45$
- $P(Y = 2) = 0.05 + 0.20 = 0.25$

b) Calculate the conditional probability $P(Y|X = 0)$ for all values of Y.

Solution: Using $P(Y|X = 0) = \frac{P(X=0,Y)}{P(X=0)}$:

- $P(Y = 0|X = 0) = 0.10/0.30 = 1/3 \approx 0.333$
- $P(Y = 1|X = 0) = 0.15/0.30 = 0.5$
- $P(Y = 2|X = 0) = 0.05/0.30 = 1/6 \approx 0.167$

c) Calculate the conditional expected value $E(Y|X = 0)$.

Solution: $E(Y|X = 0) = \sum_y y \cdot P(Y = y|X = 0)$

$$E(Y|X = 0) = (0 \cdot 1/3) + (1 \cdot 0.5) + (2 \cdot 1/6) = 0 + 0.5 + 1/3 = 5/6 \approx 0.833$$

d) Demonstrate the **Product Rule** by showing that $P(X = 0, Y = 1) = P(Y = 1|X = 0)P(X = 0)$.

Solution: From the table: $P(X = 0, Y = 1) = 0.15$.
Using components: $P(Y = 1|X = 0)P(X = 0) = 0.5 \times 0.30 = 0.15$.
The rule holds as both sides are equivalent.

Problem 2: Bayesian Inference and Terminology

A robotic vacuum cleaner uses a laser sensor to detect if it is in a "Clear" (C) or "Obstacle" (O) state. Based on previous navigation data, the robot has a **prior** belief that there is a 20% chance it is heading toward an obstacle ($P(O) = 0.2$). The sensor's performance is characterized by the following **likelihoods**:

- If there is an obstacle, the sensor correctly returns a "Red" signal (R) with probability $P(R|O) = 0.9$.
- If the path is clear, the sensor erroneously returns a "Red" signal with probability $P(R|C) = 0.15$.

a) Identify and state the numerical values for the **Prior** and the **Likelihood** of a "Red" signal given an obstacle.

Solution:

- Prior: $P(O) = 0.2$
- Likelihood: $P(R|O) = 0.9$

b) Calculate the **Evidence** (marginal likelihood) $P(R)$ using the Law of Total Probability.

Solution: $P(R) = P(R|O)P(O) + P(R|C)P(C)$
With $P(C) = 1 - 0.2 = 0.8$:

$$P(R) = (0.9)(0.2) + (0.15)(0.8) = 0.18 + 0.12 = 0.30$$

c) If the sensor flashes "Red", use Bayes' Theorem to calculate the **Posterior** probability that there is actually an obstacle, $P(O|R)$.

Solution:

$$P(O|R) = \frac{P(R|O)P(O)}{P(R)} = \frac{0.18}{0.30} = 0.6$$

d) Compare the posterior $P(O|R)$ to the prior $P(O)$. Qualitatively explain how the "Red" signal updated the robot's belief.

Solution: The posterior (0.6) is significantly higher than the prior (0.2). This indicates that the "Red" signal provided strong evidence of an obstacle, shifting the robot's belief state.

Problem 3: Frequentist vs. Bayesian Perspectives

A box contains two types of coins:

- **Type F**: A Fair coin ($P(H) = 0.5$).
- **Type B**: A Biased coin ($P(H) = 0.9$).

Suppose the box has 9 Fair coins and 1 Biased coin. You pick one coin at random and flip it twice, resulting in two Heads (HH).

a) **Frequentist View:** Calculate the Likelihood of the data $P(HH|\text{type})$ for both types of coins. Based strictly on which model makes the observed data most likely, which coin would a Frequentist prefer?

Solution:

- $P(HH|F) = 0.5 \times 0.5 = 0.25$
- $P(HH|B) = 0.9 \times 0.9 = 0.81$

The Frequentist (specifically using Maximum Likelihood) would prefer **Type B**, as it makes the observed data HH much more likely $(0.81 > 0.25)$.

b) **Bayesian View:** Using the distribution of coins in the box as your **Prior** $(P(F) = 0.9, P(B) = 0.1)$, calculate the **Posterior** probability that the coin you picked is actually the Biased one, $P(B|HH)$.

Solution: First, calculate the Evidence $P(HH)$:

$$P(HH) = P(HH|F)P(F) + P(HH|B)P(B)$$

$$P(HH) = (0.25)(0.9) + (0.81)(0.1) = 0.225 + 0.081 = 0.306$$

Now, use Bayes' Theorem:

$$P(B|HH) = \frac{P(HH|B)P(B)}{P(HH)} = \frac{0.081}{0.306} \approx 0.265$$

c) **Comparison:** Does the Bayesian believe the coin is biased? Why is this different from the Frequentist conclusion?

Solution: No, the Bayesian still believes there is only a $\approx 26.5\%$ chance the coin is biased. While the data (HH) points toward bias, the Bayesian "anchors" the conclusion with the Prior knowledge that biased coins are rare. The Frequentist conclusion ignores the rarity of the biased coin and focuses only on the evidence provided by the flips.