# Solutions

## Solution 1

**(a) Sampling technique analysis in the context of estimating probabilities from the given travel dataset.**

- **Direct Sampling** is suitable if the dataset provides a clear probability distribution for travel preferences. It allows for straightforward estimates of probabilities, such as how many people prefer air travel. However, if the distribution is complex (e.g., involving dependencies between stress levels and travel preferences), direct sampling may become less effective.

  - **Strengths:**
    1. Simplicity: Direct sampling is straightforward to implement when the probability distributions are fully known, as is the case with the given travel dataset.
    2. Efficiency: All generated samples are valid, leading to no wasted computational effort.
    3. Independent Samples: Each sample is drawn independently, which is beneficial for statistical analyses that assume independent observations.
  - **Weaknesses:**
    1. Requires full knowledge of the distribution, which might not always be available.
    2. Inflexibility: Difficult to apply if the distribution is complex or if there are dependencies not captured by the known probabilities.

  **Application to the Dataset:** In the context of the travel dataset, direct sampling involves generating samples based on the provided probabilities (e.g., 80% prefer air or train travel, 20% of air travelers travel for business). This method is efficient here because the necessary probabilities are explicitly given.

- **Rejection Sampling** could be beneficial if we want to sample from a more complex model of travel preferences that accounts for factors like stress. However, if the rejection rate is high (e.g., many travelers who prefer air travel do not feel stressed), this could lead to inefficiencies.

  - Strengths:
    1. Flexibility with complex distributions. Can handle sampling from distributions that are difficult to sample by using a proposal distribution.
    2. Useful when the target distribution is known up to a normalization constant.
  - Weaknesses:
    1. Potentially high computational cost due to low acceptance rates, especially if the proposal distribution is far from the target distribution.

  **Application to the Dataset:** For the travel dataset, rejection sampling would involve proposing samples (e.g., random travel preferences) and accepting them based on a criterion derived from the known probabilities. Given that direct sampling is feasible, rejection sampling may introduce unnecessary complexity and inefficiency without providing significant benefits.

- **Gibbs Sampling** would be useful if we want to explore the joint distribution of travel preferences and stress levels more thoroughly. It can model dependencies well, allowing for a better understanding of how one variable influences another. However, the method may require significant iterations to converge, making it computationally intensive.

  - Strengths:
    1. Effective for multivariate distributions where direct sampling is infeasible.

2. Only requires the ability to sample from the conditional distributions of each variable.

– Weaknesses:

1. Generates correlated samples, which may require more iterations to achieve a desired level of accuracy.
2. Computation time might increase due to the increase in the number of iterations.

**Application to the Dataset:**In the travel dataset, Gibbs sampling can be applied if we model the problem with multiple interdependent variables (e.g., mode of travel, purpose, stress level). It allows sampling from the joint distribution using the known conditional probabilities (e.g., probability of leisure travel given train preference). However, it is more computationally intensive and may be unnecessary if direct sampling is sufficient.

For estimating probabilities from the given travel dataset, direct sampling is the most appropriate method due to its efficiency and the availability of explicit probabilities. Rejection sampling and Gibbs sampling are less suitable in this context because they offer no significant advantages and may increase computational complexity without improving the accuracy of the estimates.

**(b) Estimating number of train travelers for Leisure in a Sample**

- $P(\text{Leisure} \mid \text{Train}) = 0.400$
- Expected number of leisure travelers $= 30 \times 0.400 = 12$ (accepting based on probability).

**(c) Probability that a random person prefers air travel and travels for business**

- $P(\text{Air Travel}) = 0.80$ and $P(\text{Business} \mid \text{Air Travel}) = 0.20$
- $P(\text{Air Travel} \cap \text{Business}) = P(\text{Air Travel}) \times P(\text{Business} \mid \text{Air Travel}) = 0.80 \times 0.20 = 0.160$

**(d) Sample Size Impacts**

- **Accuracy:** It refers to how close an estimate is to the true population parameter.
  - **Effect of increased sample size:** Increasing the sample size generally enhances the accuracy of estimates obtained through direct sampling. A larger sample is more likely to represent the true underlying population, reducing the bias associated with random fluctuations. As the sample size increases, the sample mean or proportion approaches the true population parameter due to the Law of Large Numbers.
  - **Implications for the Dataset:** For the travel preferences dataset, if the sample size increases, estimates such as the proportion of people who prefer air travel or the probability of feeling stressed while traveling become more reliable. A larger sample would likely provide a better representation of the different travel preferences (e.g., air, train, bus) across various demographics.
- **Precision:** Precision refers to the variability of the estimates;
  - **Effect of increased sample size:** With an increased sample size, the confidence intervals for estimates become narrower, indicating less variability and greater certainty about the estimates. This means that repeated samples will yield results that are closer to each other.
  - **Implications for the Dataset:** In the context of the travel dataset, a larger sample size would result in more precise estimates of the probabilities associated with different travel modes and reasons for travel (e.g., business vs. leisure). This increased precision allows for more informed decision-making, such as targeted marketing strategies or policy implementations based on accurate assessments of traveler behavior.

## Solution 2

### (a) Random Variables and Mathematical Representation

The random variables are:

- $b$: Event that a person **reads books**.
- $j$: Event that a person **accesses academic journals**.
- $c$: Event that a person **participates in book clubs**.

Using these variables, we can represent each statement mathematically:

1. $P(b \cup j) = 0.910$
2. $P(j \mid b) = 0.40 \quad \text{and} \quad P(\neg j \mid b) = 0.60$
3. $P(c \mid b) = 0.320$
4. $P(j \cap \neg b) = 0.227$
5. $P(\neg b \cap \neg j) = 0.090$
6. $P(j \mid \neg b) = 0.716$
7. $P(c \cap j) = 0.088$
8. $P(c \cup j) = 0.631$
9. $P(j \mid c) = 0.400$
10. $P(j) = 0.500$
11. $P(c \mid \neg b) = 0.0044 \quad \text{(Additional Statement)}$

### (b) Valid Probability Distribution and Listing of Axioms

To verify that these propositions create a valid probability distribution, we ensure they satisfy the three axioms of probability:

1. **Non-negativity**: $P(E) \geq 0$ for any event $E$.
2. **Normalization**: $P(\Omega) = 1$, where $\Omega$ is the sample space.
3. **Additivity**: For mutually exclusive events $A$ and $B$, $P(A \cup B) = P(A) + P(B)$.

We will compute all joint probabilities to confirm that they are non-negative and sum to 1.

### (c) Full Joint Probability Distribution Table

#### Step 1: Compute Marginal Probabilities

- From Statement 4 and 5:

$$P(\neg b) = P(\neg b \cap j) + P(\neg b \cap \neg j) = 0.227 + 0.090 = 0.317$$

  Therefore:

$$P(b) = 1 - P(\neg b) = 1 - 0.317 = 0.683$$

- From Statement 2:

$$P(b \cap j) = P(b) \times P(j \mid b) = 0.683 \times 0.40 = 0.2732$$

$$P(b \cap \neg j) = P(b) - P(b \cap j) = 0.683 - 0.2732 = 0.4098$$

- Verify $P(j)$:

$$P(j) = P(b \cap j) + P(\neg b \cap j) = 0.2732 + 0.227 = 0.5002 \approx 0.500$$

- From Statement 9:

$$P(c) = \frac{P(c \cap j)}{P(j \mid c)} = \frac{0.088}{0.400} = 0.220$$

**Step 2: Compute Joint Probabilities Involving $c$**

**For individuals who read books ($b$)**

- From Statement 3:

$$P(c \mid b) = 0.320$$

- Compute:

$$P(b \cap j \cap c) = P(b \cap j) \times P(c \mid b) = 0.2732 \times 0.320 = 0.0874$$

$$P(b \cap j \cap \neg c) = P(b \cap j) - P(b \cap j \cap c) = 0.2732 - 0.0874 = 0.1858$$

$$P(b \cap \neg j \cap c) = P(b \cap \neg j) \times P(c \mid b) = 0.4098 \times 0.320 = 0.1311$$

$$P(b \cap \neg j \cap \neg c) = P(b \cap \neg j) - P(b \cap \neg j \cap c) = 0.4098 - 0.1311 = 0.2787$$

**For individuals who do not read books ($\neg b$)**

- From Additional Statement 11:

$$P(c \mid \neg b) = 0.0044$$

- Compute:

$$P(\neg b \cap j \cap c) = P(\neg b \cap j) \times P(c \mid \neg b) = 0.227 \times 0.0044 = 0.0010$$

$$P(\neg b \cap j \cap \neg c) = P(\neg b \cap j) - P(\neg b \cap j \cap c) = 0.227 - 0.0010 = 0.2260$$

$$P(\neg b \cap \neg j \cap c) = P(\neg b \cap \neg j) \times P(c \mid \neg b) = 0.090 \times 0.0044 = 0.0004$$

$$P(\neg b \cap \neg j \cap \neg c) = P(\neg b \cap \neg j) - P(\neg b \cap \neg j \cap c) = 0.090 - 0.0004 = 0.0896$$

**Step 3: Compile the Joint Probability Distribution Table**

| $b$ | $j$ | $c$ | $P(b,j,c)$ |
|---|---|---|---|
| 1 | 1 | 1 | 0.087 |
| 1 | 1 | 0 | 0.186 |
| 1 | 0 | 1 | 0.131 |
| 1 | 0 | 0 | 0.279 |
| 0 | 1 | 1 | 0.001 |
| 0 | 1 | 0 | 0.226 |
| 0 | 0 | 1 | 0.000 |
| 0 | 0 | 0 | 0.090 |

**Verification:**

$$\sum P(b,j,c) = 0.087 + 0.186 + 0.131 + 0.279 + 0.001 + 0.226 + 0.000 + 0.090 = 1.000$$

All probabilities are non-negative and sum up to 1, satisfying the axioms of probability.

**(d) Checking for Conditional Independence**

**1. Is $c$ independent of $j$ given $b$?**

- Compute $P(c \mid b, j)$:

$$P(c \mid b, j) = \frac{P(b \cap j \cap c)}{P(b \cap j)} = \frac{0.087}{0.273} = 0.319$$

- From Statement 3:

$$P(c \mid b) = 0.320$$

- Since $P(c \mid b, j) \approx P(c \mid b)$, $c$ is conditionally independent of $j$ given $b$.

**2. Is $c$ independent of $b$ given $j$?**

- Compute $P(c \mid j, b)$:

$$P(c \mid j, b) = P(c \mid b) = 0.320$$

- Compute $P(c \mid j)$:

$$P(c \mid j) = \frac{P(c \cap j)}{P(j)} = \frac{0.088}{0.500} = 0.176$$

- Since $P(c \mid j, b) \neq P(c \mid j)$, $c$ is not conditionally independent of $b$ given $j$.

**3. Is $b$ independent of $j$ given $c$?**

- Compute $P(b \mid c, j)$:

$$P(b \mid c, j) = \frac{P(b \cap c \cap j)}{P(c \cap j)} = \frac{0.087}{0.088} = 0.989$$

- Compute $P(b \mid c)$:

$$P(b \mid c) = \frac{P(b \cap c)}{P(c)} = \frac{0.087 + 0.131}{0.220} = \frac{0.218}{0.220} = 0.991$$

- Since $P(b \mid c, j) \approx P(b \mid c)$, $b$ is conditionally independent of $j$ given $c$.

**4. Is $b$ independent of $c$ given $j$?**

- Compute $P(b \mid j, c)$:

$$P(b \mid j, c) = \frac{P(b \cap j \cap c)}{P(j \cap c)} = \frac{0.087}{0.088} = 0.989$$

- Compute $P(b \mid j)$:

$$P(b \mid j) = \frac{P(b \cap j)}{P(j)} = \frac{0.273}{0.500} = 0.546$$

- Since $P(b \mid j, c) \neq P(b \mid j)$, $b$ is not conditionally independent of $c$ given $j$.

**Conclusion:**

- $c$ is conditionally independent of $j$ given $b$.
- $c$ is not conditionally independent of $b$ given $j$.
- $b$ is conditionally independent of $j$ given $c$.
- $b$ is not conditionally independent of $c$ given $j$.

Similarly it can be verified that, j is independent of c given b and b given c.

## Solution 3

### (a) Problem Formulation

We define the following events:

- $A$: Misclassification due to an adversarial perturbation.
- $B$: Misclassification due to a backdoor attack.
- $M$: Misclassification alarm is triggered.

Initially, adversarial perturbations and backdoor attacks are considered independent, so:

$$P(A \cap B) = P(A)P(B)$$

Our goal is to compute the posterior probability that the misclassification was caused by an adversarial perturbation, given that there is new information about the increased prevalence of backdoor attacks. We need to update our belief using Bayes' Theorem:

$$P(A \mid M) = \frac{P(M \mid A)P(A)}{P(M)}$$

We also have:

$$P(M) = P(M \mid A)P(A) + P(M \mid B)P(B)$$

### (b) Probabilities Involved

- $P(A)$: The prior probability of an adversarial perturbation causing a misclassification.
- $P(B)$: The prior probability of a backdoor attack causing a misclassification.
- $P(M \mid A)$: The probability of observing a misclassification alarm given that an adversarial perturbation occurred.
- $P(M \mid B)$: The probability of observing a misclassification alarm given that a backdoor attack occurred.

### (c) Belief Change

After hearing about the prevalence of backdoor attacks, $P(B)$ increases, leading to a change in our belief about $P(A \mid M)$. Specifically, conditioning on the alarm and the new information about backdoor attacks increases the likelihood that a backdoor attack caused the misclassification, reducing our belief that it was due to an adversarial perturbation.