Homework 1: Probability and Priors - Hilde Younce

Problem 1: Basic Probability

Alice has a bag with 3 red balls, 2 green balls, and 5 blue balls.

(a) What is the probability of drawing a red ball from the bag?

(b) If Alice draws one ball and it's blue, what's the probability that the next ball she draws is also blue?

a) $\frac{3}{3+2+5} = \boxed{\frac{3}{10}}$

b) $\frac{5}{10}\left(\frac{4}{9}\right) = \frac{20}{90} = \boxed{\frac{2}{9}}$

Problem 2: Independent Events

The probability of a server being down in a data center is 0.05. The data center is designed such that server failures are independent events.

(a) What is the probability that 2 servers will be down at the same time?

(b) What is the probability that at least one of two servers will be down?

a) $P(\text{server 1 down} \cap \text{server 2 down}) = P(\text{server 1 down})P(\text{server 2 down})$
$= (0.05)(0.05)$
$= \boxed{0.0025}$

b) $P(\text{at least 1 server down}) = 1 - P(\text{no servers are down})$
$= 1 - P(\text{server 1 not down} \cap \text{server 2 not down})$
$= 1 - P(\text{server 1 not down})\, P(\text{server 2 not down})$
$= 1 - (1 - P(\text{server 1 down})) \cdot (1 - P(\text{server 2 down}))$
$= 1 - (1-0.05)(1-0.05)$
$= 1 - (0.95)(0.95)$
$= 1 - 0.9025$
$= \boxed{0.0975}$

Problem 3: Conditional Probability

In a Machine Learning company, 30% of the employees are Data Scientists, 40% of the Data Scientists have PhDs, while only 10% of non-Data Scientists have PhDs.

(a) If an employee is chosen randomly, what is the probability that the employee is a Data Scientist with a PhD?

(b) Given that an employee has a PhD, what is the probability that the employee is a Data Scientist?

a) $\frac{4}{10} = \frac{x}{30} = \frac{120}{10} = 12 \Rightarrow \boxed{0.12}$

b) $p(\text{data scientist} \mid \text{has phd}) = \frac{p(\text{data scientist} \cap \text{has phd})}{p(\text{has phd})}$

$p(\text{has phd}) = p(\text{phd} \cap \text{data scientist}) + p(\text{phd} \cap \text{not data scientist})$

$= 0.12 + 0.07 = 0.19$

$\Rightarrow p(\text{data scientist} \mid \text{has phd}) = \frac{0.12}{0.19} = \boxed{0.63}$

Problem 4: Law of Total Probability

A diagnostics test has a probability of 0.95 of giving a positive result when applied to a person suffering from a certain disease. It has a probability of 0.10 of giving a (false) positive result when applied to a non-sufferer. It is estimated that 0.5% of the population has this disease.

(a) If a person tested positive in the test, what is the probability that the person actually has the disease?

(b) What is the total probability of a person testing positive?

a) A - has disease, B - test is positive

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

- $P(B \mid A) = 0.95$
- $P(A) = 0.5$
- $P(B) = P(B \mid A)P(A) + P(B \mid A^c)P(A^c) = 0.95 \cdot 0.5 + 0.10 \cdot 0.5 = 0.525$

$\Rightarrow P(A \mid B) = \frac{0.95\,(0.5)}{0.525} = \boxed{0.905}$

b) $P(B) = P(B \mid A)P(A) + P(B \mid A^c)P(A^c) = \boxed{0.525}$

Problem 5: Bayes' Theorem

An email filter is set up to classify emails into "spam" and "not spam". It is known that 90% of all emails received are spam. The filter correctly identifies spam 95% of the time and correctly identifies "not spam" 85% of the time.

(a) If an email is picked at random, and the filter classifies it as spam, what is the probability that it is actually spam?

(b) If an email is classified as "not spam", what is the probability that it is actually spam?

a) A - actually spam, B - filter says spam

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

- $P(B \mid A) = 0.95$
- $P(A) = 0.9$
- $P(B) = P(B \mid A)P(A) + P(B \mid A^c)P(A^c) = 0.95 \cdot 0.9 + 0.15 \cdot 0.1 = 0.87$

$\Rightarrow P(A \mid B) = \frac{0.95 \cdot 0.9}{0.87} = \boxed{0.983}$

b) A - actually spam , B - filter says not spam

- $P(B|A) = 0.15$
- $P(A) = 0.9$
- $P(B) = P(B|A)P(A) + P(B|A^c)P(A^c) = 0.9 \cdot 0.15 + 0.85 \cdot 0.1 = 0.22$

$\Rightarrow P(A|B) = \frac{0.15(0.9)}{0.22} = \boxed{0.61}$

Problem 6: Expectation of a Discrete Random Variable

Consider a dice game where you roll a fair six-sided die. If a 6 appears, you win $10. If any other number appears, you lose $2.

(a) Define the random variable X that models this game.

(b) Compute the expected value of X.

a) Let 1, 2, ..., 6 be the sides of the dice
then $\boxed{X_1 = X_2 = X_3 = X_4 = X_5 = -2 \text{ and } X_6 = 10}$

b) $E[X] = \sum_{i=1}^{6} X_i P(X_i) = \frac{1}{6}(-2) + \frac{1}{6}(-2) + \frac{1}{6}(-2) + \frac{1}{6}(-2) + \frac{1}{6}(-2) + \frac{1}{6}(10) = \boxed{0}$

Problem 7: Expectation of a Continuous Random Variable

Let X be a continuous random variable representing the time (in hours) it takes for a server to process a certain type of query. Suppose the density function of X is given by $f(x) = 2e^{-2x}$ for $x \geq 0$.

(a) Compute the expected value $E[X]$ of X.

(b) Compute the variance $Var[X]$ of X.

(c) Interpret your findings from parts (a) and (b) in the context of the server's processing time.

a) $E[X] = \int_0^\infty x \cdot 2e^{-2x} \, dx$ $\quad u = x, \; dv = 2e^{-2x} \Rightarrow v = -e^{-2x}$

$= -xe^{-2x} - \int_0^\infty -e^{-2x} \, dx = -xe^{-2x} - \frac{1}{2}e^{-2x} \Big|_0^\infty$

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

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

$E[X^2] = \int_0^\infty x^2 \cdot 2e^{-2x}$ $\quad u = x^2, \; dv = 2e^{-2x} \Rightarrow v = -e^{-2x}$

$= -x^2e^{-2x} - \int_0^\infty -2xe^{-2x} \, dx = -x^2e^{-2x} - xe^{-2x} - \frac{1}{2}e^{-2x} \Big|_0^\infty$

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

$Var[X] = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{2} - \frac{1}{4} = \boxed{\frac{1}{4}}$

c) From a/b, we can say that the average time the server takes to process a query is 30 minutes. With a variance of 1/4, our standard deviation is 1/2 an hour, meaning that most compute times will fall between 30 mins of the mean.

Problem 8: Markov Chain

(a) Construct the transition matrix for this Markov chain.

After many iterations or steps, the probabilities of being in each state may stabilize to a constant value. These constant values form the *stationary distribution* of the Markov chain. To compute the stationary distribution, find the probability vector that remains unchanged after multiplication with the transition matrix.

(b) If today is sunny, what is the probability that it will be rainy two days from now?

(c) Find the stationary distribution of this Markov chain.

(d) Interpret the stationary distribution in the context of this weather model.

a) $$P = \begin{array}{c} \\ \text{sunny} \\ \text{cloudy} \\ \text{rainy} \end{array} \begin{array}{c} \begin{array}{ccc} \text{sunny} & \text{cloudy} & \text{rainy} \end{array} \\ \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{bmatrix} \end{array}$$

b) $P(\text{Day } 2 = \text{Rainy} \mid \text{Day } 0 = \text{Sunny}) =$

$$\begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{bmatrix}^2 = \begin{bmatrix} 0.57 & 0.25 & 0.18 \\ 0.39 & 0.31 & 0.3 \\ 0.33 & 0.31 & 0.36 \end{bmatrix} \Rightarrow [1 \quad 0 \quad 0]\begin{bmatrix} 0.57 & 0.25 & 0.18 \\ 0.39 & 0.31 & 0.3 \\ 0.33 & 0.31 & 0.36 \end{bmatrix}$$

$= [0.57 \quad 0.25 \quad 0.18] \Rightarrow P(\text{rainy in 2 days} \mid \text{sunny today}) = \boxed{0.18}$

c) A stationary distribution is a probability vector $\pi = (\pi_1, \pi_2, \pi_3)$ such that $\pi P = \pi$

$$\Rightarrow [\pi_1 \quad \pi_2 \quad \pi_3]\begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{bmatrix} = [\pi_1 \quad \pi_2 \quad \pi_3]$$

$$\Rightarrow \left.\begin{aligned} 0.7\pi_1 + 0.3\pi_2 + 0.2\pi_3 &= \pi_1 \\ 0.2\pi_1 + 0.4\pi_2 + 0.3\pi_3 &= \pi_2 \\ 0.1\pi_1 + 0.3\pi_2 + 0.5\pi_3 &= \pi_3 \end{aligned}\right\} \Rightarrow \begin{aligned} &0.7\pi_1 + 0.3\pi_2 + 0.2\pi_3 = \pi_1 \\ -&(0.1\pi_1 + 0.3\pi_2 + 0.5\pi_3) = \pi_3 \\ \hline &0.6\pi_1 - 0.3\pi_3 = \pi_1 - \pi_3 \end{aligned} \Rightarrow 0.4\pi_1 = 0.7\pi_3$$

$\Rightarrow \pi_1 = 1.75\pi_3 \Rightarrow 0.2(1.75\pi_3) + 0.4\pi_2 + 0.3\pi_3 = \pi_2 \Rightarrow 0.4\pi_2 + 0.65\pi_3 = \pi_2$

$\Rightarrow \pi_2 = 1.08\pi_3$ Since π is a probability vector we also have $\pi_1 + \pi_2 + \pi_3 = 1$

$\Rightarrow 1.75\pi_3 + 1.08\pi_3 + \pi_3 = 1 \Rightarrow 3.83\pi_3 = 1 \Rightarrow \underline{\pi_3 = 0.26}$

$$\left.\begin{aligned} &\Rightarrow \pi_1 = 1.75\pi_3 \Rightarrow \pi_1 = 1.75(0.26) = 0.46 \\ &\pi_2 = 1.08\pi_3 \Rightarrow \pi_2 = 1.08(0.26) = 0.28 \end{aligned}\right\} \boxed{\pi = (0.46, 0.28, 0.26)}$$

d) The stationary distribution represents the percentage of time the Markov chain spends in each state in the long run. For our model, this means that in the long run we expect about half of days to be sunny, a quarter of them cloudy, and a quarter of them rainy.

Problem 9: Conjugate Priors and Posterior Distribution

(a) Suppose the prior distribution for θ (the recovery rate) is Beta(2, 2). Calculate the posterior distribution after observing the results of the experiment.

(b) Based on the posterior distribution, provide an estimate for θ.

(c) Explain the role of the conjugate prior in simplifying the calculation of the posterior distribution.

a) posterior distribution : $\text{Beta}(\alpha+k,\ n-k+\beta)$ with prior: Beta (2, 2)

$\alpha = \beta = 2,\quad n = 100,\quad k = 30$

$\Rightarrow \text{Beta}(2+30,\ 100-30+2) = \boxed{\text{Beta}(32, 72)}$

b) $\text{EAP} = \dfrac{\alpha}{\alpha+\beta} = \dfrac{32}{32+72} = \boxed{0.308}$

c) A conjugate prior serves to simplify the calculation of the posterior distribution by complimenting the likelihood so that once the algebra has been calculated, the posterior is equal to the same distribution as the prior.