Assignment 4

jeudi 10 mars 2022 17:37

1. *Facebook interview question.* Facebook has a content team that labels pieces of content on the platform as spam or not spam. 90% of them are diligent raters and will label 20% of the content as spam and 80% as non-spam. The remaining 10% are non-diligent raters and will label 0% of the content as spam and 100% as non-spam. Assume the pieces of content are labeled independently from one another, for every rater. Given that a rater has labeled 4 pieces of content as non-spam, what is the probability that they are a diligent rater?

$D = \{\text{"Diligent"}\}$ $D^c = \{\text{"non-diligent"}\}$

$S = \{\text{"labeled spam"}\}$ $S^c = \{\text{"labeled non-spam"}\}$

$\mathbb{P}(D) = 0.9$

For 1 piece:

$\mathbb{P}(S \mid D) = 0.2$ $\mathbb{P}(S^c \mid D) = 0.8$, $\mathbb{P}(S \mid D^c) = 0$ $\mathbb{P}(S^c \mid D^c) = 1$

For 4 pieces:

$\because$ Pieces labeled individually: $\mathbb{P}(S^c \mid D)^4 = 0.4096$, $\mathbb{P}(S^c \mid D^c)^4 = 1$

$\mathbb{P}(D \mid S^c) = \dfrac{\mathbb{P}(S^c \mid D)\,\mathbb{P}(D)}{\mathbb{P}(S^c)}$ $\rightarrow \mathbb{P}(S^c) = \mathbb{P}(S^c \mid D)\mathbb{P}(D) + \mathbb{P}(S^c \mid D^c)\mathbb{P}(D^c)$
$= 0.468$

$\mathbb{P}(D \mid S^c) = 0.7866$

2. *Microsoft interview question.* Three friends in Vancouver each told you it's rainy, and each person has a 1/3 probability of lying. What is the probability that Vancouver is rainy? Assume the probability of rain on any given day in Vancouver is 0.25.

$R = \{\text{"Rainy"}\}$ $R^c = \{\text{"Not Rainy"}\}$ 0.7273

$T = \{\text{"Told rainy"}\}$ $T^c = \{\text{"Told not Rainy"}\}$

For 1 person: $\mathbb{P}(T \mid R) = 2/3$ $\mathbb{P}(T^c \mid R) = 1/3$

For 3 people: $\mathbb{P}(T \mid R)^3 = 0.296$, $\mathbb{P}(T^c \mid R)^3 = 0.037$

$\mathbb{P}(R) = 0.25$

$\mathbb{P}(R \mid T) = \dfrac{\mathbb{P}(T \mid R)\,\mathbb{P}(R)}{\mathbb{P}(T)}$ $\mathbb{P}(T) = \mathbb{P}(T \mid R)\mathbb{P}(R) + \mathbb{P}(T \mid R^c)\mathbb{P}(R^c)$
$= 0.10175$

$= 0.7272$

3. *Lyft interview question.*

   (a) You and your friend are playing a game. The two of you will continue to toss a coin until the sequence HH or TH shows up. If HH shows up first, you win. If TH shows up first, your friend wins. What is the probability of you winning?

   (b) In part (a), instead of the sequence HH or TH, we now consider the sequence HHT or HTT. Your friend wins if HHT shows up first. What is the probability of you winning?

a) H_1 = 1st head H_2 = 2nd head ...

$\mathbb{P}(H_1) = \mathbb{P}(T_1) = 0.5$, $\mathbb{P}(H_2) = \mathbb{P}(T_2) = 0.5$

↳ If T flipped, sequence TH will appear eventually

↳ If T flipped at any time, will lose eventually

$\therefore$ Can only win if 1st & 2nd flips = H

$\mathbb{P}(H_1 \cap H_2) = \mathbb{P}(H_1)\,\mathbb{P}(H_2) = (0.5)(0.5) = 0.25$

b) HHT & HTT both need H_1 → need HT vs TT

$\mathbb{P}(H) = \mathbb{P}(T) = 0.5$

→ Fail → Fails

↳ Possible combinations w/ H_1: HHH, HTT, HHT, HTH

↳ Fails eventually ↳ resets

win eventually

$\therefore \mathbb{P}(HTT) = 1/3$

$\therefore \mathbb{P}(HTT) = 1/3$

↳ fails eventually → resets

win eventually

4. *Balls and bins.* There are 10 bins numbered $1, 2, \ldots, 10$. n balls are thrown into the 10 bins. For each ball, the probability that it falls into bin i is $\frac{1}{10}$ for $i = 1, 2, \ldots, 10$. Different balls are thrown independently of each other. Let Y be the number of balls in bin 1. Let Z be the total number of balls in bins $6, 7, 8, 9, 10$.

   (a) Find the pmf of Y.

   (b) Find the pmf of Z.

   (c) Find the conditional pmf $p(z|y)$. Please specify the range of y, z.

   (d) Find the conditional pmf $p(y|z)$. Please specify the range of y, z.

a) pmf Y: $Y \sim Bin(n, p = 1/10)$

$\mathbb{P}(Y=y) = C_y^n \, p^y (1-p)^{n-y} = C_y^n (1/10)^y (9/10)^{n-y}$

b) pmf Z: $\because$ 5 bins, $p = \sum^5 p = 5/10 \rightarrow Z \sim Bin(n, p = 5/10)$

$\mathbb{P}(Z=z) = C_y^n (5/10)^y (5/10)^{n-y} = C_y^n (5/10)^n$

c) Given $Y=y$, $Z \sim Bin(n-y, p = 5/9)$

$p(z|y) = \mathbb{P}(Z=z \mid Y=y) = C_z^{n-y} (5/9)^z (4/10)^{n-y-z}$, $0 < z < n-y$

d) Given $Z=z$, $Y \sim Bin(n-z, p = 1/5)$

$p(y|z) = \mathbb{P}(Y=y \mid Z=z) = C_y^{n-z} (1/5)^y (4/5)^{n-z-y}$, $0 < y < n-z$

5. *Polya's urn.* Suppose we have an urn containing one red ball and one blue ball. We draw a ball at random from the urn. If it is red, we put the drawn ball plus another red ball into the urn. If it is blue, we put the drawn ball plus another blue ball into the urn. We then repeat this process. At the n-th stage, we draw a ball at random from the urn with $n+1$ balls, note its color, and put the drawn ball plus another ball of the same color into the urn.

   (a) Find the probability that the first ball is red.

   (b) Find the probability that the second ball is red.

   (c) Find the probability that the first three balls are all red.

   (d) Find the probability that two of the first three balls are red.

a) For urn w/ 1 R & 1 B balls, $\mathbb{P}(R) = 1/2$

b) Draw 1: R $1/2$ 2|3; B $1/2$ 2|3

Draw 2: RR $(1/2)(2/3)$ 3|4; RB $(1/2)(1/3)$ 2|4; BR $(1/2)(1/3)$ 2|4; BB $(1/2)(2/3)$ 3|4

$\therefore \mathbb{P}(R_2) = \mathbb{P}(RR) + \mathbb{P}(BR) = (1/2)(2/3) + (1/2)(1/3) = 1/2$

c) Draw 3: RRR $(1/2)(2/3)(3/4)$ $\therefore \mathbb{P}(RRR) = (1/2)(2/3)(3/4) = 1/4$

d) Draw 3: RRB $(1/2)(2/3)(1/4)$; RBR $(1/2)(1/3)(2/4)$; BRR $(1/2)(1/3)(2/4)$

$\therefore \mathbb{P}(2/3\ R) = \mathbb{P}(RRB) + \mathbb{P}(RBR) + \mathbb{P}(BRR) = 1/12 + 1/12 + 1/12 = 1/4$

6. *Coin with random bias.* You are given a coin but are not told what its bias (probability of heads) is. You are told instead that the bias is the outcome of a random variable $P \sim \text{Unif}[0, 1]$. To get more infromation about the coin bias, you flip it independently 10 times. Let X be the number of heads you get. Thus $X \sim \text{Binom}(10, P)$. Assuming that $X = 9$, find and sketch the conditional pmf of P, i.e., $f_{P|X}(p|9)$.

$P \sim \text{Unif}[0, 1]$, $X \sim \text{Binom}(10, P)$

$P(X=9) = \underbrace{C^{10}_{9}}_{\text{const.}} p^9 (1-p)^{10-9} = p^9(1-p)$

$f_{P|X}(p|9) = \beta\, p^9 (1-p) \longrightarrow 0 < p < 1$

$\therefore f_{P|X}(p|9) \sim \text{Beta}(10, 2)$