Name : DOAN QUOC SY VO

ID : 150201133

Module : CS596

HW : 1

Problem 1).

	Orange	Apple	Cherry	Total
Red	1	4	5	10
Blue	7	2	1	10
Green	3	4	3	10
Total	11	10	9	30

Box

Denote Red: R, Blue: B, Green: G, Orange: O, Apple: A, Cherry: C

(a) $P(O) = P(O|R)P(R) + P(O|G)P(G) + P(O|B)P(B)$

$= \frac{11}{30}$

$P(R|O) = P(O|R)P(R) / P(O) = \frac{1}{11}$

$P(G|O) = \frac{P(O|G)P(G)}{P(O)} = \frac{3}{11}$

$P(B|O) = \frac{P(O|B)P(B)}{P(O)} = \frac{7}{11}$

(b) $P(C) = P(C|R)P(R) + P(C|G)P(G) + P(C|B)P(B)$

$= 7/25$

$P(R|C) = P(C|R)P(R) / P(C) = \frac{5}{14}$

$P(G|C) = P(C|G)P(G) / P(C) = \frac{15}{28}$

$P(B|C) = P(C|B)P(B) / P(C) = \frac{3}{28}$

Problem 2) $P(\text{defective}) = 30\% \implies P(\sim D) = 70\%$

Test 1 result

$TP_1 = P(\text{positive} \mid D) = 80\% \implies FN_1 = P(\text{Negative} \mid D) = 20\%$

$TN_1 = P(\text{negative} \mid \sim D) = 80\% \implies FP_1 = P(\text{Positive} \mid \sim D) = 20\%$

Test 2 result:

$TP_2 = P(\text{positive} \mid D) = 90\% \implies FN_2 = P(\text{negative} \mid D) = 10\%$

$TN_2 = P(\text{negative} \mid \sim D) = 70\% \implies FP_2 = P(\text{Positive} \mid \sim D) = 30\%$

(a) Base on test **1** result:

$$P(D \mid \text{positive}) = \frac{P(\text{positive} \mid D) \times P(D)}{P(\text{positive})} = \frac{0.8 \times 0.3}{P(\text{positive})}$$

$$P(\text{positive}) = P(\text{positive}, D) + P(\text{positive}, \sim D)$$
$$= P(\text{positive} \mid D) \, P(D) + P(\text{positive} \mid \sim D) \, P(\sim D)$$
$$= 0.8 \times 0.3 + 0.2 \times 0.7 = 0.38$$

$$\implies P(D \mid \text{positive}) = \frac{0.24}{0.38} = 0.632$$

Similarly, base on test **2** result:

$$\implies P(\text{positive}) = 0.9 \times 0.3 + 0.3 \times 0.7 = 0.48$$

$$\implies P(D \mid \text{positive}) = \frac{0.9 \times 0.3}{0.48} = 0.5632$$

(b) $$P(D \mid \text{negative}) = \frac{P(\text{negative} \mid D) \, P(D)}{P(\text{negative} \mid D) \, P(D) + P(\text{negative} \mid \sim D) \, P(\sim D)}$$

Base on test result 1:

$$P(D \mid \text{negative}) = \frac{0.2 \times 0.3}{0.2 \times 0.3 + 0.8 \times 0.7} = \frac{0.06}{0.62} = 0.0967$$

Base on test result 2:

$$P(D \mid \text{negative}) = \frac{0.1 \times 0.3}{0.1 \times 0.3 + 0.7 \times 0.7} = \frac{0.03}{0.52} = 0.0576$$

(c) Assume total Walmart can sell X cameras

$$\text{Defective} = 0.3X \quad, \text{non Defective} = 0.7X.$$

The total cost of both "missdetection" & "false alarm" :

Totalcost = $12 × "miss detection" units + $6 × "false alarm" units

$$= \left[12 \times P(D|negative) \times P(negative) + 6 \times P(positive|\sim D) \times P(\sim D)\right] \times X$$

* Base on test 1 result:

$$\frac{\text{Totalcost}}{X} = 12 \times 0.0967 \times 0.62 + 6 \times 0.2 \times 0.7 = 1.554$$

$$\Rightarrow \text{Total cost}_{(1)} = \$1.559 X$$

* Base on test 2 result:

$$\frac{\text{Totalcost}}{X} = 12 \times 0.0576 \times 0.52 + 6 \times 0.3 \times 0.7 = 1.615$$

$$\Rightarrow \text{Total cost}_{(2)} = \$1.619 X$$

Base on total cost calculated of both test 1 & test 2, we found test 2 actually costs more money than test 1. Hence, **test 1 should be the preferred method.**

Problem 3)

Name	Healthy Bun.	Burgerlet	YourBurg	Fat Burger	Greesy Joint
Fat (X)	19	31	35	39	43.
Calories(Y)	410	580	570	640	660

We need to find out the slope (w_1) & the interceipt (w_0) in the linear equation $Y = w_0 + w_1 X$

Applying the formulars:

$$w_1 = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}$$

$$= \frac{5(19 \times 410 + 31 \times 580 + 35 \times 570 + 39 \times 640 + 43 \times 660) - \sum x \sum y}{5(19^2 + 31^2 + 35^2 + 39^2 + 43^2) - (19 + 31 + 35 + 39 + 43)^2}$$

$$= \frac{495300 - (19 + 31 + 35 + 39 + 43) \times (410 + 580 + 570 + 640 + 660)}{29585 - 27889}$$

$$= \frac{17680}{1696} = 10.4245283$$

$$w_0 = \frac{\sum y - w_1 \sum x}{n} = \frac{(410 + 580 + 570 + 640 + 660) - 10.4245283 \times (19 + 31 + 35 + 39 + 43)}{5}$$

$$= \frac{1119.1037739}{5} = 223.8207$$

Hence, coef = 10.42.

        intercept = 223.82.

Problem 4) Please refer to hw1_150201133.py