Name : DOAN QUOC SY VO

ID : 150201133

Module : CS596

HW : 1

Problem 1).

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

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

(a) $P(O|R) = \frac{1}{10}$; $P(O|G) = \frac{3}{10}$; $P(O|B) = \frac{7}{10}$

(b) $P(C, R) = P(C|R) \times P(R)$

$= \frac{5}{10} \times 0.2 = 0.1$

$P(C, G) = P(C|G) \times P(G)$

$= \frac{3}{10} \times 0.5 = 0.15$

$P(C, B) = P(C|B) \times P(B)$

$= 0.1 \times 0.3 = 0.03$

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

Test 1 result

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

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

Test 2 result:

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

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

(a) Base on test 1 result:

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

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

$$\Rightarrow P(D \mid \text{positive}) = \frac{0.24}{0.38} = 0.605$$

Similarly, base on test 2 result:

$$\Rightarrow P(\text{positive}) = 0.9 \times 0.3 + 0.2 \times 0.7 = 0.41$$

$$\Rightarrow P(D \mid \text{positive}) = \frac{0.9 \times 0.3}{0.41} = 0.658$$

(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:

$$P(D \mid \text{negative}) = \frac{0.2 \times 0.3}{0.2 \times 0.3 + 0.8 \times 0.7} = \frac{0.06}{0.62} = 0.096$$

Base on test result 2:

$$P(D \mid \text{negative}) = \frac{0.1 \times 0.3}{0.1 \times 0.3 + 0.7 \times 0.7} = \frac{0.03}{0.52} = 0.057$$

(c) Assume total Walmart can sell X cameras

Defective = 0.3X , non Defective = 0.7X.

The total cost of both "miss detection" & "false alarm":

Total cost = \$12 × "miss detection" units + \$6 × "false alarm" units

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

* Base on test 1 result:

$\frac{\text{Total cost}}{X} = 12 \times 0.096 \times 0.62 + 6 \times 0.2 \times 0.7 = 1.554$

$\Rightarrow \text{Total cost}_{(1)} = \$1.554X$

* Base on test 2 result:

$\frac{\text{Total cost}}{X} = 12 \times 0.057 \times 0.52 + 6 \times 0.3 \times 0.7 = 1.615$

$\Rightarrow \text{Total cost}_{(2)} = \$1.615X$

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:

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

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

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

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

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

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

Hence, coef = 10.42.

intercept = 223.82.

Problem 4) Please refer to hw1_150201133.py