Machine Learning 1: Homework week 1

Mai Xuan Bach
ID 11200489

Ngày 17 tháng 8 năm 2022

1 Problem 1

a. The marginal distributions $p(x)$ and $p(y)$ is computed below:

	0.16	0.17	0.11	0.22	0.34	
y_1	0.01	0.02	0.03	0.1	0.1	0.26
y_2	0.05	0.1	0.05	0.07	0.2	0.47
y_3	0.1	0.05	0.03	0.05	0.04	0.27
	x_1	x_2	x_3	x_4	x_5	

Hình 1: Marginal distribution

b. The conditional distributions $p(x|Y = y_1)$ and $p(x|Y = y_3)$:

$p(X = x_1|Y = y_1) = \frac{p(X=x_1,Y=y_1)}{p(Y=y_1)} = \frac{0.01}{0.26} = 0.038$

$p(X = x_2|Y = y_1) = \frac{p(X=x_2,Y=y_1)}{p(Y=y_1)} = \frac{0.02}{0.26} = 0.077$

$p(X = x_3|Y = y_1) = \frac{p(X=x_3,Y=y_1)}{p(Y=y_1)} = \frac{0.03}{0.26} = 0.115$

$p(X = x_4|Y = y_1) = \frac{p(X=x_4,Y=y_1)}{p(Y=y_1)} = \frac{0.1}{0.26} = 0.385$

$p(X = x_5|Y = y_1) = \frac{p(X=x_5,Y=y_1)}{p(Y=y_1)} = \frac{0.1}{0.26} = 0.385$

$p(X = x_1|Y = y_3) = \frac{p(X=x_1,Y=y_3)}{p(Y=y_3)} = \frac{0.1}{0.27} = 0.37$

$p(X = x_2|Y = y_3) = \frac{p(X=x_2,Y=y_3)}{p(Y=y_3)} = \frac{0.05}{0.27} = 0.185$

$p(X = x_3|Y = y_3) = \frac{p(X=x_3,Y=y_3)}{p(Y=y_3)} = \frac{0.03}{0.27} = 0.(1)$

$p(X = x_4|Y = y_3) = \frac{p(X=x_4,Y=y_3)}{p(Y=y_3)} = \frac{0.05}{0.27} = 0.185$

$p(X = x_5|Y = y_3) = \frac{p(X=x_5,Y=y_3)}{p(Y=y_3)} = \frac{0.04}{0.27} = 0.148$

2 Problem 2

$E_X[x] = \sum_x xp(X = x)$

$E_X[x|y] = \sum_x xp(X = x|Y = y) = \sum_x x\frac{p(X = x, Y = y)}{p(Y = y)} = \sum_x x\frac{p(Y = y|X = x).p(X = x)}{p(Y = y)}$

$E_Y[E_X[x|y]] = \sum_y E_X[x|y].p(Y = y) = \sum_y \sum_x x\frac{p(Y = y|X = x).p(X = x)}{p(Y = y)}.p(Y = y)$

$= \sum_x xp(X = x) \sum_y p(Y = y|X = x) = E_X[x].1 = E_X[x]$

So that, $E_X[x] = E_Y[E_X[x|y]]$.

3 Problem 3

We have:

X: event using product X, Y: event using product Y

$p(X) = 20.7\% = 0.207$

$p(Y) = 50\% = 0.5$

$p(X|Y) = 36.5\% = 0.365$

Answer:

a.

$p(X, Y) = p(X|Y).p(Y) = 0.365 * 0.5 = 0.1825$

b.

$p(Y|\bar{X}) = \frac{p(\bar{X}|Y).p(Y)}{p(\bar{X})} = \frac{(1-p(X|Y)).p(Y)}{(1-p(X))} = \frac{(1-0.365)*0.5}{(1-0.207)} = 0.4003783$

4 Problem 4

$V_X = E_X[(x - E_X[x])^2] = E_X[(x^2 - 2x.E_X[x] + (E_X[x])^2)] = E_X[x^2] - 2.E_X[x].E_X[x] + (E_X[x])^2$
$= E_X[x^2] - 2.(E_X[x])^2 + (E_X[x])^2 = E_X[x^2] - (E_X[x])^2$

So that, $V_X = E_X[x^2] - (E_X[x])^2$

5 Problem 5: Monty Hall Problem

Câu hỏi của bài toán: Giữ nguyên ô ban đầu hay đổi sang ô cửa còn lại chưa được lật mở?

Nói cách khác: Ta tính xem, xác suất chọn đúng khi giữ nguyên ô ban đầu và xác suất chọn đúng khi đổi sang ô cửa còn lại chưa được lật mở, trong hai cái đó, cái nào lớn hơn?

Gọi số thứ tự các cửa là 1,2 và 3. Kí hiệu Prize = 1 nghĩa là giải thưởng (Prize) nằm ở cửa số 1. Có tất cả 6 trường hợp liên quan đến bài toán, tương ứng với 6 xác suất ta cần tìm, với trường hợp ta lựa chọn ban đầu là cửa số 1:

1. P(Prize = 1 | Choice = 1, Monty = 2) = ?

2. P(Prize = 1 | Choice = 1, Monty = 3) = ?

3. P(Prize = 2 | Choice = 1, Monty = 2) = 0

4. P(Prize = 2 | Choice = 1, Monty = 3) = ?

5. P(Prize = 3 | Choice = 1, Monty = 2) = ?

6. P(Prize = 3 | Choice = 1, Monty = 3) = 0

Như vậy, còn 4 tình huống xác suất ta chưa biết. Song về bản chất, việc Monty mở cửa nào không quan trọng. Điều quan trọng là ta cần so sánh, trong cùng 1 trường hợp về lựa chọn mở cửa 1 trước, và Monty mở cùng 1 cửa, thì xác suất giải thưởng nằm ở cửa nào là cao hơn?

Ta xét hai tình huống đại diện sau (là tình huống 1 và 5 ở trên), chúng hoàn toàn tương tự cho cặp tình huống số 2 và 4 ở trên:

a. P(Prize = 1 | Choice = 1, Monty = 2) = ?

b. P(Prize = 3 | Choice = 1, Monty = 2) = ?

Từ tiếp theo, Pr viết tắt cho Prize, C cho Choice và M cho Monty.

Theo conditional rule: $P(A|B,C) = \frac{P(A.B|C)}{P(B|C)} = \frac{P(B|C,A).P(A|C)}{P(B|C)}$

Xét a.

$P(Pr = 1|C = 1, M = 2) = \frac{P(M=2|Pr=1,C=1).P(Pr=1|C=1)}{P(M=2|C=1)} = \frac{\frac{1}{2}*\frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}$

Xét b.

$P(Pr = 3|C = 1, M = 2) = \frac{P(M=2|Pr=3,C=1).P(Pr=3|C=1)}{P(M=2|C=1)} = \frac{1*\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$

Như vậy, xác suất giải thưởng (car) nằm ở cửa số 3 là cao hơn xác suất giải thưởng nằm ở cửa số 1, khi ta đã chọn 1 cửa và Monty chọn 1 cửa ($\frac{2}{3} > \frac{1}{3}$). Do đó, ta NÊN ĐỔI SANG Ô CÒN LẠI.