# ML Week 1

Ngo Doan Kien

August 2022

## 1 Problem 1

### 1.1 (a) The marginal distributions p(x) and p(y)

$$P(x) = \begin{cases} 0.16 & X = x_1 \\ 0.17 & X = x_2 \\ 0.11 & X = x_3 \\ 0.22 & X = x_4 \\ 0.37 & X = x_5 \end{cases}$$

$$P(y) = \begin{cases} 0.26 & Y = y_1 \\ 0.47 & Y = y_2 \\ 0.27 & Y = y_3 \end{cases}$$

### 1.2 (b) The conditional distributions $p(x|Y = y_1)$ and $p(x|Y = y_3)$

$p(x|Y = y_1) = p(x, Y = y_1)\ /\ p(Y = y_1) = p(x, Y = y_1)\ /\ 0.26$

Then,
$p(X = x_1|Y = y_1) = 0.038462$
$p(X = x_2|Y = y_1) = 0.076923$
$p(X = x_3|Y = y_1) = 0.115385$
$p(X = x_4|Y = y_1) = 0.384615$
$p(X = x_5|Y = y_1) = 0.384615$

$p(x|Y = y_3) = p(x, Y = y_3)\ /\ p(Y = y_3) = p(x, Y = y_3)\ /\ 0.27$

Then,
$p(X = x_1|Y = y_3) = 0.37037$
$p(X = x_2|Y = y_3) = 0.185185$
$p(X = x_3|Y = y_3) = 0.111111$

$p(X = x_4|Y = y_3) = 0.185185$
$p(X = x_5|Y = y_3) = 0.148148$

# 2 Problem 2

Prove: $E_x[x] = E_y[E_x[x|Y = y]]$

We have:
$E_y[E_x[x|Y = y]] = E_y[\sum_i x_i * P(X = x_i|Y = y_j)]$

$= E_y[\sum_i x_i * \frac{P(X=x_i,Y=y_j)}{P(Y=y_j)}]$ (Baye' rule)

$= \sum_j (E_y[\sum_i x_i * \frac{P(X=x_i,Y=y_j)}{P(Y=y_j)}]) * P[Y = y_j]$

$= \sum_j \sum_i x_i * P(X = x_i, Y = y_j)$ (reduction $P[Y = y_j]$)

$= \sum_i [x_i \sum_j P(X = x_i, Y = y_j)]$

$= \sum_i x_i P(X = x_i)$ (sum rule)

$= E(x)$ => proved

# 3 Problem 3

## 3.1 (a) Xác suất người ấy dùng cả X và Y

From what the problem gives, we have:
$P(X) = 0.207$
$P(Y) = 0.5$
$P(X|Y) = 0.365$

Xác suất dùng cả X và Y là P(X, Y) =>
$P(X, Y) = P(X|Y) * P(Y) = 0.365 * 0.5 = 0.1825$ (Product rule)

## 3.2 (b) Xác suất người ấy dùng Y (biết rằng người ấy không dùng X)

Gọi biến cố không dùng X là: $\overline{X}$
=> Xác suất dùng Y, biết không dùng X là: $P(Y|\overline{X})$

we have:
$P(Y|\overline{X}) = \frac{P(Y,\overline{X})}{P(\overline{X})}$ (Baye's rule) (0)

$P(Y, \overline{X}) = P(\overline{X}|Y) * P(Y)$ (Product rule) (1)

$P(\overline{X}|Y) = \frac{P(\overline{X},Y)}{P(Y)}$ (Baye's rule)
$= \frac{P(Y)-P(X,Y)}{P(Y)} = \frac{0.5-0.1825}{0.5} = 0.635$ (tổng xác suất của 2 biến cố đối) (2)

Từ (1) và (2) ta có:
$P(Y, \overline{X}) = 0.635 * P(Y) = 0.635 * 0.5 = 0.3175$

=> Xác suất cần tìm là:
$(0) = P(Y|\overline{X}) = \frac{P(Y,\overline{X})}{P(\overline{X})} = \frac{0.3175}{1-P(X)} = \frac{0.3175}{0.793} = 0.4004$

# 4 Problem 4

Prove: $V_x(x) = E_x(x^2) - [E_x(x)]^2$

We have:
$V_x(x) = \sum_i [(x-\mu)^2 * f(x)]$

and also:
$E_x(x) = \sum x * f(x)$
$E_x(x^2) = \sum x^2 * f(x)$

=> $V_x(x) = E_x(x-\mu)^2 = E_x[x^2 - 2\mu * x + (\mu)^2]$
$= E_x(x^2) - 2\mu * E_x(x) + (\mu)^2$

$\mu$ is defined as expected value, so $\mu = E_x(x)$

=> $V_x(x) = E_x(x^2) - 2 * E_x(x) * E_x(x) + [E_x(x)]^2$
$= E_x(x^2) - 2[E_x(x)]^2 + [E_x(x)]^2$
$= E_x(x^2) - [E_x(x)]^2$ => Proved

# 5 Problem 5: Monty Hall

Giả sử ta chọn ô cửa 1
Gọi A là biến cố chiếc xe nằm trong ô cửa 1
B là biến cố Monty mở ô cửa 2 có dê (sau khi ta chọn ô 1)
Khi đó P(A) = 1/3
P(B|A) = 1/2 (vì giả sử xe ở ô 1 rồi vậy nên Monty chỉ có 2 trường hợp là mở cửa 2 hoặc cửa 3)
P(A|B) = 1/3 (vì nếu ta ko đổi ô thì việc Monty mở ô thứ 2 không ảnh hưởng đến xác suất => vẫn giữ nguyên giống P(A))

Gọi C là biến cố xe nằm ở ô cửa số 3 => A và C là 2 biến cố xung khắc (do xe chỉ ở cửa 1 hoặc cửa 3) => $P(C) = 1 - P(A) = 2/3$

KẾT LUẬN: Đổi sang ô cửa còn sẽ có xác suất trúng xe lớn hơn là giữ nguyên