# ADDIS ABABA UNIVERSITY
# ADDIS ABABA INSTITUTE OF TECHNOLOGY

## Introduction to AI

### Assignment 3: Probability Questions


Group members

1. Fuad Mohammed - UGR/6052/14
2. Salman Ali - UGR/7808/14
3. Abdulmunim Jundurahman -UGR/8625/14
4. Fathiya Safi - UGR/7007/13
5. Obsu Kebede - UGR /7425/14


Submitted to: Mr. Amanuel Negash

Submission date: 29 June 2024

1. Suppose that $P(A) = 0.4$, $P(B) = 0.3$ and $P((A \cup B)^c) = 0.42$. Are A and B independent?

**Solution:**

Given: $P(A) = 0.4$

$P(B) = 0.3$

$P((A \cup B)^c) = 0.42$

$P((A \cup B)^c) = 1 - P(A \cup B)$

$0.42 = 1 - P(A \cup B)$

$P(A \cup B) = 1 - 0.42 = 0.58$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$0.58 = 0.4 + 0.3 - P(A \cap B)$

$P(A \cap B) = 0.12$

If A and B are independent, $P(A \cap B) = P(A) * P(B)$

$P(A \cap B) = 0.12, \quad P(A) * P(B) = 0.4 * 0.3 = 0.12$

**Therefore, A and B are independent.**

2. Two dice are rolled. A = ‘sum of two dice equals 3’ B = ‘sum of two dice equals 7’ C = ‘at least one of the dice shows a 1’

**Solution:**

A = ‘sum of two dice equals 3’

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 1, 1 | 1, 2 | 1, 3 | 1, 4 | 1, 5 | 1, 6 |
| 2 | 2, 1 | 2, 2 | 2, 3 | 2, 4 | 2, 5 | 2, 6 |
| 3 | 3, 1 | 3, 2 | 3, 3 | 3, 4 | 3, 5 | 3, 6 |
| 4 | 4, 1 | 4, 2 | 4, 3 | 4, 4 | 4, 5 | 4, 6 |
| 5 | 5, 1 | 5, 2 | 5, 3 | 5, 4 | 5, 5 | 5, 6 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 6 | 6, 1 | 6, 2 | 6, 3 | 6, 4 | 6, 5 | 6, 6 |

P(A) = 2 / 36

B = 'sum of two dice equals 7'

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 1, 1 | 1, 2 | 1, 3 | 1, 4 | 1, 5 | 1, 6 |
| 2 | 2, 1 | 2, 2 | 2, 3 | 2, 4 | 2, 5 | 2, 6 |
| 3 | 3, 1 | 3, 2 | 3, 3 | 3, 4 | 3, 5 | 3, 6 |
| 4 | 4, 1 | 4, 2 | 4, 3 | 4, 4 | 4, 5 | 4, 6 |
| 5 | 5, 1 | 5, 2 | 5, 3 | 5, 4 | 5, 5 | 5, 6 |
| 6 | 6, 1 | 6, 2 | 6, 3 | 6, 4 | 6, 5 | 6, 6 |

P(B) = 6 / 36

C = 'at least one of the dice shows a 1'

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 1, 1 | 1, 2 | 1, 3 | 1, 4 | 1, 5 | 1, 6 |
| 2 | 2, 1 | 2, 2 | 2, 3 | 2, 4 | 2, 5 | 2, 6 |
| 3 | 3, 1 | 3, 2 | 3, 3 | 3, 4 | 3, 5 | 3, 6 |
| 4 | 4, 1 | 4, 2 | 4, 3 | 4, 4 | 4, 5 | 4, 6 |
| 5 | 5, 1 | 5, 2 | 5, 3 | 5, 4 | 5, 5 | 5, 6 |
| 6 | 6, 1 | 6, 2 | 6, 3 | 6, 4 | 6, 5 | 6, 6 |

P(C) = 11 / 36

a) What is P(A|C)?

$$
\begin{aligned}
P(A|C) &= P(A \cap C) / P(C) \\
&= (2 / 36) / (11 / 36) \\
&= \mathbf{2 / 11}
\end{aligned}
$$

b) What is P(B|C)?

$$\begin{aligned} P(B|C) &= P(B \cap C) / P(C) \\ &= (2 / 36) / (11 / 36) \\ &= \mathbf{2 / 11} \end{aligned}$$

c) Are A and C independent? What about B and C?

$P(A \cap C) = 2 / 36$ $\qquad$ $P(A) * P(C) = (2 / 36) * (11 / 36) = 22 / 1296$

$$P(A \cap C) \neq P(A) * P(C)$$

**Therefore A and C are not independent.**

$P(B \cap C) = 2 / 36$ $\qquad$ $P(B) * P(C) = (6 / 36) * (11 / 36) = 11 / 216$

$$P(B \cap C) \neq P(B) * P(C)$$

**Therefore B and C are not independent.**

3. Let C and D be two events with $P(C) = 0.25$, $P(D) = 0.45$, and $P(C \cap D) = 0.1$. What is $P(C^c \cap D)$?

**Solution:**

Given: $P(C) = 0.25$

$P(D) = 0.45$

$P(C \cap D) = 0.1$

$$\begin{aligned} P(C^c \cap D) &= P(D) - P(C \cap D) \\ &= 0.45 - 0.1 \\ &= \mathbf{0.35} \end{aligned}$$

4. There are 3 arrangements of the word DAD, namely DAD, ADD, and DDA. How many arrangements are there of the word PROBABILITY?

**Solution:**

Total arrangements: (number of letters)! / (number of repeated letters)!
=11! / (2! * 2!)
**=9,979,200**

5. Let A and B be two events. Suppose the probability that neither A or B occurs is 2/3. What is the probability that one or both occur?

Probability of at least one occurring = 1 – probability of neither occurring
= 1 – 2 / 3
**= 1 / 3**

6. Let X denote the number of times a photocopy machine will malfunction: 0, 1, 2, or 3 times, on any given month. Let Y denote the number of times a technician is called on an emergency call. The joint p.m.f. p( x, y ) is presented in the table below:

| | $x$ | | | | |
|---|---|---|---|---|---|
| $y$ | 0 | 1 | 2 | 3 | $p_Y(y)$ |
| 0 | 0.15 | 0.30 | 0.05 | 0 | 0.50 |
| 1 | 0.05 | 0.15 | 0.05 | 0.05 | 0.30 |
| 2 | 0 | 0.05 | 0.10 | 0.05 | 0.20 |
| $p_X(x)$ | 0.20 | 0.50 | 0.20 | 0.10 | 1.00 |

a) Find the probability P( Y > X ).

**Solution:**

P(Y > X) = P(x=0, y=1) + P(x=1, y=2)
= 0.05 + 0.05 = **0.1**

b) Find p X( x ), the marginal p.m.f. of X.

**Solution:**

By summing the joint probabilities along the rows or columns of the table, we can obtain the marginal probability mass function (PMF) of each variable separately.

$$
P_x = \begin{cases} 0.2, & \text{If } x = 0 \\ 0.5, & \text{If } x = 1 \\ 0.2, & \text{If } x = 2 \\ 0.1, & \text{If } x = 3 \end{cases}
$$

c) Find p Y( y ), the marginal p.m.f. of Y.

**Solution:**

$$
P_y = \begin{cases} 0.5, & \text{If } y = 0 \\ 0.3, & \text{If } y = 1 \\ 0.2, & \text{If } y = 2 \end{cases}
$$

d) Are X and Y independent?

**Solution:**

We can say that X and Y are independent if and only if $P(X \cap Y) = P(X) \times P(Y)$ for every x and y. Let's check for x = 0 and y = 0:

$P(x=0, y=0)$ **= 0.15** $\quad\quad$ $P(X=0) * P(y=0) = 0.2 * 0.5$ **= 0.1**

Since $P(x=0, y=0) \neq P(x=0) * P(y=0)$, x and y are **not independent**

7. The following are data points with their labels:
   (1, 2, 3, 4), 1
   (5, 6, 7, 8), 0
   (9, 10, 11, 12), 1

   The following are the randomly set weights:
   w1 = 0.1

w2 = 0.2
w3 = -0.1
w4 = 0.0

Task: make three learning updates with a learning rate of 0.1 using the data points. The updates should be based on both the Perceptron and the logistic regression. Compare the two results.

**Solution:**

Using Perceptron:

| | | | | | Label |
|---|---|---|---|---|---|
| Row-1 | 1 | 2 | 3 | 4 | 1 |
| Row-2 | 5 | 6 | 7 | 8 | 0 |
| Row-3 | 9 | 10 | 11 | 12 | 1 |

| | | | | |
|---|---|---|---|---|
| weights | 0.1 | 0.2 | -0.1 | 0.0 |

| | |
|---|---|
| predicted | = weights.features |
| | = 1*0.1 + 2*0.2 + 3*(-0.1) + 4*0.0 |
| | = 0.1 + 0.4 - 0.3 + 0 |
| | = 0.3 |
| | ≈ 1 |

| | |
|---|---|
| Updated weight | = = weights – features * learning_rate * (predicted – label) |
| | = (0.1, 0.2, -0.1, 0.0) – ((1 , 2, 3, 4) * (0.1 * (1 – 1))) |
| | = (0.1, 0.2, -0.1, 0.0) – (0, 0, 0, 0) |
| | = (0.1, 0.2, -0.1, 0.0) |

| | | | | | Label |
|---|---|---|---|---|---|

| | | | | | |
|---|---|---|---|---|---|
| Row-1 | 1 | 2 | 3 | 4 | 1 |
| Row-2 | 5 | 6 | 7 | 8 | 0 |
| Row-3 | 9 | 10 | 11 | 12 | 1 |

| | | | | |
|---|---|---|---|---|
| weights | 0.1 | 0.2 | -0.1 | 0.0 |

| | |
|---|---|
| predicted | = weights.features |
| | = 5*0.1 + 6*0.2 + 7*(-0.1) + 8*0.0 |
| | = 0.5 + 1.2 - 0.7 + 0 |
| | = 1 |

| | |
|---|---|
| Updated weight | = = weights – features * learning_rate * (predicted – label) |
| | = (0.1, 0.2, -0.1, 0.0) – ((5, 6, 7, 8) * (0.1 * (1 – 0))) |
| | = (0.1, 0.2, -0.1, 0.0) – (0.5, 0.6, 0.7, 0.8) |
| | = (-0.4, -0.4, -0.8, -0.8) |

| | | | | | Label |
|---|---|---|---|---|---|
| Row-1 | 1 | 2 | 3 | 4 | 1 |
| Row-2 | 5 | 6 | 7 | 8 | 0 |
| Row-3 | 9 | 10 | 11 | 12 | 1 |

| | | | | |
|---|---|---|---|---|
| weights | -0.4 | -0.4 | -0.8 | -0.8 |

| | |
|---|---|
| predicted | = weights.features |
| | = 9*(-0.4) + 10*(-0.4) + 11*(-0.8) + 12*(-0.8) |
| | = -3.6 – 4 – 8.8 – 9.6 |
| | = -26 |
| | ≈ 0 |

| | |
|---|---|
| Updated weight | = = weights – features * learning_rate * (predicted – label) |
| | = (-0.4, -0.4, -0.8, -0.8) – ((9, 10, 11, 12) * (0.1 * (0 – 1))) |
| | = (-0.4, -0.4, -0.8, -0.8) – (-0.9, -1, -1.1, -1.2) |
| | = (0.5, 0.6, 0.3, 0.4) |

Using Logistic Regression:

| | | | | | Label |
|---|---|---|---|---|---|
| Row-1 | 1 | 2 | 3 | 4 | 1 |
| Row-2 | 5 | 6 | 7 | 8 | 0 |
| Row-3 | 9 | 10 | 11 | 12 | 1 |
| weights | 0.1 | 0.2 | -0.1 | 0.0 | |

| | |
|---|---|
| predicted | = weights.features |
| | = 1*0.1 + 2*0.2 + 3*(-0.1) + 4*0.0 |
| | = 0.1 + 0.4 - 0.3 + 0 |
| | = 0.3 |
| | ≈ ( 1 / (1 + $e^{-0.3}$) ) = 0.574 |
| | ≈ 1 |

| | |
|---|---|
| Updated weight | = = weights – features * learning_rate * (predicted – label) |
| | = (0.1, 0.2, -0.1, 0.0) – ((1 , 2, 3, 4) * (0.1 * (1 – 1))) |
| | = (0.1, 0.2, -0.1, 0.0) – (0, 0, 0, 0) |
| | = (0.1, 0.2, -0.1, 0.0) |

| | | | | | Label |
|---|---|---|---|---|---|
| Row-1 | 1 | 2 | 3 | 4 | 1 |
| Row-2 | 5 | 6 | 7 | 8 | 0 |
| Row-3 | 9 | 10 | 11 | 12 | 1 |
| weights | 0.1 | 0.2 | -0.1 | 0.0 | |

| | |
|---|---|
| predicted | = weights.features |
| | = 5*0.1 + 6*0.2 + 7*(-0.1) + 8*0.0 |

| Updated weight | = = weights – features * learning_rate * (predicted – label) |
|---|---|
| | = (0.1, 0.2, -0.1, 0.0) – ((5, 6, 7, 8) * (0.1 * (1 – 0))) |
| | = (0.1, 0.2, -0.1, 0.0) – (0.5, 0.6, 0.7, 0.8) |
| | = (-0.4, -0.4, -0.8, -0.8) |

| | | | | | Label |
|---|---|---|---|---|---|
| Row-1 | 1 | 2 | 3 | 4 | 1 |
| Row-2 | 5 | 6 | 7 | 8 | 0 |
| Row-3 | 9 | 10 | 11 | 12 | 1 |
| weights | -0.4 | -0.4 | -0.8 | -0.8 | |

| predicted | = weights.features |
|---|---|
| | = 9*(-0.4) + 10*(-0.4) + 11*(-0.8) + 12*(-0.8) |
| | = -3.6 – 4 – 8.8 – 9.6 |
| | = -26 |
| | ≈ ( 1 / (1 + $e^{-1}$) ) = 5.1 x $10^{-12}$ |
| | ≈0 |

| Updated weight | = = weights – features * learning_rate * (predicted – label) |
|---|---|
| | = (-0.4, -0.4, -0.8, -0.8) – ((9, 10, 11, 12) * (0.1 * (0 – 1))) |
| | = (-0.4, -0.4, -0.8, -0.8) – (-0.9, -1, -1.1, -1.2) |

| | |
|---|---|
| | = (0.5, 0.6, 0.3, 0.4) |