

ADDIS ABABA UNIVERSITY

ADDIS ABABA INSTITUTE OF TECHNOLOGY

Introduction to AI

Assignment 3: Part 2 - Probability

**Group Members**

1. Yohannes Solomon UGR/3323/13 Section 1

2. Dawit Belay UGR/8622/14 Section 1

3. Sifan Fita UGR/8856/14 Section 2

4. Ahmed Muhammed UGR/8920/14 Section 1

Assignment III (AI)  
  
  
**part II - Probability Questions**

1) Suppose that P(A) = 0.4, P(B) = 0.3 and P((A ∪ B)C) = 0.42.  
  
 Are A and B independent?  
  
**SOLUTION:**  
  
Given:

* P(A)=0.4P
* P(B)=0.3
* P((A∪B) C)=0.42  
    
  first find P(A∪B):  
  P(A∪B)C=1-P(A∪B)  
   0.42= 1- P(A∪B)  
  P(A∪B)=1−0.42=**0.58**  
    
  by using this union of A and B now we need to get the intersection of A and B as follows,  
  To get the insection we use this formula of union :  
  P(A∪B)=P(A)+P(B)−P(A∩B) from this => P(A∩B)= P(A)+P(B)− P(A∪B)  
     
   P(A)=0.4  
   P(B)=0.3  
   P(A∪B)= 0.58

P(A∩B)= P(A)+P(B)− P(A∪B)  
 =(0.4 + 0.3 ) -0.58  
 =**0.12  
To check the independence**

P(A)×P(B)=0.4×0.3=**0.12 =>**This is equal to the value of P(A∩B) ,  
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’

1. What is P(A|C)?
2. What is P(B|C)?
3. Are A and C independent? What about B and C?

**SOLUTION:**

**Event A**: There are only two outcomes that result in a sum of 3:

**=>**(1, 2) and (2, 1).

There are 36 possible outcomes when two dice are rolled.

P(A)=2/36=1/18

**Event B**: The outcomes that result in a sum of 7 are :

**=>**(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), and (6, 1).

P(B)=6/36=1/6 **Event C**: The outcomes where at least one die shows a 1 are:

=> (1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2, 1), (3, 1), (4, 1), (5, 1), (6, 1).

So there are 11 outcomes.  
 **P(C)=11/36**  
Then we can find P(A∩C) as follows :  
The outcomes in A∩C are: (1, 2) and (2, 1).  
Now **P(A∩C)=2/36=1/18**

* a)Then we can find the P(A∣C) as follows:

P(A∣C) = P(A∩C)​/ P(C)  
 =(1/18)/(11/36)  
 =(1/36)\*(36/11)  
 **= 2/11**

* b)Now let’s find P(B∣C):

P(B∣C) = P(B∩C)​ / P(C)

=(1/18)/(11/36)

=(1/18)\*(36 / 11)

**=2/11**

* **checking independence**c) Are A and C independent? What about B and C?

**1)**To check independence of A and Cif **P(A∩C)=P(A)P(C).  
  
 P(A∩C)=**1/18 and **P(A)\*P(C)**=(1/18)\*(11/36)  
 =11/648  
 =1/59  
since P(A∩C) **not equal** P(A)P(C) ,the A and C are **not independent**

**2**) To check the independence of B and C

we do have check equality P(B∩C)=P(B)P(C):  
  
P(B∩C)=1/18  
P(B)\*P(C)=(1/6) \* (11/36)=11/216  
 since 1/18 not equal to 11/216,they are not independent

3) Let C and D be two events with P(C) = 0.25, P(D) = 0.45, and P(C ∩ D) = 0.1.

What is P(Cc ∩ D)?

**SOLUTION:**

***G****iven:* P(C)=0.25 P(D)=0.45 P(C∩D)=0.1  
  
we can calculate P(CC∩D) as follows :  
P()=1−P(C)=1-0.25=**0.75**P( ∩D)= P(D)−P(C∩D)=0.45−0.1=**0.35**

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

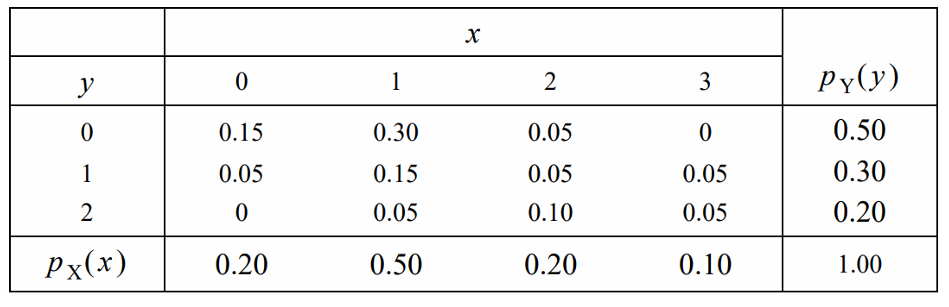
**SOLUTION:**

word : **PROBABILITY  
total letters =11  
B** occurs twice, **I** occurs twice

Number of arrangements ===**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?  **SOLUTION:**Probability that neither A nor B occurs is 2/3 this means P(   
Probability that one of A or occur means P(A∪B)  
therefore we needed to find P(A∪B) as following :

P(A∪B)=1- P(  
 =1-2/3  
 =**1/3=0.3333..**

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:  


● Find the probability P( Y > X ).

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

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

● Are X and Y independent?

**SOLUTION:**

1. Calculate P(Y>X),means sum the probabilities where Y > X from the table above :  
   P(Y>X)= p(0,1)+p(0,2)+p(0,3)+p(1,2)+p(1,3)+p(2,3)

=0.1+0.05+0.05+0.1+0.05+0.05

=**0.4**

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

pX​(x)=

therefore we calculate for each x:

pX​(0)=0.1+0.1+0.1+0.1=**0.4**

pX​(1)=0.05+0.05+0.05+0.05=**0.2**

pX​(2)=0.05+0.05+0.05+0.05=**0.2**

pX​(3)=0.05+0.05+0.05+0.05=**0.2**

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

pY​(y)=   
Calculate for each y:

pY​(0)=0.1+0.05+0.05+0.05=**0.25**

pY​(1)=0.1+0.05+0.05+0.05=**0.25**

pY​(2)=0.1+0.05+0.05+0.05**=0.25**

pY​(3)=0.1+0.05+0.05+0.05=**0.25**

**d)** Are X and Y independent?

For X and Y to be independent, p(x,y) should be equal to pX(x)⋅pY(y) for all x and y.

=>Let’s check for all pairs start from p(0,0) to p(3,3):

* For x=0

P(0,0)=**0.1**

pX​(0)⋅pY​(0)=0.4⋅0.25=**0.1**

p(0,1)=**0.1**

pX​(0)⋅pY​(1)=0.4⋅0.25=**0.1**

p(0,2)=**0.1**

pX​(0)⋅pY​(2)=0.4⋅0.25=**0.1**

p(0,3)=**0.1**

pX​(0)⋅pY​(3)=0.4⋅0.25=**0.1**

* For x=1

p(1,0)=**0.05**

pX​(1)⋅pY​(0)=0.2⋅0.25=**0.05**

p(1,1)=**0.05**

pX​(1)⋅pY​(1)=0.2⋅0.25=**0.05**

p(1,2)=**0.05**

pX​(1)⋅pY​(2)=0.2⋅0.25=**0.05**

p(1,3)=**0.05**

pX​(1)⋅pY​(3)=0.2⋅0.25=**0.05**

* For x=2

p(2,0)=**0.05**

pX​(2)⋅pY​(0)=0.2⋅0.25=**0.05**

p(2,1)=**0.05**

pX​(2)⋅pY​(1)=0.2⋅0.25=**0.05**

p(2,2)=**0.05**

pX​(2)⋅pY​(2)=0.2⋅0.25=**0.05**

p(2,3)=**0.05**

pX​(2)⋅pY​(3)=0.2⋅0.25=**0.05**

* For x=3

p(3,0)=**0.05**

pX​(3)⋅pY​(0)=0.2⋅0.25=**0.05**

p(3,1)=**0.05**

pX​(3)⋅pY​(1)=0.2⋅0.25=**0.05**

p(3,2)=**0.05**

pX​(3)⋅pY​(2)=0.2⋅0.25=**0.05**

p(3,3)=**0.05**

pX​(3)⋅pY​(3)=0.2⋅0.25=**0.05**

**since** all the are pairs satisfy the condition p(x,y)=pX​(x)⋅pY​(y), X and Y are indeed they are **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:**

**Perceptron Algorithm**

* Initial Weights:
* w1​=0.1, w2​=0.2, w3​=−0.1, w4​=0.0
* Learning rate η=0.1

Data Points:

(1, 2, 3, 4), label = 1 = y1

(5, 6, 7, 8), label = 0 = y2

(9, 10, 11, 12), label = 1 = y3

Perceptron Update Rule

* The sign function sign(z) returns +1 if z>0 and −1 otherwise
* y^​=sign(w⋅x)
* wi​=wi​+η⋅(y−y^​)⋅xi​

**Update 1:**

* Calculate the dot product:

w⋅x1​=0.1⋅1+0.2⋅2+(−0.1) ⋅3+0.0⋅4=0.1+0.4−0.3=0.2

* Determine the predicted label:

Y^1  = sign(0.2) =1

* Since y^1=y1​, no update is needed.

**Update 2:**

* Calculate the dot product:

w⋅x2​=0.1⋅5+0.2⋅6+(−0.1)⋅7+0.0⋅8=0.5+1.2−0.7=1.0

* Determine the predicted label:

Y^1  = sign(1.0) =1

* Since y^2 ≠y2​, update the weights:

w1​=0.1+0.1⋅(0−1)⋅5=0.1−0.5=−0.4

w2​=0.2+0.1⋅(0−1)⋅6=0.2−0.6=−0.4

w3​=−0.1+0.1⋅(0−1)⋅7=−0.1−0.7=−0.8

w4​=0.0+0.1⋅(0−1)⋅8=0.0−0.8=−0.8

**Update 3:**

* Calculate the dot product:

w⋅x3​= (−0.4)⋅9+(−0.4)⋅10+(−0.8)⋅11+(−0.8)⋅12=−3.6−4−8.8−9.6=−26

* Determine the predicted label:

Y^1  = sign(-26) =1

* Since y^3 ≠y3​, update the weights:

w1​= −0.4+0.1⋅(1−(−1))⋅9=−0.4+0.1⋅2⋅9=−0.4+1.8=1.4

w2​= −0.4+0.1⋅(1−(−1))⋅10=−0.4+0.1⋅2⋅10=−0.4+2.0=1.6

w3​= −0.8+0.1⋅(1−(−1))⋅11=−0.8+0.1⋅2⋅11=−0.8+2.2=1.4

w4​= −0.8+0.1⋅(1−(−1))⋅12=−0.8+0.1⋅2⋅12=−0.8+2.4=1.6

**Final Perceptron Weights:**

* w1​=1.4, w2​=1.6, w3​=1.4, w4​=1.6

**Logistic Regression**

* Initial Weights:
* w1​=0.1, w2​=0.2, w3​=−0.1, w4​=0.0

Logistic Regression Update Rule

* y^​=σ(w⋅x)
* wi​=wi​+η⋅(y−y^​)⋅xi​

**Update 1:**

* Calculate the dot product and the sigmoid:

w⋅x1​=0.1⋅1+0.2⋅2+(−0.1)⋅3+0.0⋅4=0.1+0.4−0.3=0.2

y^1 = ​ ≈ 0.55

* Update the weights:

w1​=0.1+0.1⋅(1−0.55)⋅1=0.1+0.045=0.145

w2​=0.2+0.1⋅(1−0.55)⋅2=0.2+0.09=0.29

w3​=−0.1+0.1⋅(1−0.55)⋅3=−0.1+0.135=0.035

w4​=0.0+0.1⋅(1−0.55)⋅4=0.0+0.18=0.18

**Update 2:**

* Calculate the dot product and the sigmoid:

w⋅x2= 0.145⋅5+0.29⋅6+0.035⋅7+0.18⋅8=0.725+1.74+0.245+1.44=4.15

y^2 = ​ ≈ 0.984

* Update the weights:

w1​= 0.145+0.1⋅(0−0.984)⋅5=0.145−0.492=−0.347

w2​= 0.29+0.1⋅(0−0.984)⋅6=0.29−0.5904=−0.3004

w3​= 0.035+0.1⋅(0−0.984)⋅7=0.035−0.6888=−0.6538

w4​= 0.18+0.1⋅(0−0.984)⋅8=0.18−0.7872=−0.6072

**Update 3:**

* Calculate the dot product and the sigmoid:

w⋅x3=−0.347⋅9+(−0.3004)⋅10+(−0.6538)⋅11+(−0.6072)⋅12=−3.123−3.004−7.1918−7.2864=−20.6052

y^3 = ​≈ 0

* Update the weights:

w1​= −0.347+0.1⋅(1−0)⋅9=−0.347+0.9=0.553

w2​= −0.3004+0.1⋅(1−0)⋅10=−0.3004+1.0=0.6996

w3​= −0.6538+0.1⋅(1−0)⋅11=−0.6538+1.1=0.4462

w4​= −0.6072+0.1⋅(1−0)⋅12=−0.6072+1.2=0.5928

**Final Logistic Regression Weights:**

* w1​=0.553, w2​=0.6996, w3​=0.4462, w4​=0.5928

**Comparison**

**Perceptron Weights:**

* w1​=1.4, w2​=1.6, w3​=1.4, w4​=1.6

**Logistic Regression Weights:**

* w1​=0.553, w2​=0.6996, w3​=0.4462, w4​=0.5928

The Perceptron algorithm updates weights more drastically compared to Logistic Regression because it applies the update rule based on whether the prediction is right or wrong. In contrast, Logistic Regression updates weights based on the gradient of the loss function, resulting in more gradual and smaller updates.