# Machine Learning from Data: Homework 3 - Probabilities

227367455 and 323081950

May 19, 2021

# Question 1

Given a random sample  $\{x_1, x_2, ..., x_n\}$ , derive the maximum likelihood estimator  $\hat{p}$  of the Binomial distribution.

$$B(x,p) = \binom{n}{x} p^x (1-p)^{n-x}$$

We first want to calculate the likelihood:

$$L = P(x_1, ...x_n \mid p) = \prod_{i=1}^n P(x_i \mid p)$$

$$= \prod_{i=1}^n \binom{n}{x_i} p^{x_i} (1-p)^{1-x_i}$$

$$= \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i} \prod_{i=1}^n \binom{n}{x_i}$$

$$= p^{\sum_{i=1}^n x_i} (1-p)^{n-\sum_{i=1}^n x_i} \prod_{i=1}^n \binom{n}{x_i}$$

From the likelihood we calculate the log-likelihood:

$$ln(L) = ln(p^{\sum_{i=1}^{n} x_i} (1-p)^{n-\sum_{i=1}^{n} x_i} \prod_{i=1}^{n} \binom{n}{x_i})$$

$$= ln(p^{\sum_{i=1}^{n} x_i}) + ln((1-p)^{n-\sum_{i=1}^{n} x_i}) + ln(\prod_{i=1}^{n} \binom{n}{x_i})$$

$$= ln(p) \sum_{i=1}^{n} x_i + ln(1-p)(n-\sum_{i=1}^{n} x_i) + \sum_{i=1}^{n} ln(\binom{n}{x_i})$$

We will take the derivative in respect to p our given value:

$$\frac{\partial [ln(L)]}{\partial p} = \frac{\partial [ln(p)\sum_{i=1}^{n} x_i]}{\partial p} + \frac{\partial [ln(1-p)(n-\sum_{i=1}^{n} x_i)]}{\partial p}$$
$$= \frac{\sum_{i=1}^{n} x_i}{p} - \frac{(n-\sum_{i=1}^{n} x_i)}{1-p}$$

To find the a maximum we set the derivative to 0 obtaining:

$$\frac{\sum_{i=1}^{n} x_i}{\hat{p}} - \frac{n - \sum_{i=1}^{n} x_i}{1 - \hat{p}} = 0$$

$$(1 - \hat{p}) \sum_{i=1}^{n} x_i - \hat{p}(n - \sum_{i=1}^{n} x_i) = 0$$

$$\sum_{i=1}^{n} x_i - \hat{p} \sum_{i=1}^{n} x_i - \hat{p}n + \hat{p} \sum_{i=1}^{n} x_i = 0$$

$$\sum_{i=1}^{n} x_i - \hat{p}n + = 0$$

$$\hat{p}n = \sum_{i=1}^{n} x_i$$

Thus we obtain:

$$\hat{p} = \frac{\sum_{i=1}^{n} x_i}{n}$$

#### Question 2

A student wants to know her chances to pass and fail an exam if she studies and if she doesn't study. From last year's results, she sees that P(Pass) = 60%. She also found out that  $P(Studied \mid Pass) = 95\%$ ,  $P(Studied \mid Failed) = 60\%$ . You can assume that every student either studied or didn't study, and either passed or failed.

$$P(Failed) = 1 - P(Pass) = 1 - 0.6 = 0.4$$

$$P(Studied) = P(Studied \mid Pass)P(Pass) + P(Studied \mid Failed)P(Failed)$$

$$= 0.95 \cdot 0.6 + 0.6 \cdot 0.4 = 0.81$$

$$P(\overline{Studied}) = 1 - P(Studied) = 1 - 0.81 = 0.19$$

$$P(\overline{Studied} \mid Pass) = 1 - P(Studied \mid Pass) = 1 - 0.95 = 0.05,$$

a.

What is her probability of passing the exam if she studies?

$$P(Pass|Studied) = \frac{P(Studied \mid Pass)P(Pass)}{P(Studied)} = \frac{0.95 \cdot 0.6}{0.81} = 0.7037$$

b.

What is her probability of passing if she doesn't study?

$$P(Pass \mid \overline{Studied}) = \frac{P(\overline{Studied} \mid Pass)P(Pass)}{P(\overline{Studied})} = \frac{0.05 \cdot 0.6}{0.19} = 0.1578$$

## Question 3

Find 3 random variables X, Y, C such that:

- a)  $X \perp Y \mid C$
- b) X and Y are not independent
- c) X, Y are integers such that  $3 \le X, Y \le 9$  and C is binary.
- d) The following conditions hold:

i. 
$$P(1 \le X, \le 5) = 0.4$$

ii. 
$$P(1 \le Y, \le 5) = 0.4$$

iii. 
$$P(C=0) = 0.3$$

You need to specify the value of P(X = x, Y = y, C = c). How many relevant values exist?

$$A \sim B(1, \frac{6}{7})$$

$$B \sim B(1, \frac{4}{7})$$

$$X = 5 + CA$$

$$Y = 6 - CB$$

$$C \sim B(1, 0.7)$$

Relevant values for when C = 1:

$$P(X = x, Y = y, C = 1)$$

$$= P(5 + X = x \mid C = 1)P(6 - Y = y \mid C = 1)P(C = 1)$$

$$= P(X = x - 5 \mid C = 1) \cdot P(Y = 6 - y \mid C = 1) \cdot 0.7$$

$$= P(X = x - 5 \mid C = 1) \cdot P(Y = 6 - y \mid C = 1) \cdot 0.7$$

$$= \binom{1}{x - 5} \frac{6^{x - 5}}{7} \frac{1^{1 - (x - 5)}}{7} \binom{1}{6 - y} \frac{4^{6 - y}}{7} \frac{3^{1 - (6 - y)}}{7} \cdot 0.7$$

$$= \frac{6^{x - 5}}{7} \frac{1^{1 - x + 5}}{7} \frac{4^{6 - y}}{7} \frac{3^{1 - 6 + y}}{7} \cdot 0.7$$

$$= \frac{6^{x - 5}}{7} \frac{1^{6 - x}}{7} \frac{4^{6 - y}}{7} \frac{3^{y - 5}}{7} \cdot 0.7$$

$$=\frac{6}{7}^{x-5}\frac{1}{7}^{1-x+5}\frac{4}{7}^{6-y}\frac{3}{7}^{1-6+y}\cdot 0.7$$

$$=\frac{6}{7}^{x-5}\frac{1}{7}^{6-x}\frac{4}{7}^{6-y}\frac{3}{7}^{y-5}\cdot 0.7$$

$$= \frac{6^{x-5}}{7} 4^{6-y} \frac{3^{y-5}}{7} \cdot \frac{7}{10}$$

$$= 6^{x-5} \cdot 4^{6-y} \cdot 3^{y-5} \cdot \frac{1}{70} = 2^{x-5+12-2y} \cdot 3^{x+y-10} \cdot \frac{1}{70}$$

$$= 2^{x-2y+7} \cdot 3^{x+y-10} \cdot \frac{1}{70}$$

For the values of x = 5, 6 and y = 5, 6

When C=0, the only relevant values are when X=5, Y=6, for which:  $P(X=5, Y=6, C=0) = P(X=5 \mid C=0) \\ P(Y=6 \mid C=0) \\ P(C=0) = 1 \cdot 1 \cdot 0.3 = 0.3$ 

C = 1	X = 5	X = 6	P(Y=k)
Y = 5	0.057	0.343	0.4
Y = 6	0.043	0.257	0.3
P(X = k)	0.1	0.6	0.7

C = 0	X = 5	X = 6	P(Y=k)
Y = 5	0	0	0
Y = 6	0.3	0	0.3
P(X = k)	0.3	0	0.3

### Question 4

The probability of Wolt arriving on time is 0.75.

a.

What is the probability of having 2 on-time meals in a week (7 days)?

$$X \sim B(7, 0.75)$$
  
 $P(X = 2) = \binom{7}{2} \cdot 0.75^2 \cdot (1 - 0.75)^{7-2} = \binom{7}{2} \cdot 0.75^2 \cdot 0.25^5 = 21 \cdot 0.5625 \cdot 0.000976 = 0.01154$ 

b.

What is the probability of having at least 4 on-time meals in a week?

$$P(X \ge 4) = 1 - P(X \le 3) = 1 - 0.07056 = 0.92944$$

c.

A company of 100 employees recorded the number of on-time meals they had during a particular week and averaged their results. What do you expect the value of that average to be?

Let 
$$X_i \sim B(7, 0.75)$$
,

$$Y = \sum_{i=1}^{100} X_i \sim B(7 * 100, 0.75)$$

Now we take the average as follows:

$$E(\frac{Y}{100}) \stackrel{linearity}{=} \frac{E(Y)}{100} \stackrel{\text{binomial formula}}{=} \frac{7*100*0.75}{100} = 5.25$$