BLG 454E Learning From Data (Spring 2018)
Homework I

Safa Keskin

March 6, 2018

1 Question 1

$P(A/B) = \frac{P(B/A)*P(A)}{P(B)}$

This is stated in Bayes' Theorem, and this equation means that, probability of observing A given that B is equal to multiplication of probability of observing B given that A and probability of observing B.

For our question, parameters can be expressed as:

P(A): probability of rain on Saturday ($P(A) = \frac{1}{4}$)

P(B): probability of rain on Sunday
$P(B) = P(B/A) + P(B/notA) = P(B) = \frac{1}{4} * \frac{1}{2} + \frac{3}{4} * \frac{1}{4} = \frac{5}{16}$

P(A / B): probability of rain on Saturday given that it rained on Sunday

P(B / A): probability of rain on Sunday given that it rained on Saturday ($P(B/A) = \frac{1}{2}$)

$$P(A/B) = \frac{\frac{1}{2} * \frac{1}{4}}{\frac{5}{16}} = \frac{2}{5} \tag{1}$$

2 Question 2

Let's call the event "reaching destination A in at most 2 steps" as E. E consist of 3 sub-events E_0, E_1, E_2, these are the events reaching destination A at 0,1,2 steps respectively. Let the S become the event reaching A.

$P(E) = \quad P(E_0) + P(E_1) + P(E_2)$

$P(E_0) = \quad P(S/A)$
E_0 is reaching A in 0 moves. $P(E_0\ )$ is $\frac{1}{7}$, because this case can be observed only if A is the starting point.

$P(E_1) = \quad P(S/B) + P(S/F) + P(S/G)$
E_1 is reaching A in 1 moves. $P(E_1\ )$ can be calculated as sum of the probabilities of reaching A from B,G or F, because these are adjacent points of A. The equation is:
$P(E_1) = P(S/B) + P(S/F) + P(S/G) = \frac{1}{7} * \frac{1}{3} + \frac{1}{7} * \frac{1}{6} + \frac{1}{7} * \frac{1}{3} = \frac{5}{42}$

$P(E_2) = \quad P(S/C) + P(S/E) + P(S/D) + P(S/B) + P(S/F) + P(S/G)$ E_2 is reaching A in 1 moves. $P(E_2\ )$ can be calculated as sum of the probabilities of reaching A from C,E,D,B,F or G, because A is reachable at 2 moves from these points. The equation is:
$P(E_2) =]P(S/C2) + P(S/E2) + P(S/D2) + P(S/B2) + P(S/F2) + P(S/G2)$
$= \frac{1}{42} + \frac{2}{63} + \frac{1}{63} + \frac{2}{63} = \frac{5}{63} + \frac{1}{42}$

Then,

$$\begin{aligned} P(E) &= P(E_0) + P(E_1) + P(E_2) \\ P(E) &= \frac{1}{7} + \frac{5}{42} + \frac{1}{42} + \frac{5}{63} = \frac{23}{63} \end{aligned} \tag{2}$$

3 Question 3

3.1 A-

Likelihood function is: $\prod_{i=1}^{n} P(x_i/\theta)$

Density function: $\frac{1}{\sqrt{2*\pi*\sigma^2}} * e^{\frac{-(x_i-\theta_1)^2}{2*\theta_2}}$

We will apply our likelihood function over density function in order to find value of the parameters μ and σ^2.

$L(\mu,\sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2*\pi*\sigma^2}} * e^{\frac{-(x_i-\mu)^2}{2*\sigma^2}} = \frac{1}{\sigma^n*(2\pi)^{\frac{n}{2}}} * e^{(\frac{-1}{2\sigma^2}*\sum_{i=1}^{n}(x_i-\mu)^2)}$

Then, we take the logarithm of this equation.

$lnL(\mu,\sigma^2) = ln(\frac{1}{\sigma^n*(2\pi)^{\frac{n}{2}}} * e^{(\frac{-1}{2\sigma^2}*\sum_{i=1}^{n}(x_i-\mu)^2)})$

$= -n * ln(\sigma) - \frac{n}{2} * ln(2*\pi) - \frac{1}{\sigma^2} * \sum_{i=1}^{n}(x_i-\mu)^2$

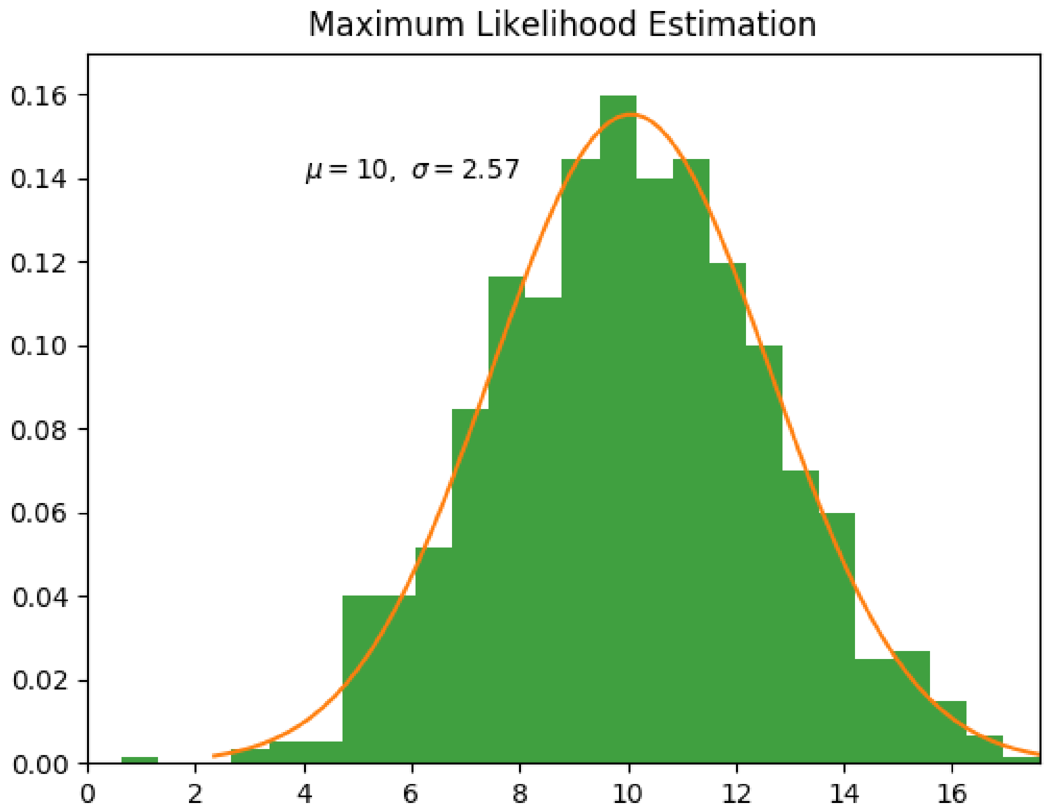
If we take partial derivative of this equation respect to μ and σ^2 respectively, we will get the equation we desire.
First, we will take derivative of equation respect to μ and set it equal to 0.

$$\begin{aligned}
&\frac{\partial lnL(\mu,\sigma^2)}{\partial\mu} = 0\\
&-2 * \frac{\sum(x_i-\mu)*(-1)}{2*\sigma^2} = 0\\
&\sum x_i = n-\mu\\
&\mu = \frac{\sum x_i}{n}
\end{aligned} \tag{3}$$

Then, we will take derivative of equation respect to σ^2 and set it equal to 0.

$$\begin{aligned}
&\frac{\partial lnL(\mu,\sigma^2)}{\partial\sigma^2} = 0\\
&-\frac{n}{2*\sigma^2} + \frac{\sum(x_i-\mu)^2}{2*\sigma^2} = 0\\
&n*\sigma^2 = \sum(x_i-\mu)^2\\
&\sigma^2 = \frac{\sum(x_i-\mu)^2}{n}
\end{aligned} \tag{4}$$

3.2 B-



4 Question 4

4.1 A-

$P(C/X_{1...n}) = \frac{\prod_{i=1}^{n} P(x_i/C)*P(C)}{P(x_{1...n})}$

$P(C/X) = \frac{P(X/C)*P(C)}{P(X)}$

We know that, $P(C=+) = \frac{1}{2}$ and $P(C=-) = \frac{1}{2}$ according to data.

As mentioned in Naive Bayes' Theorem, three different parameters can be considered as separate parameters that affect output class. So, we can construct classes as:

$P(x_1 = 1/y = +) = \frac{3}{5}$ $P(x_1 = 1/y = -) = \frac{2}{5}$

$P(x_2 = 1/y = +) = \frac{2}{5}$ $P(x_2 = 1/y = -) = \frac{2}{5}$

$P(x_3 = 1/y = +) = \frac{4}{5}$ $P(x_3 = 1/y = -) = \frac{1}{5}$

4.2 B-

We know that $P(C/X)\ \alpha\ P(X/C) * P(C)$ according to Bayes' Theorem formula. And also because $P(C=+) = P(C=-) = \frac{1}{2}$, there is no effect by term $P(C)$ in this calculation for this specific example. So,

$P(X_1 = 1, X_2 = 1, X_3 = 1/C = -) = P(X_1 = 1/C = -) * P(X_2 = 1/C = -) * P(X_3 = 1/C = -)$

$= \frac{2}{5} * \frac{2}{5} * \frac{1}{5} = 0.032$

$P(X_1 = 1, X_2 = 1, X_3 = 1/C = +) = P(X_1 = 1/C = +) * P(X_2 = 1/C = +) * P(X_3 = 1/C = +)$

$= \frac{3}{5} * \frac{2}{5} * \frac{4}{5} = 0.192$

$$\begin{aligned} &P(C = +/X_{1,2,3} = 1) > P(C = -/X_{1,2,3} = 1) \\ &\implies \quad C = + \end{aligned} \tag{5}$$

4.3 C-

If $P(A/B) = P(A)$ and $P(B/A) = P(B)$, then these two events are independent.
According to data in the table, $P(A) = \frac{1}{2}$ and $P(B) = \frac{2}{5}$. Now, we should calculate below probabilities in order to prove that these two events are independent.

$$\begin{aligned} P(A/B) &= \frac{P(B/A) * P(A)}{P(B)} \\ &= \frac{\frac{2}{5} * \frac{1}{2}}{\frac{2}{5}} \\ &= \frac{1}{2} \\ \\ P(B/A) &= \frac{P(A/B) * P(B)}{P(A)} \\ &= \frac{\frac{2}{4} * \frac{2}{5}}{\frac{1}{2}} \\ &= \frac{2}{5} \end{aligned} \tag{6}$$

As it is seen, $P(A/B) = P(A)$ and $P(B/A) = P(B)$, these two events are independent.