Partial exam 2 - Wednesday 2 May 2018 - Duration : 60 min

No document, no phone, no computing machine.

Name : First name : Signature :

Exercise 1 :	Exercise 2 :	Grade /20 :

Exercise 1 (Gaussian Bayes Classifier, ≃ 8 pts)

Suppose you have the following training set with one real-valued input X and a categorical output Y that has two values A and B.

X	0	2	3	4	5	6	7
Y	A	A	B	B	B	B	B

1. You must learn the Gaussian Bayes Classifier from this data. Write the parameters of the classifiers in this table :

$\mu_A =$	$\sigma_A^2 =$	$\Pr(Y = A) =$
$\mu_B =$	$\sigma_B^2 =$	$\Pr(Y = B) =$

Justify your calculation hereafter :

2. Calculate $\alpha = f_{X|Y}(X = 2|Y = A)$ and $\beta = f_{X|Y}(X = 2|Y = B)$. Do not propose any numerical approximation ; just give a simplified closed form expression.

3. What is the joint probability $f_{X,Y}(X=2, Y=A)$? The answer must be given in terms of α and β only.

4. What is the joint probability $f_{X,Y}(X=2, Y=B)$? The answer must be given in terms of α and β only.

5. What is $f_X(X=2)$? The answer must be given in terms of α and β only.

6. What is the conditional probability $\Pr(Y=A|X=2)$?

7. Consider the figure 1. If you trained a new Bayes classifier on this data, what class would you predicted for the query location indicated with " ?" ? Explain carefully your answer.

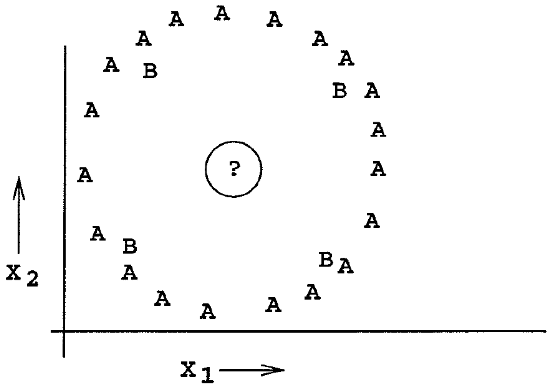

FIGURE 1 – Training data set and query location indicated with " ?".

Exercise 2 (Test and p-value, $\simeq$ 12 pts)

Assume that x is a sample of a random variable X following an exponential distribution with the unknown parameter θ. The exponential probability density function with parameter θ is

$$f_\theta(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & \text{for } x > 0, \\ 0 & \text{elsewhere.} \end{cases}$$

We want to test $H_0 : \{\theta = \theta_0\}$ versus $H_1 : \{\theta = \theta_1\}$ with $0 < \theta_1 < \theta_0$.

1. Calculate the cumulative distribution function $F_\theta(x)$ associated to $f_\theta(x)$.

2. Calculate the decision function $d(x)$ of the Neyman-Pearson test of size α. Simplify it such that the final decision function consists in comparing x to a threshold h_α.

3. Calculate the threshold h_α of the test. The threshold must be given in closed-form.

4. Describe carefully the critical region C_α of the test.

5. Calculate the power of the test, i.e., the probability γ to reject H_0 when H_1 is true.

6. Show that $C_\alpha \subset C_{\alpha'}$ if $\alpha < \alpha'$.

7. Calculate the p-value $\hat{p}(x)$ of the sample x from the definition of the p-value.

8. Show that $\hat{p}(x)$ is uniformly distributed over $[0,1]$ when x follows the exponential distribution with pdf $f_{\theta_0}(\cdot)$.

9. Propose a test equivalent to the Neyman-Pearson test of question 2 whose decision function is $\hat{p}(x)$. Precise clearly the threshold of the test.