Partial exam 2 - Jeudi 4 mai 2017 - Duration : 60 min
No document, no phone, no computing machine.

Name :   First name :   Signature :

Exercise 1 :	Exercise 2 :	Grade /20 :

Exercise 1 (Naive Bayes Test, ≃ 10 pts)
Consider the following data set with two predictive attributes
   — "T=Temperature" taking on the possible values : low, medium and high,
   — "W=Wind" taking on the possible values : weak and strong,
and the Boolean classification "S=Swim" with two outcomes "yes" or "no" :

T	W	S
low	weak	no
low	strong	no
medium	weak	yes
high	strong	no

1. Calculate the look up tables from the training dataset. Show your calculation.

2. We now encounter a new example : "T=high" and "W=weak". How should this example be classified using the Naive Bayes test ? Show your calculation.

3. When you encounter the situation "T=high" and "W=weak", which intuitive decision seems reasonable to take? Does the prediction be consistent with the intuition? Explain.

Exercise 2 (Time series, ≃ 10 pts)

1. Give the mathematical definition of an ARMA(p,q) time series. All variables and notations must be defined carefully.

2. An ARMA(p,q) time series is a weakly stationary stochastic process. What does it mean? Give a mathematical answer.

3. Give the definition (all notations must be introduced) of the autocorrelation function $(\rho_k)_{k\in\mathbb{Z}}$ for a weakly stationary time series $(X_t)_{t\in\mathbb{Z}}$ where $\mathbb{Z}$ is the set of negative and positive integers.

4. Calculate the autocorrelation function $(\rho_k)_{k\in\mathbb{Z}}$ of a MA(1) time series with coefficients β_0 and β_1.
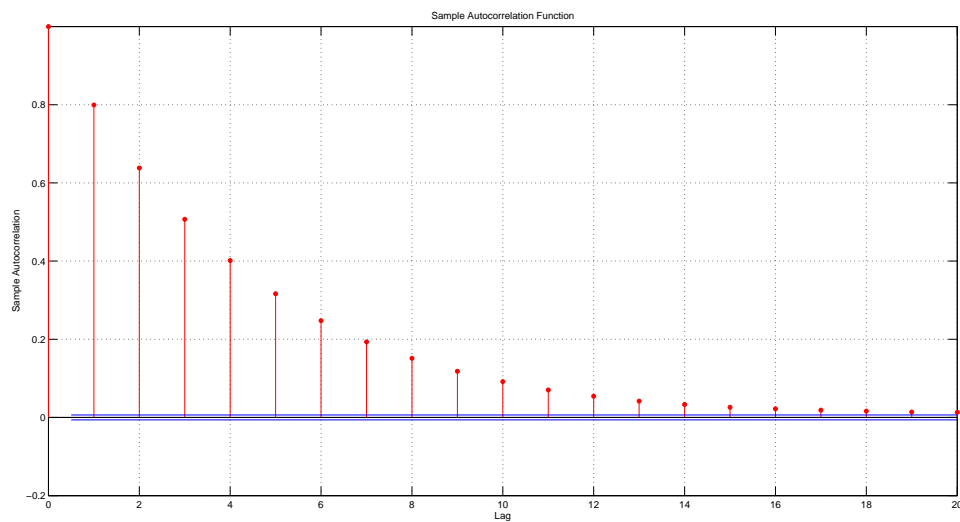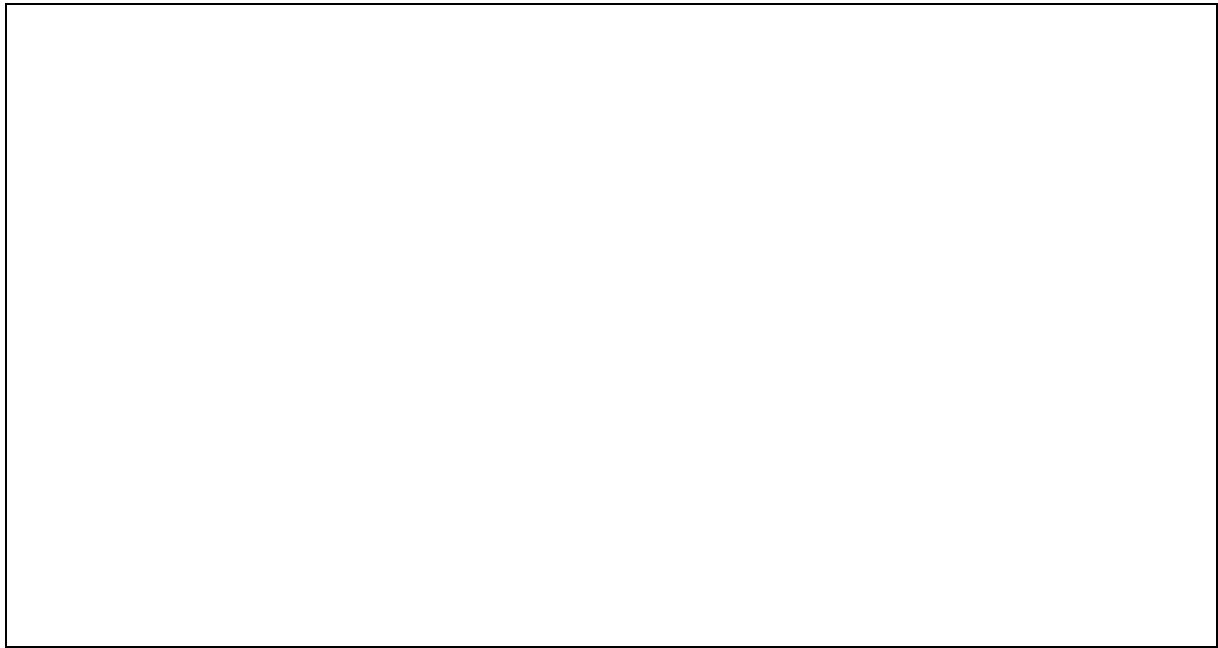
FIGURE 1 – Sample autocorrelation of a time series.

5. Figure 1 shows the autocorrelation of a time series. Do you think that this times series could be well modeled by MA(1)? If not, which model could be more relevant?