Columbia University
Statistical Analysis and Time Series

IEOR-4709

A. Capponi — **Problem Set #1** — Issued: January 27, 2025
Spring 2025 — Due: **BEFORE CLASS** February 5, 2025

Note: Please put the number of hours that you spent on this homework set on top of the first page of your homework. The TA in charge of this homework is Boxuan Li. The CA in charge of this homework is Kishore Kuppusamy.

Ex. 1.
Suppose $\{X_n\}_{n\geq 1}$ is a sequence of independent and identically distributed random variables with mean μ_X and variance σ_X^2. Additionally, suppose $\{Y_n\}_{n\geq 1}$ is a sequence of independent and identically distributed random variables with mean μ_Y and variance σ_Y^2. Assume that the sequences $\{X_n\}_{n\geq 1}$ and $\{Y_n\}_{n\geq 1}$ are independent. Let $\mu_{XY} := \mathbb{E}\left[XY\right]$ and $\sigma_{XY} = Cov(X,Y)$. Consider the estimator $s_{XY} := \frac{1}{n-1}\sum_{i=1}^n \left(X_i - \bar{X}\right)(Y_i - \bar{Y})$, where $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$, and $\bar{Y} = \frac{1}{n}\sum_{i=1}^n Y_i$.

- Show that s_{XY} is a consistent estimator of σ_{XY}.
- Show that s_{XY} is an unbiased estimator of σ_{XY}.

Ex. 2.
Let $X_1, X_2, \ldots, X_n$ be random variables with mean μ and variance σ^2. What are the method of moment estimators of the mean μ and variance σ^2?

Ex. 3.
Consider a mixture distribution with two components. Let X have the density function:

$$f(x;\theta) = \frac{1}{2}f_1(x;\theta_1) + \frac{1}{2}f_2(x;\theta_2)$$

where $f_1(x;\theta_1)$ and $f_2(x;\theta_2)$ are two different probability density functions with parameters θ_1 and θ_2, respectively. Assume

$$f_1(x;\theta_1) = \frac{1}{\sqrt{2\pi\sigma_1^2}}\exp\left(-\frac{(x-\theta_1)^2}{2\sigma_1^2}\right)$$

and

$$f_2(x;\theta_2) = \frac{1}{\sqrt{2\pi\sigma_2^2}}\exp\left(-\frac{(x-\theta_2)^2}{2\sigma_2^2}\right).$$

Assume σ_1 and σ_2 are know. Find the method of moment estimator for θ_1 and θ_2.

Ex. 4.
Let $X_1, X_2, \ldots, X_n$ be i.i.d. Gaussian random variables with mean μ and standard deviation σ. Do the following

- Is the sample mean $M_n := \frac{1}{n}\sum_{i=1}^{n} X_i$ a Gaussian random variable? Compute the mean and variance of the sample mean.
- Find the value of n such that $\mathbb{P}(79 \leq M_n(X) \leq 81) = 0.99$. Assume $\mu = 80$ and $\sigma = 22$.

Ex. 5.
Go on Google or Yahoo finance and download the Banco Bilbao Vizcaya Argentaria SA, and Microsoft time series from January 18, 2024 back to January 22, 2022. Do the following:

- Produce a scatterplot of each time return time series (compute the returns using open prices). What do you observe? Comment on where they are centered, and how spread out the graphs are.
- Produce a histogram of each time return time series. Does the histogram confirm your findings from the scatterplot?
- Compute the mean and variance of each return time series. Do these statistics reflect your perception from the graphs?
- Compute the covariance of the two return time series. What do you observe? Can you explain your findings using common arguments? (for instance the fact that they belong to difference sectors, the fact that the two companies are located in different countries, etc...)