Université Paris I, Panthéon - Sorbonne

Master 2 M.O. 2020 – 2021

Time Series Tutorial n^0 1 : How to manipulate and generate time series?

The aims of this tutorial is a to provide a first overview on the objects and commands of the R software relative to time series.

R is a free software of numerical applied mathematics and more precisely statistics. It is continuously developed from the more recent researches of mathematicians. It can be downloaded from `http://www.r-project.org/` as well for Windows, Mac or Linux. An accurate version of R, called R-studio, could also be downloaded on `https://www.rstudio.com/`.

In the sequel, you may read at the left side the R commands which could be written on your computer, comments on these commands could be found on the right-side.

An interesting way of using this software is to open a R-script (on File) and write all your commands indeed. You could save the R-script and directly make run the commands on the command window.

Manipulations of R objects

x=c(4,3,7,1.7,3,8.4,5,12,9,12)	Generate a numerical vector (`c` as collector).
xx=rep(c(-1,4,-log(2)),3)	To repeat a sequence of numbers.
x=c(x,xx)	New vector x from its previous affectation.
x	To print the value of x.
is.ts(x)	Is x a "time series" object?
y=as.ts(x)	Transform x to a "time series" object..
y	Changes?
time(y)	See the vector of time of this time series
is.ts(y)	Just to verify...
	More generally, R software run with several types of objects:
	Vectors (commands *is.vector* and *as.vector*);
	Matrices (commands *is.matrix* and *as.matrix*);
	Data tables (commands *is.data.frame* and *as.data.frame*);
	Time series (commands *is.ts* and *as.ts*);
	Lists (commandes *is.list* and *as.list*);
t=tsp(y)	Associate to y a vector of times t (by default from 1 to 1, with frequency 1).
plot.ts(y)	Draw the path of the time series y..
plot.ts(y,type="b")	To indicate the points of the time series.
z=ts(y,freq=4)	Divide the time series in trimesters (run also with divisions in months where $freq = 12$ ...).
z	Verification.
time(z)	Associate vector of time
z=ts(x,freq=4,1991+1/4, 1993)	A new vector of times, with the beginning and the end.
time(z)	
plot.ts(z)	New graph.
frequency(z)	To specify the frequency.
length(z)	Length of the time series z.

Generation of a time series

	In the help desk, write the keyword "distributions".
	Note that prefix "r" is used to generate pseudo-random variables, "d" for density,
	"p" for cumulative distribution function and "q" for quantile.
qnorm(0.95)	95% percentile of standard Gaussian distribution.
pnorm(2)	Cumulative distribution function in 2 of standard Gaussian distribution.
x=rnorm(300)	On génère un bruit blanc.
mean(x)	Empirical mean.
sd(x)	Standard deviation.
var(x)	Empirical deviation. Which is the renormalization?
cov(x,x ∧2)	Empirical covariance between both the vectors.
cor(x,x)	Empirical correlation. Explain the result?
cor(x,x∧2)	Empirical correlation. Explain the result?
par(mfrow=c(2,2))	Share the window in 4 windows.
plot.ts(x)	Graph.
hist(x,nclass=6)	Histogram with a specified number of classes.
	Generate 100 independent realizations of $\mathcal{N}(-1,3)$ random variables.
	Compute the empirical mean, the standard deviation and draw an histogram.
	Explain the results...
x=rbinom(30,10,0.3)	Generation of another vector of independent binomial random variables.
	Compute the usual statistics relative to this vector.
	Center and normalize this time series. Let z be this new vector.
y=as.ts(x)	Transform y in a time series.
y=sort(y)	To order y.
plot.ts(x,y)	Explain this graph.
acf(x)	Correlogram of x. Explain the result.
acf(y)	Correlogram of y. Explain the result.

Exercices

Exercice 1: Let $(\varepsilon_i)_{i\in\mathbf{Z}}$ be a white noise of $[-3,3]$-uniform distribution.

1. Generate a realization of $(\varepsilon_1,\cdots,\varepsilon_{100})$.
2. Let $X_i=\varepsilon_{i+1}-2\,\varepsilon_i$ for $i\in\mathbf{Z}$. Generate $(X_1,\cdots,X_{100})$.
3. Draw the correlogram of $(\varepsilon_1,\cdots,\varepsilon_{100})$ and $(X_1,\cdots,X_{100})$. Explain the results.

Exercice 2: Let $Z=(Z_i)_{1\leq i\leq n}$ be a vector of independent standard Gaussian random variables.

1. Let $A=(A_1,\cdots,A_n)$ be a (deterministic) vector of real numbers and Σ be a (n,n) definite positive matrix. Show that $A+\Sigma^{1/2}\,Z$ is a realization of a $\mathcal{N}(A\,,\,\Sigma)$ random vector. In the sequel $n=100$.
2. Use R software to generate a realization X of a $\mathcal{N}(A\,,\,\Sigma)$ random vector with $A=0$ and $\Sigma_{ii}=5$, $\Sigma_{ij}=-2$ if $|j-i|=1$ and $\Sigma_{ij}=0$ if $|j-i|>1$ for $1\leq i,j\leq n$ (you may use the command `matrix` and `chol`).
3. Prove that the distribution of $(X_i)_i$ is the same than the distribution of $(Z_{i+1}-2\,Z_i)_i$.
4. Compute the autocorrelation of (X_i) and draw the correlogram of $(X_1,\ldots,X_n)$. What's happening when n increases?
5. Using Monte-Carlo experiments, establish numerically the $\sqrt{n}\,(\widehat{\rho}_X(1)+0.4)\ \xrightarrow[n\to+\infty]{\mathcal{L}}\ \mathcal{N}(0\,,\,\gamma^2)$ and give an approximation of γ^2. Provide a theoretical prove of this result and the exact value of γ^2.