Στοχαστική Ανάλυση Χρονοσειρών

Περιεχόμενο:

Στατιστικά και Διαγράμματα που Χρησιμοποιούμε στην Ανάλυση Χρονοσειρών.
Μοντέλα AR(p) και MA(q).

Επαμεινώνδας Διαμαντόπουλος
Πέμπτη 4 Απριλίου 2024

Στατιστικά και Διαγράμματα που Χρησιμοποιούμε στην Ανάλυση Χρονοσειρών (συνέχεια)

Συντελεστής Συσχέτισης Pearson

Η συνδιακύμανση Cov(X, Y) είναι μία στατιστική ποσότητα που μας επιτρέπει να ποσοτικοποιήσουμε τον τρόπο με τον οποίο συμμεταβάλλονται δύο τυχαίες μεταβλητές X, Y. Η κανονικοποιημένη εκδοχή της συνδιακύμανσης είναι ο συντελεστής συσχέτισης Pearson $\rho_{X,Y}$, ο οποίος παίρνει τιμές από -1 έως +1.

$$\rho_{X,Y} = \operatorname{corr}(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\mathrm{E}[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y}$$

Δειγματικός Συντελεστής Συσχέτισης

$$r_{xy} \overset{\text{def}}{=} \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{(n-1)s_x s_y} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2 \sum_{i=1}^{n}(y_i-\bar{y})^2}},$$

Μερικός Συντελεστής Συσχέτισης (Partial Correlation Coefficient)

Ο μερικός συντελεστής συσχέτισης (partial correlation coefficient) υπολογίζει τη συσχέτιση δύο μεταβλητών X, Y ύστερα από την απομάκρυνση της μεταβλητότητας των X, Y που εξηγείται από μία τρίτη μεταβλητή Z.

$$\rho_{X,Y|Z} = \frac{\mathrm{Cov}(X, Y | Z)}{\sigma_{X|Z} \cdot \sigma_{Y|Z}}$$

Ο μερικός συντελεστής συσχέτισης δίνει περισσότερο ακριβή εικόνα για τη σχέση δύο μεταβλητών στην περίπτωση που είναι γνωστή μία τρίτη μεταβλητή που συσχετίζεται και με τις δύο. Ο υπολογισμός του μερικού συντελεστή συσχέτισης μπορεί να γενικευθεί για περισσότερες της μίας μεταβλητής, Z_1, Z_2, …, Z_n.

$$\rho_{X,Y|Z_1,Z_2,\ldots,Z_n} = \frac{\mathrm{Cov}(X, Y | Z_1, Z_2, \ldots, Z_n)}{\sigma_{X|Z_1,Z_2,\ldots,Z_n} \cdot \sigma_{Y|Z_1,Z_2,\ldots,Z_n}}$$

Στατιστικά και Διαγράμματα που Χρησιμοποιούμε στην Ανάλυση Χρονοσειρών

Αυτοδιακύμανση και Αυτοσυσχέτιση

Η αυτοδιακύμανση (autocovariance) και η αυτοσυσχέτιση (autocorrelation) είναι δύο στατιστικές ποσότητες που μας επιτρέπουν να ανιχνεύσουμε εξάρτηση των τιμών μίας χρονοσειράς $\{X_n, n \in N^*\} = \{X_1, X_2, X_3, \ldots\}$ με ένα μετρήσιμο τρόπο. Ειδικότερα, με τον τρόπο αυτό ανιχνεύουμε περιοδικότητα.

Η **αυτοδιακύμανση** είναι η συνδιακύμανση της μεταβλητής X_t των τιμών της χρονοσειράς με ένα αντίγραφό της X_{t-h} χρονικά μετατοπισμένο κατά πλήθος χρονικών στιγμών h = 1, 2, …

h = 1: $\mathrm{Cov}(X_t, X_{t-1}) = \mathrm{Cov}(\{X_1, X_2, X_3, \ldots\}, \{X_2, X_3, X_4, \ldots\})$

h = 2: $\mathrm{Cov}(X_t, X_{t-2}) = \mathrm{Cov}(\{X_1, X_2, X_3, \ldots\}, \{X_3, X_4, X_5, \ldots\})$

….

h = n: $\mathrm{Cov}(X_t, X_{t-n}) = \mathrm{Cov}(\{X_1, X_2, X_3, \ldots\}, \{X_{n+1}, X_{n+2}, X_{n+3}, \ldots\})$

Σημείωση: Φανερά, κάθε μία αυτοδιακύμανση υπολογίζεται σε πλήθος παρατηρήσεων μειωμένο κατά ένα από την προηγούμενη.

Στατιστικά και Διαγράμματα που Χρησιμοποιούμε στην Ανάλυση Χρονοσειρών

Αυτοδιακύμανση και Αυτοσυσχέτιση

Η αυτοσυσχέτιση είναι η κανονικοποιημένη ως προς το πιθανό εύρος τιμών εκδοχή της αυτοδιακύμανσης.

$$\rho_1 = \mathrm{Corr}(X_t, X_{t-1}) = \mathrm{Cov}(X_t, X_{t-1}) / [\mathrm{SD}(X_t)\mathrm{SD}(X_{t-1})]$$

$$\rho_2 = \mathrm{Corr}(X_t, X_{t-2}) = \mathrm{Cov}(X_t, X_{t-2}) / [\mathrm{SD}(X_t)\mathrm{SD}(X_{t-2})]$$

$$\dots$$

$$\rho_n = \mathrm{Corr}(X_t, X_{t-n}) = \mathrm{Cov}(X_t, X_{t-n}) / [\mathrm{SD}(X_t)\mathrm{SD}(X_{t-n})]$$

Στην περίπτωση όπου $\{x_1, x_2, x_3, \dots\}$ είναι συγκεκριμένες αριθμητικές τιμές τότε με αυτές εκφράζεται μία εκδοχή της στοχαστικής διεργασίας $\{X_n, n \in N^*\} = \{X_1, X_2, X_3, \dots\}$ που τις παράγει.

Στην περίπτωση αυτή η $\mathrm{Cov}(x_t, x_{t-n})$ ονομάζεται **δειγματική αυτοδιακύμανση** και η $\mathrm{Corr}(x_t, x_{t-n})$ **δειγματική αυτοσυσχέτιση**.

Στατιστικά και Διαγράμματα που Χρησιμοποιούμε στην Ανάλυση Χρονοσειρών

Δειγματική Συνάρτηση Αυτοσυσχέτισης (ACF)

Ο υπολογισμός της αυτοσυσχέτισης για όλες τις υστερήσεις Lag = 1, 2, 3, …, οδηγεί στον ορισμό της δειγματικής συνάρτησης αυτοσυσχέτισης (Autocorrelation Function – ACF). Πιο συγκεκριμένα, αν x_1, x_2, x_3, …, x_n είναι τα χρονικά δεδομένα, τότε η ACF είναι η συνάρτηση $r(h) = r_h: N \to R$, με

$$r_0 = \mathrm{Corr}(\{x_1, x_2, x_3, \ldots, x_n\}, \{x_1, x_2, x_3, \ldots, x_n\}),$$

$$r_1 = \mathrm{Corr}(\{x_1, x_2, x_3, \ldots, x_{n-1}\}, \{x_2, x_3, x_4, \ldots, x_n\}),$$

$$r_2 = \mathrm{Corr}(\{x_1, x_2, x_3, \ldots, x_{n-2}\}, \{x_3, x_4, x_5, \ldots, x_n\}),$$

….

$$r_h = \mathrm{Corr}(\{x_1, x_2, x_3, \ldots, x_{n-h}\}, \{x_{h+1}, x_{h+2}, x_{h+3}, \ldots, x_n\}),$$

Στατιστικά και Διαγράμματα που Χρησιμοποιούμε στην Ανάλυση Χρονοσειρών

Δειγματική Συνάρτησης Μερικής Αυτοσυσχέτισης (PACF)

Αν $x_1, x_2, x_3, \ldots, x_n$ είναι τα χρονικά δεδομένα, τότε η συνάρτηση μερικής αυτοσυσχέτισης φ_h με υστέρηση (lag) h, (Partial Autocorrelation Function – PACF) είναι η αυτοσυσχέτιση μεταξύ των $x_1, x_2, x_3, \ldots, x_{n-h}$ και $x_{h+1}, x_{h+2}, x_{h+3}, \ldots, x_n$ από την οποία αφαιρείται η γραμμική εξάρτηση όλων των ενδιάμεσων υστερήσεων 1, 2, …, h – 1.

Δηλαδή:

$$\varphi_1 = \text{Corr}(x_{t+1}, x_t) = r_1, \quad \varphi_2 = \text{Corr}\left(x_{t+2} - \hat{x}_{t+2}, x_t - \hat{x}_t\right), \quad \varphi_3 = \text{Corr}\left(x_{t+3} - \hat{x}_{t+3}, x_t - \hat{x}_t\right), \ldots$$

Όπου

α) Για το φ_2, $\hat{x}_{t+2}$ και $\hat{x}_t$ είναι εκτιμητές ελαχίστων τετραγώνων με ανεξάρτητη μεταβλητή το x_{t+1}

β) Για το φ_3, $\hat{x}_{t+3}$ και $\hat{x}_t$ είναι εκτιμητές ελαχίστων τετραγώνων με ανεξάρτητες μεταβλητές τις x_{t+1}, x_{t+2},

κλπ...

Στατιστικά και Διαγράμματα που Χρησιμοποιούμε στην Ανάλυση Χρονοσειρών

ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΡΑΚΗΣ | DEMOCRITUS UNIVERSITY OF THRACE

Δραστηριότητα

Τι μπορούμε να καταλάβουμε για μία χρονοσειρά με τα εξής στοιχεία;

$x_{n+1} = a \cdot x_n + b$

$x_{n+2} = a \cdot x_n + b$

$x_{n+3} = a \cdot x_n + b$

lag	acf	pacf
1	0.9	0.9
2	0.85	0.4
3	0.7	0.1

x_{n+1}, x_{n+2}

x_n ---- x_{n+3}

ns

Απάντηση

Από τις τιμές του ACF υποδεικνύεται πως κάθε όρος x_t, έχει σημαντική γραμμική σχέση με τους τρεις προηγούμενούς του x_{t-1}, x_{t-2}, x_{t-3}. Όμως, από τις τιμές του PACF, προκύπτει πως η γραμμική σχέση μεταξύ των x_t, x_{t-3}, παύει να είναι σημαντική ύστερα από την αφαίρεση της επιρροής των ενδιάμεσων δύο x_{t-1}, x_{t-2}. Συμπεραίνουμε, πως ένα μοντέλο που θα μπορούσε να προσαρμοστεί στη χρονοσειρά αυτή είναι το εξής: $x_t = \mu + \alpha \cdot x_{t-1} + \beta \cdot x_{t-2}$.

Στατιστικά και Διαγράμματα που Χρησιμοποιούμε στην Ανάλυση Χρονοσειρών

Άσκηση Εμπέδωσης

Τι μπορούμε να καταλάβουμε για μία χρονοσειρά με τα εξής στοιχεία;

lag	acf	pacf
1	0.8	0.8
2	0.75	0.15
3	0.65	0.1

$X_n = a \cdot X_{n-1} + b$

$\cancel{X_n = a \cdot X_{n-2} + b}$

$\cancel{X_n = a \cdot X_{n-3} + b}$

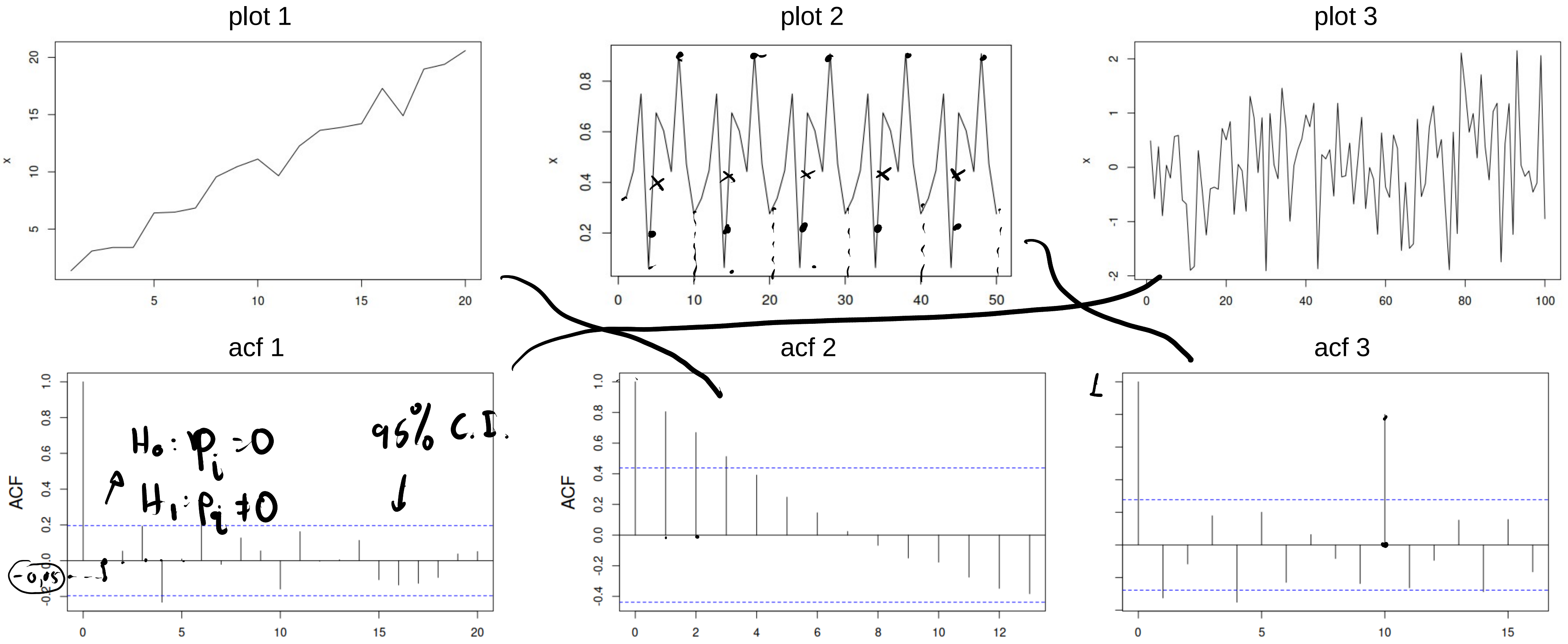
X_n --- ns --- X_{n-2}

X_{n-1}

Στατιστικά και Διαγράμματα που Χρησιμοποιούμε στην Ανάλυση Χρονοσειρών

$\rho_0 = Corr(X_t, X_t) = 1.$

Δραστηριότητα Αντιστοίχισης



ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΡΑΚΗΣ | DEMOCRITUS UNIVERSITY OF THRACE

Τύποι αναλυτικών μοντέλων

Τύποι αναλυτικών μοντέλων

Υπάρχουν πολλοί τύποι αναλυτικών μοντέλων. Δύο κύριες περιπτώσεις είναι οι εξής:

• Μοντέλα γραμμικής παλινδρόμησης στα οποία ο χρόνος τοποθετείται στη θέση της ανεξάρτητης μεταβλητής x. Στην περίπτωση αυτή είναι δυνατόν να χρησιμοποιηθεί η μέθοδος των ελαχίστων τετραγώνων για τον υπολογισμό των συντελεστών.

• Μοντέλα ARIMA (Autoregressive Integrated Moving Average) με τα οποία συσχετίζεται κάθε τιμή μιας χρονοσειράς με προηγούμενες τιμές και προηγούμενα σφάλματα πρόβλεψης.

ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΡΑΚΗΣ | DEMOCRITUS UNIVERSITY OF THRACE

Μοντέλα ARIMA

AR : Auto Regressive
I : Integrated
MA : Moving Average

Αυτοπαλινδρομικό μοντέλο AR(p)

Γενική εξίσωση AR(p)

Έστω B ο τελεστής μετατόπισης της ακολουθίας μία θέση προς τα πίσω (ή τελεστής υστέρησης ή τελεστής ολίσθησης, backward shift) που στέλνει κάθε στοιχείο της ακολουθίας στο αμέσως προηγούμενό του:

$$\mathbf{B(X_n) = X_{n-1}} \qquad B^2(X_n) = B(B(X_n)) = B(X_{n-1}) = X_{n-2}.$$

Φανερά $\mathbf{B^K(X_n) = X_{n-K}}$. Με χρήση του τελεστή υστέρησης B, η γενική εξίσωση ενός AR(p) μοντέλου

$$\mathbf{X_t = \delta + \alpha_1 \cdot X_{t-1} + \alpha_2 \cdot X_{t-2} + \ldots + \alpha_p X_{t-p} + \varepsilon_t,\ \ \varepsilon_t \sim IWN(0, \sigma^2),} \qquad \varepsilon_t : \text{iid} \sim N(0, \sigma^2).$$

indep. white noise.

εκφράζεται ως εξής:

$$(1 - \alpha_1 B - \alpha_2 B^2 - \ldots - \alpha_p B^p) X_t = \delta + \varepsilon_t,$$

ή ακόμα και ως

$$\mathbf{\Phi(B)X_t = \delta + \varepsilon_t,}$$

όπου $\mathbf{\Phi(\omega) = 1 - \alpha_1\omega - \alpha_2\omega^2 - \ldots - \alpha_p\omega^p}$, το **χαρακτηριστικό πολυώνυμο** της AR(p).

Χαρακτηριστικό Πολυώνυμο AR(p) και Στασιμότητα

μ : σταθ $< \infty$, Var = σταθ $< \infty$.

Οι ρίζες του Φ(ω) προσδιορίζουν αν θα είναι στάσιμη ή όχι η AR(p) διεργασία. Αποδεικνύεται το:

Θεώρημα

Έστω $\mathbf{\Phi(B)X_t = \delta + \varepsilon_t}$, και $\mathbf{\Phi(\omega) = 1 - \alpha_1\omega - \alpha_2\omega^2 - \ldots - \alpha_p\omega^p}$. Αν όλες οι ρίζες του Φ βρίσκονται έξω από τον μοναδιαίο δίσκο, τότε η AR(p) είναι στάσιμη.

Σημείωση:

Θέτοντας ω = 1/z, η εξίσωση Φ(ω) = 0, μετασχηματίζεται στην

$$\mathbf{z^p - \alpha_1 z^{p-1} - \alpha_2 z^{p-2} - \ldots - \alpha_p = 0}$$

Αυτή, αντίστοιχα οφείλει να έχει όλες τις ρίζες της μέσα στο μοναδιαίο δίσκο.

Χαρακτηριστικό Πολυώνυμο AR(p) και Στασιμότητα

Περίπτωση AR(1): $X_t = \delta + \alpha X_{t-1} + \varepsilon_t$,

Είναι $\Phi(\omega) = 1 - \alpha\omega = 0$ ή $\omega = 1/\alpha$. Άρα: Το AR(1) είναι στάσιμο αν και μόνο αν $|\alpha| < 1$,

Περίπτωση AR(2): $X_t = \delta + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \varepsilon_t$,

Είναι $\Phi(\omega) = 1 - \alpha_1\omega - \alpha_2\omega^2 = 0$ ή $z^2 - \alpha_1 z - \alpha_2 = 0$. (2)

Οι ρίζες είναι $z_{1,2} = \alpha_1 \pm (\alpha_1^2 - 4\alpha_2)^{1/2}$. Αποδεικνύεται(1) ότι αυτές είναι στο εσωτερικό του μοναδιαίου δίσκου, όταν

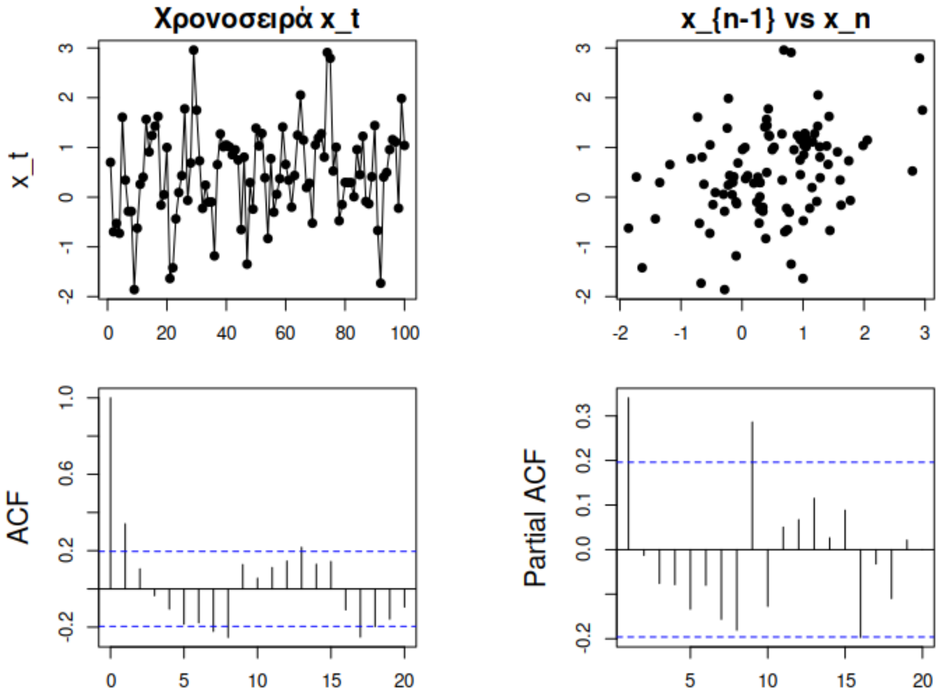
$$\alpha_1 + \alpha_2 < 1, \quad \alpha_2 - \alpha_1 < 1, \quad |\alpha_2| < 1.$$

(1) Μία απόδειξη είναι διαθέσιμη εδώ: https://stats.stackexchange.com/questions/118019/a-proof-for-the-stationarity-of-an-ar2

Χαρακτηριστικό Πολυώνυμο AR(p) και Στασιμότητα

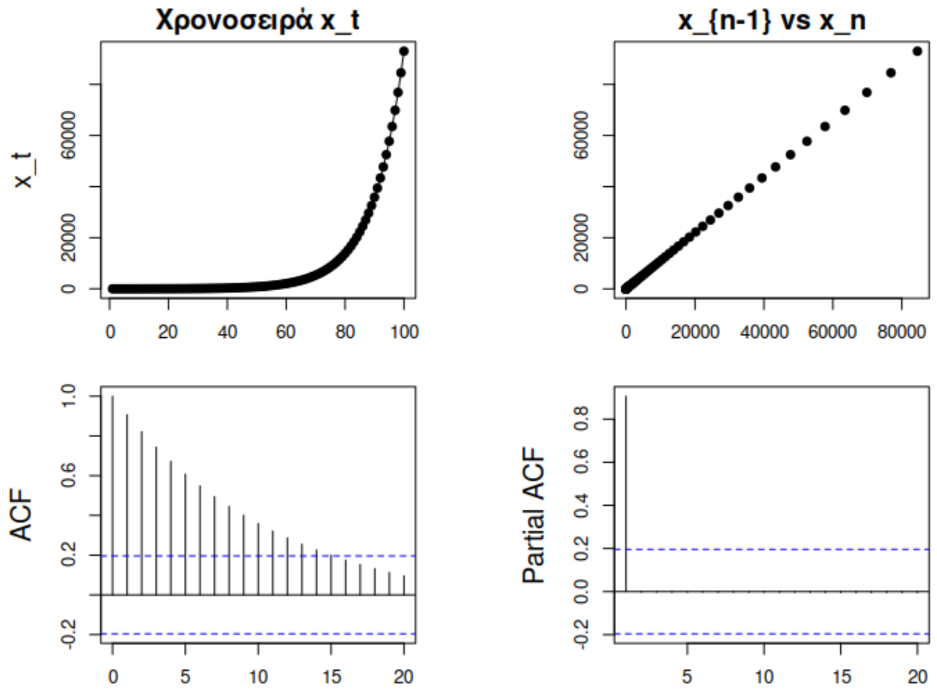
$0,3 = \theta < 1$

$$X_t = 0.4 + 0.3\, X_{t-1} + \varepsilon_t$$



$X_1 \qquad \theta = 1,1 > 1$

$$X_t = 0.4 + 1.1 \cdot X_{t-1} + \varepsilon_t,$$



Χαρακτηριστικό Πολυώνυμο AR(p) και Στασιμότητα

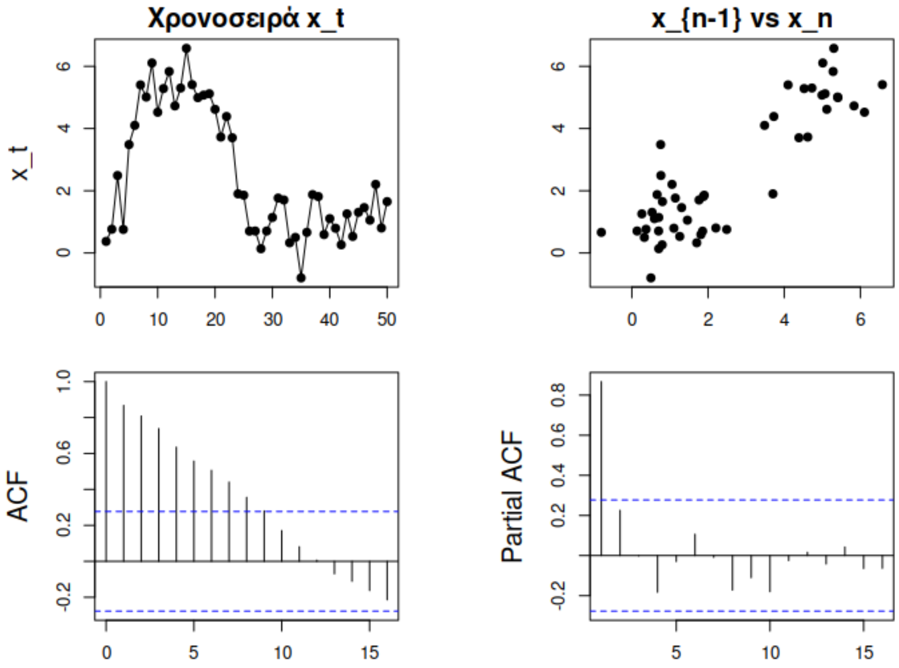
$\alpha_1 + \alpha_2 < 1 \quad \alpha_2 - \alpha_1 < 1, \ |\alpha_2| < 1$

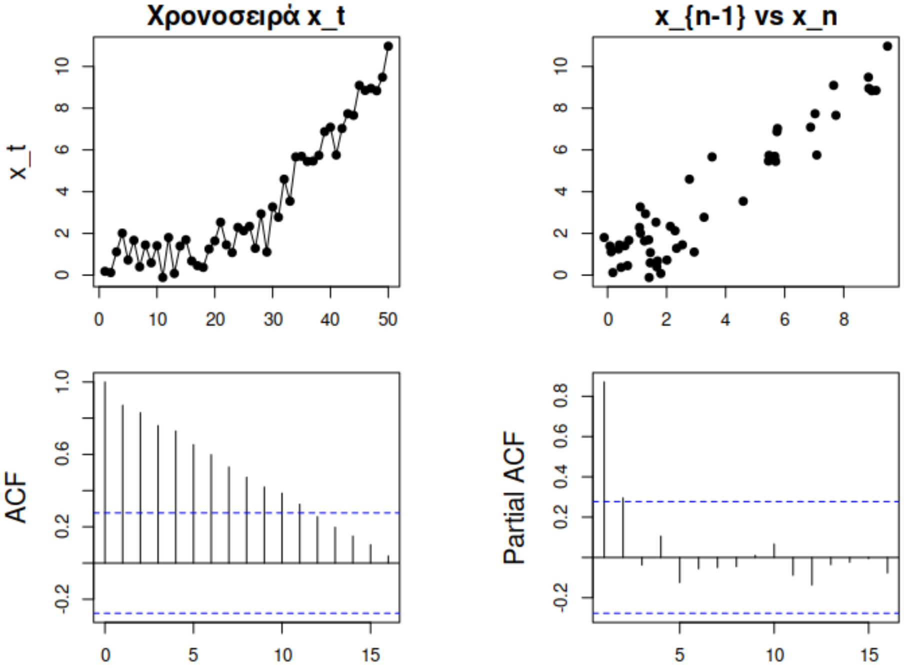
$\alpha_1 \qquad \alpha_2$

$X_t = 0.4 + 0.6\, X_{t-1} + 0.3\, X_{t-2} + \varepsilon_t$

$\alpha_1 + \alpha_2 = 1.$

$X_t = 0.4 + 0.7\, X_{t-1} + 0.3\, X_{t-2} + \varepsilon_t$





Χαρακτηριστικό Πολυώνυμο AR(p) και Στασιμότητα

Κώδικας R

```
# X_t = 0.4 + 1.1X_{t−1} + e_t
et = rnorm(100)
xt = runif(100, 0, 1)
for(i in 1:99){
  xt[i+1] = 0.4 +  1.1*xt[i] + et[i]
}
my.plot.ts.short(xt)
# η) X_t = 0.4 + 0.7X_{t−1} + 0.3X_{t−2} + e_t
et = rnorm(50)
xt = runif(50, 0, 1)
for(i in 1:48){
  xt[i+2] = 0.4 + 0.6*xt[i+1] + 0.3*xt[i] + et[i]
}
my.plot.ts.short(xt)
```

```
my.plot.ts.short = function(x){
  library(tseries)
  library(Hmisc)
  par(mfrow=c(2,2))
  par(mar = c(2, 4.5, 2, 4.5))
  plot(x, main = "Χρονοσειρά x_t", xlab = "t", ylab = "x_t", cex =
1.5, cex.lab = 1.5, cex.main = 1.5, pch=20);
  lines(1:length(x), x, pch=20)
  plot(na.omit(Lag(x, 1)), x[2:length(x)], main = paste0('x_{n-1}
vs x_n'), ylab = '', xlab = paste0('x_{t-1}'), cex = 1.5, cex.main
= 1.5,cex.lab = 1.5, pch=20)
  acf(x, pl=T, cex.lab = 1.5, cex.main = 1.5)
  pacf(x, pl=T, cex.lab = 1.5, cex.main = 1.5)
  par(mfrow=c(1,1))
}
```

Χαρακτηριστικό Πολυώνυμο AR(p) και Στασιμότητα

Παράδειγμα I

$z^2 \quad -2z \quad +0.5$

Να βρεθεί αν η AR(2) χρονοσειρά $X_t = 3 + 2X_{t-1} – 0.5X_{t-2} + w_t$, είναι στάσιμη.

Λύση

Θεωρούμε την εξίσωση

$$\mathbf{z^2 – 2z + 0.5 = 0.}$$

Οι ρίζες της εξίσωσης είναι $z_{1,2} = 1 \pm 0.5^{½}$. Παρατηρούμε, πως η $z_1 = 1 + 0.5^{½}$, δεν ανήκει στο μοναδιαίο δίσκο ($|z_1| > 1|$), άρα η χρονοσειρά δεν είναι στάσιμη.

Χαρακτηριστικό Πολυώνυμο AR(p) και Στασιμότητα

Παράδειγμα ΙΙ

Να βρεθεί αν η AR(2) χρονοσειρά $X_t = 5 + 0.4\,X_{t-1} – 0.1\,X_{t-2} + w_t$, είναι στάσιμη.

Λύση

Θεωρούμε την εξίσωση

$$\mathbf{z^2 – 0.4z + 0.1 = 0.}$$

Οι ρίζες της εξίσωσης είναι $z_{1,\,2} = 0.2 \pm i\cdot 0.06^{½}$. Παρατηρούμε πως

$$|z_{1,2}| = (0.2^2 + 0.06)^{½} = 0.1^{½} < 1,$$

άρα η χρονοσειρά είναι στάσιμη.

Χαρακτηριστικό Πολυώνυμο AR(p) και Στασιμότητα

Άσκηση Εμπέδωσης

Να βρεθεί αν η AR(2) χρονοσειρά $X_t = 5 + 1.3\, X_{t-1} – 0.4\, X_{t-2} + w_t$, είναι στάσιμη.

$$z^2 - 1{,}3z + 0{,}4 = 0 \qquad \Delta = (-1{,}3)^2 - 4 \cdot 1 \cdot 0{,}4 = 0{,}09$$

$$z_{1,2} = \frac{1{,}3 \pm 0{,}3}{2} \begin{cases} 0{,}8 \\ 0{,}5 \end{cases}$$

$$|z_{1,2}| < 1 \Rightarrow \text{Στάσιμη}$$

Αντιστρεψιμότητα ενός AR(p) μοντέλου

Στις χρονοσειρές είναι σημαντικό να μπορούμε να εκφράσουμε τα υπόλοιπα του μοντέλου από τις τιμές της ίδιας της σειράς. Στην περίπτωση αυτή το μοντέλο ονομάζεται **αντιστρέψιμο**.

Παρατηρούμε, πως στην περίπτωση ενός AR(p) μοντέλου, αυτό είναι πάντα εφικτό:

$$X_t = \delta + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \varepsilon_t \Leftrightarrow$$

$$\varepsilon_t = X_t - \delta - \alpha_1 X_{t-1} - \alpha_2 X_{t-2} - \dots - \alpha_p X_{t-p}.$$

Δηλαδή:

Ένα AR(p) μοντέλο είναι πάντα αντιστρέψιμο.

Χαρακτηριστικά αυτοπαλινδρομικού μοντέλου 1ης τάξης (AR(1))

Αυτοπαλινδρομικό μοντέλο 1ης τάξης (AR(1))

ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΡΑΚΗΣ DEMOCRITUS UNIVERSITY OF THRACE

Ένα από τα απλούστερα μοντέλα είναι το αυτοπαλινδρομικό μοντέλο τάξης 1, AR(1) (Autoregressive model of order 1) με το οποίο συσχετίζεται κάθε μία τιμή της σειράς

$\{X_t, t \in N^*\}$ με την αμέσως προηγούμενη της. Η αναλυτική του έκφραση είναι

$$\underline{X_t = \delta + \varphi_1 \cdot X_{t-1} + w_t,} \rightarrow \text{WN}$$

όπου w_t είναι τα ανεξάρτητα και ισόνομα σφάλματα των προβλέψεων με κανονική κατανομή $N(0, \sigma^2)$ τα οποία επιπλέον, δεν εξαρτώνται από την αντίστοιχη τιμή της σειράς X_t.

Μοντέλο AR(1) για μία ασθενώς στάσιμη χρονοσειρά

Υποθέτουμε ότι η χρονοσειρά $\{X_t,\ t \in N^*\}$ αποτελεί μία ασθενώς στάσιμη στοχαστική διεργασία, δηλαδή ότι, για κάθε $t \in N^*$:

- $E(X_t) = \mu =$ σταθερό,
- $Var(X_t) = \sigma^2 =$ σταθερό,
- Η συνδιακύμανση $Cov(X_t, X_{t-h})$ εξαρτάται μόνο από το $h = 1, 2, \ldots$και όχι από το t (ομοιογένεια συνδιακύμανσης στο χρόνο).

Στην περίπτωση αυτή, η συνάρτηση αυτοσυσχέτισης για τη χρονοσειρά γράφεται

$$\rho_h = Corr(X_t, X_{t-h}) = Cov(X_t, X_{t-h}) \,/\, [\sigma(X_t)\,\sigma(X_{t-h})]$$

$$= Cov(X_t, X_{t-h}) \,/\, Var(X_t), \qquad (\sigma(X_t) = \sigma(X_{t-h}) = Var(X_t))$$

δηλαδή, μία ασθενώς στάσιμη στοχαστική διεργασία αντιστοιχεί σε χρονοσειρά που για δεδομένη υστέρηση h, έχει σταθερή αυτοσυσχέτιση σε όλο το εύρος του χρονικού διαστήματος που περιγράφει.

Μοντέλο AR(1) για μία ασθενώς στάσιμη χρονοσειρά

Αναμενόμενη τιμή ασθενούς στάσιμης χρονοσειράς AR(1)

Από την εξίσωση $X_t = \delta + \varphi_1 \cdot X_{t-1} + w_t$, παίρνουμε

$E(X_t) = E(\delta + \varphi_1 \cdot X_{t-1} + w_t)$

$= \delta + E(\varphi_1 \cdot X_{t-1}) + E(w_t)$

$= \delta + \varphi_1 \cdot E(X_{t-1})$ $\quad (E(w_t) = 0)$

Όμως, $E(X_t) = E(X_{t-1}) = \mu$, άρα

$$\mu = \delta + \varphi_1 \cdot \mu$$

ή

$$\mathbf{E(X_t) = \mu = \delta / (1 - \varphi_1)}$$

Μοντέλο AR(1) για μία ασθενώς στάσιμη χρονοσειρά

Διακύμανση ασθενούς στάσιμης χρονοσειράς AR(1)

Από την εξίσωση $X_t = \delta + \varphi_1 \cdot X_{t-1} + w_t$, παίρνουμε

$w_t \sim N(0, \sigma^2)$

$Var(X_t) = Var(\delta + \varphi_1 \cdot X_{t-1} + w_t)$

$= Var(\varphi_1 \cdot X_{t-1}) + Var(w_t)$ (τα σφάλματα w_t είναι ανεξάρτητα από τις τιμές X_{t-1})

$= \varphi_1^2 \cdot Var(X_{t-1}) + \sigma^2$

Όμως, $Var(X_t) = Var(X_{t-1})$, άρα

$$Var(X_t) = \varphi_1^2 \cdot Var(X_t) + \sigma^2$$

ή

$$\mathbf{Var(X_t) = \sigma^2 / (1 - \varphi_1^2)}$$

Μοντέλο AR(1) για μία ασθενώς στάσιμη χρονοσειρά

Αυτοδιακύμανση και αυτοσυσχέτιση (υστέρησης h) ασθενούς στάσιμης χρονοσειράς AR(1)

Υποθέτουμε ότι δ = 0, δηλαδή ότι $E(X_t) = 0$.

(αν $\delta \neq 0$, τότε θεωρούμε τη $X'_t = X_t - \delta / (1 - \varphi_1)$).

Αν δ = 0, τότε $X_t = \varphi_1 \cdot X_{t-1} + w_t$ και $Cov(X_t, X_{t-h}) = E(X_t \cdot X_{t-h})$. Είναι:

$$X_t = \varphi_1 \cdot X_{t-1} + w_t \Leftrightarrow X_{t-h} \cdot X_t = \varphi_1 \cdot X_{t-h} \cdot X_{t-1} + X_{t-h} \cdot w_t \text{ ή}$$

$$E(X_{t-h} \cdot X_t) = E(\varphi_1 \cdot X_{t-h} \cdot X_{t-1}) + E(X_{t-h} \cdot w_t)$$

$$= \varphi_1 \cdot E(X_{t-h} \cdot X_{t-1})$$

($E(X_{t-h} \cdot w_t) = 0$, γιατί τα σφάλματα w_t είναι ανεξάρτητα από τις X_{t-h} και $E(X_{t-h}) = 0$)

Αν $\gamma_h = E(X_{t-h} \cdot X_t)$ η παραπάνω σχέση γράφεται

$$\boldsymbol{\gamma_h = \varphi_1 \cdot \gamma_{h-1}}$$

Μοντέλο AR(1) για μία ασθενώς στάσιμη χρονοσειρά

Παρατηρούμε ότι $\gamma_1 = E(X_{t-1} \cdot X_t)$

$$= E(X_{t-1} \cdot (\varphi_1 \cdot X_{t-1} + w_t))$$

$$= E(\varphi_1 \cdot X^2_{t-1} + X_{t-1} \cdot w_t)$$

$$= \varphi_1 \cdot E(X^2_{t-1}) = \varphi_1 \cdot Var(X_{t-1})$$

$$= \varphi_1 \cdot Var(X_t).$$

Από τις σχέσεις $\gamma_h = \varphi_1 \cdot \gamma_{h-1}$ και $\gamma_1 = \varphi_1 \cdot Var(X_t)$ με αναδρομικό συλλογισμό βρίσκουμε ότι

$$\mathbf{E(X_{t-h} \cdot X_t) = \gamma_h = \varphi^h_1 \cdot Var(X_t)}$$

Μοντέλο AR(1) για μία ασθενώς στάσιμη χρονοσειρά

Τώρα,

$$\rho_h = Cov(X_t, X_{t-h}) / Var(X_t)$$

$$= E(X_t \cdot X_{t-h}) / Var(X_t)$$

$$= \varphi^h{}_1 \cdot Var(X_t) / Var(X_t)$$

$$= \varphi^h{}_1$$

Βρήκαμε ότι,

AR(1)

$$\underline{\boldsymbol{\rho_h = \varphi^h{}_1}}$$

$\rho_1 = \varphi_1$

$\rho_2 = \varphi_1^2$

$\rho_3 = \varphi_1^3$

Μοντέλο AR(1) για μία ασθενώς στάσιμη χρονοσειρά

Σύνοψη

Σε κάθε ασθενώς στάσιμο AR(1) μοντέλο χρονοσειράς $X_t = \delta + \varphi_1 \cdot X_{t-1} + w_t$, ισχύει ότι:

$E(X_t) = \mu = \delta / (1 - \varphi_1)$.

$Var(X_t) = \sigma^2 / (1 - \varphi^2{}_1)$.

$Cov(X_t, X_{t-h}) = \rho_h = \varphi^h{}_1$.

Από την ισότητα $\rho_h = \varphi^h{}_1$ προκύπτει ότι $\rho_2 = \rho^2{}_1$, $\rho_3 = \rho^3{}_1$, ..., σχέσεις που προσφέρουν μία απλή μέθοδο αξιολόγησης μίας χρονοσειράς ως προς τη δυνατότητα προσαρμογής ενός AR(1) μοντέλου σε αυτήν.

Μοντέλο AR(1) για μία ασθενώς στάσιμη χρονοσειρά

Δραστηριότητα Ι

Να βρεθούν τα $E(X_t)$, $Var(X_t)$, ρ_h της AR(1) χρονοσειράς $X_t = 0.3 + 0.6 X_{t-1} + \varepsilon_t$, $\varepsilon_t \sim N(0, 1)$

Μοντέλο AR(1) για μία ασθενώς στάσιμη χρονοσειρά

Δραστηριότητα ΙΙ

Να βρεθούν τα $E(X_t)$, $Var(X_t)$, ρ_h της AR(1) χρονοσειράς $X_t = 0.3 – 0.6 X_{t-1} + \varepsilon_t$, $\varepsilon_t \sim N(0, 1)$

Μοντέλο AR(1) για μία ασθενώς στάσιμη χρονοσειρά

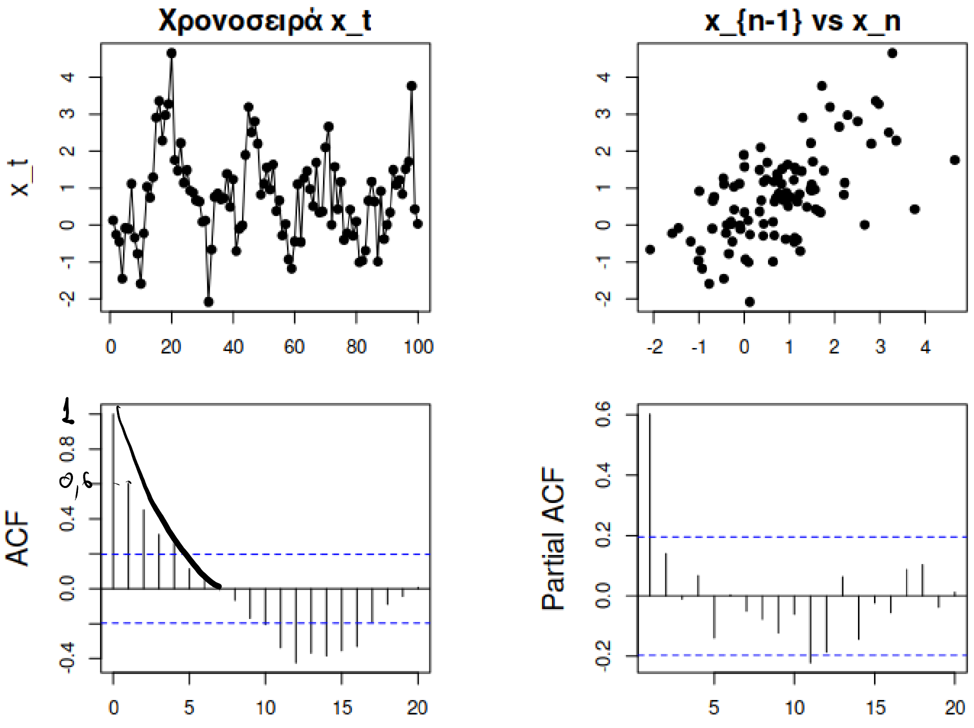
$\rho_h = \varphi_1^h$

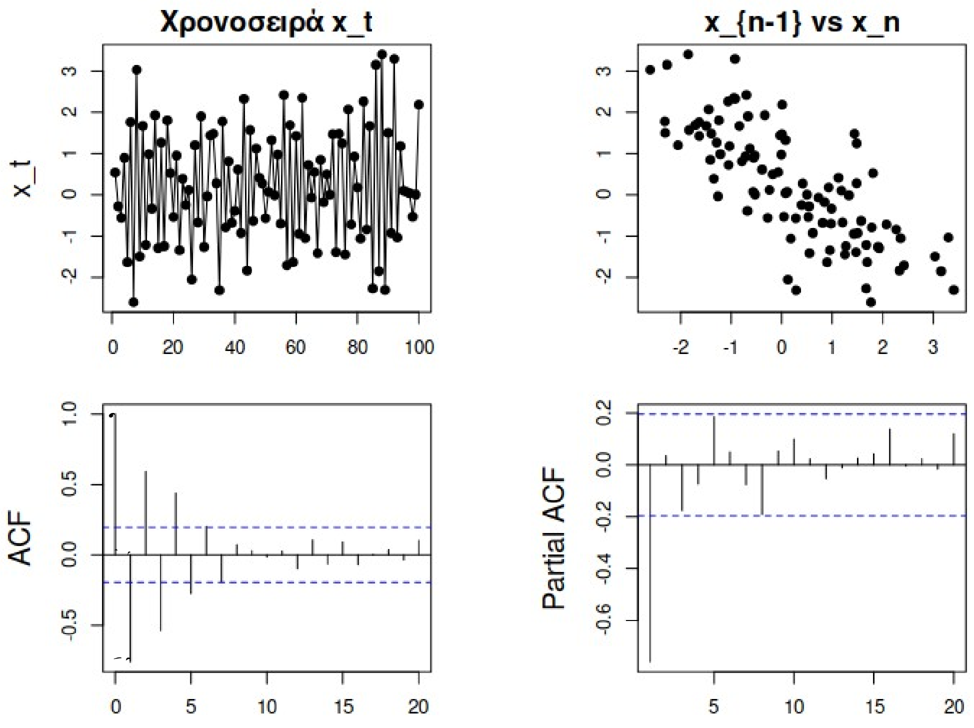
δ $\quad \varphi_1 = 0{,}6 \quad \rho_h = 0{,}6^h$

$X_t = 0.3 + 0.6\, X_{t-1} + \varepsilon_t,\ \varepsilon_t \sim N(0, 1)$

$\varphi_1 = -0{,}6 \quad \rho_h = (-0{,}6)^h$

$X_t = 0.3 - 0.6\, X_{t-1} + \varepsilon_t,\ \varepsilon_t \sim N(0, 1)$





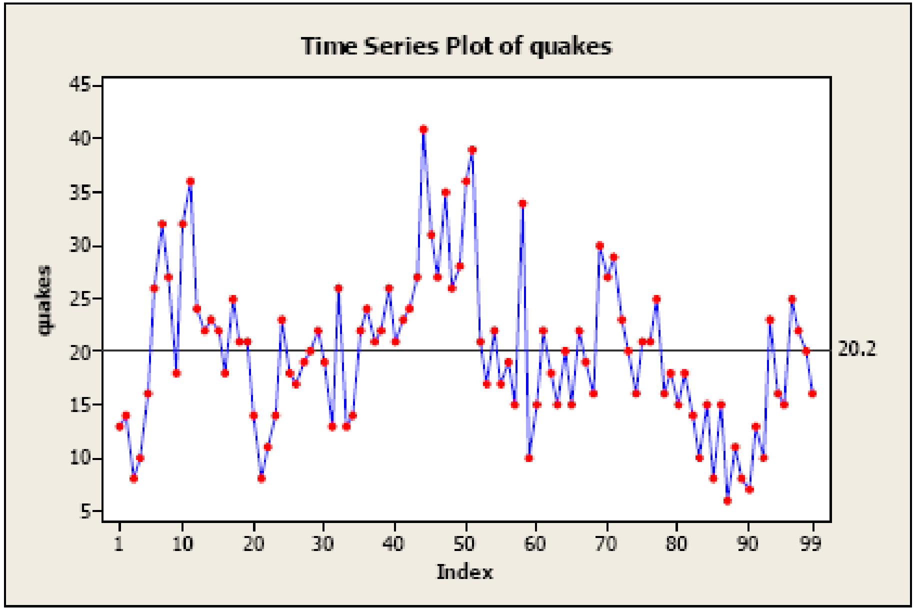
$E(X_t) = \delta / (1 - \varphi_1) = 0.75$. $Var(X_t) = \sigma^2 / (1 - \varphi^2_1) = 1.56$. $\rho_h = 0.6^h$.

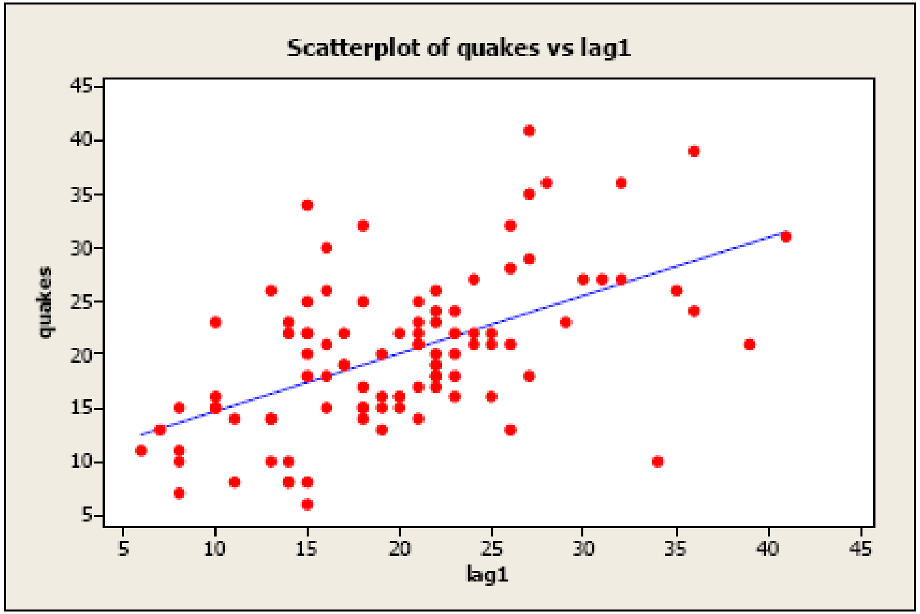
$E(X_t) = \delta / (1 - \varphi_1) = 0.188$. $Var(X_t) = \sigma^2 / (1 - \varphi^2_1) = 1.56$. $\rho_h = (-0.6)^h$.

Παράδειγμα Ι

Δεδομένα: Πλήθος σεισμών ανά έτος σε μία περιοχή. Το διάγραμμα διασποράς της χρονοσειράς με την lag(1) εκδοχή της υποδεικνύει γραμμικότητα στη σχέση μεταξύ x_t, x_{t-1}, υποδεικνύοντας ως κατάλληλο ένα AR(1) μοντέλο. Βρίσκουμε:

$$\mathbf{x_t = 9{,}19 + 0{,}543 \cdot x_{t-1}.}$$

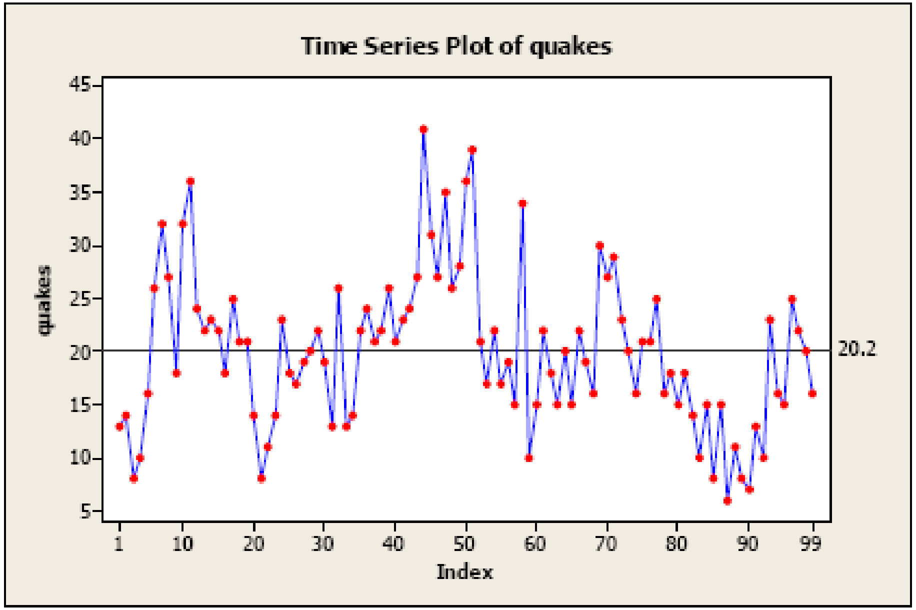


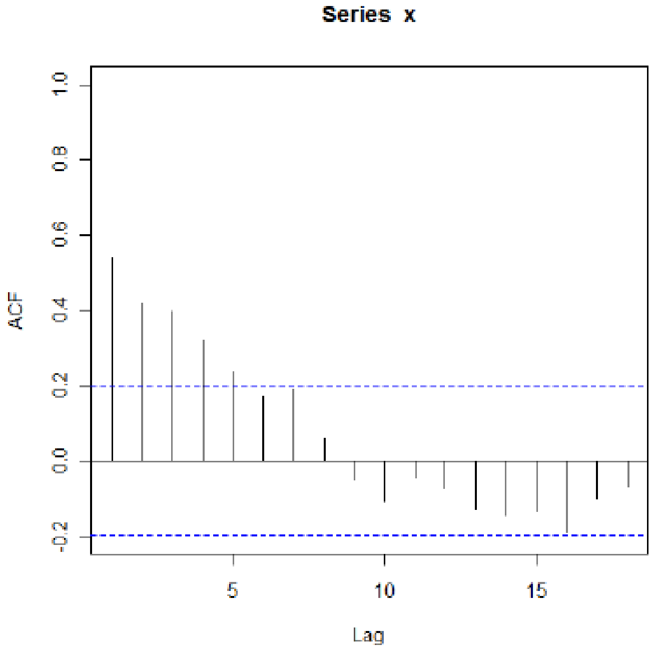


Παράδειγμα Ι

Δεδομένα: Πλήθος σεισμών ανά έτος σε μία περιοχή. Το διάγραμμα διασποράς της χρονοσειράς με την lag(1) εκδοχή της υποδεικνύει γραμμικότητα στη σχέση μεταξύ x_t, x_{t-1}, υποδεικνύοντας ως κατάλληλο ένα AR(1) μοντέλο. Βρίσκουμε:

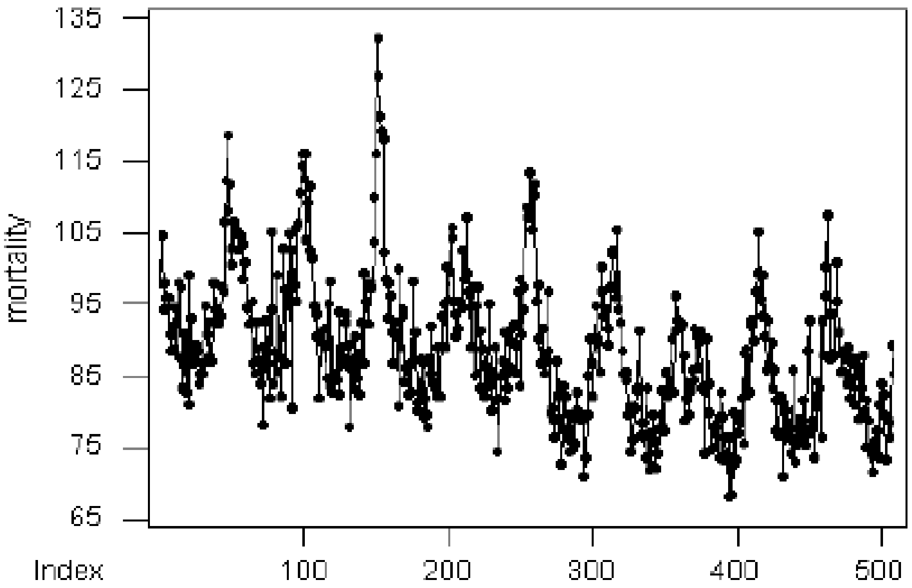
$$\mathbf{x_t = 9{,}19 + 0{,}543 \cdot x_{t-1}}.$$

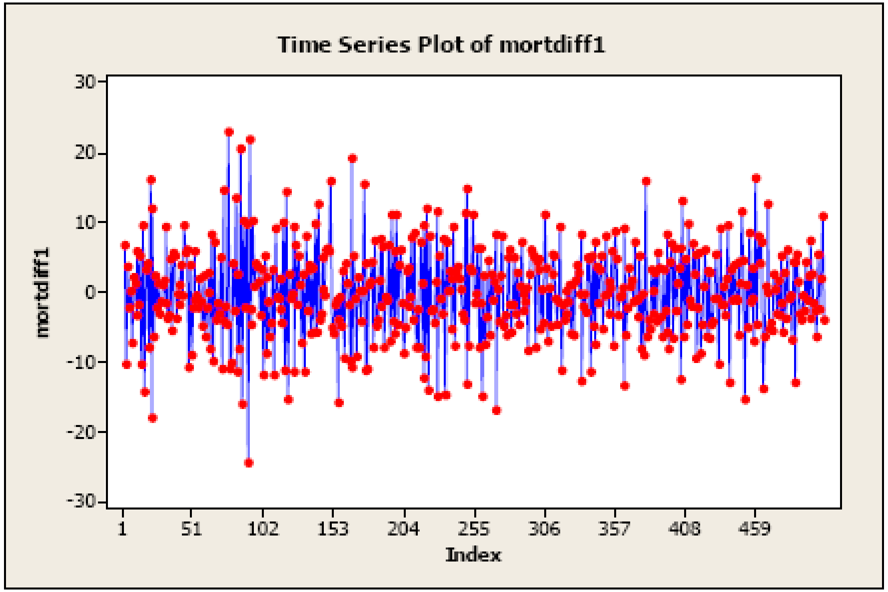




Παράδειγμα ΙΙ

Δεδομένα: Θάνατοι ανά ημέρα από καρδιακό επεισόδιο στο LA για 500 ημέρες. Η χρονοσειρά έχει αρνητική κλίση, άρα δεν είναι στάσιμη. αν αντί των ίδιων των τιμών $\{x_1, x_2, x_3, \ldots\}$ πάρουμε τις διαφορές πρώτης τάξης $\{z_1, z_2, z_3, \ldots\} = \{x_2 - x_1, x_3 - x_2, x_4 - x_3, \ldots\}$ η χρονοσειρά $\{z_1, z_2, z_3, \ldots\}$ που προκύπτει έχει σταθερή τάση.

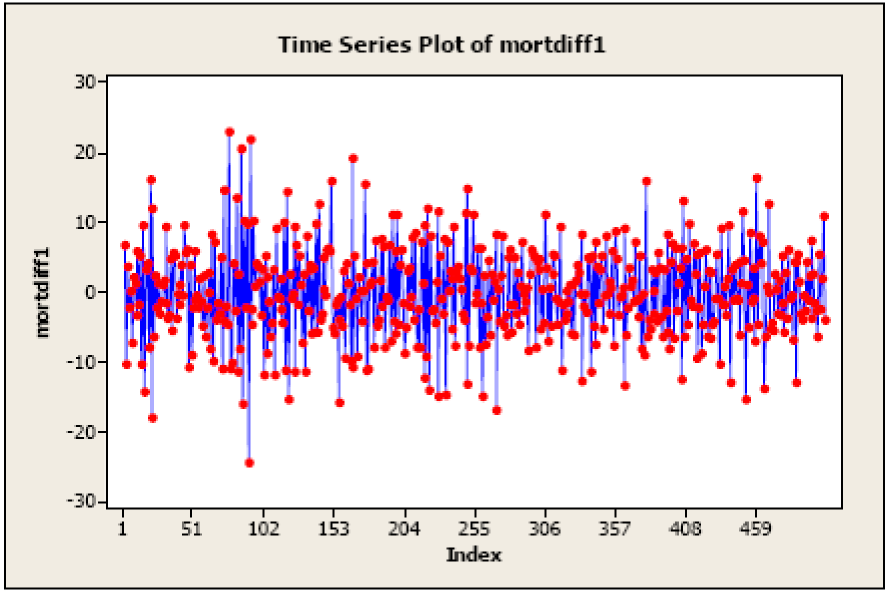


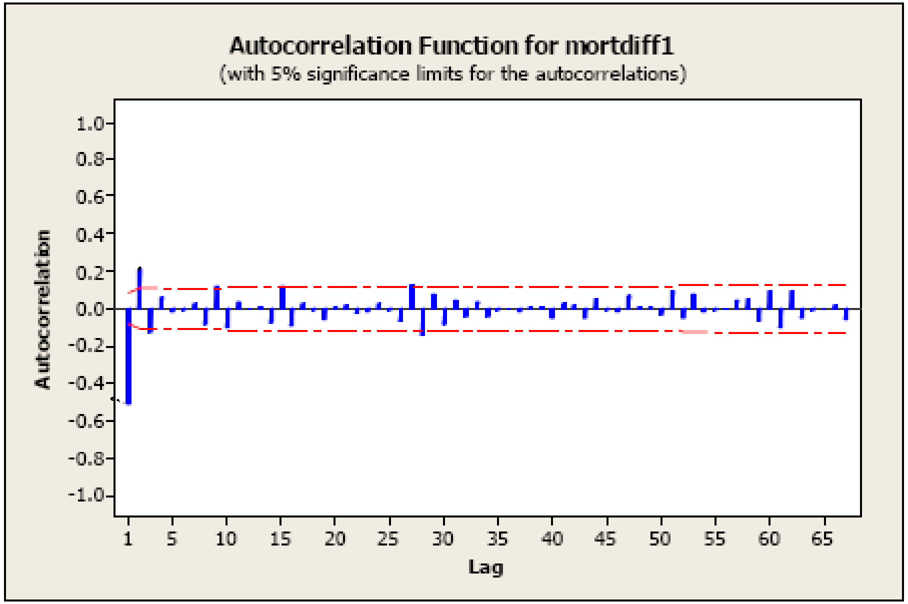


Παράδειγμα ΙΙ

Η καταλληλότητα του AR(1) μοντέλου για την ακολουθία των διαφορών αναδεικνύεται και από τις τιμές της αυτοσυσχέτισης οι οποίες στις πρώτες τους τιμές ικανοποιούν τις σχέσεις $\rho_2 = \rho^2{}_1$, $\rho_3 = \rho^3{}_1$, $\rho_4 = \rho^4{}_1$. Βρίσκουμε: $\mathbf{z_t = -0,046 - 0,506\, z_{t-1}}$.

$x_t - x_{t-1}$ $\quad (x_{t-1} - x_{t-2})$ $\quad r_h = \varphi_1^h$ $\quad \varphi_1 < 0$





Χαρακτηριστικά αυτοπαλινδρομικού μοντέλου 2ης τάξης (AR(2))

Μοντέλο AR(2) για μία ασθενώς στάσιμη χρονοσειρά

Με το αυτοπαλινδρομικό μοντέλο τάξης 2, AR(2) (Autoregressive model of order 2) συσχετίζεται κάθε μία τιμή της σειράς $\{X_t, t \in N^*\}$ με την αμέσως προηγούμενη και την παρατήρηση που βρίσκεται δύο θέσεις πριν.

Η αναλυτική του έκφραση είναι

$$X_t = \delta + \varphi_1 \cdot X_{t-1} + \varphi_2 \cdot X_{t-2} + w_t,\ w_t \sim N(0, \sigma^2),$$

όπου $w_t \sim N(0, \sigma^2)$, είναι τα ανεξάρτητα και ισόνομα σφάλματα των προβλέψεων με κανονική κατανομή $N(0, \sigma^2)$, τα οποία επιπλέον, δεν εξαρτώνται από την αντίστοιχη τιμή της σειράς x_t.

Μοντέλο AR(2) για μία ασθενώς στάσιμη χρονοσειρά

ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΡΑΚΗΣ | DEMOCRITUS UNIVERSITY OF THRACE

Αποδεικνύεται ότι, $Var(X_t) = \sigma^2_W / (1 - \rho_1 \cdot \varphi_1 - \rho_2 \cdot \varphi_2)$, ή

$$\mathbf{Var(X_t) = (1 - \varphi_2)\sigma^2_W / [(1 + \varphi_2)(1 - \varphi_1 - \varphi_2)(1 + \varphi_1 - \varphi_2)]}$$

Επιπλέον, οι αυτοσυσχετίσεις με υστέρηση 1, 2 και 3 αντίστοιχα δίνονται από τους τύπους

$\rho_1 = \varphi_1 / (1 - \varphi_2)$.

$\rho_2 = \varphi^2_1 / (1 - \varphi_2) + \varphi_2$.

$\rho_3 = (\varphi^3_1 + \varphi_1 \cdot \varphi_2) / (1 - \varphi_2) + \varphi_1 \cdot \varphi_2$.

Αντίστοιχα, ισχύει:

$\varphi_1 = \rho_1(1 - \rho_2) / (1 - \rho^2_1)$,

$\varphi_2 = (\rho_2 - \rho^2_1) / (1 - \rho^2_1)$.

Μοντέλο AR(2) για μία ασθενώς στάσιμη χρονοσειρά

Παράδειγμα

Να βρεθούν τα $E(X_t)$, $Var(X_t)$, ρ_1, ρ_2, ρ_3, της χρονοσειράς $X_t = 0.1\,X_{t-1} + 0.5\,X_{t-2} + \varepsilon_t$, $\varepsilon_t \sim N(0, 1)$

Λύση

$E(X_t) = 0$

$Var(X_t) = (1 - \varphi_2)\sigma^2 / [(1 + \varphi_2)(1 - \varphi_1 - \varphi_2)(1 + \varphi_1 - \varphi_2)] = 0.5/(1.5\ 0.4\ 0.6) = 1.39$

$\rho_1 = \varphi_1 / (1 - \varphi_2) = 0.1/0.5 = 0.2.$

$\rho_2 = \varphi^2{}_1 / (1 - \varphi_2) + \varphi_2 = 0.1^2 / (1 - 0.5) + 0.5 = 0.52$

$\rho_3 = (\varphi^3{}_1 + \varphi_1 \cdot \varphi_2) / (1 - \varphi_2) + \varphi_1 \cdot \varphi_2 = (0.1^3 + 0.1\ 0.5) / (1 - 0.5) + 0.1\ 0.5 = 0.152$

Παράδειγμα

$X_t = 0.1 \cdot X_{t-1} + 0.5 \cdot X_{t-2} + \varepsilon_t,\ \varepsilon_t \sim N(0, 1)$

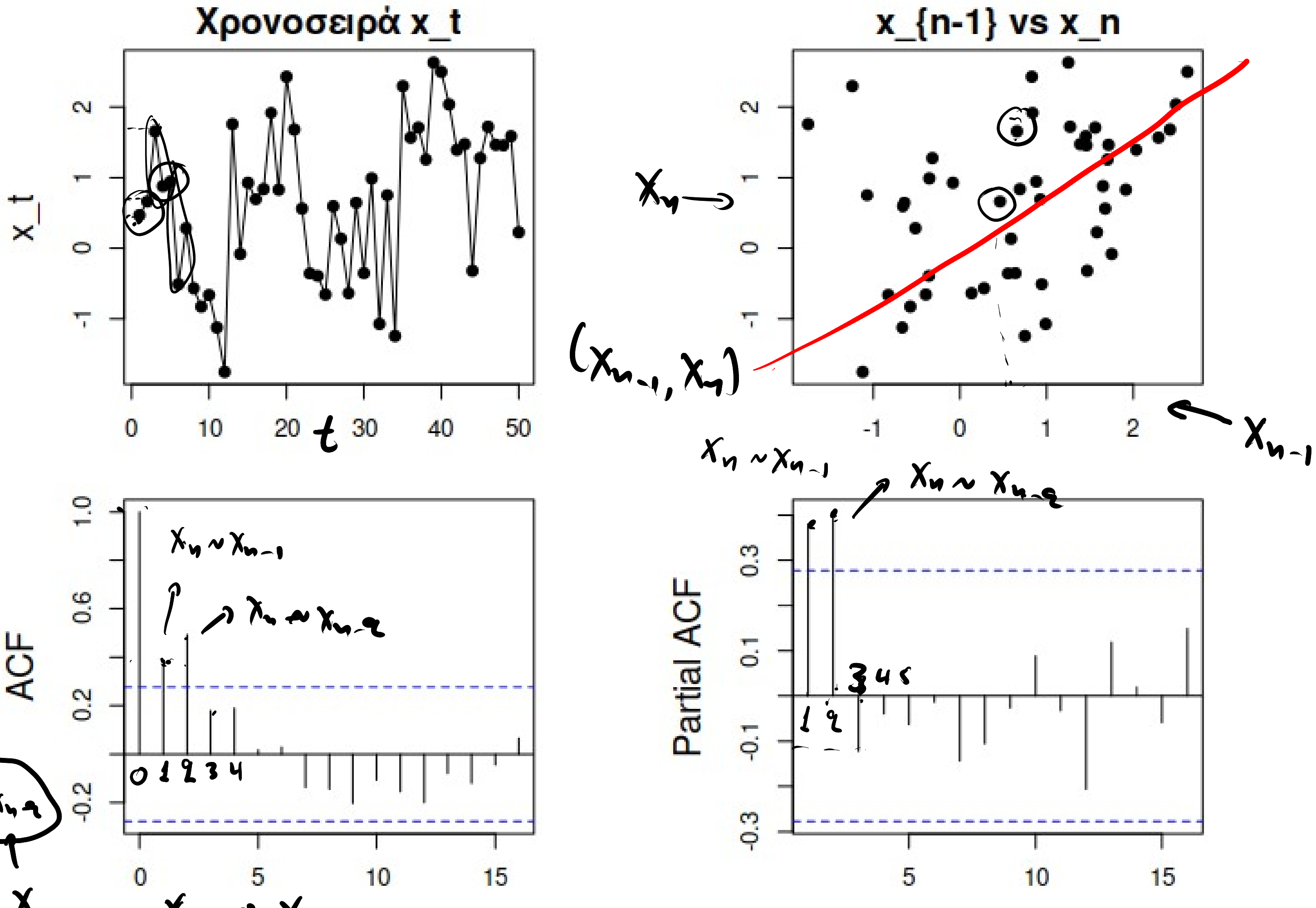
$E(X_t) = 0$

$Var(X_t) = 1.39$

$\rho_1 = 0.2$

$\rho_2 = 0.52$

$\rho_3 = 0.152$



X_{n-1}

$X_n \xrightarrow{sign} X_{n-2}$

$\underbrace{X_n = a \cdot X_{n-2}} + b$

$X_n = a \cdot X_{n-1} + b \cdot X_{n-2} + \gamma.$

(X_{n-1}, X_{n-2})

$X_n \sim X_{n-3}$

$X_n = a \cdot X_{n-3}$

Μοντέλο AR(2) για μία ασθενώς στάσιμη χρονοσειρά

Άσκηση Εμπέδωσης

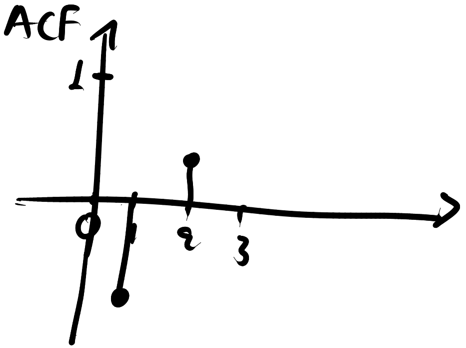
Να βρεθούν τα $E(X_t)$, $Var(X_t)$, ρ_1, ρ_2, ρ_3, της χρονοσειράς $X_t = -0.3\,X_{t-1} + 0.2\,X_{t-2} + \varepsilon_t$, $\varepsilon_t \sim N(0, 1)$

$a_1 = -0{,}3 \quad , \quad a_2 = 0{,}2$

$E(X_t) = 0$

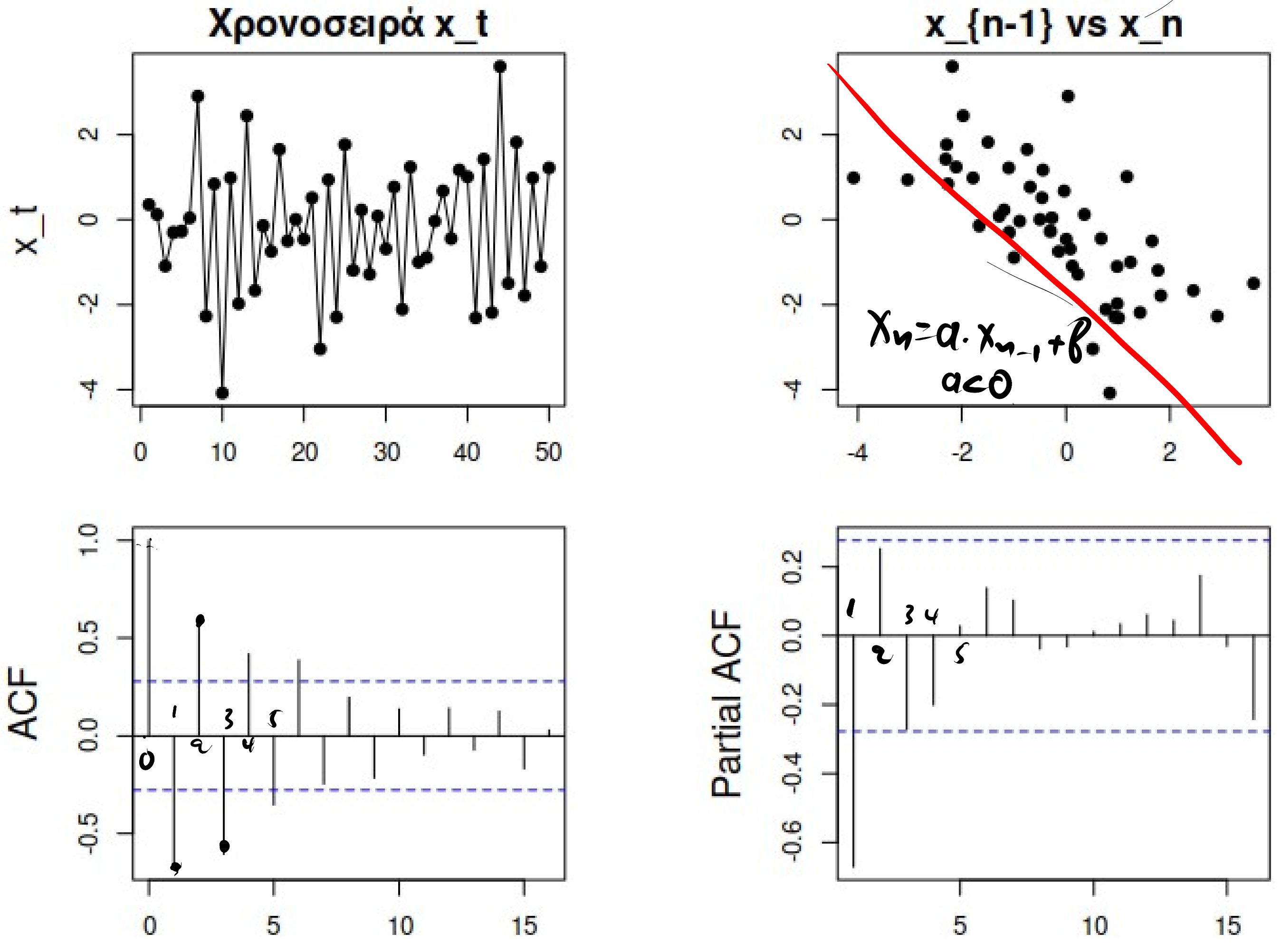
$$\rho_1 = \frac{a_1}{1-a_2} = \frac{-0{,}3}{1-0{,}2} = -\frac{3}{8} < 0$$

$$\rho_2 = \frac{a_1^2}{1-a_2} + a_2 = \frac{0{,}09}{0{,}8} + 0{,}2 > 0$$



Άσκηση Εμπέδωσης

$X_t = -0.3 \cdot X_{t-1} + 0.2 \cdot X_{t-2} + \varepsilon_t,\ \varepsilon_t \sim N(0, 1)$



$E(X_t) =$

$Var(X_t) =$

$\rho_1 =$

$\rho_2 =$

$\rho_3 =$

$\rho_h = \varphi_1^h$

Μοντέλο κινούμενου μέσου όρου (Moving Average)

Γέννεση μοντέλων Moving Average

Παράδειγμα Ι

Υποθέτουμε ότι κάθε χρονική περίοδο σκοπεύουμε να παράγουμε **m** μονάδες από κάποιο αγαθό. Ωστόσο, αναμένονται διάφορα i.i.d. τυχαία σφάλματα $\boldsymbol{\varepsilon_t}$ με $E(\varepsilon_t) = 0$, τα οποία θα αυξομειώνουν την παραγωγή. Άρα, η παραγωγή $\boldsymbol{x_t}$ τη χρονική στιγμή t είναι $\boldsymbol{x_t = m + \varepsilon_t}$**.** Έστω τώρα θ є (0, 1). Κάθε περίοδο σκοπεύουμε να διαθέτουμε το **1 – θ** μέρος της παραγωγής και να αποθηκεύουμε το μέρος **θ** για να το πουλήσουμε την επόμενη περίοδο. Θέλουμε να προσδιορίσουμε το απόθεμα $\boldsymbol{y_t}$ που υπάρχει στην αποθήκη τη στιγμή t.

Είναι $\boldsymbol{y_t}$ **= (παραγωγή της περιόδου t) + (θ μέρος παραγωγής περιόδου t – 1).**

Γράφουμε: $y_t = x_t + \theta y_{t-1} = (m + \varepsilon_t) + \theta(m + \varepsilon_{t-1}) = (\theta + 1)m + \theta\varepsilon_{t-1} + \varepsilon_t$

Δηλαδή:

$$\boldsymbol{y_t = \mu + \theta\varepsilon_{t-1} + \varepsilon_t,}$$

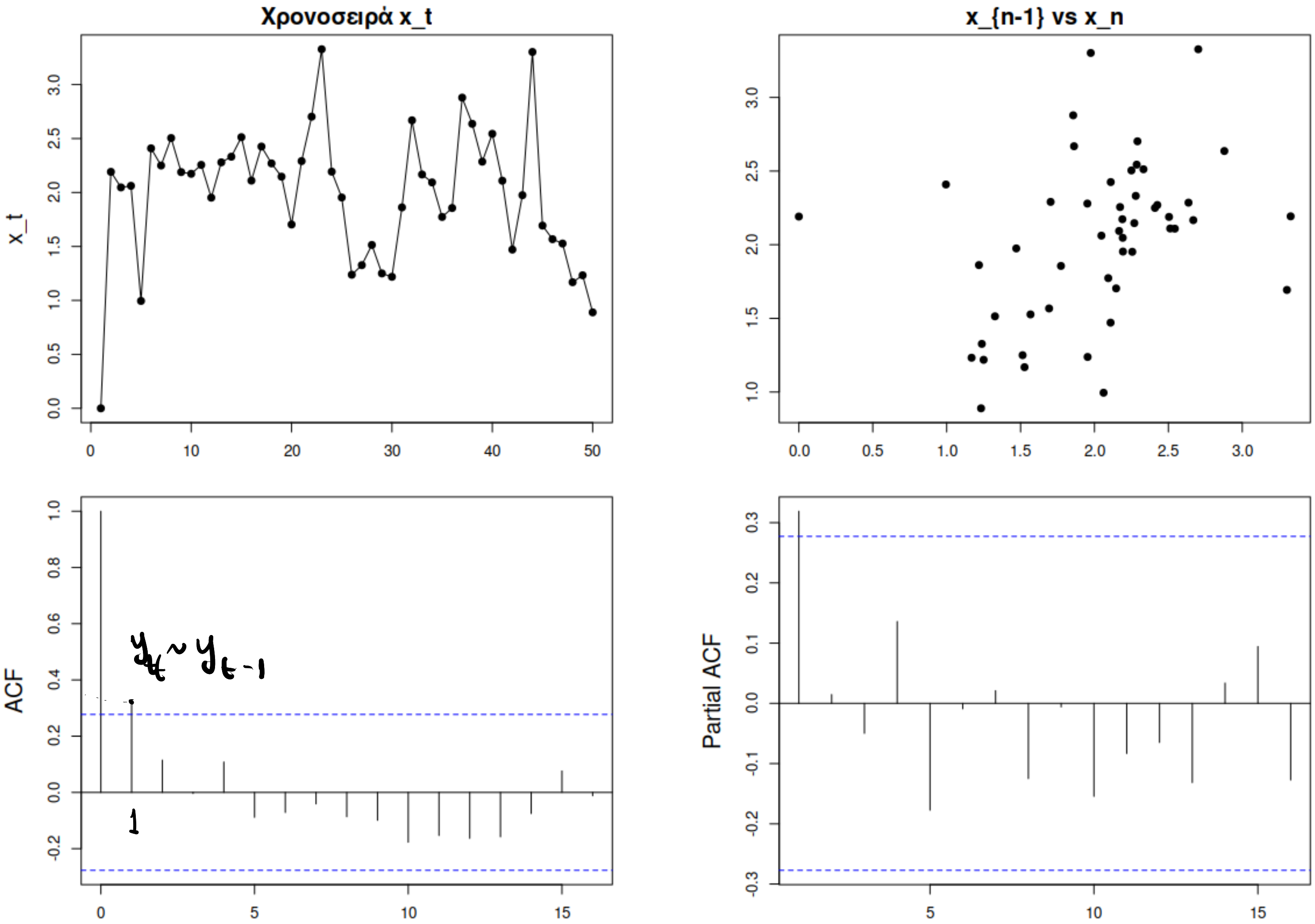
όπου μ = (θ + 1)m.

Γέννεση μοντέλων Moving Average

$Y_t = 2 + 0.3\varepsilon_{t-1} + \varepsilon_t$

$\varepsilon_t \sim N(0, 0.5^2)$

N = 50



Γέννεση μοντέλων Moving Average

$Y_t = 2 + 0.6\varepsilon_{t-1} + \varepsilon_t$

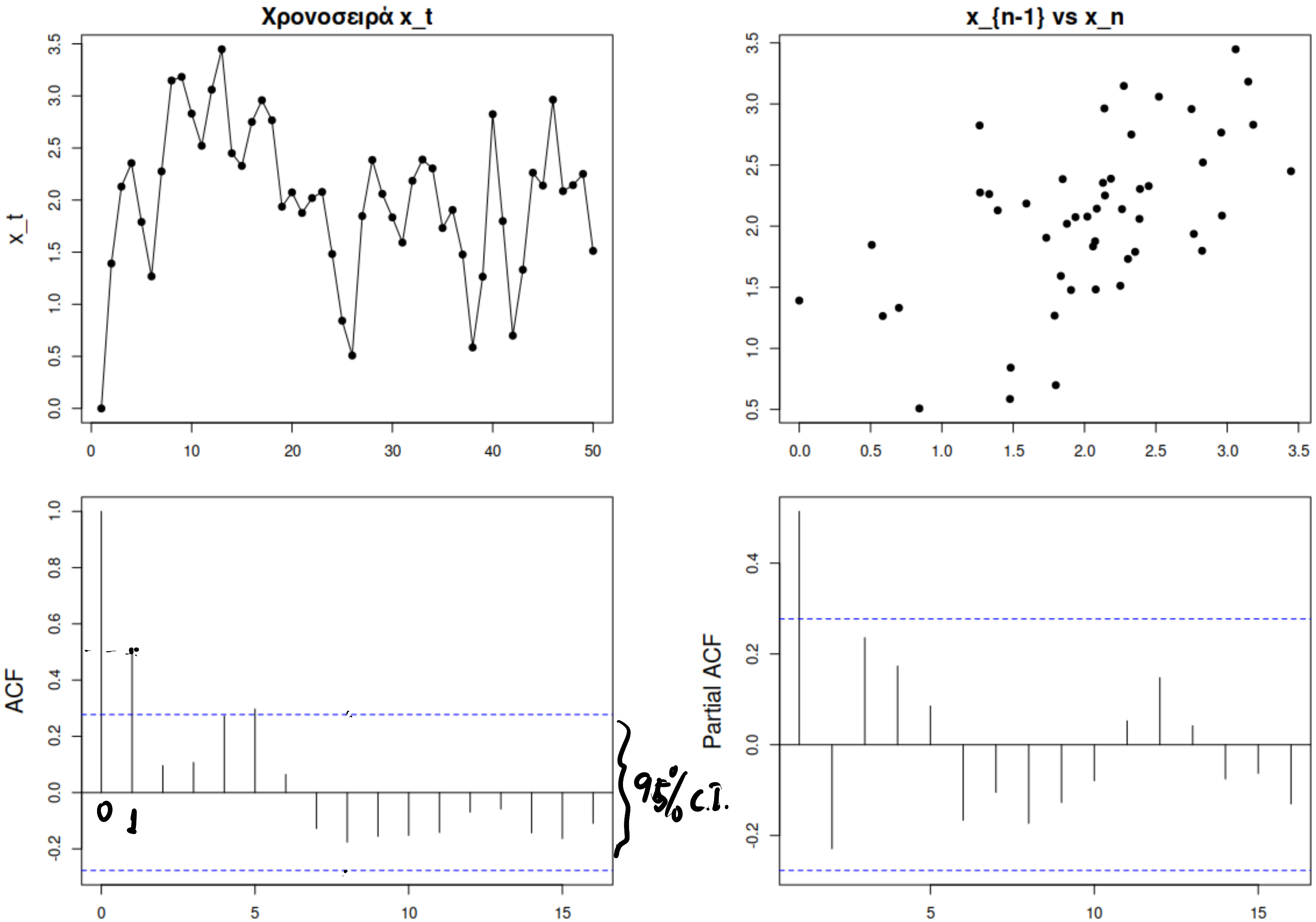
$\varepsilon_t \sim N(0, 0.5^2)$

$N = 50$

ΠΛΗΘΥΣΜΟΣ ρ_1

N=50 Δείγμα

$r_1 \approx 0{,}5$

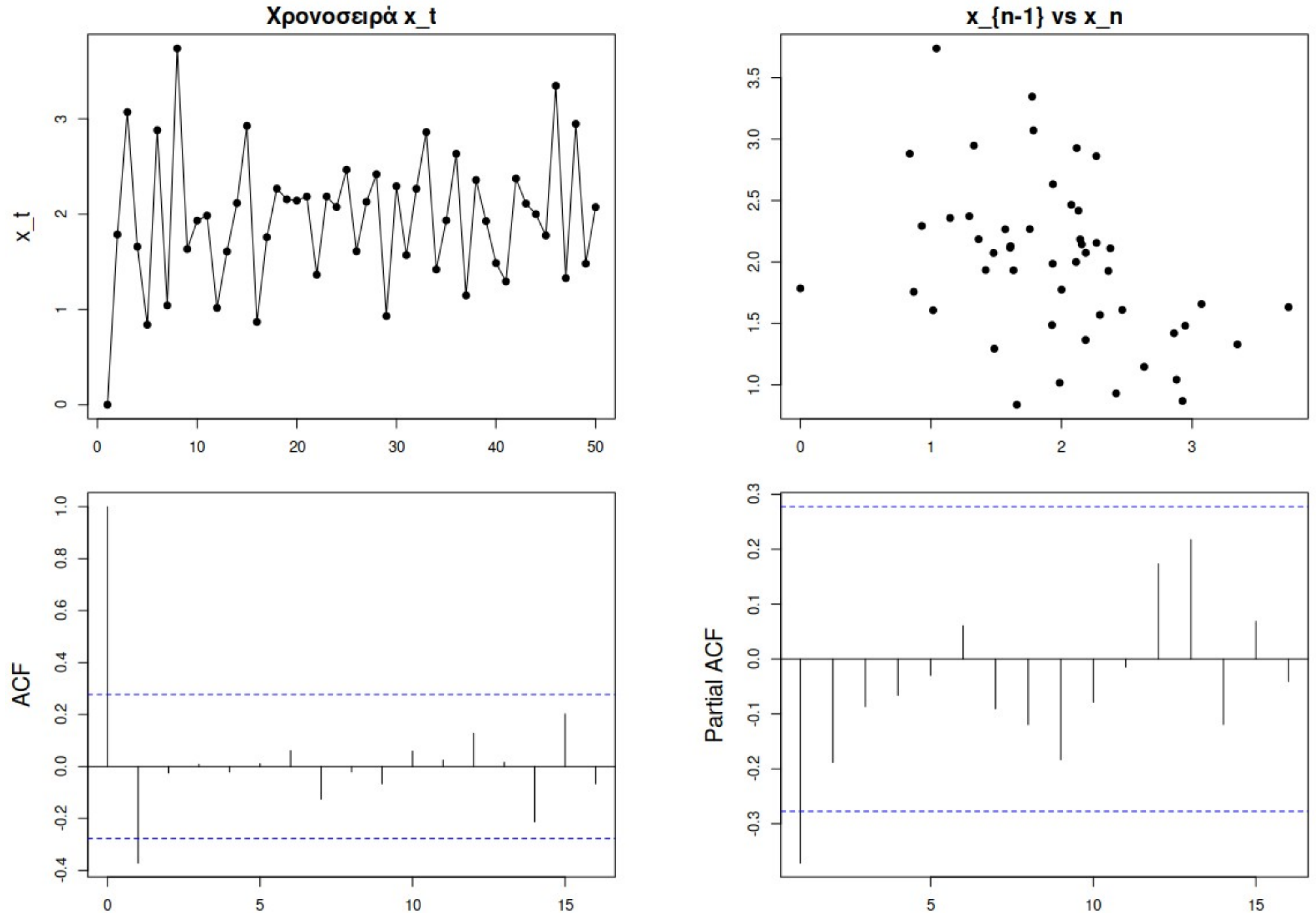


Γέννεση μοντέλων Moving Average

$Y_t = 2 - 0.4\varepsilon_{t-1} + \varepsilon_t$

$\varepsilon_t \sim N(0, 0.5^2)$

$N = 50$



Γέννεση μοντέλων Moving Average

Παράδειγμα II

Έστω ότι ο ίδιος παραγωγός σκοπεύει κάθε περίοδο να αποθηκεύει μέρος θ_1 της παραγωγής για να το διαθέσει την επόμενη περίοδο και μέρος θ_2 για να το διαθέσει τη μεθεπόμενη περίοδο.

Στην περίπτωση αυτή, το απόθεμα y_t, είναι

y_t = (παραγωγή της περιόδου t) + (θ_1 μέρος t – 1 παραγωγής) + (θ_2 μέρος t – 2 παραγωγής).

Γράφουμε: $y_t = x_t + \theta_1 y_{t-1} + \theta_2 y_{t-1} = (m + \varepsilon_t) + \theta_1(m + \varepsilon_{t-1}) + \theta_2(m + \varepsilon_{t-2})$

$$= (\theta_1 + \theta_2 + 1)m + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \varepsilon_t$$

Δηλαδή:

$$\boldsymbol{y_t = \mu + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \varepsilon_t,}$$

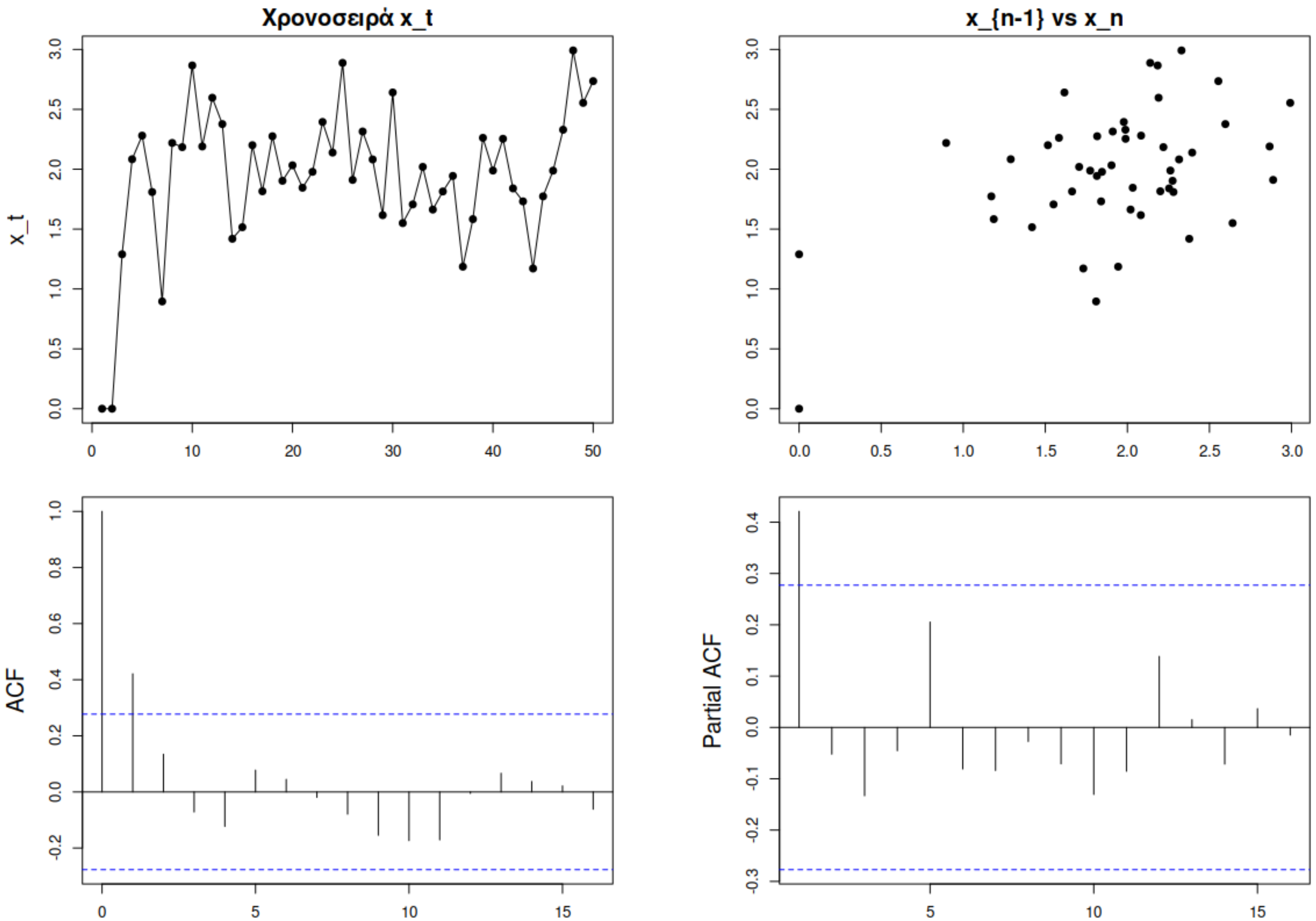
όπου $\mu = (\theta_1 + \theta_2 + 1)m$.

Γέννεση μοντέλων Moving Average

$Y_t = 2 + 0.2\varepsilon_{t-1} + 0.1\varepsilon_{t-2} + \varepsilon_t$

$\varepsilon_t \sim N(0, 0.5^2)$

$N = 50$

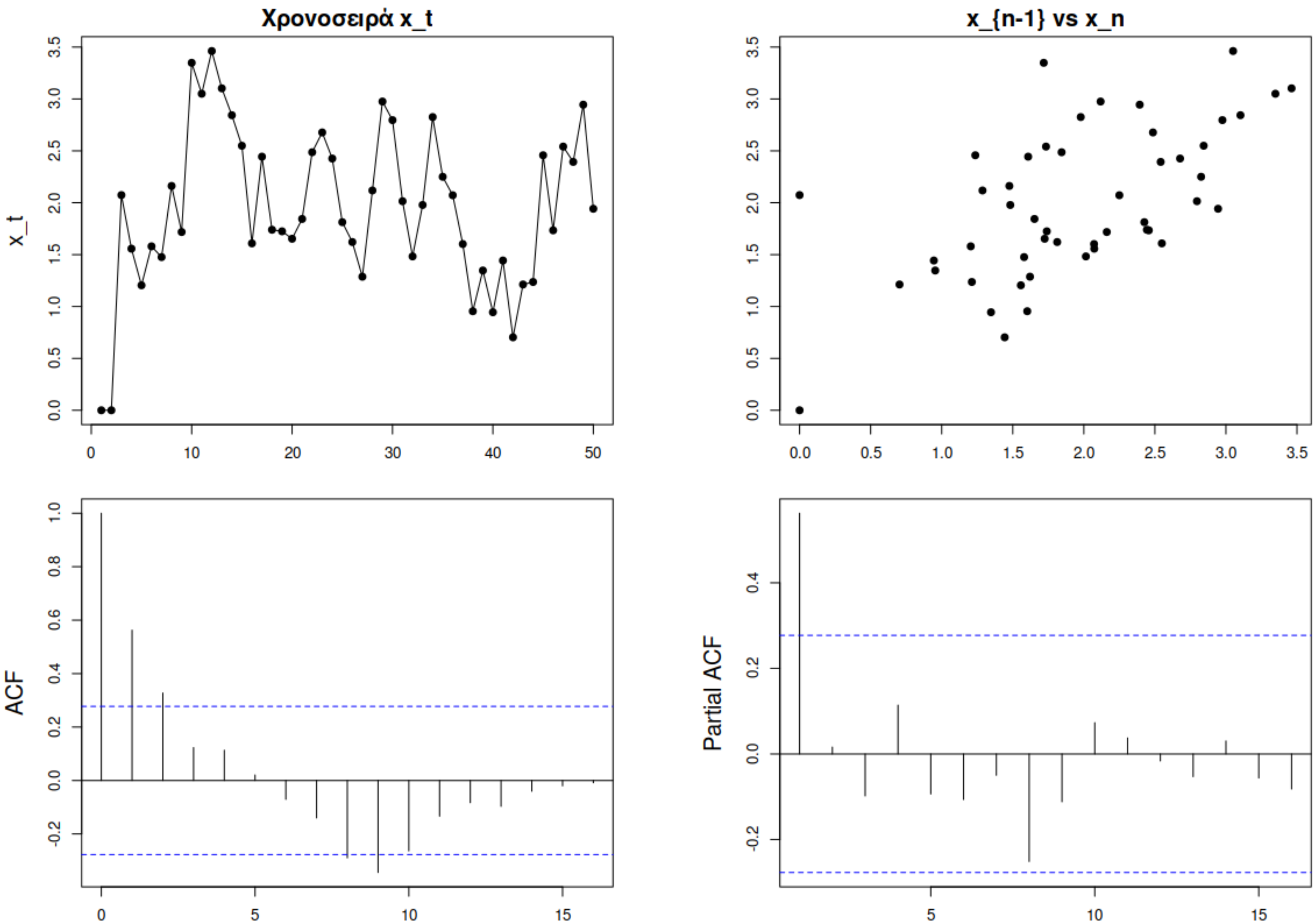


Γέννεση μοντέλων Moving Average

$Y_t = 2 + 0.6\varepsilon_{t-1} + 0.4\varepsilon_{t-2} + \varepsilon_t$

$\varepsilon_t \sim N(0, 0.5^2)$

$N = 50$

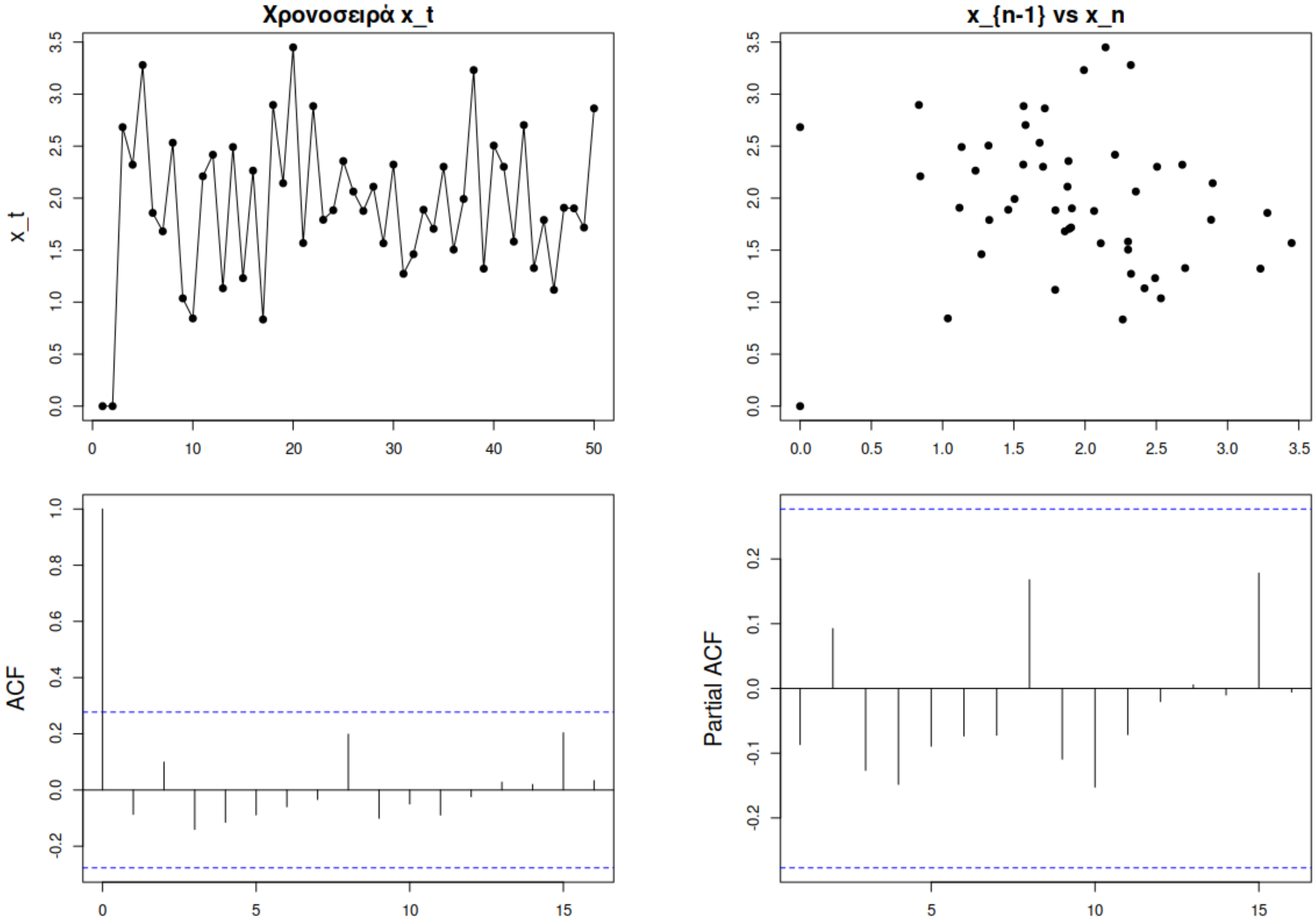


Γέννεση μοντέλων Moving Average

$Y_t = 2 - 0.2\varepsilon_{t-1} + 0.4\varepsilon_{t-2} + \varepsilon_t$

$\varepsilon_t \sim N(0, 0.5^2)$

$N = 50$



Γέννεση μοντέλων Moving Average

Κώδικας R

```
ma.sim = function(mu = 2, theta = 0.4, number = 100){
  ma_vals = rep(0, times = number)
  en = rnorm(number,0, mu/4)
  for(i in 2:number){ma_vals[i] = mu +  en[i] + theta * en[i-1] }
  return(ma_vals)
}
x=ma.sim(mu = 2, theta = -0.4, number=50)
my.plot.ts.short(x)

ma2.sim = function(mu = 2, theta1 = 0.2, theta2 = 0.1, number = 100){
  ma_vals = rep(0, times = number)
  en = rnorm(number,0, mu/4)
  for(i in 3:number){ma_vals[i] = mu +  en[i] + theta1 * en[i-1] + theta2 * en[i-2] }
  return(ma_vals)
}
x=ma2.sim(mu = 2, theta1 = -0.2, theta2 = 0.4, number=50)
my.plot.ts.short(x)
```

Μοντέλο κινούμενου μέσου όρου (Moving Average)

Έστω $\{X_t, t \geq 1\}$ μία χρονοσειρά και $\{w_t\}$, $t \geq 1$, μία οικογένεια ισόνομων και ανεξάρτητων τυχαίων μεταβλητών που ακολουθούν την $N(0, \sigma^2)$ κατανομή (δηλαδή WN).

Το μοντέλο κινητού μέσου 1ης τάξης, που συμβολίζεται με MA(1) είναι:

$$X_t = \mu + w_t + \theta_1 \cdot w_{t-1}.$$

Το μοντέλο κινητού μέσου 2ης τάξης, που συμβολίζεται με MA(2) είναι:

$$X_t = \mu + w_t + \theta_1 \cdot w_{t-1} + \theta_2 \cdot w_{t-2}.$$

Το μοντέλο κινητού μέσου όρου q^{th} τάξης, που συμβολίζεται με MA(q) είναι:

$$X_t = \mu + w_t + \theta_1 \cdot w_{t-1} + \theta_2 \cdot w_{t-2} + \ldots + + \theta_q \cdot w_{t-q}.$$

Τα μοντέλα MA είναι πάντα στάσιμα.

$$E(X_t) = \mu, \quad Var(X_t) = (1 + \theta_1^2 + \theta_2^2 + \cdots + \theta_q^2)\sigma^2.$$

Μοντέλο MA(1)

Μοντέλο MA(1)

Το μοντέλο κινητού μέσου 1ης τάξης, που συμβολίζεται με MA(1) είναι:

$$X_t = \mu + w_t + \theta \cdot w_{t-1},\ w_t \sim N(0, \sigma^2)$$

Εύκολα αποδεικνύεται ότι:

$E(X_t) = \mu$.

$Var(X_t) = (1 + \theta^2) \cdot \sigma^2$. $Var(X_t) = Var(\mu + w_t + \theta \cdot w_{t-1}) = Var(w_t) + Var(\theta \cdot w_{t-1}) = (1+\theta^2) \cdot \sigma^2$

Επιπλέον, μπορεί να δειχθεί ότι **$\rho_1 = \theta / (1 + \theta^2)$** και **$\rho_h = 0,\ h \geq 2$.**

Πράγματι: $Cov(X_t, X_{t-h}) = E[(X_t - \mu)\cdot(X_{t-h} - \mu)]$

$$= E[(w_t + \theta \cdot w_{t-1})\cdot(w_{t-h} + \theta \cdot w_{t-h-1})]$$

$$= E(w_t w_{t-h} + \theta w_{t-1} w_{t-h} + \theta w_t w_{t-h-1} + \theta w_{t-1} w_{t-h-1})$$

Παρατηρούμε, ότι για $h = 1$, είναι $Cov(X_t, X_{t-1}) = \theta \cdot \sigma^2_W$, ενώ $Cov(X_t, X_{t-h}) = 0$, για $h > 1$. Ο τύπος της της συνδιακύμανσης προκύπτει άμεσα.

Μοντέλο MA(1)

Άσκηση Εμπέδωσης

Έστω η MA(1) χρονοσειρά $X_t = 0.05 + w_t + 0.9\, w_{t-1}$, όπου $\{w_t\}$ iid και $\{w_t\} \sim N(0, 0.1^2)$. Να υπολογιστούν τα $E(X_t)$, $Var(X_t)$, $r_1 = cor(X_t, X_{t-1})$.

$\mu = 0{,}05$, $\theta = 0{,}9$, $\sigma^2 = 0{,}1^2$ ή $\sigma = 0{,}1$

$E(X_t) = \mu = 0{,}05$

$Var(X_t) = (1+\theta^2)\sigma^2 = (1+0{,}9^2)\cdot 0{,}1^2 = 0{,}0181$

$r_1 = \dfrac{0{,}9}{1+0{,}9^2} = \dfrac{0{,}9}{1{,}81}$, $r_h = 0$, $h \geq 2$

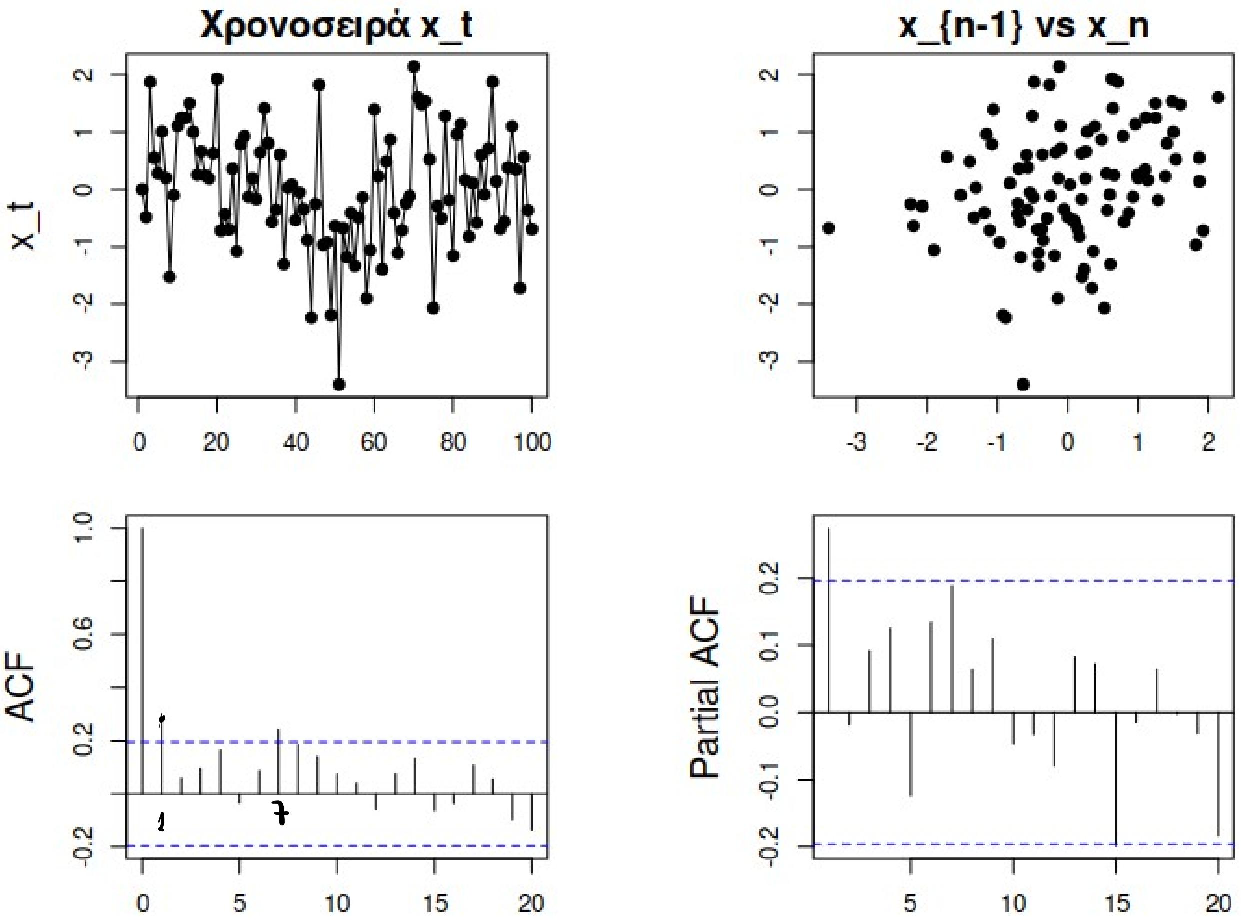
Μοντέλο MA(1)

Άσκηση Εμπέδωσης

Έστω η MA(1) χρονοσειρά $X_t = w_t + 0.4\, w_{t-1}$, όπου $\{w_t\}$ iid και $\{w_t\} \sim N(0, 0.1^2)$. Να υπολογιστούν τα $E(X_t)$, $Var(X_t)$, $r_1 = cor(X_t, X_{t-1})$.

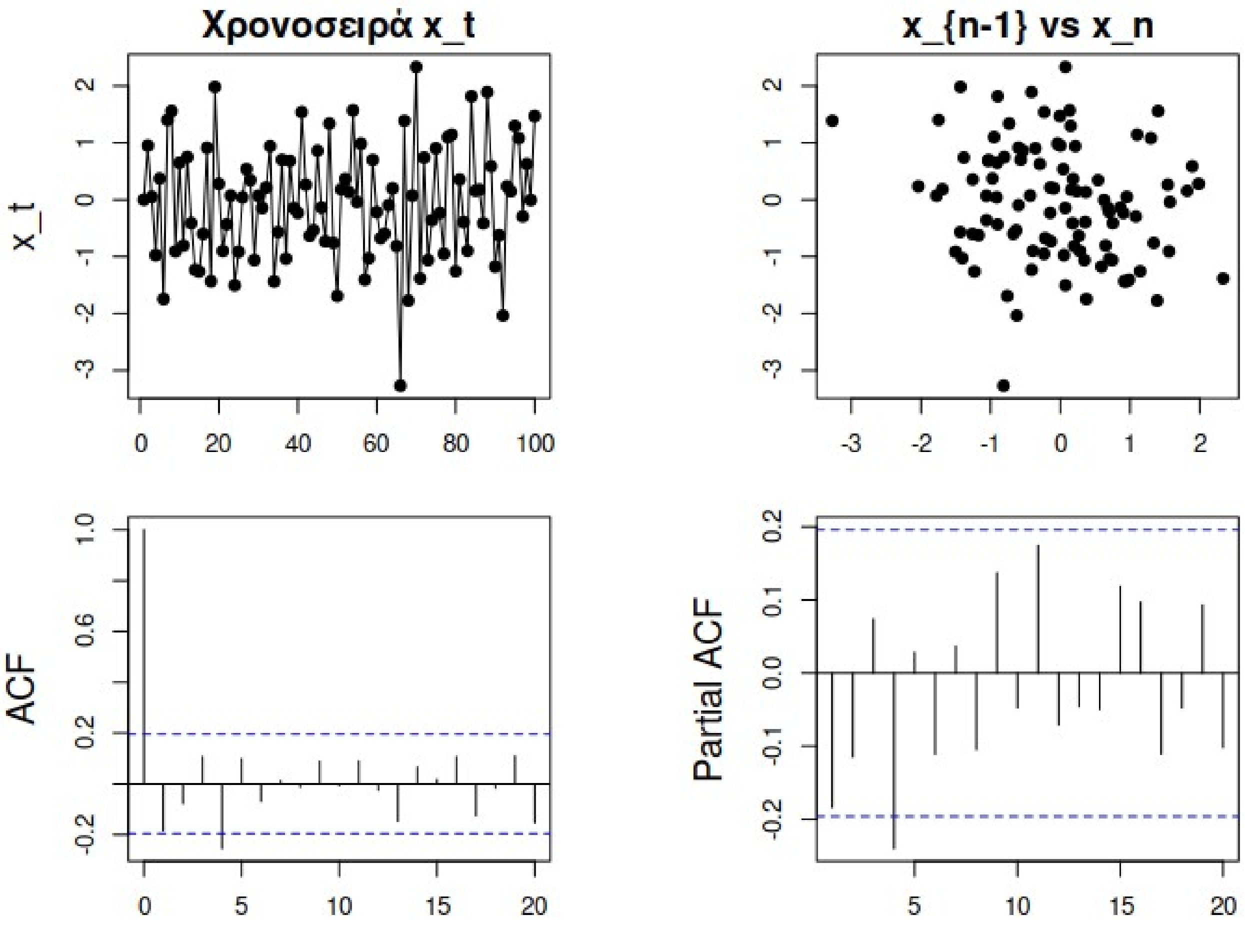
Παράδειγμα I

$X_t = w_t + 0.4 \cdot w_{t-1}$, $w_t \sim N(0, 1)$



$X_t = w_t - 0.3 \cdot w_{t-1}$, $w_t \sim N(0, 1)$

Παράδειγμα II



Αντιστροφή μοντέλου MA(1)

Αν $X_t = w_t + \theta \cdot w_{t-1}$.~ MA(1) τότε $\rho_1 = \theta / (1 + \theta^2)$ και $\rho_h = 0$, $h \geq 2$.

Παρατηρούμε ότι η αυτοσυσχέτιση ρ_1, παραμένει αμετάβλητη αν στη θέση του θ, τοποθετηθεί το 1/θ.

Πιο συγκεκριμένα, τα μοντέλα $\theta = 1/5$ $\sigma = 5$

$$X_t = w_t + 1/5 \cdot w_{t-1},\ w_t \sim N(0, 5^2),$$

$$X_t = w_t + 5 \cdot w_{t-1},\ w_t \sim N(0, 1),$$

$\theta = 5$ $\sigma = 1$

Θα έχουν και τα δύο:

- $E(X_t) = \mu = 0$.
- $Var(X_t) = (1 + \theta^2) \cdot \sigma^2 = 26$
- $\rho_1 = 5/26$. , $\rho_h = 0$, $h \geq 2$.

AR(p)

$w_t \sim X_t, X_{t-1}, \ldots$

Αντιστροφή μοντέλου MA(1)

Τίθεται το ερώτημα: Πού διαφέρουν τα μοντέλα:

$$X_t = w_t + 1/5 \cdot w_{t-1},\ w_t \sim N(0, 5^2),$$

$$X_t = w_t + 5 \cdot w_{t-1},\ w_t \sim N(0, 1),$$

Απάντηση: Μόνο ένα από τα δύο μπορεί να γραφεί στη μορφή AR(∞).

Θεώρημα

Έστω $X_t = w_t + \theta \cdot w_{t-1}$.~ MA(1) και $|\theta| < 1$. Τότε είναι δυνατόν τα υπόλοιπα να εκφραστούν ως ένα AR(∞) μοντέλο της χρονοσειράς: $w_t = X_t + \theta X_{t-1} + \theta^2 X_{t-2} + \ldots$

Στην περίπτωση αυτή, το μοντέλο ονομάζεται **αντιστρέψιμο**.

Παράδειγμα

Από τα δύο μοντέλα που αναφέρθηκαν, το $X_t = w_t + 1/5 \cdot w_{t-1},\ w_t \sim N(0, 5^2)$, είναι αντιστρέψιμο ενώ το $X_t = w_t + 5 \cdot w_{t-1},\ w_t \sim N(0, 1)$, δεν είναι.

Αντιστροφή μοντέλου MA(1)

Θεώρημα

Έστω $X_t = w_t + \theta \cdot w_{t-1} \sim MA(1)$ και $|\theta| < 1$. Τότε είναι δυνατόν τα υπόλοιπα να εκφραστούν ως ένα $AR(\infty)$ μοντέλο της χρονοσειράς: $w_t = X_t + \theta X_{t-1} + \theta^2 X_{t-2} + \dots$. Στην περίπτωση αυτή, το μοντέλο ονομάζεται **αντιστρέψιμο**.

Απόδειξη:

$$X_t = w_t + \theta w_{t-1} \Leftrightarrow$$

$$X_t = (I - \theta B) w_t \Leftrightarrow$$

$$w_t = (I - \theta B)^{-1} X_t \Leftrightarrow$$

$$w_t = (I + \theta B + \theta^2 B^2 + \dots) X_t \Leftrightarrow$$

$$w_t = X_t + \theta X_{t-1} + \theta^2 X_{t-2} + \dots$$

$$X_t = w_t + \theta \cdot w_{t-1} \Leftrightarrow w_t = X_t - \theta \cdot w_{t-1} \overset{(1)}{=} X_t - \theta X_{t-1} + \theta^2 w_{t-2}$$

$$w_{t-1} = X_{t-1} - \theta \cdot w_{t-2} \quad (1)$$

$$w_{t-2} = X_{t-2} - \theta \cdot w_{t-3} \quad (2)$$

$$\overset{(2)}{=} X_t - \theta \cdot X_{t-1} + \theta^2 X_{t-2} - \theta^3 \cdot w_{t-3}$$

$$= \sum_{n=0}^{+\infty} (-\theta)^n \cdot X_{t-n}$$

$\downarrow |\theta| < 1$.

συγκλίνει γιατί $|\theta| < 1 \Rightarrow$

$\Rightarrow (-\theta)^n \to 0$

Σημείωση

Μία ισοδύναμη διατύπωση είναι η εξής:

MA(1) ≡ AR(∞) αν και μόνο αν η ρίζα του Θ(z) = 1 + θz, βρίσκεται εκτός του μοναδιαίου δίσκου.

Αντιστροφή μοντέλου MA(1)

Η αντιστρεψιμότητα είναι επιθυμητή γιατί δίνει την ευκαιρία αντίληψης του MA(1) μοντέλου $X_t = w_t + \theta \cdot w_{t-1}$ ως ένα οριακό AR(p) μοντέλο πρόβλεψης των σφαλμάτων w_t.

$$w_t = \left(I + \theta B + \theta^2 B^2 + \dots\right) X_t = \sum_{i=1}^{\infty} \theta^i X_{t-i}$$

Ο παραπάνω τύπος δηλώνει πως το σφάλμα κάθε παρατήρησης εκφράζεται ως άθροισμα όλων των προηγούμενων παρατηρήσεων. Ειδικότερα, μπορούμε να υπολογίζουμε αναδρομικά τα σφάλματα από τις παρατηρούμενες τιμές της χρονοσειράς, έχοντας τη ρεαλιστική παραδοχή πως το μέρος της συνεισφοράς κάθε μίας παρατήρησης στο υπόλοιπο w_t, μειώνεται με την απόσταση της από τη θέση αυτή.

Στη συνέχεια μπορούμε να ελέγξουμε εάν τα σφάλματα συμπεριφέρονται σαν λευκός θόρυβος (π.χ. με τη μέθοδο Box–Jenkins) και επομένως εάν το μοντέλο είναι μια καλή προσέγγιση με τα δεδομένα.

Αντιστροφή μοντέλου MA(1)

Άσκηση Εμπέδωσης

Έστω η MA(1) χρονοσειρά $X_t = w_t - 0.7\, w_{t-1}$, όπου $\{w_t\}$ iid. Να γραφεί η χρονοσειρά σε AR(∞) μορφή.

$$X_t = w_t - 0{,}7\, w_{t-1} = (I - 0{,}7 \cdot B) \cdot w_t \Leftrightarrow$$

$$\Leftrightarrow \quad w_t = (I - 0{,}7 \cdot B)^{-1} \cdot X_t = \sum_{n=0}^{+\infty} (-0{,}7)^n \cdot B^n \cdot X_t =$$

$$= \sum_{n=0}^{+\infty} (-0{,}7)^n \cdot X_{t-n} = X_t - 0{,}7 \cdot X_{t-1} + 0{,}7^2 \cdot X_{t-2} - \cdots$$

Αντιστροφή μοντέλου MA(1)

Άσκηση Εμπέδωσης

Δείξτε ότι οι MA(1) χρονοσειρές $X_t = w_t - 0.5\, w_{t-1}$, και $X_t = w_t - 2\, w_{t-1}$ έχουν την ίδια αυτοσυσχέτιση 1ης τάξης. Βρείτε ποια από τις δύο είναι αντιστρέψιμη.

$$\rho_1 = \frac{\theta}{1+\theta^2}, \quad \theta = -0{,}5 \Rightarrow \rho_1 = \frac{-0{,}5}{1+0{,}5^2} = \frac{-0{,}5}{1{,}25} = -\frac{50}{125} = -\frac{2}{5}.$$

$$\theta = -2 \Rightarrow \rho_1 = \frac{-2}{1+(-2)^2} = \frac{-2}{5}$$

Το $X_t = w_t - 0{,}5\, w_{t-1}$, είναι αντιστρέψιμο γιατί $|\theta| = |-0{,}5| < 1$.

Μοντέλο MA(2)

Μοντέλο MA(2)

Το μοντέλο κινητού μέσου 2ης τάξης, που συμβολίζεται με MA(2) είναι:

$$X_t = \mu + w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2},\ w_t \sim N(0, \sigma^2)$$

Εύκολα αποδεικνύεται ότι:

$$\mathbf{E(X_t) = \mu.}$$

$$\mathbf{Var(X_t) = (1 + \theta^2_1 + \theta^2_2) \cdot \sigma^2.}$$

Επιπλέον, μπορεί να δειχθεί ότι

$$\rho_1 = (\theta_1 + \theta_1 \cdot \theta_2) / (1 + \theta^2_1 + \theta^2_2),$$

$$\rho_2 = \theta_2 / (1 + \theta^2_1 + \theta^2_2),$$

$$\rho_h = 0,\ h \geq 3.$$

Μοντέλο MA(2)

Άσκηση Εμπέδωσης

Έστω η MA(2) χρονοσειρά $X_t = w_t + 0.5\, w_{t-1} + 0.3\, w_{t-1}$, όπου $\{w_t\}$ iid και $\{w_t\} \sim N(0, 1)$.

Να υπολογιστούν τα $E(X_t)$, $Var(X_t)$, $r_1 = cor(X_t, X_{t-1})$.

$\theta_1 = 0{,}5$, $\theta_2 = 0{,}3$

$E(X_t) = 0$

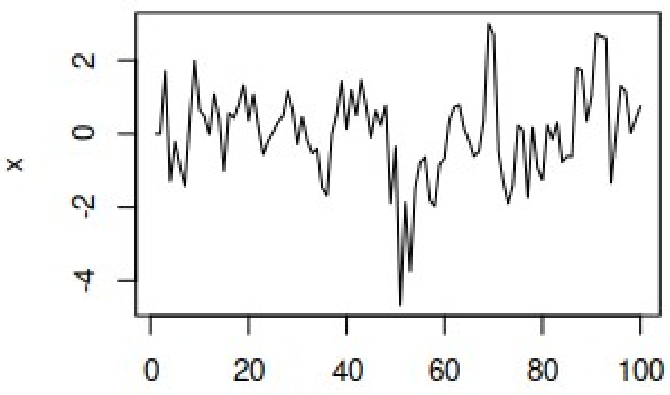
$Var(X_t) = (1 + 0{,}5^2 + 0{,}3^2) \cdot 1^2 = 1{,}34$

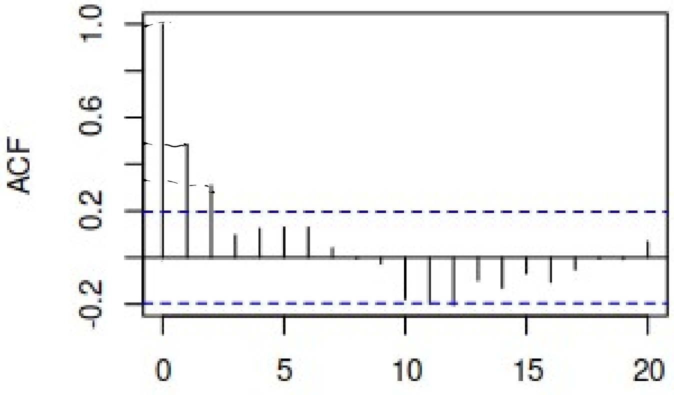
$r_1 > 0$

$r_2 > 0$

$r_h = 0, h \geq 3.$

Παράδειγμα Ι



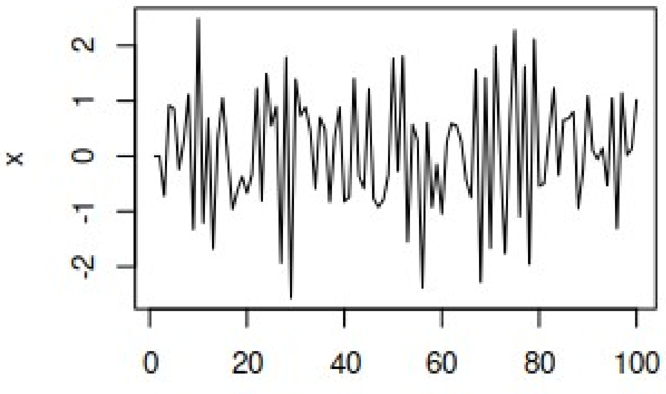


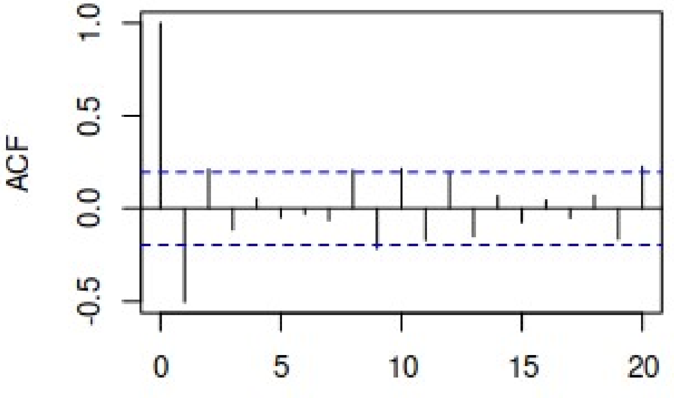
Δεδομένα $X_t = w_t + 0.5 \cdot w_{t-1} + 0.3 \cdot w_{t-2}$

```
ma.sim2 = function(theta1 = 0.5, theta2 = 0.3, number = 100){
  ma_vals = rep(0, times = number)
  wn = rnorm(number)
  for(i in 3:number){
    ma_vals[i] = wn[i] + theta1 * wn[i-1] + theta2 * wn[i-2]}
  return(ma_vals)
}

x=ma.sim2(theta1 = 0.5, theta2 = 0.3, number=100)
par(mfrow = c(2, 2))
plot(x, type = 'l')
acf(x)
pacf(x)
```

Παράδειγμα ΙΙ





Δεδομένα $X_t = w_t - 0.5 \cdot w_{t-1} + 0.1 \cdot w_{t-2}$

```
ma.sim2 = function(theta1 = 0.5, theta2 = 0.3, number = 100){ma_vals = rep(0, times = number)
  wn = rnorm(number)
  for(i in 3:number){
    ma_vals[i] = wn[i] + theta1 * wn[i-1] + theta2 * wn[i-2]}
  return(ma_vals)
}

x=ma.sim2(theta1 = -0.3, theta2 = 0.1, number=100)
par(mfrow = c(2, 2))
plot(x, type = 'l')
acf(x)
pacf(x)
```

Αντιστροφή μοντέλου MA(2)

Έστω $X_t = w_t + \theta_1 \cdot w_{t-1} + \theta_2 \cdot w_{t-2}$

$= (I + \theta_1 B + \theta_2 B^2)\, w_t$ (το $1 + \theta_1 z + \theta_2 z^2$ παραγοντοποιείται σε όρους 1ου βαθμού)

$= (I - \varphi_1 . B)(I - \varphi_2 . B)\, w_t \Leftrightarrow$

$w_t = (I - \varphi_2 B)^{-1}(I - \varphi_1 B)^{-1} X_t.$

Αν $|\varphi_i| < 1$, i = 1, 2 τότε $(1 - \varphi_i B)^{-1} = \sum_{n=0}^{\infty} \varphi_i^n B^n,\ i = 1, 2,$ και είναι δυνατή η έκφραση των w_t στη μορφή AR(∞) από τις τιμές X_t.

Στην περίπτωση αυτή, το MA(2) μοντέλο ονομάζεται **αντιστρέψιμο**.

Παρατηρούμε ότι οι ρίζες του $1 + \theta_1 z + \theta_2 z^2 = (I - \varphi_1 z)(I - \varphi_2 z)$ είναι οι αριθμοί $1/\varphi_1$, $1/\varphi_2$.

Συμπεραίνουμε ότι:

MA(2) ≡ AR(∞) αν και μόνο αν οι ρίζες του $1 + \theta_1 z + \theta_2 z^2$ είναι έξω από το μοναδιαίο δίσκο.

Αντιστροφή μοντέλου MA(2)

Θεώρημα

Έστω $X_t = w_t + \theta_1 \cdot w_{t-1} + \theta_2 \cdot w_{t-2}$, και έστω ότι οι ρίζες της εξίσωσης $1 + \theta_1 z + \theta_2 z^2 = 0$, βρίσκονται έξω από το μοναδιαίο δίσκο. Τότε είναι δυνατόν τα υπόλοιπα του μοντέλου να εκφραστούν στη μορφή AR(∞) από τιμές της σειράς. Στην περίπτωση αυτή, το μοντέλο ονομάζεται **αντιστρέψιμο**.

Παράδειγμα

Να βρεθεί αν η MA(2) χρονοσειρά $X_t = w_t + 0.5w_{t-1} - 0.3w_{t-2}$, είναι αντιστρέψιμη.

Λύση

Ελέγχουμε τις ρίζες της εξίσωσης

$1 + 0.5z – 0.3z^2 = 0$. Οι ρίζες είναι $z_{1,2} = 0.5 \pm 1.45^{½}$, και παρατηρούμε πως και οι δύο έχουν μέτρο μεγαλύτερο της μονάδας. Άρα, το μοντέλο είναι αντιστρέψιμο.

Αντιστροφή μοντέλου MA(2)

Παράδειγμα

(α) Να δείξετε ότι η MA(2) χρονοσειρά $X_t = w_t - w_{t-1} + 0.21 w_{t-2}$, $w_t \sim N(0, \sigma^2)$, είναι στάσιμη και αντιστρέψιμη.

(β) Να βρεθεί η έκφραση των υπολοίπων ως σειρά AR(∞) των τιμών X_t.

Λύση

(α) Η στασιμότητα είναι προφανής. $E(X_t) = 0$, $Var(X_t) = (1 + (-1)^2 + 0.21^2) \cdot \sigma^2 = \sigma\tau\alpha\theta$

Είναι $\Theta(z) = 1 - z + 0.21 z^2$. Επαληθεύουμε ότι οι ρίζες της είναι εκτός του μοναδιαίου δίσκου άρα η σειρά είναι αντιστρέψιμη.

$\Delta = (-1)^2 - 4 \cdot 0{,}21 \cdot 1 = 1 - 0{,}84 = 0{,}16$, $z_{1,2} = \dfrac{1 \pm 0{,}4}{0.42}$

$\dfrac{1.4}{0.42} = \dfrac{0.7}{0.21} = \dfrac{70}{21} = \dfrac{10}{3} > 1$

$\dfrac{0.6}{0.42} = \dfrac{30}{21} = \dfrac{10}{7} > 1$

Αντιστροφή μοντέλου MA(2)

Παράδειγμα

(α) Να δείξετε ότι η MA(2) χρονοσειρά $X_t = w_t - w_{t-1} + 0.21 w_{t-2}$, wt ~ N(0, σ^2), είναι στάσιμη και αντιστρέψιμη.

(β) Να βρεθεί η έκφραση των υπολοίπων ως σειρά AR(∞) των τιμών X_t.

Λύση

(β) Απαιτείται ο εντοπισμός των φ_1, φ_2, στην έκφραση $1 - z + 0.21 z^2 = (1 - \varphi_1 \cdot z)(1 - \varphi_2 z)$.

Παρατηρούμε ότι πρέπει $\varphi_1 \cdot \varphi_2 = 0.21$ και $\varphi_1 + \varphi_2 = 1$, από όπου βρίσκουμε $\varphi_1 = 0.7$, $\varphi_2 = 0.3$.

Άρα, $w_t = (1 - 0.7B)^{-1}(1 - 0.3B)^{-1} X_t = \left(\sum_{n=0}^{\infty} 0.7^n B^n\right)\left(\sum_{n=0}^{\infty} 0.3^n B^n\right) X_t = \sum_{n=0}^{+\infty} \delta_n X_{t-n}$,

όπου $\delta_0 = 1$, $\delta_1 = 0.7 + 0.3 = 1$, και $\delta_n = \sum_{i=0}^{n} 0.7^i \cdot 0.3^{n-i}$.

$(I + 0{,}7 \cdot B + 0{,}7^2 \cdot B^2 + \ldots) \cdot (I + 0{,}3 \cdot B + 0{,}3^2 \cdot B^2 + \ldots) = I + (0{,}7 + 0{,}3) B + (0{,}7^2 + 0{,}7 \cdot 0{,}3 + 0{,}3^2) \cdot B^2 + \ldots$

Αντιστροφή μοντέλου MA(2)

Άσκηση

(α) Να δείξετε ότι η MA(2) χρονοσειρά $X_t = w_t – 0.1w_{t-1} – 0.56w_{t-2}$, $w_t \sim N(0, \sigma^2)$, είναι στάσιμη και αντιστρέψιμη.

(β) Να βρεθεί η έκφραση των υπολοίπων ως σειρά AR(∞) των τιμών X_t.

Λύση

Γενική εξίσωση MA(q)

Η γενική εξίσωση ενός MA(q) μοντέλου είναι ($\varepsilon_t \sim IWN(0, \sigma^2)$):

$$X_t = \mu + \varepsilon_t + \theta_1 \cdot \varepsilon_{t-1} + \theta_2 \cdot \varepsilon_{t-2} + \ldots + \theta_q \cdot \varepsilon_{t-q},$$

Χωρίς βλάβη της γενικότητας μπορεί να γραφεί

$$\mathbf{X_t = \mu + \varepsilon_t - \theta_1 \cdot \varepsilon_{t-1} - \theta_2 \cdot \varepsilon_{t-2} - \ldots - \theta_q \cdot \varepsilon_{t-q}.}$$

Με χρήση του τελεστή B η τελευταία σχέση εκφράζεται ως εξής:

$$X_t = \mu + (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q)\varepsilon_t,$$

ή ακόμα και ως

$$\mathbf{X_t = \mu + \Theta(B)\varepsilon_t,}$$

όπου $\mathbf{\Theta(\omega) = 1 - \theta_1\omega - \theta_2\omega^2 - \ldots - \theta_q\omega^q}$, το **χαρακτηριστικό πολυώνυμο** της MA(q).

Γενική εξίσωση MA(q)

Θεώρημα

Έστω το MA(q) μοντέλο $X_t = \mu + \Theta(B)\varepsilon_t$, όπου $\Theta(\omega) = 1 - \theta_1\omega - \theta_2\omega^2 - \dots - \theta_q\omega^q$. Αν οι ρίζες της $\Theta(\omega) = 0$, βρίσκονται έξω από το μοναδιαίο δίσκο τότε το μοντέλο είναι αντιστρέψιμο.

Απόδειξη

Ανάλογη με την περίπτωση $q = 2$.

Μοντέλα ARMA(p, q)

Μοντέλα ARMA(p, q)

Τα μοντέλα χρονοσειρών που είναι γνωστά ως μοντέλα ARMA μπορεί να περιλαμβάνουν αυτοπαλινδρομικούς όρους, ~~διαφορές πρώτης ή μεγαλύτερης τάξης μεταξύ των παρατηρήσεων~~ και όρους κινητού μέσου όρου.

Τα στοιχεία στο μοντέλο καθορίζονται με τη σειρά (AR, MA).

Ένα μοντέλο με δύο όρους AR μόνο, θα προσδιορίζεται ως ARMA τάξης (2,0).

Ένα μοντέλο MA(2) θα προσδιορίζεται ως ARMA τάξης (0,2).

Μοντέλα ARMA(p, q)

Η γενική εξίσωση ενός ARMA(p, q) μοντέλου είναι (ε_t ~ IWN(0, σ^2)):

$$X_t = \delta + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q},$$

Χωρίς βλάβη της γενικότητας μπορεί να γραφεί

$$X_t = \delta + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q},$$

Με χρήση των τελεστών B, Θ, η τελευταία σχέση εκφράζεται ως εξής:

$$\mathbf{\Phi(B)X_t = \delta + \Theta(B)\varepsilon_t}$$

όπου **Φ(ω) = 1 – α1ω – α2ω^2 – … – αpω^p**, το χαρακτηριστικό πολυώνυμο της AR(p),

και **Θ(ω) = 1 – θ1ω – θ2ω^2 – … – θqω^q**, το χαρακτηριστικό πολυώνυμο της MA(q).

Μοντέλα ARMA(p, q)

Παράδειγμα

Ένα ARMA(1, 1) μοντέλο με σταθερά $\delta = 0$, έχει εξίσωση $X_t = \varphi X_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1}$.

Αποδεικνύεται ότι:

$\rho_1 = (\varphi + \theta)(1 + \varphi\theta) / (1 + 2\varphi\theta + \theta^2)$ και

$\rho_h = \varphi^{h-1}\rho_1$, $h \geq 2$.

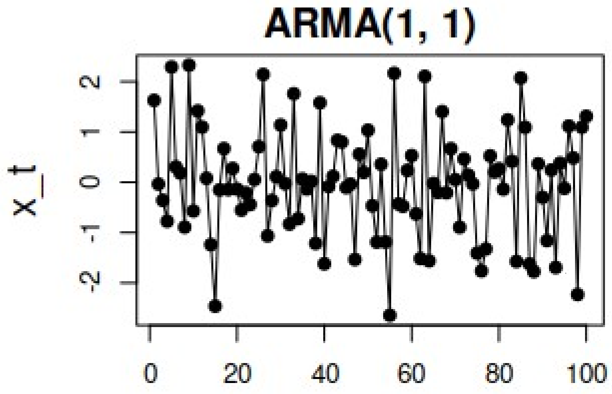
Ειδικότερα, αν $\varphi = \theta$, τότε ARMA(1, 1) = WN.

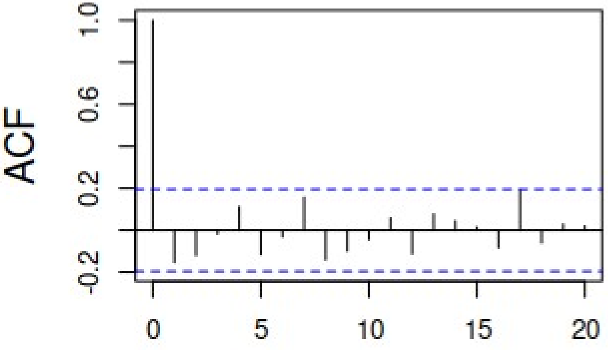
Επιπλέον, αποδεικνύεται ότι ένα ARMA(1, 1) μοντέλο $X_t = \varphi X_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1}$ είναι:

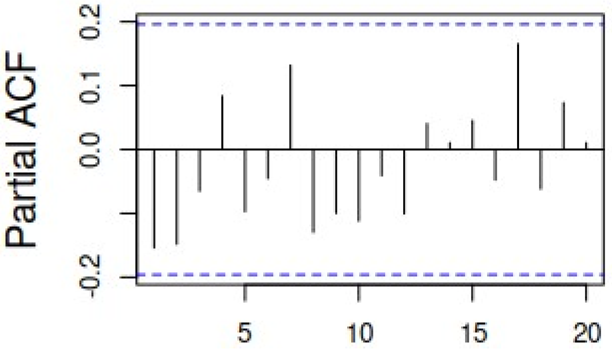
- Στάσιμο αν το AR(1) κομμάτι του είναι στάσιμο: $|\varphi| < 1$.
- Αντιστρέψιμο αν το MA(1) κομμάτι του είναι αντιστρέψιμο: $|\theta| < 1$.

Παράδειγμα ARMA(1, 1)

Στο διάγραμμα παρουσιάζεται η χρονοσειρά $X_t = 0.5\ X_{t-1} + \varepsilon_t - 0.5\ \varepsilon_{t-1}$ με δείγμα 100 τιμών.



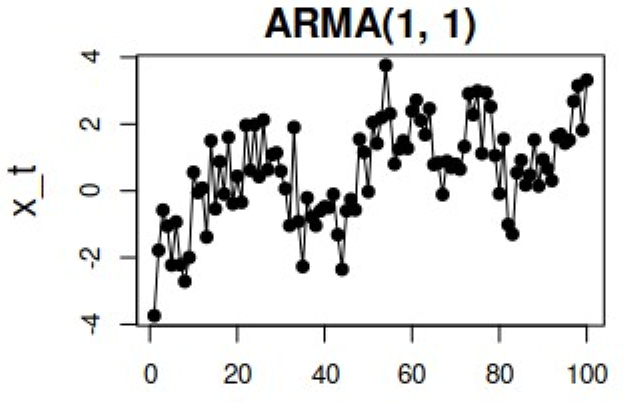


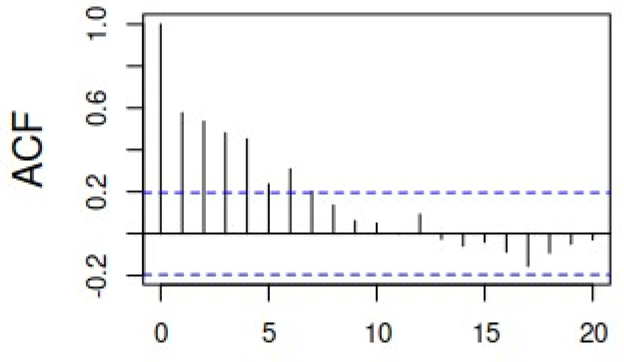


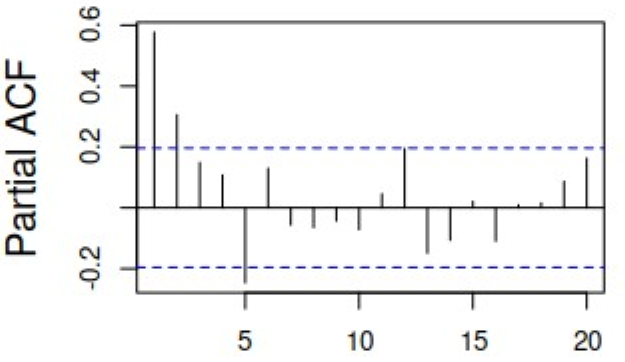

```
x <- arima.sim(n=100, model=list(ar=0.5, ma=-0.5))
```

Παράδειγμα ARMA(1, 1)

Στο διάγραμμα παρουσιάζεται η χρονοσειρά $X_t = 0.9\, X_{t-1} + \varepsilon_t - 0.5\, \varepsilon_{t-1}$ με δείγμα 100 τιμών.



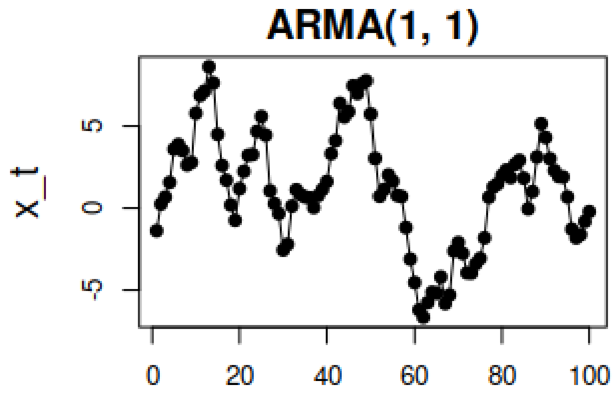


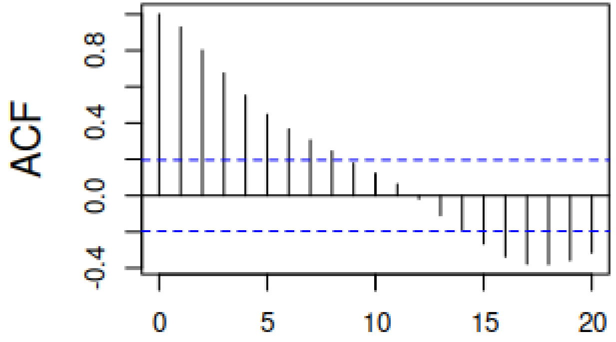


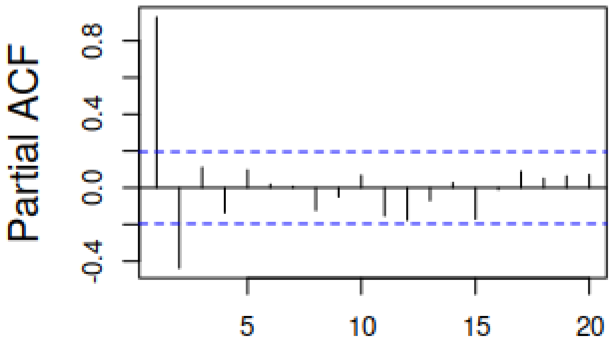

```
x <- arima.sim(n=100, model=list(ar=0.9, ma=-0.5))
```

Παράδειγμα ARMA(1, 1)

Στο διάγραμμα παρουσιάζεται η χρονοσειρά $X_t = 0.9\ X_{t-1} + \varepsilon_t + 0.5\ \varepsilon_{t-1}$ με δείγμα 100 τιμών.








```
x <- arima.sim(n=100, model=list(ar=0.9, ma=0.5))
```

Παράδειγμα ARMA(1, 1)

```
my.plot.ts.simple = function(x, title = "Χρονοσειρά x_t"){
  library(tseries)
  library(Hmisc)
  x = as.vector(x)
  par(mfrow=c(2,2))
  par(mar = c(2, 4.5, 2, 4.5))
  plot(x, main = title, xlab = "t", ylab = "x_t", cex = 1.5, cex.lab = 1.5, cex.main = 1.5, pch=20);
  lines(1:length(x), x, pch=20)
  acf(x, pl=T, cex.lab = 1.5, cex.main = 1.5)
  pacf(x, pl=T, cex.lab = 1.5, cex.main = 1.5)
  par(mfrow=c(1,1))
}

x <- arima.sim(n=100, model=list(ar=0.9, ma=0.5))
my.plot.ts.simple(x, title = 'ARMA(1, 1)')
```

Μοντέλα ARIMA(p, d, q)

Μοντέλα ARIMA(p, d, q)

Γενική εξίσωση I(d)

Παρατηρούμε ότι η διαφορά δύο διαδοχικών όρων μίας χρονοσειράς μπορεί να γραφεί ως:

$$\Delta X_t = X_t - X_{t-1} = X_t - B(X_t) = (I - B)X_t.$$

Άρα, η διαφορά 2ης τάξης των στοιχείων της χρονοσειράς γράφεται ως:

$$\Delta^2 X_t = Z_t - Z_{t-1} = Z_t - B(Z_t) = (I - B)Z_t = (I - B)^2 X_t.$$

Συμπεραίνουμε ανάλογα ότι ο υπολογισμός των διαφορών d τάξης μίας χρονοσειράς αποδίδει την ακολουθία

$$\Delta^d X_t = (I - B)^d X_t.$$

Μοντέλα ARIMA(p, d, q)

Γενική εξίσωση ARIMA(p, d, q)

Ένα ARIMA(p, d, q) μοντέλο προσαρμόζει ένα μοντέλο ARMA(p, q) στις διαφορές d τάξης της χρονοσειράς. Η γενική εξίσωση του ARIMA(p, d, q) μοντέλου είναι:

$$\Phi(B)\{(I - B)^d X_t\} = \delta + \Theta(B)\varepsilon_t,$$

όπου $\Phi(\omega) = 1 - \alpha_1\omega - \alpha_2\omega^2 - \ldots - \alpha_p\omega^p$, το χαρακτηριστικό πολυώνυμο της AR(p) και

$\Theta(\omega) = 1 - \theta_1\omega - \theta_2\omega^2 - \ldots - \theta_q\omega^q$, το χαρακτηριστικό πολυώνυμο της MA(q).

Μοντέλα ARIMA(p, d, q)

Τα μοντέλα ARIMA μπορεί να περιλαμβάνουν αυτοπαλινδρομικούς όρους, διαφορές πρώτης ή μεγαλύτερης τάξης μεταξύ των παρατηρήσεων και όρους κινητού μέσου όρου.

Τα στοιχεία στο μοντέλο καθορίζονται με τη σειρά (AR, διαφορές, MA).

Ένα μοντέλο με δύο όρους AR μόνο, θα προσδιορίζεται ως ARIMA τάξης (2,0,0).

Ένα μοντέλο MA(2) θα προσδιορίζεται ως ARIMA τάξης (0,0,2).

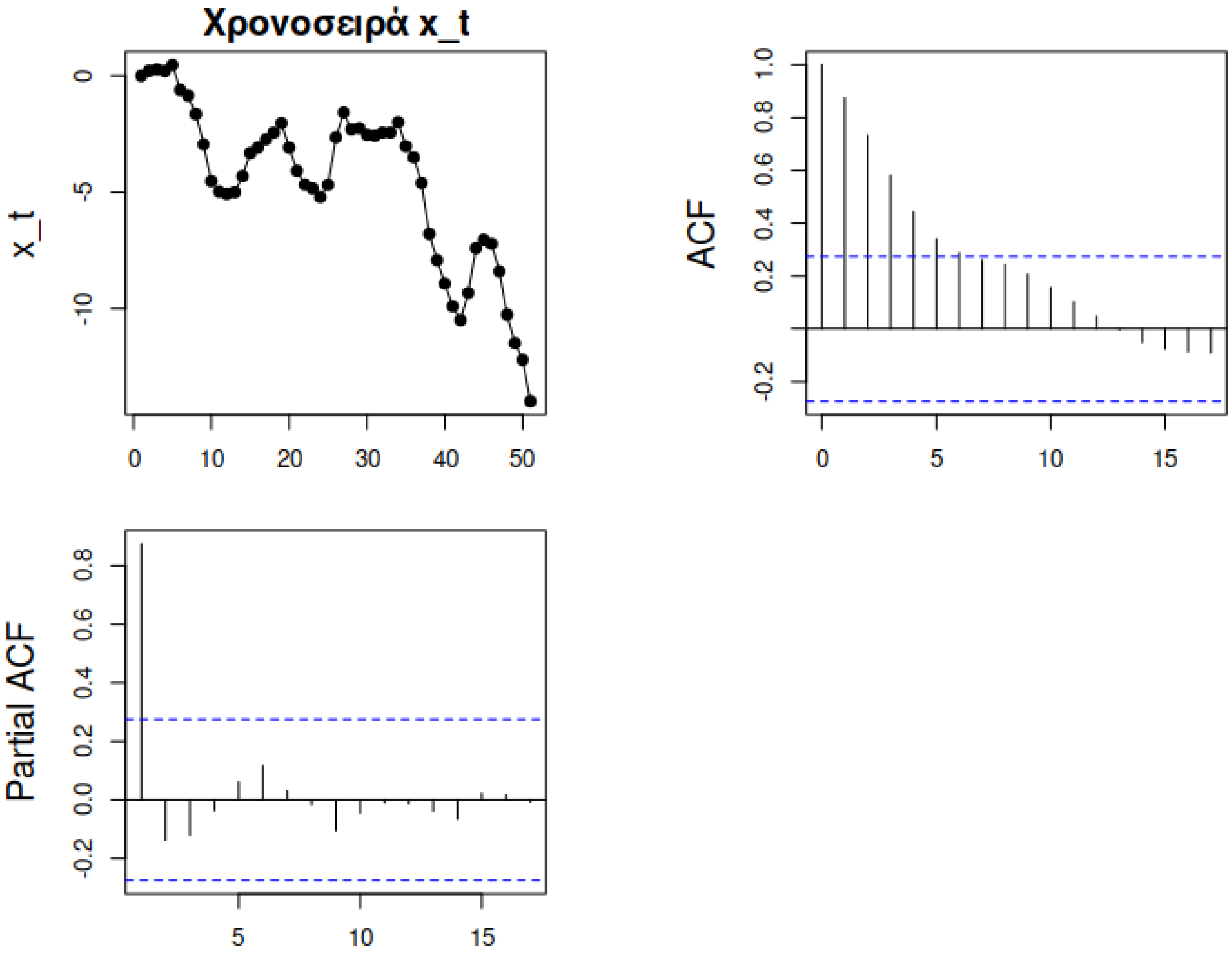
Ένα μοντέλο με έναν όρο AR, μια διαφορά πρώτης τάξης μεταξύ των τιμών και έναν όρο MA θα προσδιορίζεται ως (1,1,1).

Παράδειγμα ARIMA(1, 1, 1): $Z_t = \mu + \alpha Z_{t-1} + \theta \varepsilon_{t-1}$, $Z_t = X_t - X_{t-1}$ ή

$$\mathbf{X_t - X_{t-1} = \mu + \alpha(X_{t-1} - X_{t-2}) + \theta\varepsilon_{t-1},}$$

Παράδειγμα ARIMA(1, 1, 1)

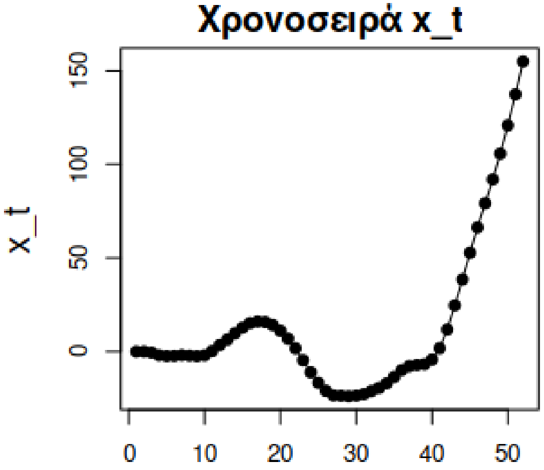
$X_t - X_{t-1} = 0.7(X_{t-1} - X_{t-2}) + 0.2\varepsilon_{t-1}.$

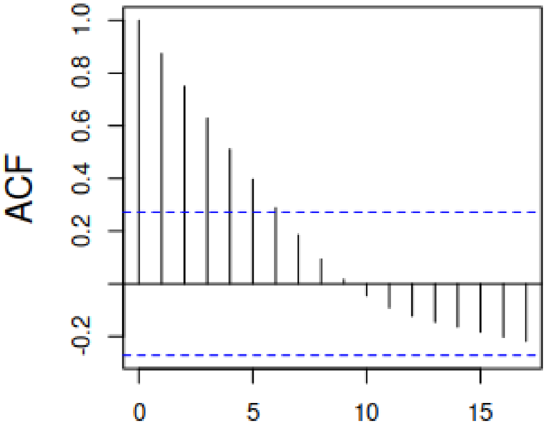


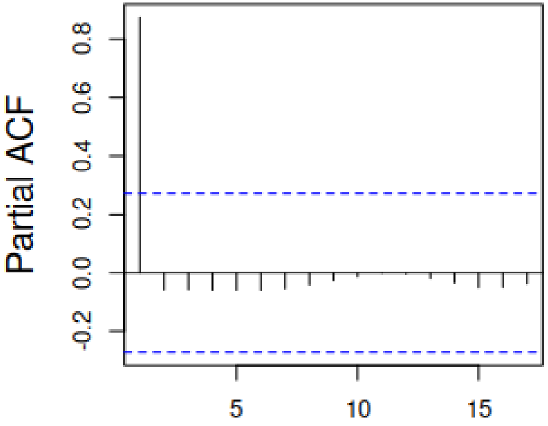
x = arima.sim(list(order = c(1,1,1), ar = 0.7, ma = 0.2), n = 50)

Παράδειγμα ARIMA(1, 2, 1)

$\Delta X_t - \Delta X_{t-1} = 0.7(\Delta X_{t-1} - \Delta X_{t-2}) + 0.2\varepsilon_{t-1}.$







x = arima.sim(list(order = c(1,2,1), ar = 0.7, ma = 0.2), n = 50)