



ΔΗΜΟΚΡΙΤΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΡΑΚΗΣ

DEMOCRITUS
UNIVERSITY
OF THRACE

Στοχαστική Ανάλυση Χρονοσειρών

Περιεχόμενο:

Πρόβλεψη

Θεώρημα Αποσύνθεσης του Wold

Επιλέγοντας ένα μοντέλο

Μοντέλα SARIMA(p, d, q)(P, D, Q)_m



ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΡΑΚΗΣ

DEMOCRITUS
UNIVERSITY
OF THRACE

Σύνοψη των αποτελεσμάτων (1^η και 2^η διάλεξη)



Στασιμότητα και Αντιστρεψιμότητα

MA(q): $X_t = \mu + \theta(B)\varepsilon_t$, $\varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2)$, $\theta(z) = 1 - \theta_1 z - \theta_2 z^2 - \dots - \theta_q z^q$.

Μία σειρά MA(q) είναι πάντα στάσιμη. Αν οι ρίζες του $\theta(z)$ είναι εκτός του μοναδιαίου δίσκου τότε η MA(q) είναι αντιστρέψιμη, δηλαδή είναι δυνατόν τα υπόλοιπα να εκφραστούν από τις τιμές της χρονοσειράς.

AR(p): $\varphi(B)X_t = \delta + \varepsilon_t$, $\varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2)$, $\varphi(\omega) = 1 - \alpha_1 \omega - \alpha_2 \omega^2 - \dots - \alpha_p \omega^p$.

Μία σειρά AR(p) είναι πάντα αντιστρέψιμη. Αν οι ρίζες του $\varphi(z)$ είναι εκτός του μοναδιαίου δίσκου τότε η AR(p) γράφεται στη μορφή MA(∞), ειδικότερα είναι στάσιμη.

ARMA(p, q): $\varphi(B)X_t = \delta + \theta(B)w_t$

Είναι στάσιμο αν-ν το AR(p) μέρος του είναι. Είναι αντιστρέψιμο αν-ν το MA(q) είναι.

Augmented Dickey-Fuller (ADF):

H_0 : Στο AR(p) υπάρχει μοναδιαία ρίζα vs H_1 : Δεν υπάρχει μοναδιαία ρίζα.

KPSS Level: $H_{0,\text{Level}}$: $X_t = \mu + u_t$, u_t στάσιμη vs $H_{1,\text{Level}}$: $X_t = \mu + r_t + u_t$.

KPSS Trend: $H_{0,\text{Trend}}$: $X_t = \mu + bt + u_t$ vs $H_{1,\text{Trend}}$: $X_t = \mu + bt + r_t + u_t$.



Αποτελέσματα για τα στάσιμα ARMA μοντέλα

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

- $\rho_h = \varphi_1^h$, $h \geq 0$.

AR(2): $X_t = \delta + \varphi_1 \cdot X_{t-1} + \varphi_2 \cdot X_{t-2} + w_t$.

- $\rho_1 = \varphi_1 / (1 - \varphi_2)$.
- $\rho_2 = \varphi_1^2 / (1 - \varphi_2) + \varphi_2$.
- $\rho_3 = (\varphi_1^3 + \varphi_1 \cdot \varphi_2) / (1 - \varphi_2) + \varphi_1 \cdot \varphi_2$.

ARMA(1, 1): $X_t = \varphi X_{t-1} + w_t - \theta w_{t-1}$.

- $\rho_1 = (\varphi + \theta)(1 + \varphi\theta) / (1 + 2\varphi\theta + \theta^2)$.
- $\rho_h = \varphi^{h-1} \rho_1$, $h \geq 2$.

AR(p): $\rho_k \sim C_1 \lambda_1^k + \dots + C_m \lambda_m^k$, $|\lambda_i| < 1$, λ_i : αντίστροφοι των ριζών του $\varphi(z)$.

ARMA(p, q): Τα ρ_k έχουν μορφή που προκύπτει από το συνδυασμό των επιρροών του AR και του MA μέρους του μοντέλου.

MA(1): $X_t = \mu + w_t + \theta w_{t-1}$.

- $\rho_1 = \theta / (1 + \theta^2)$ και $\rho_h = 0$, $h \geq 2$.

MA(2): $X_t = \mu + w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2}$.

- $\rho_1 = (\theta_1 + \theta_1 \cdot \theta_2) / (1 + \theta_1^2 + \theta_2^2)$,
- $\rho_2 = \theta_2 / (1 + \theta_1^2 + \theta_2^2)$,
- $\rho_h = 0$, $h \geq 3$.

Σημείωση: w_t i.i.d. και $w_t \sim N(0, \sigma^2)$



Θεωρητικά Αποτελέσματα: Γενική Εξίσωση ARIMA (p. d. q)

$$\text{ARIMA}(p, d, q): \varphi(B)\Delta^d X_t = \delta + \theta(B)w_t, \Delta = I - B.$$

π.χ.

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

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

$$\text{ARIMA}(2, 0, 1): (1 - \varphi_1 B - \varphi_2 B^2)X_t = (1 + \theta_1 B)\varepsilon_t \text{ ή}$$

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$



Προσομοίωση ARIMA με τη γλώσσα R

WN `ts.sim <- arima.sim(list(order = c(0,0,0)), n = 500)`

AR(1) `ts.sim <- arima.sim(list(order = c(1,0,0), ar = 0.8), n = 500)`

AR(2) `ts.sim <- arima.sim(list(order = c(2,0,0), ar = c(0.4, 0.5)), n = 500)`

MA(1) `ts.sim <- arima.sim(list(order = c(0,0,1), ma = 0.7), n = 500)`

MA(2) `ts.sim <- arima.sim(list(order = c(0,0,2), ma = c(0.4, 0.5)), n = 500)`

ARIMA(0,1,0) `ts.sim <- arima.sim(list(order = c(0,1,0)), n = 500)`

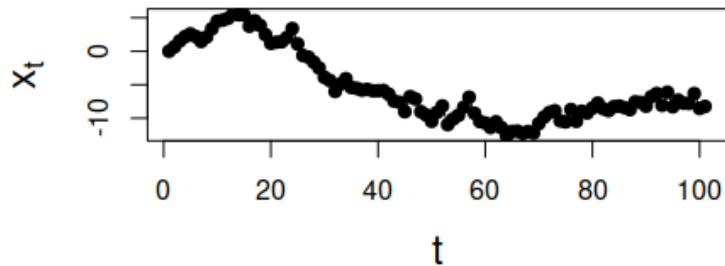
Σημείωση: Παρατηρούμε την αναμενόμενη δομή στα διαγράμματα ACF και PACF.

`par(mfrow=c(3, 1))`

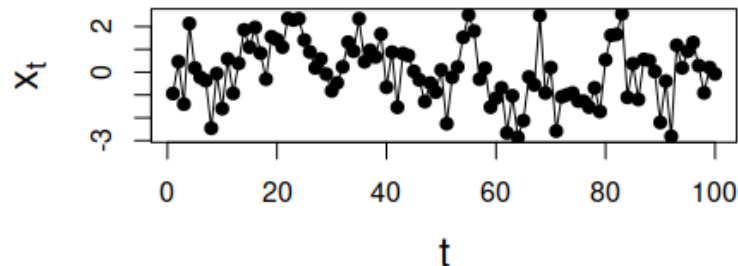
`plot.ts(ts.sim)`

`acf(ts.sim)`

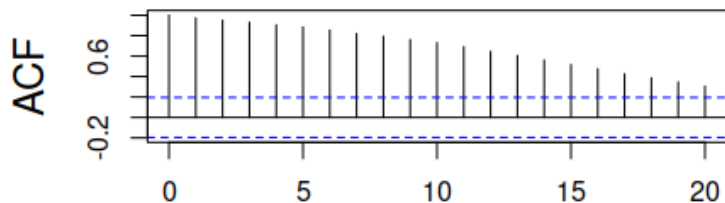
`pacf((ts.sim))`



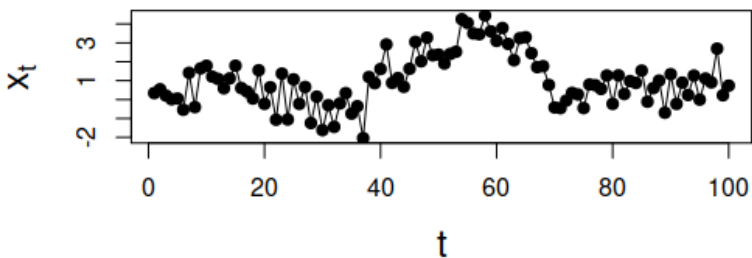
Διάγραμμα 1



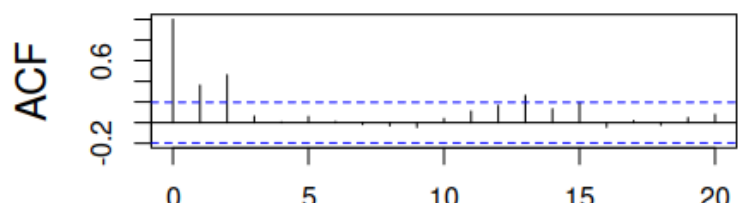
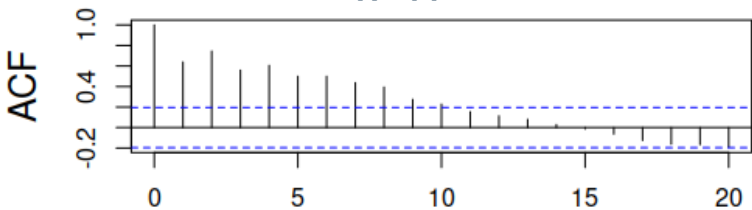
Διάγραμμα 2



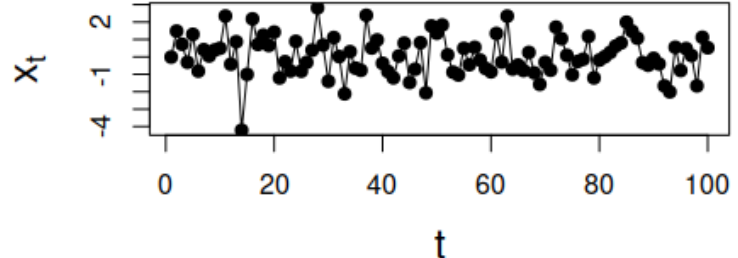
Διάγραμμα 3



Διάγραμμα 3



Διάγραμμα 4



MA(2) (2)

ARIMA(0,1,0) (RW)

AR(2) (3)

WN (4)

Δραστηριότητα

Αντιστοιχίστε τα 4
χρονοδιαγράμματα με τα
αντίστοιχα μοντέλα.



ΔΗΜΟΚΡΙΤΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΡΑΚΗΣ

DEMOCRITUS
UNIVERSITY
OF THRACE

Θεώρημα Αποσύνθεσης του Wold



Στάσιμο $\text{ARMA}(p, q) \rightarrow \text{MA}(\infty)$

Θεώρημα I

Έστω η στάσιμη $\text{ARMA}(p, q)$ σειρά $\varphi(B)X_t = \theta(B)\varepsilon_t$ (ρίζες $\varphi(z) = 0$ εκτός του δίσκου). Τότε υπάρχει μοναδική ακολουθία $\{b_j\} \in l^2$ τέτοια ώστε

$$X_t = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}.$$

Επιπλέον, οι συντελεστές ικανοποιούν την εξίσωση:

$$b_j = \varphi_1 b_{j-1} + \dots + \varphi_p b_{j-p}, \quad j \geq p.$$

Η γενική λύση της τελευταίας ομογενούς γραμμικής εξίσωσης διαφορών είναι $b_j = C_1 \lambda_1^j + \dots + C_m \lambda_m^j$, όπου $\lambda_1, \dots, \lambda_m$ είναι οι αντίστροφοι των ριζών της $\varphi(z) = 0$.

Συνάγουμε το

Πόρισμα

Οι συντελεστές της $\text{MA}(\infty)$ μορφής του $\text{ARMA}(p, q)$ είναι της μορφής $b_j = C_1 \lambda_1^j + \dots + C_m \lambda_m^j$, όπου $\lambda_1, \dots, \lambda_m$ είναι οι αντίστροφοι των ριζών της $\varphi(z) = 0$.



$MA(\infty) \nrightarrow$ Στάσιμο ARMA(p, q)

Παρατηρήσεις

1. Κάθε $MA(\infty)$ σειρά είναι στάσιμη. Πράγματι, αν $Y_t = \sum_{j=0}^{+\infty} b_j \varepsilon_{t-j}$, $\varepsilon_t: WN$, $b_j \in l^2$, από την ανεξαρτησία των ε_t συνάγουμε άμεσα τη στασιμότητα.
2. Θεώρημα I: Κάθε στάσιμη ARMA(p, q) σειρά μπορεί να γραφεί και στη μορφή $MA(\infty)$.
3. Κάθε $MA(\infty)$ σειρά δεν είναι απαραίτητα ARMA(p, q), π.χ. $Y_t = \sum_{j=0}^{+\infty} \frac{1}{j+1} \varepsilon_{t-j}$, $\varepsilon_t: WN$.



Θεώρημα Αποσύνθεσης του Wold

Θεώρημα (Wold)

Κάθε στάσιμη χρονοσειρά μπορεί να γραφτεί ως άθροισμα ενός ντετερμινιστικού μέρους και ενός στοχαστικού μέρους που έχει μοναδική αναπαράσταση $MA(\infty)$:

$$Y_t = D_t + \sum_{j=0}^{+\infty} b_j \varepsilon_{t-j}, \quad \varepsilon_t: WN, \quad b_j \in l^2.$$

Το D_t , είναι ντετερμινιστικό υπό την έννοια πως προσδιορίζεται πλήρως ως γραμμικός συνδυασμός των προηγούμενων τιμών

(π.χ. $D_t = AR(p)$ χωρίς το σφάλμα ή $D_t = A\cos(\omega t) + B\sin(\omega t)$)

Καθώς κάθε στάσιμο $ARMA(p, q)$ έχει αναπαράσταση σε $MA(\infty)$ μορφή, τα μοντέλα αυτά αναδεικνύονται ως καλές επιλογές για την αναπαράσταση κάθε στάσιμης σειράς.



Επιλογή Υποδείγματος - Μέθοδος Box–Jenkins

George Box: “All models are wrong, but some are useful.”

“...in applying mathematics to subjects such as physics or statistics, we make tentative assumptions about the real world which we know are false but which we believe may be useful nonetheless.”

George E. P. Box. (1976). Science and Statistics. Journal of the American Statistical Association, 71(356), 791–799. <https://doi.org/10.2307/2286841>

Ένα μοντέλο ARMA μπορεί να είναι ή να μην είναι ο πραγματικός μηχανισμός πίσω από τη γέννηση μίας σειράς. Ωστόσο, είναι::

- αρκετά ευέλικτο για short-memory σειρές.
- αρκετά απλό για estimation.
- αρκετά αποδοτικό για πρόβλεψη.



Επιλογή Υποδείγματος - Μέθοδος Box–Jenkins

Η μεθοδολογία Box – Jenkins περιγράφεται από τρία βασικά στάδια, τα οποία είναι:

- Η ταυτοποίηση (identification)

Στο στάδιο της ταυτοποίησης γίνεται ο καθορισμός των τιμών p , d και q του υποδείγματος ARIMA (p,d,q) με βάση τις πληροφορίες που προέρχονται από το δείγμα (ACF, PACF).

- Η εκτίμηση (estimation)

Μετά το στάδιο της ταυτοποίησης, ακολουθεί το στάδιο της εκτίμησης των συντελεστών του υποδείγματος ARIMA(p,d,q).

- Ο διαγνωστικός έλεγχος (diagnostic checking)

Το υπόδειγμα ελέγχεται για την καταλληλότητά του με βάση τους παρακάτω άξονες:

- Σημαντικότητα των συντελεστών του υποδείγματος
- Ιδιότητες των υπολοίπων
- Προβλεπτική ικανότητα του υποδείγματος

Αν το υπόδειγμα θεωρηθεί κατάλληλο, δηλαδή περιγράφει ικανοποιητικά τη διαδικασία από την οποία προέρχονται τα δεδομένα, τότε τα υπόλοιπα θα πρέπει να συμπεριφέρονται ως μια διαδικασία λευκού θορύβου.

- Αν αποτύχει ο διαγνωστικός έλεγχος, επιστρέφουμε στο στάδιο της ταυτοποίησης και τροποποιούμε το μοντέλο.



Επιλογή Υποδείγματος - Μέθοδος Box-Jenkins

1ο Στάδιο: Ταυτοποίηση.

Σε αυτό γίνεται μια προσπάθεια να αναγνωριστούν τα πιθανά υποδείγματα που ταιριάζουν με τα δεδομένα. Η αναγνώριση γίνεται μέσα από τα διαγράμματα αυτοσυσχέτισης και μερικής αυτοσυσχέτισης.

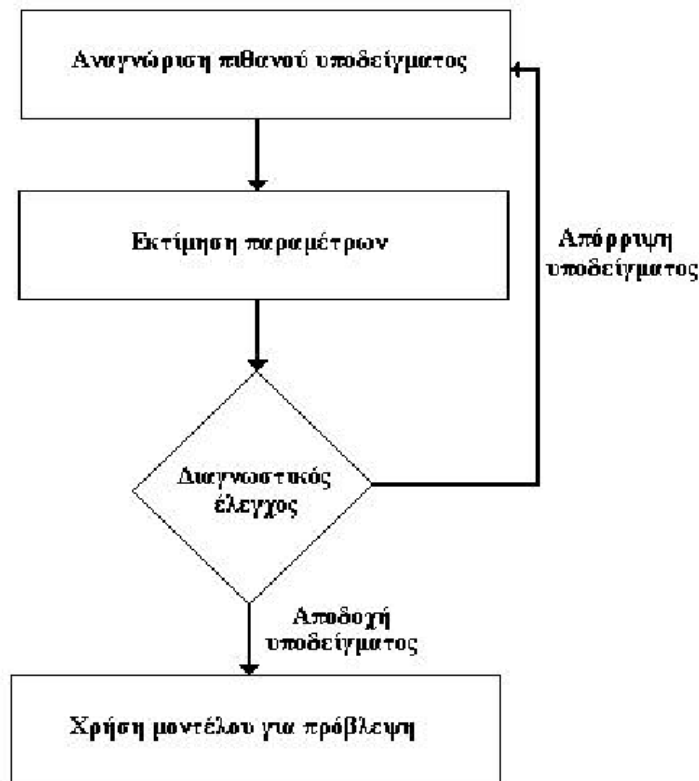
2ο Στάδιο: Εκτίμηση και δοκιμή.

Στο στάδιο αυτό γίνεται η εκτίμηση των παραμέτρων του υποδείγματος που αναγνωρίστηκε. Στη συνέχεια γίνεται η δοκιμή, ο έλεγχος προσαρμογής και ικανότητας πρόβλεψης του υποδείγματος. Σε περίπτωση μη αποδοχής του υποδείγματος, επαναλαμβάνεται η διαδικασία για να επιλεγεί ένα πιο κατάλληλο υπόδειγμα.

3ο Στάδιο: Χρήση.

Σε περίπτωση που το υπόδειγμα γίνει αποδεκτό ακολουθεί το τρίτο στάδιο, δηλαδή η χρήση του μοντέλου για πρόβλεψη.

Βήματα ανάλυσης για την κατασκευή μοντέλων ARIMA





Ταυτοποίηση: Ενα δύσκολο πρόβλημα

Η αναγνώριση του μοντέλου μίας χρονοσειράς, δεν είναι εύκολη υπόθεση. Υπάρχει περίπτωση το μοντέλο να έχει δομή ARIMA αλλά η ανταγωνιστική φύση των όρων τους να προσδιορίζει τη σειρά ως Λευκό Θόρυβο.

Ενδεικτικά, παρουσιάζεται το παρακάτω παράδειγμα ενός μοντέλου ARIMA(1, 0, 1) το οποίο όμως δεν μπορεί να αναγνωρισθεί από τα διαγράμματα ACF και PACF.

```
x = arima.sim(list(order = c(1,0,1), ar = 0.6, ma = c(-0.5)), n = 500)
par(mfrow=c(3, 1))
plot.ts(x)
acf(x)
pacf(x)
```



Ελεγχος ότι η σειρά αποτελείται από εξαρτημένες παρατηρήσεις

Πριν την προσαρμογή ενός μοντέλου, πρέπει να ελέγξουμε ότι η σειρά δεν αποτελείται από τυχαίες τιμές χωρίς εξάρτηση μεταξύ τους.

Για το σκοπό αυτό, μπορούμε να εφαρμόσουμε τη δοκιμασία Box–Pierce, την όμοια Ljung–Box ή ακόμα και την Breusch–Godfrey.

Με αυτές ελέγχεται η στατιστική υπόθεση:

H_0 : Οι τιμές της χρονοσειράς είναι NWN.

έναντι της ερευνητικής υπόθεσης:

H_1 : όχι η H_0 .



Ελεγχος ότι η σειρά αποτελείται από εξαρτημένες παρατηρήσεις

```
set.seed(112)
sample.wn = ts(rnorm(100))
plot(sample.wn, type = 'l')
Box.test(sample.wn, lag = 20, type = "Ljung-Box")
Box.test(sample.wn, lag = 20, type = "Box-Pierce")
```

```
ts.value = c(-0.43,-0.92,-0.31,0.09,-0.54,3.35,-3.43,2.32,0.37,-2.65,-0.6,-0.43,0.27,-0.63,0.52,-
0.25,0.78,2.12,-1.06,-0.44,-1.91,1.97,-1.18,-0.73,-0.95,2.9,-0.21,-2.39,1.04,-0.45,1.95,-0.64,0.02,-
0.71,1.3,-0.61,0.86,-1.87,0.77,-0.14,1.88,-0.93,-0.62,0.59,-1.26,2.66,-1.52,0.88,-0.74,-0.51,2.39,-0.48,-
1.01,0.15,-0.68,1.78,0.2,-2.4,2.32,-2.97,1.66,0.46,0.47,2.05,-2.41,-0.84,1.48,0.18,-0.99,0.6,-0.31,1.92,-
2.72,1.7,1.02,-0.13,-0.68,1.18,-1.79,0.21,0.57,1.28,-0.72,-0.42,0.66,-0.12,0.81,-2.21,1.09,-0.24,0,-
0.37,0.72,-1.48,0.68,0.54,-1.13,1.07,-0.81)
plot(ts.value, type = 'l')
Box.test(ts.value, lag = 20, type = "Ljung-Box")
Box.test(ts.value, lag = 20, type = "Box-Pierce")
```



Η συνάρτηση `forecast::auto.arima`

Η συνάρτηση `auto.arima` της βιβλιοθήκης `forecast` προσφέρει τη δυνατότητα αυτόματης επιλογής ενός μοντέλου βάσει των δεικτών προσαρμογής AIC, AICc και BIC. Είναι ένα σημαντικό εργαλείο, ωστόσο δεν αποδίδει πάντα τα σωστά αποτελέσματα, όπως δείχνουν τα επόμενα παραδείγματα:

Παράδειγμα: Σειρά ARMA(1, 1) που αναγνωρίζεται ως ARMA(1, 5)

```
set.seed(11101)
ts.sim = arima.sim(list(order = c(1,0,1), ar = 0.6, ma = c(1.4)), n = 500)
auto.arima(ts.sim)
```

Παράδειγμα: Σειρά ARMA(0, 3) που αναγνωρίζεται ως ARMA(3, 2)

```
set.seed(11101)
ts.sim = arima.sim(list(order = c(0,0,3), ma = c(1.4, 1, 0.5)), n = 500)
auto.arima(ts.sim)
```

Γραφική Αναπαράσταση

```
par(mfrow=c(3, 1))
plot.ts(ts.sim)
acf(ts.sim)
pacf((ts.sim))
```



Η συνάρτηση forecast::auto.arima

Ωστόσο, είναι επίσης πιθανό να αναγνωρίσει την πραγματική δομή της σειράς, όπως στο παρακάτω παράδειγμα

```
set.seed(1234)
```

```
ts.sim = arima.sim(list(order = c(1,0,1), ar = 0.6, ma = c(0.4)), n = 500)
```

```
auto.arima(ts.sim)
```

Γραφική Αναπαράσταση

```
par(mfrow=c(3, 1))
```

```
plot.ts(ts.sim)
```

```
acf(ts.sim)
```

```
pacf((ts.sim))
```



Η συνάρτηση forecast::Arima

Για την προσαρμογή ενός ARMA(p, q) μοντέλου σε μία χρονοσειρά, θα αξιοποιηθεί η συνάρτηση Arima από τη βιβλιοθήκη forecast. Στα πλαίσια της βιβλιοθήκης αυτής, όταν αναφέρεται σταθερά ενός μοντέλο AR(p) τότε υποτίθεται ότι η σταθερά αφαιρείται από την τιμή της χρονοσειράς κάθε φορά.

Για παράδειγμα, ένα μοντέλο AR(1) με σταθερά, έχει την εξής εξίσωση:

$$(x_t - c) = \varphi(x_{t-1} - c) + e_t$$

Στην περίπτωση ενός MA(q) μοντέλου η σταθερά είναι η μέση τιμή.

Ο λόγος για την συγκεκριμένη επιλογή διαχείρισης της σταθεράς είναι πως καθιστά πιο εύκολη την ερμηνεία της στην περίπτωση όπου υπάρχουν και άλλες επεξηγηματικές μεταβλητές στο μοντέλο (όπως π.χ. σε ένα μοντέλο XARIMA).

Ο δημιουργός της βιβλιοθήκης forecast επεξηγεί τις θεωρητικές λεπτομέρειες στο blog του:

<https://robjhyndman.com/hyndsight/arimax/>



Η συνάρτηση forecast::Arima

Με την εντολή print, το output της εντολής Arima μεταξύ άλλων αναφέρει:

Τη διακύμανση σ^2 της ακολουθίας wt.

Το λογάριθμο της πιθανοφάνειας log likelihood.

και τους παρακάτω δείκτες προσαρμογής:

Τον δείκτη AIC (Akaike information criterion).

Τον δείκτη AICc, ο οποίος είναι μια διορθωμένη ως προς το μέγεθος του δείγματος εκδοχή του AIC.

Τον δείκτη BIC (Bayesian information criterion).

Σημείωση: Δεν υπάρχει βέλτιστος δείκτης απόδοσης ενός μοντέλου. Περισσότερα σχόλια μπορούν να βρεθούν εδώ:

<https://stats.stackexchange.com/questions/577/is-there-any-reason-to-prefer-the-aic-or-bic-over-the-other>



Κριτήριο AIC (Akaike information criterion)

Το κριτήριο πληροφοριών Akaike (AIC) (Hirotugu Akaike(1927 – 2009)) είναι ένας εκτιμητής του σφάλματος πρόβλεψης ενός μοντέλου. Το AIC εκτιμά την ποιότητα ενός μοντέλου σε σχέση με άλλα υποψήφια μοντέλα, λαμβάνοντας υπόψη τόσο την καλή προσαρμογή (fit) όσο και την πολυπλοκότητα (complexity) του μοντέλου. Ο τύπος είναι:

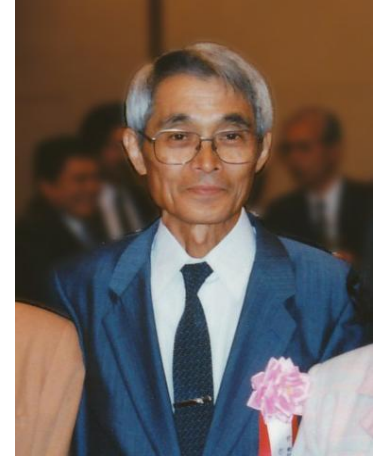
$$AIC = -2\ln(L) + 2k$$

όπου:

L: Η μέγιστη πιθανότητα (likelihood) του μοντέλου.

k: Ο αριθμός των παραμέτρων του μοντέλου.

Μεγαλύτερη τιμή πιθανοφάνειας, αντιστοιχεί σε μικρότερη τιμή του AIC.



Hirotugu Akaike(1927 – 2009)



AICc και BIC

Το AICc (Akaike information criterion – corrected) είναι μια διορθωμένη έκδοση του AIC που λαμβάνει υπόψη μικρά δείγματα. Ο τύπος του είναι:

$$\text{AICc} = \text{AIC} + \frac{2k(k+1)}{n - k - 1}$$

Για μεγάλα δείγματα το AICc συγκλίνει στο AIC. Είναι προτιμότερο από το AIC όταν το μέγεθος δείγματος n είναι συγκρίσιμο ή μικρότερο από τον αριθμό των παραμέτρων k .

Το BIC (Bayesian Information Criterion) βασίζεται σε μια Bayesian προσέγγιση και περιλαμβάνει έναν αυστηρότερο όρο για την πολυπλοκότητα του μοντέλου. Ο τύπος του είναι:

$$\text{BIC} = -2\ln(L) + k \cdot \ln(n)$$

Όσο μικρότερη η τιμή του BIC τόσο καλύτερη η προσαρμογή του μοντέλου στα δεδομένα.



Παράδειγμα

```
set.seed(1234)
```

```
ts.sim <- arima.sim(list(order = c(1,0,1), ar = 0.7, ma = -0.6), n = 300)
```

```
model1 = Arima(ts.sim,order=c(1,0,0))
```

```
model2 = Arima(ts.sim,order=c(2,0,0))
```

```
model3 = Arima(ts.sim,order=c(1,0,1))
```

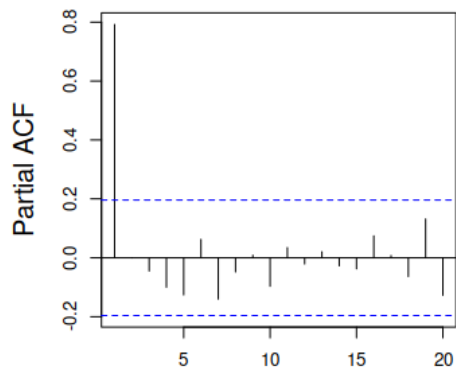
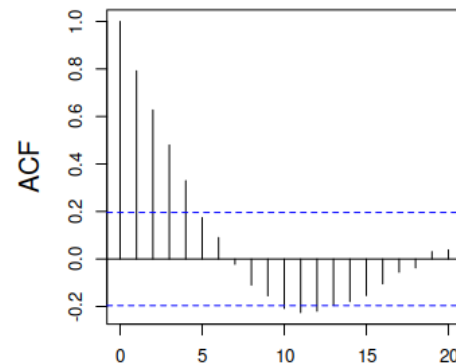
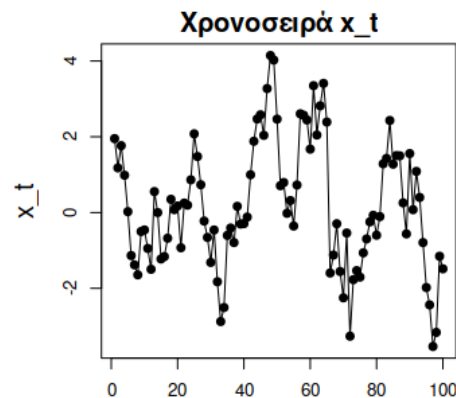
```
model4 = Arima(ts.sim,order=c(0,0,2))
```

```
print(model1$aic)
```

```
print(model2$aic)
```

```
print(model3$aic)
```

```
print(model4$aic)
```



```
> print(model1$aic)
```

```
[1] 854.7152
```

```
> print(model2$aic)
```

```
[1] 854.8835
```

```
> print(model3$aic)
```

```
[1] 854.2488
```

```
> print(model4$aic)
```

```
[1] 855.5619
```




Η συνάρτηση Arima

Αν εκτελεστεί με την εντολή `summary`, παρουσιάζει επιπλέον τα παρακάτω στατιστικά που περιγράφουν την ποιότητα προσαρμογής του μοντέλου στα ίδια τα δεδομένα.

- Το μέσο σφάλμα (ME: Mean Error)
- Το μέσο τετραγωνικό σφάλμα (RMSE: root mean squared error)
- Το μέσο απόλυτο σφάλμα (MAE: mean absolute error)
- Το μέσο ποσοστιαίο σφάλμα (MPE: mean percentage error)
- Το μέσο απόλυτο ποσοστιαίο σφάλμα (MAPE: mean absolute percentage error)
- Το μέσο απόλυτο τυποποιημένο σφάλμα (MASE: mean absolute scaled error)
- Την αυτοσυσχέτιση 1ης τάξης (ACF1).



Αξιολόγηση της προσαρμογής

Μετά τον υπολογισμό των συντελεστών του μοντέλου υπολογίζονται τα in-sample σφάλματα ε_t . Από αυτά υπολογίζονται τα παρακάτω μέτρα προσαρμογής:

$$\text{Μεσο Σφάλμα } ME = \frac{1}{n} \sum_{i=1}^n \varepsilon_i$$

$$\text{Ρίζα Μέσου Τετραγωνικού Σφάλματος } RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \varepsilon_i^2}$$

$$\text{Μεσο Απόλυτο Σφάλμα } MAE = \frac{1}{n} \sum_{i=1}^n |\varepsilon_i|$$

$$\text{Μεσο Ποσοστιαίο Σφάλμα } MPE = \frac{1}{n} \sum_{i=1}^n \frac{\varepsilon_i}{x_i} 100$$

$$\text{Μεσο Απολυτο Ποσοστιαίο Σφάλμα } MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|\varepsilon_i|}{x_i} 100$$

$$\text{Μεσο Απολυτο Τυποποιημένο (Scaled) Σφάλμα } MASE = \frac{MAE_{\text{model}}}{MAE_{\text{naive}}}, \quad MAE_{\text{naive}} = \frac{1}{n-1} \sum_{i=1}^n |\Delta x_t|$$



ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΡΑΚΗΣ

DEMOCRITUS
UNIVERSITY
OF THRACE

Πρόβλεψη



Πρόβλεψη

Γνωρίζοντας τα στοιχεία $X_{1:n} = X_1, X_2, \dots, X_n$, θα εκτιμήσουμε την X_{n+k} . Την πρόβλεψη της επόμενης τιμής τη συμβολίζουμε $X_{n+k|n}$ και προσδιορίζεται μοναδικά ως συνάρτηση των τιμών $X_{1:n} = (X_1, X_2, \dots, X_n)$.

Ως βέλτιστη πρόβλεψη επιλέγουμε εκείνη που ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα: $E[(X_{n+k} - X_{n+k|n})^2 | X_{1:n}]$.

Είναι

$$\begin{aligned} E[(X_{n+k} - X_{n+k|n})^2 | X_{1:n}] &= E[X_{n+k}^2 - 2 X_{n+k} X_{n+k|n} + X_{n+k|n}^2 | X_{1:n}] \\ &= E(X_{n+k}^2 | X_{1:n}) - 2 E(X_{n+k} X_{n+k|n} | X_{1:n}) + E(X_{n+k|n}^2 | X_{1:n}) \\ &= E(X_{n+k}^2 | X_{1:n}) - 2 X_{n+k|n} E(X_{n+k} | X_{1:n}) + X_{n+k|n}^2 \end{aligned}$$



Πρόβλεψη

$$X_{n+k|n}^2 - 2 X_{n+k|n} E(X_{n+k} | X_{1:n}) + E(X_{n+k}^2 | X_{1:n}) = E[(X_{n+k} - X_{n+k|n})^2 | X_{1:n}]$$

Η τελευταία σχέση είναι τριώνυμο ως προς $X_{n+k|n}$.

Υπολογίζοντας την παράγωγο (ως προς $X_{n+k|n}$) και εξισώνοντας με το 0, βρίσκουμε ότι η ελάχιστη τιμή του λαμβάνεται όταν:

$$X_{n+k|n} = E(X_{n+k} | X_{1:n}).$$

Δηλαδή, ο MSE εκτιμητής της τιμής X_{n+k} είναι η τιμή $E(X_{n+k} | X_{1:n})$



Σφάλμα Πρόβλεψης

Το σφάλμα της πρόβλεψης είναι:

$$\text{Error}(k) = X_{n+k} - X_{n+k|n}$$

Το σφάλμα $\text{Error}(k)$ είναι μία τυχαία μεταβλητή. Παρατηρούμε ότι

$$E(X_{n+k} - X_{n+k|n} \mid X_{1:n}) = E(X_{n+k} \mid X_{1:n}) - X_{n+k|n} = 0$$

$$\text{Var}(X_{n+k} - X_{n+k|n} \mid X_{1:n}) = \text{Var}(X_{n+k} \mid X_{1:n})$$



Πρόβλεψη AR(1)

Με MLE (ή LS) διαδικασία εκτιμώνται από τα δεδομένα τα δ , φ και σ^2 του AR(1) μοντέλου

$$X_t = \delta + \varphi X_{t-1} + w_t, \quad w_t \sim WN(0, \sigma^2).$$

Η βέλτιστη (MSE) πρόβλεψη της $n+1$ τιμής από τις πρώτες n είναι η

$$\begin{aligned} X_{n+1|n} &= E(X_{n+1} \mid X_{1:n}) \\ &= E(\delta + \varphi X_n + w_{n+1} \mid X_{1:n}) \\ &= \delta + \varphi E(X_n \mid X_{1:n}) \\ &= \delta + \varphi X_n \end{aligned}$$

Το σφάλμα της πρόβλεψης είναι $X_{n+1} - X_{n+1|n} = w_{n+1}$ και η διακύμανσή του είναι σ^2 .



Πρόβλεψη AR(1)

Η εκτίμηση της $n+2$ τιμής από τις πρώτες n είναι

$$X_{n+2|n} = E(X_{n+2} | X_{1:n}) = E(\delta + \varphi X_{n+1} + w_{n+2} | X_{1:n}) = \delta + \varphi X_{n+1|n} = \delta(1 + \varphi) + \varphi^2 X_n$$

Το σφάλμα της εκτίμησης είναι $w_{n+2} + \varphi w_{n+1}$ με διακύμανση $(1 + \varphi^2)\sigma^2$.

Αντίστοιχα, ακολουθεί ο υπολογισμός των $X_{n+k|n} = \delta(1 + \varphi + \varphi^2 + \dots + \varphi^{k-1}) + \varphi^k X_n$, $k \geq 2$.
Παρατηρούμε ότι:

- Το AR(1) έχει μη τετριμμένες προβλέψεις για όλα τα k .
- Η επίδραση της τελευταίας παρατήρησης φθίνει γεωμετρικά.
- Οι προβλέψεις συγκλίνουν στη μέση τιμή μ .



Πρόβλεψη MA(1)

Με την MLE διαδικασία από τα δεδομένα εκτιμώνται τα δ , θ και σ^2 του MA(1) μοντέλου

$$X_t = \delta + \theta w_{t-1} + w_t, w_t \sim WN(0, \sigma^2).$$

Στη συνέχεια, από το μοντέλο εκτιμούνται αναδρομικά οι w_t , $t = 1, 2, \dots, n$.

Η εκτίμηση της $n+1$ τιμής από τις πρώτες n είναι η

$$\begin{aligned} X_{n+1|n} &= E(X_{n+1} \mid X_{1:n}) \\ &= E(\delta + \theta w_n + w_{n+1} \mid X_{1:n}) \\ &= \delta + \theta E(w_n \mid X_{1:n}) + E(w_{n+1} \mid X_{1:n}) \\ &= \delta + \theta w_n \end{aligned}$$

Το σφάλμα της πρόβλεψης είναι $X_{n+1} - X_{n+1|n} = w_{n+1}$ και η διακύμανσή του είναι σ^2 .



Πρόβλεψη MA(1)

Η πρόβλεψη της $n+2$ τιμής από τις πρώτες n είναι η $X_{n+2|n} = \delta + \theta w_{n+1}$.

Δεν υπάρχει πληροφορία για το w_{n+1} , άρα η πιο απλή εκτίμησή του είναι η $E(w_{n+1}) = 0$.

Δηλαδή, τελικά: $X_{n+2|n} = \delta$.

Αντίστοιχα, $X_{n+k|n} = \delta$ για κάθε $k \geq 2$.



Πρόβλεψη MA(2)

Με την MLE διαδικασία από τα δεδομένα εκτιμώνται τα δ , θ_1 , θ_2 και σ^2 του MA(1) μοντέλου

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

Στη συνέχεια, από το μοντέλο εκτιμούνται αναδρομικά οι w_t , $t = 1, 2, \dots, n$.

Η βέλτιστη (MSE) πρόβλεψη της $n+1$ τιμής από τις πρώτες n είναι η $X_{n+1|n} = \delta + \theta_1 w_n$



Πρόβλεψη MA(2)

Πρόβλεψη X_{n+2} .

Δεν υπάρχει πληροφορία για το w_{n+1} , άρα η πιο απλή εκτίμησή του είναι η $E(w_{n+1}) = 0$.

Αντίστοιχα, η πρόβλεψη της $n+2$ τιμής από τις πρώτες n είναι η

$$X_{n+2|n} = \delta + \theta_1 w_{n+1} + \theta_2 w_n = \delta + \theta_2 w_n$$

Ανάλογα, $X_{n+k|n} = \delta$ για κάθε $k \geq 3$.

Γενικότερα:

Ένα MA(q) παράγει μη τετριμμένες εκτιμήσεις για τις πρώτες q προβλέψεις.



Πρόβλεψη ARMA(p, q)

Με την MLE διαδικασία από τα δεδομένα εκτιμώνται τα δ , ϕ_i , θ_j , σ^2 του ARMA(p, q) μοντέλου

$$X_t = \delta + \sum \phi_i X_{t-i} + \sum \theta_j w_{t-j} + w_t, w_t \sim WN(0, \sigma^2).$$

Στη συνέχεια, από το μοντέλο εκτιμούνται αναδρομικά οι w_t , $t = 1, 2, \dots, n$.

Η βέλτιστη (MSE) πρόβλεψη της $n+1$ τιμής από τις πρώτες n είναι η

$$\hat{X}_{n+1|n} = \delta + \sum \phi_i X_{t-i} + \sum \theta_j w_{t-j}.$$

Στη συνέχεια:

- Το MA(q) μέρος επηρεάζει τη μέση πρόβλεψη μόνο μέχρι q βήματα μπροστά.
- Το AR(p) μέρος επηρεάζει τη μέση πρόβλεψη για όλα τα μελλοντικά βήματα.



Παράδειγμα Ι

ts.value = c(-0.43,-0.92,-0.31,0.09,-0.54,3.35,-3.43,2.32,0.37,-2.65,-0.6,-0.43,0.27,-0.63,0.52,-0.25,0.78,2.12,-1.06,-
0.44,-1.91,1.97,-1.18,-0.73,-0.95,2.9,-0.21,-2.39,1.04,-0.45,1.95,-0.64,0.02,-0.71,1.3,-0.61,0.86,-1.87,0.77,-
0.14,1.88,-0.93,-0.62,0.59,-1.26,2.66,-1.52,0.88,-0.74,-0.51,2.39,-0.48,-1.01,0.15,-0.68,1.78,0.2,-2.4,2.32,-
2.97,1.66,0.46,0.47,2.05,-2.41,-0.84,1.48,0.18,-0.99,0.6,-0.31,1.92,-2.72,1.7,1.02,-0.13,-0.68,1.18,-
1.79,0.21,0.57,1.28,-0.72,-0.42,0.66,-0.12,0.81,-2.21,1.09,-0.24,0,-0.37,0.72,-1.48,0.68,0.54,-1.13,1.07,-0.81,-0.28,-
1.33,1.19,-0.2,0.94,0.36,1.7,-3.26,0.96,-0.87,1.23,0.33,-0.06,0.18,-0.25,-1.27,1.29,-0.48,2.04,-1.31,0.69,-
0.17,1.38,0.32,-1.64,0.3,-0.02,-1.37,0,0.24,-0.37,-2.33,2.05,0.51,-0.16,-0.27,0.07,0.92,-0.57,-1.16,1.01,-0.85,-
0.55,0.51,0.57,0.02,-1.2,1.63,-0.92,-1.27,1.55,-1.59,1.41,1.73,-2.39,1.63,-0.53,1.87,0.07,-2.63,0.65,-0.01,0.97,-
1.59,1.55,0.67,0.49,-1.1,-0.31,1.49,-0.73,-0.75,-1.24,1.97,0.23,0.69,-0.83,-0.04,-1.46,-1.36,0.55,1.46,1.29,-2.16,-
2.26,1.15,0.37,-1.13,2.84,-2.6,0.68,1.6,-0.56,-0.35,-0.11,0.11,0.67,-1.15,-1.03,0.54,-0.15,-1.12,1.2,0.46,-1.58,1.08,-
1.98,-0.29,0.97,1.91,-0.24,-0.03,-0.62,-1.22,0.6,1.35,-1.28,0.86,-1.19,-0.35,-0.15,2.02,-0.56,0.72,-2.6,1.59,-0.18,-
1.33,2.84,-1.09,-0.48,1.53,-1.6,-1.87,1.45,-0.06,0.08,1.01,-0.92,-2.51,2.03,-1.06,0.52,1.2,-0.2,-0.5,-1.07,2.44,-
1.33,1.28,-1.63,-0.22,-0.43,1.61,-1.07,1.5,1.42,-1.61,-0.52,1.49,-0.03,-0.28,-0.71,1.26,-0.79,0.9,1.19,-
2.46,1.51,0.19,0.55,-1.14,-0.24,0.53,0.61,-0.22,1.36,0.05,-0.77,0.08,0.24,-0.13,1.46,-1.83,1.42,-1.55,-
0.43,0.02,0.66,1.05,-2.06,4.25,-3.06,1.57,-1.45,1.68,-0.26,-1.64,0.85,1.59,-1.26)



Παράδειγμα I

```
par(mfrow=c(3, 1)); plot.ts(ts.value); acf(ts.value); pacf(ts.value)
```

Παρατηρούμε ότι μόνο η αυτοσυσχέτιση 1ης τάξης είναι στατιστικώς σημαντική ($\rho_1 = -0.470$), ενώ όλες οι υπόλοιπες δεν διαφέρουν σημαντικά από το 0.

Στα πλαίσια ενός μοντέλου ARIMA, η παρατήρηση αυτή υποδεικνύει το μοντέλο MA(1) ως ένα μοντέλο που μπορεί να προσαρμοστεί καλά στη χρονοσειρά.

Προχωρούμε στον υπολογισμό του μοντέλου με την εντολή:

```
library(forecast)  
my.model = Arima(ts.value, order = c(0, 0, 1))  
print(my.model)
```



Παράδειγμα I

Το MA(1) μοντέλο που υπολογίζεται είναι το:

$$X_t = 0.0074 + w_t - 0.7331 w_{t-1}.$$

Σε ένα μοντέλο MA(1) η σταθερά είναι η μέση τιμή της χρονοσειράς. Παρατηρούμε ότι $E(X_t) = 0.0074 \approx 0$. Επιδιώκοντας ένα απλούστερο μοντέλο, υπολογίζουμε εκ νέου το μοντέλο χωρίς την σταθερά:

```
library(forecast)
my.model = Arima(ts.value, order = c(0, 0, 1), include.constant = FALSE)
print(my.model)
```

Παρατηρούμε, ότι το MA(1) μοντέλο χωρίς σταθερά που υπολογίζεται είναι το:

$$X_t = w_t - 0.7319 w_{t-1}.$$



Παράδειγμα I

Έχουμε τη δυνατότητα να δημιουργήσουμε προβλέψεις για τις επόμενες τιμές της χρονοσειράς με την εντολή:

```
my.forecast = forecast(my.model)  
summary(my.forecast)
```

Μία γραφική αναπαράσταση των προβλέψεων μπορούμε να αποκτήσουμε με την εντολή:

```
plot(my.forecast, showgap = FALSE)
```



Παράδειγμα I

Φανερά, οι προβλέψεις δεν είναι πολύ ικανοποιητικές. Ωστόσο, αυτό είναι που μπορεί να μας δώσει ένα τόσο απλό μοντέλο.

Επαληθεύουμε την καλή προσαρμογή του στη σειρά, εξετάζοντας τα υπόλοιπα των προβλέψεων, με την εντολή:

```
checkresiduals(my.model)
```

Παρατηρούμε, ότι τα υπόλοιπα είναι κανονικά κατανεμημένα και πως δεν έχουν μεταξύ τους αυτοσυσχέτιση, γεγονός που επιβεβαιώνει πως το MA(1) μοντέλο προσαρμόζεται ικανοποιητικά στα δεδομένα της σειράς.



Δραστηριότητα I

```
ts.value = c(0,-1.23,-0.37,-1.3,-1.7,-1.36,-2.3,-1.55,-1.46,-2.83,-0.77,-4.43,-3.2,-3.75,-3.82,-4.98,-3.9,-4.34,-4.07,-  
3.11,-5.23,-1.98,-4.83,-3.37,-4.9,-4.21,-4.41,-3.38,-6.26,-3.23,-6.13,-1.85,-5.28,-2.75,-5.3,-4.34,-5.72,-3.74,-5.48,-  
4.41,-5.09,-3.02,-4.08,-3.15,-3.97,-2.42,-3.12,-3.69,-2,-3.82,-3.96,-3.03,-3.98,-2.85,-3.07,-3.08,-3.65,-3.77,-5.29,-  
6.35,-4.59,-6.73,-3.44,-5.98,-4.07,-6.35,-4.41,-5.31,-3.96,-5.15,-5.76,-3.91,-4.61,-5.49,-5.28,-6.18,-5.8,-6.29,-5.39,-  
6.29,-6.61,-5.11,-4.41,-3.36,-4.89,-4.09,-4.33,-2.66,-3.3,-3.81,-3.78,-2.64,-4.77,-3.12,-4.6,-3.28,-5.29,-2.48,-4.73,-  
3.65,-3.85,-3.04,-3.81,-5.01,-3.52,-5.41,-4.75,-3.66,-5.13,-3.71,-4.91,-2.18,-5.36,-2.79,-5.78,-2.62,-3.88,-3.57,-4.05,-  
3.3,-4.34,-3.76,-3.07,-3.8,-3.81,-5.11,-3.09,-4.86,-4.19,-5.71,-3.31,-5.29,-1.91,-4.85,-3.49,-4.68,-3.29,-3.71,-2.78,-  
3.88,-1.61,-3.61,-1.54,-3.7,-1.39,-2.4,-1.32,-0.28,-2.67,-1.02,-3.64,0.1,-3.63,0.58,-4.48,0.66,-2.61,-0.94,-2.57,-1.72,-  
0.74,0.23,0.58,0.28,0.61,2.05,0.92,1.45,0.14,0.19,0.37,-1.01,-1.22,-1.3,-2.82,-1.21,-2.75,-1.75,-2.58,-0.52,-3.35,-  
1.52,-1.48,-0.94,-0.03,0.25,-0.4,-1,0.23,-0.44,1.71,-0.95,1.64,1.38,0.96,2.5,0.39,2.5,0.27,2.05,0.73,1.03,0.19,1.61,-  
0.51,0.14,-0.72,-0.54,-1.81,-2.5,-1.91,-2.06,-1.89,-3.32,-1.13,-4.16,-1.02,-3.28,-2.07,-2.75,-2.5,-2.19,-1.61,-0.41,-2.2,-  
2.97,-2.55,-4.27,-4.23,-4.54,-5.38,-3.36,-6.76,-3.61,-6,-4.44,-7.21,-2.91,-6.53,-3.96,-7.04,-3.25,-5.66,-3.31,-6.06,-  
6.17,-5.96,-5.73,-5.03,-6.12,-4.71,-7.42,-4.05,-7.25,-3.54,-4.45,-4.09,-4.18,-6.42,-3.92,-5.66,-2.62,-4.15,-3.07,-4.99,-  
3.87,-4.65,-3.93,-4.7,-1.71,-5.45,-2.03,-3.25,-3.56,-4.13,-2.08,-2.6,-0.97,-1.26,-1.04,-0.35,0.36,-0.75,-0.88,-1.7,-0.52,-  
0.76,-0.17,0.21,1.15,0.51,2.41,-0.14,2.86,0.08,2.3,-0.32,3.33,-0.61,2.73,0.66)
```

Υπόδειξη: Δοκιμάστε την προσαρμογή τόσο χωρίς όσο και με διαφοροποίηση.



ΔΗΜΟΚΡΙΤΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΡΑΚΗΣ

DEMOCRITUS
UNIVERSITY
OF THRACE

Μέθοδοι Υπολογισμού των Συντελεστών



Μέθοδοι Υπολογισμού των Συντελεστών

Για τον υπολογισμό των συντελεστών ενός μοντέλου που προσαρμόζεται σε μία χρονοσειρά, υπάρχουν πολλές μέθοδοι. Οι κυριότερες είναι οι εξής:

- Μέθοδος ελαχίστων τετραγώνων (Least Squares - LS).
- Εκτιμητής μέγιστης πιθανοφάνειας (Maximum Likelihood Estimator – MLE).



ΔΗΜΟΚΡΙΤΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΡΑΚΗΣ

DEMOCRITUS
UNIVERSITY
OF THRACE

Μέθοδος των ελαχίστων τετραγώνων



Υπολογισμός συντελεστών 1/3

Έστω (x_i, y_i) , $i = 1, 2, \dots, n$ τα n ζεύγη παρατηρήσεων και έστω

$$\hat{y} = \hat{\alpha} + \hat{b}x$$

η εξίσωση της ευθείας που αναζητούμε. Τα σφάλματα της εκτίμησης για κάθε ένα x_i , $i = 1, 2, \dots, n$, είναι:

$$\text{resid}_i(\hat{\alpha}, \hat{b}) = y_i - \hat{y}_i = y_i - \hat{\alpha} - \hat{b}x_i$$

Αναζητούμε την εξίσωση ευθείας που ελαχιστοποιεί το άθροισμα τετραγώνων αυτών των σφαλμάτων.

$$\hat{\alpha}, \hat{b} = ;, \text{ ώστε } S(\hat{\alpha}, \hat{b}) = \sum_{i=1}^n \text{resid}_i^2(\hat{\alpha}, \hat{b}) = \sum_{i=1}^n (y_i - \hat{\alpha} - \hat{b}x_i)^2 = \text{Minimum}$$

Για τον υπολογισμό των α, b , αρκεί να βρούμε τις κρίσιμες τιμές: $\nabla S(\hat{\alpha}, \hat{b}) = 0 \Leftrightarrow \frac{\partial S}{\partial \hat{\alpha}} = 0$ και $\frac{\partial S}{\partial \hat{b}} = 0$

Είναι:

$$\frac{\partial S}{\partial \hat{\alpha}} = \sum_{i=1}^n -2(y_i - \hat{\alpha} - \hat{b}x_i) = -2 \left[\sum_{i=1}^n y_i - \sum_{i=1}^n \hat{\alpha} - \hat{b} \sum_{i=1}^n x_i \right] = -2[n\bar{y} - n\hat{\alpha} - \hat{b}n\bar{x}]$$

Και:

$$\frac{\partial S}{\partial \hat{\alpha}} = 0 \Leftrightarrow \bar{y} - \hat{\alpha} - \hat{b}\bar{x} = 0 \text{ ή } \hat{\alpha} = \bar{y} - \hat{b}\bar{x}$$

Σημείωση

Από την τελευταία σχέση προκύπτει ειδικότερα ότι το σημείο $(\bar{x}, \bar{y})$ ανήκει στην ευθεία ελαχίστων τετραγώνων.



Υπολογισμός συντελεστών 2/3

Περαιτέρω, από την σχέση $\frac{\partial S}{\partial \hat{b}} = 0$ παίρνουμε:

$$\sum_{i=1}^n x_i(y_i - \hat{\alpha} - \hat{b}x_i) = 0 \Leftrightarrow \sum_{i=1}^n x_i(y_i - \bar{y} + \hat{b}\bar{x} - \hat{b}x_i) = 0 \Leftrightarrow \sum_{i=1}^n x_i(y_i - \bar{y}) - \hat{b} \sum_{i=1}^n x_i(x_i - \bar{x}) = 0$$

$$\hat{b} = \frac{\sum_{i=1}^n x_i(y_i - \bar{y})}{\sum_{i=1}^n x_i(x_i - \bar{x})}$$



Υπολογισμός συντελεστών 3/3

Εύκολα, αποδεικνύεται ότι

$$\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y}) = \sum_{i=1}^n x_i (y_i - \bar{y}) - \sum_{i=1}^n \bar{x} y_i + n \bar{x} \bar{y} = \sum_{i=1}^n x_i (y_i - \bar{y})$$

$$\sum_{i=1}^n (x_i - \bar{x}) \cdot (x_i - \bar{x}) = \sum_{i=1}^n x_i (x_i - \bar{x}) - \sum_{i=1}^n \bar{x} x_i + n \bar{x} \bar{x} = \sum_{i=1}^n x_i (x_i - \bar{x})$$

Από τις τελευταίες δύο ισότητες, συνάγεται ότι μία ισοδύναμη έκφραση των συντελεστών είναι:

$$\hat{\alpha} = \bar{y} - \hat{\mathbf{b}} \bar{x}, \quad \hat{\mathbf{b}} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{s_{XY}^2}{s_X^2} = r_{XY} \frac{s_Y}{s_X}$$



Σημαντικά αθροίσματα τετραγώνων

Αποδεικνύεται ότι:
$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Βασικά βήματα απόδειξης:

$$(y_i - \bar{y})^2 = (y_i - \hat{y}_i + \hat{y}_i - \bar{y})^2 = (\hat{y}_i - \bar{y})^2 + (y_i - \hat{y}_i)^2 + 2(\hat{y}_i - \bar{y})(y_i - \hat{y}_i)$$

$$\hat{y}_i - \bar{y} = \hat{b}(x_i - \bar{x}), \text{ από } \bar{y} = \hat{a} + \hat{b}\bar{x} \text{ και } \hat{y}_i = \hat{a} + \hat{b}x_i$$

$$y_i - \hat{y}_i = (y_i - \bar{y}) + (\bar{y} - \hat{y}_i) = (y_i - \bar{y}) - \hat{b}(x_i - \bar{x}) = (y_i - \bar{y}) - \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} (x_i - \bar{x})$$

$$\sum_{i=1}^n (\hat{y}_i - \bar{y})(y_i - \hat{y}_i) = \hat{b} \sum_{i=1}^n (x_i - \bar{x})(y_i - \hat{y}_i) = \hat{b} \sum_{i=1}^n (x_i - \bar{x}) \left[(y_i - \bar{y}) - \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} (x_i - \bar{x}) \right] = 0$$



$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Στα πλαίσια της γραμμικής παλινδρόμησης τα παραπάνω αθροίσματα τετραγώνων έχουν ιδιαίτερη σημασία και ονομασία.

α) $TSS = \sum_{i=1}^n (y_i - \bar{y})^2$: συνολικό άθροισμα τετραγώνων (total sum of squares)

Το συνολικό άθροισμα τετραγώνων (TSS) εκφράζει τη συνολική μεταβλητότητα της εξαρτημένης μεταβλητής.

β) $ESS = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$: επεξηγημένο άθροισμα τετραγώνων (explained sum of squares)

Το επεξηγημένο άθροισμα τετραγώνων (ESS) είναι το μέρος της μεταβλητότητας της εξαρτημένης μεταβλητής που εξηγείται από τις επεξηγηματικές μεταβλητές του γραμμικού μοντέλου.

γ) $RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2$: άθροισμα τετραγώνων υπολοίπων (residual sum of squares)

Το άθροισμα τετραγώνων υπολοίπων (RSS) εκφράζει το μέρος της μεταβλητότητας της εξαρτημένης μεταβλητής που δεν εξηγείται από τις επεξηγηματικές μεταβλητές του γραμμικού μοντέλου.



Η σχέση $\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2$ γράφεται:

$$TSS = ESS + RSS$$

Διαιρώντας, με το συνολικό άθροισμα τετραγώνων (TSS) παίρνουμε:

$$\frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$



Συντελεστής Προσδιορισμού (R^2) 1/2

Το πόσο καλά προσαρμόζεται το γραμμικό μοντέλο στα δεδομένα ποσοτικοποιείται και από το συντελεστή προσδιορισμού (coefficient of determination) ο οποίος ορίζεται ως:

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Αν το μοντέλο προβλέπει χωρίς σφάλμα όλες τις παρατηρήσεις, τότε

$$y_i - \hat{y}_i = 0 \Leftrightarrow R^2 = 1 = 100\%.$$

Αν τα σφάλματα προβλέψεων του μοντέλου είναι όσο η διαφορά από την μέση τιμή (baseline model), τότε

$$y_i - \hat{y}_i = y_i - \bar{y} \Leftrightarrow R^2 = 0 = 0\%.$$

Αν τα σφάλματα προβλέψεων ξεπερνούν την απόσταση των παρατηρήσεων από τη μέση τιμή, τότε $R^2 < 0$.



Συντελεστής Προσδιορισμού (R^2) 2/2

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Ο συντελεστής R^2 συνήθως παρουσιάζεται ως ποσοστό με μεγαλύτερες τιμές να καταδεικνύουν καλύτερη προσαρμογή στα δεδομένα. Τα όρια αποδοχής της αποτελεσματικότητας του μοντέλου ποικίλουν ανάλογα με την επιστημονική περιοχή.

Στην απλή γραμμική παλινδρόμηση αποδεικνύεται ότι ο συντελεστής προσδιορισμού είναι το τετράγωνο του συντελεστή συσχέτισης Pearson μεταξύ των x_i και των y_i .

$$R^2 = r^2$$

Ειδικότερα: Στην απλή γραμμική παλινδρόμηση είναι αδύνατον για το R^2 να πάρει αρνητικές τιμές.

Σημείωση

Μία απόδειξη μπορεί να βρεθεί εδώ:

<https://math.stackexchange.com/questions/129909/correlation-coefficient-and-determination-coefficient>



Τυπικό Σφάλμα Συντελεστών

$$\hat{\alpha} = \bar{y} - \hat{b}\bar{x}, \quad \hat{b} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{s_{xy}^2}{s_x^2} = r_{xy} \frac{s_y}{s_x}$$

Αποδεικνύεται ότι το τυπικό σφάλμα των παραπάνω υπολογισμών δίνεται από τους τύπους

$$SE(\hat{\alpha}) = \sqrt{\frac{1}{n(n-2)} \sum_{i=1}^n x_i^2 \sum_{i=1}^n (y_i - \hat{y})^2 / \sum_{i=1}^n (x_i - \bar{x})^2} = \sqrt{\frac{n-1}{n(n-2)} \sum_{i=1}^n x_i^2 \cdot \frac{RSS}{s_x^2}}$$

$$SE(\hat{b}) = \sqrt{\frac{1}{n-2} \sum_{i=1}^n (y_i - \hat{y})^2 / \sum_{i=1}^n (x_i - \bar{x})^2} = \sqrt{\frac{n-1}{n-2} \cdot \frac{RSS}{s_x^2}}$$



Τυπικό Σφάλμα Συντελεστών

Η γνώση του τυπικού σφάλματος του γραμμικού μοντέλου, μας επιτρέπει:

(α) Να ελέγχουμε αν απορρίπτεται ή όχι η υπόθεση $H_0: b = 0$ (αντ. η $H_0: \alpha = 0$) υπολογίζοντας το στατιστικό

$$H_0: b=0: t = \frac{\hat{b}}{SE(\hat{b})} \sim t_{n-2}, \quad p = P(|t| > |t_0|)$$

(β) Να υπολογίζουμε το 95% διάστημα εμπιστοσύνης για τους συντελεστές του μοντέλου

$$\left(\hat{b} - t_{n-2;0.025} SE(\hat{b}), \hat{b} + t_{n-2;0.025} SE(\hat{b}) \right)$$

Σημείωση:

Μία απόδειξη για το γεγονός ότι οι συντελεστές ακολουθούν την κατανομή t μπορεί να βρεθεί εδώ:

<https://stats.stackexchange.com/questions/117406/proof-that-the-coefficients-in-an-ols-model-follow-a-t-distribution-with-n-k-d>



Υπολογισμός συντελεστών – Γενική Περίπτωση 1/3

Η γενική περίπτωση μπορεί να περιέχει πολλές επεξηγηματικές μεταβλητές.

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \cdots + \beta_k x_{ki} + \varepsilon_i \quad (i = 1, \dots, n).$$

Ορίζουμε:

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, X = \begin{pmatrix} 1 & x_{21} & \cdots & x_{k1} \\ \vdots & \vdots & & \vdots \\ 1 & x_{2n} & \cdots & x_{kn} \end{pmatrix}, \beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}, \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Με το νέο συμβολισμό, η αρχική εξίσωση γράφεται:

$$y = X\beta + \varepsilon.$$

Με β συμβολίζεται το διάνυσμα των άγνωστων παραμέτρων και με ε το διάνυσμα των σφαλμάτων που αντιστοιχούν σε κάθε μία παρατήρηση.



Υπολογισμός συντελεστών – Γενική Περίπτωση 2/3

Αν το διάνυσμα b είναι η εκτίμηση του β , τότε μπορούμε να γράψουμε:

$$y = Xb + e.$$

Στην εξίσωση αυτή, e δηλώνει τις παρατηρήσιμες διαφορές, δηλαδή τα σφάλματα εκτίμησης του γραμμικού μοντέλου. Ειδικότερα:

$$e = y - Xb.$$

Στο σημείο αυτό ορίζουμε τη συνάρτηση:

$$\begin{aligned} S(b) &= \sum e_i^2 = e'e = (y - Xb)'(y - Xb) \\ &= y'y - y'Xb - b'X'y + b'X'Xb. \end{aligned}$$



Υπολογισμός συντελεστών – Γενική Περίπτωση 3/3

Με λίγη άλγεβρα πινάκων, βρίσκουμε ότι:

$$\frac{\partial S}{\partial b} = -2X'y + 2X'Xb.$$

Η ελάχιστη τιμή ως προς b , λαμβάνεται όταν η παράγωγος είναι ίση με 0.
Συμπεραίνουμε, ότι, ο εκτιμητής ελαχίστων τετραγώνων ικανοποιεί την εξίσωση:

$$X'Xb = X'y.$$

Λύνοντας ως προς b βρίσκουμε:

$$b = (X'X)^{-1}X'y$$

Η μόνη προϋπόθεση υλοποίησης της παραπάνω διαδικασίας είναι ο πίνακας $X'X$ να είναι αντιστρέψιμος ή ισοδύναμα το πλήθος παρατηρήσεων να είναι μεγαλύτερο ή ίσο από το πλήθος των αγνώστων συντελεστών του μοντέλου.



Εφαρμογή ελαχίστων τετραγώνων σε μοντέλα της μορφής $X_t = M(t, \theta)$

Η μέθοδος αυτή είναι δυνατόν να εφαρμοστεί σε μοντέλα της μορφής:

$$X_t = M(t, \theta) + \varepsilon_t$$

όπου:

- $M(t, \theta)$ είναι μία γνωστή συνάρτηση που εξαρτάται από το t ,
- θ είναι το διάνυσμα των αγνώστων παραμέτρων και
- $\varepsilon_t \text{ iid} \sim N(0, \sigma^2)$, το τυχαίο σφάλμα της πρόβλεψης.

Ελαχιστοποιώντας το άθροισμα τετραγώνων των ε_t :

$$\sum_{t=1}^n [X_t - M(t, \theta)]^2$$

είναι δυνατόν να υπολογίσουμε εκτιμητές ελαχίστων τετραγώνων για τις παραμέτρους του διανύσματος θ .



Μέθοδος των ελαχίστων τετραγώνων στις χρονοσειρές

Παράδειγμα 1

Στον πίνακα παρουσιάζεται η παραγωγή ζαχαρότευτλου (σε τόνους) σε μία περιοχή ανά έτος. Να γίνει γραμμικό μοντέλο πρόβλεψης της εξέλιξης της παραγωγής στο χρόνο.

Έτος t	2010	2011	2012	2013	2014	2015	2016
Παραγωγή	40	45	46	42	47	50	46

Τυπολόγιο: $\hat{y} = \hat{\alpha} + \hat{b}t$

$$\hat{b} = \frac{\sum_{i=1}^v (t_i - \bar{t})(y_i - \bar{y})}{\sum_{i=1}^v (t_i - \bar{t})^2}$$
$$\hat{\alpha} = \bar{y} - \hat{b}\bar{t}$$



Μέθοδος των ελαχίστων τετραγώνων στις χρονοσειρές

Παράδειγμα 1

Στον πίνακα παρουσιάζεται η παραγωγή ζαχαρότευτλου (σε τόνους) σε μία περιοχή ανά έτος. Να γίνει γραμμικό μοντέλο πρόβλεψης της εξέλιξης της παραγωγής στο χρόνο.

Έτος t	2010	2011	2012	2013	2014	2015	2016
Παραγωγή	40	45	46	42	47	50	46

Τυπολόγιο: $\hat{y} = \hat{\alpha} + \hat{b}t$

$$\hat{b} = \frac{\sum_{i=1}^n (t_i - \bar{t})(y_i - \bar{y})}{\sum_{i=1}^n (t_i - \bar{t})^2} \quad \hat{\alpha} = \bar{y} - \hat{b}\bar{t}$$

Αν οι υπολογισμοί γίνονται με το χέρι, τότε είναι πρακτικό το έτος t να αντικατασταθεί με [Έτος – Μέση τιμή] = {-3, -2, -1, 0, 1, 2, 3}, καθώς τότε η μέση τιμή του t είναι 0 και οι τύποι απλοποιούνται:

$$\hat{b} = \sum_{i=1}^n t_i y_i / \sum_{i=1}^n t_i^2 \quad \hat{\alpha} = \bar{y}$$



Μέθοδος των ελαχίστων τετραγώνων στις χρονοσειρές

Παράδειγμα 1

Έτος	Παραγωγή (y)	t = Έτος – 2013	ty	t ²
2010	40	-3	-120	9
2011	45	-2	-90	4
2012	46	-1	-46	1
2013	42	0	0	0
2014	47	1	47	1
2015	50	2	100	4
2016	46	3	138	9
Άθροισμα	316	0	29	28

$$\hat{b} = \sum_{i=1}^n t_i y_i / \sum_{i=1}^n t_i^2 = 29/28 = 1,036$$

$$\hat{\alpha} = \bar{y} = 316/7 = 45,143$$



Μέθοδος των ελαχίστων τετραγώνων

Παράδειγμα 1

$$\hat{b} = \sum_{i=1}^n t_i y_i / \sum_{i=1}^n t_i^2 = 29/28 = 1,036$$

$$\hat{\alpha} = \bar{y} = 316/7 = 45,143$$

Εξίσωση πρόβλεψης:

$$\text{Παραγωγή} = 1,036 \cdot (\text{Έτος} - 2013) + 45,143.$$

Συμπεραίνουμε ότι:

Για κάθε έτος που περνάει αναμένουμε αυξημένη παραγωγή κατά 1,036 τόνους.



Μέθοδος των ελαχίστων τετραγώνων

Παράδειγμα 1

Υλοποίηση στο Excel / Calc

Συναρτήσεις $\text{SLOPE}(y, x)$ και $\text{INTERCEPT}(y, x)$ (ακολουθεί επίδειξη)

Υλοποίηση στη γλώσσα R

```
years = c(-3, -2, -1, 0, 1, 2, 3)
```

```
production = c(40, 45, 46, 42, 47, 50, 46)
```

```
lm(production ~ years)
```

(ακολουθεί επίδειξη)



Εφαρμογή ελαχίστων τετραγώνων σε AR(p) μοντέλα

Η μέθοδος αυτή είναι δυνατόν να εφαρμοστεί σε μοντέλα AR(p): $x_n = \sum_{k=1}^p a_k x_{n-k} + \epsilon_n$
όπου $\epsilon_n \sim N(0, \sigma^2)$.

Αν $x_1, x_2, \dots, x_N$, είναι οι τιμές της χρονοσειράς που έχουμε παρατηρήσει, τότε μπορούμε να βρούμε τους συντελεστές, λύνοντας το σύστημα των $M = N + 1 - p$ εξισώσεων:

$$\begin{bmatrix} x_0 & x_1 & \cdots & x_{p-1} \\ x_1 & x_2 & \cdots & x_p \\ x_2 & x_3 & \cdots & x_{p+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N-p} & x_{N-p+1} & \cdots & x_{N-1} \end{bmatrix} \begin{bmatrix} a_p \\ a_{p-1} \\ \vdots \\ a_1 \end{bmatrix} = \begin{bmatrix} x_p \\ x_{p+1} \\ x_{p+2} \\ \vdots \\ x_N \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} \mathbf{x}_{N-p,M} & \mathbf{x}_{N-p+1,M} & \cdots & \mathbf{x}_{N-1,M} \end{bmatrix} \mathbf{a} = \mathbf{x}_{N,M}$$
$$\downarrow$$
$$\mathbf{X}_{N,M} \mathbf{a} = \mathbf{x}_{N,M}$$

Ο εκτιμητής ελαχίστων τετραγώνων είναι:

$$\hat{\mathbf{a}} = (\mathbf{X}_{N,M}^T \mathbf{X}_{N,M})^{-1} \mathbf{X}_{N,M}^T \mathbf{x}_{N,M}$$



Εφαρμογή ελαχίστων τετραγώνων σε AR(p) μοντέλα

Παράδειγμα AR(1) $X_t = \phi X_{t-1} + \mu + Z_t$

Αρκεί να ελαχιστοποιήσουμε το άθροισμα τετραγωνικών σφαλμάτων:

$$E = \frac{1}{2n} \sum_{t=2}^n (X_t - \phi X_{t-1} - \mu)^2$$

Υπολογίζουμε τις μερικές παραγώγους:

$$\frac{\partial E}{\partial \mu} = \frac{1}{n} \sum_{t=2}^n (\phi X_{t-1} + \mu - X_t) = \phi \langle X_{t-1} \rangle + \mu - \langle X_t \rangle$$

$$\frac{\partial E}{\partial \phi} = \frac{1}{n} \sum_{t=2}^n (\phi X_{t-1} + \mu - X_t) X_{t-1} = \phi \langle X_{t-1}^2 \rangle + \mu \langle X_{t-1} \rangle - \langle X_t X_{t-1} \rangle$$



Εφαρμογή ελαχίστων τετραγώνων σε AR(p) μοντέλα

Παράδειγμα AR(1)

Για να βρούμε τα ακρότατα, αρκεί να εξισώσουμε τις δύο εξισώσεις με το 0:

$$\frac{\partial E}{\partial \mu} = \frac{1}{n} \sum_{t=2}^n (\phi X_{t-1} + \mu - X_t) = \phi \langle X_{t-1} \rangle + \mu - \langle X_t \rangle = 0$$

$$\frac{\partial E}{\partial \phi} = \frac{1}{n} \sum_{t=2}^n (\phi X_{t-1} + \mu - X_t) X_{t-1} = \phi \langle X_{t-1}^2 \rangle + \mu \langle X_{t-1} \rangle - \langle X_t X_{t-1} \rangle = 0$$

Παίρνουμε το σύστημα:
$$\begin{pmatrix} 1 & \langle X_{t-1} \rangle \\ \langle X_{t-1} \rangle & \langle X_{t-1}^2 \rangle \end{pmatrix} \begin{pmatrix} \mu \\ \phi \end{pmatrix} = \begin{pmatrix} \langle X_t \rangle \\ \langle X_{t-1} X_t \rangle \end{pmatrix} .$$

Η λύση του είναι:
$$\begin{pmatrix} \mu \\ \phi \end{pmatrix} = \frac{1}{\text{var}(X_{t-1})} \begin{pmatrix} \text{cov}(X_t X_{t-1}, X_{t-1}) - \text{cov}(X_t, X_{t-1}^2) \\ \text{cov}(X_{t-1}, X_t) \end{pmatrix} .$$

Εκτιμητές μέγιστης πιθανοφάνειας



Εκτιμητές μέγιστης πιθανοφάνειας

Στην Στατιστική, η εκτίμηση μέγιστης πιθανότητας (MLE) είναι μια μέθοδος εκτίμησης των παραμέτρων μιας άγνωστης κατανομής πιθανότητας, από ένα σύνολο παρατηρούμενων δεδομένων.

Αυτό επιτυγχάνεται με τη μεγιστοποίηση μιας συνάρτησης πιθανότητας, έτσι ώστε, σύμφωνα με το υποτιθέμενο στατιστικό μοντέλο, τα παρατηρούμενα δεδομένα να είναι πιο πιθανά να εμφανιστούν.

Το σημείο στο χώρο των παραμέτρων που μεγιστοποιεί τη συνάρτηση πιθανότητας ονομάζεται εκτιμητής μέγιστης πιθανότητας.

Η διαδικασία είναι εφικτή τόσο για συνεχείς όσο και για διακριτές τμ. Η μέθοδος προτάθηκε από τον Fisher το 1912.



Εκτιμητές μέγιστης πιθανοφάνειας

Έστω ότι τα δεδομένα $x = (x_1, \dots, x_n)$, προέρχονται από μία κατανομή της οποίας γνωρίζουμε το είδος της κατανομής αλλά όχι τις παραμέτρους αυτής. Έστω επίσης πως οι παρατηρήσεις είναι ανεξάρτητες.

Ορίζουμε τη συνάρτηση πιθανοφάνειας L ως εξής :

$$L(\theta) = L(X, \theta) = f(x_1, \dots, x_n | \theta)$$

Αναζητούμε την τιμή της παραμέτρου θ που μεγιστοποιεί τη συνάρτηση L . Η τιμή που βρίσκουμε με την παραπάνω διαδικασία ονομάζεται

Εκτιμητής Μέγιστης Πιθανοφάνειας (Maximum Likelihood Estimation (MLE)).



Εκτιμητές μέγιστης πιθανοφάνειας

Οι παρατηρήσεις θεωρούνται ανεξάρτητες, άρα

$$f(x_1, \dots, x_n | \theta) = f(x_1 | \theta) \cdot f(x_2 | \theta) \cdot \dots \cdot f(x_n | \theta)$$

Η συνάρτηση πιθανοφάνειας L γράφεται:

$$L(\theta) = L(X, \theta) = f(x_1 | \theta) \cdot f(x_2 | \theta) \cdot \dots \cdot f(x_n | \theta)$$

Για να βρούμε την τιμή της παραμέτρου θ που μεγιστοποιεί τη συνάρτηση L , αρκεί να βρούμε την τιμή που μεγιστοποιεί την

$$l(\theta) = \ln(L(\theta)) = \sum \ln(f(x_i | \theta)).$$

Ο εκτιμητής μέγιστης πιθανοφάνειας προκύπτει από την επίλυση της εξίσωσης

$$d/d\theta l(\theta) = 0$$



Εκτιμητές μέγιστης πιθανοφάνειας

Άσκηση 1. Έστω X μία ΤΜ με κατανομή

x	0	1	2	3
$f_X(x)$	$2\theta/3$	$\theta/3$	$2(1 - \theta)/3$	$(1 - \theta)/3$

Για να εκτιμήσουμε την άγνωστη παράμετρο θ , υλοποιούμε δειγματοληψία 10 ανεξάρτητων τιμών και παίρνουμε το δείγμα (3, 0, 2, 1, 3, 2, 1, 0, 2, 1). Να βρεθεί η τιμή της παραμέτρου θ με τη μέθοδο του Εκτιμητή Μέγιστης Πιθανοφάνειας.



Εκτιμητές μέγιστης πιθανοφάνειας

Λύση Άσκησης 1.

$$\begin{aligned} L(X, \theta) &= f((3, 0, 2, 1, 3, 2, 1, 0, 2, 1) | \theta) \\ &= f(3|\theta) \cdot f(0|\theta) \cdot f(2|\theta) \cdot f(1|\theta) \cdot f(3|\theta) \cdot f(2|\theta) \cdot f(1|\theta) \cdot f(0|\theta) \cdot f(2|\theta) \cdot f(1|\theta) \\ &= [f(0|\theta)]^2 \cdot [f(1|\theta)]^3 \cdot [f(2|\theta)]^3 \cdot [f(3|\theta)]^2 \\ &= [2\theta/3]^2 \cdot [\theta/3]^3 \cdot [2(1-\theta)/3]^3 \cdot [(1-\theta)/3]^2 \\ &= \theta^5 \cdot (1-\theta)^5 \cdot 2^5 / 3^{10}. \end{aligned}$$

$$l(\theta) = \ln L(X, \theta) = 5 \ln \theta + 5 \ln(1-\theta) + c.$$

Η συνάρτηση $l(\theta)$ μεγιστοποιείται για $\theta = 1/2$. Αυτή είναι η τιμή του εκτιμητή μέγιστης πιθανοφάνειας. Δηλαδή, η τιμή $\theta = 1/2$ είναι αυτή που μεγιστοποιεί την πιθανότητα το δείγμα που έχουμε να έχει προέλθει από την υποτιθέμενη κατανομή.



Εκτιμητές μέγιστης πιθανοφάνειας

Άσκηση 2. Έστω ότι η τ.μ. X μετράει το χρόνο μεταξύ αφίξεων σε μία διαδικασία Poisson και πως $X \sim \text{Exp}(\lambda)$. Παρακολουθούμε δείγμα 4 τιμών της X και βρίσκουμε τις τιμές 1.23, 3.32, 1.98, 2.12. Να βρεθεί ο Εκτιμητής Μέγιστης Πιθανοφάνειας της παραμέτρου λ .

Λύση Άσκησης 2

Από την υπόθεση έχουμε πως $X \sim \text{Exp}(\lambda)$, δηλαδή ότι $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$.

Ορίζουμε $L(X, \theta) = f(1.23, 3.32, 1.98, 2.12 | \lambda)$

$$= f(1.23|\lambda) \cdot f(3.32|\lambda) \cdot f(1.98|\lambda) \cdot f(2.12|\lambda)$$

$$= \lambda e^{-1.23\lambda} \cdot \lambda e^{-3.32\lambda} \cdot \lambda e^{-1.98\lambda} \cdot \lambda e^{-2.12\lambda}$$

$$= \lambda^4 e^{-8.65\lambda}.$$

Είναι $l(\theta) = \ln L(X, \theta) = 4 \ln \lambda - 8.65 \lambda$.

Η συνάρτηση $l(\theta)$ μεγιστοποιείται για $\lambda = 0.46$, τιμή που μεγιστοποιεί την πιθανότητα το δείγμα που έχουμε να έχει προέλθει από την κατανομή $\text{Exp}(\lambda)$.



Αμερόληπτοι και συνεπείς εκτιμητές

Πόσο καλές είναι οι εκτιμήσεις των παραμέτρων που προκύπτουν από τους τύπους που χρησιμοποιούμε;

Στην Στατιστική είναι επιθυμητό ο εκτιμητής θ_n να είναι

- **αμερόληπτος** (unbiased)

$$\text{Bias}(\theta_n, \theta) = E_{x|\theta}(\theta_n - \theta) = 0.$$

(ανεξάρτητα από το μέγεθος του δείγματος, ο εκτιμητής δίνει τη σωστή τιμή της παραμέτρου ως αναμενόμενη τιμή)

- **συνεπής** (consistent)

$$\text{plim}_{n \rightarrow \infty} \theta_n = \lim_{n \rightarrow \infty} P(|\theta_n - \theta| > \varepsilon) = 0, \text{ για κάθε } \varepsilon > 0.$$

(η αύξηση του μεγέθους του δείγματος οδηγεί σε καλύτερη πρόβλεψη της παραμέτρου)

Και οι δύο έννοιες αποκτούν νόημα με την προϋπόθεση πως γνωρίζουμε το είδος της κατανομής που ακολουθεί ο πληθυσμός από τον οποίο πήραμε το δείγμα.



Μέγιστη Πιθανοφάνεια ~~→~~ Αμεροληψία

Παράδειγμα 1 (με αφορμή την άσκηση 2)

Εστω X μία μεταβλητή που γνωρίζουμε ότι ακολουθεί την $\text{Exp}(\lambda)$ αλλά δεν γνωρίζουμε το λ (= πλήθος αφίξεων / μονάδα χρόνου). Για το λόγο αυτό παρατηρούμε n διαφορετικές διάρκειες μεταξύ αφίξεων $X_1, X_2, \dots, X_n$ και εκτιμούμε τον ρυθμό λ από τον τύπο:

$$\hat{\lambda} = \frac{n}{X_1 + X_2 + \dots + X_n}$$

Ο εκτιμητής αυτός **δεν είναι αμερόληπτος**. Πράγματι, για $n = 1$, είναι

$$E(\hat{\lambda}) = E\left(\frac{1}{X_1}\right) = \int_0^{+\infty} \frac{\lambda}{x} e^{-\lambda/x} dx = +\infty$$

ενώ για $n > 1$ είναι $X_1 + X_2 + \dots + X_n \sim \text{Γάμμα}(n, \lambda)$, άρα $1/(X_1 + X_2 + \dots + X_n) \sim \text{Inverse Γάμμα}(n, \lambda)$ και

$$E(\hat{\lambda}) = E\left(\frac{n}{X_1 + X_2 + \dots + X_n}\right) = n \frac{\lambda}{n-1} \neq \lambda.$$

Ιδιαίτερα, συμπεραίνουμε, ότι ο $\hat{\lambda} = \frac{n-1}{X_1 + X_2 + \dots + X_n}$ είναι ένας αμερόληπτος εκτιμητής του λ .

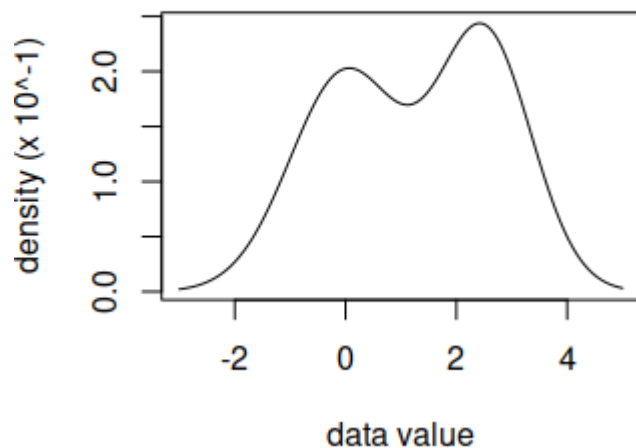


Μέγιστη Πιθανοφάνεια $\nrightarrow$ Συνέπεια

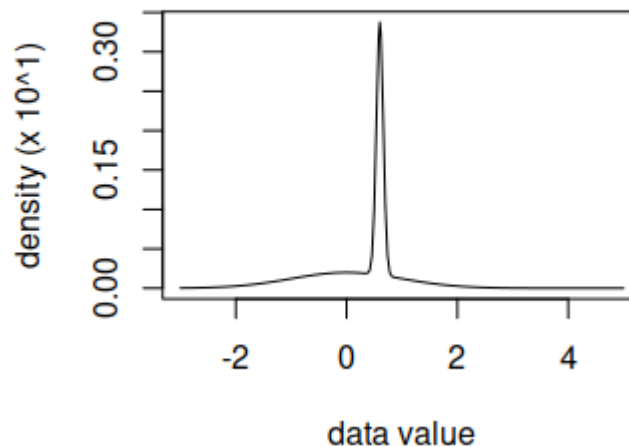
Παράδειγμα 2

Εστω X τμ για την οποία γνωρίζουμε ότι $X \sim 1/2 N(0, 1) + 1/2 N\left(t, e^{-\frac{2}{t^2}}\right)$

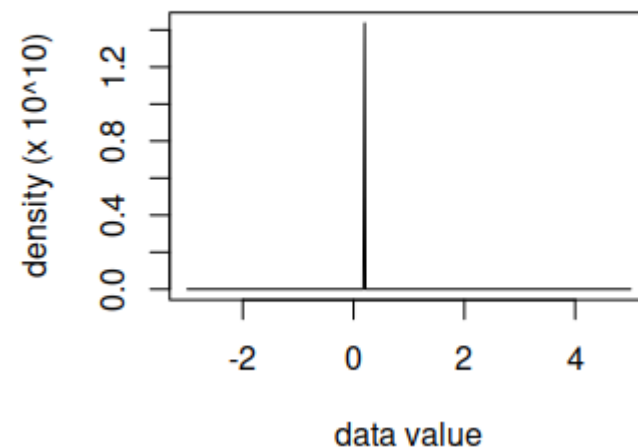
Density with parameter 2.5



Density with parameter 0.6



Density with parameter 0.2



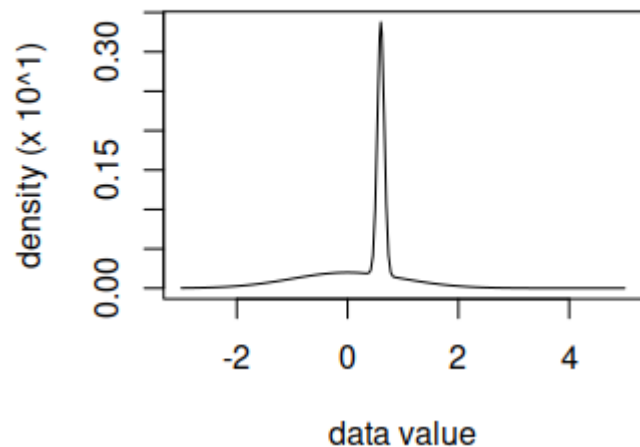


Μέγιστη Πιθανοφάνεια $\nrightarrow$ Συνέπεια

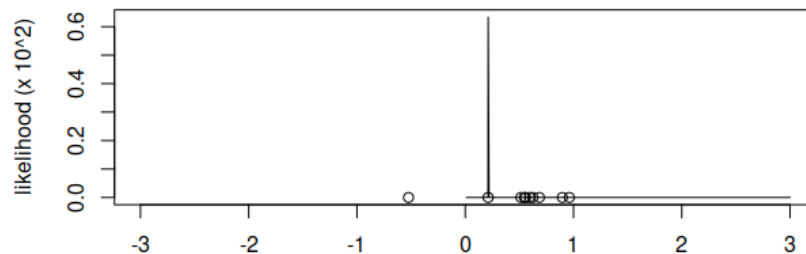
Παράδειγμα 2

Για $t = 0,6$, και $\{x_i\}_{i \leq n}$, $n = 10$ ή 30 ή 100 παρατηρήσεις της X , ο Ε.Μ.Π. της παραπάνω κατανομής **δεν είναι συνεπής**.

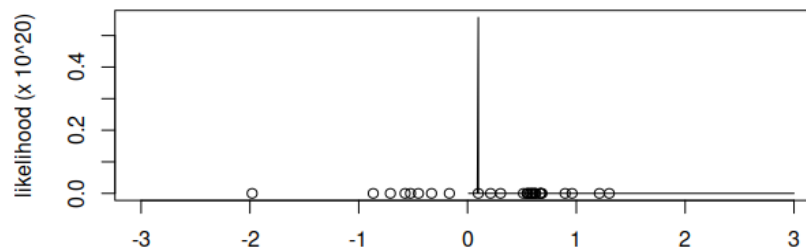
Density with parameter 0.6



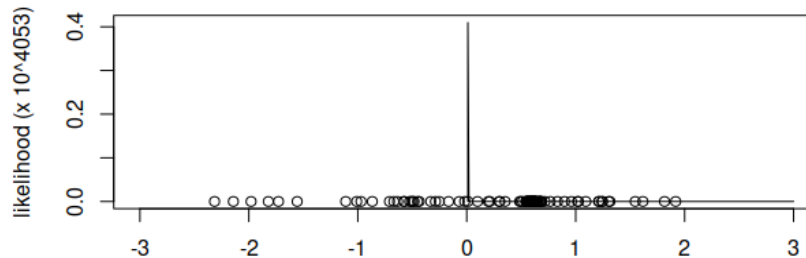
Likelihood - 10 data points, MLE approximately 0.2117



Likelihood - 30 data points, MLE approximately 0.0989



Likelihood - 100 data points, MLE approximately 0.0102





Μέγιστη Πιθανοφάνεια $\nrightarrow$ Μοναδικότητα

Παράδειγμα 3

Αν $X \sim \text{Laplace}(\mu, \beta)$ με $f(x) = \frac{1}{2\beta} \exp\left(-\frac{|x - \mu|}{\beta}\right)$, με άγνωστες παραμέτρους μ, β

και $x_1, x_2, \dots, x_n$, ένα δείγμα τιμών, τότε αποδεικνύεται ότι ο εκτιμητής μέγιστης πιθανοφάνειας για την παράμετρο μ είναι η διάμεσος του δείγματος⁽¹⁾.

Καθώς η διάμεση τιμή δεν είναι μοναδική όταν το n είναι άρτιος, καταλαβαίνουμε ότι ο εκτιμητής δεν ορίζεται με μοναδικό τρόπο.

(1) Μία απόδειξη είναι διαθέσιμη εδώ:

<https://math.stackexchange.com/questions/240496/finding-the-maximum-likelihood-estimator>



Εκτιμητές μέγιστης πιθανοφάνειας

Εκτίμηση AR(1): MLE

Μία στάσιμη χρονοσειρά που ακολουθεί ένα AR(1) μοντέλο $X_t = \delta + \phi X_{t-1} + w_t$, $w_t \sim N(0, \sigma^2)$ αποτελείται από εξαρτημένες παρατηρήσεις.

Ωστόσο, τα σφάλματα w_t , είναι iid ακολουθία με σ.π.π.
$$f_w(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}.$$

Αρα, αν $W = (w_1, w_2, \dots, w_n)$ η ακολουθία των σφαλμάτων, τότε η κοινή συνάρτηση πυκνότητας θα είναι το γινόμενο των επιμέρους

$$L(W, \sigma) = f(w_1, w_2, \dots, w_n | \sigma) = f(w_1 | \sigma) f(w_2 | \sigma) \dots f(w_n | \sigma) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n w_i^2\right)$$

Ομως, $w_t = X_t - \delta - \phi X_{t-1}$ και η τελευταία σχέση ξαναγράφεται

$$L(X, \sigma) = f(x_2, x_3, \dots, x_n | x_0, \sigma) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \delta - \phi x_{i-1})^2\right)$$

Συμπεραίνουμε, ότι η συνάρτηση πιθανοφάνειας είναι

$$l(\sigma) = -\ln L(X, \sigma) = \frac{n}{2} \ln(\sigma^2) + \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \delta - \phi x_{i-1})^2 + c$$



Εκτιμητές μέγιστης πιθανοφάνειας

Εκτίμηση AR(1): MLE

Η βελτιστοποίηση της $I(\sigma) = -\ln L(X, \sigma) = \frac{n}{2} \ln(\sigma^2) + \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \delta - \phi x_{i-1})^2 + c$
οδηγεί στη λύση:

$$\hat{\phi} = \frac{\sum_{i=1}^n (x_{i-1} - \bar{x})(x_i - \bar{x}_+)}{\sum_{i=0}^{n-1} (x_i - \bar{x})^2}, \quad \hat{\delta} = \bar{x}_+ - \hat{\phi} \bar{x}, \quad \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{\delta} - \hat{\phi} x_{i-1})^2$$

όπου

$$\bar{x} = \frac{1}{n} \sum_{i=0}^{n-1} x_i, \quad \bar{x}_+ = \frac{1}{n} \sum_{i=1}^n x_i$$

Σημείωση: Η παραπάνω διαδικασία γενικεύεται στην περίπτωση των AR(p) μοντέλων.



Εκτιμητές μέγιστης πιθανοφάνειας

Εναλλακτική Εκτίμηση AR(1): Exact MLE

Από τη θεωρία πιθανοτήτων γνωρίζουμε ότι για $\theta = (\delta, \phi, \sigma^2)$ είναι:

$$f(x_0, x_1, x_2, \dots, x_n | \theta) = f(x_0 | \theta) f(x_1 | x_0, \theta) \dots f(x_n | x_{n-1}, x_{n-2}, \dots, x_0, \theta) = f(x_0 | \theta) \prod_{i=1}^n f(x_i | x_{i-1}, x_{i-2}, \dots, x_0, \theta)$$

Καθώς η χρονοσειρά $X_t = \delta + \phi X_{t-1} + w_t$ είναι στάσιμη, είναι $X_0 \sim N(\mu, \text{Var}(X_t))$.

Ομως, $\mu = \delta / (1 - \phi)$, και $\text{Var}(X_t) = \sigma^2 / (1 - \phi^2)$, δηλαδή $X_0 \sim N(\delta / (1 - \phi), \sigma^2 / (1 - \phi^2))$ με

$$f(x_0 | \theta) = \frac{\sqrt{1 - \phi^2}}{\sqrt{2\pi} \sigma} \exp\left(-\frac{(x_0 - \delta / (1 - \phi))^2}{2 \sigma^2 / (1 - \phi^2)}\right)$$

Περαιτέρω, $x_n | x_{n-1}, x_{n-2}, \dots, x_1 = x_n | x_{n-1} = N(\delta + \phi x_{n-1}, \sigma^2)$ και

$$f(x_n | x_{n-1}, x_{n-2}, \dots, x_1, \sigma) = f(x_n | x_{n-1}, \sigma) = \frac{1}{\sqrt{2\pi} \sigma} \exp\left(-\frac{(x_n - \delta - \phi x_{n-1})^2}{2 \sigma^2}\right)$$



Εκτιμητές μέγιστης πιθανοφάνειας

Εναλλακτική Εκτίμηση AR(1): Exact MLE

Η συνάρτηση πιθανοφάνειας είναι η:

$$l(\sigma) = -\ln L(\sigma | X) = \frac{1}{2} \ln \frac{\sigma^2}{1 - \varphi^2} + \frac{1}{2} n \ln \sigma^2 + \frac{(x_0 - \delta / (1 - \varphi))^2}{2 \sigma^2 (1 - \varphi^2)} + \frac{1}{2 \sigma^2} \sum_{i=1}^n (x_i - \delta - \varphi x_{i-1})^2 + c$$

Ο εντοπισμός των ακρότατων στην περίπτωση αυτή γίνεται με αριθμητικές – προσεγγιστικές μεθόδους.



Σχέση μεταξύ MLE και Least Squares

Για το παράδειγμα AR(1), η συνάρτηση πιθανοφάνειας προέκυψε να είναι, μία από τις παρακάτω:

$$l(\sigma) = -\ln L(X, \sigma) = \frac{n}{2} \ln(\sigma^2) + \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \delta - \phi x_{i-1})^2 + c$$

$$l(\sigma) = -\ln L(\sigma | X) = \frac{1}{2} \ln \frac{\sigma^2}{1 - \phi^2} + \frac{1}{2} n \ln \sigma^2 + \frac{(x_0 - \delta / (1 - \phi))^2}{2\sigma^2(1 - \phi^2)} + \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \delta - \phi x_{i-1})^2 + c$$

Παρατηρούμε πως οι συναρτήσεις αυτές αποκτούν ελάχιστη τιμή της όταν ελαχιστοποιείται το

$$\sum_{i=1}^n (x_i - \delta - \phi x_{i-1})^2$$

Το τελευταίο άθροισμα το αναγνωρίζουμε ως το άθροισμα των τετραγωνικών σφαλμάτων το οποίο ελαχιστοποιείται στην μέθοδο των ελαχίστων τετραγώνων.

Συμπεραίνουμε πως, όταν $w_i \sim N(0, \sigma^2)$, τότε οι μέθοδοι μέγιστης πιθανοφάνειας και ελαχίστων τετραγώνων, οδηγούν στο ίδιο αποτέλεσμα.



ΔΗΜΟΚΡΙΤΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΡΑΚΗΣ

DEMOCRITUS
UNIVERSITY
OF THRACE

ARIMA με περιοδικότητα (ή SARIMA)



Μοντέλα SARIMA(p, d, q)(P, D, Q)_m

Η αναγνώριση της περιοδικότητας σε μία σειρά μπορεί να γίνει:

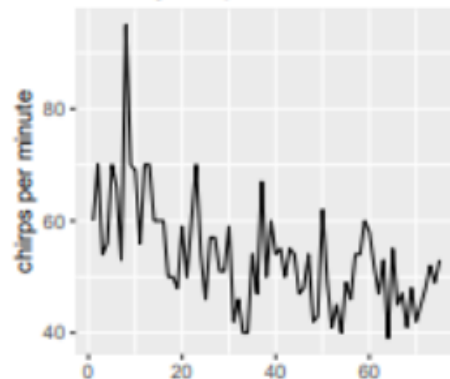
(α) Από το χρονοδιάγραμμα των τιμών με απλή παρατήρηση.

(β) Από το διάγραμμα αυτοσυσχετίσεων ACF: Θετική αυτοσυσχέτιση σε σταθερό πλήθος διαφορών δείχνει περιοδικότητα αντίστοιχης περιόδου.

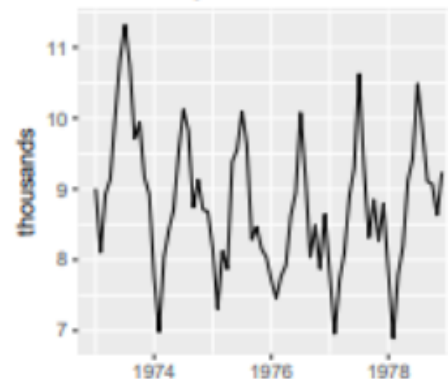
Στην περίπτωση όπου υπάρχει περιοδικότητα στη χρονοσειρά τότε αυτή δεν είναι στάσιμη καθώς η μέση τιμή δεν είναι διαχρονικά σταθερή. Καθώς η προσαρμογή ενός μοντέλου ARMA προϋποθέτει τη στασιμότητα, η πρώτη ενέργεια στην περίπτωση αυτή είναι ο εντοπισμός της εξίσωσης που θα εξαλείψει από τα δεδομένα την περιοδικότητα.



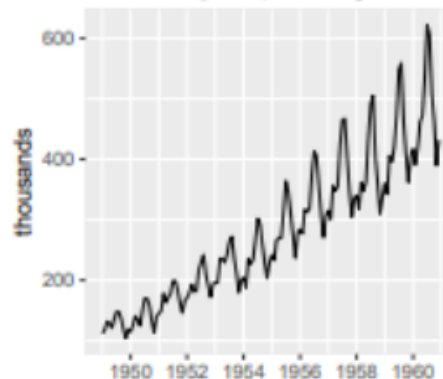
1. Daily temperature of cow



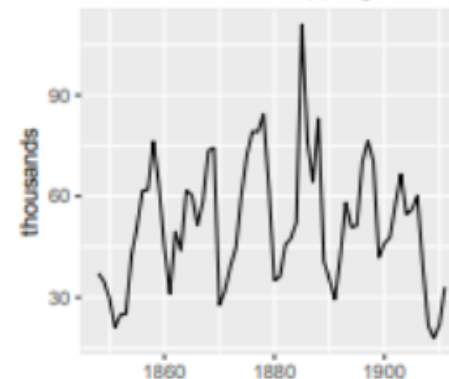
2. Monthly accidental deaths



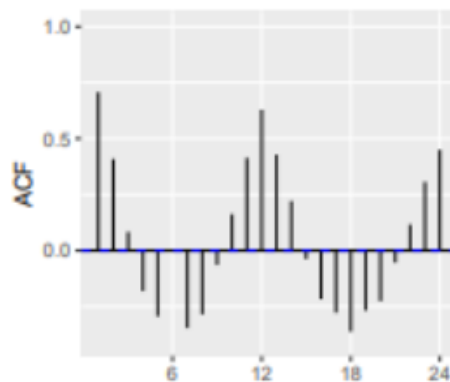
3. Monthly air passengers



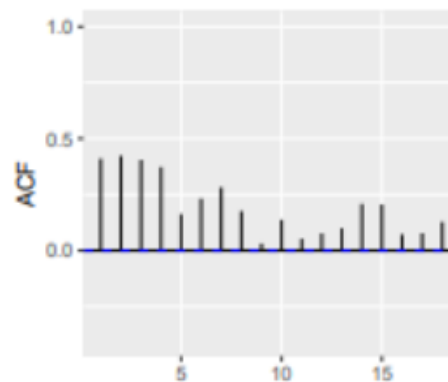
4. Annual mink trappings



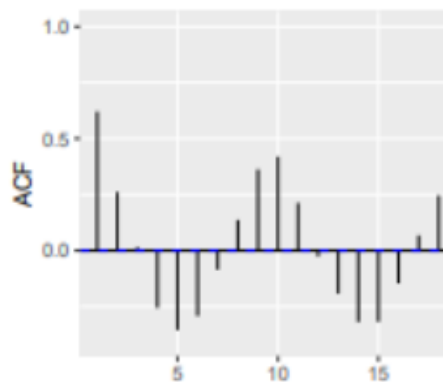
A



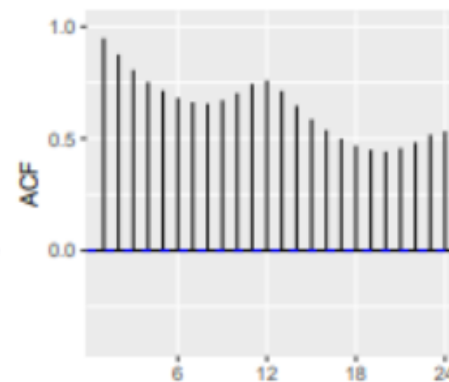
B



C



D





Μοντέλα SARIMA(p, d, q)(P, D, Q)_m

Ένας τρόπος για την αντιμετώπιση της εποχικότητας σε μια χρονοσειρά είναι η εφαρμογή εποχιακής διαφοροποίησης, η οποία αποτελεί ενσωματωμένο στοιχείο των εποχικών ARIMA μοντέλων (SARIMA).

Με την εποχιακή διαφορά αφαιρούμε την τιμή ενός σημείου από την αντίστοιχη τιμή της προηγούμενης περιόδου, δηλαδή:

$$\Delta_s X_t = X_t - X_{t-s}.$$

Με τον τρόπο αυτό εξαλείφεται το επαναλαμβανόμενο εποχικό μοτίβο.



Μοντέλα SARIMA(p, d, q)(P, D, Q)_m

Παράδειγμα 1

$$X = (X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, \dots) = (1, 2, 3, 4, 1, 2, 3, 4, \dots),$$

Παρατηρούμε ότι $T = 4$ και

$$\Delta_4 X = (X_5 - X_1, X_6 - X_2, X_7 - X_3, X_8 - X_4, \dots) = (0, 0, 0, 0, \dots)$$

Παράδειγμα 2

$$x = (1, 2, 3, 3, 2, 1, 1, 2, 3, 3, 2, 1, \dots),$$

Παρατηρούμε ότι $T = 6$ και

$$\Delta_6 X = (X_7 - X_1, X_8 - X_2, X_9 - X_3, X_{10} - X_4, \dots) = (0, 0, 0, 0, \dots)$$

```
x = rep(c(1, 2, 3, 4), 20)
plot.ts(x)
library(Hmisc)
D4x = Lag(x, 4) - x
plot.ts(D4x)
```



Μοντέλα SARIMA(p, d, q)(P, D, Q)_m

Στην περίπτωση που η εποχικότητα έχει μεταβαλλόμενα χαρακτηριστικά, τότε μπορεί να χρειαστεί εποχική διαφοροποίηση 2^{ης} ή μεγαλύτερης τάξης.

π.χ. αν

$$X = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, \dots) = (1, 2, 3, 4, 2, 3, 4, 5, 4, 5, 6, 7, \dots),$$

Τότε

$$\Delta_4 X = (x_5 - x_1, x_6 - x_2, x_7 - x_3, x_8 - x_4, \dots) = (1, 1, 1, 1, 2, 2, 2, 2, \dots)$$

Και

$$\Delta_4^2 X = (\Delta_4 x_5 - \Delta_4 x_1, \Delta_4 x_6 - \Delta_4 x_2, \Delta_4 x_7 - \Delta_4 x_3, \Delta_4 x_8 - \Delta_4 x_4, \dots) = (1, 1, 1, 1, \dots)$$



Μοντέλα $SARIMA(p, d, q)(P, D, Q)_m$

Για τη μοντελοποίηση της εποχικότητας μπορούν να χρησιμοποιηθούν εποχικοί αυτοπαλινδρομικοί όροι (πλήθος P), εποχιακές διαφορές (πλήθος D) και εποχικοί όροι κινούμενου μέσου (πλήθος Q).

Το αντίστοιχο εποχικό $ARIMA$ μοντέλο συμβολίζεται:

$$SARIMA(p, d, q)(P, D, Q)_m,$$

όπου m η περίοδος της χρονοσειράς.



Μοντέλα SARIMA(p, d, q)(P, D, Q)_m

Εξίσωση SARIMA(p, d, q)(P, D, Q)_m.

Συμβολίζουμε με

- $\varphi(\omega) = 1 - \varphi_1\omega - \varphi_2\omega^2 - \dots - \varphi_p\omega^p$, το πολυώνυμο της AR(p).
- $\Phi(\omega) = 1 - \Phi_1\omega - \Phi_2\omega^2 - \dots - \Phi_P\omega^P$, το πολυώνυμο της Seasonal AR(P).
- $\theta(\omega) = 1 - \theta_1\omega - \theta_2\omega^2 - \dots - \theta_q\omega^q$, το πολυώνυμο της MA(q).
- $\Theta(\omega) = 1 - \Theta_1\omega - \Theta_2\omega^2 - \dots - \Theta_Q\omega^Q$, το πολυώνυμο της Seasonal MA(Q).

Τότε:

Εξίσωση ARMA(p, q)(P, Q)_m: $\varphi(B)\Phi(B^m)X_t = \theta(B)\Theta(B^m)w_t$.

Εξίσωση ARIMA(p, d, q)(P, D, Q)_m: $\varphi(B)\Phi(B^m)\Delta^d \Delta_m^D X_t = \theta(B)\Theta(B^m)w_t$.

$$\Delta X_t = X_t - X_{t-1}, \quad \Delta^d X_t = \Delta \cdot \Delta \cdot \Delta \cdot \dots \cdot \Delta X_t, \quad \Delta_m X_t = X_t - X_{t-m}, \quad \Delta_m^D X_t = \Delta_m \cdot \Delta_m \cdot \Delta_m \cdot \dots \cdot \Delta_m X_t.$$



Πρόβλεψη $SARIMA(p, d, q)(P, D, Q)_m$

Η πρόβλεψη των τιμών ακολουθεί τους ίδιους κανόνες με αυτές των $ARIMA(p, d, q)$ μοντέλων.

- Το $SMA(Q)_m$ μέρος επιδρά μέχρι τις επόμενες $m \cdot Q$ εκτιμήσεις.
- Το $SAR(P)_m$ μέρος επιδρά στο σύνολο των μελλοντικών προβλέψεων.



Αναγνώριση ACF και PACF στο SARIMA(p, d, q)(P, D, Q)_m

Η εποχικότητα περιόδου m αντανακλάται σε σημαντικές συσχετίσεις στα lags $m, 2m, 3m, \dots$

Η αναγνώριση των εποχικών $AR(P)_m$ και $MA(Q)_m$ όρων γίνεται με την εξέταση του PACF και ACF αντίστοιχα, λαμβάνοντας υπόψη μόνο τα lags που είναι πολλαπλάσια του m .

Ενδεικτικά:

- $SAR(1)_m$: το PACF εμφανίζει σημαντική αιχμή στο lag m , ενώ το ACF φθίνει γεωμετρικά στα lags $m, 2m, 3m, \dots$
- $SMA(1)_m$: το ACF εμφανίζει απότομο cutoff στο lag m , ενώ το PACF φθίνει στα lags $m, 2m, 3m, \dots$

Η εποχική αναγνώριση γίνεται πάντα μετά την εφαρμογή της εποχικής διαφοράς D . Πριν από αυτήν, τα ACF/PACF διαγράμματα είναι παραπλανητικά.



Προσομοίωση SARIMA με τη γλώσσα R

```
library(astsa)
```

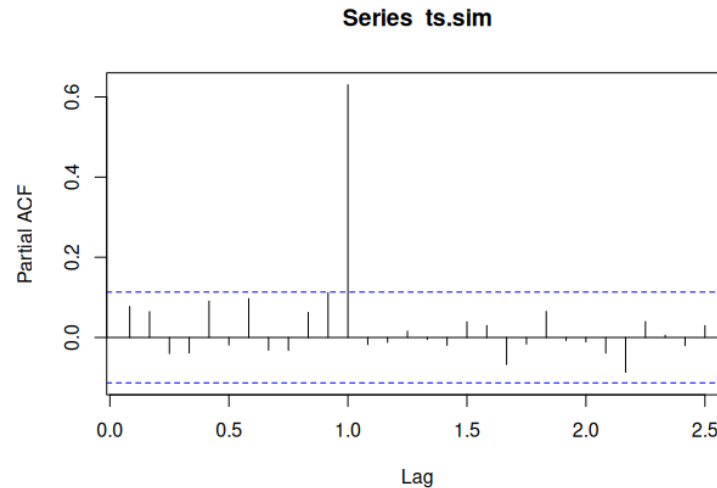
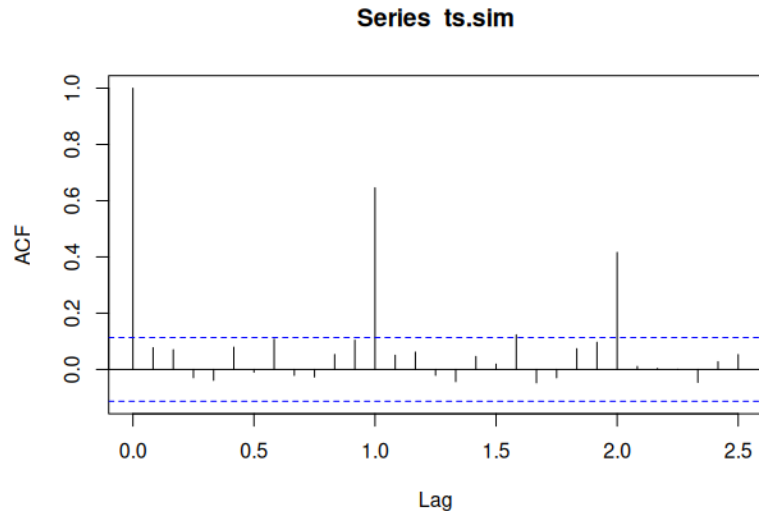
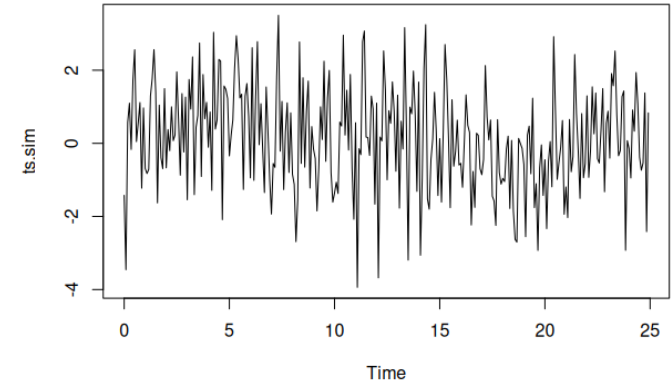
```
# SARIMA(0,0,0)(1,0,0)12
```

```
ts.sim = sarima.sim(D=0, sar = 0.6, S=12, n=300)
```

```
plot.ts(ts.sim)
```

```
acf(ts.sim, lag.max = 30) # ή acf(as.numeric(ts.sim), lag.max = 30)
```

```
pacf(ts.sim, lag.max = 30)
```





Προσομοίωση SARIMA με τη γλώσσα R

```
library(astsa)
```

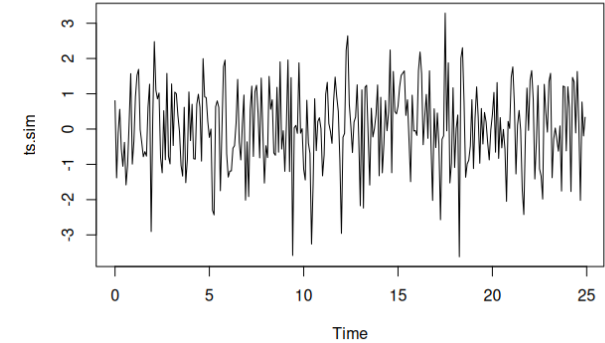
```
# SARIMA(0,0,0)(0,0,1)12
```

```
ts.sim = sarima.sim(D=0, sma = 0.6, S=12, n=300)
```

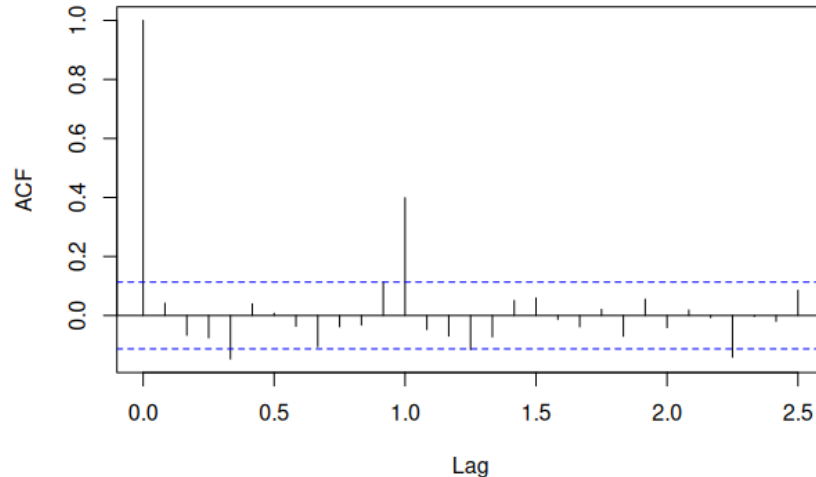
```
plot.ts(ts.sim)
```

```
acf(ts.sim, lag.max = 30) # ή acf(as.numeric(ts.sim), lag.max = 30)
```

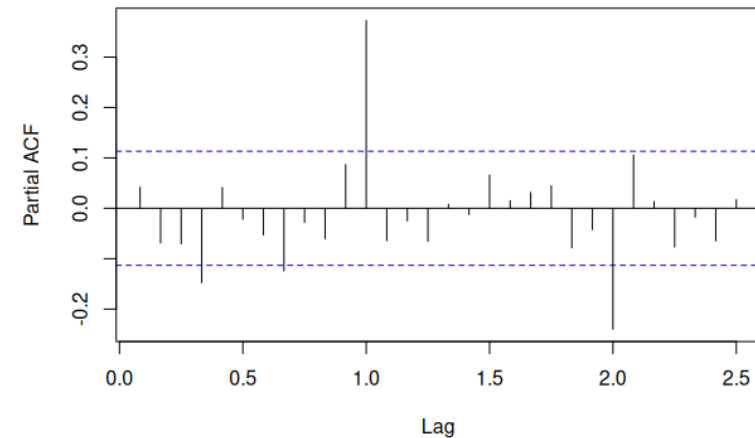
```
pacf(ts.sim, lag.max = 30)
```



Series ts.sim



Series ts.sim





Προσομοίωση SARIMA με τη γλώσσα R

```
library(astsa)
```

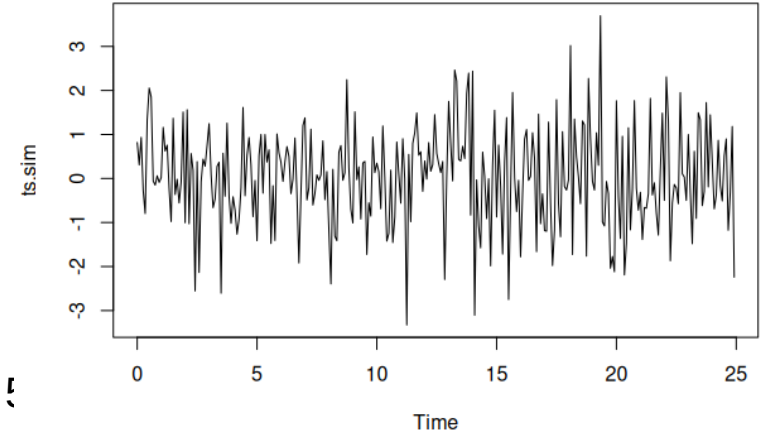
```
# SARIMA(0,0,0)(0,0,2)12
```

```
ts.sim = sarima.sim(D=0, sma = c(-0.45, -0.45), S=12, n=300)
```

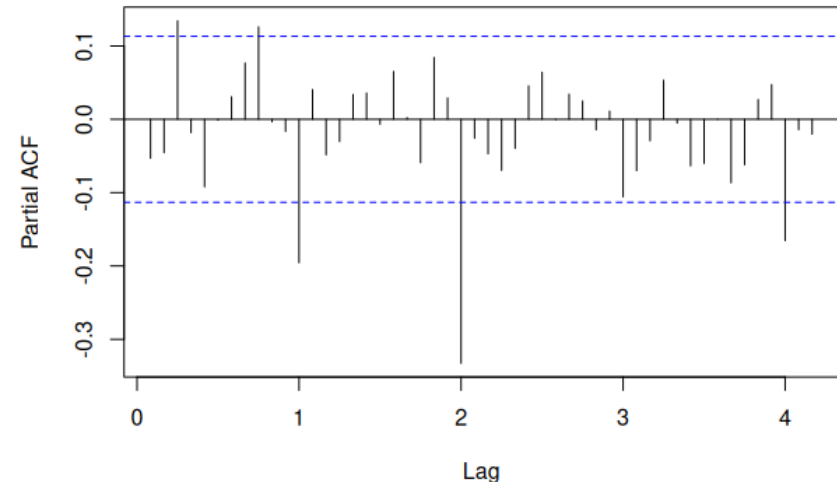
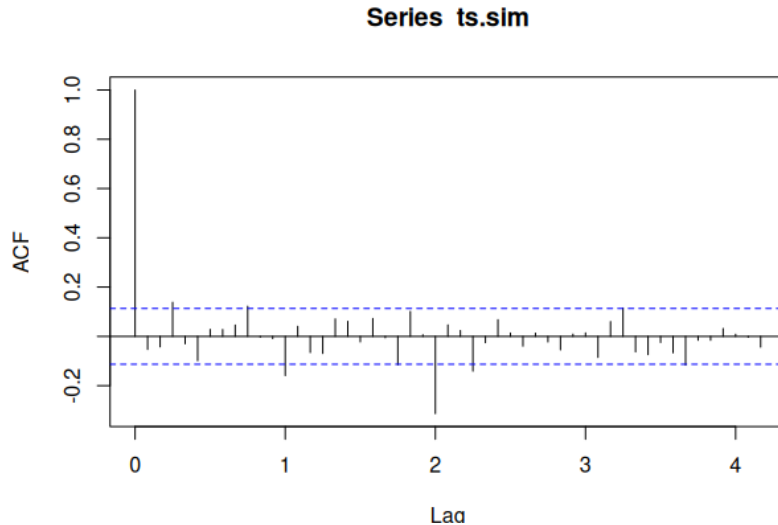
```
plot.ts(ts.sim)
```

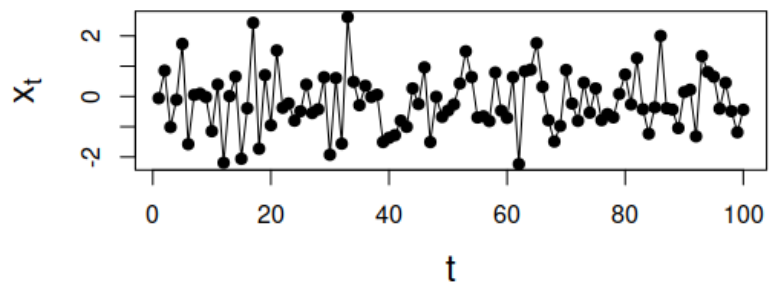
```
acf(ts.sim, lag.max = 50) # ή acf(as.numeric(ts.sim), lag.max = 50)
```

```
pacf(ts.sim, lag.max = 50)
```

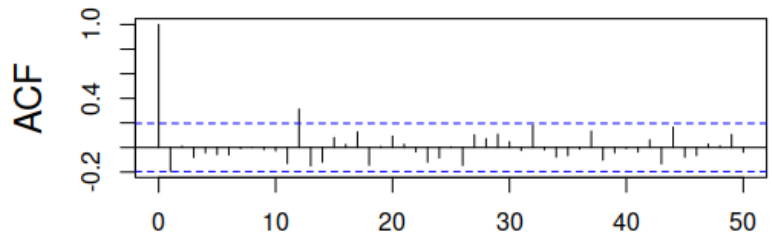


Series ts.sim

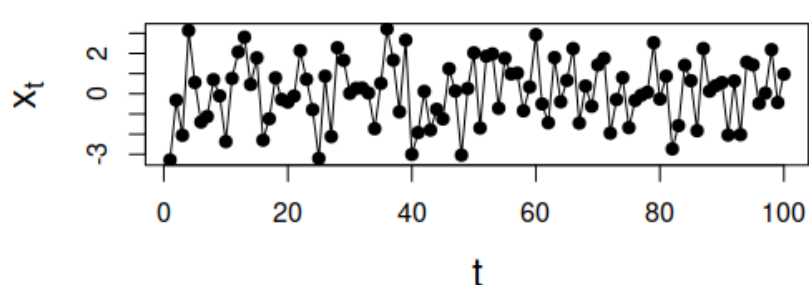
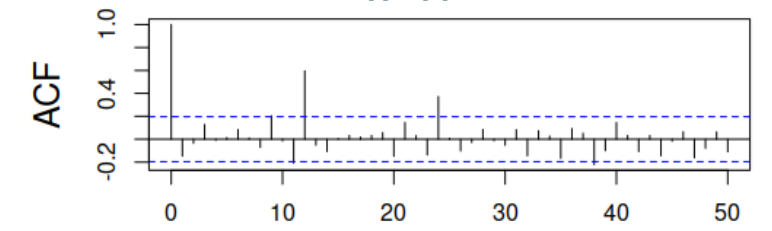




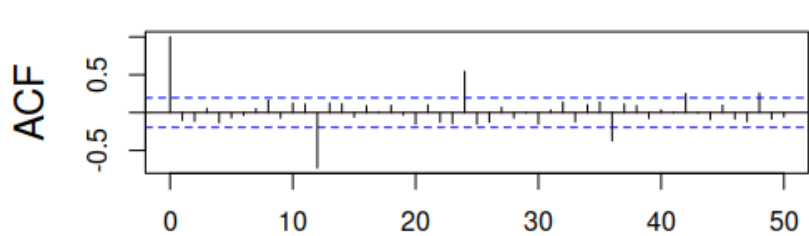
Διάγραμμα 1



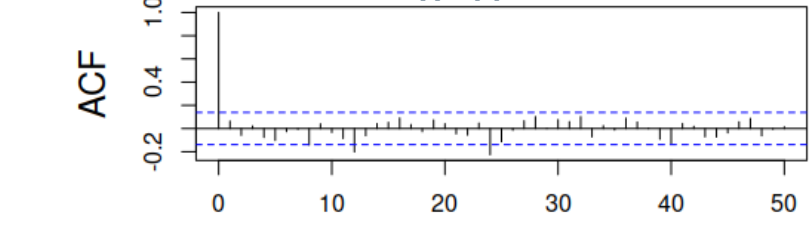
Διάγραμμα 3



Διάγραμμα 2



Διάγραμμα 4



SARIMA(0,0,0)(1,0,0)₁₂

SARIMA(0,0,0)(0,0,1)₁₂

SARIMA(0,0,0)(0,0,2)₁₂

Δραστηριότητα

Αντιστοιχίστε τα 4
χρονοδιαγράμματα με τα
αντίστοιχα μοντέλα.



SARIMA(0, 0, 0)(0, 1, 0)₁₂

Παράδειγμα: Ένα SARIMA(0, 0, 0)(0, 1, 0)₁₂ μοντέλο, χωρίς σταθερά, έχει εξίσωση της μορφής

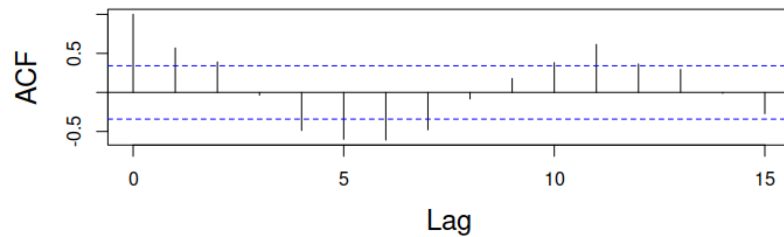
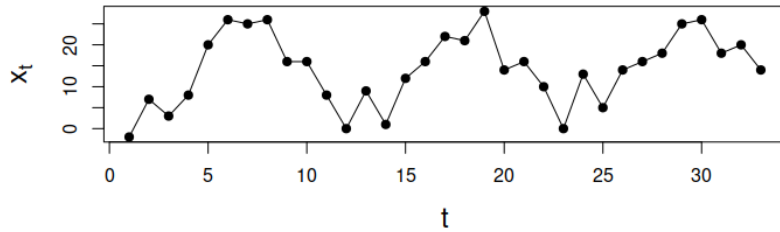
$$\Delta^{12}X_t = w_t \text{ ή } (1 - B^{12})X_t = w_t \text{ ή } X_t - X_{t-12} = w_t.$$

temp = c(-2, 7, 3, 8, 20, 26, 25, 26, 16, 16, 8, 0, 9, 1, 12, 16, 22, 21, 28, 14, 16, 10, 0, 13, 5, 14, 16, 18, 25, 26, 18, 20, 14)

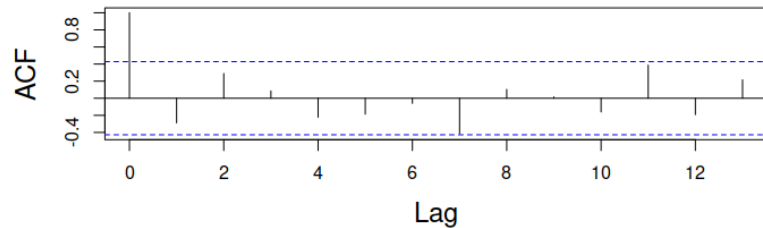
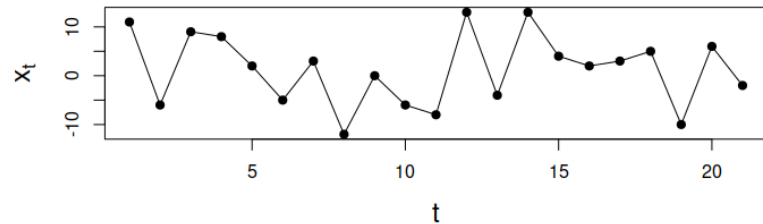
Οι διαφορές των παραπάνω ανά 12 είναι οι εξής:

$X_{13} - X_1 = 11$, $X_{14} - X_2 = -6, \dots$ 9, 8, 2, -5, 3, -12, 0, -6, -8, 13, -4, 13, 4, 2, 3, 5, -10, 6, -2.

Χρονοδιάγραμμα



Χρονοδιάγραμμα





SARIMA(0, 0, 0)(0, 0, 1)₁₂

Παράδειγμα: Ένα SARIMA(0, 0, 0)(0, 0, 1)₁₂ μοντέλο, χωρίς σταθερά, έχει εξίσωση της μορφής

$$X_t = (I + \Theta B^{12})w_t \text{ ή}$$

$$X_t = (I + \Theta B^{12})w_t \text{ ή}$$

$$X_t = w_t + \Theta w_{t-12} .$$



SARIMA(0, 0, 0)(0, 1, 1)₁₂

Παράδειγμα: Ένα SARIMA(0, 0, 0)(0, 1, 1)₁₂ μοντέλο, χωρίς σταθερά, έχει εξίσωση της μορφής

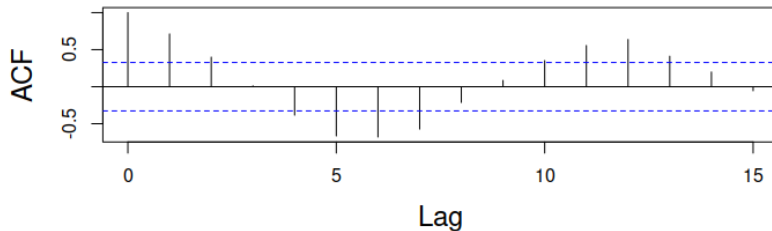
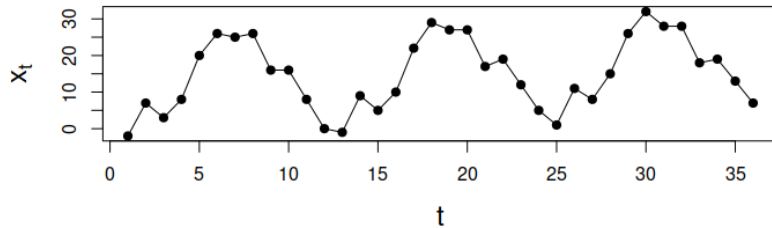
$$\Delta^{12}X_t = (I + \Theta B^{12})w_t \text{ ή } (1 - B^{12})X_t = (I + \Theta B^{12})w_t \text{ ή } X_t - X_{t-12} = w_t + \Theta w_{t-12}.$$

temp2 = c(-2, 7, 3, 8, 20, 26, 25, 26, 16, 16, 8, 0, -1, 9, 5, 10, 22, 29, 27, 27, 17, 19, 12, 5, 1, 11, 8, 15, 26, 32, 28, 28, 18, 19, 13, 7)

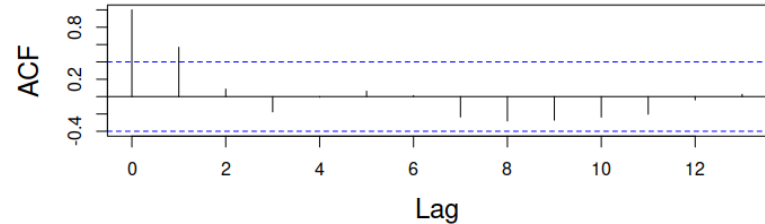
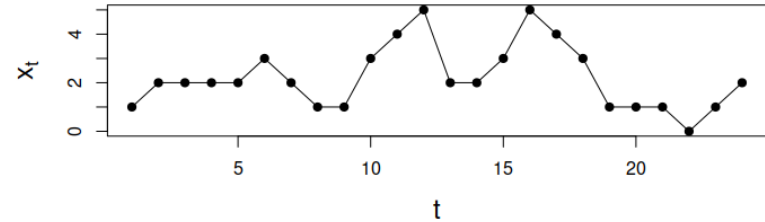
Οι διαφορές των παραπάνω ανά 12 είναι οι εξής:

$X_{13} - X_1 = 1$, $X_{14} - X_2 = 2, \dots$ 2, 2, 2, 3, 2, 1, 1, 3, 4, 5, 2, 2, 3, 5, 4, 3, 1, 1, 1, 0, 1, 2

Χρονοδιάγραμμα



Χρονοδιάγραμμα





SARIMA(1, 0, 0)(0, 1, 1)₁₂

Παράδειγμα

Ένα SARIMA(1, 0, 0)(0, 1, 1)₁₂ μοντέλο, χωρίς σταθερά, έχει εξίσωση της μορφής

$$(1 - B^{12})X_t = \varphi(1 - B^{12})X_{t-1} + (1 + \Theta B^{12})w_t$$

ή

$$X_t - X_{t-12} = \varphi X_{t-1} - \varphi X_{t-13} + w_t + \Theta w_{t-12},$$



SARIMA(0, 2, 1)(0, 0, 1)₁₂

Παράδειγμα

Ένα SARIMA(0, 2, 1)(0, 0, 1)₁₂ μοντέλο, χωρίς σταθερά, έχει εξίσωση της μορφής

$$(1 - B)^2 X_t = (1 + \theta B)(1 + \Theta B^{12}) w_t,$$

ή

$$X_t - 2X_{t-1} + X_{t-2} = w_t + \Theta w_{t-12} + \theta w_{t-1} + \theta \Theta w_{t-13}.$$

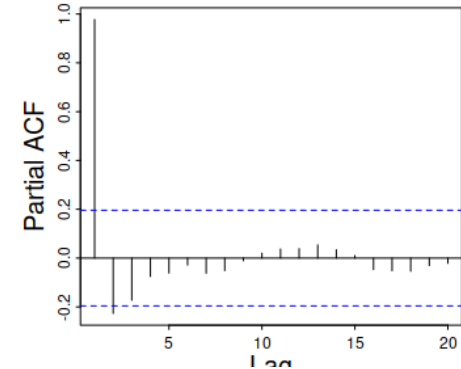
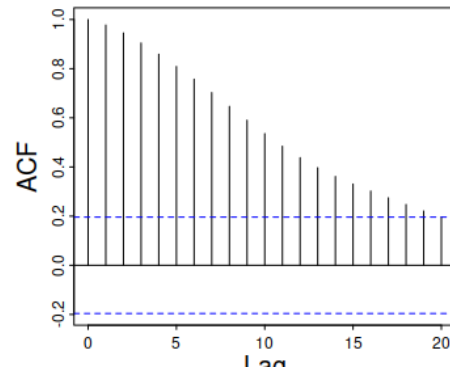
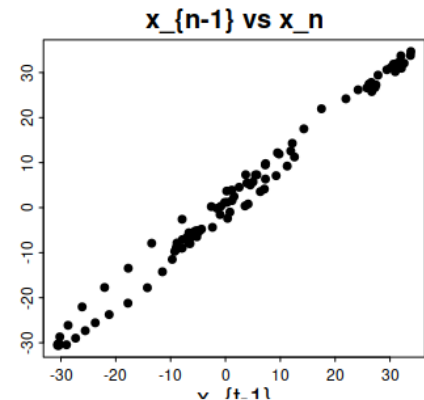
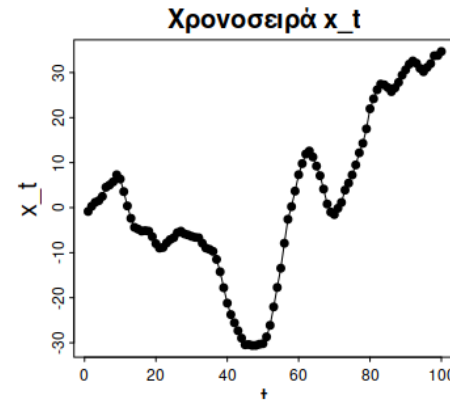


SARIMA(0, 2, 1)(0, 0, 1)₁₂

ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΡΑΚΗΣ
DEMOCRITUS UNIVERSITY OF THRACE

Παράδειγμα

$$X_t - 2X_{t-1} + X_{t-2} = \varepsilon_t - 0.2\varepsilon_{t-12} + 0.1\varepsilon_{t-1} - 0.02\varepsilon_{t-13}.$$



```
library(astsa)
x = sarima.sim(ar = 0, d=2, ma=0.1, sar = 0, D=0, sma=-0.2, S=12, n=100)
my.plot.ts.simple(x)
```



SARIMA(2, 1, 0)(0, 2, 2)₁₂

Παράδειγμα

Ένα ARIMA(2, 1, 0)(0, 2, 2)₁₂ μοντέλο, χωρίς σταθερά, έχει εξίσωση της μορφής

$$(1 - \phi_1 B - \phi_2 B^2)(1 - B)(1 - B^{12})^2 Y_t = (1 + \Theta_1 B^{12} + \Theta_2 B^{24}) e_t$$

$$(1 - \phi_1 B - \phi_2 B^2)(1 - B)(1 - 2B^{12} + B^{24}) Y_t = (1 + \Theta_1 B^{12} + \Theta_2 B^{24}) e_t$$

$$(1 - \phi_1 B - \phi_2 B^2 - B + \phi_1 B^2 + \phi_2 B^3)(1 - 2B^{12} + B^{24}) Y_t \\ = (1 + \Theta_1 B^{12} + \Theta_2 B^{24}) e_t$$

$$(1 - \phi_1 B - \phi_2 B^2 - B + \phi_1 B^2 + \phi_2 B^3 - 2B^{12} + 2\phi_1 B^{13} + 2\phi_2 B^{14} + 2B^{13} \\ - 2\phi_1 B^{14} - 2\phi_2 B^{15} + B^{24} - \phi_1 B^{25} - \phi_2 B^{26} - B^{25} + \phi_1 B^{26} + \phi_2 B^{27}) Y_t \\ = (1 + \Theta_1 B^{12} + \Theta_2 B^{24}) e_t$$

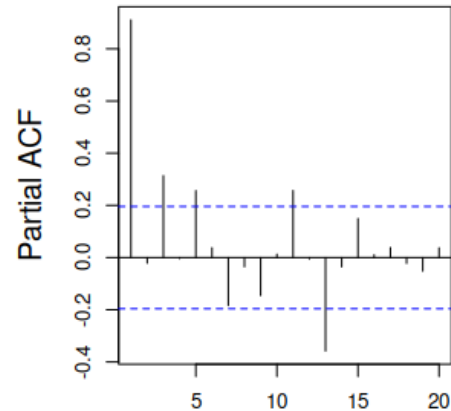
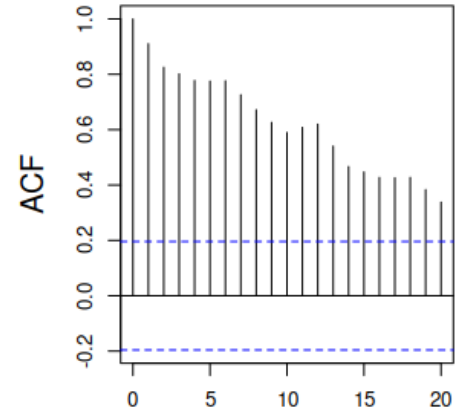
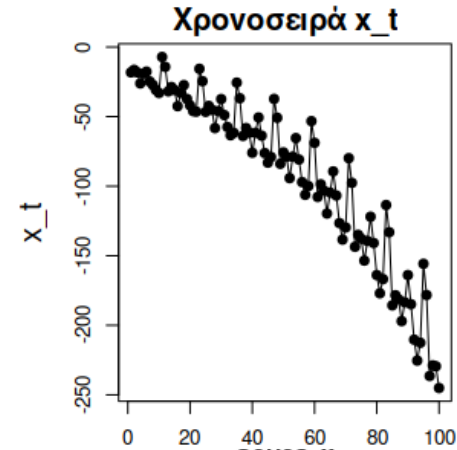
$$Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - Y_{t-1} + \phi_1 Y_{t-2} + \phi_2 Y_{t-3} - 2Y_{t-12} + 2\phi_1 Y_{t-13} + 2\phi_2 Y_{t-14} \\ + 2Y_{t-13} - 2\phi_1 Y_{t-14} - 2\phi_2 Y_{t-15} + Y_{t-24} - \phi_1 Y_{t-25} - \phi_2 Y_{t-26} - Y_{t-25} + \phi_1 Y_{t-26} \\ + \phi_2 Y_{t-27} = e_t + \Theta_1 e_{t-12} + \Theta_2 e_{t-24}$$



SARIMA(2, 1, 0)(0, 2, 2)₁₂

ΔΗΜΟΚΡΙΤΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΡΑΚΗΣ
DEMOCRITUS UNIVERSITY OF THRACE

$$Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - Y_{t-1} + \phi_1 Y_{t-2} + \phi_2 Y_{t-3} - 2Y_{t-12} + 2\phi_1 Y_{t-13} + 2\phi_2 Y_{t-14} + 2Y_{t-13} - 2\phi_1 Y_{t-14} - 2\phi_2 Y_{t-15} + Y_{t-24} - \phi_1 Y_{t-25} - \phi_2 Y_{t-26} - Y_{t-25} + \phi_1 Y_{t-26} + \phi_2 Y_{t-27} = e_t + \Theta_1 e_{t-12} + \Theta_2 e_{t-24}$$



```
library(astsa)
```

```
x = sarima.sim(ar = c(0.1, -0.3), d=1, ma=0, sar = 0, D=2, sma=c(-0.2, 0.1), S=12, n=100)
```

```
my.plot.ts.simple(x)
```



Παράδειγμα I

```
ts.value = c(1661980, 1357840, 1678840, 1859050, 1859540, 1814700, 2280190, 2240690, 2072170,  
1872380, 1693170, 1857710, 1609670, 1523130, 1919710, 1879330, 1723990, 2160040, 2318030,  
2308970, 2092730, 1975510, 1907070, 1904990, 1686480, 1511710, 1840750, 2083950, 1946310,  
1917100, 2315390, 2393510, 2087530, 1961500, 1807630, 1877010, 1709370, 1621490, 1946690,  
2010600, 2069190, 2031980, 2318730, 2451960, 2213610, 2196030, 1880220)
```

Από το χρονοδιάγραμμα είναι εμφανής η ύπαρξη τάσης και περιοδικότητα με περίοδο 12. Άρα, θα προχωρήσουμε σε διαφορά με lag 12.

```
ts.value.D1 <- diff(ts.value, lag = 12, differences = 1)
```



Παράδειγμα Ι

Από το νέο ACF διάγραμμα των διαφορών $\Delta_{12}X_t = X_{t+12} - X_t = (I - B^{12})X_t$ παρατηρούμε μία αρνητική αυτοσυσχέτιση σε lag 12, πρότυπο που ταιριάζει σε μοντέλο MA(1) και θα μπορούσε να εξηγηθεί από έναν MA(12) όρο στην αρχική χρονοσειρά. Δηλαδή, υποδεικνύεται πως ένα μοντέλο που πιθανώς να ταιριάζει στη χρονοσειρά είναι το

$$\Delta_{12}X_t = w_t + \Theta w_{t-12}.$$

Εφαρμόζουμε το μοντέλο με τον κώδικα:

```
library(forecast)
my.model = Arima(ts.value, order = c(0, 0, 0), seasonal = list(order = c(0, 1, 1), period = 12))
print(my.model)

checkresiduals(my.model)
```



Παράδειγμα Ι

Η εφαρμογή του παραπάνω μοντέλου, οδηγεί στην εξίσωση:

$$X_{t+12} - X_t = w_t - 0.3507w_{t-12}$$

Ωστόσο, φαίνεται πως τα υπόλοιπα εξακολουθούν να έχουν κάποια αυτοσυσχέτιση σε lag 12.

Πειραματιζόμενοι, καταλήγουμε πως μία αποδεκτή προσαρμογή παρουσιάζει ένα

ARIMA(1,0,0)(0,1,1)₁₂ μοντέλο: $(1-B^{12})X_t = \phi(1-B^{12})X_{t-1} + (1+\Theta B^{12})w_t$

ή

$$X_t - X_{t-12} = \phi X_{t-1} - \phi X_{t-13} + w_t + \Theta w_{t-12},$$

```
library(forecast)
```

```
my.model = Arima(ts.value, order = c(0, 0, 1), seasonal = list(order = c(0, 1, 1), period = 12))
```

```
print(my.model)
```

```
checkresiduals(my.model)
```



Παράδειγμα I

Παρατηρούμε, ότι το $ARIMA(1,0,0)(0,1,1)_{12}$ μοντέλο που υπολογίζεται είναι το:

$$X_t - X_{t-12} = 0.3855(X_{t-1} - X_{t-13}) + w_t - 0.7336w_{t-12}$$

Δημιουργούμε τα διάγραμμα ACF, PACF των υπολοίπων του μοντέλου.

```
acf(my.model$residuals, lag.max = 55)  
pacf(my.model$residuals, lag.max = 55)
```

Μία γραφική αναπαράσταση των προβλέψεων μπορούμε να αποκτήσουμε με τις εντολές:

```
my.forecast = forecast(my.model)  
summary(my.forecast)
```

```
plot(my.forecast, showgap = FALSE)
```



Δραστηριότητα Ι

ts.value = c(-0.25,1.82,0.13,3.27,-2.44,-2.77,-1.32,-0.44,3.61,3.31,1.78,-0.1,-2.71,1.29,0.2,2.9,-2.55,-4.96,-1.11,-
0.68,4.09,1.99,2.28,0.02,-2.83,2.64,0.58,2.7,-1.06,-5.09,1.04,-1.21,1.56,1.15,1.04,0.24,-0.61,3.94,0.86,0.7,-0.73,-
1.38,2.12,-2.78,-1.22,1.38,0.97,0.02,-1.07,2.23,2.55,-1.47,-1.24,-0.47,1.54,-1.8,-2.14,0.67,-0.51,0.31,-2.74,-0.38,4.7,-
1.46,-0.94,-0.91,2.24,0.53,-1.85,0.71,-3.4,0.31,-3.3,-2.06,2.96,-1.26,-2.28,-2.21,3.22,1.16,-3.45,0.39,-3.17,0.73,-
3.83,-1.11,0.36,1.24,-3.57,-3.46,3.33,0.01,-4.51,-1.33,-1.6,1.31,-2.9,0.98,-1.19,1.95,-1.18,-3.35,0.9,-1.01,-3.71,-2.29,-
1.49,0.72,-0.65,1.14,-2.36,1.29,1.79,-4.67,-1.34,2.12,-3.42,-3.14,-0.28,1.92,-0.27,-0.17,-1.63,2.07,2.41,-5.52,-
4.06,1.87,-4.26,-3.66,-0.01,2.46,-0.48,0.58,-0.91,2.48,2.12,-5.39,-5.7,0.97,-4.93,-2.46,-1.13,0.27,-1.27,0.47,-
0.31,2.08,0.72,-3.81,-5.14,-0.11,-2.93,-2.17,-0.94,-1.35,-2.24,-0.14,0.58,2.2,-0.53,-2.19,-2.71,-1.78,0.08,-1.46,-2.33,-
1.98,-2.26,0.88,2.81,0.78,-1.28,-2.08,-1.13,-1.06,0.63,-2.33,-1.01,-1.4,-2.38,0.24,2.22,-0.01,-1.38,-2.37,-1.89,-
0.53,0.24,-3.08,-0.51,0.58,-2.76,-1.62,2.98,0.11,0.33,-0.57,-2.38,1.25,-1.19,-2.34,-0.56,0.15,-0.43,-1.78,3.67,-
0.31,0.1,-0.04,-2.68,1.18,-2.2,-1.16,-0.34,-1.03,1.08,-0.59,4.89,-0.62,-2.02,-2.27,-1.26,-0.39,-1.66,-0.57,-1.12,-1.8,-
0.76,-0.17,5.43,-0.26,-1.95,-3.47,1.16,-0.15,-2.1,-0.42,-2.88,-1.87,-0.98,1.19,2.62,-0.18,-0.92,-4.05,1.88,-0.87,-3.32,-
0.46,-4.36,-0.94,-1.69,2.5,0.47,-2.99,-1.26,-3.57,1.71,-0.37,-4.65,1.23,-3.61,-1.53,-1.88,0.94,0.46,-4.96,-1.87,-
0.62,1.23,0.04,-4.65,2.56,-2.74,-0.88,0.49,0.09,1.29,-3.51,-2.14,1.86,0.6,-0.65,-4.66,1.56,-0.53,-0.45,2.23,2.44,1.6,-
2.82,-1.15,1.34,-0.09,-1.23,-4.8,1.09,1.31,1.54)



Δραστηριότητα II

rainXanthi = c(8, 4.4, 9.4, 29.4, 135.6, 65, 70.9, 15.2, 38.8, 87.2, 5.2, 0.2, 7.8, 46.4, 35.9, 77.4, 34.2, 157.4, 46.4, 19, 43.8, 41.2, 96.8, 25.6, 40.2, 111.8, 111.8, 49.8, 39.4, 9.6, 11.6, 8.6, 44.2, 41, 4.6, 36.2, 12.4, 45.8, 0, 30.8, 134.6, 167.2, 13.8, 56.6, 112.3, 13.6, 18.8, 18.2, 7, 40.4, 30.8, 130.2, 234.6, 106.2, 87.2, 15, 7.6, 63, 18, 2.6, 28, 24, 153.2, 24.4, 69.6, 27, 134, 181.6, 46, 85.4, 18.6, 32, 83.6, 42.6, 32.8, 127.5, 92.8, 122, 129.6, 35.4, 20.6, 88, 14.8, 12.8, 33.8, 58.6, 104.2, 0.2, 86.8, 35.6, 64.8, 7.8, 65, 76.6, 8.8, 1.2, 9.8, 21.2, 46, 4.6, 25, 44.4, 66.8, 29.4, 72.4, 7.8, 29.2, 4.6, 29.4, 54.6, 82.6, 103.6, 13, 51.2, 94, 1.2, 17.6, 180.4, 170.2, 0, 25.8, 46.8, 119.4, 123.6)



Δραστηριότητα III

Προσπαθήστε να δημιουργήσετε μοντέλο πρόβλεψης της κατανάλωσης μπίρας στην Αυστραλία, αξιοποιώντας ιστορικά δεδομένα ανά τρίμηνο από το 1992 έως το 2008:

```
library(fpp2)  
data(ausbeer)  
ts.value <- window(ausbeer, start=1992)  
head(ts.value)
```



Δεύτερη εργασία μαθήματος

Δίνονται τα δεδομένα

st1 = c(1196, 2795, 2959, 2868, 2953, 2724, 2349, 1930, 2217, 2163, 1441, 2192, 2045, 3331, 4546, 3697, 3438, 3118, 2582, 2515, 2727, 1950, 1187, 836, 1023, 2410, 3274, 3510, 3143, 3700, 2820, 1780, 2980, 2175, 1616, 1260, 1672, 2795, 3255, 3946, 2094, 1850, 1730, 1490, 1665, 2120, 1747, 1388, 1660, 1970, 1560, 2440, 1890, 1690, 2397, 1718, 2600, 1770, 580, 365, 1232, 2561, 2935, 3479, 3289, 2966, 2940, 3451, 3001, 1866, 1028, 1590, 1877, 1687, 3037, 2342, 4315, 3199, 2266, 1677, 1668, 1835, 1188, 1093, 1203, 2423, 2060, 2150, 1911, 2203, 2230, 2185, 2421, 1975, 1391, 2196, 1571, 1745, 1377, 1343, 1515, 1761, 2371, 2260, 2534, 1716, 1576, 1155, 1353, 1623, 1790, 1414, 1210, 1330, 1016, 1460, 1570, 1097, 1543, 1664, 1684, 1472, 1608, 1841, 2101, 1662, 1390, 1106, 1355, 1564, 1586, 1635, 1842, 1445, 1038, 1172, 1114, 934, 823, 1195, 1621, 1996, 1652, 1421, 2340, 1656, 1310, 1726, 1482, 1207, 816, 775, 897, 2202, 1577, 1951, 1834, 1408, 1452, 1542, 1563, 1029, 710, 698, 1145, 1617, 1590, 1812, 1549, 1245, 1202, 1183, 1379, 889, 515, 622, 1101, 1327, 1468, 1706, 1628, 1344, 1038, 1487, 1436, 984, 749, 726, 1007, 2727, 1588, 1723, 1768, 1404, 1728, 1369, 2728, 1589, 1723, 1769, 1404, 1373, 2035, 1722, 1279, 788, 888, 1526, 1804, 1836, 3461, 2323, 1070, 1013, 894, 1553, 916, 701, 374, 1175, 1804, 1834, 1116, 1758, 1441, 1260, 1565, 1897, 983, 469, 234, 1392, 3259, 1557, 2400, 2211, 1803, 1312, 1338, 2126, 1384, 733, 84, 587, 1868, 1438, 1897, 1970, 1352, 1013, 1051, 1987, 1267, 650, 432, 492)



Δεύτερη εργασία μαθήματος

(α) Αναπαραστήστε τη σειρά σε χρονοδιάγραμμα.

```
plot(st1, type = 'l')
```

(β) Χρησιμοποιήστε τη δοκιμασία Box–Pierce και επιβεβαιώστε πως πρόκειται για χρονοσειρά εξαρτημένων τιμών.

```
Box.test(st1, lag = 10, type = "Ljung-Box")
```

Διατυπώστε το αποτέλεσμα ακολουθώντας το υπόδειγμα:

“Η δοκιμασία Box–Pierce κατέδειξε πως το δείγμα των τιμών της χρονοσειράς προέρχεται από χρονοσειρά εξαρτημένων μεταβλητών ($\chi^2(\text{_____}) = \text{_____}, p = \text{_____}$)”

(γ) Δημιουργήστε το συσχετόγραμμα και προσδιορίστε τυχόν περιοδικότητα.

```
acf(st1, lag.max = 100)
```

(δ) Υπολογίστε τη διαφοροποιημένη σειρά και επαληθεύστε ότι η περιοδικότητα αφαιρέθηκε.

```
st1.D1 <- diff(st1, lag = 12, differences = 1)
```

```
plot(st1.D1, type = 'l')
```

```
acf(st1.D1, lag.max = 100)
```



Δεύτερη εργασία μαθήματος

Βάσει των θεωρητικών προβλέψεων για την αυτοσυσχέτιση ενός MA(1) μοντέλου, τι μπορείτε να συμπεράνετε για τη διαφοροποιημένη σειρά από το τελευταίο ACF;

(ε) Προσαρμόστε κατάλληλο μοντέλο SARIMA(p, d, q)(P, D, Q)m και ελέγξτε την ποιότητα της προσαρμογής του.

Ενδεικτικός κώδικας:

```
library(forecast)
```

```
my.model = Arima(st1, order = c(1, 0, 0), seasonal = list(order = c(0, 1, 1), period = 12))
```

```
print(my.model)
```

```
checkresiduals(my.model)
```

(στ) Πειραματιστείτε με τους δείκτες και βρείτε ένα μοντέλο που προσαρμόζεται όσο το δυνατόν καλύτερα στα δεδομένα, ελαχιστοποιώντας το δείκτη AIC.

(ζ) Προχωρήστε σε πρόβλεψη με τις εντολές:

```
my.forecast = forecast(my.model)
```

```
summary(my.forecast)
```

```
plot(my.forecast, showgap = FALSE)
```