

Χρονικές σειρές

6^ο μάθημα: Αυτοπαλίνδρομα μοντέλα (2)

Εαρινό εξάμηνο 2018-2019

Τμήμα Μαθηματικών ΑΠΘ

Διδάσκουσα: **Αγγελική Παπάνα**

Μεταδιδακτορική Ερευνήτρια

Πολυτεχνική σχολή, Α.Π.Θ. & Οικονομικό Τμήμα, Πανεπιστήμιο Μακεδονίας

<http://users.auth.gr/~agrapana/>

Μερική αυτοσυσχέτιση

Οι συντελεστές ρ_{ss} για διάφορες τιμές του s είναι η συνάρτηση μερικής αυτοσυσχέτισης:

$$\rho_{ss} = \frac{\begin{vmatrix} 1 & \rho_1 & \dots & \rho_1 \\ \rho_1 & 1 & \dots & \rho_2 \\ \dots & \dots & \dots & \dots \\ \rho_{s-1} & \rho_{s-2} & \dots & \rho_s \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \dots & \rho_{s-1} \\ \rho_1 & 1 & \dots & \rho_{s-2} \\ \dots & \dots & \dots & \dots \\ \rho_{s-1} & \rho_{s-2} & \dots & 1 \end{vmatrix}}$$

AR(1)

$$\rho_{11} = \rho_1 = \alpha_1$$

$$\rho_{ss} = 0, \text{ για } s > 1$$

AR(2)

$$\rho_{11} = \rho_1$$

$$\rho_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

$$\rho_{ss} = 0, \text{ για } s > 2$$

AR(p)

$$\rho_{11} = \rho_1 \neq 0$$

$$\rho_{22} \neq 0$$

...

$$\rho_{pp} \neq 0$$

$$\rho_{ss} = 0, \text{ για } s > p$$

Έλεγχος σημαντικότητας συντελεστών αυτοσυσχετίσης

Στην πράξη, επειδή οι αληθινές μερικές αυτοσυσχετίσεις ρ_{ss} και οι αληθινές αυτοσυσχετίσεις ρ_s δεν είναι γνωστές, χρησιμοποιούνται οι αντίστοιχες εκτιμήσεις τους από το δείγμα.

Με βάση τις εκτιμήσεις αυτές, μπορεί να γίνει έλεγχος σημαντικότητας των παραμέτρων στον πληθυσμό.

Για μεγάλα δείγματα οι εκτιμήσεις $\hat{\rho}_s$ των αυτοσυσχετίσεων ρ_s κατανέμονται κανονικά με μέσο μηδέν και διακύμανση $1/T$, όπου T το μέγεθος του δείγματος.

Το ίδιο ισχύει και για τις εκτιμήσεις $\hat{\rho}_{ss}$ των μερικών αυτοσυσχετίσεων ρ_{ss} , για υστερήσεις s μεγαλύτερες από την τάξη p της AR διαδικασίας.

Δηλαδή:

$$\hat{\rho}_s \sim N\left(0, \frac{1}{T}\right)$$

$$\hat{\rho}_{ss} \sim N\left(0, \frac{1}{T}\right) \text{ για } s > p$$

Ο έλεγχος της στατιστικής σημαντικότητας του συντελεστή ρ_s , δηλαδή ο έλεγχος της υποθέσεως:

$$H_0: \rho_s = 0$$

$$H_1: \rho_s \neq 0$$

γίνεται με την στατιστική:

$$t_s = \frac{\hat{\rho}_s}{\sqrt{\frac{1}{T}}} = \hat{\rho}_s \sqrt{T}$$

Για δεδομένο επίπεδο σημαντικότητας α , η μηδενική υπόθεση απορρίπτεται αν η τιμή του στατιστικού $t_s >$ κρίσιμη τιμή του t .

Επειδή για $T > 30$ η κρίσιμη τιμή του t για $\alpha = 5\%$ είναι περίπου ± 2 , η μηδενική υπόθεση απορρίπτεται αν

$$|t_s| > 2$$

Το 95% διάστημα εμπιστοσύνης για το ρ_s είναι:

$$\hat{\rho}_s - \frac{2}{\sqrt{T}} \leq \rho_s \leq \hat{\rho}_s + \frac{2}{\sqrt{T}}$$

Έλεγχος στατιστικής σημαντικότητας του μερικού συντελεστή αυτοσυσχέτισης

Ο συντελεστής ρ_{ss} είναι σημαντικός αν

$$|\hat{\rho}_{ss}\sqrt{T}| > 2$$

Με βάση τον παραπάνω έλεγχο σημαντικότητας των μερικών συντελεστών αυτοσυσχέτισης μπορεί να καθοριστεί η τάξη μιας AR διαδικασίας.

Δηλαδή εξετάζοντας την ακολουθία των τιμών t_s για $s = 1, 2, \dots$ επιλέγεται ως τάξη της σειράς αυτή που αντιστοιχεί στην τελευταία σημαντική τιμή του t_s .

Παράδειγμα

Έστω δείγμα 100 παρατηρήσεων και με βάση την σχέση

$$y_t = \rho_{1s}y_{t-1} + \rho_{2s}y_{t-2} + \rho_{3s}y_{t-3} + \dots + \rho_{ss}y_{t-s} + \varepsilon_t$$

προέκυψαν οι παλινδρομήσεις:

$$\hat{y}_t = 0.464y_{t-1}$$

$$\hat{y}_t = 0.212y_{t-1} + 0.493y_{t-2}$$

$$\hat{y}_t = 0.221y_{t-1} + 0.498y_{t-2} - 0.017y_{t-3}$$

Επομένως:

$$\hat{\rho}_{11} = 0.464 \quad \text{και} \quad t_1 = 0.464\sqrt{100} = 4.64 > 2$$

$$\hat{\rho}_{22} = 0.493 \quad \text{και} \quad t_2 = 0.493\sqrt{100} = 4.93 > 2$$

$$\hat{\rho}_{33} = -0.017 \quad \text{και} \quad t_3 = -0.017\sqrt{100} = -0.17 < 2$$

Τα $\hat{\rho}_{11}$, $\hat{\rho}_{22}$ στατιστικά σημαντικά, ενώ ο μερικός συντελεστής συσχέτισης $\hat{\rho}_{33}$ δεν είναι σημαντικός.

Επομένως μπορούμε να θεωρήσουμε ότι το δείγμα προέρχεται από μια AR(2) διαδικασία.

Εκτίμηση υποδειγμάτων AR

Αν ένα δείγμα προέρχεται από μια διαδικασία AR, με βάση την δειγματική (εκτιμημένη) συνάρτηση μερικής αυτοσυσχέτισης μπορεί να καθοριστεί η τάξη p του μοντέλου.

Αν υποθέσουμε ότι γνωρίζουμε την τάξη του μοντέλου, το ερώτημα είναι πως μπορούμε να εκτιμήσουμε τις παραμέτρους του.

1ος τρόπος Χρησιμοποιούμε τις **εξισώσεις Yule-Walker** ή την

σχέση $\mathbf{A} = \mathbf{\Pi}^{-1}\mathbf{R}$ αντικαθιστώντας τις αυτοσυσχετίσεις ρ_s με τις εκτιμήσεις $\hat{\rho}_s$ από το δείγμα, οι οποίες προκύπτουν από την σχέση:

$$\hat{\rho}_s = \frac{\sum_{t=1}^T (Y_t - \bar{Y})(Y_{t+s} - \bar{Y})}{\sum_{t=1}^T (Y_t - \bar{Y})^2}$$

ή

$$\hat{\rho}_s = \frac{T-1}{T-s} \frac{\sum_{t=1}^T (Y_t - \bar{Y})(Y_{t+s} - \bar{Y})}{\sum_{t=1}^T (Y_t - \bar{Y})^2}$$

2ος τρόπος Χρησιμοποιούμε την **μέθοδο ελαχίστων τετραγώνων**. Το μοντέλο AR(p) στην γενική μορφή του

$$Y_t = \delta + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + \varepsilon_t$$

μπορεί να θεωρηθεί ως ένα γραμμικό μοντέλο με p ανεξάρτητες μεταβλητές. Μόνο που οι μεταβλητές αυτές είναι στοχαστικές.

Οι εκτιμητές που προκύπτουν από την μέθοδο ελαχίστων τετραγώνων έχουν τις ιδιότητες μεγάλων δειγμάτων. Είναι δηλαδή συνεπείς και ακολουθούν την κανονική κατανομή.

Για ένα δείγμα T παρατηρήσεων έχουμε το ακόλουθο σύστημα $T - p$ εξισώσεων:

$$Y_{p+1} = \delta + \alpha_1 Y_p + \alpha_2 Y_{p-1} + \dots + \alpha_p Y_1 + \varepsilon_{p+1}$$

$$Y_{p+2} = \delta + \alpha_1 Y_{p+1} + \alpha_2 Y_p + \dots + \alpha_p Y_2 + \varepsilon_{p+2}$$

.....

$$Y_T = \delta + \alpha_1 Y_{T-1} + \alpha_2 Y_{T-2} + \dots + \alpha_p Y_{T-p} + \varepsilon_T$$

Με την χρήση πινάκων, το παραπάνω σύστημα γράφεται **$Y = X\beta + \varepsilon$**

όπου

$$Y = \begin{pmatrix} Y_{p+1} \\ Y_{p+2} \\ \dots \\ Y_T \end{pmatrix}, \beta = \begin{pmatrix} \delta \\ \alpha_1 \\ \dots \\ \alpha_p \end{pmatrix}, \varepsilon = \begin{pmatrix} \varepsilon_{p+1} \\ \varepsilon_{p+2} \\ \dots \\ \varepsilon_T \end{pmatrix}, X = \begin{pmatrix} 1 & Y_p & Y_{p-1} & Y_{p-2} & \dots & Y_1 \\ 1 & Y_{p+1} & Y_p & Y_{p-1} & \dots & Y_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & Y_{T-1} & Y_{T-2} & Y_{T-3} & \dots & Y_{T-p} \end{pmatrix}$$

Οι εκτιμητές ελαχίστων τετραγώνων δίνονται από την σχέση

$$\hat{\beta} = (X'X)^{-1}X'Y$$

όπου $\hat{\beta} = \begin{pmatrix} \hat{\delta} \\ \hat{\alpha}_1 \\ \dots \\ \hat{\alpha}_p \end{pmatrix}$.

Ο πίνακας των διακυμάνσεων – συνδιακυμάνσεων των εκτιμητών είναι:

$$s^2 = \frac{(Y - X\hat{\beta})'(Y - X\hat{\beta})}{T - p - (p + 1)} = \frac{\sum \hat{\varepsilon}^2}{T - 2p - 1}$$

Ένας εκτιμητής του μέσου της AR(p) διαδικασίας δίνεται από την σχέση

$$\hat{\mu} = \frac{\hat{\delta}}{1 - \hat{\alpha}_1 - \dots - \hat{\alpha}_p}$$

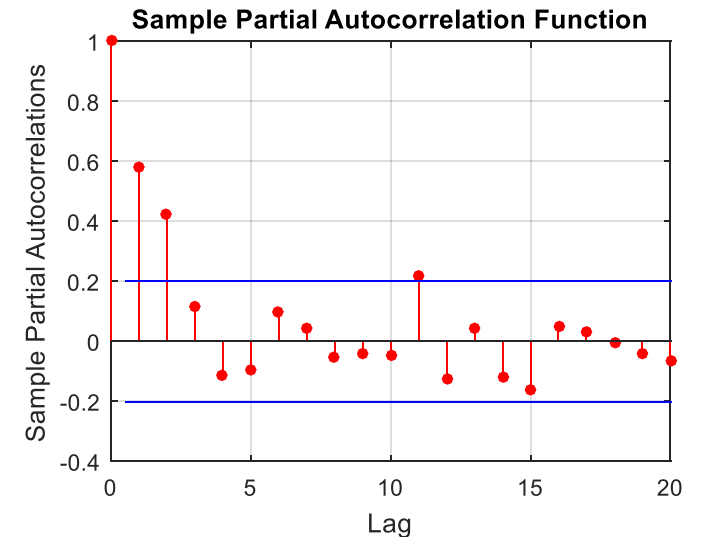
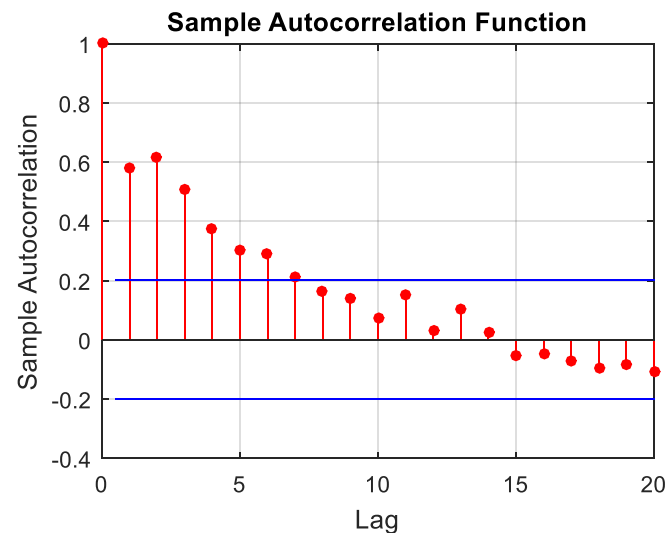
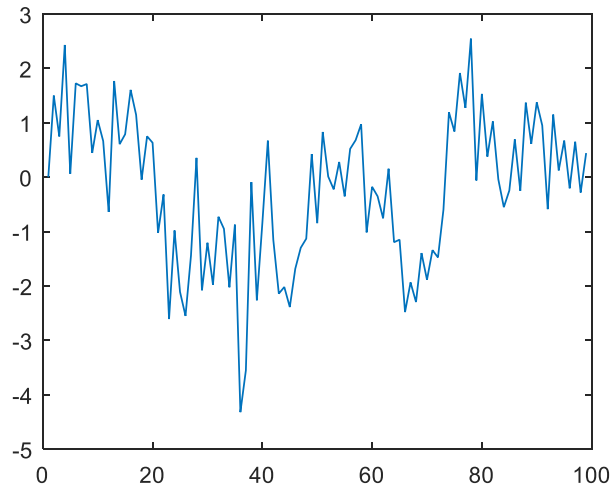
Παράδειγμα

Δίνονται 100 παρατηρήσεις μιας χρονολογικής σειράς Y . Οι παρατηρήσεις είναι «κατασκευασμένες» με βάση το μοντέλο

$$Y_t = 0.3Y_{t-1} + 0.5Y_{t-2} + \varepsilon_t$$

υποθέτοντας ότι ε_t λευκός θόρυβος. Η σειρά είναι στάσιμη.

Παρακάτω φαίνονται οι γραφικές παραστάσεις της χρονοσειράς, της συνάρτησης αυτοσυσχέτισης και μερικής αυτοσυσχέτισης.



```
>> [acM,lags,bounds]=autocorr(xV)
```

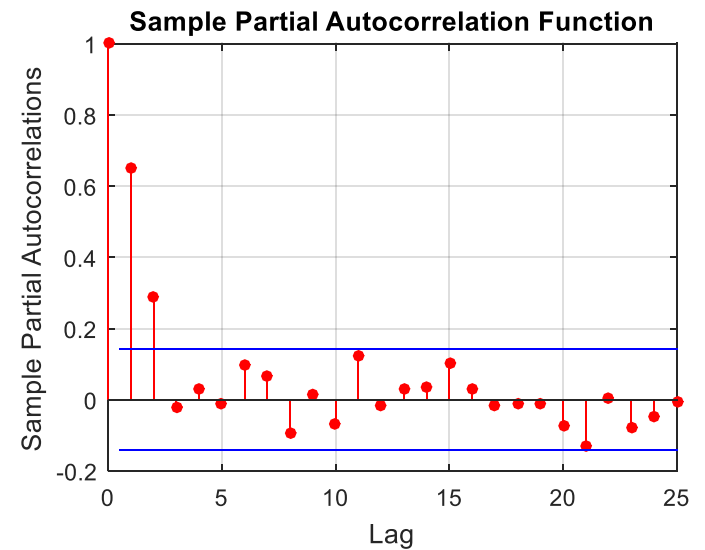
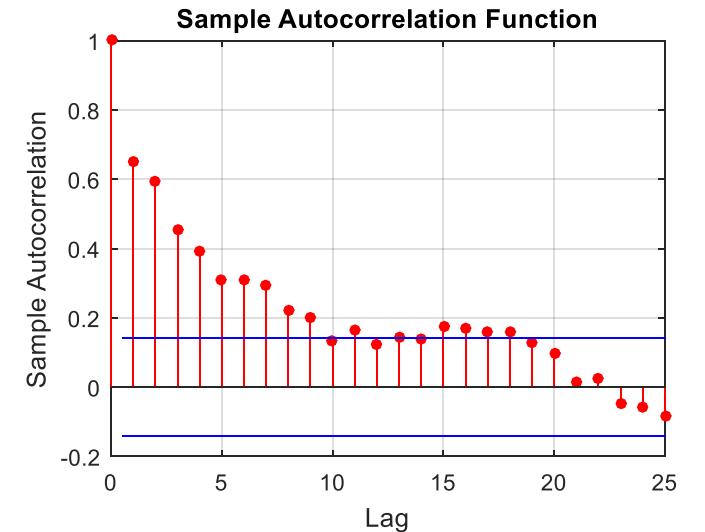
```
>> [pacM,lags,bounds]=parcorr(xV)
```

Υστέρηση	Συνάρτηση αυτοσυσχέτισης	Συνάρτηση μερικής αυτοσυσχέτισης
1	0.5836	0.5852
2	0.6277	0.4384
3	0.4832	0.0476
4	0.4415	0.0213

95% δ.ε. για τις αυτοσυσχετίσεις: [-0.20, 0. 20]

95% δ.ε. για τις μερικές αυτοσυσχετίσεις:

[-0.20, 0. 20]

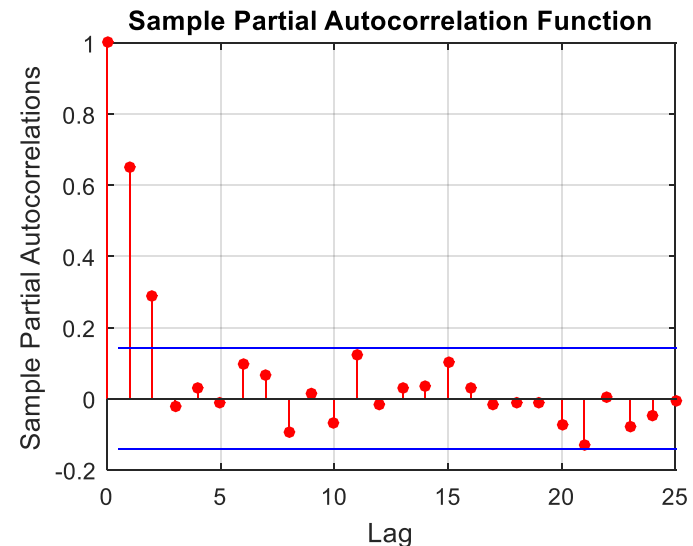
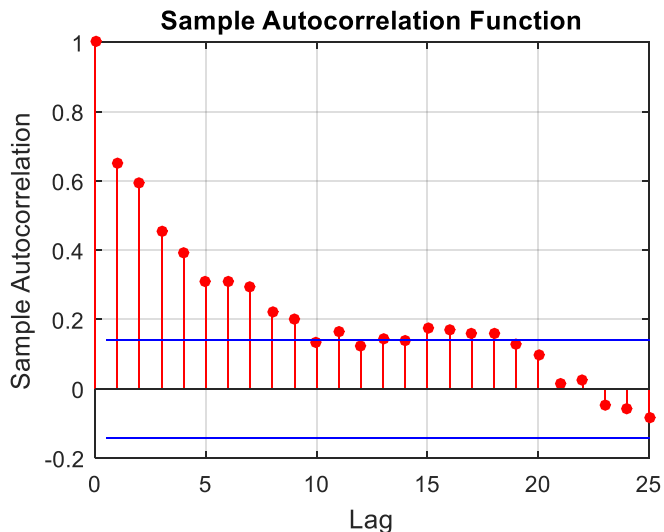


Οι πρώτες 9 αυτοσυσχετίσεις είναι σημαντικές.

Οι πρώτες 2 μερικές αυτοσυσχετίσεις είναι σημαντικές.

Για να εκτιμήσουμε το μοντέλο AR(2), θα χρησιμοποιήσουμε την μέθοδο των ελαχίστων τετραγώνων, δηλαδή θα εκτιμήσουμε το υπόδειγμα

$$Y_t = \delta + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \varepsilon_t$$



Εκτίμηση συντελεστών μοντέλου (Μέθοδος Yule-Walker)

>> MODEL = ar(xV,2,'yw')

MODEL =

Discrete-time AR model: $A(z)y(t) = e(t)$

$$A(z) = 1 - 0.3324 z^{-1} - 0.4432 z^{-2}$$

Sample time: 1 seconds

Parameterization:

Polynomial orders: na=2

Number of free coefficients: 2

Status:

Estimated using AR ('yw/ppw') on "xV".

Fit to estimation data: 27.89%

FPE: 0.9585, MSE: 0.9046

Εκτίμηση συντελεστών μοντέλου (Μέθοδος ελαχίστων τετραγώνων)

>> MODEL = ar(xV,2,'ls')

MODEL =

Discrete-time AR model: $A(z)y(t) = e(t)$

$$A(z) = 1 - 0.332 z^{-1} - 0.4439 z^{-2}$$

Sample time: 1 seconds

Parameterization:

Polynomial orders: na=2

Number of free coefficients: 2

Status:

Estimated using AR ('ls/now') on "xV".

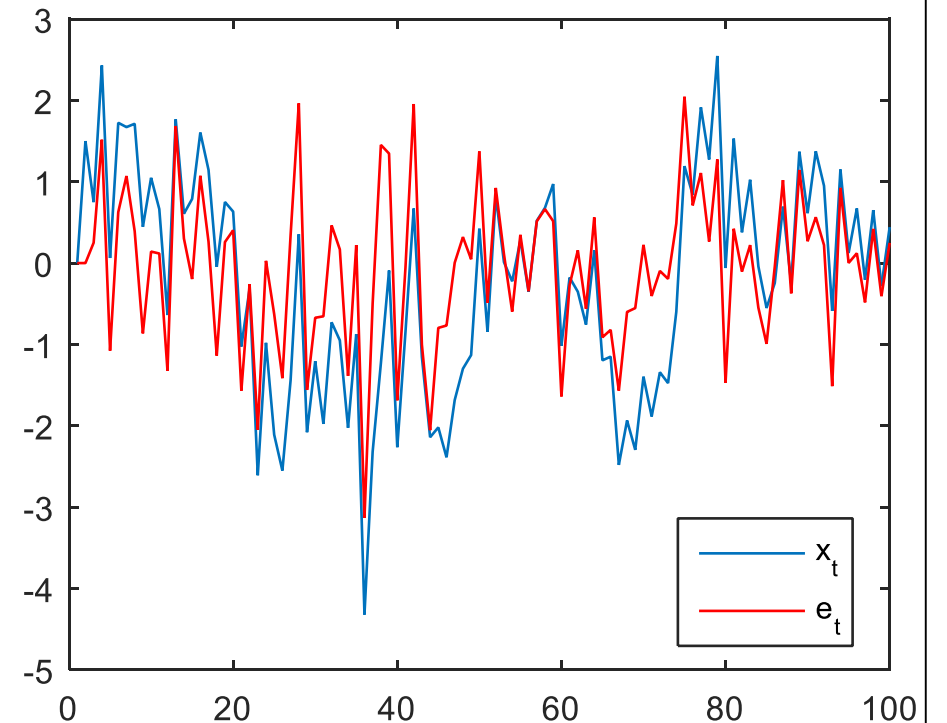
Fit to estimation data: 27.89%

FPE: 0.9599, MSE (mean square error): 0.9046

Οι εκτιμήσεις των συντελεστών του μοντέλου (και με τις 2 μεθόδους) δεν συμπίπτουν με τις πραγματικές, αλλά είναι πολύ κοντά.

Εκτίμηση υπολοίπων / διαταρακτικού όρου

```
>> et = resid(MODEL,xV)
```



Άσκηση 1

Έστω η ακόλουθη AR(1) διαδικασία $Y_t = 5 + 0.5Y_{t-1} + \varepsilon_t$. Το ε_t είναι λευκός θόρυβος με διακύμανση ίση με $\sigma_\varepsilon^2 = 4$.

α) Είναι η διαδικασία στάσιμη;

β) Να διατυπωθεί η σειρά ως γραμμικό φίλτρο.

γ) Να βρεθούν ο μέσος, οι αυτοσυνδιακυμάνσεις και αυτοσυσχετίσεις για υστερήσεις $s = 1, 2, 3$.

δ) Να γίνει το διάγραμμα αυτοσυσχέτισης.

Λύση

α) Η AR(1) διαδικασία $Y_t = 5 + 0.5Y_{t-1} + \varepsilon_t$ είναι στάσιμη γιατί ο συντελεστής της Y_{t-1} είναι μικρότερος από την μονάδα.

β) Ως γραμμικό φίλτρο η δοθείσα σειρά διατυπώνεται ως εξής:

$$Y_t = 5 + 0.5Y_{t-1} + \varepsilon_t \quad (\text{Θέτω } y_t = Y_t - \mu)$$

$$y_t = 0.5y_{t-1} + \varepsilon_t$$

$$y_t = 0.5(0.5y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t$$

$$y_t = 0.5^2 y_{t-2} + 0.5\varepsilon_{t-1} + \varepsilon_t$$

$$y_t = 0.5^2 (0.5y_{t-3} + \varepsilon_{t-2}) + 0.5\varepsilon_{t-1} + \varepsilon_t$$

$$y_t = 0.5^3 y_{t-3} + 0.5^2 \varepsilon_{t-2} + 0.5\varepsilon_{t-1} + \varepsilon_t$$

$$Y_t = \varepsilon_t + 0.5\varepsilon_{t-1} + 0.5^2 \varepsilon_{t-2} + 0.5^3 \varepsilon_{t-3} + \dots$$

γ) Ο μέσος δίνεται από την σχέση:

$$\mu = \frac{\alpha_0}{1-\alpha_1} = \frac{5}{1-0.5} = 10.$$

Οι αυτοσυνδιακυμάνσεις για υστερήσεις $s = 1,2,3$ είναι:

$$\gamma_0 = \sigma^2 \frac{1}{1-\alpha_1^2} = 4 \frac{1}{1-0.5^2} = 5.33$$

$$\gamma_1 = \alpha_1 \gamma_0 = 0.5 * 5.33 = 2.665$$

$$\gamma_2 = \alpha \gamma_0 = 0.5^2 * 5.33 = 1.3325$$

Οι αυτοσυσχετίσεις για υστερήσεις $s = 1,2,3$ είναι:

$$\rho_1 = \alpha_1 = 0.5$$

$$\rho_2 = \alpha_1^2 = 0.5^2 = 0.25$$

$$\rho_3 = \alpha_1^3 = 0.5^3 = 0.125$$

δ) Διάγραμμα αυτοσυσχέτισης

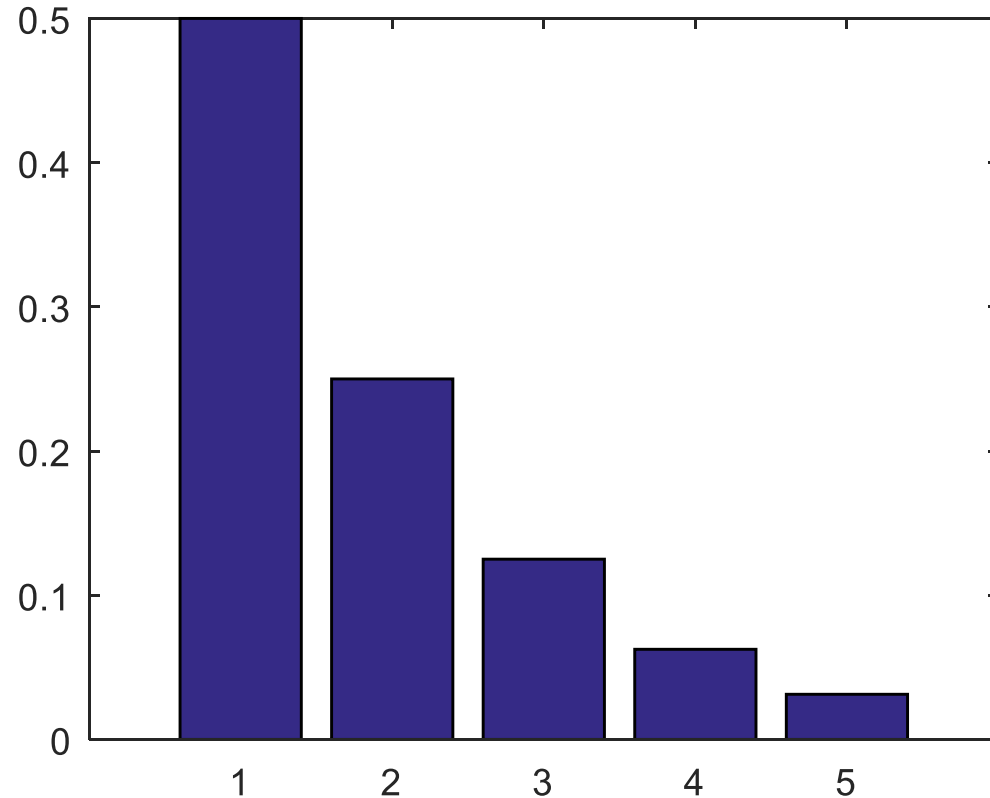
$$\rho_1 = \alpha_1 = 0.5$$

$$\rho_2 = \alpha_1^2 = 0.5^2 = 0.25$$

$$\rho_3 = \alpha_1^3 = 0.5^3 = 0.125$$

$$\rho_4 = \alpha_1^4 = 0.5^4 = 0.0625$$

$$\rho_5 = \alpha_1^5 = 0.5^5 = 0.0313$$



Άσκηση 2

Έστω η ακόλουθη στοχαστική διαδικασία

$$Y_t = 10 + 1.5Y_{t-1} - 0.6Y_{t-2} + \varepsilon_t.$$

Το ε_t είναι λευκός θόρυβος με διακύμανση ίση με $\sigma_\varepsilon^2 = 1$.

- α) Να διατυπωθεί η παραπάνω διαδικασία με τον συμβολισμό του τελεστή υστερήσεως L.
- β) Είναι η σειρά στάσιμη;
- γ) Ποιος είναι ο μέσος της σειράς;
- δ) Να διατυπωθούν και να λυθούν οι εξισώσεις Yule-Walker.
- ε) Να γίνει το διάγραμμα αυτοσυσχετίσης.
- στ) Να βρεθούν οι μερικές αυτοσυσχετίσεις και να γίνει διάγραμμα μερικών αυτοσυσχετίσεων.

Λύση

$$\alpha) Y_t = 10 + 1.5Y_{t-1} - 0.6Y_{t-2} + \varepsilon_t$$

$$\text{Θέτω } y_t = Y_t - \mu$$

$$y_t = 1.5y_{t-1} - 0.6y_{t-2} + \varepsilon_t$$

$$y_t - 1.5y_{t-1} + 0.6y_{t-2} = \varepsilon_t$$

$$(1 - 1.5L + 0.6L^2)y_t = \varepsilon_t$$

β) Η σειρά είναι στάσιμη επειδή ικανοποιούνται οι σχέσεις

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

$$\alpha_1 + \alpha_2 = 1.5 + (-0.6) = 0.9 < 1$$

$$-\alpha_1 + \alpha_2 = -1.5 + (-0.6) = -2.1 < 1$$

$$-1 < \alpha_2 = -0.6 < 1$$

γ) Ο μέσος της σειράς είναι:

$$\mu = \frac{\alpha_0}{1 - \alpha_1 - \alpha_2} = \frac{10}{1 - 1.5 - (-0.6)} = 100$$

δ) Οι εξισώσεις Yule-Walker είναι:

$$\begin{array}{l} \rho_1 = \alpha_1 + \alpha_2 \rho_1 \\ \rho_2 = \alpha_1 \rho_1 + \alpha_2 \end{array} \left. \vphantom{\begin{array}{l} \rho_1 = \alpha_1 + \alpha_2 \rho_1 \\ \rho_2 = \alpha_1 \rho_1 + \alpha_2 \end{array}} \right\} \begin{array}{l} \rho_1 = 1.5 - 0.6 \rho_1 \\ \rho_2 = \alpha_1 \rho_1 + \alpha_2 \end{array} \left. \vphantom{\begin{array}{l} \rho_1 = 1.5 - 0.6 \rho_1 \\ \rho_2 = \alpha_1 \rho_1 + \alpha_2 \end{array}} \right\} \begin{array}{l} \rho_1 = 0.94 \\ \rho_2 = 0.81 \end{array}$$

Για $s > 2$: $\rho_s = a_1 \rho_{s-1} + a_2 \rho_{s-2}$

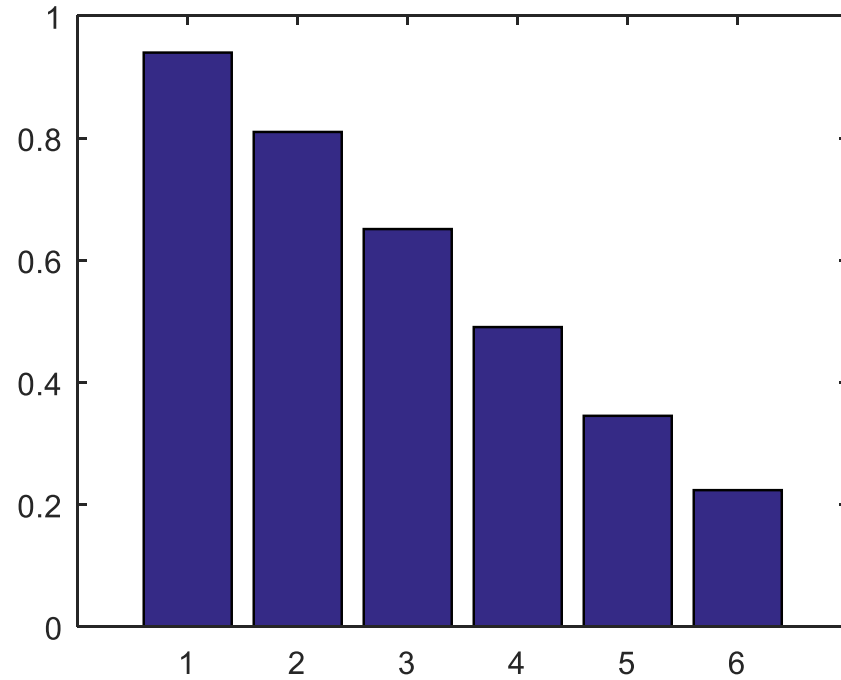
$$\rho_3 = a_1 \rho_2 + a_2 \rho_1 = 0.6510$$

$$\rho_4 = a_1 \rho_3 + a_2 \rho_2 = 0.4905$$

$$\rho_5 = a_1 \rho_4 + a_2 \rho_3 = 0.3452$$

$$\rho_6 = a_1 \rho_5 + a_2 \rho_4 = 0.2235$$

ε) Διάγραμμα αυτοσυσχέτισης.



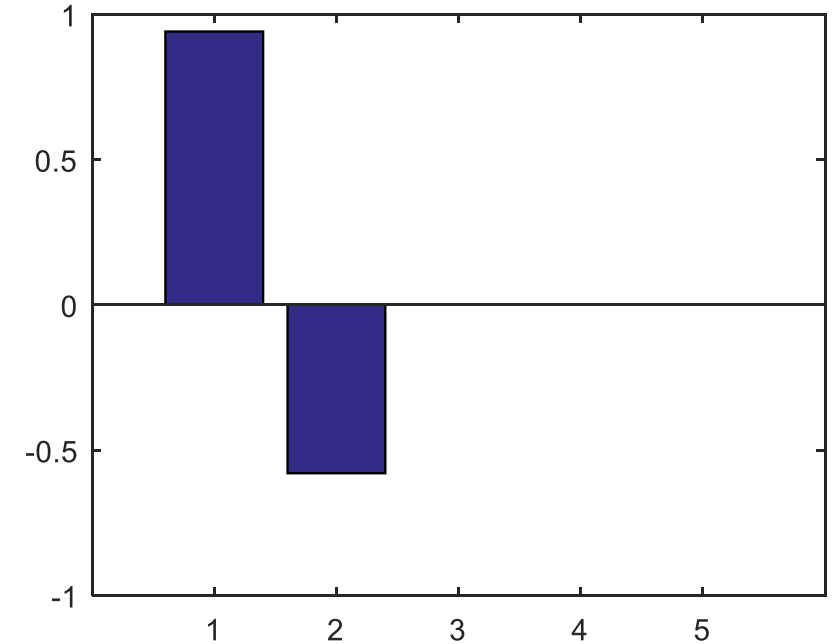
στ) Μερικές αυτοσυσχετίσεις

$$\rho_{11} = \rho_1 = 0.94$$

$$\rho_{22} = \frac{\rho_2 - \rho_1^2}{1 + \rho_1^2} = -0.58$$

$$\rho_{ss} = 0, \text{ για } s > 2$$

Διάγραμμα μερικών
αυτοσυσχετίσεων



Άσκηση 3

Από ένα δείγμα 100 παρατηρήσεων για την μεταβλητή Y προέκυψαν οι εξής παλινδρομήσεις

$$\hat{y}_t = 0.556y_{t-1}$$

$$\hat{y}_t = 0.325y_{t-1} + 0.442y_{t-2}$$

$$\hat{y}_t = 0.112y_{t-1} + 0.385y_{t-2} + 0.130y_{t-3}$$

Μπορείτε με βάση τα παραπάνω, να ισχυριστείτε ότι το δείγμα προέρχεται από μια AR(3) διαδικασία;

$$\hat{y}_t = 0.556y_{t-1}$$

$$\hat{y}_t = 0.325y_{t-1} + 0.442y_{t-2}$$

$$\hat{y}_t = 0.112y_{t-1} + 0.385y_{t-2} + 0.130y_{t-3}$$

Πρέπει να κάνουμε έλεγχο στατιστικής σημαντικότητας των μερικών αυτοσυσχετίσεων με τη στατιστική t .

Είναι

$$\hat{\rho}_{11} = 0.556, t_1 = 0.556\sqrt{100} = 5.56 > 2 \text{ στατιστικά σημαντικό}$$

$$\hat{\rho}_{22} = 0.442, t_2 = 0.442\sqrt{100} = 4.42 > 2 \text{ στατιστικά σημαντικό}$$

$$\hat{\rho}_{33} = 0.130, t_2 = 0.130\sqrt{100} = 1.30 < 2 \text{ δεν είναι στατιστικά σημαντικό}$$

Άρα δεν μπορούμε να θεωρήσουμε ότι το δείγμα προέρχεται από μια AR(3) διαδικασία.

Βιβλιογραφία

1. Ε. Μπόρα – Σέντα, Χ. Μωυσιάδης. Εφαρμοσμένη στατιστική, Β' έκδοση, Εκδόσεις Ζήτη, 1995.
2. Γ. Κ. Χρήστου. Εισαγωγή στην Οικονομετρία, Β τόμος (Γ' έκδοση), Εκδόσεις Gutenberg, 2007.
3. Δ. Κουγιουμτζής. Σημειώσεις μαθήματος Χρονοσειρών. Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, ΑΠΘ.
4. Γ.Ε. Κοκολάκης. Σημειώσεις ανάλυσης Χρονοσειρών. Σχολή Εφαρμοσμένων Μαθηματικών & Φυσικών Επιστημών, Αθήνα.