

### 3 Συσχετίσεις σε χρονοσειρές

Η χρονοσειρά ενός χρηματιστηριακού δείκτη  $\{y_0, y_1, \dots, y_N\}$  μπορεί να θεωρηθεί ως πραγματοποίηση μιας στοχαστικής διαδικασίας  $\{Y_t\}_{t=-\infty}^{\infty}$ . Αντίστοιχα η χρονοσειρά της μεταβολής των τιμών του δείκτη  $\{x_1, x_2, \dots, x_N\}$  μπορεί να θεωρηθεί ως πραγματοποίηση μιας άλλης στοχαστικής διαδικασίας  $\{X_t\}_{t=-\infty}^{\infty}$ . Σημειώνεται ότι η μεταβολή των τιμών του δείκτη μπορεί να ορισθεί από τις πρώτες διαφορές των τιμών του δείκτη  $x_t = y_t - y_{t-1}$  ή από τις πρώτες διαφορές των λογαρίθμων των τιμών  $x_t = \log y_t - \log y_{t-1}$  (που είναι ισοδύναμο με τη σχετική μεταβολή ή όπως λέγεται στα χρηματοοικονομικά απόδοση,  $x_t = \frac{y_t - y_{t-1}}{y_{t-1}}$ ). Στο προηγούμενο κεφάλαιο

ασχοληθήκαμε με τη στατική περιγραφή των χρονοσειρών (και κατ' επέκταση των αντίστοιχων στοχαστικών διαδικασιών) και η μελέτη περιορίστηκε στην περιθώρια κατανομή των  $\{y_0, y_1, \dots, y_N\}$  και  $\{x_1, x_2, \dots, x_N\}$ . Στη συνέχεια θα ασχοληθούμε με κάποια θέματα της δυναμικής περιγραφής των χρονοσειρών αυτών και ειδικότερα τις συσχετίσεις. Πρώτα θα παρουσιαστούν κάποια βασικά στοιχεία για την κατανομή, τις ροπές και τη στασιμότητα μιας στοχαστικής διαδικασίας. Στη συνέχεια θα μελετηθούν συσχετίσεις βραχείας και μακράς κλίμακας και θα δοθούν μέθοδοι για την εκτίμηση του δείκτη μακράς συσχέτισης Hurst.

#### 3.1 Κατανομές και ροπές στοχαστικής διαδικασίας

Η πλήρης περιγραφή μιας στοχαστικής διαδικασίας  $\{Y_t\}_{t=-\infty}^{\infty}$  απαιτεί ότι οι κοινές κατανομές όλων των τάξεων είναι γνωστές για κάθε χρονική στιγμή  $t$ . Ξεκινώντας από τάξη ένα, η (περιθώρια) κατανομή της  $\{Y_t\}_{t=-\infty}^{\infty}$  στη γενική της μορφή είναι

$$\forall t \in Z, \quad f_{Y_t}(y) = f_Y(y, t),$$

δηλαδή ορίζεται ως συνάρτηση όχι μόνο της κάθε τιμής  $y$  αλλά και του χρόνου  $t$ . Κατά τον ίδιο τρόπο η κοινή κατανομή δύο μεταβλητών της  $\{Y_t\}_{t=-\infty}^{\infty}$  (κατανομή τάξης 2) είναι

$$\forall t_1, t_2 \in Z, \quad f_{Y_{t_1}, Y_{t_2}}(y_1, y_2) = f_Y(y_1, y_2, t_1, t_2),$$

η κοινή κατανομή τριών μεταβλητών (κατανομή τάξης 3) είναι

$$\forall t_1, t_2, t_3 \in Z, \quad f_{Y_{t_1}, Y_{t_2}, Y_{t_3}}(y_1, y_2, y_3) = f_Y(y_1, y_2, y_3, t_1, t_2, t_3)$$

και αντίστοιχα ορίζονται οι κατανομές μεγαλύτερων τάξεων.

Αντίστοιχα με τις κατανομές ορίζονται και οι ροπές της στοχαστικής διαδικασίας, δηλαδή ως συναρτήσεις του χρόνου. Η **μέση τιμή** (ροπή πρώτης τάξης) είναι

$$\forall t \in Z, \quad \mu_t = E[Y_t] = \int_{-\infty}^{\infty} y f_Y(y, t) dy.$$

Η ροπή δεύτερης τάξης είναι

$$\forall t_1, t_2 \in Z, \quad \kappa(t_1, t_2) = E[Y_{t_1} Y_{t_2}] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y_1 y_2 f_Y(y_1, y_2, t_1, t_2) dy_1 dy_2$$

και η κεντρική ροπή δεύτερης τάξης που ονομάζεται **αυτοδιασπορά** (autocovariance) είναι

$$\gamma(t_1, t_2) = E[(Y_{t_1} - \mu_{t_1})(Y_{t_2} - \mu_{t_2})] = E[Y_{t_1}Y_{t_2}] - \mu_{t_1}\mu_{t_2} = \kappa(t_1, t_2) - \mu_{t_1}\mu_{t_2}.$$

Για  $t = t_1 = t_2$  ορίζεται η **διασπορά**  $\sigma_t^2 \equiv \sigma_{Y_t}^2 = E[(Y_t - \mu_t)^2]$ . Αντίστοιχα ορίζονται οι ροπές και οι κεντρικές ροπές μεγαλύτερης τάξης για δύο μεταβλητές και οι ροπές γενικεύονται για περισσότερες μεταβλητές. Σημειώνεται ότι από τις ροπές για κάθε τάξη για δύο ή περισσότερες μεταβλητές μπορεί να οριστεί η αντίστοιχη κοινή κατανομή.

Σε αυτήν τη γενική περιγραφή της στοχαστικής διαδικασίας οι κατανομές και οι ροπές είναι συναρτήσεις των χρονικών στιγμών, δηλαδή μπορούν να μεταβάλλονται με το χρόνο.

### 3.2 Στασιμότητα

Η στατιστική περιγραφή της στοχαστικής διαδικασίας απλουστεύεται αν θεωρήσουμε ότι οι στατιστικές της ιδιότητες παραμένουν σταθερές στο χρόνο και τότε η στοχαστική διαδικασία ορίζεται ως στάσιμη. Αυτή είναι μια υπόθεση που δύσκολα μπορεί να υποτεθεί σε χρηματο-οικονομικά προβλήματα, αλλά μπορεί να χρησιμοποιηθεί ως υπόθεση εργασίας για την εξαγωγή χρήσιμων συμπερασμάτων, ακόμα και σε χρηματο-οικονομικά προβλήματα.

Ειδικότερα ορίζονται δύο μορφές στασιμότητας. Η στοχαστική διαδικασία  $\{Y_t\}_{t=-\infty}^{\infty}$  είναι **αυστηρά στάσιμη** (strict-sense stationary) όταν οι κατανομές της για κάθε τάξη (ή ισοδύναμα όλες οι ροπές) είναι σταθερές στο χρόνο, δηλαδή όταν ισχύει

$$\forall t \in Z, \quad f_{Y_t}(y) = f_Y(y, t) = f_Y(y),$$

$$\forall t_1, t_2 \in Z, \quad f_{Y_{t_1}, Y_{t_2}}(y_1, y_2) = f_{Y_t, Y_{t-\tau}}(y_1, y_2),$$

$$\forall t_1, t_2, t_3 \in Z, \quad f_{Y_{t_1}, Y_{t_2}, Y_{t_3}}(y_1, y_2, y_3) = f_{Y_t, Y_{t-\tau_1}, Y_{t-\tau_2}}(y_1, y_2, y_3),$$

και αντίστοιχα για κατανομές μεγαλύτερης τάξης.

Για ροπές τάξης μεγαλύτερης του ένα, οι κατανομές δίνονται ως συνάρτηση όχι των χρονικών στιγμών, π.χ.  $t_1, t_2$ , αλλά της υστέρησης μεταξύ των χρονικών στιγμών, π.χ.  $\tau = t_2 - t_1$ , δηλαδή για οποιεσδήποτε δύο χρονικές στιγμές που απέχουν μεταξύ τους  $\tau$  χρονικά βήματα. Ο έλεγχος της αυστηρής στασιμότητας απαιτεί τη διερεύνηση κοινών κατανομών ή ροπών όλων των τάξεων και δεν αποτελεί μια πρακτικά χρήσιμη ιδιότητα. Για αυτό συχνά χαλαρώνουμε τη συνθήκη στασιμότητας περιορίζοντας την στις δύο πρώτες ροπές.

Η στοχαστική διαδικασία  $\{Y_t\}_{t=-\infty}^{\infty}$  είναι **ασθενής στάσιμη** (weak ή wide-sense stationary) όταν οι ροπές πρώτης και δεύτερης τάξης είναι σταθερές στο χρόνο, δηλαδή

$$\alpha) \text{ η μέση τιμή είναι σταθερή : } \forall t \in Z, \quad E[Y_t] = \mu,$$

β) η αυτοδιασπορά ορίζεται μόνο ως προς την υστέρηση και όχι τις χρονικές στιγμές:  $\forall t_1, t_2 \in Z, \quad \gamma(t_1, t_2) = \gamma(t, t - \tau) = \gamma(\tau)$ .

Το β) προκύπτει από τη συνθήκη ότι η δεύτερη ροπή είναι σταθερή, δηλαδή ισχύει  $E[Y_{t_1}Y_{t_2}] = E[Y_t, Y_{t-\tau}] = \kappa(t, t - \tau) = \kappa(\tau)$ . Από τις συνθήκες α) και β) προκύπτει ότι η διασπορά είναι επίσης σταθερή. Πράγματι για  $\tau = 0$ , ισχύει  $E[Y_t^2] = \kappa(0)$  και άρα

$$\sigma_Y^2 = \gamma(0) = E[Y_t^2] - (E[Y_t])^2 = \kappa(0) - \mu^2.$$

Στην πράξη, η συνθήκη ασθενούς στασιμότητας ερμηνεύεται συχνά ως σταθερή μέση τιμή και διασπορά (απλή ροπή δεύτερης τάξης), που δεν είναι σωστό αφού η συνθήκη αναφέρεται στην κοινή ροπή δεύτερης τάξης (αυτοδιασπορά).

### 3.3 Αυτοσυσχέτιση

Για τη μελέτη συσχετίσεων σε στάσιμες χρονοσειρές χρησιμοποιείται η αυτοσυσχέτιση, που είναι η κανονικοποίηση της αυτοδιασποράς με την διασπορά.

Σε μια (ασθενής) στάσιμη στοχαστική διαδικασία  $\{Y_t\}_{t=-\infty}^{\infty}$  ορίζεται η **αυτοσυσχέτιση** για υστέρηση  $\tau$  ως

$$\rho(\tau) = \frac{\gamma(\tau)}{\gamma(0)}.$$

Η αυτοσυσχέτιση μετράει τη συσχέτιση μεταβλητών της  $\{Y_t\}_{t=-\infty}^{\infty}$  που βρίσκονται σε χρονική υστέρηση  $\tau$  και είναι ένα χρήσιμο μέτρο της «μνήμης» της στοχαστικής διαδικασίας.

Η εκτίμηση της αυτοδιασποράς από μια πραγματοποίηση (χρονοσειρά)  $\{y_1, \dots, y_N\}$  της στοχαστικής διαδικασίας  $\{Y_t\}_{t=-\infty}^{\infty}$  είναι

$$c(\tau) \equiv \hat{\gamma}(\tau) = \frac{1}{N-\tau} \sum_{t=\tau+1}^N (y_t y_{t-\tau} - \bar{y}^2).$$

Αντίστοιχα, η εκτίμηση της αυτοσυσχέτισης είναι

$$r(\tau) \equiv \hat{\rho}(\tau) = \frac{c(\tau)}{c(0)} = \frac{c(\tau)}{s_Y^2}.$$

Η αυτοσυσχέτιση μας επιτρέπει να ορίσουμε μια *χαρακτηριστική χρονική κλίμακα* για τη μνήμη της στοχαστικής διαδικασίας. Αυτή ορίζεται από την υστέρηση αποσυσχέτισης  $\tau_c$  για την οποία η αυτοσυσχέτιση φθίνει στο μηδέν. Εναλλακτικά χρησιμοποιείται η υστέρηση για την οποία η αυτοσυσχέτιση πέφτει στο επίπεδο του  $1/e$ .

#### Παράδειγμα: Εκτίμηση αυτοσυσχέτισης για χρονοσειρά από AR(1)

Η αυτοπαλινδρομούμενη στοχαστική διαδικασία τάξης ένα, AR(1), δίνεται ως

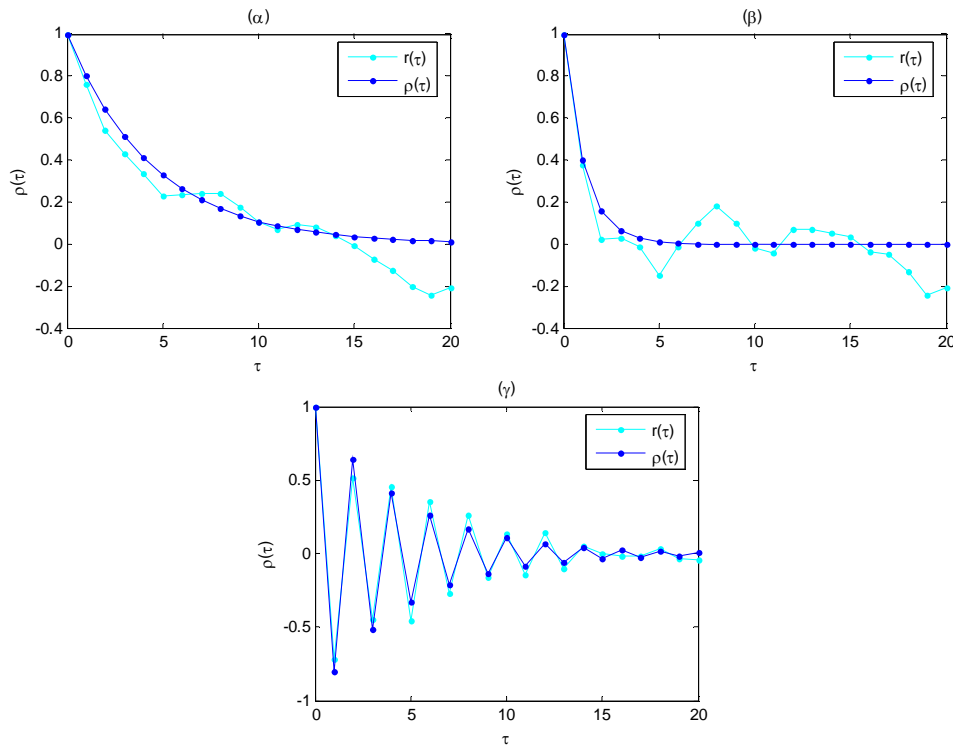
$$Y_t = \phi_0 + \phi Y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{iid}, \quad E[\varepsilon_t] = 0, \quad (12)$$

δηλαδή η μεταβλητή σε κάποιο χρονικό βήμα  $t$  ορίζεται από τη μεταβλητή στο προηγούμενο βήμα σταθμισμένη με την παράμετρο  $\phi$ , όπου  $-1 < \phi < 1$ , και μια τυχαία συνιστώσα  $\varepsilon_t$  που ακολουθεί την ίδια κατανομή με διασπορά  $\sigma_\varepsilon^2$  σε κάθε χρονικό βήμα αλλά είναι ανεξάρτητη από τις τυχαίες συνιστώσες σε προηγούμενα βήματα. Μπορεί ναδειχθεί πως η διασπορά της  $\{Y_t\}_{t=-\infty}^{\infty}$  είναι  $\sigma_Y^2 = \gamma(0) = \frac{\sigma_\varepsilon^2}{1-\phi^2}$ , η αυτοδιασπορά είναι  $\gamma(\tau) = \phi^\tau \sigma_Y^2$  και άρα η αυτοσυσχέτιση της AR(1) είναι

$$\rho(\tau) = \phi^\tau. \quad (13)$$

Στο Σχήμα 14 δίνεται η συνάρτηση αυτοσυσχέτισης για την AR(1) στοχαστική διαδικασία και η εκτίμηση της από μια πραγματοποίηση της. Η αυτοσυσχέτιση φθίνει

εκθετικά με ρυθμό που καθορίζεται από το συντελεστή  $\phi$ . Επίσης φθίνει μονότονα αν  $\phi > 0$  και εναλλασσόμενα γύρω από το 0 αν  $\phi < 0$ . Για  $\phi = 0$  η στοχαστική διαδικασία εκφυλίζεται σε λευκό θόρυβο και  $\rho(\tau) = 0$  για  $\tau > 0$ .



Σχήμα 14 Συνάρτηση αυτοσυσχέτισης για AR(1) στοχαστική διαδικασία και εκτίμηση της από χρονοσειρά 100 σημείων: (α)  $\phi = 0.8$ , (β)  $\phi = 0.4$ , (γ)  $\phi = -0.8$ .

### 3.3.1 Αυτοσυσχέτιση βραχείας και μακράς κλίμακας

Η συνάρτηση αυτοσυσχέτισης μετράει τη «μνήμη» της στοχαστικής διαδικασίας. Για παράδειγμα, ο λευκός θόρυβος έχει μηδενική μνήμη και η αυτοσυσχέτιση μηδενίζεται για όλες τις μη-μηδενικές υστερήσεις. Μαρκοβιανές στοχαστικές διαδικασίες κάποιας τάξης έχουν πεπερασμένη μνήμη και η αυτοσυσχέτιση φθίνει εκθετικά.

Η χρονική μνήμη της στοχαστικής διαδικασίας χαρακτηρίζεται από την *χαρακτηριστική κλίμακα μνήμης* της συνάρτησης αυτοσυσχέτισης. Υπάρχουν όμως και στοχαστικές διαδικασίες που δεν έχουν χαρακτηριστική κλίμακα και η αυτοσυσχέτιση φθίνει αργά ακολουθώντας κάποιο νόμο δύναμης. Σε στάσιμες στοχαστικές διαδικασίες μπορούμε να εξετάσουμε την ύπαρξη χαρακτηριστικής κλίμακας από το ολοκλήρωμα της συνάρτησης αυτοσυσχέτισης  $\int_0^\infty \rho(\tau) d\tau$ . Αν το ολοκλήρωμα είναι πεπερασμένο υπάρχει χαρακτηριστικός χρόνος μνήμης που δηλώνεται από το *χρόνος συσχέτισης* (ή *αποσυσχέτισης*)  $\tau_0$  της στοχαστικής διαδικασίας. Η μορφή της συνάρτησης αυτοσυσχέτισης είναι

$$\rho(\tau) = \exp(-\tau^\nu / \tau_0), \quad (14)$$

δηλαδή φθίνει εκθετικά με ρυθμό που δίνεται από τον εκθέτη  $\nu$  και το χρόνο συσχέτισης  $\tau_0$ . Επίσης ο χαρακτηριστικός χρόνος μνήμης μπορεί να υποδηλώνεται από την τιμή του  $\tau_c = \int_0^\infty \rho(\tau) d\tau$ , δηλαδή από την επιφάνεια κάτω από την καμπύλη της συνάρτησης αυτοσυσχέτισης. Για υστερήσεις μεγαλύτερες του  $\tau_c$  η

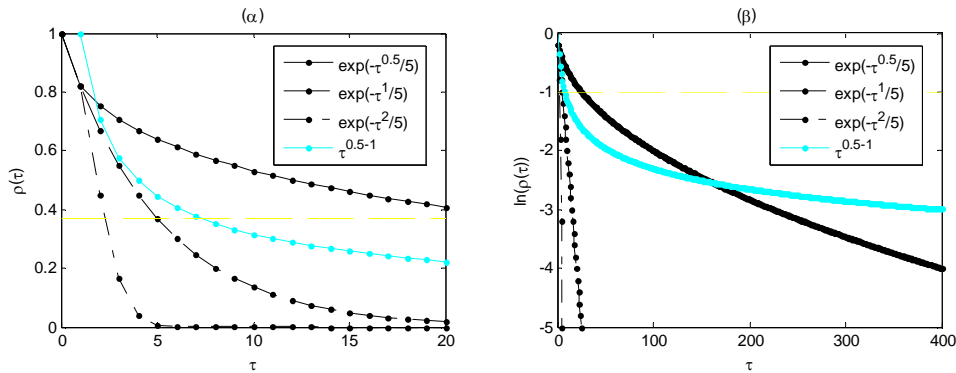
αυτοσυσχέτιση μηδενίζεται. Τέτοιες στοχαστικές διαδικασίες έχουν **συσχετίσεις βραχείας κλίμακας**.

Υπάρχουν φθίνουσες συναρτήσεις αυτοσυσχέτισης που δεν έχουν όμως πεπερασμένο ολοκλήρωμα  $\int_0^\infty \rho(\tau) d\tau$ . Για τη συνάρτηση αυτοσυσχέτισης της μορφής

$$\rho(\tau) \sim \tau^{-\eta-1} \quad (15)$$

για  $0 < \eta < 1$  και μεγάλες υστερήσεις, το ολοκλήρωμα απειρίζεται. Για  $\eta=1$  καθώς και για  $1 < \eta < 2$  η αυτοσυσχέτιση για μεγάλες υστερήσεις δίνεται από πιο πολύπλοκη μορφή που περιέχει τον ίδιο νόμο δύναμης και πάλι το ολοκλήρωμα απειρίζεται.

Δεν υπάρχει λοιπόν κάποιος χαρακτηριστικός χρόνος  $\tau_c$  που να χωρίζει σε χρόνους ύπαρξης μνήμης (υστερήσεις μη-μηδενικής αυτοσυσχέτισης) και μη ύπαρξης μνήμης (υστερήσεις μηδενικής αυτοσυσχέτισης). Μεταβλητές της στοχαστικής διαδικασίας παραμένουν συσχετισμένες όσο μακριά χρονικά και αν βρίσκονται και μια τέτοια στοχαστική διαδικασία έχει **συσχετίσεις μακράς κλίμακας**. Στο Σχήμα 15 δίνονται αντιπροσωπευτικά παραδείγματα συναρτήσεων αυτοσυσχέτισης που φθίνουν εκθετικά και με νόμο δύναμης για στοχαστικές διαδικασίες με συσχετίσεις βραχείας και μακράς κλίμακας, αντίστοιχα. Η συνάρτηση αυτοσυσχέτισης με εκθετική μορφή φθίνει πιο γρήγορα από τη συνάρτηση αυτοσυσχέτισης που ακολουθεί νόμο δύναμης ανεξάρτητα από τον εκθέτη της εκθετικής συνάρτησης. Αυτό μπορεί να μη φαίνεται παρά μόνο αν παρατηρήσουμε τις τιμές αυτοσυσχέτισης (που είναι πολύ κοντά στο 0) για μεγάλες υστερήσεις, όπως στο Σχήμα 15β.



Σχήμα 15 (α) Αυτοσυσχέτιση με εκθετική μορφή για τρεις διαφορετικούς εκθέτες και με νόμο δύναμης όπως δίνεται στο ένθετο. (β) Το ίδιο όπως στο (α) αλλά η αυτοσυσχέτιση δίνεται σε λογαριθμική κλίμακα και για μεγαλύτερο εύρος υστερήσεων.

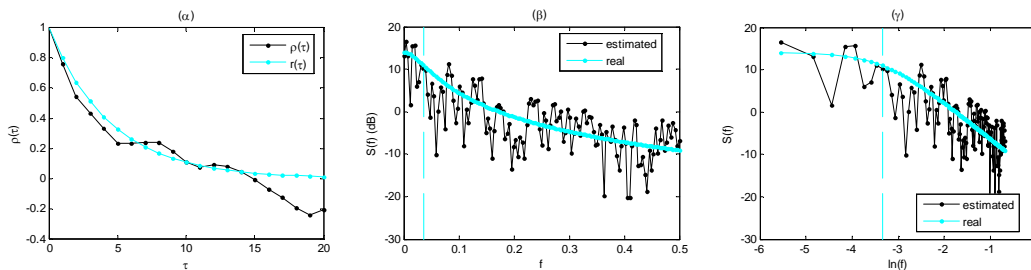
### 3.3.2 Αυτοσυσχέτιση και φάσμα ισχύος για συσχετίσεις βραχείας κλίμακας

Οι στατιστικές ιδιότητες που χαρακτηρίσαμε με την αυτοσυσχέτιση μπορούν να χαρακτηριστούν ισοδύναμα με το φάσμα ισχύος, που δίνεται ως  $S(f) = \int_{-\infty}^{\infty} \gamma(\tau) e^{-2\pi f \tau} d\tau$ , όπου  $\gamma(\tau)$  είναι η αυτοδιασπορά. Ας θεωρήσουμε στοχαστικές διαδικασίες με συσχετίσεις βραχείας κλίμακας που δίνονται από την συνάρτηση αυτοσυσχέτισης

$$\rho(\tau) = \exp(-\tau / \tau_0). \quad (16)$$

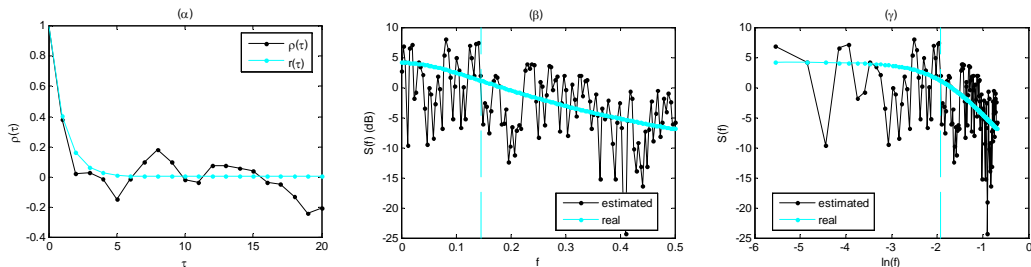
Μπορεί κάποιος να παρατηρήσει ότι  $\tau_c = \tau_0 = \int_0^\infty \exp(-\tau/\tau_0) d\tau$ . Αντικαθιστώντας την αυτοδιασπορά  $\gamma(\tau) = \sigma^2 \exp(-\tau/\tau_0)$  στον τύπο του φάσματος ισχύος παίρνουμε  $S(f) = \frac{2\sigma^2\tau_c}{1+(2\pi f\tau_c)^2}$ . Παρατηρούμε ότι για συχνότητες  $f \ll 1/(2\pi\tau_c)$  το φάσμα ισχύος έχει σταθερή τιμή ανεξάρτητη του  $f$ . Για συχνότητες  $f \gg 1/(2\pi\tau_c)$  το φάσμα ισχύος φθίνει ως  $S(f) \sim 1/f^2$ .

Την παραπάνω ιδιότητα μπορούμε να τη δούμε από την αυτοπαλινδρομούμενη στοχαστική διαδικασία τάξης ένα, AR(1), για την οποία η αυτοσυσχέτιση δίνεται στην (13). Η αυτοσυσχέτιση αυτή μπορεί να γραφεί όπως στη (16) θέτοντας  $\tau_c = 1/\ln(1/\phi)$ . Στο Σχήμα 16 δίνεται η αυτοσυσχέτιση και το φάσμα ισχύος για την AR(1) με συντελεστή  $\phi = 0.8$ . Από το γράφημα του φάσματος ισχύος ως προς το λογάριθμο της συχνότητας (Σχήμα 16γ) φαίνεται η μορφή του φάσματος ισχύος να σταθεροποιείται για συχνότητες μικρότερες του  $1/(2\pi\tau_c)$  και να φθίνει με νόμο δύναμης και εκθέτη 2 για συχνότητες μεγαλύτερες του  $1/(2\pi\tau_c)$ .



**Σχήμα 16** Αυτοσυσχέτιση και φάσμα ισχύος για AR(1) στοχαστική διαδικασία με συντελεστή  $\phi = 0.8$  και εκτίμηση τους από πραγματοποίηση 100 παρατηρήσεων: (α) αυτοσυσχέτιση, (β) φάσμα ισχύος προς συχνότητα και (γ) φάσμα ισχύος προς λογάριθμο συχνότητας.

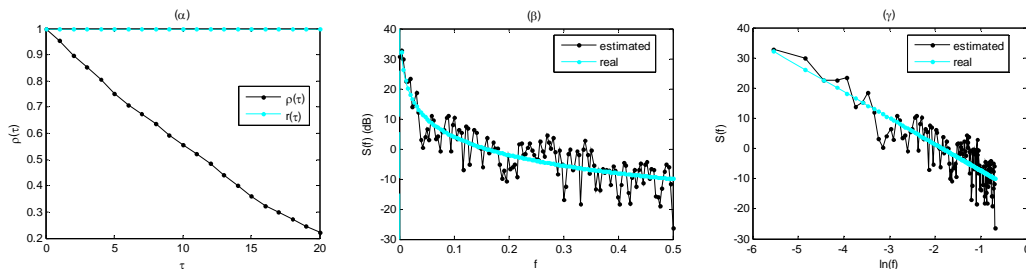
Όταν η αυτοσυσχέτιση φθίνει πιο γρήγορα, όπως για AR(1) με συντελεστή  $\phi = 0.4$ , ο χαρακτηριστικός χρόνος μνήμης είναι μικρότερος με αποτέλεσμα η αυτοσυσχέτιση να φθίνει πιο γρήγορα και το φάσμα ισχύος να είναι επίπεδο για μεγαλύτερο εύρος συχνοτήτων, όπως δίνεται στο Σχήμα 17.



**Σχήμα 17** Όπως στο Σχήμα 16 αλλά για  $\phi = 0.4$ .

Όταν ο συντελεστής του AR(1) είναι  $\phi = 0$ , ο χαρακτηριστικός χρόνος είναι 0, η στοχαστική διαδικασία γίνεται iid η αυτοσυσχέτιση μηδενίζεται για όλες τις υστερήσεις και το φάσμα ισχύος είναι επίπεδο. Από την άλλη, όταν  $\phi = 1$  η

στοχαστική διαδικασία είναι τυχαίος περίπατος. Τυπικά η αυτοσυσχέτιση  $\rho(\tau)$  δεν ορίζεται αφού η στοχαστική διαδικασία δεν είναι στάσιμη και η εκτίμηση της  $r(\tau)$  φθίνει αργά με ρυθμό που εξαρτάται από το μήκος της χρονοσειράς. Για το φάσμα ισχύος ισχύει  $S(f) \sim 1/f^2$  για ολόκληρο το εύρος συχνοτήτων (δες Σχήμα 18).

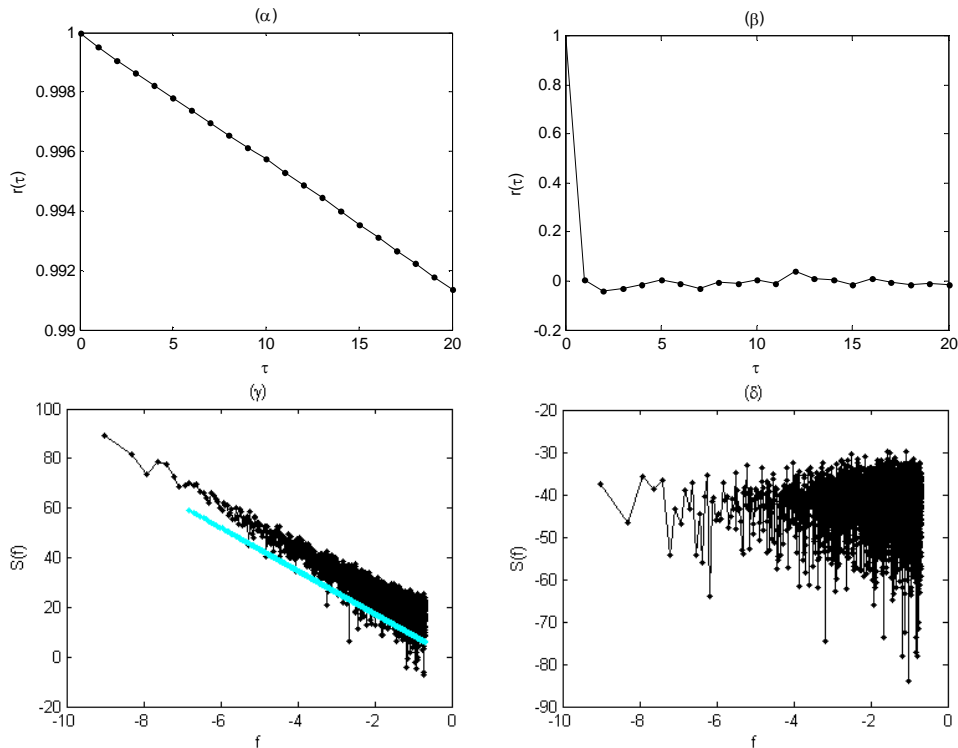


Σχήμα 18 Όπως στο Σχήμα 16 αλλά για  $\phi = 1$  (τυχαίος περίπατος).

Μια στοχαστική διαδικασία με συσχετίσεις βραχείας κλίμακας μπορεί να έχει περισσότερες από μια χαρακτηριστικές κλίμακες, όπως π.χ. η στοχαστική διαδικασία που προκύπτει από το άθροισμα δύο στοχαστικών διαδικασιών με διαφορετικές χαρακτηριστικές κλίμακες. Γενικά όμως θα περιμένουμε μια στάσιμη στοχαστική διαδικασία που έχει συσχετίσεις βραχείας κλίμακας να έχει αυτοσυσχέτιση που φθίνει γρήγορα (εκθετικά) και φάσμα ισχύος που είναι επίπεδο για μικρές συχνότητες. Αντίστροφα, όταν μελετάμε μια χρονοσειρά, μπορούμε να συμπεράνουμε ότι είναι πραγματοποίηση στοχαστικής διαδικασίας με συσχετίσεις βραχείας κλίμακας όταν η εκτιμώμενη αυτοσυσχέτιση φθίνει γρήγορα (εκθετικά) στο μηδέν. Οι χρονοσειρές μεταβολών ή αποδόσεων των χρηματιστηριακών δεικτών μπορεί να προέρχονται από τέτοιου τύπου στοχαστικές διαδικασίες, και μάλιστα με ασήμαντες συσχετίσεις. Από την άλλη, η χρονοσειρά του χρηματιστηριακού δείκτη σχηματίζεται από την ολοκλήρωση των μεταβολών και έχει αυτοσυσχέτιση που φθίνει πολύ αργά και φάσμα ισχύος που ακολουθεί νόμο δύναμης ως προς τη συχνότητα με εκθέτη κοντά στο -2, δηλαδή μπορεί να εξηγηθεί ως τυχαίος περίπατος.

Παράδειγμα: Εκτίμηση συσχέτισης βραχείας κλίμακας για τη χρονοσειρά S&P500

Στο Σχήμα 19 παρουσιάζονται οι εκτιμήσεις της αυτοσυσχέτισης και του φάσματος ισχύος για τη χρονοσειρά του δείκτη S&P 500 την περίοδο 1982-2008 και τη χρονοσειρά των αποδόσεων του δείκτη. Η αυτοσυσχέτιση και το φάσμα ισχύος έχουν τα χαρακτηριστικά λευκού θορύβου (iid) για τις αποδόσεις και τυχαίου περιπάτου για το δείκτη.



Σχήμα 19 Εκτίμηση της αυτοσυσχέτισης στο (α) και του φάσματος ισχύος στο (β) για τη χρονοσειρά των αποδόσεων του ημερήσιου δείκτη S&P 500 την περίοδο 1982-2008. Στο (γ) και (δ) είναι οι ίδιες εκτιμήσεις για τη χρονοσειρά του δείκτη.

### 3.3.3 Αυτοσυσχέτιση και φάσμα ισχύος για συσχετίσεις μακράς κλίμακας

Ο λευκός θόρυβος  $X_t$  και οι αυτοπαλινδρομούμενες στοχαστικές διαδικασίες είναι παραδείγματα στοχαστικών διαδικασιών που έχουν συσχετίσεις βραχείας κλίμακας. Ο τυχαίος περίπατος  $Y_t$  μπορεί να έχει συσχετίσεις μακράς κλίμακας αλλά αυτές εξηγούνται από την ολοκλήρωση των βημάτων λευκού θορύβου σε κάθε βήμα της στοχαστικής διαδικασίας.

Θεωρώντας τη γενική έκφραση για το φάσμα ισχύος  $S(f) \sim 1/f^\eta$ , ο λευκός θόρυβος αντιστοιχεί σε  $\eta = 0$  και ο τυχαίος περίπατος σε  $\eta = 2$ . Μπορούμε όμως να θεωρήσουμε και στοχαστικές διαδικασίες για διαφορετικό εκθέτη  $\eta$ . Για τιμές του εκθέτη  $0 < \eta < 2$  η στοχαστική διαδικασία που δίνει φάσμα ισχύος για μικρές συχνότητες

$$S(f) \sim 1/f^\eta \quad (17)$$

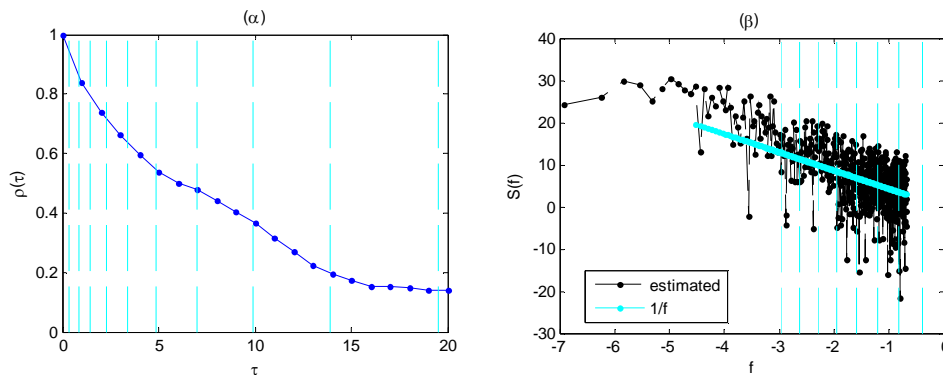
έχει συσχετίσεις μακράς κλίμακας και αναφέρεται ως **έγχρωμος θόρυβος** (colored noise) ή **κλασματική κίνηση Brown** (fractional Brownian motion).

Μια γνωστή κλάση έγχρωμου θορύβου είναι ο **1/f θόρυβος** ή **ροζ θόρυβος** (pink noise) για  $\eta = 1$ , δηλαδή για φάσμα ισχύος  $S(f) \sim 1/f$ . Αυτός ο τύπος στοχαστικής διαδικασίας έχει χρησιμοποιηθεί για να εξηγήσει συσχετίσεις μακράς διάρκειας σε διάφορα πραγματικά φαινόμενα στην υδρολογία, κλιματολογία, στα δίκτυα υπολογιστών, αλλά και στα χρηματο-οικονομικά.

Ένα σημαντικό ερώτημα σχετικά με το χαρακτηρισμό μιας χρονοσειράς ως προς την κλίμακα των συσχετίσεων είναι κατά πόσο μπορούμε να διαχωρίσουμε μια στοχαστική διαδικασία συσχετίσεων βραχείας κλίμακας με πολλούς



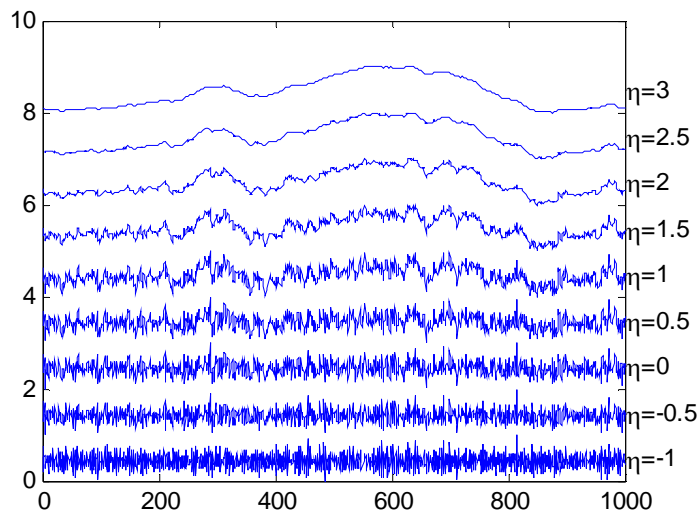
χαρακτηριστικούς χρόνους από έγχρωμο θόρυβο. Η δυσκολία αυτού του διαχωρισμού φαίνεται στο παρακάτω παράδειγμα. Θεωρούμε τη στοχαστική διαδικασία που προκύπτει από το άθροισμα 10 AR(1) στοχαστικών διαδικασιών με διαφορετικούς χρόνους κλίμακας  $\tau_c$ . Η αυτοσυσχέτιση και το φάσμα ισχύος της στοχαστικής διαδικασίας δίνονται στο Σχήμα 20, όπου φαίνονται και οι χρόνοι  $\tau_c$ . Η αυτοσυσχέτιση φθίνει αργά και το φάσμα ισχύος σχηματίζει κλίση που εξηγείται ικανοποιητικά από  $1/f$  ακόμα και για μικρές συχνότητες. Τα χαρακτηριστικά αυτά είναι γνωρίσματα έγχρωμου θορύβου.



**Σχήμα 20** Αυτοσυσχέτιση στο (α) και φάσμα ισχύος στο (β) για το άθροισμα 10 AR(1) στοχαστικών διαδικασιών με διαφορετικούς χαρακτηριστικούς χρόνους που δίνονται με τις κατακόρυφες γραμμές στο (α) και στο (β) για τις αντίστοιχες συχνότητες. Στο (β) δίνεται και το γράφημα του  $1/f$ .

Πράγματι η παραπάνω διαδικασία για μεγάλο πλήθος διαφορετικών χρόνων  $\tau_c$  αποτελεί μέθοδο δημιουργίας χρονοσειρών ροζ θορύβου. Εναλλακτικά, τέτοιες χρονοσειρές μπορούν να σχηματιστούν με τον αντίστροφο μετασχηματισμό Fourier για πλάτη που δίνονται από το  $S(f) \sim 1/f$  (ή  $S(f) \sim 1/f^\eta$  γενικά) και τυχαίες φάσεις.

Κλείνοντας την παράγραφο αυτή για τις συσχετίσεις μακράς κλίμακας παραθέτουμε στο Σχήμα 21 το διάγραμμα διασποράς χρονοσειρών από στοχαστικές διαδικασίες με φάσμα ισχύος τύπου  $S(f) \sim 1/f^\eta$  για διαφορετικές τιμές του  $\eta$ . Οι χρονοσειρές έχουν δημιουργηθεί με χρήση του αντίστροφου μετασχηματισμού Fourier. Για να συγκριθεί η επίδραση του συντελεστή  $\eta$  στη μορφή της χρονοσειράς οι φάσεις είναι ίδιες για όλες τις χρονοσειρές. Παρατηρούμε ότι για  $\eta < 0$  η ενέργεια της χρονοσειράς αυξάνει προς τις υψηλές συχνότητες και οι αυτοσυσχετίσεις φθίνουν με την υστέρηση εναλλάσσοντας σε θετικές και αρνητικές τιμές. Για  $\eta = 0$  η ενέργεια είναι ίδια σε όλες τις συχνότητες και δεν υπάρχουν συσχετίσεις (λευκός θόρυβος). Για  $\eta = 0.5, 1, 1.5$  η ενέργεια φθίνει με σταθερό ρυθμό που δημιουργεί μακρές συσχετίσεις χωρίς όμως η χρονοσειρά να μπορεί να χαρακτηριστεί μη-στάσιμη. Αυτό συμβαίνει για  $\eta = 2$ , όπου έχουμε τυχαίο περίπατο και η χρονοσειρά αυτή αντιστοιχεί σε ολοκλήρωση (άθροισμα) των βημάτων του λευκού θορύβου για  $\eta = 0$ . Αντίστοιχα για  $\eta = 2.5$  και  $\eta = 3$  οι χρονοσειρές αποτελούν ολοκληρώσεις των χρονοσειρών για  $\eta = 0.5$  και  $\eta = 1$  αντίστοιχα. Για τις τελευταίες είναι φανερό ότι το γράφημα είναι αρκετά ομαλό αφού η ενέργεια φθίνει με ταχύ ρυθμό με τη συχνότητα και ουσιαστικά η ενέργεια συγκεντρώνεται μόνο στις πολύ χαμηλές συχνότητες (αργές τάσεις).



Σχήμα 21 Χρονοσειρές από στοχαστικές διαδικασίες με φάσμα ισχύος τύπου  $S(f) \sim 1/f^\eta$  για τιμές του εκθέτη  $\eta = -1.0 : 0.5 : 3.0$ .

Ως τώρα ορίσαμε τη συσχέτιση μακράς κλίμακας με νόμο δύναμης για την αυτοσυσχέτιση,  $\rho(\tau) \sim \tau^{\eta-1}$ , για  $0 < \eta < 1$  (δες (15)) και για το φάσμα ισχύος,  $S(f) \sim 1/f^\eta$  για  $0 < \eta < 2$  (δες (17)). Παρατηρούμε ότι ο ίδιος εκθέτης  $\eta$  εμφανίζεται και στους δύο νόμους δύναμης, για την αυτοσυσχέτιση και το φάσμα ισχύος. Επίσης σημειώνεται ότι για τη χρονοσειρά  $Y_t$  που προκύπτει από την ολοκλήρωση της στάσιμης χρονοσειράς  $X_t$  ο νόμος δύναμης για το φάσμα ισχύος έχει εκθέτη  $\eta+2$ ,  $S(f) \sim 1/f^{\eta+2}$  για  $0 < \eta < 2$  (δες Σχήμα 21). Στη συνέχεια παρατίθενται μέθοδοι εκτίμησης του εκθέτη συσχέτισης μακρά κλίμακας.

### 3.4 Εκτίμηση του εκθέτη συσχέτισης μακράς κλίμακας

Για τη μέτρηση της συσχέτισης μακρά κλίμακας θα αναφερθούμε όχι στον εκθέτη  $\eta$  αλλά σε έναν άλλο εκθέτη, τον εκθέτη Hurst  $H$ , που σχετίζεται με τον εκθέτη  $\eta$  ως  $\eta=2H-1$  για τη (στάσιμη) χρονοσειρά των μεταβολών  $X_t$  και  $\eta=2H+1$  για τη (μη-στάσιμη) χρονοσειρά  $Y_t$  που προκύπτει από την ολοκλήρωση της  $X_t$ . Ο δείκτης Hurst δηλώνει το βαθμό και τύπο της αυτο-συνάφειας (self-affinity) της στοχαστικής διαδικασίας. Η έννοια ή ιδιότητα της αυτο-συνάφειας είναι συγγενική αλλά διαφορετική από την ιδιότητα της αυτο-ομοιότητας (self-similarity) που αναφέρεται σε (πολύ-) μορφοκλασματικές δομές. Η αυτό-ομοιότητα αναφέρεται στη γεωμετρία του χώρου που βρίσκονται τα σημεία που παράγει το σύστημα (ή η διαδικασία) ενώ η αυτό-συνάφεια αναφέρεται στη σχέση της κλίμακας του χρόνου και της κλίμακας των τιμών της παρατηρούμενης διαδικασίας. Συγκεκριμένα, αν αλλάξουμε την κλίμακα του χρόνου  $t$  με έναν παράγοντα  $a$ , π.χ. από λεπτό σε ώρα ή από ημέρα σε βδομάδα ή μήνα, για να διατηρήσουμε την ίδια (στατιστικά) εικόνα διακυμάνσεων του δείκτη (ή των μεταβολών του δείκτη)  $X_t$  θα πρέπει να αλλάξουμε την κλίμακα των τιμών του  $X_t$  με παράγοντα  $a^H$ . Σε αυτήν την περίπτωση η χρονοσειρά έχει την ιδιότητα της αυτο-συνάφειας και ισχύει η σχέση κλίμακας

$$X_t \rightarrow a^H X_{at} \quad (18)$$

για οποιοδήποτε παράγοντα  $a$ .

Για παράδειγμα, για μια στοχαστική διαδικασία που χαρακτηρίζεται από  $H=0.5$ , αν η κλίμακα του χρόνου τετραπλασιαστεί τότε η κλίμακα της θέσης των τιμών της θα πρέπει να διπλασιαστεί, δηλαδή η μεταβολή των τιμών της στοχαστικής διαδικασίας διπλασιάζεται όταν το παράθυρο χρόνου παρατήρησης της τετραπλασιάζεται. Αυτό είναι το γνώρισμα του τυχαίου περιπάτου. Πράγματι στην Παρ. 2.3 είχε αναφερθεί ότι η διασπορά του τυχαίου περιπάτου αυξάνει γραμμικά με τον αριθμό των χρονικών βημάτων  $n$ , δηλαδή  $E[Y_n^2] = n\sigma^2$  και  $\sigma_Y^2(n) \equiv \text{Var}[Y_n] = n\sigma^2$  (όπου θεωρούμε πως  $E[Y_n] = 0$ ) όπου  $\sigma^2$  είναι η διασπορά των βημάτων και  $\sigma_Y^2(n)$  είναι η διασπορά του τυχαίου περιπάτου σε χρονικό παράθυρο  $n$  βημάτων. Άρα για τη διασπορά του τυχαίου περιπάτου ισχύει ο νόμος κλίμακας  $\sigma_Y^2(n) \sim n$  και αντίστοιχα για την τυπική απόκλιση  $\sigma_Y(n) \sim \sqrt{n}$ , που δηλώνει ότι ο εκθέτης του  $n$  είναι  $H=0.5$ .

Η τιμή  $H=0.5$  χαρακτηρίζει τυχαίο περίπατο  $Y_t$  με τυχαία βήματα iid  $X_t$  ή με βραχείας κλίμακας συσχετίσεις. Όταν τα τυχαία βήματα έχουν μακράς κλίμακας αρνητικές συσχετίσεις (αντι-συσχετίσεις) και η ενέργεια είναι σε υψηλές συχνότητες, ο συντελεστής  $\eta$  είναι αρνητικός και άρα από τη σχέση  $\eta=2H-1$  προκύπτει ότι  $H<0.5$ . Αντίστοιχα για μακράς κλίμακας θετικές συσχετίσεις όπου  $\eta>0$  ισχύει  $H>0.5$ . Αν αντί των τυχαίων βημάτων  $X_t$  θεωρήσουμε την ολοκλήρωσή τους  $Y_t$ , αντίστοιχα με το  $\eta$  που αυξάνει κατά 2, ο εκθέτης  $H$  αυξάνει κατά 1. Το αντίστροφο φυσικά ισχύει αν από μια χρονοσειρά  $Y_t$  θεωρήσουμε τη χρονοσειρά των πρώτων διαφορών: το  $\eta$  μειώνεται κατά 2 και ο εκθέτης  $H$  κατά 1.

### 3.4.1 Μέθοδος της διασποράς

Γενικά θα περιμένουμε για μια στοχαστική διαδικασία με μακρές συσχετίσεις να ισχύει για την τυπική απόκλιση  $\sigma_Y(n)$  σε  $n$  βήματα ο νόμος κλίμακας  $\sigma_Y(n) \sim n^H$  και αντίστοιχα για τη διασπορά  $\sigma_Y^2(n) \sim n^{2H}$ . Σε αυτήν την ιδιότητα βασίζεται η πρώτη μέθοδος για την εκτίμηση του εκθέτη Hurst.

Έστω η χρονοσειρά  $\{y_1, y_2, \dots, y_N\}$  ενός (χρηματο-οικονομικού) δείκτη. Ο υπολογισμός του εκθέτη Hurst  $H$  με τη μέθοδο της (αυξανόμενης) διασποράς δίνεται από την εκτίμηση του νόμου κλιμάκωσης

$$\sigma_Y^2(n) \sim n^{2H}, \quad (19)$$

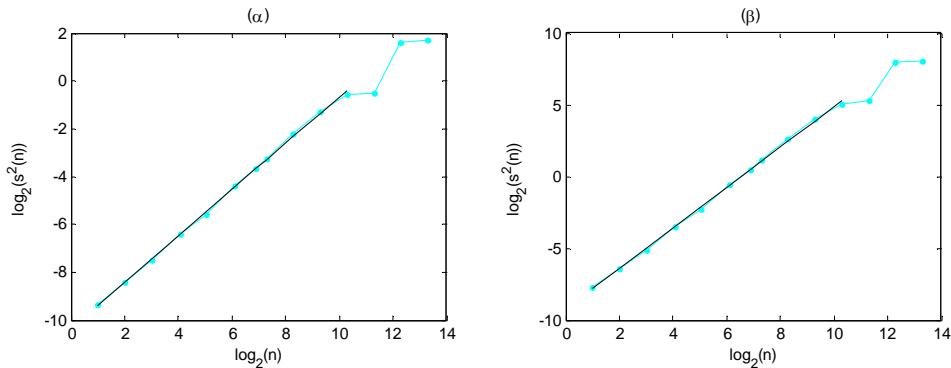
για χρονικά παράθυρα αυξανόμενων βημάτων  $n$ .

Ο υπολογισμός του  $H$  γίνεται ως εξής. Αρχίζοντας από κάποιο μικρό  $n$  (όπως για  $n=3$  και  $n=4$  στο Σχήμα 22), χωρίζεται η χρονοσειρά σε τμήματα μήκους  $n$ , υπολογίζονται οι διασπορές σε κάθε τμήμα και κατόπιν υπολογίζεται ο μέσος όρος αυτών των διασπορών  $s^2(n)$ .

$$\begin{array}{c}
 \{y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8, y_9, \dots, y_N\} \\
 \underbrace{\qquad\qquad\qquad}_{s_1^2} \quad \underbrace{\qquad\qquad\qquad}_{s_2^2} \quad \underbrace{\qquad\qquad\qquad}_{s_3^2} \quad \dots \\
 n=3 \qquad\qquad\qquad s^2(3) \\
 \\
 \{y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8, y_9, \dots, y_N\} \\
 \underbrace{\qquad\qquad\qquad}_{s_1^2} \quad \underbrace{\qquad\qquad\qquad}_{s_2^2} \quad \dots \\
 n=4 \qquad\qquad\qquad s^2(4)
 \end{array}$$

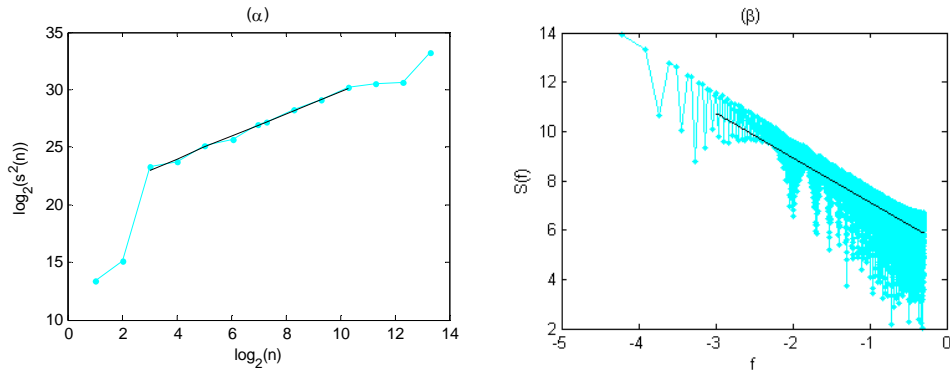
Σχήμα 22 Διάγραμμα του χωρισμού της χρονοσειράς σε  $n$  τμήματα και υπολογισμού της διασποράς  $s^2(n)$ .

Η αύξηση των βημάτων  $n$  είναι εκθετική ως το μήκος της χρονοσειράς  $N$ . Έχοντας υπολογίσει τις διασπορές  $s^2(n)$  για κάθε  $n$ , υπολογίζεται ο εκθέτης Hurst  $H$  από την κλίση της προσαρμοσμένης ευθείας στο γράφημα  $\log s^2(n)$  vs  $\log n$ . Στο Σχήμα 23 δίνεται το παραπάνω γράφημα (σε λογάριθμο με βάση το 2) καθώς και η προσαρμοσμένη ευθεία για μια χρονοσειρά 10000 παρατηρήσεων Γκαουσιανού τυχαίου περιπάτου ( $H=0.5$ ) και τυχαίου περιπάτου με βήματα που έχουν μακρές συσχετίσεις που δίνονται από  $H=0.75$ . Η εκτίμηση του Hurst από την κλίση της ευθείας είναι 0.4845 για την πρώτη περίπτωση και 0.7043 για τη δεύτερη. Σημειώνεται εδώ πως οι χρονοσειρές έχουν δημιουργηθεί από τον αντίστροφο μετασχηματισμό Fourier για φάσμα ισχύος με νόμο δύναμης ως προς τη συχνότητα και εκθέτη  $\eta=2$  και  $\eta=2.5$  αντίστοιχα (όπου ισχύει  $\eta=2H+1$ ).



Σχήμα 23 Εκτίμηση του εκθέτη Hurst από το γράφημα του λογαρίθμου της αυξανόμενης διασποράς από χρονοσειρά μήκους  $N=10000$  με εκθέτη Hurst  $H=0.5$  στο (a) και  $H=0.75$  στο (β).

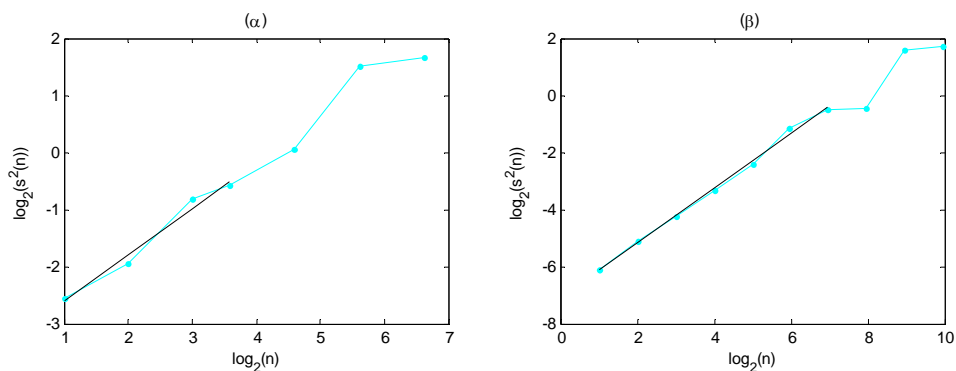
Παρατηρούμε στο Σχήμα 23 πως για μεγάλα χρονικά παράθυρα που πλησιάζουν το μήκος της χρονοσειράς δε διατηρείται ο νόμος κλίμακας. Γενικά η εκτίμηση του εκθέτη Hurst από το γράφημα του λογαρίθμου της αυξανόμενης διασποράς πρέπει να γίνει με προσοχή και έχουν προταθεί πιο πολύπλοκες συναρτήσεις προσαρμογής που αντιμετωπίζουν τις ιδιαιτερότητες της κλιμάκωσης για πολύ μικρά και πολύ μεγάλα  $n$ .



**Σχήμα 24** Εκτίμηση του εκθέτη Hurst από την αυξανόμενη διασπορά στο (α) και το φάσμα ισχύος στο (β) από χρονοσειρά τυχαίου περιπάτου μήκους  $N=10000$  με αλλά με βήματα που ακολουθούν κατανομή Cauchy.

Η εκτίμηση του εκθέτη μακράς συσχέτισης σε μια χρονοσειρά δεν αλλάζει με την περιθώρια κατανομή της χρονοσειράς. Στο Σχήμα 24 δίνεται η εκτίμηση του εκθέτη Hurst από το γράφημα της αυξανόμενης διασποράς και από το γράφημα του φάσματος ισχύος για μια χρονοσειρά τυχαίου περιπάτου με iid βήματα από κατανομή Cauchy. Η εκτίμηση από το γράφημα της αυξανόμενης διασποράς είναι  $H=0.4889$  και από το φάσμα ισχύος  $H=0.4004$  ( $\eta=1.8009$ ). Η εκτίμηση από το φάσμα ισχύος δεν είναι τόσο ακριβής όσο από την αυξανόμενη διασπορά, όπου όμως η προσαρμογή της ευθείας επιλέχθηκε προσεκτικά αγνοώντας τις δύο πιο μικρές και τρεις μεγαλύτερες τιμές του  $n$ .

Ένα άλλο πρόβλημα στην εκτίμηση του Hurst εκθέτη με τη μέθοδο της αυξανόμενης διασποράς είναι το μήκος της χρονοσειράς. Όταν το μήκος της χρονοσειράς είναι μικρό, όπως στο Σχήμα 25 για  $N=100$  και  $N=1000$ , τα παράθυρα που μπορούμε να σχηματίσουμε για να υπολογίσουμε τη διασπορά είναι λιγότερα. Επίσης το γράφημα  $\log s^2(n)$  vs  $\log n$  παρουσιάζει αποκλίσεις από την ευθεία και συνεπώς η εκτίμηση του  $H$  είναι πιο ασταθής. Τα αποτελέσματα της εκτίμησης είναι για  $N=100$   $H=0.4064$  με μέσο τετραγωνικό σφάλμα προσαρμογής της ευθείας  $MSE=0.0145$ , για  $N=1000$   $H=0.4779$  με  $MSE=0.0109$ , ενώ είχαμε βρει για  $N=10000$   $H=0.4845$  με  $MSE=0.0066$ . Για να βελτιώσουμε την ακρίβεια της εκτίμησης (μικρό  $MSE$ ) θα πρέπει το μήκος της χρονοσειράς να αυξηθεί σημαντικά.



**Σχήμα 25** Εκτίμηση του εκθέτη Hurst από το γράφημα του λογαρίθμου της αυξανόμενης διασποράς από χρονοσειρά με εκθέτη Hurst  $H=0.5$  μήκους  $N=100$  στο (α) και  $N=1000$  στο (β).

Τα προβλήματα εκτίμησης του εκθέτη Hurst σχετικά με το μήκος της χρονοσειράς καθώς και τη μη διατήρηση του νόμου κλίμακας για μικρά και μεγάλα χρονικά

παράθυρα υφίστανται και για τις άλλες μεθόδους εκτίμησης του εκθέτη Hurst που θα δούμε στη συνέχεια.

### 3.4.2 Ανάλυση αλλαγής κλίμακας του εύρους

Η κλασική εκτίμηση του εκθέτη Hurst που προτάθηκε από τον ίδιο τον Hurst είναι με την *ανάλυση κλίμακας του εύρους* (rescaled range analysis, R/S analysis). Θεωρούμε ότι έχουμε τη χρονοσειρά  $\{x_1, x_2, \dots, x_N\}$  των μεταβολών του δείκτη (και όχι τη χρονοσειρά του δείκτη, όπως στην προηγούμενη μέθοδο). Η εκτίμηση γίνεται στα παρακάτω βήματα:

1. Όπως και για τη μέθοδο της αυξανόμενης διασποράς, χωρίζουμε τη χρονοσειρά σε τμήματα μήκους  $n$  για αυξανόμενα  $n$ . Για κάποιο μήκος χρονικού παραθύρου  $n$  έχουμε  $N_n = \lfloor N/n \rfloor$  τμήματα, όπου  $\lfloor \cdot \rfloor$  δηλώνει το ακέραιο μέρος.

2. Σε κάθε τμήμα  $v = 0, \dots, N_n - 1$  υπολογίζουμε το προφίλ από την ολοκλήρωση

$$Y_v(j) = \sum_{i=1}^j (x_{vn+i} - \bar{x}_{vn})$$

όπου  $\bar{x}_{vn} = \frac{1}{n} \sum_{i=1}^n x_{vn+i}$  είναι ο (τοπικός) μέσος όρος σε κάθε τμήμα  $v$ .

3. Σε κάθε τμήμα  $v$  υπολογίζουμε το εύρος και την τυπική απόκλιση του προφίλ που δημιουργήσαμε στο βήμα 2.

$$R_v(n) = \max_{j=1}^n Y_v(j) - \min_{j=1}^n Y_v(j) \quad S_v(n) = \sqrt{\frac{1}{n} \sum_{j=1}^n Y_v^2(j)}$$

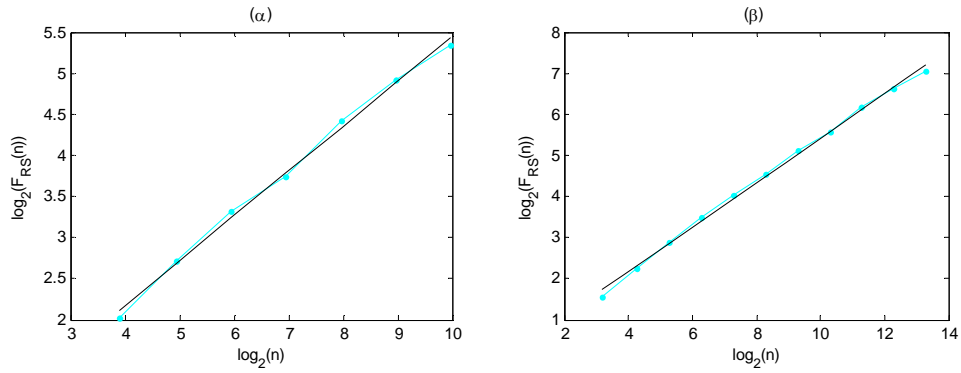
4. Υπολογίζουμε το μέσο όρο του λόγου εύρους προς τυπική απόκλιση από όλα τα τμήματα

$$F_{RS}(n) = \frac{1}{N_n} \sum_{v=0}^{N_n-1} \frac{R_v(n)}{S_v(n)}. \quad (20)$$

Η  $F_{RS}(n)$  μετράει τη μεταβολή των τιμών του προφίλ της χρονοσειράς (δηλαδή του δείκτη  $Y_t$ ) για παράθυρα αυξανόμενου μήκους  $n$ . Η  $F_{RS}(n)$  είναι αντίστοιχη της αυξανόμενης τυπικής απόκλισης (ρίζα της διασποράς) αλλά κατά κάποιο τρόπο προσαρμοσμένη για να υπολογίζει συσχετίσεις μακράς διάρκειας. Για μεγάλα παράθυρα μήκους  $n$  περιμένουμε να ισχύει ο νόμος κλίμακας

$$F_{RS}(n) \sim n^H. \quad (21)$$

Στο Σχήμα 26 παρουσιάζεται η εκτίμηση του εκθέτη Hurst από το γράφημα  $\log_2 F_{RS}(n)$  vs  $\log_2 n$  για χρονοσειρά 1000 και 10000 Γκαουσιανών βημάτων iid (ή αντίστοιχα προφίλ Γκαουσιανού τυχαίου περιπάτου). Οι εκτιμήσεις του εκθέτη Hurst για  $N=1000$  είναι  $H=0.5476$  με  $MSE=0.0047$  και για  $N=10000$  είναι  $H=0.5413$  με  $MSE=0.0074$ . Εδώ φαίνεται η προσαρμογή της ευθείας να είναι καλύτερη από την προσαρμογή της ευθείας για το γράφημα της αυξανόμενης διασποράς (δες Σχήμα 23α και Σχήμα 25β), όπου οι χρονοσειρές είναι ίδιες. Παρατηρούμε επίσης ότι η προσαρμογή είναι το ίδιο καλή για 1000 και 10000 παρατηρήσεις (μάλιστα στο παράδειγμα αυτό το MSE τυχάνει να είναι μικρότερο για  $N=1000$ ).



Σχήμα 26 Εκτίμηση του εκθέτη Hurst από το γράφημα του λογαρίθμου της συνάρτησης ανάλυση κλίμακας του εύρους από χρονοσειρά με εκθέτη Hurst  $H=0.5$  μήκους  $N=1000$  στο (α) και  $N=10000$  στο (β).

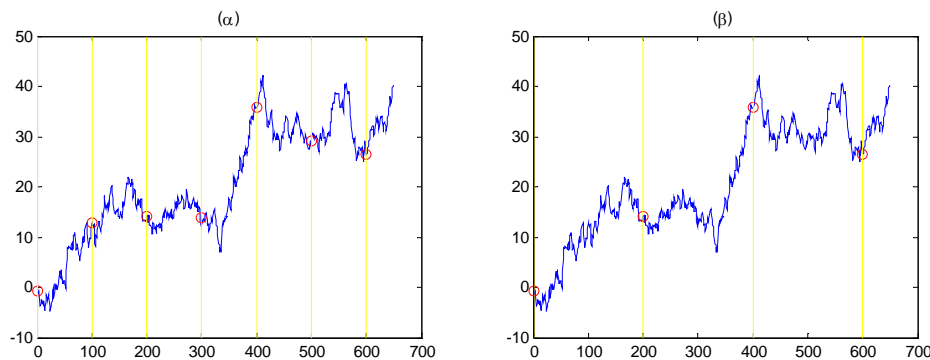
### 3.4.3 Ανάλυση διακυμάνσεων

Στη μέθοδο της αυξανόμενης διασποράς υπολογίσαμε το μέσο όρο της διασποράς (διακύμανσης) σε τμήματα της χρονοσειράς  $Y_t$  μήκους  $n$ . Στην *ανάλυση των διακυμάνσεων ή αυξομειώσεων* (fluctuation analysis, FA) εκτιμούμε τη διασπορά ή διακύμανση από το τετράγωνο της διαφοράς των ακραίων τιμών των τμημάτων. Θεωρούμε πάντα τη χρονοσειρά του προφίλ (ολοκλήρωση) των τυχαίων βημάτων, δηλαδή τη χρονοσειρά του χρηματο-οικονομικού δείκτη  $Y_t$ . Η εκτίμηση γίνεται στα παρακάτω βήματα:

1. Όπως και για τη μέθοδο της αυξανόμενης διασποράς, χωρίζουμε τη χρονοσειρά  $\{y_1, y_2, \dots, y_N\}$  σε τμήματα μήκους  $n$  για αυξανόμενα  $n$ . Για κάποιο μήκος χρονικού παραθύρου  $n$  έχουμε  $N_n = [N/n]$  τμήματα, όπου  $[.]$  δηλώνει το ακέραιο μέρος.

2. Σε κάθε τμήμα  $v = 0, \dots, N_n - 1$  υπολογίζουμε τη διακύμανση ή αυξομείωση από τα άκρα του τμήματος ως (δες Σχήμα 27)

$$F_v^2(n) = (y_{vn} - y_{(v+1)n})^2.$$



Σχήμα 27 Διάγραμμα που δείχνει τα άκρα των τμημάτων της χρονοσειράς μήκους  $n=100$  παρατηρήσεων στο (α) και  $n=200$  παρατηρήσεων στο (β).

3. Υπολογίζουμε τη τετραγωνική ρίζα του μέσου όρου των διακυμάνσεων από όλα τα τμήματα

$$F_{FA}(n) = \left( \frac{1}{N_n} \sum_{v=0}^{N_n-1} F_v^2(n) \right)^{1/2}. \quad (22)$$

Για  $n < N/10$  ο εκθέτης Hurst υπολογίζεται από το νόμο δύναμης

$$F_{FA}(n) \sim n^H. \quad (23)$$

Όταν  $N_n \neq N/n$  τα τελευταία  $N - nN_n$  δε χρησιμοποιούνται. Για αυτό τα βήματα 1 και 2 επαναλαμβάνονται με τμήματα που σχηματίζονται από το τέλος προς την αρχή της χρονοσειράς. Για κάθε τμήμα  $v = N_n, \dots, 2N_n - 1$  η αυξομείωση υπολογίζεται πάλι από τα άκρα ως  $F_v^2(n) = \left( y_{N-(v-N_n)n} - y_{N-(v+1-N_n)n} \right)^2$ . Η έκφραση (22) για τη μέση διακύμανση λαμβάνοντας υπόψη τα τμήματα από την αρχή προς το τέλος αλλά και από το τέλος προς την αρχή γίνεται

$$F_{FA}(n) = \left( \frac{1}{2N_n} \sum_{v=0}^{2N_n-1} F_v^2(n) \right)^{1/2}. \quad (24)$$

Η παραπάνω διαδικασία μπορεί να εφαρμοσθεί και εφαρμόζεται και για την R/S ανάλυση.

Γενικά αυτή η μέθοδος δεν προσφέρει καλή εκτίμηση του εκθέτη Hurst και χρησιμοποιείται εδώ ως μια πρώτη προσέγγιση που επεκτάθηκε στην παρακάτω μέθοδο.

#### 3.4.4 Ανάλυση διακυμάνσεων με απαλοιφή τάσης

Η ανάλυση διακυμάνσεων ή αυξομειώσεων με απαλοιφή τάσης (detrended fluctuation analysis, DFA) διαφέρει από την FA στον υπολογισμό της διακύμανσης (αυξομείωσης) στο βήμα 2. Αντί να υπολογιστεί αυτό από τη διαφορά στα άκρα του κάθε τμήματος, υπολογίζεται από τη διασπορά των υπολοίπων από τη προσαρμογή της τάσης στα δεδομένα του κάθε τμήματος, η οποία γίνεται με πολυώνυμο κάποιου βαθμού  $m$ . Η απαλοιφή της μονότονης τάσης σε κάθε τμήμα βοηθά στην πιο ακριβή εκτίμηση της διακύμανσης σε κάθε τμήμα και κατ' επέκταση στην πιο ακριβή εκτίμηση του εκθέτη Hurst.

Αναλυτικά η εκτίμηση του εκθέτη Hurst με τη μέθοδο DFA γίνεται στα παρακάτω βήματα:

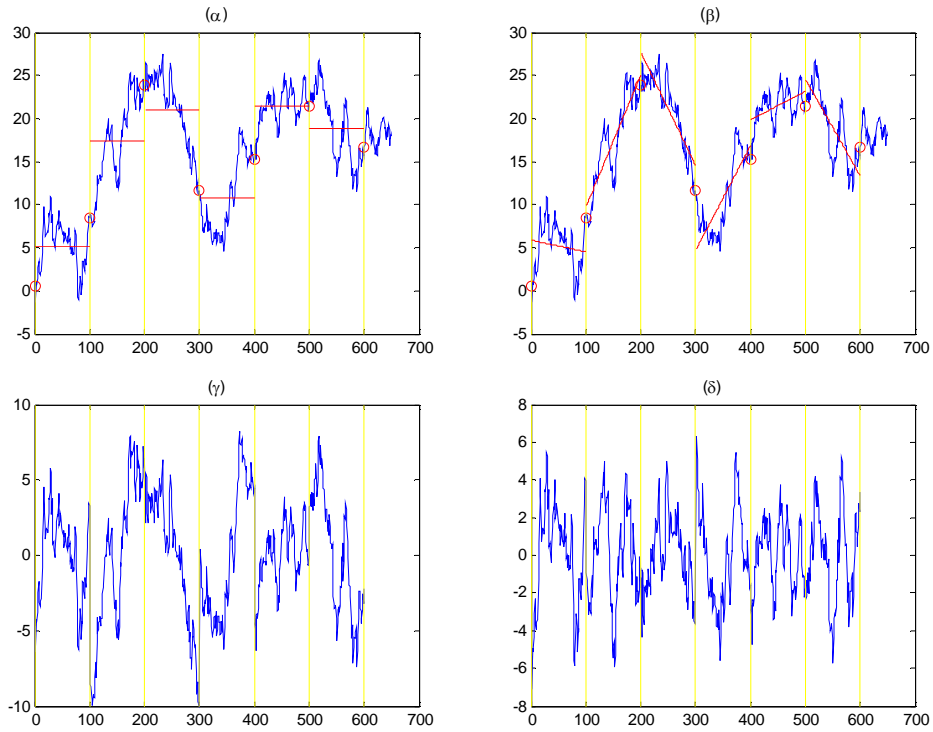
1. Όπως και για τη μέθοδο FA, χωρίζουμε τη χρονοσειρά  $\{y_1, y_2, \dots, y_N\}$  σε τμήματα μήκους  $n$  για αυξανόμενα  $n$  από την αρχή προς το τέλος της χρονοσειράς και από το τέλος προς την αρχή. Για κάποιο μήκος χρονικού παραθύρου  $n$  έχουμε  $N_n = \lceil N/n \rceil$  τμήματα, όπου  $\lceil \cdot \rceil$  δηλώνει το ακέραιο μέρος.

2. Σε κάθε τμήμα  $v = 0, \dots, N_n - 1$  εκτιμούμε την τάση βαθμού  $m$  με προσαρμογή πολυωνύμου βαθμού  $m$  στα δεδομένα,  $y_{v,n}^m(j)$ , για  $j=1, \dots, n$  (δες Σχήμα 27 για πολυώνυμο μηδενικού και πρώτου βαθμού). Το προφίλ απαλλαγμένο από την τάση είναι  $\tilde{y}_j = y_j - y_{v,n}^m(j)$ .

3. Η διακύμανση (αυξομείωση) σε κάθε τμήμα  $v$  δίνεται ως

$$F_v^2(n) = \frac{1}{n} \sum_{j=1}^n \tilde{y}_j^2.$$





**Σχήμα 28** Διάγραμμα που δείχνει την προσαρμογή πολυωνμικής τάσης βαθμού 0 και 1 σε τμήματα της χρονοσειράς μήκους  $n=100$ . (α) Προσαρμογή πολυωνύμου μηδενικού βαθμού (μέση τιμή). (β) Προσαρμογή πολυωνύμου πρώτου βαθμού. (γ) Υπόλοιπα από την προσαρμογή στο (α). (δ) Υπόλοιπα από την προσαρμογή στο (β).

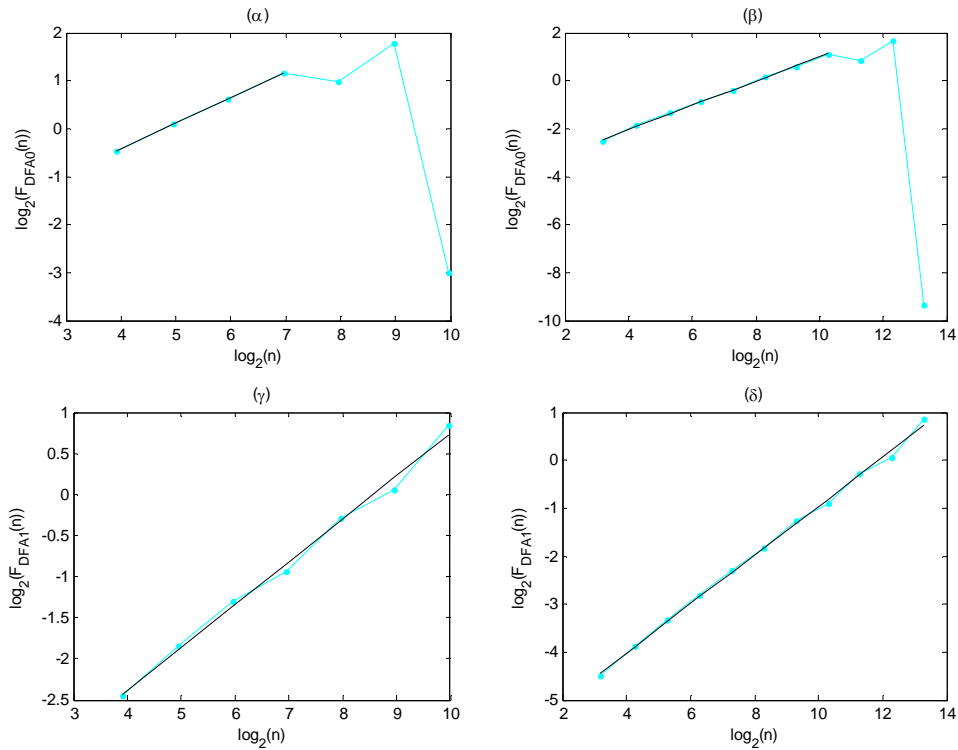
4. Υπολογίζουμε τη τετραγωνική ρίζα του μέσου όρου των διακυμάνσεων από όλα τα τμήματα

$$F_{DFAm}(n) = \left( \frac{1}{2N_n} \sum_{v=0}^{2N_n-1} F_v^2(n) \right)^{1/2}. \quad (25)$$

Η μέθοδος DFA είναι πιο ακριβής από τη μέθοδο FA και μπορεί να αντιμετωπίσει διαφορετικές τάσεις στη χρονοσειρά επιλέγοντας αντίστοιχα διαφορετικό βαθμό πολυωνύμου. Για βαθμούς  $m > 1$  τα πολυώνυμα που προσαρμόζονται σε κάθε τμήμα της χρονοσειράς θα πρέπει να είναι μονότονα γιατί η ιδέα της μεθόδου DFA είναι η απαλοιφή της μονότονης τάσης και όχι η βέλτιστη προσαρμογή σε κάθε τμήμα. Η μέθοδος DFA για κάποιο βαθμό πολυωνύμου  $m$  συνήθως συμβολίζεται ως  $DFAm$  (π.χ. DFA1 για γραμμική προσαρμογή τάσης). Η εκτίμηση του εκθέτη Hurst με DFA0 (μηδενικό βαθμό πολυωνύμου) είναι πρακτικά κοντά στην εκτίμηση με τη μέθοδο FA και ο νόμος κλίμακας δε διατηρείται για μεγάλα  $n$ .

Στο Σχήμα 29 παρουσιάζεται η εκτίμηση του εκθέτη Hurst με DFA0 και DFA1 για χρονοσειρά Γκαουσιανού τυχαίου περιπάτου 1000 και 10000 βημάτων. Οι εκτιμήσεις του εκθέτη Hurst για  $N=1000$  είναι για το DFA0  $H=0.5313$  με  $MSE=0.0002$  και το DFA1  $H=0.5211$  με  $MSE=0.0073$ , δηλαδή περίπου ίδιες εκτιμήσεις αλλά με καλύτερη προσαρμογή για το DFA0, όπου όμως δε συμπεριλήφθηκαν μεγάλα  $n$  στην προσαρμογή για τα οποία ο νόμος κλίμακας δε διατηρείται. Τα αποτελέσματα για  $N=10000$  είναι όμοια, για DFA0 είναι  $H=0.5036$  με  $MSE=0.0013$  και για DFA1 είναι  $H=0.5105$  με  $MSE=0.0054$ . Παρατηρούμε ότι για μεγαλύτερο μήκος χρονοσειράς η εκτίμηση γίνεται πιο ακριβής με DFA είτε μηδενικού ή πρώτου βαθμού. Γενικά η εκτίμηση του εκθέτη Hurst με DFA είναι πιο

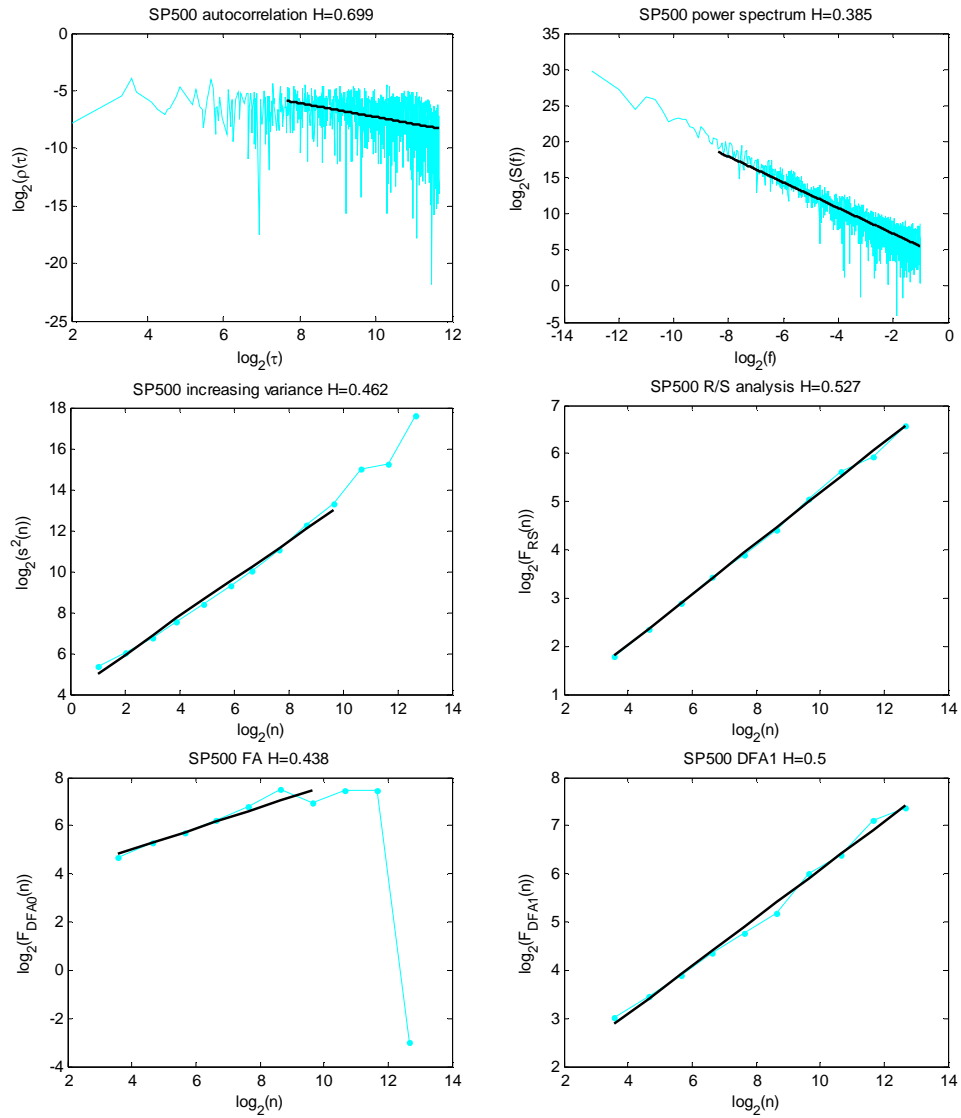
ακριβής από την εκτίμηση με την R/S ανάλυση, τη μέθοδο της αυξανόμενης διασποράς καθώς και από το φάσμα ισχύος. Αυτό φαίνεται συγκρίνοντας τις εκτιμήσεις με τις διάφορες μεθόδους για το παράδειγμα του τυχαίου περιπάτου.



**Σχήμα 29** Εκτίμηση του εκθέτη Hurst από το γράφημα του λογαρίθμου της συνάρτησης DFA0 από χρονοσειρά με εκθέτη Hurst  $H=0.5$  μήκους  $N=1000$  στο (α) και  $N=10000$  στο (β). Το ίδιο για DFA1 στο (γ) και (δ).

Παράδειγμα: Εκτίμηση του εκθέτη Hurst για τη χρονοσειρά S&P500

Κλείνουμε το κεφάλαιο αυτό για του τύπους στοχαστικών διαδικασιών και τις συσχετίσεις μακράς κλίμακας εξετάζοντας την ύπαρξη μακρών συσχετίσεων στον ημερήσιο δείκτη S&P500 (τιμή κλεισίματος) την περίοδο 20/10/1982 – 13/3/2008. Η χρονοσειρά του δείκτη S&P500 είναι  $\{y_1, y_2, \dots, y_N\}$  και  $N=6424$ . Στο Σχήμα 30 δίνονται τα γραφήματα της λογαριθμικής συνάρτησης αυτοσυσχέτισης, φάσματος ισχύος, αυξανόμενης διασποράς, R/S, DFA0 και DFA1, καθώς και η προσαρμογή της ευθείας που δίνει ως κλίση τον εκθέτη μακράς συσχέτισης. Η αυτοσυσχέτιση υπολογίστηκε για τις πρώτες διαφορές του δείκτη SP500 και η κλίση είναι  $1-\eta$  και ο εκθέτης Hurst υπολογίστηκε από τη σχέση  $\eta=2H-1$ . Το φάσμα ισχύος υπολογίστηκε στη χρονοσειρά του δείκτη, η κλίση είναι  $-\eta$  και ο εκθέτης Hurst υπολογίστηκε από τη σχέση  $\eta=2H+1$ . Παρατηρούμε ότι οι δύο αυτές εκτιμήσεις είναι πολύ διαφορετικές: η πρώτη εκτιμά ισχυρές μακρές συσχετίσεις ( $H \approx 0.7$ ) ενώ η άλλη εκτιμά αντισυσχετίσεις ( $H \approx 0.38$ ). Οι διαφορές για τις εκτιμήσεις του εκθέτη Hurst με τις άλλες μεθόδους είναι μικρότερες. Οι εκτιμήσεις της R/S ανάλυσης και DFA1 είναι πολύ κοντά στο 0.5 (το DFA1 εκτιμά ακριβώς 0.5!) και συνιστούν πως οι αυξομειώσεις στο δείκτη SP500 ακολουθούν το μοντέλο του τυχαίου περιπάτου, μη Γκαουσιανού όπως έδειξε η ανάλυση στο προηγούμενο κεφάλαιο.



**Σχήμα 30** Εκτίμηση του εκθέτη Hurst για τη χρονοσειρά του δείκτη SP500 με τις διαφορετικές μεθόδους όπως δίνεται στον τίτλο του κάθε διαγράμματος, όπου δίνεται και η τιμή της εκτίμησης του  $H$ .