Στοχαστικές διαδικασίες και συσχέτιση μακράς κλίμακας

Εκτίμηση του εκθέτη Hurst με ανάλυση διακυμάνσεων Fluctuation Analysis (FA)

$\{y_0, y_1, \ldots, y_N\}$ χρονοσειρά

Υπόθεση τυχαίου περιπάτου
τυχαία βήματα $\{x_1, x_2, \ldots, x_N\}$

$$y_i = \sum_{j=1}^{i} x_i$$

Κλιμάκωση των τετραγώνων των διακυμάνσεων του τυχαίου περιπάτου με το χρονικό παράθυρο (αριθμό βημάτων) *n*

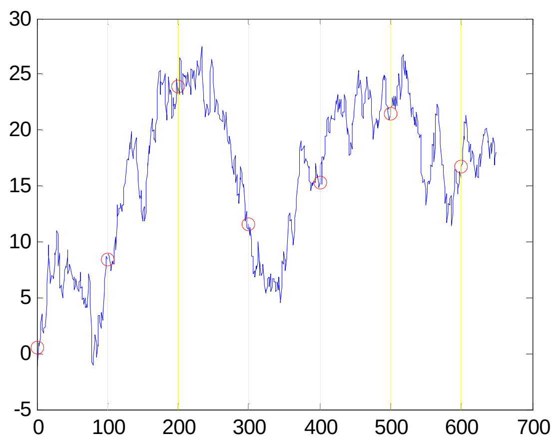
1. Χωρίζουμε τη χρονοσειρά σε τμήματα μήκους *n* (κλίμακα χρόνου *n*) :

π.χ. $n = 3$ $\{\underbrace{y_1, y_2, y_3}, \underbrace{y_4, y_5, y_6}, \underbrace{y_7, y_8, y_9}, \ldots, y_N\}$ $N_n = \lceil N/n \rceil$

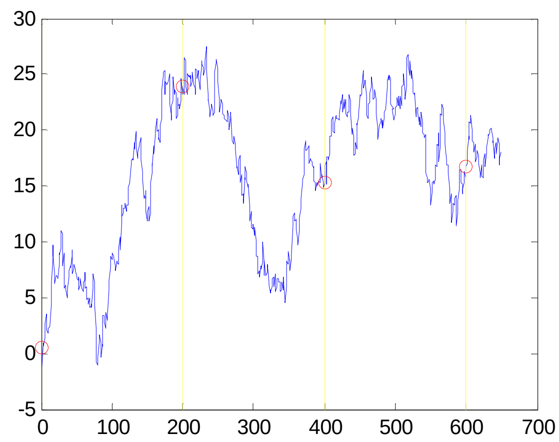
2. Σε κάθε τμήμα $\nu = 0, \ldots, N_n - 1$ υπολογίζουμε τη διακύμανση από τα άκρα του τμήματος

$$F_\nu^2(n) = \left(y_{\nu n} - y_{(\nu+1)n}\right)^2$$

n=100



n=200



3. Υπολογίζουμε τη τετραγωνική ρίζα του μέσου όρου των διακυμάνσεων

$$F_{\mathrm{FA}}(n) = \left(\frac{1}{N_n} \sum_{\nu=0}^{N_n - 1} F_\nu^2(n) \right)^{1/2}$$

Για n<N/10 ο εκθέτης Hurst υπολογίζεται από το νόμο δύναμης $F_{\mathrm{FA}}(n) \sim n^H$

Όταν $N_n = \lceil N/n \rceil \neq N/n \;\Rightarrow\; N - nN_n$ δε χρησιμοποιούνται

Επανάληψη των βημάτων 1 και 2 με τμήματα που σχηματίζονται από το τέλος προς την αρχή της χρονοσειράς.

$$\nu = N_n, \dots, 2N_n - 1 \qquad F_\nu^2(n) = \left(y_{N-(\nu - N_n)n} - y_{N-(\nu+1-N_n)n} \right)^2$$

Το βήμα 3 γίνεται

$$F_{\mathrm{FA}}(n) = \left(\frac{1}{2N_n} \sum_{\nu=0}^{2N_n - 1} F_\nu^2(n) \right)^{1/2}$$

Βελτίωση της μεθόδου FA: **ανάλυση διακυμάνσεων με απαλοιφή τάσης**
Detrended Fluctuation Analysis (DFA)

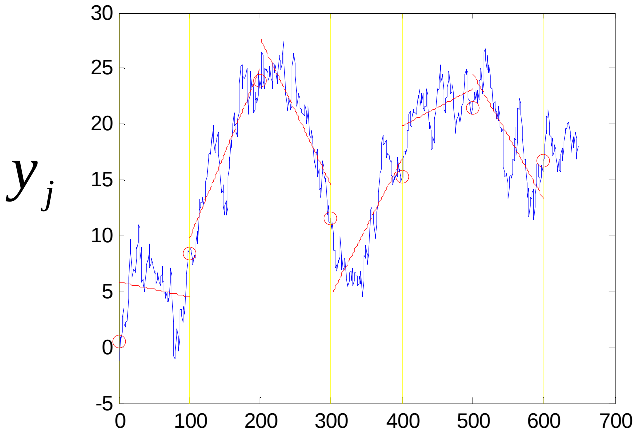
1. Χωρίζουμε τη χρονοσειρά σε τμήματα μήκους *n*:
   - από την αρχή προς το τέλος της χρονοσειράς $\{y_0, y_1, \ldots, y_N\}$
   - από το τέλος προς την αρχή

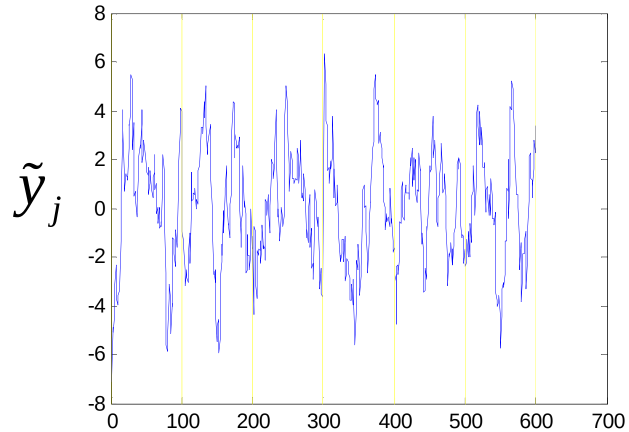
2. Σε κάθε τμήμα *ν* εκτιμούμε πολυωνυμική τάση βαθμού *m* $y^m_{\nu,n}(j)$
   (προσαρμογή πολυωνύμου στα δεδομένα)

   Το προφίλ απαλλαγμένο από τάση $\tilde{y}_j = y_j - y^m_{\nu,n}(j)$

3. Η διακύμανση στο κάθε τμήμα *ν* $F^2_\nu(n) = \frac{1}{n}\sum_{j=1}^{n} \tilde{y}_j^2$

4. Η μέση διακύμανση $F_{\mathrm{DFA}m}(n) = \left(\frac{1}{2N_n} \sum_{\nu=0}^{2N_n-1} F^2_\nu(n) \right)^{1/2}$





Στάσιμες χρονοσειρές και μοντέλα ARMA

Η χρονοσειρά των διακυμάνσεων του δείκτη $\{x_i\}$: → στάσιμη **?** → έχει συσχετίσεις μικρής διάρκειας **?**

αυτοπαλινδρομούμενα κινούμενου μέσου (AutoRegressive Moving Average)

ARMA μοντέλα τάξης *p* και *q* → μικρής διάρκειας

ARMA(*p*,*q*) $$x_{t+1} = \phi_0 + \phi_1 x_t + \cdots + \phi_p x_{t+1-p} + z_t - \theta_1 z_t - \cdots - \theta_q z_{t+1-q}$$

εξωγενείς τυχαίες επιδράσεις → $z_t \sim \text{iid},\ \mathrm{E}[z_t] = 0\ \ \mathrm{Var}[z_t] = \sigma_z^2$

AR(*p*) $$x_t = \phi_0 + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + z_t$$

Η χρονοσειρά των διακυμάνσεων του δείκτη $\{y_t\}$: → μη-στάσιμη

$\{y_t\}$ ~ ARIMA(*p*,1,*q*) όταν $\{x_t\}$ ~ ARMA(*p*,*q*) $\quad x_t = y_t - y_{t-1} \quad\quad x_t = \dfrac{y_t - y_{t-1}}{y_t}$

ολοκληρωμένο αυτοπαλινδρομούμενο μοντέλο κινούμενου μέσου ή **ολοκληρωμένο μικτό μοντέλο** (autoregressive integrated moving average model, ARIMA)

d επαναλήψεις των πρώτων διαφορών → ARIMA(*p*,*d*,*q*)

AR(1) $$x_t = \phi_0 + \phi_1 x_{t-1} + z_t$$ $$x'_t = x_t - \mu$$

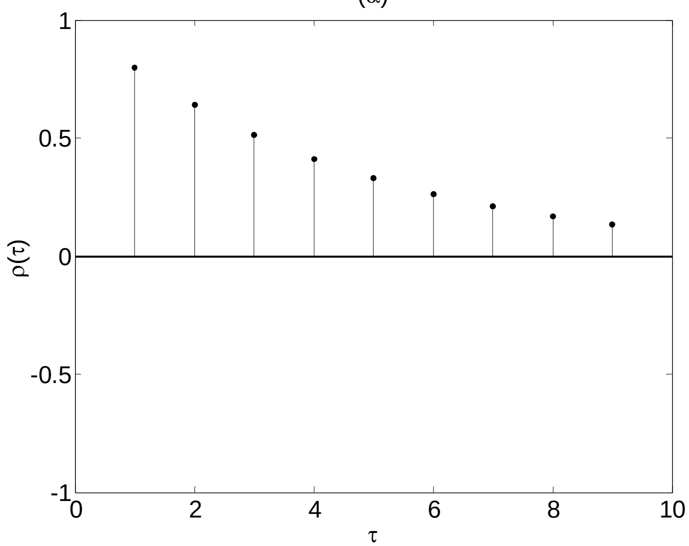
$$\left.\begin{matrix} x_t = \phi_0 + \phi_1 x_{t-1} + z_t \\ x'_t = x_t - \mu \end{matrix}\right\} \quad x'_t = \phi x'_{t-1} + z_t$$

$$\rho_\tau = \phi^\tau$$

$$\gamma_0 = \sigma_x^2 = \frac{\sigma_z^2}{1-\phi^2}$$

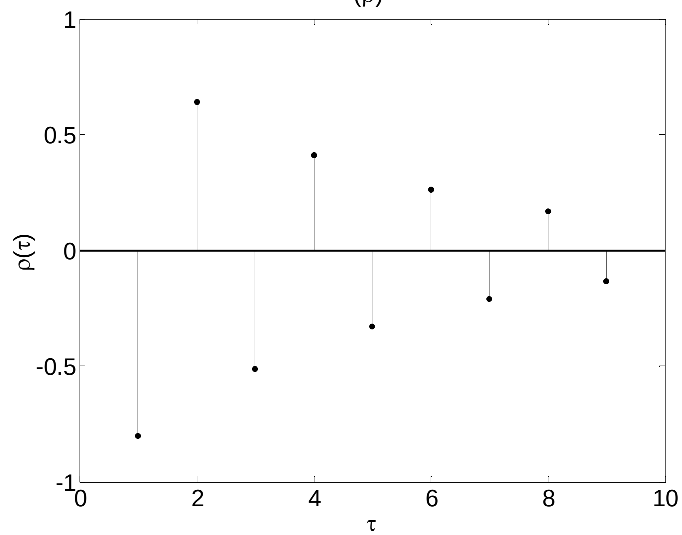
$\phi = 0.8$

(α)



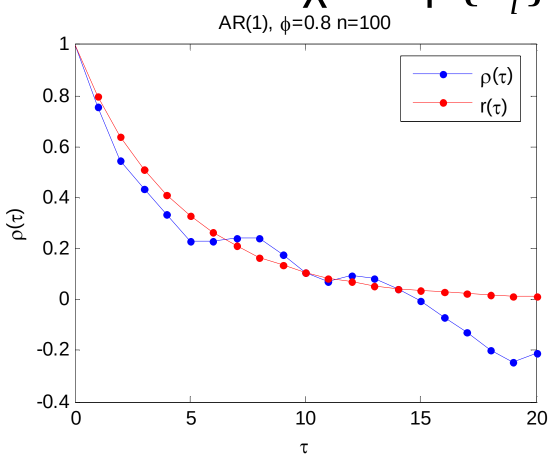
$\phi = -0.8$

(β)

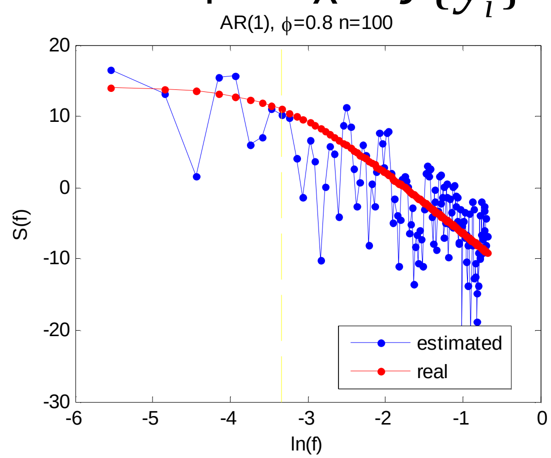


Η αυτοσυσχέτιση φθίνει εκθετικά προς το 0

αυτοσυσχέτιση $\{x_i\}$



Φάσμα ισχύος $\{y_i\}$



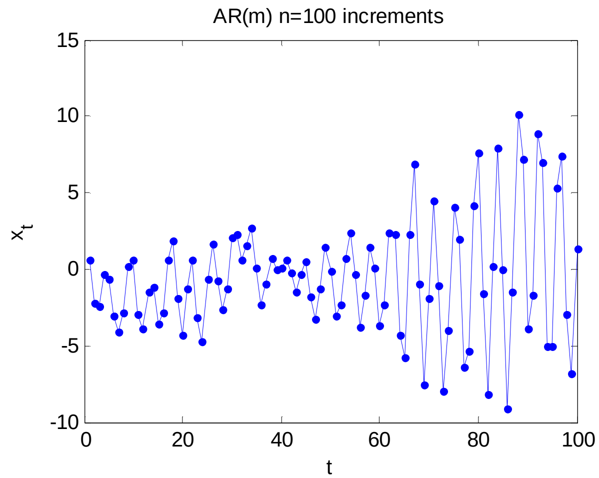
AR(m) μπορεί να δώσει ποικίλες μορφές της συνάρτησης αυτοσυσχέτισης
… αλλά πάντα φθίνει εκθετικά προς το 0

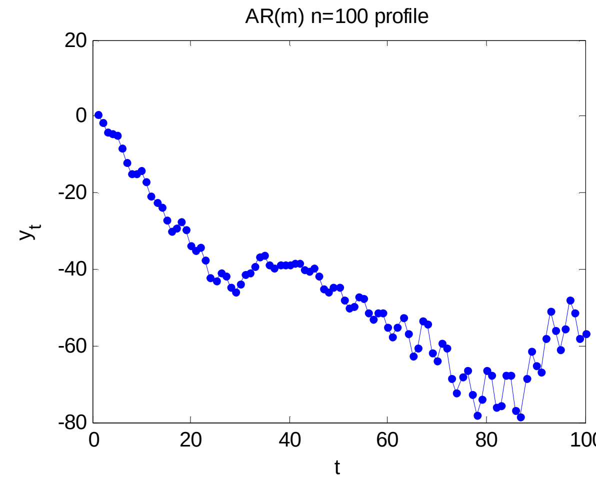
$$x_t = 1.4x_{t-1} - 1.53x_{t-2} + 1.228x_{t-3} - 0.3104x_{t-4} + z_t \qquad \sigma_z^2 = 1$$

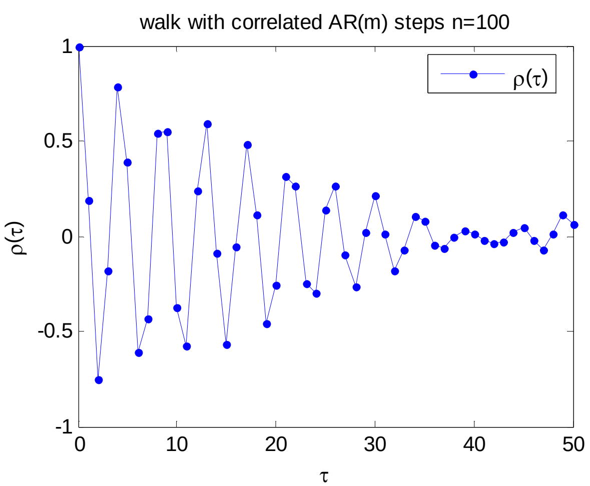
ARIMA(m,1,0)

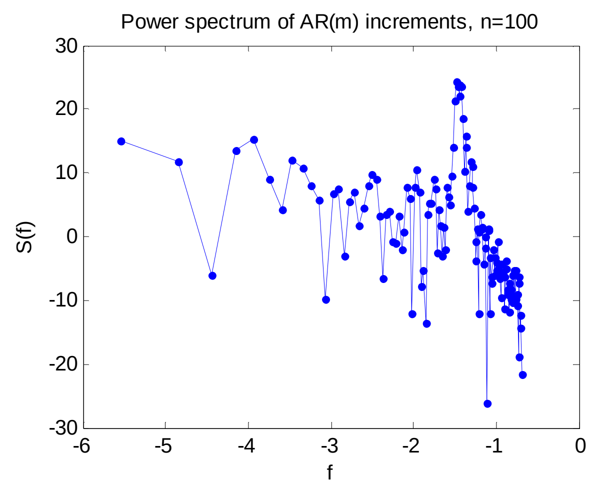
ρίζες χαρακτηριστικού πολυωνύμου

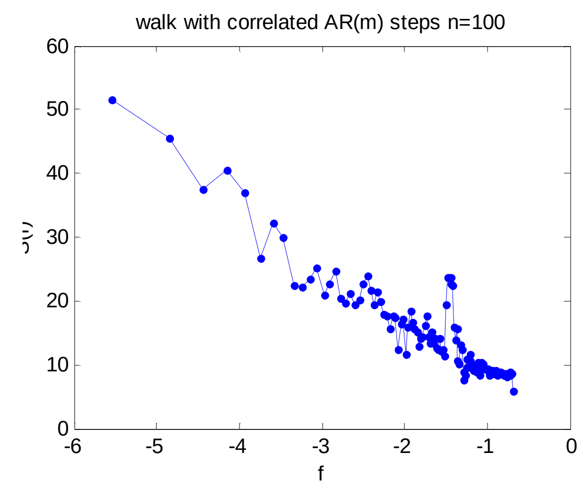
0.1000 + 0.9798i
0.1000 - 0.9798i
0.4000
0.8000



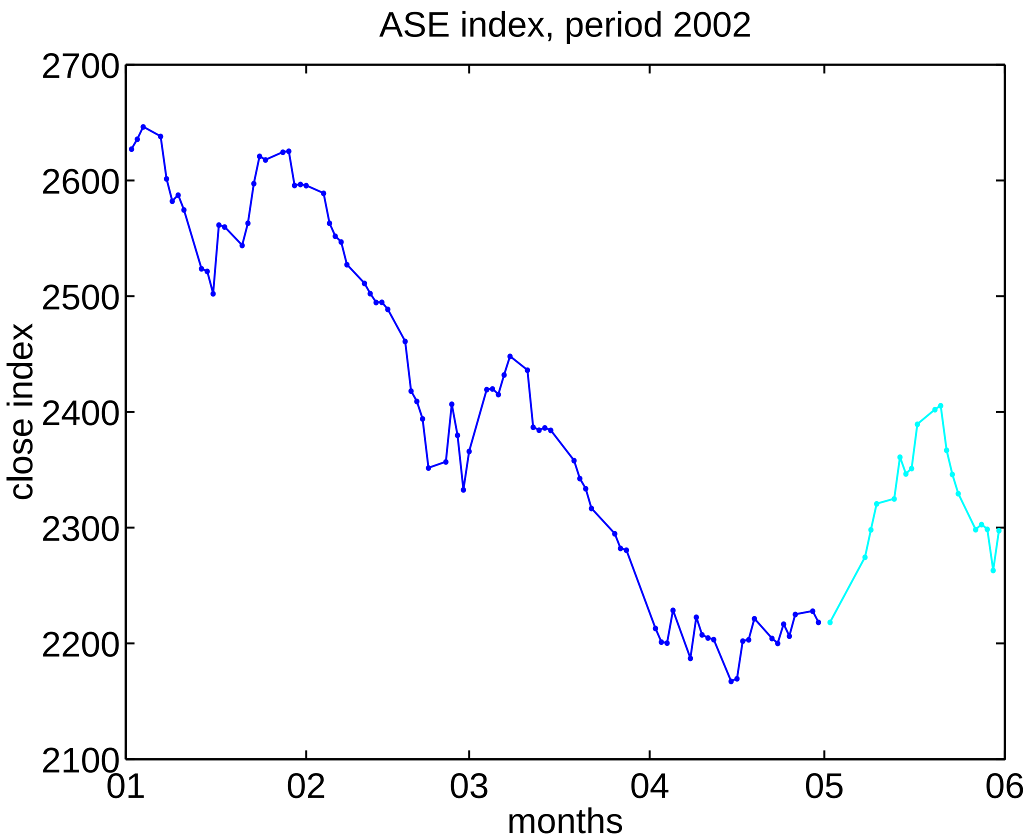








Πρόβλεψη τιμών του δείκτη ?



$\{y_0, y_1, \ldots, y_N\}$ χρονοσειρά

τυχαία βήματα $\{x_1, x_2, \ldots, x_N\}$

$$y_i = \sum_{j=1}^{i} x_i$$

Πρόβλεψη για T βήματα μπροστά

$$\hat{y}_{n+T} = y_n(T)$$

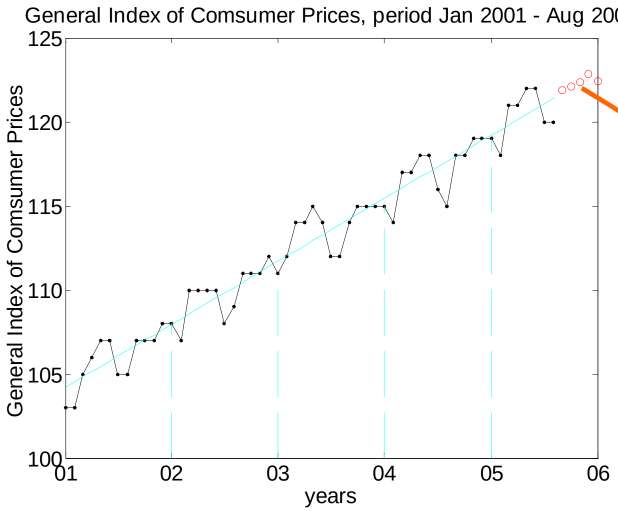
Ανάλυση χρονοσειράς σε συνιστώσες

Συνιστώσα αιτιοκρατικής τάσης $y_t = \mu_t + z_t$

Συνιστώσα περιοδικότητας / εποχικότητας $y_t = s_t + z_t$

Συνιστώσες αιτιοκρατικής τάσης και περιοδικότητας $y_t = \mu_t + s_t + z_t$

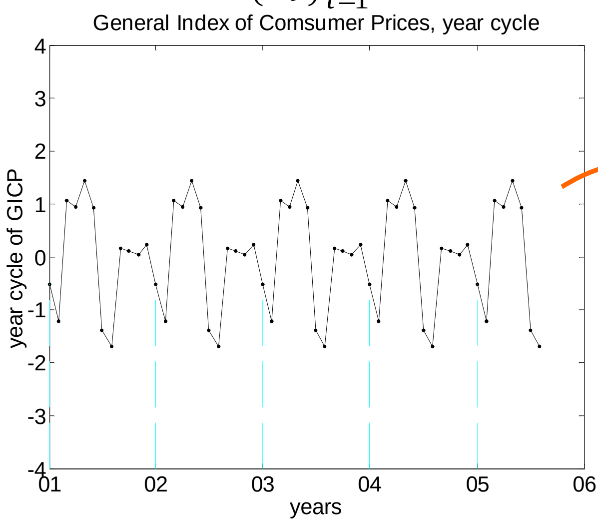
$\{y_t\}_{t=1}^{56}$ GICP, Ιανουάριος 2001 – Αύγουστος 2005 $y'_t = y_t - \mu_t$



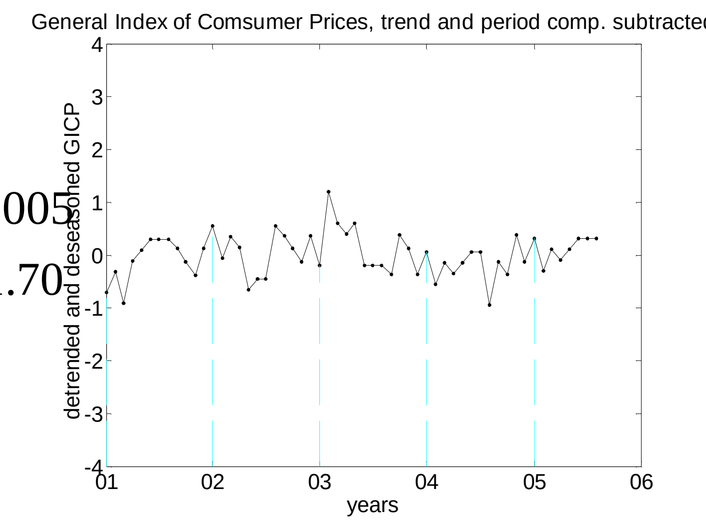
$$\mu_t = 103.9 + 0.31t$$

$$y_t = \mu_t + s_t + z_t$$

$\{s_t\}_{t=1}^{n}$



$$z_t = y'_t - s_t = y_t - \mu_t - s_t$$



$$y_n(k) = \mu_{n+k} + s_{n+k}$$

Πρόβλεψη Σεπτεμβρίου 2005

$$\mu_{57} = 103.9 + 0.31 * 57 = 121.70$$

$$s_9 = 0.16$$

$$x_{56}(1) = 121.86$$