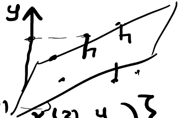
d - ελεύθερες μεταβλητές

1 - εξαρτημένη μεταβλητή.




$$\{(x_1^{(1)}, x_1^{(2)}, y_1), (x_2^{(1)}, x_2^{(2)}, y_2), \ldots, (x_n^{(1)}, x_n^{(2)}, y_n)\}$$

$$\hat{p} = (X^T X)^{-1} X^T y \in \mathbb{R}^3$$

$$X = \begin{bmatrix} 1 & x_1^{(1)} & x_1^{(2)} \\ 1 & x_2^{(1)} & x_2^{(2)} \\ \vdots & \vdots & \\ 1 & x_n^{(1)} & x_n^{(2)} \end{bmatrix} \in \mathbb{R}^{n \times 3}$$

# MEM-205 Περιγραφική Στατιστική

## Τμήμα Μαθηματικών και Εφ. Μαθηματικών, Πανεπιστήμιο Κρήτης


Κώστας Σμαραγδάκης (kesmarag@pm.me)


$$\hat{p} = \begin{bmatrix} \alpha \\ b^{(1)} \\ b^{(2)} \end{bmatrix}$$

$$X^T \in \mathbb{R}^{3 \times n} \Rightarrow X^T X \in \mathbb{R}^{3 \times 3}$$

$$\hat{y} = \alpha + b^{(1)} X^{(1)} + b^{(2)} X^{(2)}$$

27-03-2023

## Παράδειγμα

- $\mathrm{Y}$ - Ο τελικός βαθμός σε ένα συγκεκριμένο μάθημα του 4ου έτους σπουδών
- $\mathrm{X}^{(1)}$ - Ο βαθμός στη πρόοδο του μαθήματος
- $\mathrm{X}^{(2)}$ - Ο μέσος όρος βαθμολογίας του φοιτητή/τριας
- Το τμήμα του φοιτητή/τριας (πχ. tem, math, csd) 3
- Παρακολούθηση τουλάχιστον των μισών μαθημάτων μετά τη πρόοδο 2

| | tem | math | csd |
|---|---|---|---|
| | 1 | 0 | 0 |
| $d^{(1)}$ | 0 | 1 | 0 |
| $d^{(2)}$ | 0 | 0 | 1 |

| | NAI | OXI |
|---|---|---|
| $d^{(3)}$ | 1 | 0 |
| | 0 | 1 |

$$(8, 6.7, csd, NAI)$$

$$\downarrow$$

$$(8, 6.7, \underbrace{0, 1}_{csd}, \underset{NAI}{1})$$

$$\hat{\beta} = (\alpha, b^{(1)}, b^{(2)}, b^{(3)}, b^{(4)}, b^{(5)})$$

$$\hat{y} = \alpha + 8b^{(1)} + 6.7b^{(2)} + b^{(4)} + b^{(5)}$$

## Εύρεση του $\hat{b}$

$$\{(7, 5.5, \underbrace{1, 0}_{\text{math}}, 0, 7.5), (5, 7, \underbrace{0, 0}_{\text{term}}, 1, 6), (8, 8, \underbrace{0, 1}_{\text{csd}}, 0, 7.5)\}$$

$$X = \begin{bmatrix} 1 & 7 & 5.5 & 1 & 0 & 0 \\ 1 & 5 & 7 & 0 & 0 & 1 \\ 1 & 8 & 8 & 0 & 1 & 0 \end{bmatrix} \qquad y = \begin{bmatrix} 7.5 \\ 6 \\ 7.5 \end{bmatrix} \qquad b = \begin{bmatrix} \alpha \\ b^{(1)} \\ \vdots \\ b^{(d)} \end{bmatrix}$$

$$(X^T X)^{-1} X^T y = \hat{b}$$

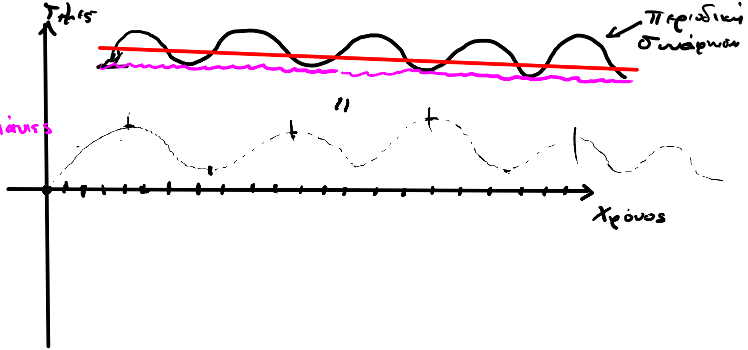
$$y = 1 \cdot \alpha + b^{(1)} x^{(1)} + b^{(2)} x^{(2)} \dots + b^{(d)} x^{(d)}$$

$$= [1, x^{(1)}, \dots, x^{(d)}] \begin{bmatrix} \alpha \\ b^{(1)} \\ \vdots \\ b^{(d)} \end{bmatrix}$$

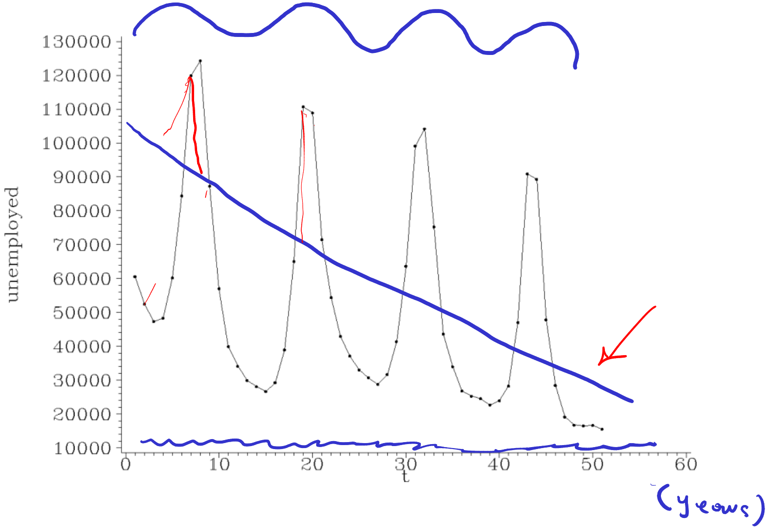
$\{x_1, x_2, \dots, x_n\} \longrightarrow x_{n+1} = ;$

Μια χρονολογική σειρά είναι ένα σύνολο παρατηρήσεων που παρουσιάζονται σε χρονολογική διάταξη.

- Μακροχρόνια τάση
- Εποχικές κυμάνσεις
- Κυκλικές κυμάνσεις
- Τυχαίες κυμάνσεις



# Χρονολογικές Σειρές (Time Series)



## Το προσθετικό μοντέλο για χρονολογικές σειρές

$$Y_t = T_t + \cancelto{0}{C_t} + S_t + R_t, \quad t = 1, \dots, N$$

- $T_t$ : Η Μακροχρόνια τάση για την $t$-χρονική περίοδο.
- $S_t$ : Ο δείκτης εποχικότητας για την $t$-χρονική περίοδο.
- $C_t$ : Η κυκλική κύμανση για την $t$-χρονική περίοδο.
- $R_t$ : Η τυχαία κύμανση για την $t$-χρονική περίοδο.

μοντέλο για μακροχ. τάση

## Απλουστευμένο μοντέλο

$(C_t = S_t = 0)$

$$Y_t = T_t + R_t, \quad t = 1, \dots, N$$

$$\mathbb{E}\{Y_t\} = \mathbb{E}\{T_t\} + \cancel{\mathbb{E}\{R_t\}}^{\,0}$$

$$\mathbb{E}\{R_t\} = 0, \quad \mathbb{E}\{Y_t\} = T_t \equiv f(t)$$

- $f(t) = f(t; \beta_1, \beta_2, \dots, \beta_p)$
- Έυρεση εκτιμήσεων $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p$ των παραμέτρων της $f$.

$$y_t = \overbrace{f(t; \underbrace{\beta_1, \beta_2, \dots, \beta_p}_{\text{παράμετροι}})}^{\text{μακροχρ. τάση}} + r(t)$$

Τυχαίο σφάλμα.

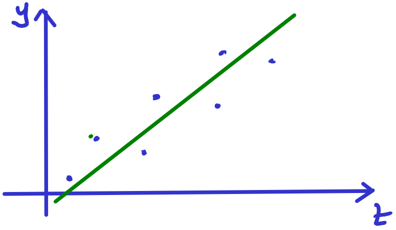
$$\hat{y}_t = f(t; \underbrace{\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_p}_{\text{εκτίμηση των παραμέτρων}})$$

## Linear function

$\alpha + bx$

$$f(t) = f(t; \beta_1, \beta_2) = \beta_1 + \beta_2\, t, \quad \beta_1, \beta_2 \in \mathbb{R}$$

Συνολο δεδομενων : $\{Y_1, Y_2, \ldots, Y_n\} \longrightarrow \beta_1, \beta_2$




$\{(1, y_1), (2, y_2), \ldots, (n, y_n)\}$

$\hat{\beta}_2 = \dfrac{SS_{ty}}{SS_{tt}}$ $\qquad$ $\hat{\beta}_1 = \bar{Y} - \hat{\beta}_2 \bar{t}$

$b$ $\qquad$ $\alpha$

$t = \{1, 2, \ldots, n\}$

$Y = \{y_1, \ldots, y_n\}$

Παράδειγμα :

$Y_t = \{2, \; 2, \; 3\} \longrightarrow \{(1,2),(2,2),(3,3)\} \longrightarrow (4, 3.33)$

| t | y | $t^2$ | ty |
|---|---|---|---|
| 1 | 2 | 1 | 2 |
| 2 | 2 | 4 | 4 |
| 3 | 3 | 9 | 9 |
| 6 | 7 | 14 | 15 |

$$\hat{\beta}_2 = \frac{SS_{ty}}{SS_{tt}} = \frac{\sum t_i y_i - \frac{1}{3}\sum t_i \sum y_i}{\sum t_i^2 - \frac{1}{3}(\sum t_i)^2} =$$

$$= \frac{15 - \frac{1}{3}6\cdot 7}{14 - \frac{1}{3}6^2} = \frac{1}{2}$$

$$\hat{\beta}_1 = \frac{7}{3} - \frac{1}{2}\cdot 2 = \frac{7}{3} - 1 = \frac{4}{3}$$

$$T_t = f(t; 4/3, 1/2) = 4/3 + t/2 \quad \Rightarrow \quad \hat{T}_4 = 4/3 + 4/2 = 2 + 4/3 = 3.33$$

**Logistic function**

$\beta_1, \beta_2, \beta_3$

$\frac{1}{f(t)} = g\left(\frac{1}{f(t-1)}\right)$

$$\mathrm{f}(\mathrm{t}) = \mathrm{f}(\mathrm{t}; \beta_1, \beta_2, \beta_3) = \frac{\beta_3}{1 + \beta_2 \exp(-\beta_1 \mathrm{t})}, \quad \beta_1, \beta_2 > 0,\ \beta_3 \in \mathbb{R} - \{0\}$$

$$\frac{1}{f(t)} = \frac{1 + \beta_2 \exp(-\beta_1 t)}{\beta_3} = \frac{1 + \beta_2 \exp(-\beta_1) \exp(-\beta_1 (t-1))}{\beta_3} =$$

$$= \frac{\exp(-\beta_1)\left[\exp(\beta_1) + \beta_2 \exp(-\beta_1 (t-1))\right]}{\beta_3} =$$

$$= \frac{\exp(-\beta_1)\left[\exp(\beta_1) - 1 + 1 + \beta_2 \exp(-\beta_1 (t-1))\right]}{\beta_3} \Rightarrow$$

$$\overbrace{\frac{1}{f(t)}}^{y(t)} = \overbrace{\frac{\exp(-\beta_1)\left[\exp(\beta_1)-1\right]}{\beta_3}}^{\alpha} + \overbrace{\exp(-\beta_1)}^{b}\ \overbrace{\frac{1}{f(t-1)}}^{y(t-1)}$$

$$y_t \approx f(t;\beta_1,\beta_2,\beta_3) = T_t$$

$$\{y_1, y_2, \dots, y_n\} \to \{1/y_1,\ 1/y_2,\ \dots,\ 1/y_n\}$$

$$\to \Big\{ (\overset{x_1}{1/y_1},\ \overset{y_1}{1/y_2}),\ (\overset{x_2}{1/y_2},\ \overset{y_2}{1/y_3}),\ \dots,\ (\overset{x_{n-1}}{1/y_{n-1}},\ \overset{y_{n-1}}{1/y_n}) \Big\}$$

$n-1$ στοιχεία.

Από απλή γραμμική παλινδρόμηση. $\Rightarrow \hat{b} = \dfrac{SS_{xy}}{SS_{xx}} = \exp(-\hat{\beta}_1)$

$$\hat{\alpha} = \frac{\exp(-\hat{\beta}_1)\left[\exp(\hat{\beta}_1)-1\right]}{\hat{\beta}_3} = \bar{Y} - \hat{b}\bar{X}$$

$\hat{b}, \hat{\alpha}$ υπολογισμένα από απλή γραμμική παλινδρόμηση

$$-\hat{\beta}_1 = \ln\hat{b} \Rightarrow \hat{\beta}_1 = -\ln\hat{b}$$

$$\hat{\beta}_3 = \frac{\exp(-\hat{\beta}_1)\left[\exp(\hat{\beta}_1) - 1\right]}{\hat{\alpha}}$$

---

$\hat{\beta}_2$ ελεύθερο $\qquad \{y_1, y_2, \ldots, y_n\} \rightarrow \{(y_1, y_2), (y_2, y_3), \ldots, (y_{n-1}, y_n)\}$

$1 \rightarrow y_1$

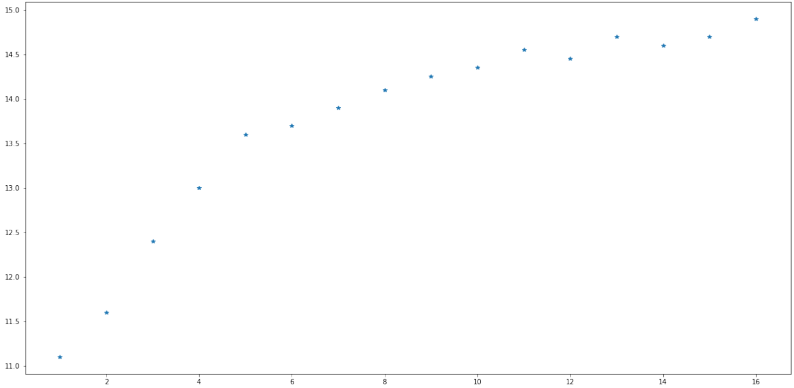
Επιβάλλουμε την συνθήκη.

$$f(1) = y_1 \Rightarrow \frac{\hat{\beta}_3}{1 + \underline{\hat{\beta}_2} \exp(-\hat{\beta}_1)} = y_1 \Rightarrow \text{λύνω ως προς } \hat{\beta}_2$$

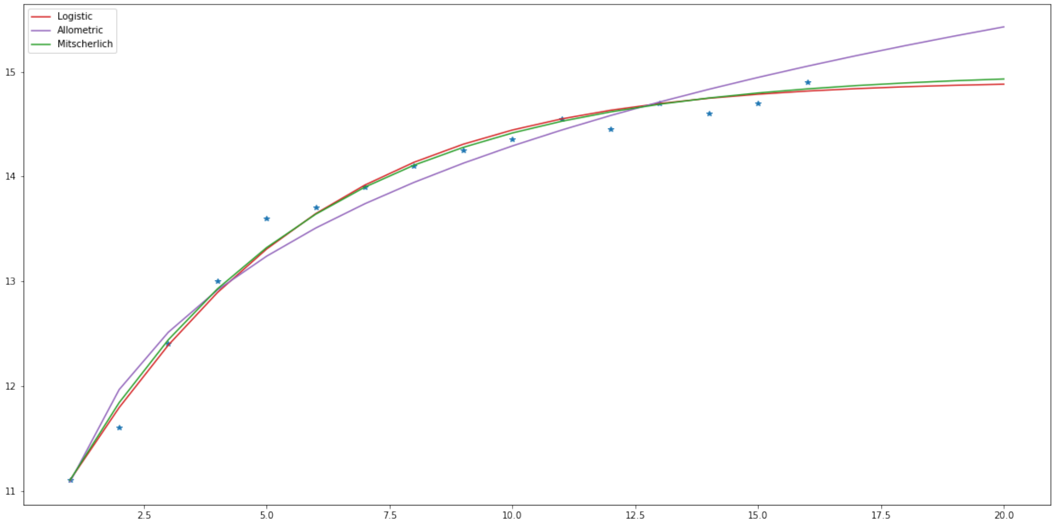
## Παράδειγμα

$$\left\{ \left(\frac{1}{11.1}, \frac{1}{11.6}\right), \left(\frac{1}{11.6}, \frac{1}{12.4}\right), \ldots, \left(\frac{1}{14.7}, \frac{1}{14.9}\right) \right\}$$

$$\{11.1, 11.6, 12.4, 13.0, 13.6, 13.7, 13.9, 14.1, 14.25, 14.35, 14.55, 14.45, 14.7, 14.6, 14.7, 14.9\}$$

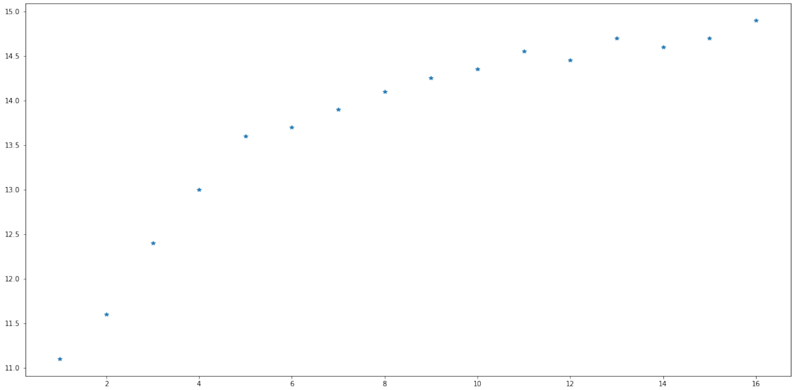


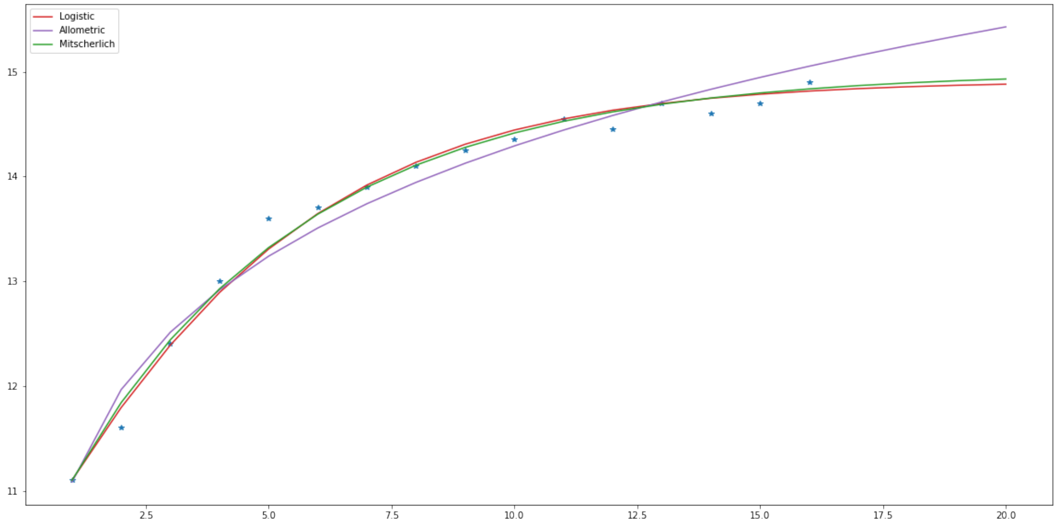
# Χρονολογικές Σειρές (Time Series)



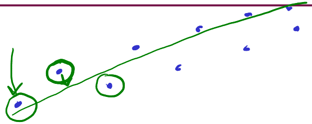
## Παράδειγμα

$$\{11.1, 11.6, 12.4, 13.0, 13.6, 13.7, 13.9, 14.1, 14.25, 14.35, 14.55, 14.45, 14.7, 14.6, 14.7, 14.9\}$$





Logistic
Allometric
Mitscherlich
15
14
13
12
11
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0



$s = 1$

## Εφαρμογή γραμμικού φίλτρου στη χρονολογική σειρά

$$\mathbf{a} = [a_{-s}, \dots, a_s]^T, \quad \mathbf{a}^T\mathbf{a} = 1,\ a_u \geq 0$$

$$Y_t^* = \sum_{u=-s}^{s} a_u Y_{t+u} =$$

$$= \alpha_{-s} Y_{t-s} + \cdots + \alpha_s Y_{t+s}$$

- Απλός κινητός μέσος τάξης $2s + 1$

$3 = 2 \cdot 1 + 1$ $\alpha = \left[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right]$

$$a_u = \frac{1}{2s+1}, \quad u = -s, \dots, s$$

- Απλός κινητός μέσος τάξης $2s$

$2 = 2 \cdot 1$ $\alpha = \left[\frac{1}{4}\ \frac{1}{2}\ \frac{1}{4}\right]$

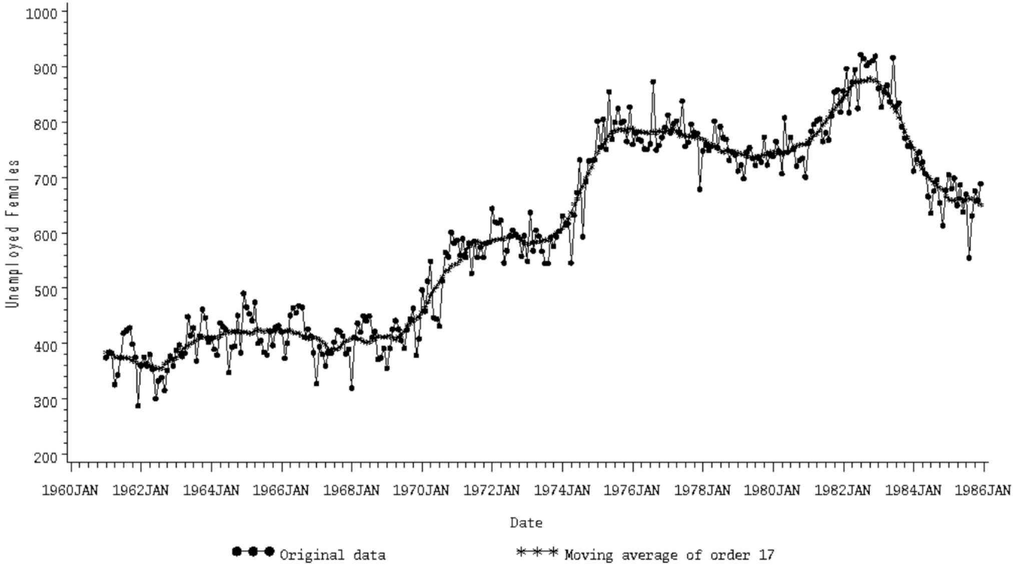
$$a_u = \frac{1}{2s}, \quad u = -s+1, \dots, s-1, \quad a_{-s} = a_s = \frac{1}{4s}$$

## Παράδειγμα

Ποιά είναι τα διανύσματα συντελεστών για τα γραμμικά φίλτρα που αντιστοιχούν στους κινητούς μέσους με τάξεις 4 και 5;

$4 = 2 \cdot 2$ $\alpha = \left[\frac{1}{8}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{8}\right]$

$5 = 2 \cdot 2 + 1$ $\alpha = \left[\frac{1}{5}, 1/5, 1/5, 1/5, 1/5\right]$


1000
900
800
700
600
500
400
300
200
Unemployed Females
1960JAN
1962JAN
1964JAN
1966JAN
1968JAN
1970JAN
1972JAN
1974JAN
1976JAN
1978JAN
1980JAN
1982JAN
1984JAN
1986JAN
Date
Original data
Moving average of order 17

# Απλός Κινητός Μέσος (Simple Moving Average)

## Παράδειγμα

Εφαρμόστε το φίλτρο για τον απλό κινητό μέσο 3ης τάξεως στην παρακάτω χρονολογική σειρά

$$\{1,\ 3,\ 5,\ 4,\ 6,\ 5,\ 7\}$$

$\alpha = [1/3, 1/3, 1/3]$

$\{\underset{1}{\sqcup}, 3, 4, 5, 5, 6, \underset{7}{\sqcup}\}$

# Απλός Κινητός Μέσος (Simple Moving Average)