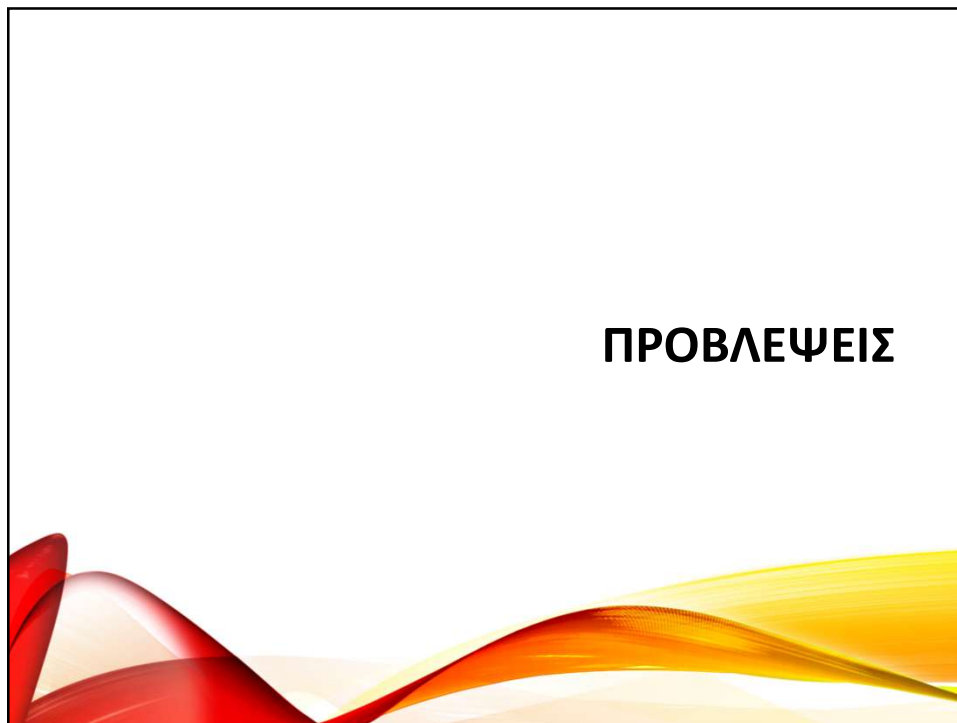




1



2

1

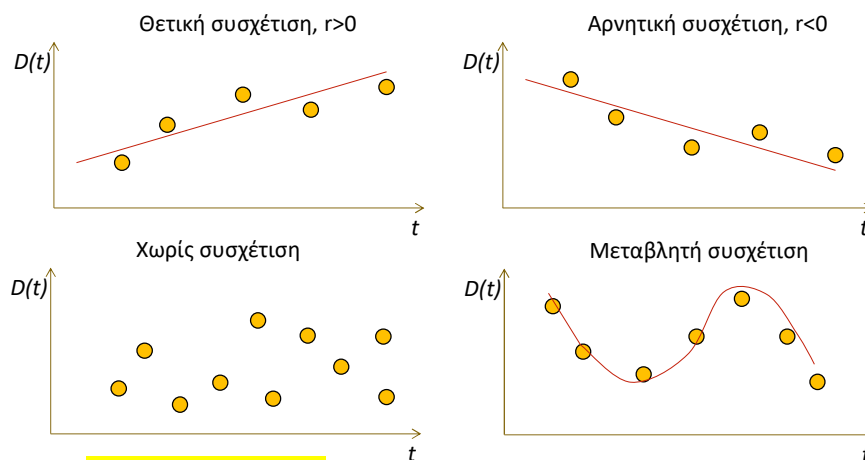
## ΣΥΣΧΕΤΙΣΗ ΚΑΙ ΠΑΛΙΝΔΡΟΜΗΣΗ

- Η συσχέτιση (correlation) εξετάζει εάν υπάρχει σχέση μεταξύ δύο μεταβλητών και εφόσον υπάρχει, σε τι έκταση
- Η παλινδρόμηση (regression) θεσπίζει την κατάλληλη σχέση μεταξύ δύο μεταβλητών
- Ο συντελεστής συσχέτισης μεταξύ δύο μεταβλητών  $X$  και  $Y$  ορίζεται ως:

$$r = \frac{\sum (X - \bar{X}) \cdot (Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 \cdot \sum (Y - \bar{Y})^2}} \quad \text{ή} \quad \frac{\frac{1}{N} \sum (X - \bar{X}) \cdot (Y - \bar{Y})}{\sigma_X \cdot \sigma_Y}$$

3

## ΔΙΑΓΡΑΜΜΑΤΑ ΔΙΑΣΠΟΡΑΣ



Παραδείγματα στο Excel

4

2

## ΠΑΡΑΔΕΙΓΜΑ ΥΠΟΛΟΓΙΣΜΟΥ ΣΥΣΧΕΤΙΣΗΣ

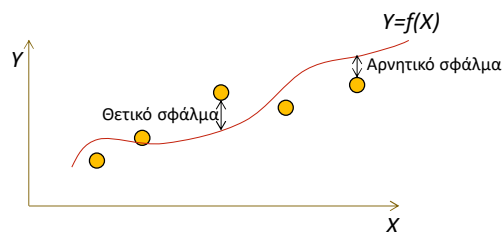
| ΕΤΗ                    | X          | Y            | $X-\mu_X$   | $Y-\mu_Y$ | $(X-\mu_X)^2$ | $(Y-\mu_Y)^2$  | $(X-\mu_X)(Y-\mu_Y)$ |
|------------------------|------------|--------------|-------------|-----------|---------------|----------------|----------------------|
| 1                      | 50         | 700          | 21          | 274       | 441           | 75.076         | 5.754                |
| 2                      | 50         | 650          | 21          | 224       | 441           | 50.176         | 4.704                |
| 3                      | 50         | 600          | 21          | 174       | 441           | 30.276         | 3.654                |
| 4                      | 40         | 500          | 11          | 74        | 121           | 5.476          | 814                  |
| 5                      | 30         | 450          | 1           | 24        | 1             | 576            | 24                   |
| 6                      | 20         | 400          | -9          | -26       | 81            | 676            | 234                  |
| 7                      | 20         | 300          | -9          | -126      | 81            | 15.876         | 1.134                |
| 8                      | 15         | 250          | -14         | -176      | 196           | 30.976         | 2.464                |
| 9                      | 10         | 210          | -19         | -216      | 361           | 46.656         | 4.104                |
| 10                     | 5          | 200          | -24         | -226      | 576           | 51.076         | 5.424                |
| <b>Σύνολο</b>          | <b>290</b> | <b>4.260</b> | <b>0</b>    | <b>0</b>  | <b>2.740</b>  | <b>306.840</b> | <b>28.310</b>        |
| X: Διαφημιστική δαπάνη |            |              | Y: Πωλήσεις |           |               |                |                      |
| $\mu_X$                | 29         |              | $r$         | 0,976     |               |                |                      |
| $\mu_Y$                | 426        |              | $r^2$       | 0,953     |               |                |                      |

Στο Excel μπορούμε να χρησιμοποιήσουμε τη συνάρτηση CORREL

5

## ΤΙ ΕΙΝΑΙ ΠΑΛΙΝΔΡΟΜΗΣΗ;

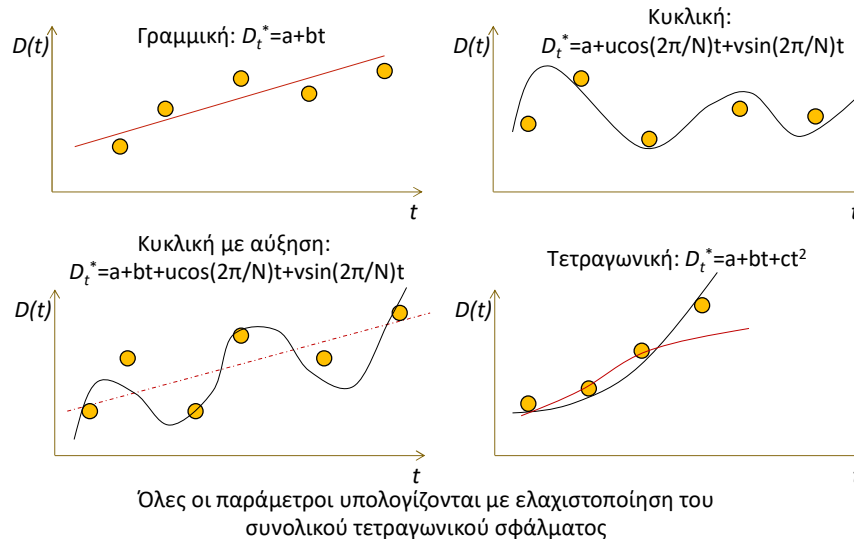
- Ο προσδιορισμός του πώς μία εξαρτώμενη μεταβλητή (Y) σχετίζεται με μία ή περισσότερες ανεξάρτητες μεταβλητές (X)
- Ως κριτήριο καλύτερης προσαρμογής είναι η ελαχιστοποίηση κάποιου στατιστικού δείκτη (ΜΣ, ΜΑΑ, ΜΤΣ, κλπ.)· το προτιμώμενο κριτήριο είναι η ελαχιστοποίηση του ΜΤΣ ή του αθροιστικού ΤΣ



6

3

## ΣΥΝΗΘΕΙΣ ΣΥΝΑΡΤΗΣΕΙΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ



7

## ΑΝΑΛΥΣΗ ΠΑΛΙΝΔΡΟΜΗΣΗΣ

- Στην *ανάλυση συσχέτισης ή παλινδρόμησης* χρησιμοποιείται μια **εξίσωση συσχέτισης ή παλινδρόμησης (regression)** που εκφράζει την εξαρτημένη μεταβλητή συναρτήσει των ανεξάρτητων μεταβλητών
- Αν η ανεξάρτητη μεταβλητή είναι μία, η σχέση μεταξύ των δύο βρίσκεται χρησιμοποιώντας την ανάλυση απλής παλινδρόμησης, ενώ όταν είναι περισσότερες χρησιμοποιείται η πολλαπλή παλινδρόμηση
- Η **γραμμική παλινδρόμηση (linear regression)** βασίζεται στην υπόθεση ότι η προς πρόβλεψη μεταβλητή εξαρτάται από μία ανεξάρτητη μεταβλητή με γραμμική σχέση της μορφής  $Y = \alpha + \beta X$
- Μια σχέση που δεν είναι γραμμική είναι δυνατό να μπορεί να μελετηθεί με την ανάλυση παλινδρόμησης, με κατάλληλο μετασχηματισμό, π.χ.:

$$Y = \alpha + \beta/X \quad \rightarrow \quad Y = \alpha + \beta Z$$

$$Y = \alpha + X^\beta \quad \rightarrow \quad \log Y = \log \alpha + \beta \log X$$

8

4

## ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

- Έστω ένα σύνολο  $n$  ζευγών δεδομένων  $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$  για τις τιμές των δύο μεταβλητών  $X$  και  $Y$
- Με δεδομένο αυτό το σύνολο, θα μπορούσε να προσδιοριστεί μια ευθεία, που θα εξέφραζε μια γραμμική σχέση μεταξύ των δύο μεταβλητών, σε σχέση με την οποία οι αποκλίσεις των σημείων θα ήταν ελάχιστες, δηλαδή ένα «ιδανικό» θεωρητικό πρότυπο
- Η εύρεση της ευθείας μπορεί να γίνει με τη **μέθοδο των ελαχίστων τετραγώνων (least-squares method)**
- Μια ιδιότητα της ευθείας που προκύπτει με τη μέθοδο των ελαχίστων τετραγώνων είναι ότι το άθροισμα των αποκλίσεων από την ευθεία ισούται με μηδέν, δηλαδή το άθροισμα των θετικών αποκλίσεων ισούται με το άθροισμα των αρνητικών αποκλίσεων

9

## ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

- Η πρόβλεψη της τιμής της εξαρτημένης μεταβλητής  $Y_0$ , όταν γνωρίζουμε (ή προβλέπουμε) ότι η τιμή της ανεξάρτητης μεταβλητής θα είναι ίση με  $X_0$ , θα προκύπτει από την παραπάνω σχέση, αν αντικαθιστούμε το  $X$  με  $X_0$  και υπολογίζουμε την αντίστοιχη τιμή  $Y_0$  του  $Y$  (ή  $\hat{Y}$ )
- Η υπογράμμιση (ή το «καπελάκι») συμβολίζει την εκτιμώμενη τιμή
- Στην πράξη είναι πιθανό η πραγματική τιμή της  $Y$  να αποδειχθεί διαφορετική από την προβλεπόμενη
- Η απόκλιση της προβλεπόμενης από την πραγματική τιμή θα είναι τόσο μεγαλύτερη όσο μεγαλύτερη είναι η διασπορά των σημείων που αντιστοιχούν στα ζεύγη  $(X_i, Y_i)$  σε σχέση με την ευθεία  $\underline{Y} = \alpha + \beta X$

10

5

## ΠΡΟΣΑΡΜΟΓΗ ΕΥΘΕΙΑΣ ΓΡΑΜΜΗΣ

- Προσαρμογή ευθείας γραμμής = Linear fitting
- $F(t)=a+bt$  είναι η εξίσωση της ευθείας που θα προσαρμοσθεί
- $F(t)$  είναι η προσαρμοσμένη συνάρτηση την περίοδο  $t$  και εκφράζει την πρόβλεψη για τη χρονική στιγμή  $t$
- $D(t)$  είναι η πραγματική ζήτηση την περίοδο  $t$
- Θα πρέπει να υπάρχουν διαθέσιμα προηγούμενα δεδομένα για  $n$  περιόδους
- Οι παράμετροι  $a$  και  $b$  υπολογίζονται από τα δεδομένα χρησιμοποιώντας το κριτήριο ελαχιστοποίησης του αθροίσματος των τετραγώνων των σφαλμάτων (minimization of sum of square errors)

11

## ΚΑΝΟΝΙΚΕΣ ΕΞΙΣΩΣΕΙΣ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ

$$SSE = \sum_{t=1}^N (D_t - F_t)^2 = \sum_{t=1}^N (D_t - a - bt)^2$$

$$\left. \begin{aligned} \frac{d(SSE)}{da} &= \sum_{t=1}^N 2(D_t - a - bt)(-1) = 0 \\ \frac{d(SSE)}{db} &= \sum_{t=1}^N 2(D_t - a - bt)(-t) = 0 \end{aligned} \right\} \Rightarrow \begin{aligned} n \cdot a + b \sum_{t=1}^n t &= \sum_{t=1}^n D_t \\ a \cdot \sum_{t=1}^n t + b \sum_{t=1}^n t^2 &= \sum_{t=1}^n t \cdot D_t \end{aligned}$$

$$a = \frac{\begin{vmatrix} \sum D_t & \sum t \\ \sum t \cdot D_t & \sum t^2 \end{vmatrix}}{\begin{vmatrix} n & \sum t \\ \sum t & \sum t^2 \end{vmatrix}} = \frac{\sum D_t \cdot \sum t^2 - \sum t \cdot \sum t \cdot D_t}{n \cdot \sum t^2 - (\sum t)^2} = \dots = \mu_{D_t} - b \cdot \mu_t$$

$$b = \frac{\begin{vmatrix} n & \sum D_t \\ \sum t & \sum t \cdot D_t \end{vmatrix}}{\begin{vmatrix} n & \sum t \\ \sum t & \sum t^2 \end{vmatrix}} = \frac{n \cdot \sum t \cdot D_t - \sum t \cdot \sum D_t}{n \cdot \sum t^2 - (\sum t)^2} = \frac{\sum t \cdot D_t - n \cdot \mu_t \cdot \mu_{D_t}}{\sum t^2 - n \cdot \mu_t^2}$$

12

## ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ – ΛΥΜΕΝΟ ΠΑΡΑΔΕΙΓΜΑ #1

| Μήνας       | t <sub>i</sub> | Ζήτηση (D <sub>i</sub> ) | t <sub>i</sub> D <sub>i</sub> | t <sub>i</sub> <sup>2</sup> |
|-------------|----------------|--------------------------|-------------------------------|-----------------------------|
| Jan         | 1              | 199                      | 199                           | 1                           |
| Feb         | 2              | 202                      | 404                           | 4                           |
| Mar         | 3              | 199                      | 597                           | 9                           |
| Apr         | 4              | 208                      | 832                           | 16                          |
| May         | 5              | 212                      | 1060                          | 25                          |
| Jun         | 6              | 194                      | 1164                          | 36                          |
| Jul         | 7              | 214                      | 1498                          | 49                          |
| Aug         | 8              | 220                      | 1760                          | 64                          |
| Sep         | 9              | 219                      | 1971                          | 81                          |
| Oct         | 10             | 234                      | 2340                          | 100                         |
| Nov         | 11             | 219                      | 2409                          | 121                         |
| Dec         | 12             | 233                      | 2796                          | 144                         |
| <b>SUM</b>  | <b>78</b>      | <b>2.553</b>             | <b>17.030</b>                 | <b>650</b>                  |
| <b>MEAN</b> | <b>7</b>       | <b>213</b>               |                               |                             |

a 192,95

b 3,05

$$F_t = 193 + 3t$$

13

## ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ – ΛΥΜΕΝΟ ΠΑΡΑΔΕΙΓΜΑ #1

| Μήνας | Ζήτηση | Γραμμική<br>παλινδρόμηση |
|-------|--------|--------------------------|
| Jan   | 199    | 196                      |
| Feb   | 202    | 199                      |
| Mar   | 199    | 202                      |
| Apr   | 208    | 205                      |
| May   | 212    | 208                      |
| Jun   | 194    | 211                      |
| Jul   | 214    | 214                      |
| Aug   | 220    | 217                      |
| Sep   | 219    | 220                      |
| Oct   | 234    | 223                      |
| Nov   | 219    | 226                      |
| Dec   | 233    | 229                      |
|       |        | 232                      |

$$r = 0,8439$$

14

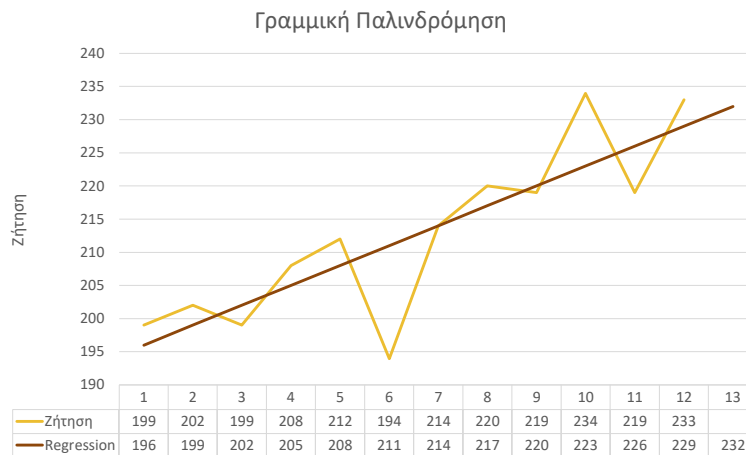
## ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ – ΛΥΜΕΝΟ ΠΑΡΑΔΕΙΓΜΑ #1

- $D_t$ =πραγματική ζήτηση την περίοδο  $t$
- $F_t$ =πρόβλεψη για την περίοδο  $t$
- $N$ = πλήθος δεδομένων
- Τυπικό σφάλμα εκτίμησης:  $\sigma = \sqrt{\left(\frac{1}{N-f}\right) \sum_1^N (D_t - F_t)^2} = 7,32$
- $f$ =αριθμός χαμένων βαθμών ελευθερίας (για linear regression  $f=2$  καθότι προσδιορίζουμε δύο παραμέτρους,  $a$  και  $b$ )
- Μοντέλο πρόβλεψης:  $F_t=193+3t$
- **Βεβαιότητα 95% για την πρόβλεψη του επόμενου Ιανουαρίου:  $232 \pm 14,6$  (δηλαδή,  $\pm 2\sigma$ )\***
- Η ζήτηση θα είναι μεταξύ 218 και 246 με βεβαιότητα 95%

\* Εξήγηση με το Excel

15

## ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ – ΛΥΜΕΝΟ ΠΑΡΑΔΕΙΓΜΑ #1



16

## ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ – ΛΥΜΕΝΟ ΠΑΡΑΔΕΙΓΜΑ #2

Μία τσιμεντοβιομηχανία έχει ένα εργοστάσιο με μέγιστη δυνατότητα παραγωγής 160 εκατομμύρια τόνους ετησίως. Από τα προηγούμενα έτη, οι τιμές των τριμηνιαίων δεικτών εποχικότητας είναι οι ακόλουθες:

$$I_1=0.85 \quad I_2=1.05 \quad I_3=1.2 \quad I_4=0.9$$

Η διοίκηση πιστεύει ότι οι πωλήσεις αυξάνονται κατά ένα σταθερό ποσό ετησίως. Ένα γραμμικό μοντέλο ανάλυσης παλινδρόμησης που αναπτύχθηκε για την πρόβλεψη των πωλήσεων είναι το ακόλουθο:

$$Y=a+bX=127.35+4.68X$$

όπου Y: η πρόβλεψη ετήσιων πωλήσεων σε εκατομμύρια τόνους

X: το έτος για το οποίο προβλέπονται οι πωλήσεις.

Το έτος εκκίνησης είναι το 2007.

1. Χρησιμοποιώντας την ευθεία παλινδρόμησης να εξαχθεί πρόβλεψη για τις πωλήσεις το έτος 2011.
2. Χρησιμοποιώντας τις τιμές των δεικτών εποχικότητας υπολογίστε μία πρόβλεψη για κάθε τρίμηνο του 2011.

17

## ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ – ΛΥΜΕΝΟ ΠΑΡΑΔΕΙΓΜΑ #2

1. Το έτος 2011 είναι το 5<sup>ο</sup> έτος της ανάλυσης (2007, 8, 9, 10 και 11). Επομένως, στην εξίσωση παλινδρόμησης πρέπει να θέσουμε X=5. Τότε:

$$Y_{2011}=127.35+4.68 \times 5=150.75 \text{ Mtons}$$

2. Για να υπολογισθούν οι προβλέψεις για κάθε τρίμηνο του 2011, θα χρησιμοποιηθεί ως βάση η μέση πρόβλεψη για κάθε τρίμηνο του 2011 και θα πολλαπλασιαστεί με τους αντίστοιχους συντελεστές εποχικότητας. Έτσι:

$$Y_{2011,1}=0.85 \times 150.75 / 4 = 32.03 \text{ Mtons}$$

$$Y_{2011,2}=1.05 \times 150.75 / 4 = 39.57 \text{ Mtons}$$

$$Y_{2011,3}=1.2 \times 150.75 / 4 = 45.23 \text{ Mtons}$$

$$Y_{2011,4}=0.9 \times 150.75 / 4 = 33.92 \text{ Mtons}$$

18

## ΑΝΑΛΥΣΗ ΓΡΑΜΜΙΚΗΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ

- Ένα μέτρο του μεγέθους της διασποράς των σημείων γύρω από την ευθεία, κατά συνέπεια και της αξιοπιστίας της πρόβλεψης που γίνεται με βάση αυτή, μπορεί να βρεθεί χρησιμοποιώντας το **τυπικό σφάλμα συσχέτισης ΤΣΣ**
- Αν τα σημεία είναι διεσπαρμένα κατά τυχαίο τρόπο γύρω από την ευθεία παλινδρόμησης, τότε κατά προσέγγιση τα 2/3 των σημείων θα βρίσκονται μέσα σε μια ζώνη πλάτους ίσου με  $2 \cdot \text{ΤΣΣ}$
- Ο **συντελεστής συσχέτισης  $r$**  μετράει το βαθμό της σχέσης που υπάρχει μεταξύ των δύο μεταβλητών και οι τιμές του κυμαίνονται μεταξύ  $-1$  και  $+1$ : όταν  $r \rightarrow 0$  τόσο μικρότερη είναι η συσχέτιση των δύο μεταβλητών
- Εάν  $r = +1$  (θετική συσχέτιση) ή  $r = -1$  (αρνητική συσχέτιση) οι μεταβλητές έχουν απόλυτη συσχέτιση: στην πρώτη περίπτωση, όταν αυξάνεται η τιμή της ανεξάρτητης μεταβλητής, αυξάνεται αντίστοιχα η τιμή της εξαρτημένης μεταβλητής: το αντίθετο συμβαίνει στη δεύτερη περίπτωση

19

## ΑΝΑΛΥΣΗ ΓΡΑΜΜΙΚΗΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ

- Η γραμμικότητα της σχέσης, δηλαδή ο βαθμός στον οποίο ισχύει η υπόθεση ότι υπάρχει μια σχέση μεταξύ  $Y$  και  $X$  που είναι γραμμική, μετριέται με το **συντελεστή προσδιορισμού  $r^2$**
- Ο συντελεστής προσδιορισμού παίρνει τιμές μεταξύ 0 και 1 και αποτελεί μέτρο της καταλληλότητας χρήσης της γραμμικής σχέσης που προκύπτει με βάση τα παραπάνω
- Δίνει το ποσοστό της συνολικής απόκλισης που εξηγείται από την ευθεία συσχέτισης
- Αν η απόκλιση μεταξύ των πραγματικών τιμών του  $Y$  και των αντίστοιχων εκτιμήσεων ήταν 0, τότε θα προκύπτει  $r^2 = 1$ : αν, αντίθετα, ο συντελεστής προσδιορισμού έχει τιμή κοντά στο 0, τότε η ευθεία συσχέτισης  $\hat{Y} = \alpha + \beta X$  εξηγεί ελάχιστα τις αποκλίσεις, δηλαδή η ευθεία (και η αντίστοιχη γραμμική σχέση που αυτή εκφράζει) είναι ακατάλληλη για να εξηγήσει τη σχέση που συνδέει τις δύο μεταβλητές

20

10

## ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ – ΛΥΜΕΝΟ ΠΑΡΑΔΕΙΓΜΑ #3

- Ένας χρηματιστηριακός αναλυτής ενδιαφέρεται για τη σχέση ανάμεσα στην επί τοις εκατό μέση μηνιαία απόδοση μιας (Y) μετοχής και την αντίστοιχη απόδοση (X) του Γενικού Δείκτη Τιμών του Χρηματιστηρίου. Για το σκοπό αυτό συνέλεξε τα μηνιαία στοιχεία που συνοψίζονται στον παρακάτω πίνακα. Με βάση τα στοιχεία αυτά και αν επιπλέον υποτεθεί ότι οι μεταβλητές X και Y συνδέονται με γραμμική σχέση να εκτιμηθεί:

| Μήνας | 1  | 2  | 3 | 4 | 5  |
|-------|----|----|---|---|----|
| X (%) | -2 | -3 | 0 | 1 | 1  |
| Y (%) | -1 | 0  | 0 | 1 | -1 |

- Η ευθεία παλινδρόμησης της Y πάνω στη X
- Ο συντελεστής συσχέτισης και να ερμηνευτεί
- Ο συντελεστής προσδιορισμού και να ερμηνευτεί
- Η απόδοση της συγκεκριμένης μετοχής αν είναι γνωστό ότι ο Γενικός Δείκτης Τιμών έχει απόδοση 1%

21

## ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ – ΛΥΜΕΝΟ ΠΑΡΑΔΕΙΓΜΑ #3

| $X_i$ | $Y_i$ | $X_i^2$ | $Y_i^2$ | $X_i Y_i$ |
|-------|-------|---------|---------|-----------|
| -2    | -1    | 4       | 1       | 2         |
| -3    | 0     | 9       | 0       | 0         |
| 0     | 0     | 0       | 0       | 0         |
| 1     | 1     | 1       | 1       | 1         |
| 1     | -1    | 1       | 1       | -1        |
| -3    | -1    | 15      | 3       | 2         |

$$SSE = \sum_{i=1}^N (Y_i - F_i)^2 = \sum_{i=1}^N (Y_i - a - bX_i)^2$$

$$a = \frac{\sum Y_i \cdot \sum X_i^2 - \sum X_i \cdot \sum X_i \cdot Y_i}{n \cdot \sum X_i^2 - (\sum X_i)^2} = \dots = \mu_{Y_i} - b \cdot \mu_{X_i} = \frac{(-1) \cdot 15 - (-3) \cdot 2}{5 \cdot 15 - (-3)^2} = \frac{-9}{66} = -0,136$$

$$b = \frac{n \cdot \sum X_i \cdot Y_i - \sum X_i \cdot \sum Y_i}{n \cdot \sum X_i^2 - (\sum X_i)^2} = \frac{\sum X_i \cdot Y_i - n \cdot \mu_{X_i} \cdot \mu_{Y_i}}{\sum X_i^2 - n \cdot \mu_{X_i}^2} = \frac{5 \cdot 2 - (-3) \cdot (-1)}{5 \cdot 15 - (-3)^2} = \frac{7}{66} = 0,106$$

$$\hat{Y} = -0,136 + 0,106 \cdot X$$

22

11

## ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ – ΛΥΜΕΝΟ ΠΑΡΑΔΕΙΓΜΑ #3

2. Ο συντελεστής συσχέτισης είναι:  $r = \frac{\sum(X-\bar{X}) \cdot (Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2 \cdot \sum(Y-\bar{Y})^2}} = 0,2383$

Η συσχέτιση επομένως μεταξύ Χ και Υ είναι ασθενής

3. Ο συντελεστής προσδιορισμού  $r^2$  είναι 0.0567, άρα η μεταβολή του Γενικού Δείκτη Τιμών ερμηνεύει το 5.6% της μεταβλητότητας της μέσης μηνιαίας απόδοσης της συγκεκριμένης μετοχής
4. Η ευθεία παλινδρόμησης είναι  $Y = -0.136 + 0.106X$ . Για  $X=1$  προκύπτει  $Y = -0.03$ . Αυτό σημαίνει ότι όταν ΓΔΤ έχει απόδοση 1%, η μετοχή έχει απόδοση -0.03%.

23

## ΑΙΤΙΟΚΡΑΤΙΚΑ ΜΟΝΤΕΛΑ

- Η ζήτηση σχετίζεται με αιτιατές (ή αιτιοκρατικές-causal) μεταβλητές
- Η ζήτηση για ελαστικά, π.χ., μπορεί να μοντελοποιηθεί ως συνάρτηση της παραγωγής νέων αυτοκινήτων, των αντικαταστάσεων παλαιών, της κυβερνητικής πολιτικής για αυτοκίνητα, κλπ), ως  $D_t = \alpha P_t + \beta P_{t-5} + \gamma$ , όπου  $\alpha$ ,  $\beta$ ,  $\gamma$  παράμετροι που προκύπτουν από μοντέλο παλινδρόμησης
- Για να είναι χρήσιμο ένα αιτιοκρατικό μοντέλο, θα πρέπει οι μεταβλητές του να είναι προηγούμενες των προβλέψεων και να παρουσιάζουν ισχυρή συσχέτιση με την κύρια μεταβλητή ενδιαφέροντος

24

## ΑΙΤΙΟΚΡΑΤΙΚΑ ΜΟΝΤΕΛΑ

- Δυο τύποι αιτιοκρατικών μοντέλων, η *ανάλυση συσχέτισης ή παλινδρόμησης* και οι *οικονομετρικές μέθοδοι*, χρησιμοποιούνται συχνά στην πράξη.
- Οι οικονομετρικές μέθοδοι αποτελούν επέκταση της ανάλυσης παλινδρόμησης και περιλαμβάνουν ένα σύστημα εξισώσεων παλινδρόμησης, π.χ.:
  - Πωλήσεις =  $f_1$  (ΑΕΠ, τιμή πωλήσεων, διαφήμιση)
  - Κόστος παραγωγής =  $f_2$  (επίπεδα παραγωγής και αποθεμάτων)
  - Έξοδα πωλήσεων =  $f_3$  (διαφήμιση, άλλα έξοδα πωλήσεων)
  - Τιμή πωλήσεων =  $f_4$  (κόστος και έξοδα πωλήσεων)

25

## ΑΝΑΛΥΣΗ ΧΡΟΝΟΣΕΙΡΩΝ

- Οι χρονοσειρές αναλύονται σύμφωνα με:
    - Τάση
    - Εποχικότητα
    - Κυκλικότητα
    - Τυχαιότητα
- και αναπτύσσεται η πρόβλεψη από τη σύνθεση αυτών των συνιστωσών

26

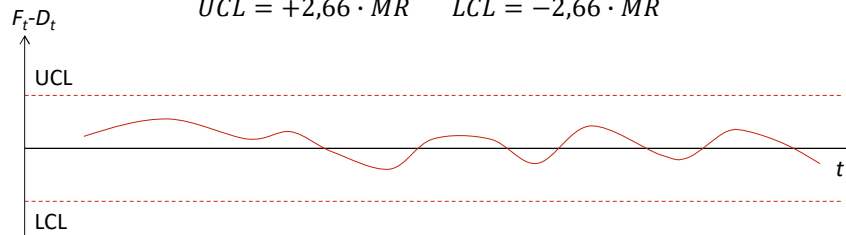
## ΑΞΙΟΛΟΓΗΣΗ ΜΟΝΤΕΛΟΥ ΠΡΟΒΛΕΨΗΣ

- Διάγραμμα **κινούμενου εύρους** (moving range) για τον έλεγχο των προβλέψεων:

$$MR = |(F_t - D_t) - (F_{t-1} - D_{t-1})|$$

$$\overline{MR} = \frac{1}{N-1} \sum MR$$

$$UCL = +2,66 \cdot \overline{MR} \quad LCL = -2,66 \cdot \overline{MR}$$



27

## ΑΞΙΟΛΟΓΗΣΗ ΜΟΝΤΕΛΟΥ ΠΡΟΒΛΕΨΗΣ

- Η στατιστική U του Theil, χρησιμοποιείται για την αξιολόγηση μιας συγκεκριμένης μεθόδου πρόβλεψης σε σχέση με την απλούστερη δυνατή

$$U = \sqrt{\frac{\sum_{t=1}^{n-1} (FPE_{t+1} - APE_{t+1})^2}{\sum_{t=1}^{n-1} (APE_{t+1})^2}}$$

$$FPE_{t+1} = \frac{F_{t+1} - Y_t}{Y_t} \quad \text{και} \quad APE_{t+1} = \frac{Y_{t+1} - Y_t}{Y_t}$$

- Η τιμή  $U=1$  επιτυγχάνεται όταν η μέθοδος πρόβλεψης που χρησιμοποιούμε είναι ισοδύναμη με την απλούστερη δυνατή, όταν δηλαδή η πρόβλεψη για την επόμενη περίοδο είναι ίση με την τρέχουσα τιμή της υπό πρόβλεψη μεταβλητής· όσο μικρότερη της μονάδας είναι η στατιστική U τόσο καλύτερη η μέθοδος σε σχέση με την απλούστερη

28

## ΑΞΙΟΛΟΓΗΣΗ ΜΟΝΤΕΛΟΥ ΠΡΟΒΛΕΨΗΣ

- Η στατιστική των Durbin και Watson χρησιμοποιείται για την ανίχνευση της ύπαρξης σχέσεων, μεταξύ δύο διαδοχικών σφαλμάτων μετά την εφαρμογή κάποιας μεθόδου πρόβλεψης

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}$$

- Αν τα διαδοχικά σφάλματα είναι τυχαία, τότε  $DW \cong 2$

29

## ΣΥΝΟΨΗ

- Εξηγήθηκε η σημασία των προβλέψεων στον προγραμματισμό
- Παρουσιάστηκαν μέθοδοι πρόβλεψης
  - Υποκειμενικές μέθοδοι όπως δημοσκοπήσεις και Delphi
  - Κινούμενοι μέσοι και εκθετική εξομάλυνση
  - Προεκβολή τάσεων με παλινδρόμηση
  - Αιτιοκρατικά μοντέλα
  - Ανάλυση χρονοσειρών
- Παρουσιάστηκε η μέθοδος ελέγχου των προβλέψεων

30