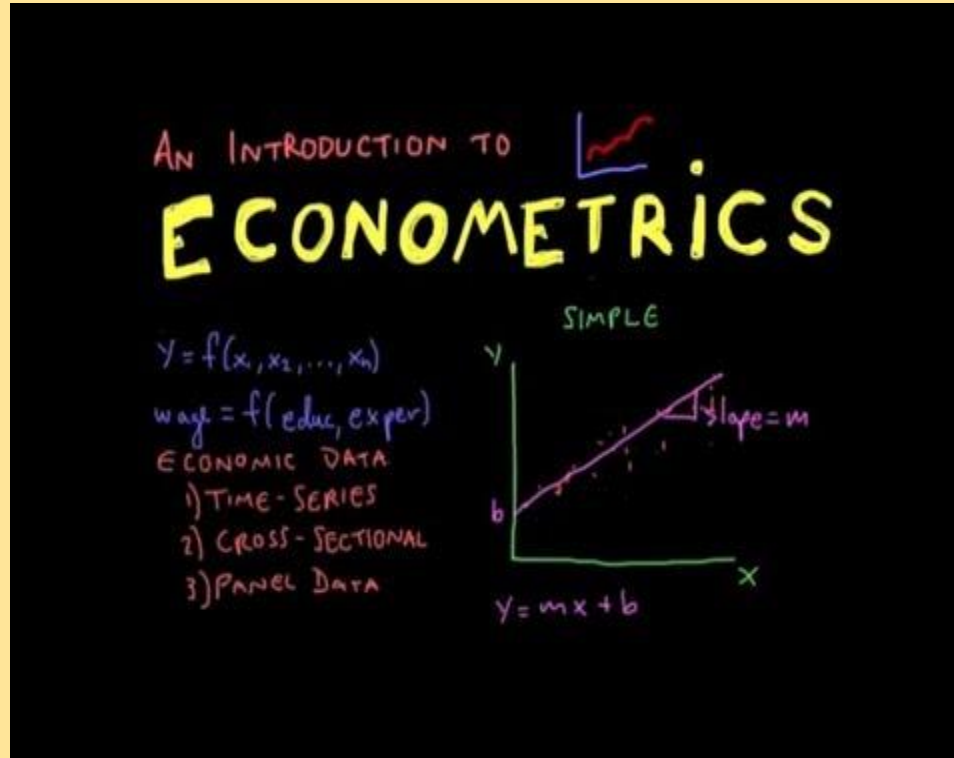


Οικονομικό Πανεπιστήμιο Αθηνών

Μάθημα: "Εφαρμογές Στατιστικών Μεθόδων σε Επιχειρησιακά Προβλήματα"

Συνάντηση: 4η



Δρ. Τρύφωνας Λεμοντζόγλου

**OLS**

**Οι τύποι που πρέπει να ξέρεις ! ! !**

*Συντελεστές Παλινδρόμησης (Μέθοδος Ελαχίστων Τετραγώνων)*

$$\hat{\beta} = \frac{n \sum \psi_i X_i - \sum X_i \sum \psi_i}{n \sum X_i^2 - (\sum X_i)^2}$$

ή

$$\frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

$$\hat{\alpha} = \bar{\psi}_i - \hat{\beta} \bar{X}_i$$

**OLS****Ερμηνεία των συντελεστών**

πχ1 Κατανάλωση(i) = 0,28 + 0,78 \* Εισόδημα(i)

$$\hat{\alpha} \quad \hat{\beta}$$

$$\widehat{\Psi}_i = \hat{a} + \hat{\beta} * X_i$$

Μια αύξηση στην τιμή της ανεξάρτητης μεταβλητής **κατά 1 μονάδα** θα οδηγήσει στην αύξηση της τιμής της εξαρτημένης μεταβλητής **κατά 0,78 μονάδες**.

Έχει νόημα η ερμηνεία του συντελεστή  $\hat{a}$  ;

“Ακόμη και με το εισόδημα να λαμβάνει την τιμή μηδέν θα υπάρχει **ένα ελάχιστο επίπεδο κατανάλωσης ίσο με 0,28 μονάδες**” (βλ. κρατική στήριξη, περιουσία, άλλοι πόροι και έσοδα)

**OLS****Ερμηνεία των συντελεστών**

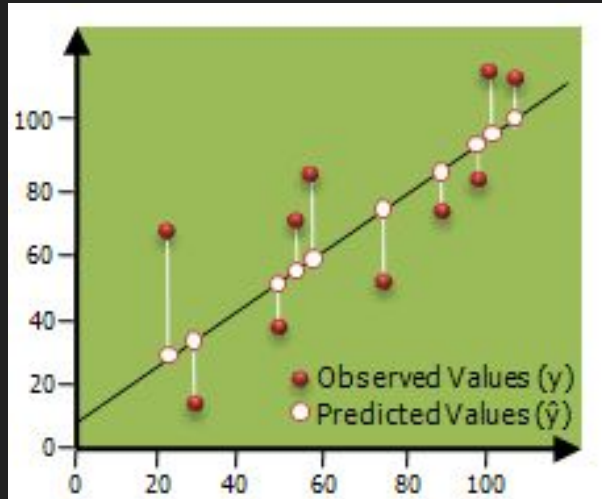
$$\text{πχ2} \quad \underset{\hat{\alpha}}{\text{Κρατήσεις}(t)} = 0,54 - \underset{\hat{\beta}}{1,25} * \text{Σεισμοί}(t)$$

$$\widehat{\Psi}_i = \hat{a} + \hat{\beta} * X_i$$

Μια αύξηση στην τιμή της ανεξάρτητης μεταβλητής **κατά 1 μονάδα** θα οδηγήσει στη μείωση της τιμής της εξαρτημένης μεταβλητής **κατά 1,25 μονάδες**.

Έχει νόημα η ερμηνεία του συντελεστή  $\hat{a}$  ;

“Επίπεδο κρατήσεων ίσο με **0,54 μονάδες**”  
(ανεξάρτητα από τη συχνότητα των σεισμών)



# ΕΦΑΡΜΟΓΕΣ (χρήση πραγματικών δεδομένων)

# 1. OLS (παράδειγμα: κατανάλωση & εισόδημα)

Ας δούμε ένα παράδειγμα με τη χρήση πραγματικών δεδομένων:

**Ψ:** επίπεδο κατανάλωσης (εξαρτημένη μεταβλητή)

**X:** ύψος εισοδήματος (ανεξάρτητη μεταβλητή)

Πηγή άντλησης δεδομένων: AMECO (Ευρωπαϊκή Επιτροπή)

# 1. OLS (παράδειγμα: κατανάλωση & εισόδημα)

Κατανάλωση =  $\Psi$  (εξαρτημένη μεταβλητή)

Εισόδημα =  $X$  (ανεξάρτητη μεταβλητή)

Απλό Γραμμικό Υπόδειγμα ( $k=1$ )  $\rightarrow$  1 ανεξάρτητη μεταβλητή

$\Psi_i = \alpha + \beta * X_i + \epsilon_i$   $\rightarrow$  πραγματική σχέση

Οι πραγματικοί συντελεστές  $\alpha$  και  $\beta$  είναι **άγνωστοι** !!!

Μπορούμε όμως να υπολογίσουμε τους εκτιμητές των συντελεστών !!!

στρογγυλοποίηση ...

| Χώρα   | Έτη (t)<br><i>τιμές εξαρτημένης μεταβλητής</i> | Ιδιωτική Κατανάλωση<br>(σε δισ. ευρώ) | Ακαθάριστο Εθνικό Εισόδημα<br>(σε δισ. ευρώ) |
|--------|------------------------------------------------|---------------------------------------|----------------------------------------------|
| Ελλάδα | t1 = 2006                                      | Y1 = 148                              | X1 = 226                                     |
| Ελλάδα | t2 = 2007                                      | Y2 = 155                              | X2 = 232                                     |
| Ελλάδα | t3 = 2008                                      | Y3 = 161                              | X3 = 231                                     |
| Ελλάδα | t4 = 2009                                      | Y4 = 159                              | X4 = 223                                     |
| Ελλάδα | t5 = 2010                                      | Y5 = 147                              | X5 = 211                                     |
| Ελλάδα | t6 = 2011                                      | Y6 = 135                              | X6 = 189                                     |
| Ελλάδα | t7 = 2012                                      | Y7 = 123                              | X7 = 178                                     |
| ...    | ...                                            | ...                                   | ...                                          |
| Ελλάδα | tn = 2026                                      | Yn = 148                              | Xn = 206                                     |

*τιμές ανεξάρτητης μεταβλητής*

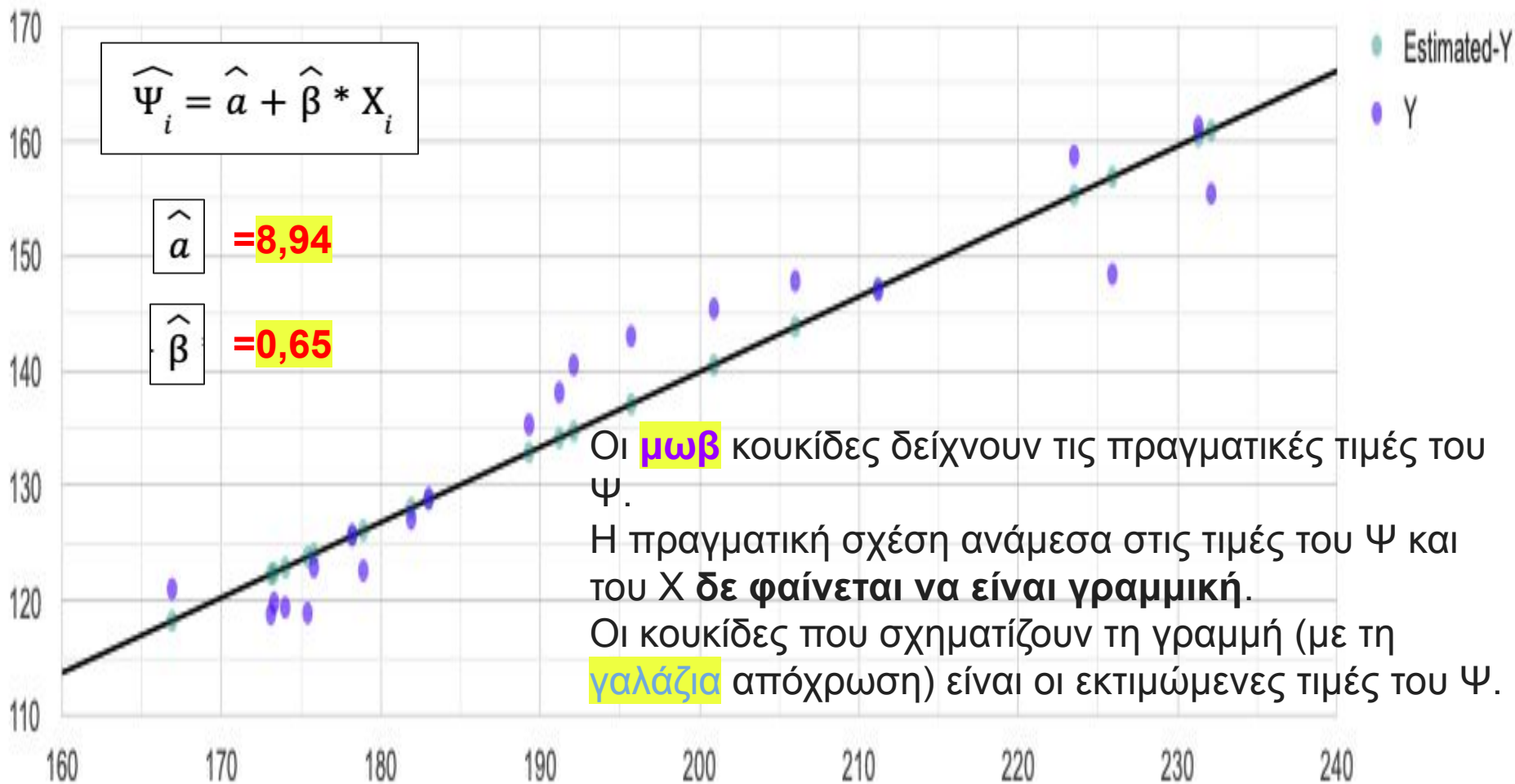


| Variable | Private FCE, at constant prices (OCPH)                                       |       | GNI, at constant prices (OVGN) |       |
|----------|------------------------------------------------------------------------------|-------|--------------------------------|-------|
| Country  | Greece                                                                       |       | Greece                         |       |
| 2026     |                                                                              | 147,8 |                                | 206,0 |
| 2025     |                                                                              | 145,4 |                                | 200,9 |
| 2024     |                                                                              | 143,0 |                                | 195,7 |
| 2023     |                                                                              | 140,5 |                                | 192,1 |
| 2022     |                                                                              | 138,1 |                                | 191,2 |
| 2021     | τιμές εξαρτημένης μεταβλητής                                                 | 127,1 |                                | 181,9 |
| 2020     |                                                                              | 121,0 | τιμές ανεξάρτητης μεταβλητής   | 166,9 |
| 2019     |                                                                              | 128,9 |                                | 183,0 |
| 2018     |                                                                              | 125,7 |                                | 178,2 |
| 2017     |                                                                              | 122,9 |                                | 175,8 |
| 2016     |                                                                              | 119,8 |                                | 173,3 |
| 2015     |                                                                              | 119,4 |                                | 174,0 |
| 2014     |                                                                              | 118,9 |                                | 175,4 |
| 2013     |                                                                              | 118,8 |                                | 173,1 |
| 2012     |                                                                              | 122,6 |                                | 178,9 |
| 2011     |                                                                              | 135,3 |                                | 189,3 |
| 2010     |                                                                              | 147,0 |                                | 211,2 |
| 2009     | πραγματικά δεδομένα<br>από τη βάση δεδομένων<br>της Ευρωπαϊκής Επιτροπής ... | 158,7 |                                | 223,5 |
| 2008     |                                                                              | 161,2 |                                | 231,3 |
| 2007     |                                                                              | 155,4 |                                | 232,1 |
| 2006     |                                                                              | 148,4 |                                | 225,9 |

| Variable | Private FCE, at constant prices (OCPH) |                         | GNI, at constant prices (OVGN) |                      | $\Sigma(X_i \Psi_i)$ | $\Sigma(X_i)^2$ |
|----------|----------------------------------------|-------------------------|--------------------------------|----------------------|----------------------|-----------------|
| Country  | Greece                                 |                         | Greece                         |                      |                      |                 |
| 2026     |                                        | 147,8                   |                                | 206,0                | 30.446,8             | 42.436,0        |
| 2025     |                                        | 145,4                   |                                | 200,9                | 29.210,9             | 40.360,8        |
| 2024     |                                        | 143,0                   |                                | 195,7                | 27.985,1             | 38.298,5        |
| 2023     |                                        | 140,5                   |                                | 192,1                | 26.990,1             | 36.902,4        |
| 2022     |                                        | 138,1                   |                                | 191,2                | 26.404,7             | 36.557,4        |
| 2021     |                                        | 127,1                   |                                | 181,9                | 23.119,5             | 33.087,6        |
| 2020     |                                        | 121,0                   |                                | 166,9                | 20.194,9             | 27.855,6        |
| 2019     |                                        | 128,9                   |                                | 183,0                | 23.588,7             | 33.489,0        |
| 2018     |                                        | 125,7                   |                                | 178,2                | 22.399,7             | 31.755,2        |
| 2017     |                                        | 122,9                   |                                | 175,8                | 21.605,8             | 30.905,6        |
| 2016     |                                        | 119,8                   |                                | 173,3                | 20.761,3             | 30.032,9        |
| 2015     |                                        | 119,4                   |                                | 174,0                | 20.775,6             | 30.276,0        |
| 2014     |                                        | 118,9                   |                                | 175,4                | 20.855,1             | 30.765,2        |
| 2013     |                                        | 118,8                   |                                | 173,1                | 20.564,3             | 29.963,6        |
| 2012     |                                        | 122,6                   |                                | 178,9                | 21.933,1             | 32.005,2        |
| 2011     |                                        | 135,3                   |                                | 189,3                | 25.612,3             | 35.834,5        |
| 2010     |                                        | 147,0                   |                                | 211,2                | 31.046,4             | 44.605,4        |
| 2009     |                                        | 158,7                   |                                | 223,5                | 35.469,5             | 49.952,3        |
| 2008     |                                        | 161,2                   |                                | 231,3                | 37.285,6             | 53.499,7        |
| 2007     |                                        | 155,4                   |                                | 232,1                | 36.068,3             | 53.870,4        |
| 2006     |                                        | 148,4                   |                                | 225,9                | 33.523,6             | 51.030,8        |
|          |                                        | $\Sigma \Psi_i$ 2.845,9 |                                | $\Sigma X_i$ 4.059,7 | 555.841,2            | 793.484,2       |
|          |                                        | 119.165,0               |                                |                      | 135,5                | 193,3           |
|          |                                        | 182.004,3               |                                |                      |                      | 8,9459          |
|          |                                        | 0,6547                  |                                |                      |                      |                 |

$$\hat{\beta} = \frac{n \Sigma \Psi_i X_i - \Sigma X_i \Sigma \Psi_i}{n \Sigma X_i^2 - (\Sigma X_i)^2}$$

$$\hat{\alpha} = \bar{\Psi}_i - \hat{\beta} \bar{X}_i$$





## 2. OLS (παράδειγμα: αστικοποίηση & ηλεκτρισμός)

Ας δούμε ένα παράδειγμα με τη χρήση πραγματικών δεδομένων:

**Ψ:** πρόσβαση στον ηλεκτρισμό (**εξαρτημένη** μεταβλητή)

**X:** μη αστικός πληθυσμός (**ανεξάρτητη** μεταβλητή)

Πηγή άντλησης δεδομένων: World Bank (Παγκ. Τράπεζα)

## Η περίπτωση της Γκάνας (time-series analysis)

|      | $\Psi$ (εξαρτημένη μεταβλητή):<br>Πρόσβαση στον<br>ηλεκτρισμό (%) | $X$ (ανεξάρτητη μεταβλητή):<br>Μη αστικός<br>πληθυσμός (%) |
|------|-------------------------------------------------------------------|------------------------------------------------------------|
| 1997 | $\Psi_1 = 36,6$                                                   | $X_1 = 58,3$                                               |
| 1998 | $\Psi_2 = 42,6$                                                   | $X_2 = 57,6$                                               |
| 1999 | $\Psi_3 = 40,7$                                                   | $X_3 = 56,8$                                               |
| 2000 | $\Psi_4 = 43,7$                                                   | $X_4 = 56,1$                                               |
| 2001 | $\Psi_5 = 44,8$                                                   | $X_5 = 55,4$                                               |
| 2002 | $\Psi_6 = 46,8$                                                   | $X_6 = 54,7$                                               |
| ...  | ...                                                               | ...                                                        |
| 2022 | $\Psi_{26} = 85,1$                                                | $X_{26} = 41,4$                                            |

## Η περίπτωση της Γκάνας (time-series analysis)

$$\hat{\beta} = \frac{n \sum \Psi_i X_i - \sum X_i \sum \Psi_i}{n \sum X_i^2 - (\sum X_i)^2}$$

Στο παράδειγμα αυτό το αθροίσματα δίνονται έτοιμα ...

**n=26** (αριθμός παρατηρήσεων)

**$\sum(\Psi_i * X_i)$  = 78.221,14**

**$\sum X_i$  = 1.291,7**

**$\sum \Psi_i$  = 1.616**

**$\sum(X_i^2)$  = 64.841,94**

**$(\sum X_i)^2$  = 1.668.504,39**

συντελεστές  
παλινδρόμησης

$$\hat{\beta} = \frac{26 * (78.221,14) - (1.291,7 * 1.616)}{(26 * 64.841,94) - (1.668.504,39)} = -3,09$$

$$\hat{a} = 62,3 - (-3,09 * 49,7) = 215,6$$

$$\hat{\alpha} = \overline{\Psi_i} - \hat{\beta} \overline{X_i}$$

## Η περίπτωση της Γκάνας (time-series analysis)

Ερμηνεία του συντελεστή b:

Αν αυξηθεί κατά 1% ο μη αστικός πληθυσμός τότε η πρόσβαση του πληθυσμού στον ηλεκτρισμό θα μειωθεί κατά **3,09%** !!!

Έχει εδώ νόημα η ερμηνεία του συντελεστή a της παλινδρόμησης ;;;

συντελεστές  
παλινδρόμησης



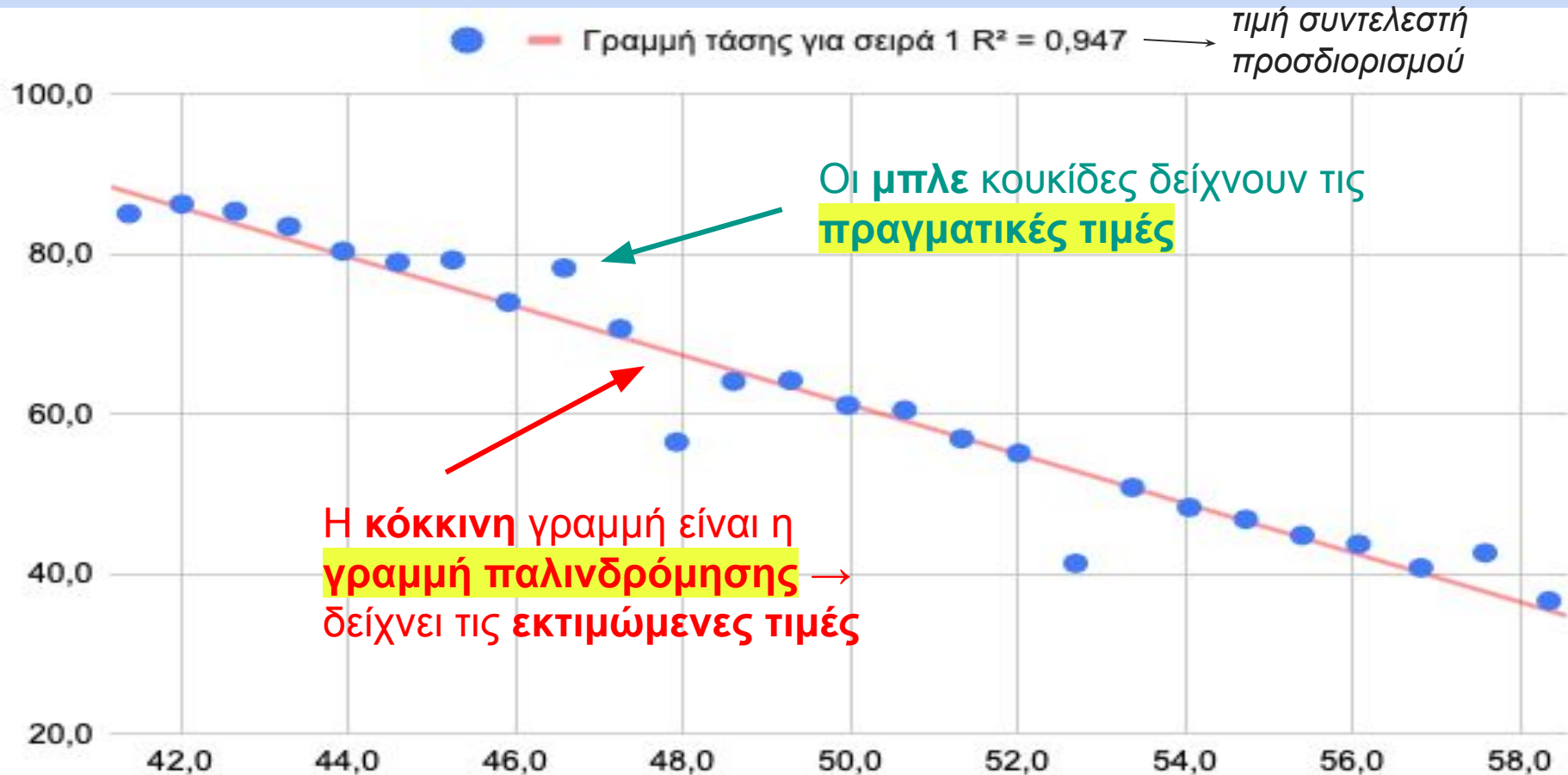
$$b = \frac{26 \cdot (78.221,14) - (1.291,7 \cdot 1.616)}{(26 \cdot 64.841,94) - (1.668.504,39)}$$

$$= \mathbf{-3,09}$$

$$a = 62,3 - (-3,09 \cdot 49,7) = \mathbf{215,6}$$



## Η περίπτωση της Γκάνας (time-series analysis)



### 3. OLS (παράδειγμα: Ανισότητα & Εκπαίδευση)

Ας δούμε ένα παράδειγμα με τη χρήση πραγματικών δεδομένων:

**Ψ:** *Gini Index* (**εξαρτημένη** μεταβλητή)

**X:** *γυναικείος εγγραμματισμός* (**ανεξάρτητη** μεταβλητή)

Πηγή άντλησης δεδομένων: World Bank (Παγκ. Τράπεζα)

## Διαστρωματικά Δεδομένα (cross-sectional data)

| Χώρα     | Έτος | Ψ: δείκτης οικονομικών ανισοτήτων | Χ: εγγράμματες γυναίκες (15-24 ετών) |
|----------|------|-----------------------------------|--------------------------------------|
| $i = 1$  | 2020 | <b>Ψ1</b>                         | <b>X1</b>                            |
| $i = 2$  | 2020 | <b>Ψ2</b>                         | <b>X2</b>                            |
| $i = 3$  | 2020 | <b>Ψ3</b>                         | <b>X3</b>                            |
| $i = 4$  | 2020 | <b>Ψ4</b>                         | <b>X4</b>                            |
| $i = 5$  | 2020 | <b>Ψ5</b>                         | <b>X5</b>                            |
| ...      | ...  | ...                               | ...                                  |
| $i = 18$ | 2020 | <b>Ψn</b>                         | <b>Xn</b>                            |

## Διαστρωματικά Δεδομένα (cross-sectional data)

$$\hat{\beta} = \frac{n \sum \Psi_i X_i - \sum X_i \sum \Psi_i}{n \sum X_i^2 - (\sum X_i)^2}$$

Στο παράδειγμα αυτό το αθροίσματα δίνονται έτοιμα ...

**n=18** (αριθμός παρατηρήσεων)

**$\sum \Psi_i X_i = 70.893,87$**

**$\sum X_i = 1.737,91$**

**$\sum \Psi_i = 734$**

**$\sum X_i^2 = 168.304,39$**

**$(\sum X_i)^2 = 3.020.326,62$**

συντελεστές  
παλινδρόμησης

$\hat{\beta}$

$$b = \frac{18 \cdot (70.893) - (1.737 \cdot 734)}{(18 \cdot 168.304) - (3.020.326)}$$

$$= \mathbf{0,051} \rightarrow \text{σχεδόν μηδενική επίδραση ...}$$

$\hat{a}$

$$a = \mathbf{35,88}$$

$$\hat{\alpha} = \overline{\Psi_i} - \hat{\beta} \overline{X_i}$$

## Συντελεστής Προσδιορισμού (R-squared)

Μας δείχνει το κομμάτι (%) της συνολικής μεταβλητότητας των τιμών της εξαρτημένης μεταβλητής  $\Psi$  που εξηγείται από τη μεταβλητότητα των τιμών της ανεξάρτητης μεταβλητής  $X$  !!!

Με πιο απλά λόγια, μας δείχνει την **ερμηνευτική ικανότητα του μοντέλου μας**.

Για παράδειγμα, μια τιμή του R-squared ίση με 0,9 σημαίνει πως το 90% της συνολικής μεταβλητότητας των τιμών της εξαρτημένης μεταβλητής εξηγείται από τη μεταβλητότητα των τιμών της ανεξάρτητης μεταβλητής (*υψηλή ερμηνευτική ικανότητα του μοντέλου μας*).

# Συντελεστής Προσδιορισμού (R-squared)

## Άθροισμα Τετραγώνων Συνολικό (SST)

$$\text{SST} = \sum_{i=1}^n (\Psi_i - \bar{\Psi})^2 \longrightarrow \text{δείχνει τη συνολική μεταβλητότητα των τιμών της εξαρτημένης μεταβλητής}$$

## Άθροισμα Τετραγώνων Παλινδρόμησης (SSR)

$$\text{SSR} = \sum_{i=1}^n (\widehat{\Psi}_i - \bar{\Psi})^2 \longrightarrow \text{δείχνει το κομμάτι που συνολικής μεταβλητότητας που εξηγείται από την παλινδρόμηση}$$

# Συντελεστής Προσδιορισμού (R-squared)

## Άθροισμα Τετραγώνων Σφάλματος (SSE)

$$\mathbf{SSE} = \sum_{i=1}^n (\Psi_i - \widehat{\Psi}_i)^2 \longrightarrow \text{δείχνει το κομμάτι που συνολικής μεταβλητότητας που εξηγείται από παράγοντες που δε λαμβάνονται υπόψη στο υπόδειγμα}$$

## Συντελεστής Προσδιορισμού (R-squared)

$$R^2 = \frac{\sum_{i=1}^n (\widehat{\Psi}_i - \bar{\Psi})^2}{\sum_{i=1}^n (\Psi_i - \bar{\Psi})^2} = 1 - \frac{\sum_{i=1}^n (\Psi_i - \widehat{\Psi}_i)^2}{\sum_{i=1}^n (\Psi_i - \bar{\Psi})^2}$$

Λαμβάνει τιμές μέσα στο διάστημα [0 , 1]

**Στο απλό γραμμικό υπόδειγμα ο συντελεστής προσδιορισμού ισούται με το τετράγωνο του συντελεστή συσχέτισης**



# 1. OLS (παράδειγμα: κατανάλωση & εισόδημα)

$$\hat{\beta} = \frac{n \sum \Psi_i X_i - \sum X_i \sum \Psi_i}{n \sum X_i^2 - (\sum X_i)^2} = 0,65$$

$$\hat{\alpha} = \overline{\Psi_i} - \hat{\beta} \overline{X_i} = 8,95$$

$$R^2 = \frac{\sum_{i=1}^n (\hat{\Psi}_i - \bar{\Psi})^2}{\sum_{i=1}^n (\Psi_i - \bar{\Psi})^2} = 1 - \frac{\sum_{i=1}^n (\Psi_i - \hat{\Psi}_i)^2}{\sum_{i=1}^n (\Psi_i - \bar{\Psi})^2} = 0,92 \text{ (92\%)} \quad \text{πολύ υψηλό}$$

## 2. OLS (παράδειγμα: αστικοποίηση & ηλεκτρισμός)

$$\hat{\beta} = \frac{n \sum \Psi_i X_i - \sum X_i \sum \Psi_i}{n \sum X_i^2 - (\sum X_i)^2} = -3,09$$

$$\hat{\alpha} = \overline{\Psi_i} - \hat{\beta} \overline{X_i} = 215,6$$

$$R^2 = \frac{\sum_{i=1}^n (\hat{\Psi}_i - \bar{\Psi})^2}{\sum_{i=1}^n (\Psi_i - \bar{\Psi})^2} = 1 - \frac{\sum_{i=1}^n (\Psi_i - \hat{\Psi}_i)^2}{\sum_{i=1}^n (\Psi_i - \bar{\Psi})^2} = 0,95 \text{ (95\%)} \quad \text{πολύ υψηλό}$$

### 3. OLS (παράδειγμα: Ανισότητα & Εκπαίδευση)

$$\hat{\beta} = \frac{n \sum \Psi_i X_i - \sum X_i \sum \Psi_i}{n \sum X_i^2 - (\sum X_i)^2} = 0,051$$

$$\hat{\alpha} = \bar{\Psi}_i - \hat{\beta} \bar{X}_i = 35,88$$

$$R^2 = \frac{\sum_{i=1}^n (\hat{\Psi}_i - \bar{\Psi})^2}{\sum_{i=1}^n (\Psi_i - \bar{\Psi})^2} = 1 - \frac{\sum_{i=1}^n (\Psi_i - \hat{\Psi}_i)^2}{\sum_{i=1}^n (\Psi_i - \bar{\Psi})^2} = 0,002 \text{ (0,2\%)} \quad \text{πάρα πολύ χαμηλό}$$