

## **ΧΡΟΝΙΚΕΣ ΣΕΙΡΕΣ**

**7ο Μάθημα: Απλή παλινδρόμηση  
(ΕΠΑΝΑΛΗΨΗ)**

**Παπάνα Αγγελική**

**Μεταδιδακτορική ερευνήτρια, ΑΠΘ & ΠΑΜΑΚ**

**E-mail: [angeliki.papana@gmail.com](mailto:angeliki.papana@gmail.com), [agrapana@auth.gr](mailto:agrapana@auth.gr)**

**Webpage: <http://users.auth.gr/agrapana>**

## Κλασσική γραμμική παλινδρόμηση

Μας ενδιαφέρει να μελετήσουμε την άγνωστη επίδραση της μεταβολής μιας μεταβλητής  $X$  (εξαρτημένη μεταβλητή), σε μια άλλη μεταβλητή, την  $Y$  (ανεξάρτητη μεταβλητή).

Το υπόδειγμα γραμμικής παλινδρόμησης ορίζει μια γραμμική σχέση μεταξύ της μεταβλητής  $X$  και της  $Y$ .

Η κλίση της γραμμής που συνδέει την  $X$  με την  $Y$  είναι η επίδραση μιας μεταβολής της  $X$  κατά μια μονάδα στην  $Y$ .

Το πρόβλημα είναι να εκτιμήσουμε αυτήν την κλίση, χρησιμοποιώντας ένα δείγμα από δεδομένα από τις δύο μεταβλητές,  $X$  και  $Y$ .

π.χ. Θέλουμε να ερευνήσουμε εμπειρικά την σχέση που υπάρχει ανάμεσα στις δαπάνες καταναλώσεως ( $Y$ ) και στο διαθέσιμο εισόδημα ( $X$ ) με βάση τις πληροφορίες από ένα δείγμα με  $T$  οικογένειες για μια δεδομένη χρονική περίοδο.

Έχουμε δηλαδή  $T$  ζεύγη από παρατηρήσεις  $(Y_t, X_t)$  όπου

$Y_t$ : δαπάνες καταναλώσεως της οικογένειας  $t$ ,  $t = 1, 2, \dots, T$

$X_t$ : διαθέσιμο εισόδημα της οικογένειας  $t$ ,  $t = 1, 2, \dots, T$

Αν η μαθηματική μορφή της συναρτησιακής σχέσης ανάμεσα στις δύο μεταβλητές είναι γραμμική τότε μπορούμε να γράψουμε:

$$Y_t = \beta_0 + \beta_1 X_t \quad (1)$$

Η σχέση (1) είναι προσδιοριστική και σημαίνει ότι όλες οι οικογένειες με το ίδιο διαθέσιμο εισόδημα έχουν τις ίδιες δαπάνες καταναλώσεως.

Στην πραγματικότητα η σχέση αυτή δεν μπορεί να ικανοποιείται από όλα τα ζεύγη παρατηρήσεων ( $Y_t, X_t$ ).

Οι διαφορές ή αποκλίσεις από την ευθεία που ορίζει η σχέση (1) μπορούν να ληφθούν υπόψη με την προσθήκη μιας τυχαίας μεταβλητής ή αλλιώς ενός διαταρακτικού όρου  $u_t$ , οπότε η προσδιοριστική σχέση γίνεται στοχαστική:

$$Y_t = \underbrace{\beta_0 + \beta_1 X_t}_{\text{συστηματικό μέρος}} + \underbrace{u_t}_{\text{μη συστηματικό μέρος}} \quad (2)$$

συστηματικό  
μέρος

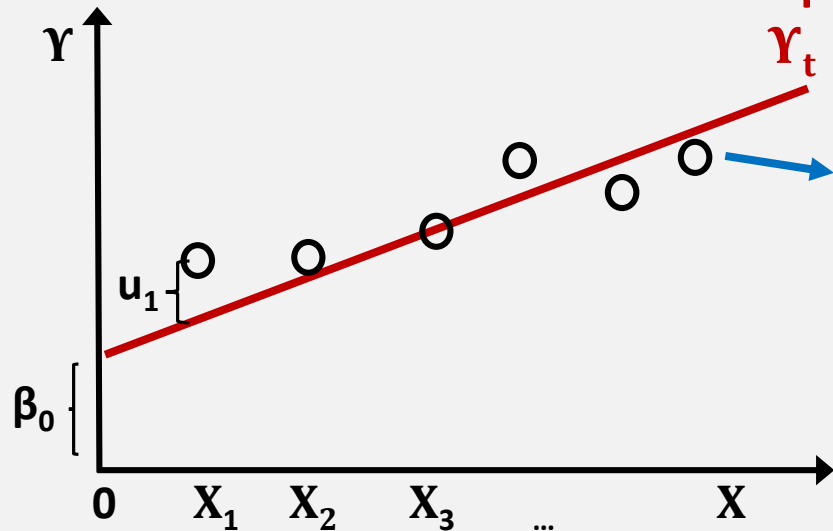
μη συστηματικό  
μέρος

## Στοχαστική συνάρτηση κατανάλωσης

$$Y_t = \beta_0 + \beta_1 X_t + u_t$$

Ευθείας παλινδρόμησης  
Προσδιοριστική σχέση

$$Y_t = \beta_0 + \beta_1 X_t$$



Πραγματικές τιμές  $Y_t$   
από δεδομένα

Προσθήκη διαταρακτικού  
όρου στην συνάρτηση  
παλινδρόμησης

- Ύπαρξη άγνωστων ή μη μετρήσιμων μεταβλητών που δεν συμπεριλαμβάνονται στο υπόδειγμα, σκόπιμη παράλειψη μεταβλητών μικρής σημασίας
- Αστάθμητη ανθρώπινη συμπεριφορά
- Σφάλματα μετρήσεως

## Ορολογία για το υπόδειγμα γραμμικής παλινδρόμησης με μια ερμηνευτική μεταβλητή

$$Y_t = \beta_0 + \beta_1 X_t + u_t$$

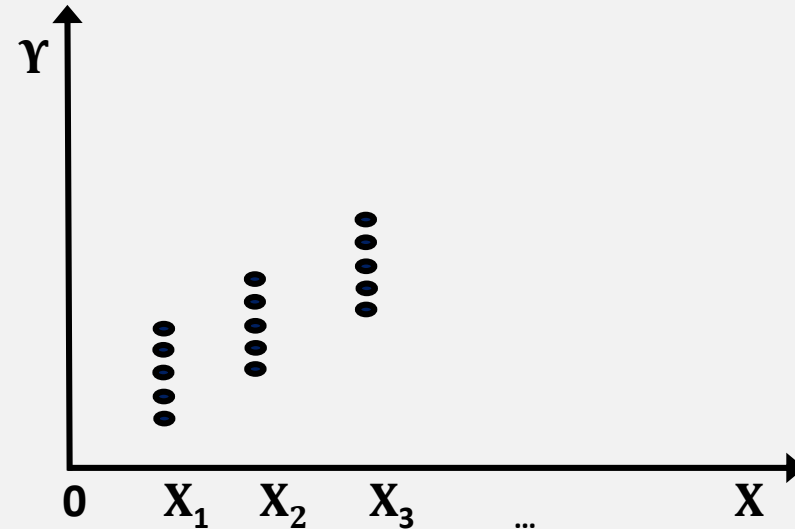
- Ο δείκτης  $t$  αναφέρεται στις παρατηρήσεις,  $t = 1, 2, \dots, T$ .
- $Y_t$  είναι η εξαρτημένη μεταβλητή
- $X_t$  είναι η ανεξάρτητη μεταβλητή
- $\beta_0 + \beta_1 X_t$  είναι η γραμμή παλινδρόμησης
- $\beta_0$  είναι ο σταθερός όρος της γραμμή παλινδρόμησης
- $\beta_1$  είναι η κλίση της γραμμή παλινδρόμησης
- $u_t$  είναι ο όρος σφάλματος ή τυχαίος όρος ή διαταρακτικός όρος

## Βασικές υποθέσεις του απλού γραμμικού υποδείγματος παλινδρόμησης

Η στοχαστική φύση της σχέσης

$$Y_t = \beta_0 + \beta_1 X_t + u_t \quad (2)$$

συνεπάγεται πως για κάθε τιμή της  $X$  δεν υπάρχει μια μόνο τιμή για την  $Y$ , αλλά μια κατανομή τιμών, που εξαρτάται από το  $u_t$ .



Για την εκτίμηση της (2) έχουμε ένα δείγμα με  $T$  ζεύγη παρατηρήσεων  $(Y_t, X_t)$ , αλλά δεν έχουμε παρατηρήσεις για τον διαταρακτικό όρο  $u_t$ , πράγμα που σημαίνει ότι θα κάνουμε κάποιες υποθέσεις σχετικά με την κατανομή της τυχαίας μεταβλητής  $u_t$ .

## Ολοκληρωμένη εξειδίκευση του γραμμικού στοχαστικού υποδείγματος

□  $Y_t = \beta_0 + \beta_1 X_t + u_t$

□  $u_t \sim (0, \sigma^2)$

α)  $u_t$  είναι τυχαία μεταβλητή

β)  $E u_t = 0$

γ)  $\text{Var}(u_t) = E u_t^2 = \sigma^2$

$u_t$ : ομοσκεδαστικός όρος, δηλ. έχει σταθερή διακύμανση

□  $\text{Cov}(u_t, u_s) = E u_t u_s = 0$  για  $t \neq s$  (ανεξαρτησία τυχαίων όρων)

□ Η μεταβλητή  $X$  δεν είναι στοχαστική. Οι τιμές της παραμένουν σταθερές και δεν είναι όλες ίσες μεταξύ τους.



## Η κατανομή της $Y$ και η Γραμμή Παλινδρομήσεως

Η μεταβλητή  $Y$  είναι συνάρτηση της τυχαίας μεταβλητής  $u_t$  και επομένως είναι και αυτή τυχαία μεταβλητή.

Η κατανομή της  $Y$  είναι κατανομή υπό συνθήκη, δεδομένης της τιμής της  $X$ .

Ισχύουν τα παρακάτω:

$EY_t = \beta_0 + \beta_1 X_t$ : γραμμή παλινδρομήσεως στον πληθυσμό

$\text{Var}(Y_t) = \sigma^2$

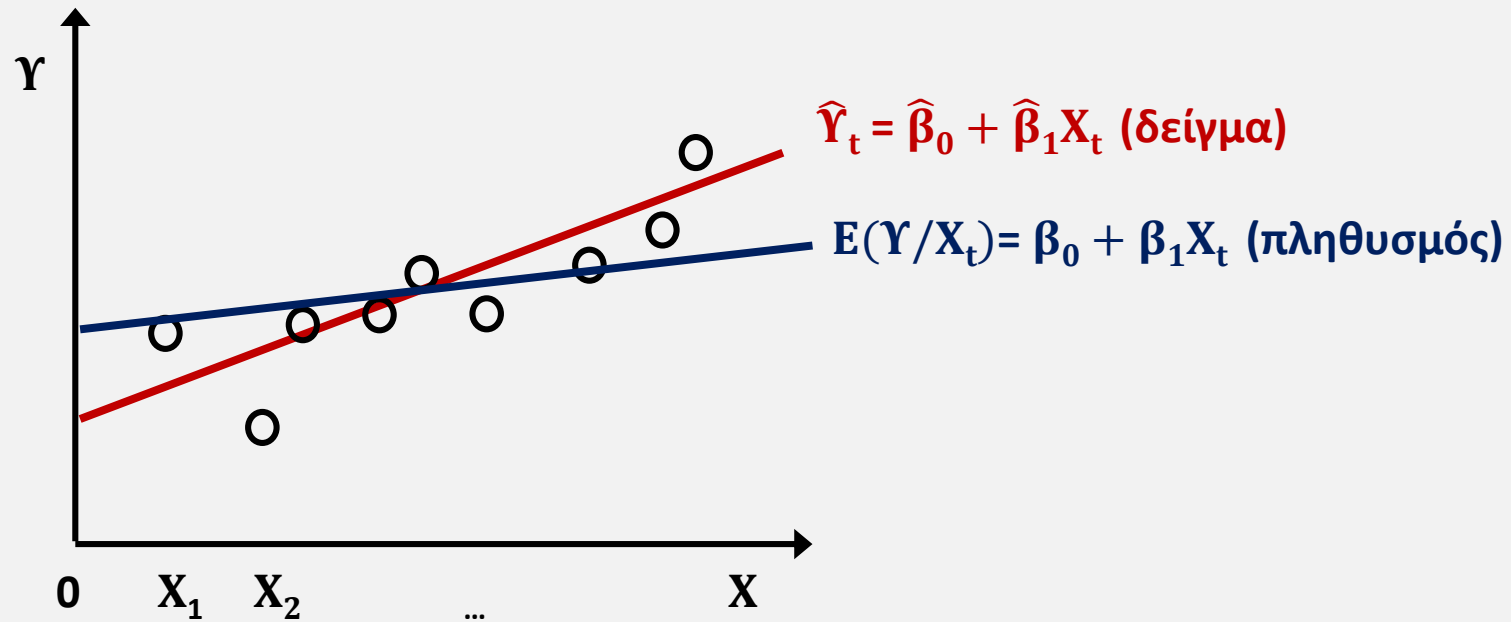
Οι συντελεστές  $\beta_0, \beta_1$  δεν είναι γνωστοί άρα πρέπει να εκτιμηθούν από το δείγμα των παρατηρήσεων των μεταβλητών  $X, Y$ .

Έστω  $\hat{\beta}_0, \hat{\beta}_1$  είναι οι εκτιμήσεις των συντελεστών  $\beta_0, \beta_1$ , αντίστοιχα.

Οπότε:

- $\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 X_t$ : γραμμή παλινδρομήσεως στο δείγμα
- $\hat{Y}_t$ : η τιμή της  $Y$  που υπολογίζουμε από την γραμμή παλινδρομήσεως του δείγματος
- $\hat{u}_t = Y_t - \hat{Y}_t$ : κατάλοιπο (residual) ή απόκλιση, η διαφορά μεταξύ των πραγματικών τιμών  $Y_t$  και των εκτιμώμενων τιμών  $\hat{Y}_t$

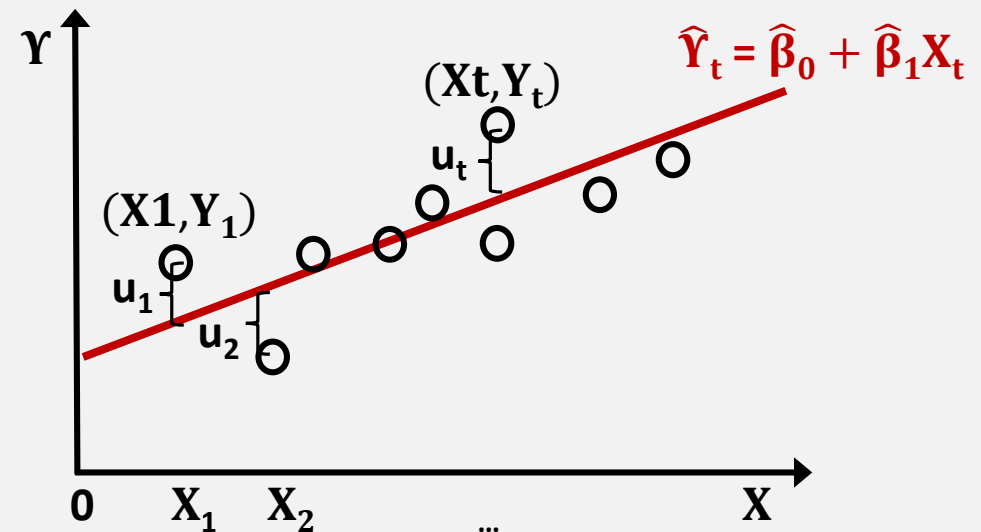
## Γραμμή παλινδρομήσεως πληθυσμού και δείγματος



## Μέθοδος ελαχίστων τετραγώνων (Ordinary Least Square - OLS)

Εκτίμηση των συντελεστών  $\beta_0, \beta_1$  ώστε να ελαχιστοποιήσουμε τα κατάλοιπα  $\hat{u}_t$  και συγκεκριμένα η OLS αποσκοπεί στην ελαχιστοποίηση του αθροίσματος των τετραγώνων των καταλοίπων (minimum sum of squared residuals):

$$\min \sum_{t=1}^T u_t^2 = \min \sum_{t=1}^T (Y_t - \hat{Y}_t)^2$$



## Εκτιμητές ελαχίστων τετραγώνων, προβλεφθείσες τιμές και κατάλοιπα

- Εκτιμητής ελαχίστων τετραγώνων της κλίσης  $\hat{\beta}_1$ :

$$\hat{\beta}_1 = \frac{\sum_{t=1}^T (X_t - \bar{X})(Y_t - \bar{Y})}{\sum_{t=1}^T (X_t - \bar{X})^2} = \frac{S_{XY}}{S_X^2} \quad \text{ή} \quad \hat{\beta}_1 = \frac{\sum XY - T\bar{X}\bar{Y}}{\sum X^2 - T\bar{X}^2}$$

- Εκτιμητής ελαχίστων τετραγώνων του σταθερού όρου  $\hat{\beta}_0$ :

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

- Προβλεφθείσες τιμές ελαχίστων τετραγώνων των  $\hat{Y}_t$ :

$$\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 X_t, \quad t = 1, \dots, T$$

- Προβλεφθείσες τιμές ελαχίστων τετραγώνων των καταλοίπων  $\hat{u}_t$ :

$$\hat{u}_t = Y_t - \hat{Y}_t, \quad t = 1, \dots, T$$

Οι εκτιμήσεις των παραπάνω γίνονται με βάση το δείγμα των  $T$  παρατηρήσεων των  $X_t$  και  $Y_t$ ,  $t = 1, \dots, T$ . Είναι εκτιμήσεις των πραγματικών παραμέτρων του πληθυσμού.

## Θεώρημα των Gauss-Markov

Εφόσον ισχύουν οι υποθέσεις του απλού γραμμικού υποδείγματος της παλινδρόμησης, οι εκτιμητές  $\hat{\beta}_0$ ,  $\hat{\beta}_1$  αποτελούν τους καλύτερους γραμμικούς αμερόληπτους εκτιμητές, δηλαδή:

- α) Είναι γραμμικές συναρτήσεις των παρατηρήσεων της εξαρτημένης μεταβλητής  $Y_t$ .
- β) Είναι αμερόληπτοι εκτιμητές.

Ένας εκτιμητής ονομάζεται αμερόληπτος όταν η αναμενόμενη τιμή του είναι ίση με την τιμή της υπό εκτίμηση άγνωστης τιμής του συντελεστή του πληθυσμού, δηλαδή  $E(\hat{\beta}_0) = \beta_0$  και  $E(\hat{\beta}_1) = \beta_1$ .

- γ) Είναι αποτελεσματικοί εκτιμητές, δηλαδή μεταξύ όλων των γραμμικών αμερόληπτων εκτιμητών έχουν την μικρότερη διακύμανση.

**Παράδειγμα.** Τα στοιχεία του Πίνακα αναφέρονται στην αξία (σε δισεκ. δραχμές) των εισαγωγών καταναλωτικών αγαθών ( $Y$ ) και στο διαθέσιμο εισόδημα ( $X$ ) για την Ελληνική Οικονομία για την περίοδο 1958-1973. Να βρεθεί η γραμμή παλινδρομήσεως του δείγματος, οι προβλεφθείσες τιμές  $\hat{Y}$  και τα κατάλοιπα  $\hat{u}$ .

Έτος	Y	X	Έτος	Y	X
1958	5,121	105,508	1966	8,625	182,420
1959	4,134	107,497	1967	9,204	192,895
1960	4,653	111,875	1968	9,647	204,164
1961	5,622	124,676	1969	10,167	221,908
1962	5,499	130,118	1970	9,961	240,471
1963	6,453	142,140	1971	10,580	267,849
1964	7,093	155,338	1972	10,658	289,450
1965	8,907	171,456	1973	13,139	318,550

Αρχικά υπολογίζουμε την  
δειγματική μέση τιμή των  
μεταβλητών  $X$  και  $Y$ ,  
αντίστοιχα:

$$\bar{X} = \frac{1}{T} \sum_{t=1}^T X_t$$

$$\bar{Y} = \frac{1}{T} \sum_{t=1}^T Y_t$$

Έπειτα υπολογίζουμε τους  
συντελεστές της γραμμής  
παλινδρόμησης του  
δείγματος

$$\hat{\beta}_1 = \frac{\sum XY - T\bar{X}\bar{Y}}{\sum X^2 - T\bar{X}^2} \text{ και}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}$$

Είναι:

$$T = \mathbf{16} \text{ (πλήθος παρατηρήσεων)}$$

$$\bar{X} = \frac{1}{16} \sum_{t=1}^{16} X_t =$$

$$= \frac{1}{16} (105,508 + 107,497 + \dots + 318,550) =$$

$$= \frac{1}{16} 2966,315 = \mathbf{185,394}$$

$$\bar{Y} = \frac{1}{16} \sum_{t=1}^{16} Y_t =$$

$$= \frac{1}{16} (5,121 + 4,134 + \dots + 13,139) =$$

$$= \frac{1}{16} 129,463 = \mathbf{8,091}$$



$$\begin{aligned}\sum XY &= \sum_{t=1}^{16} X_t Y_t = \\ &= (105,508 \times 5,121 + 107,497 \times 4,134 + \dots + 318,550 \times 13,139) = \mathbf{26541,949}\end{aligned}$$

$$\sum X^2 = \sum_{t=1}^{16} X_t^2 = (105,508^2 + 107,497^2 + \dots + 318,550^2) = \mathbf{617645,622}$$

**Οπότε:**

$$\hat{\beta}_1 = \frac{\sum XY - T\bar{X}\bar{Y}}{\sum X^2 - T\bar{X}^2} = \frac{26541,949 - 16 \times 185,394 \times 8,091}{617645,622 - 16 \times 185,394^2} = \mathbf{0,0375}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = 8,091 - 0,0375 \times 185,394 = \mathbf{1,136}$$

$\hat{Y}_t = 1,136 + 0,0375X_t$   
γραμμή παλινδρομήσεως στο  
δείγμα

Οπότε οι προβλεφθείσες τιμές  
ελαχίστων τετραγώνων των  $\hat{Y}_t$   
υπολογίζονται από την σχέση:

$$\hat{Y}_t = \hat{\beta}_0 + \hat{\beta}_1 X_t, t = 1, \dots, T$$

Οι προβλεφθείσες τιμές ελαχίστων  
τετραγώνων των καταλοίπων  $\hat{u}_t$   
υπολογίζονται από την σχέση:

$$\hat{u}_t = Y_t - \hat{Y}_t, t = 1, \dots, T$$

Προβλεφθείσες τιμές ελαχίστων  
τετραγώνων των  $\hat{Y}_t$ :

$$\begin{aligned} t = 1: \hat{Y}_1 &= 1,136 + 0,0375X_1 = \\ &= 1,136 + 0,0375 \times 105,508 = \\ &= \mathbf{5,094} \end{aligned}$$

$$\begin{aligned} t = 2: \hat{Y}_2 &= 1,136 + 0,0375X_2 = \\ &= 1,136 + 0,0375 \times 107,497 = \\ &= \mathbf{5,169} \quad \dots \end{aligned}$$

Προβλεφθείσες τιμές ελαχίστων  
τετραγώνων των καταλοίπων  $\hat{u}_t$ :

$$\begin{aligned} t = 1: \hat{u}_1 &= Y_1 - \hat{Y}_1 = \\ &= 5,121 - 4,134 = \mathbf{0,026} \end{aligned}$$

$$\begin{aligned} t = 2: \hat{u}_2 &= Y_2 - \hat{Y}_2 = \\ &= 4,134 - 5,169 = \mathbf{-1,035} \quad \dots \end{aligned}$$

Έτος	Y	X	$\hat{Y}$	$\hat{u}$
1958	5,121	105,508	5,094	0,026
1959	4,134	107,497	5,169	-1,035
1960	4,653	111,875	5,333	-0,679
1961	5,622	124,676	5,814	-0,191
1962	5,499	130,118	6,018	-0,518
1963	6,453	142,140	6,469	-0,016
1964	7,093	155,338	6,964	0,129
1965	8,907	171,456	7,568	1,338
1966	8,625	182,420	7,980	0,644
1967	9,204	192,895	8,373	0,831
1968	9,647	204,164	8,796	0,851
1969	10,167	221,908	9,461	0,705
1970	9,961	240,471	10,158	-0,197
1971	10,580	267,849	11,185	-0,605
1972	10,658	289,450	11,995	-1,337
1973	13,139	318,550	13,087	0,052

Η συνάρτηση εισαγωγών καταναλωτικών αγαθών είναι:

$$\hat{Y}_t = 1,136 + 0,0375X_t$$

Ο συντελεστής παλινδρομήσεως  $\beta_1$  παριστάνει τη μεταβολή στην προσδοκώμενη τιμή της εξαρτημένης μεταβλητής όταν η ερμηνευτική μεταβλητή ( $X$ ) μεταβάλλεται κατά μια μονάδα ή εναλλακτικά είναι η παράγωγος της  $E(Y_t)$  ως προς  $X_t$ :

$$\beta_1 = \frac{dE(Y_t)}{dX_t}$$

Στο συγκεκριμένο παράδειγμα,  $\hat{\beta}_1 = 0,0375$ , το οποίο παριστάνει την οριακή ροπή για εισαγωγές καταναλωτικών αγαθών, δηλαδή όταν το διαθέσιμο εισόδημα αυξάνεται κατά ένα δισεκ. δραχμές, οι εισαγωγές καταναλωτικών αγαθών θα αυξηθούν κατά 0,0375 δισεκ. δραχμές.

Με άλλα λόγια, το 3,75% της αύξησης του διαθέσιμου εισοδήματος απορροφάτε από τις εισαγωγές καταναλωτικών αγαθών.

## Στατιστική επαγωγή

Η γραμμή παλινδρόμησης του δείγματος είναι μια εκτίμηση της γραμμής παλινδρομήσεως του πληθυσμού και άρα υπόκειται σε σφάλματα, παρόλο που οι εκτιμητές ελαχίστων τετραγώνων ικανοποιούν τις επιθυμητές ιδιότητες (γραμμικοί, αμερόληπτοι, αποτελεσματικοί).

- Περιγράφει 'καλά' τα δεδομένα η γραμμή παλινδρόμησης του δείγματος;
- Η ανεξάρτητη/ερμηνευτική μεταβλητή ( $X$ ) ερμηνεύει μεγάλο ή μικρό τμήμα της μεταβολής της εξαρτημένης μεταβλητής ( $Y$ );
- Οι παρατηρήσεις (δεδομένα δείγματος) είναι κοντά συγκεντρωμένες γύρω από την γραμμή παλινδρόμησης του δείγματος ή είναι διάσπαρτες;

Θέλουμε να εξετάσουμε πόσο 'καλή' είναι η εκτίμηση που κάναμε, δηλαδή πόσο 'καλοί' είναι οι συντελεστές  $\hat{\beta}_0$ ,  $\hat{\beta}_1$  και τι κριτήρια μπορούμε να χρησιμοποιήσουμε για να αξιολογήσουμε τα αποτελέσματα της εκτιμήσεως.

**Συντελεστής προσδιορισμού  $R^2$  (regression  $R^2$  ή R squared)**

Έστω δείγμα  $X_t, Y_t, t = 1, \dots, T$ .

Η μεταβλητότητα της  $Y$  ορίζεται σε σχέση με τον μέσο του δείγματος:

Μεταβλητότητα της  $Y$ :  $\sum_{t=1}^T (Y_t - \bar{Y})^2$

Θέλουμε να εξετάσουμε πόση από την μεταβλητότητα που παρατηρείται στις τιμές της  $Y$  ερμηνεύεται από την παλινδρόμηση.

- $Y_t - \bar{Y}$  : απόκλιση της τιμής του δείγματος από τον μέσο
- $\hat{Y}_t - \bar{Y}$ : απόκλιση της προβλεφθείσας/εκτιμώμενης τιμής από τον μέσο
- $\hat{u}_t = Y_t - \hat{Y}_t$ : απόκλιση της τιμής του δείγματος από την γραμμή παλινδρόμησης

- Συνολική μεταβλητότητα της  $Y$ :

$$SST = \sum_{t=1}^T (Y_t - \bar{Y})^2 \text{ συνολικό άθροισμα τετραγώνων} \\ \text{(sum of squares, total)}$$

- Μεταβλητότητα της  $Y$ , που εξηγείται από την παλινδρόμηση:

$$SSR = \sum_{t=1}^T (\hat{Y}_t - \bar{Y})^2 \text{ ερμηνευόμενο άθροισμα τετραγώνων} \\ \text{(sum of squares, regression)}$$

- Μεταβλητότητα της  $Y$ , που μένει ανεξήγητη:

$$SSE = \sum_{t=1}^T (Y_t - \hat{Y}_t)^2 = \sum_{t=1}^T \hat{u}_t^2 \text{ άθροισμα τετραγώνων των καταλοίπων} \\ \text{(sum of squares, error)}$$

Ισχύει:  $SST = SSR + SSE$

Συνολική  
μεταβλητότητα της  $Y$

Εξηγείται από τις μεταβολές της  $X$

Οφείλεται στους τυχαίους παράγοντες

**Συντελεστής προσδιορισμού  $R^2$**  είναι η αναλογία ή το ποσοστό της διακύμανσης του δείγματος της  $Y$  που ερμηνεύεται / προβλέπεται από την  $X$ :

$$R^2 = \frac{SSR}{SST} \quad \text{ή}$$

$$R^2 = 1 - \frac{SSE}{SST} \quad \text{ή}$$

$$R^2 = \frac{\hat{\beta}_1 \sum xy}{\sum y^2}, \text{ όπου } x = X - \bar{X}, y = Y - \bar{Y}$$

Ο συντελεστής προσδιορισμού παίρνει τιμές από 0 έως 1:  $0 \leq R^2 \leq 1$ .

Όσο μεγαλύτερη η τιμή του  $R^2$ , τόσο καλύτερη είναι η προσαρμογή του υποδείγματος στα δεδομένα του δείγματος.



**Τυπικό σφάλμα της παλινδρόμησης  $s$  (standard error of the regression)** είναι ένας εκτιμητής της τυπικής απόκλισης του σφάλματος της παλινδρόμησης  $u_t$ .

Επειδή τα  $u_t$  δεν είναι γνωστά, η εκτίμηση προκύπτει με βάση τα  $\hat{u}_t$ .

Εκτίμηση της διακύμανσης

$$s^2 = \frac{1}{T-2} \sum_{t=1}^T \hat{u}_t^2 \quad \text{ή}$$

$$s^2 = \frac{\sum y^2 - \hat{\beta}_1 \sum xy}{T-2}, \quad \text{όπου } x = X - \bar{X}, y = Y - \bar{Y}$$

Όσο πιο μικρή τιμή προκύπτει για το  $s$  (τετραγωνική ρίζα της διακύμανσης), τόσο καλύτερη η προσαρμογή του υποδείγματος στα δεδομένα του δείγματος.

## Έλεγχος υποθέσεων των συντελεστών $\beta_0, \beta_1$

Είναι ο στατιστικός έλεγχος για τους συντελεστές  $\beta_0, \beta_1$  ώστε να διαπιστώσουμε αν πράγματι οι τιμές που εκτιμήσαμε είναι 'καλές'.

### Ανασκόπηση βασικών εννοιών

- Υπόθεση ονομάζεται η απόφαση που παίρνουμε για τα θέματα σχετικά με τους πληθυσμούς, βασιζόμενοι στις πληροφορίες που παίρνουμε από τα δείγματα των πληθυσμών.
- Έλεγχος υποθέσεων ή έλεγχος σημαντικότητας ή στατιστικός έλεγχος ονομάζεται η διαδικασία που χρησιμοποιείται ώστε να αποφασίσουμε αν θα δεχτούμε ή θα απορρίψουμε τις υποθέσεις που έχουμε κάνει.
- Μηδενική υπόθεση  $H_0$  ορίζεται ως η υπόθεση που κάνουμε αρχικά με σκοπό να την απορρίψουμε.
- Εναλλακτική υπόθεση  $H_1$  ορίζεται ως η ασυμβίβαστη υπόθεση σε σχέση με την μηδενική υπόθεση.

### 3. Έλεγχος υποθέσεων των συντελεστών $\hat{\beta}_0, \hat{\beta}_1$

- Η απόφαση αν θα γίνει δεκτή ή αν θα απορριφθεί η μηδενική υπόθεση  $H_0$  στηρίζεται σε ένα στατιστικό που ονομάζεται στατιστικό του τεστ, το οποίο υπολογίζεται από τα δεδομένα του δείγματος.
- Απορριπτική περιοχή  $R$  της  $H_0$  ονομάζεται η περιοχή στα σημεία της οποίας η  $H_0$  απορρίπτεται.
- Σφάλμα τύπου I: Απορρίπτω  $H_0$  όταν η  $H_0$  είναι αληθής, με πιθανότητα  $\alpha$
- Σφάλμα τύπου II: Δέχομαι την  $H_0$  όταν η  $H_1$  αληθής, με πιθανότητα  $\beta$

Οι υποθέσεις που θέλουμε να ελέγξουμε είναι:

Για το  $\beta_0$

$$H_0: \beta_0 = 0$$

$$H_1: \beta_0 \neq 0$$

Στατιστικό ελέγχων

$$t = \frac{\hat{\beta}_0 - \beta_0}{s_{\hat{\beta}_0}} = \frac{\hat{\beta}_0}{s_{\hat{\beta}_0}}$$

Για το  $\beta_1$

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

$$t = \frac{\hat{\beta}_1 - \beta_1}{s_{\hat{\beta}_1}} = \frac{\hat{\beta}_1}{s_{\hat{\beta}_1}}$$

Για επίπεδο σημαντικότητας  $\alpha$ , η κάθε μια μηδενική υπόθεση απορρίπτεται αν:

$$|t| \geq t_{T-2, \alpha/2}$$

Όπου  $t_{T-2, \alpha/2}$  είναι κριτική ή κρίσιμη τιμή που δίνεται από τους πίνακες της κατανομής Student.

## **Βιβλιογραφία**

**Χρήστου Κ. Γεώργιος (2007) Εισαγωγή στην Οικονομετρία, Τόμος 1, Εκδότης: Γ. ΔΑΡΔΑΝΟΣ - Κ. ΔΑΡΔΑΝΟΣ Ο.Ε.**

**Stock H. James, Watson W. Mark, επιμέλεια Πραγγίδης Ιωάννης - Χρυσόστομος (2017) Εισαγωγή στην Οικονομετρία, Εκδότης: Γ. ΔΑΡΔΑΝΟΣ - Κ. ΔΑΡΔΑΝΟΣ Ο.Ε.**

**Χρήστου Κ. Γεώργιος (2006) Εισαγωγή στην Οικονομετρία Ασκήσεις, Εκδόσεις Gutenberg.**

**Δριτσάκη Ν. Χάιδω, Δριτσάκη Ν. Μελίνα (2013) Εισαγωγή στην Οικονομετρία με τη Χρήση του Λογισμικού EViews, Κλειδάριθμος ΕΠΕ Εκδόσεις.**