

ΑΣΚΗΣΕΙΣ ΣΤΗΝ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

ΑΣΚΗΣΗ 1. Έστω το μοντέλο της απλής γραμμικής παλινδρόμησης $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, $i=1, \dots, n$, με τα σφάλματα $\varepsilon_i \sim N(0, \sigma^2)$ και ασυσχέτιστα ανά δύο. Ναδειχθεί ότι οι εκτιμητές ελαχίστων τετραγώνων των β_0 και β_1 του μοντέλου αυτού, συμπίπτουν με τους εκτιμητές μέγιστης πιθανοφάνειάς τους. Να προσδιορισθεί ο εκτιμητής μέγιστης πιθανοφάνειας της παραμέτρου σ^2 .

ΑΣΚΗΣΗ 2. Έστω το μοντέλο της απλής γραμμικής παλινδρόμησης $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, $i=1, \dots, n$, με $E(\varepsilon_i) = 0$, $Var(\varepsilon_i) = \sigma^2$ και τα ε_i , $i=1, \dots, n$, είναι ασυσχέτιστα. Αν $\hat{\beta}_0$ και $\hat{\beta}_1$ οι εκτιμητές ελαχίστων τετραγώνων των β_0 και β_1 , ναδειχθούν τα ακόλουθα:

$$\alpha) E(\hat{\beta}_0) = \beta_0, E(\hat{\beta}_1) = \beta_1, Var(\hat{\beta}_0) = \sigma^2 \left\{ \frac{1}{n} + \frac{\bar{X}^2}{\sum (X_i - \bar{X})^2} \right\} = \frac{\sigma^2 \sum X_i^2}{n \sum (X_i - \bar{X})^2}, Var(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2}.$$

$$\beta) \sum_{i=1}^n X_i e_i = 0, \text{ με } e_i = Y_i - \hat{Y}_i, i=1, \dots, n, \text{ τα υπόλοιπα.}$$

$$\gamma) Cov(\hat{\beta}_0, \hat{\beta}_1) = -\frac{\sigma^2 \bar{X}}{\sum_{i=1}^n (X_i - \bar{X})^2}.$$

ΑΣΚΗΣΗ 3. Τα παρακάτω δεδομένα αναφέρονται στην επίδραση της θερμοκρασίας στην απόδοση μιας χημικής αντίδρασης.

Θερμοκρασία	-5	-4	-3	-2	-1	0	1	2	3	4	5
Απόδοση	1	5	4	7	10	8	9	13	14	13	18

α) Αφού ορισθεί η εξαρτημένη και η ανεξάρτητη μεταβλητή να βρεθούν οι εκτιμητές ελαχίστων τετραγώνων των παραμέτρων του μοντέλου.

β) Να κατασκευαστεί ο πίνακας ANADIA και να γίνει ένα αρχικό τεστ για τον έλεγχο της υπόθεσης $H_0: \beta_1 = 0$.

γ) Κάνοντας τις κατάλληλες υποθέσεις

ι) Να ελεγχθεί η υπόθεση $H_0: \beta_1 = 0$ σε επίπεδο σημαντικότητας $\alpha = 0.05$.

ii) Να βρεθεί ένα 95% διάστημα εμπιστοσύνης για το β_1 .

ΑΣΚΗΣΗ 4. Για το μοντέλο της απλής γραμμικής παλινδρόμησης $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, $i=1, \dots, n$, ναδειχθεί ότι:

α) $r_{\hat{X}, Y}^2 = R^2$, με $r_{X, Y}$ το δειγματικό συντελεστή συσχέτισης του Pearson και R^2 το συντελεστή προσδιορισμού.

β) Αν $r_{Y, \hat{Y}}$ είναι ο δειγματικός συντελεστής συσχέτισης του Pearson μεταξύ της παρατηρούμενης και της εκτιμώμενης τιμής του Y , ναδειχθεί ότι $r_{Y, \hat{Y}}^2 = R^2$.

ΑΣΚΗΣΗ 5. Τα παρακάτω δεδομένα συγκεντρώθηκαν σε μια μελέτη που σκοπό είχε να εξετασθεί ποιοι παράγοντες επιδρούν στο βάρος ενός νεογέννητου μωρού.

Βάρος κατά τη γέννηση (gr)	1361	1588	1815	2087	2268	2404	3402	3629	3765	4083
Κοινωνικο-οικονομική κατάσταση	8	7	4	5	5	4	3	3	2	1
Σειρά γεννήσεως	4	3	4	3	2	2	2	1	1	1

i) Να ορισθεί η εξαρτημένη και οι ανεξάρτητες μεταβλητές.

ii) Να βρεθούν οι πίνακες $X'X$, $(X'X)^{-1}$, $X'Y$.

iii) Να βρεθούν οι εκτιμητές ελαχίστων τετραγώνων του μοντέλου $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \varepsilon_i$, και να δοθεί ο πίνακας διακυμάνσεων-συνδιακυμάνσεών τους. Να υπολογισθεί ο δειγματικός συντελεστής συσχέτισης μεταξύ των $\hat{\beta}_1$ και $\hat{\beta}_2$.

iv) Να δοθεί ο πίνακας ANADIA και να βρεθεί ένας αμερόληπτος εκτιμητής για τη διακύμανση σ^2 των σφαλμάτων του μοντέλου.

v) Υποθέτοντας ότι τα σφάλματα ακολουθούν κανονική κατανομή να ελεγχθούν οι υποθέσεις,

α) $\beta_1 = \beta_2 = 0$, β) $\beta_0 = 0$, γ) $\beta_1 = 0$, δ) $\beta_2 = 0$.

vi) Να δοθεί η φυσική ερμηνεία του εκτιμώμενου μοντέλου.

ΑΣΚΗΣΗ 6. Το γραφείο υποδοχής ασθενών ενός νοσοκομείου, θέλοντας να εξετάσει ποιοι παράγοντες μπορεί να επιδρούν στη διάρκεια παραμονής των ασθενών, που υποβάλλονται σε κάποιο είδος εγχείρησης στο νοσοκομείο, συγκέντρωσε τα παρακάτω δεδομένα για 20 τέτοιους ασθενείς.

A/A	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Ημέρες παραμονής Μετά την εγχείρηση	6	6	11	9	16	16	4	8	11	13	13	9	17	17	12
Αριθ. Ιατρικών Προβλημάτων	1	2	2	1	3	1	1	3	2	3	1	1	3	2	4
Ημέρες Παραμονής Πριν την εγχείρηση	1	1	2	3	3	5	1	1	2	2	4	2	3	4	1
	16	17	18	19	20										
	6	5	12	8	9										
	1	1	3	1	2										
	1	1	2	2	2										

Κάνοντας τις κατάλληλες υποθέσεις να αναλυθούν τα δεδομένα και να διατυπωθούν τα συμπεράσματα.

ΑΣΚΗΣΗ 7. Έστω Y_1 και Y_2 δύο ανεξάρτητες τυχαίες μεταβλητές με μέσες τιμές θ και 2θ αντίστοιχα, με θ πραγματικό αριθμό. Χρησιμοποιώντας το μοντέλο της γραμμικής παλινδρόμησης να προσδιορισθεί ο εκτιμητής ελαχίστων τετραγώνων της παραμέτρου θ και να υπολογισθεί το άθροισμα τετραγώνων των υπολοίπων.

ΑΣΚΗΣΗ 8. Για το μοντέλο της πολλαπλής γραμμικής παλινδρόμησης ναδειχθεί ότι

$$(Y - X\beta)'(Y - X\beta) = (Y - X\hat{\beta})'(Y - X\hat{\beta}) + (\hat{\beta} - \beta)'X'X(\hat{\beta} - \beta).$$