

ΕΞΕΤΑΣΕΙΣ**Μάθημα : ΑΝΑΛΥΣΗ ΠΑΛΙΝΔΡΟΜΗΣΗΣ****ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ****ΦΛΕΒΑΡΗΣ 2006**

***** Διάρκεια Εξέτασης : 2.30 ώρες *****

ΖΗΤΗΜΑ 1

Δίνεται ότι μέσω των σχέσεων $w_{ij} = \frac{x_{ij} - \bar{x}_j}{\sqrt{S_{jj}}}$ και $y_i^* = \frac{y_i - \bar{y}}{\sqrt{S_{yy}}}$, $S_{jj} = \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2$, $j=1, 2, 3$ και

$S_{yy} = \sum_{i=1}^n (y_i - \bar{y})^2$, το μοντέλο $Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$, $X_1 = 1$, γράφεται στη μορφή

$$Y^* = b_2 W_2 + b_3 W_3 + \varepsilon^*$$

i) Δείξτε ότι $(W'W)^{-1} = 1/(1-r^2) \begin{bmatrix} 1 & -r \\ -r & 1 \end{bmatrix}$, όπου r ο δειγματικός συντελεστής συσχέτισης μεταξύ των νέων μεταβλητών W_2 και W_3 .

ii) Δείξτε ότι το τυπικό σφάλμα των εκτιμητριών των νέων συντελεστών είναι $S_{y^*} (1-r^2)^{-1}$.

iii) Με βάση τα αποτελέσματα των (i) και (ii) τότε έχουμε ένδειξη πολυσυγγραμμικότητας και ποιες οι επιπτώσεις αυτής; (Βαθμ. 3.5)

ΖΗΤΗΜΑ 2

Εταιρία θέλει να εξετάσει το συνολικό χρόνο που αναλώνει υπάλληλός της για την παράδοση αναψυκτικών σε αυτόματη μηχανή ψυγείου σε σχέση με τον αριθμό κουτιών (X_1) και απόσταση που διανύει ο υπάλληλος με τα πόδια (X_2).

Y (min)	16.7	11.5	12.0	14.9	13.8	18.1	8.0	17.8	79.2	21.5	40.3	21.0	13.5
X_1	7	3	3	4	6	7	2	7	30	5	16	10	4
X_2 (m)	187	73	113	27	50	110	37	70	487	200	230	72	85

- Na προσδιοριστεί ο δείκτης προσδιορισμού R_{x_1, x_2}^2 και ο συντελεστής συσχέτισης r_{x_1, x_2} .
- Χρησιάζεται η μεταβλητή X_2 στο μοντέλο;
- Υπάρχει ένδειξη πολυσυγγραμμικότητας;
- Na κατασκευαστεί ένα 0.95 διάστημα εμπιστοσύνης της παραμέτρου β_1 στο μοντέλο $Y = \beta_0 + \beta_1 X_1$. Αποτελεί η παρατήρηση «9» σημείο επιρροής ($\beta_0 = 2.55$, $\beta_1 = 2.45$);

Η απόσταση Cook προσδιορίζεται ως $D_i = \frac{e_i^2 h_{ii}}{p S_{yy}^2 (1 - h_{ii})^2}$ (p ο αριθμός παραμέτρων στο μοντέλο,

$i=1, \dots, n$).

$$(\hat{\beta}_2 = 0.041, c_{22} = 0.0000268, SSR_{x_1, x_2} = 4191.2, \sum_{i=1}^n x_{1i}^2 = 1518, \sum_{i=1}^n x_{1i} x_{2i} = 417143, \sum_{i=1}^n y_i^2 = 10658.3,$$

$$\sum_{i=1}^n x_{1i} x_{2i} = 23959, \sum_{i=1}^n x_{1i} y_i = 3989.4, C_p(x_1, x_2) = 9.5, C_p(x_1) = 3, h_{99} = 0.782462, SSR_{x_1} = 4129)$$

(Βαθμ. 3.5)

ΖΗΤΗΜΑ 3

A) Δείγματα από δύο ορυχεία άνθρακα μιας περιοχής μετρήθηκαν ως προς την περιεκτικότητά τους σε θείο και έδωσαν τα παρακάτω αποτελέσματα :

Ορυχείο 1	1.51	1.92	1.08	2.04	2.14	1.76	1.17	
Ορυχείο 2	1.69	0.64	0.90	1.41	1.01	0.84	1.28	1.59

Na εξεταστεί αν υπάρχουν διαφοροποιήσεις μεταξύ των δύο ορυχείων σε ότι αφορά την περιεκτικότητά τους σε θείο μέσω ενός απλού γραμμικού μοντέλου.

B) Δώστε τους ορισμούς τριών τύπων υπολοίπων : συνήθη, τυποποιημένα και "deleted". Ποια είναι η χρήση τους;

(Βαθμ. 3.0)