

**ΕΞΕΤΑΣΕΙΣ**

**Μάθημα : ΑΝΑΛΥΣΗ ΠΑΛΙΝΔΡΟΜΗΣΗΣ**

**ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ**

**Εξετάσεις χειμερινού εξαμήνου 2007**

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

**ΖΗΤΗΜΑ 1**

(i) Υποθέτοντας ότι  $\epsilon \sim N_n(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ , να βρεθεί η εκτιμήτρια της  $\beta$  καθώς και της  $\sigma^2$  με τη μέθοδο μέγιστης πιθανοφάνειας στο γενικό γραμμικό μοντέλο  $y = X\beta + \epsilon$ , όπου  $X$  ο πίνακας σχεδιασμού.

(ii) Να βρεθεί η  $E(\hat{\beta})$  και η  $E(\hat{\sigma}^2)$ . Είναι η  $\hat{\beta}$  αμερόληπτη εκτιμήτρια της  $\beta$ ; Βρείτε μια αμερόληπτη εκτιμήτρια της  $\sigma^2$ . (iii) Δείξτε ότι για το απλό γραμμικό μοντέλο  $Y = \beta_0 + \beta_1 X + \epsilon$  η αμερόληπτη εκτιμήτρια της  $\sigma^2$  ανάγεται στην  $S_{Y_x}^2 = \frac{1}{n-2} \left\{ S_{yy} - \frac{S_{xy}^2}{S_{xx}} \right\} = \frac{1}{n-2} S_{yy} \{1 - r_{xy}^2\}$ . (Βαθμ. 3,5)

**ΖΗΤΗΜΑ 2**

Τα δεδομένα του παρακάτω πίνακα αφορούν την ταχύτητα  $X$  δύο τύπων ηλεκτρικών εργαλείων και την αντίστοιχη διάρκεια ζωής των  $Y$  (ώρες)

	Τύπος I						Τύπος II					
X	610	950	720	840	980	530	670	770	880	1000	760	590
Y	18.7	14.5	17.4	14.5	13.4	24.4	30.2	27.1	25.4	26.1	33.5	35.6

- (i) Να κατασκευαστεί ένα διάγραμμα διασποράς της  $Y$  έναντι της  $X$ .
- (ii) Να προσαρμοστεί το μοντέλο  $Y = \beta_0 + \beta_1 X + \epsilon$  στα δεδομένα και να βρεθεί ο συντελεστής προσδιορισμού  $R^2$ .
- (iii) Να γίνει ο έλεγχος  $H_0: \beta_1 = 0$  με εναλλακτική την  $H_1: \beta_1 \neq 0$ .
- (iv) Πιστεύετε ότι υπάρχει διαφοροποίηση μεταξύ των δύο τύπων εργαλείων ως προς τη διάρκεια ζωής τους με δεδομένη τη συμμετοχή της  $X$  στο μοντέλο; (δίνεται ότι  $\hat{\beta}_2 = -438.31$ ,  $c_{22} = 1.25$ . Ο νέος συντελεστής προσδιορισμού είναι  $R^2 = 0.743$  και  $SSE = 70466$ ). Τι συμπεραίνετε; (Βαθμ. 4,0)

**ΖΗΤΗΜΑ 3**

A) Δώστε μια σύντομη περιγραφή ΔΥΟ διαδικασιών με βήματα για την επιλογή ενός μοντέλου παλινδρόμησης. Θα καταλήξουν στο ίδιο μοντέλο;

B) Εξετάζεται η γραμμική παλινδρόμηση της  $Y$  σε σχέση με τις  $X_1, X_2, X_3, X_4$ . Με βάση τον παρακάτω πίνακα και για μέγεθος δείγματος  $n=13$  να βρεθεί το καταλληλότερο μοντέλο.  $\epsilon: x_1, \dots$

Πλήθος μεταβλητών στο μοντέλο	$R^2$	$C_p$ (Mallows)	$S_{Y_x}$	$X_1$	$X_2$	$X_3$	$X_4$
1	67.5	138.7	8.9639				X
1	66.6	142.5	9.0771		X		
2	97.9	2.7	2.4063	X	X		
2	97.2	5.5	2.7343	X			X
3	98.2	3.0	2.3087	X	X		X
3	98.2	3.0	2.3121	X	X	X	
4	98.2	5.0	2.4460	X	X	X	X

$(S_{Y_x} = (e'e / (n-k-1))^{1/2})$

(Βαθμ. 2,5)