

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

ΖΗΤΗΜΑ 1 (Βαθμ. 3.0)

A) (i) Δείξτε ότι στο απλό γραμμικό μοντέλο $E(Y_x) = \beta_0 + \beta_1 x$, ισχύει $R^2 = r_{xy}^2$, όπου r_{xy} ο δειγματικός συντελεστής συσχέτισης των Y και X και R^2 ο δείκτης προσδιορισμού του μοντέλου. (ii) Δείξτε ότι η ελεγχοσυνάρτηση $F \sim F_{1, (n-2)}$ για τον έλεγχο $H_0: \beta_1 = 0$ έναντι $H_1: \beta_1 \neq 0$ σε ένα απλό γραμμικό μοντέλο

γράφεται και ως $F = \frac{R^2}{(1-R^2)/(n-2)}$.

B) Εννέα μεταλλικοί σωλήνες με αντιδιαβρωτικό επίστρωμα τύπου A, και οκτώ τύπου B τέθηκαν σε διαβρωτικό περιβάλλον για το ίδιο χρονικό διάστημα. Τα παρακάτω δεδομένα αφορούν το βαθμό διάβρωσής τους.

Τύπου A	57	52	55	76	57	58	63	56	71
Τύπου B	44	42	46	69	58	54	40	46	

(i) Μέσω ενός απλού γραμμικού μοντέλου $E(Y_x) = \beta_0 + \beta_1 x$, να εξεταστεί αν υπάρχουν διαφοροποιήσεις μεταξύ των δύο σωλήνων σε ότι αφορά το βαθμό διάβρωσής τους. (ii) Ερμηνεύστε το συντελεστή $\hat{\beta}_1$.

ΖΗΤΗΜΑ 2 (Βαθμ. 2.5)

(i) Υποθέτοντας ότι $\varepsilon \sim N_n(0, \sigma^2 I_n)$, να βρεθεί η εκτιμήτρια της β καθώς και της σ^2 με τη μέθοδο ελαχίστων τετραγώνων στο γενικό γραμμικό μοντέλο $y = X\beta + \varepsilon$, όπου X ο πίνακας σχεδιασμού.

(ii) Να βρεθεί η $E(\hat{\beta})$ και ο πίνακας διασποράς-συνδιασποράς $V(\hat{\beta})$. Είναι η $\hat{\beta}$ αμερόληπτη εκτιμήτρια της β ;

(iii) Κάνοντας χρήση του πίνακα $H = X(X'X)^{-1}X'$, να βρεθεί η κατανομή της \hat{y} . Στη συνέχεια βρείτε τη διασπορά της εκτιμημένης παρατήρησης \hat{y}_i και τη συνδιασπορά μεταξύ των \hat{y}_i και \hat{y}_j , $i \neq j$.

ΖΗΤΗΜΑ 3 (Βαθμ. 2.5)

A) Δώστε τον ορισμό του συντελεστή C_p (Mallows). Πώς μας χρησιμεύει;

B) Εξετάζεται η γραμμική παλινδρόμηση της Y σε σχέση με τις X_1, X_2, X_3, X_4 . Σύμφωνα με τον παρακάτω πίνακα να βρεθεί το καταλληλότερο μοντέλο. Ενισχύστε την απάντησή σας με έναν έλεγχο F ($n=13$).

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

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

ΖΗΤΗΜΑ 4 (Βαθμ. 2.0)

A) Δώστε τον ορισμό ενός μοντέλου παλινδρόμησης Poisson. Δώστε ΔΥΟ στατιστικούς ελέγχους με τους οποίους εξετάζουμε την καταλληλότητα του μοντέλου αυτού.

B) Γιατρός θέλει να εξετάσει σε δείγμα μεγέθους $n=10$ ομάδες παρατηρήσεων, αν ο αριθμός θανάτων /ομάδα (Y) εξαρτάται από την ηλικία (x_1) και από το αν τα άτομα της κάθε ομάδας ήταν καπνιστές ($x_2=1$, αν ναι, και $x_2=0$, αν όχι).

(i) Επιβεβαιώστε ότι η μεταβλητή x_1 χρειάζεται στο μοντέλο $\ln \mu_x = \beta_0 + \beta_1 x_1$. (ii) Στη συνέχεια ελέγξτε αν και το κάπνισμα x_2 επιδρά σημαντικά στον αριθμό θανάτων/ομάδα.

(Δίνονται αποτελέσματα για τα μοντέλα:

(I) $\ln \mu_x = \beta_0 + \beta_1 x_1$, $\hat{\beta}_1 = 0.20617$, $se(\hat{\beta}_1) = 0.02687$,

(II) $\ln \mu_x = \beta_0 + \beta_1 x_1 + \beta_2 x_2$, $\hat{\beta}_2 = 1.1$, $se(\hat{\beta}_2) = 0.3$).