

Μάθημα : ΑΝΑΛΥΣΗ ΠΑΛΙΝΔΡΟΜΗΣΗΣ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΕΠΑΝΑΛΗΠΤΙΚΕΣ ΕΞΕΤΑΣΕΙΣ 2006-07 (Νοέμβριος 2007)

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

ΖΗΤΗΜΑ 1 Έστω το γενικό γραμμικό μοντέλο $Y=X\beta+\varepsilon$ (i) δείξτε ότι για το άθροισμα τετραγώνων λόγω σφάλματος ισχύει $SSE = y'(I - H)y$, όπου $H = X(X'X)^{-1}X'$ ο πίνακας προβολής.

(ii) Όταν στο γενικό γραμμικό μοντέλο $Y=X\beta+\varepsilon$ υπάρχει η σταθερά β_0 , τότε ισχύει $\sum_{i=1}^n y_i = \sum_{i=1}^n \hat{y}_i$. Να

γίνει η απόδειξη όταν υπάρχουν δύο επεξηγηματικές μεταβλητές στο μοντέλο.

(Υπόδ. $X'X\hat{\beta} = X'y$).

(iii) Θεωρώντας ότι ισχύει $\varepsilon_i \sim N(0, \sigma^2)$, $i=1, \dots, n$ ανεξάρτητα μεταξύ τους, δείξτε ότι

$\sum_{i=1}^n V(\hat{y}_i) = \sigma^2(k+1)$ όπου k ο αριθμός επεξηγηματικών μεταβλητών στο μοντέλο. (Βαθμ. 3.0)

ΖΗΤΗΜΑ 2

Πειραματιστής θέλει να εξετάσει τη σχέση μεταξύ της μεταβλητής Y (διάρκεια ζωής ενός εργαλείου σε ώρες) και της X (ταχύτητα), δύο τύπων εργαλείων A, B . Προσαρμόζοντας το απλό μοντέλο $Y=\beta_0+\beta_1X+\varepsilon$ στα δεδομένα του σχήματος, βρίσκει ότι η X συμβάλλει οριακά-σημαντικά στο μοντέλο (p -τιμή=0.058, $\hat{\beta}_1=-0.025$, $\sqrt{c_{11}}=0.0015$, $SSE=1282.08$, $R^2=0.17$)

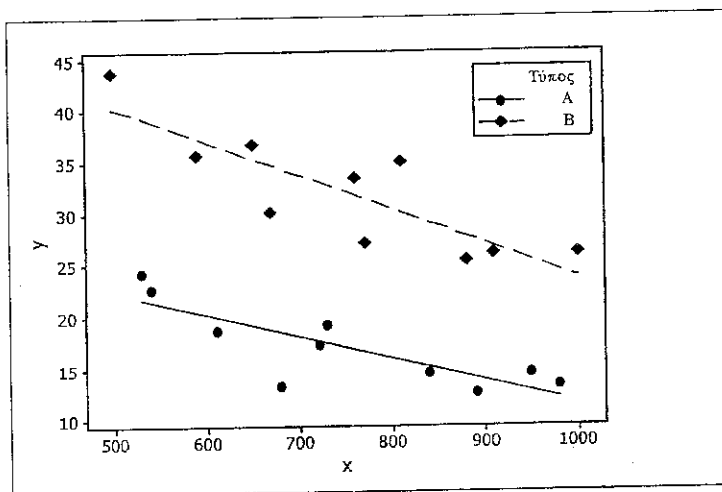
(i) Μέσω μιας ψευδομεταβλητής Z ελέγξτε αν η διάρκεια ζωής Y εξαρτάται από τον τύπο του εργαλείου. Σχολιάστε τα αποτελέσματα αυτά σε σχέση με τα πρώτα αποτελέσματα του πειραματιστή. Διαφοροποιείται η διάρκεια ζωής ως προς τους δύο τύπους εργαλείων;

($Z=1$ αν τύπου A , $Z=0$ αν τύπου B , $SSE=157.05$, $\hat{\beta}_1=-0.027$, $se(\hat{\beta}_1)=0.005$, p -τιμή του $\hat{\beta}_1 < 0.001$, $\hat{\beta}_2=-15.004$, $c_{22}=0.2$, $R^2=0.9$).

(ii) Να κατασκευαστεί ένα 0.95 - διάστημα εμπιστοσύνης για την παράμετρο β_1 του μοντέλου που προκύπτει από το ερώτημα (i).

(iii) Αν υποθέσουμε ότι η Z χρειάζεται στο μοντέλο, τότε πως εισάγεται η αλληλεπίδραση μεταξύ της X (διάρκεια ζωής) και των δύο τύπων εργαλείων στο μοντέλο παλινδρόμησης και

(iv) ελέγξτε την αλληλεπίδραση ως προς τη σημαντικότητά της ($SSE=140.98$, $R^2=0.91$).



(Βαθμ. 4.0)

ΖΗΤΗΜΑ 3

A) Η ποιότητα του κρασιού δοκιμάστηκε σε τρεις περιοχές οίνο-καλλιέργειας και έδωσαν τις ακόλουθες παρατηρήσεις :

Περιοχή 1	9.8	12.6	11.9	11.1	13.3	12.8	12.0		
Περιοχή 2	12.3	7.9	10.8	9.5	11.6	11.9	10.8	8.5	10.2
Περιοχή 3	13.6	14.4	16.1	15.5	13.8	12.7			

(i) Πως μπορούμε να εξετάσουμε αν υπάρχουν διαφοροποιήσεις μεταξύ των τριών περιοχών ως προς την ποιότητα του κρασιού μέσω ενός πολλαπλού γραμμικού μοντέλου;

(ii) Να γίνει αυτός ο έλεγχος ($SSR=56.5$).

B) Πως συγκρίνουμε δύο μοντέλα M_A και M_B της Poisson παλινδρόμησης όταν $M_A \subset M_B$;

(Βαθμ. 3.0)