Μάθημα : ΑΝΑΛΥΣΗ ΠΑΛΙΝΔΡΟΜΗΣΗΣ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
Εξετάσεις Φεβρουαρίου 2007-08

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

cov=0

ΖΗΤΗΜΑ 1 (i) Έστω το γενικό γραμμικό μοντέλο $\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$ με $\varepsilon_i \sim N(0, \sigma^2)$, $i = 1,\ldots,n$ ανεξάρτητα μεταξύ τους. Κάνοντας χρήση της ιδιότητας $E(\underset{\sim}{y}'A\underset{\sim}{y}) = tr(AV) + \underset{\sim}{\mu}'A\underset{\sim}{\mu}$, με $\underset{\sim}{\mu} = E(\underset{\sim}{y})$, $V = Var(\underset{\sim}{y})$ και $A = (I - H)$, $H = X(X'X)^{-1}X'$, δείξτε ότι $E(SSE) = (n-p)\sigma^2$ όπου SSE είναι το άθροισμα τετραγώνων λόγω σφάλματος και p ο αριθμός παραμέτρων στο μοντέλο.
(ii) Έστω το γενικό γραμμικό μοντέλο $E(\underset{\sim}{y}) = X_1\underset{\sim}{\beta}_1 + X_2\underset{\sim}{\beta}_2$ (ο πρώτος όρος περιλαμβάνει τη σταθερά β_0 και q_1 μεταβλητές, ο δεύτερος q_2 μεταβλητές). Γράψτε την ελεγχοσυνάρτηση και την κατανομή της για τον έλεγχο $H_0 : \underset{\sim}{\beta}_2 = 0$ με εναλλακτική $H_0 : \underset{\sim}{\beta}_2 \neq 0$.
(iii) Έστω το απλό γραμμικό μοντέλο $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ με $V(\varepsilon_i) = \sigma^2 x_i^2$, i=1, … , n. Πως θα μετασχηματίσουμε το μοντέλο έτσι ώστε να σταθεροποιηθεί η διασπορά του τυχαίου σφάλματος;
(Βαθμ. 3)

ΖΗΤΗΜΑ 2 Α) Δείξτε ότι όταν το πολλαπλό γραμμικό μοντέλο δεν έχει επεξηγηματικές μεταβλητές τότε $\hat{\beta}_0 = \bar{y}$ και κατά συνέπεια SST=SSE.
Β) Πειραματιστής θέλει να εξετάσει τη σχέση μεταξύ της μεταβλητής Υ και τριών επεξηγηματικών μεταβλητών X_1, X_2 και X_3.
(i) Αξιοποιώντας τον έλεγχο-F καθώς και τα κριτήρια R^2, C_p, να βρείτε το καλύτερο μοντέλο μέσω της τεχνικής της προς τα εμπρός επιλογής.
Στη συνέχεια θεωρώντας το απλό γραμμικό μοντέλο με τη σταθερά και τη μεταβλητή X_2 μόνο
(ii) να κατασκευαστεί ένα 0.99 - διάστημα εμπιστοσύνης για την παράμετρό της,
$[\hat{\beta} = 0.02955, \sum_{i=1}^{n}(x_i - \bar{x})^2 = 10452812]$ και (iii) με βάση το 11-οστό διαγώνιο στοιχείο του πίνακα **H** ελέγξτε αν η 11η παρατήρηση αποτελεί σημείο επιρροής.

$h_{ii} > \frac{2p}{n}$

Δίνονται: $h_{(11,11)}$=0.822, n =27, SST= 13792, $S_{Y_x} = \left(e'e/_{n-k-1}\right)^{1/2}$

Μεταβλητές	R^2	Mallows C_p	S_{Yx}	x 1	x 2	x 3
1	66.2	9.2	13.661		X 1η	
1	48.6	26.0	16.846			X
1	17.1	56.0	21.390	X		
2	74.7	3.2	12.069		X	X 2η
2	70.0	7.6	13.125	X	X	
2	57.2	19.8	15.688	X		X
3	75.9	4.0	12.029	X 3η	X	X

(Βαθμ. 4)

ΖΗΤΗΜΑ 3 Α) Μέσω ενός μοντέλου παλινδρόμησης εξετάζεται η περιεκτικότητα σε ασήμι 27 νομισμάτων τεσσάρων εποχών του Βυζαντίου. (i) Ελέγξτε αν υπάρχουν διαφοροποιήσεις μεταξύ των τεσσάρων εποχών. (ii) Με βάση τις τιμές των $\hat{\underset{\sim}{\beta}}$ και τη σημαντικότητά τους ερμηνεύστε τις επί μέρους διαφοροποιήσεις μεταξύ των εποχών και την επίδρασή τους στην περιεκτικότητα των νομισμάτων σε ασήμι. [Κωδικοποίηση: (X_1 X_2 X_3)=(1,0,0) αν εποχή 1, (0,1,0) αν εποχή 2, (0,0,1) αν εποχή 3 και (0,0,0) αν εποχή 4 (ομάδα αναφοράς)].

Δίνονται $\hat{b}$ $se(\hat{b})$

Χ- μεταβλητές	Coef	SE Coef	T	P
Σταθερά	5.6143	0.2616	21.46	0.000
x1	1.1302	0.3488	3.24	0.004
x2	2.6286	0.3699	7.11	0.000
x3	-0.7393	0.4338	-1.70	

R-Sq = 77.4%

Ανάλυση Διασποράς

Πηγή	ΑΤ
Παλινδρόμηση	37.748
Σφάλμα υπολοίπων	11.015
Σύνολο	48.763

Β) Διατυπώστε το μοντέλο της Poisson παλινδρόμησης. Τι είναι συνάρτηση σύνδεσης και ποια είναι αυτή για το μοντέλο της Poisson παλινδρόμησης;

(Βαθμ. 3)