

**ΕΞΕΤΑΣΗ ΑΝΑΛΥΣΗ ΠΑΛΙΝΔΡΟΜΗΣΗΣ, ΣΑΧΜ,
16/09/21**

1. (50 μονάδες) Θεωρήστε το γραμμικό μοντέλο :

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \epsilon_i, \quad i = 1, \dots, 100,$$

με ομοσκεδαστικά και ασυσχέτιστα σφάλματα που ακολουθούν κανονική κατανομή με διασπορά ίση με 2^2 .

- a. (5 μονάδες) Γράψτε την κατανομή του $\hat{\beta}$, του εκτιμητή ελαχίστων τετραγώνων του διανύσματος των συντελεστών, β .
- b. (15 μονάδες) Υπολογίστε την $P(s^2 > 4.75958)$ όπου s^2 είναι η αμερόληπτη εκτιμήτρια της διακύμανσης των σφαλμάτων.
- c. (10 μονάδες) Αποδείξτε πως $\hat{\beta}_1 - \hat{\beta}_3 + \hat{\beta}_4$ και s^2 είναι στατιστικά ανεξάρτητες τυχαίες μεταβλητές. Δείξτε ΟΛΑ τα βήματα.
- d. (20 μονάδες) Έστω $W_i = \frac{\hat{Y}_i}{s}$. Υποθέτοντας πως $\beta = 0$ υπολογίστε το:

$$E\left(\sum_{i=1}^{100} W_i^2\right)$$

2. (50 μονάδες) Χρησιμοποιήστε το output στο τέλος της εξέτασης. Η απόκριση είναι ο λογάριθμος της τιμής της σπιρομέτρησης και οι συμμεταβλητή είναι το ύψος (ht, σε ίντσες) , η ηλικία (age, σε έτη) και το φύλο (sex, 0 για τα κορίτσια κι 1 για τα αγόρια) . Τα δεδομένα αναφέρονται σε μηκαπνιστές (αγόρια και κορίτσια). Υποθέσετε πως τα σφάλματα είναι τυχαίο δείγμα από κανονική κατανομή. Στο πλήρες μοντέλο έχουν συμπεριληφθεί και αλληλεπιδράσεις του φύλου με την ηλικία και το ύψος.

- a. (10 μονάδες) Χρησιμοποιώντας το μερικό (partial) F έλεγχο, ελέγξτε την μηδενική υπόθεση που λέει πως το περιορισμένο μοντέλο με μοναδική συμμεταβλητή την ηλικία είναι επαρκές (σε σχέση με το πλήρες μοντέλο). Γράψτε αναλυτικά τις υποθέσεις, δώστε τη τιμή της σ.σ.ε., προσδιορίσετε το p -value, αναφέρετε την απόφαση και γράψτε το συμπέρασμα σε απλά ελληνικά.

- b. (5 μονάδες) Ελέγξτε την υπόθεση πως, βάσει του πλήρους μοντέλου, το φύλο δεν επηρεάζει την επίδραση της ηλικίας στην απόκριση. Γράψτε αναλυτικά τις υποθέσεις, δώστε τη τιμή της σ.σ.ε., προσδιορίσετε το p – $value$, αναφέρετε την απόφαση και γράψτε το συμπέρασμα σε απλά ελληνικά.
- c. (15 μονάδες) Υπολογίστε ένα 99% διάστημα πρόβλεψης για τη τιμή της απόκρισης ενός τυχαία επιλεγμένου 14-χρονου κοριτσιού με ύψος 62 ίντσες.
- d. (20 μονάδες) Βάσει του πλήρους μοντέλου **Γράψτε** το (προσεγγιστικό) 99% Δ.Ε. για την μέση διαφορά στη τιμή της σπιρομέτρησης μεταξύ 14-χρονων αγοριών με ύψος 64 ίντσες και 15-χρονων αγοριών με ύψος 66 ίντσες.

Οι πράξεις να γίνονται με ακρίβεια 6 δεκαδικών.

ΔΙΑΡΚΕΙΑ ΕΞΕΤΑΣΗΣ: 2 ΩΡΕΣ

Output για 3^η άσκηση

NON-SMOKERS

Regression Analysis: ln(FEV) versus age; ht; sex; age*sex; ht*sex

Regression Equation

$$\ln(\text{FEV}) = -1,888 + 0,02708 \text{ age} + 0,04129 \text{ ht} - 0,007 \text{ sex} - 0,00263 \text{ age*sex} + 0,00101 \text{ ht*sex}$$

Coefficients

Term	Coef	SE Coef	95% CI	T-Value	P-Value	VIF
Constant	-1,888	0,138	(-2,158; -1,617)	-13,71	0,000	
age	0,02708	0,00504	(0,01717; 0,03698)	5,37	0,000	5,51
ht	0,04129	0,00287	(0,03565; 0,04692)	14,40	0,000	7,63
sex	-0,007	0,175	(-0,350; 0,337)	-0,04	0,969	220,33
age*sex	-0,00263	0,00741	(-0,01718; 0,01191)	-0,36	0,722	43,51
ht*sex	0,00101	0,00374	(-0,00634; 0,00835)	0,27	0,788	390,07

Model Summary

S	R-sq	R-sq(adj)	PRESS	R-sq(pred)	AICc	BIC
0,142745	81,63%	81,48%	12,1552	81,21%	-613,54	-583,08

Analysis of Variance

Source	DF	Seq SS	Contribution	Adj SS	Seq MS	F-Value	P-Value
Regression	5	52,8025	81,63%	52,8025	10,5605	518,28	0,000
age	1	40,6713	62,88%	0,5872	40,6713	1996,03	0,000
ht	1	12,0111	18,57%	4,2259	12,0111	589,47	0,000
sex	1	0,1175	0,18%	0,0000	0,1175	5,77	0,017
age*sex	1	0,0011	0,00%	0,0026	0,0011	0,05	0,815
ht*sex	1	0,0015	0,00%	0,0015	0,0015	0,07	0,788
Error	583	11,8793	18,37%	11,8793	0,0204		
Lack-of-Fit	307	6,3938	9,89%	6,3938	0,0208	1,05	0,346
Pure Error	276	5,4854	8,48%	5,4854	0,0199		
Total	588	64,6818	100,00%				

Tests use the sequential sums of squares

$$(X^T X)^{-1} =$$

0,930937	0,0211478	-0,0188811	-0,93094	-0,0211478	0,0188811
0,021148	0,0012485	-0,0005510	-0,02115	-0,0012485	0,0005510
-0,018881	-0,0005510	0,0004033	0,01888	0,0005510	-0,0004033
-0,930937	-0,0211478	0,0188811	1,50046	0,0403152	-0,0311044
-0,021148	-0,0012485	0,0005510	0,04032	0,0026919	-0,0010898
0,018881	0,0005510	-0,0004033	-0,03110	-0,0010898	0,0006869