

## Regression Analysis: ln(FEV) versus age; ht; smoke; ht\*smoke; age\*smoke

### Regression Equation

$$\ln(\text{FEV}) = -1,9333 + 0,02469 \text{ age} + 0,04267 \text{ ht} - 0,264 \text{ smoke} + 0,00556 \text{ ht*smoke} - 0,01163 \text{ age*smoke}$$

### Coefficients

Term	Coef	SE Coef	95% CI	T-Value	P-Value	VIF
Constant	-1,9333	0,0833	(-2,0967; -1,7698)	-23,22	0,000	
age	0,02469	0,00372	(0,01739; 0,03200)	6,64	0,000	3,70
ht	0,04267	0,00180	(0,03914; 0,04620)	23,73	0,000	3,22
smoke	-0,264	0,388	(-1,026; 0,499)	-0,68	0,497	413,88
ht*smoke	0,00556	0,00609	(-0,00640; 0,01752)	0,91	0,361	443,95
age*smoke	-0,01163	0,00877	(-0,02885; 0,00559)	-1,33	0,185	39,87

### Model Summary

S	R-sq	R-sq(adj)	PRESS	R-sq(pred)	AICc	BIC
0,146044	80,94%	80,80%	14,0953	80,57%	-652,27	-621,06

### Analysis of Variance

Source	DF	Seq SS	Contribution	Adj SS	Seq MS	F-Value	P-Value
Regression	5	58,7048	80,94%	58,7048	11,7410	550,47	0,000
age	1	43,2101	59,58%	0,9391	43,2101	2025,89	0,000
ht	1	15,3263	21,13%	12,0111	15,3263	718,57	0,000
smoke	1	0,1236	0,17%	0,0098	0,1236	5,79	0,016
ht*smoke	1	0,0073	0,01%	0,0178	0,0073	0,34	0,558
age*smoke	1	0,0375	0,05%	0,0375	0,0375	1,76	0,185
Error	648	13,8211	19,06%	13,8211	0,0213		
Lack-of-Fit	281	6,1949	8,54%	6,1949	0,0220	1,06	0,297
Pure Error	367	7,6262	10,52%	7,6262	0,0208		
Total	653	72,5259	100,00%				

Tests use the sequential sums of squares

### Θέματα Πολλαπλής Επιλογής

- Το  $SSR(ht, smoke|age)$  ισούται με
  - 15,3263
  - 0,1236
  - 15,3263+0,1236**
  - 13,8211
- Η εκτίμηση της διασποράς των σφαλμάτων της παλινδρόμησης ισούται με
  - 13,8211
  - 0,0213**
  - 0,146044
  - 6,1949
- Βάσει του output, κι αναφερόμενοι στο πλήρες μοντέλο
  - Εκτιμάται πως το κάπνισμα μειώνει τη μέση τιμή της ln(FEV) κατά 0,1236 ανεξαρτήτως ύψους και ηλικίας
  - Εκτιμάται πως το κάπνισμα μειώνει τη μέση τιμή της ln(FEV) κατά 0,02469, ανεξαρτήτως ύψους και ηλικίας
  - Εκτιμάται πως για τους καπνιστές η αναμενόμενη διαφορά της ln(FEV) σε δύο πληθυσμούς με ένα έτος διαφοράς ηλικίας = 0,02469-0,01163.**
  - Εκτιμάται πως για τους καπνιστές η αναμενόμενη διαφορά της ln(FEV) σε δύο πληθυσμούς με ένα έτος διαφοράς ηλικίας = -0,264-0,01163

4. Το μέγεθος του δείγματος ισούται με
- 654**
  - 653
  - Άπειρο
  - 5
5. Σε ένα μοντέλο με μόνο ύψος και ηλικία, ο συντελεστής του ύψους ισούται με
- 0,02469
  - 0,04267
  - 15,3263
  - Τίποτα από το παραπάνω**

### Υπολογιστικές ερωτήσεις

1. Στο μοντέλο με μόνο ύψος και ηλικία, ελέγξτε αν ο συντελεστής της ηλικίας είναι στατιστικά σημαντικός σε  $\alpha=0,05$ . Δώστε τιμή σ.σ.ε., βαθμούς ελευθερίας, p-value, συμπέρασμα (σε απλά ελληνικά)
- Το ερώτημα δεν μπορεί να απαντηθεί βάσει των output καθότι στα Sequential Sum of Squares η ηλικία εισέρχεται 1<sup>η</sup> στη σειρά.**
2. Στο μοντέλο με μόνο ύψος και ηλικία, ελέγξτε αν ο συντελεστής του ύψους είναι στατιστικά σημαντικός σε  $\alpha=0,05$ . Δώστε τιμή σ.σ.ε., βαθμούς ελευθερίας, p-value, συμπέρασμα (σε απλά ελληνικά)
- F= 718,57. Κατανομή αναφοράς κάτω από τη μηδενική υπόθεση  $F_{1,654-3}$ . P-value=0<0,05. Απορρίπτω την μηδενική υπόθεση. Υπάρχει επαρκής ένδειξη σε ε.σ.  $\alpha=0,05$  πως το ύψος χρειάζεται να ληφθεί υπόψη στη εκτίμηση και πρόβλεψη της  $\ln(\text{FEV})$ , επιπλέον της ηλικίας.**
3. Στο πλήρες μοντέλο ελέγξτε αν το κάπνισμα επηρεάζει την επίδραση της ηλικίας στην μέση  $\ln(\text{FEV})$ . Δώστε **τιμή του σχετικού συντελεστή**, τιμή σ.σ.ε., βαθμούς ελευθερίας, p-value, συμπέρασμα (σε απλά ελληνικά)

Τελευταία γραμμή από τον πίνακα με τα coefficients.

Term	Coef	SE Coef	95% CI	T-Value	P-Value	VIF
age*smoke	-0,01163	0,00877	(-0,02885; 0,00559)	-1,33	0,185	39,87

Τιμή του σχετικού συντελεστή:  $\beta_5 = -0,01163$ ,  $t = -1,33$ ,  $df = 648$  (Κατανομή αναφοράς κάτω από τη μηδενική υπόθεση  $t_{648}$ ),  $p = 0,185 > 0,05$ . Δεν απορρίπτω την μηδενική υπόθεση και καταλήγω στο συμπέρασμα πως δεν έχω αρκετή ένδειξη σε ε.σ.  $\alpha = 0,05$  πως το κάπνισμα επηρεάζει την επίδραση της ηλικίας στην μέση  $\ln(\text{FEV})$ .

4. Στο πλήρες μοντέλο ελέγξτε τις εξής δύο μηδενικές υποθέσεις ταυτόχρονα α) το κάπνισμα δεν επηρεάζει την επίδραση της ηλικίας στην μέση  $\ln(\text{FEV})$ , β) το κάπνισμα δεν επηρεάζει την επίδραση του ύψους στην μέση  $\ln(\text{FEV})$ . Δώστε τιμή σ.σ.ε., βαθμούς ελευθερίας, p-value, συμπέρασμα (σε απλά ελληνικά)

$$SSR(\text{age} * \text{smoke}, \text{age} * \text{ht} | \text{age}, \text{ht}, \text{smoke}) = 0,0073 + 0,0375 = 0,0448$$

$$F = \frac{0,0448/2}{0,0213} = 1,05164$$

Κατανομή αναφοράς κάτω από τη μηδενική υπόθεση  $F_{2,648}$ . P-value>0,10 (βάσει πίνακα). Δεν απορρίπτω την μηδενική υπόθεση. Δεν έχω αρκετή ένδειξη σε ε.σ.  $\alpha=0,05$  πως το κάπνισμα επηρεάζει είτε την επίδραση της ηλικίας είτε του ύψους στην μέση  $\ln(\text{FEV})$ .