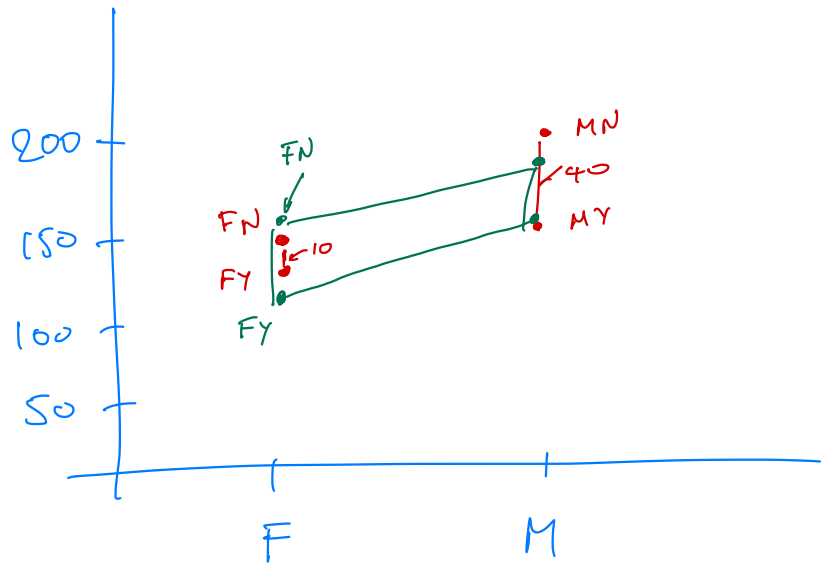
23-3-2023

dataframe phys'

	μ
FN	150.2
FY	142
MN	200
MY	160



$$Y = b_0 + b_1 X_M + b_2 X_Y$$

$$\hat{b}_0 = 156, \quad \hat{b}_1 = 34, \quad \hat{b}_2 = -22$$

Sex	Smoking	X_M	X_Y	Y	$\hat{\mu}$
Fem	No	0	0	$\mu_{FN} = b_0$	156
Fem	Yes	0	1	$\mu_{FY} = b_0 + b_2$	134
Male	No	1	0	$\mu_{MN} = b_0 + b_1$	190
Male	Yes	1	1	$\mu_{MY} = b_0 + b_1 + b_2$	168

Αλληλεπίδραση (Interaction)

Η επίδραση ενός παράγοντα στο $E(Y)$ εξαρτάται από τα επίπεδα των άλλων παραγόντων.

$$Y = b_0 + b_1 X_M + b_2 X_Y + b_3 X_M X_Y$$

main effects

interaction effects

main effects + interaction model

Κατ	X_M	X_Y	$Y = b_0 + b_1 X_M + b_2 X_Y + b_3 X_M X_Y$
ΓΝ	0	0	b_0
ΓΥ	0	1	$b_0 + b_2$
ΜΝ	1	0	$b_0 + b_1$
ΜΥ	1	1	$b_0 + b_1 + b_2 + b_3$

Επ. κατ. στις γυναίκες $\mu_{ΓΥ} - \mu_{ΓΝ} = b_2$

" " στους άνδρες $\mu_{ΜΥ} - \mu_{ΜΝ} = b_2 + b_3$

συνε. αλληλεπίδρασης

$$b_3 = (\mu_{ΜΥ} - \mu_{ΜΝ}) - (\mu_{ΓΥ} - \mu_{ΓΝ})$$

Επ. Φύλου στο μι κατ. $= \mu_{ΜΝ} - \mu_{ΓΝ} = b_1$

" " στους κατ. $= \mu_{ΜΥ} - \mu_{ΓΥ} = b_1 + b_3$ $\Rightarrow$

$$\Rightarrow b_3 = (\mu_{ΜΥ} - \mu_{ΓΥ}) - (\mu_{ΜΝ} - \mu_{ΓΝ})$$

Παραδ. 2 1) Color $\{W, B, R, G\}$ (ref: W) X_B, X_R, X_G

2) Quality $\{L, M, H\}$ (ref: L) X_M, X_H

Συνδυασμοί

μ_{WL}
μ_{WM}
⋮
μ_{GH}

$4 \times 3 = 12$

Main Effects Model

$$Y = b_0 + b_1 X_B + b_2 X_R + b_3 X_G + b_4 X_M + b_5 X_H$$

(6 παράμ.)

$n \geq 30$

$n \geq 60$

Επίπ. χρώματος (Blue)

$$\mu_{BL} - \mu_{WL} = \mu_{BM} - \mu_{WM} = \mu_{BH} - \mu_{WH}$$

όλα ίσα

Επίπεδα (Red) $\mu_{RL} - \mu_{WL} = \mu_{RM} - \mu_{WM} = \mu_{RH} - \mu_{WH}$

Όμοια (Green)

Αλληλεπιδράσεις Γινόμενα μεταξύ X

main effect Color | main effects Quality

$$Y = b_0 + b_1 X_B + b_2 X_R + b_3 X_G + b_4 X_M + b_5 X_H +$$

$$+ b_6 X_B X_M + b_7 X_B X_H + b_8 X_R X_M + b_9 X_R X_H + b_{10} X_G X_M + b_{11} X_G X_H$$

interactions

δεν επιτρέπονται $X_R^2 = X_R$ $(X_R = 0, 1)$ $+ b_{12} X_R^2$

Sample size $n \geq 60$

n $n \geq 120$ (συνήθως)

X_1 : Επίπεδα ref (1), 2, 3, 4 — binary X_{12}, X_{13}, X_{14}

X_2 : " ref (1), 2, 3, 4, $X_{22}, X_{23}, X_{24}, X_{25}$

X_3 : " ref (1), 2, 3, 4 X_{32}, X_{33}, X_{34}

Περιπτώσεις $4 \times 5 \times 4 = 80$ μ

$\mu_{111}, \mu_{112}, \ldots, \mu_{454}$

Όροι αλληλεπίδρασης:

Επιστροφή στο phys:

$Y = b_0 + b_1 X_M + b_2 X_Y + b_3 X_{MY}$

$\hat{b}_0 = 151$
$\hat{b}_1 = 50$
$\hat{b}_2 = -8.4$
$\hat{b}_3 = -31.4$

Καιρ	X_M	X_Y	$Y = b_0 + b_1 X_M + b_2 X_Y + b_3 X_M X_Y$
FN	0	0	$b_0 = 151$
FY	0	1	$b_0 + b_2 = 141.2$
MN	1	0	$b_0 + b_1 = 201$
MY	1	1	$b_0 + b_1 + b_2 + b_3 \simeq 160$

dataframe

	μ
FN	150.2
FY	142
MN	200
MY	160

Analysis of Covariance (ANACOVA)

Παράγοντες { κατηγορικές μεταβλητές + ποσοτικές μεταβλητές } → πιό χρήσιμη περίπτωση

Παράδειγμα : apart2.Rdata

Y = price

S = size (scale)

City : Boston (ref)

$X_{CH} = 1$ (Chicago)

$X_{NY} = 1$ (NY)

Main Effects Model

$$Y = b_0 + b_1 \cdot S + b_2 \cdot X_{CH} + b_3 X_{NY}$$

$b_1 \cdot S$: effect size

$b_2 \cdot X_{CH} + b_3 X_{NY}$: effects city

4 παράμετροι

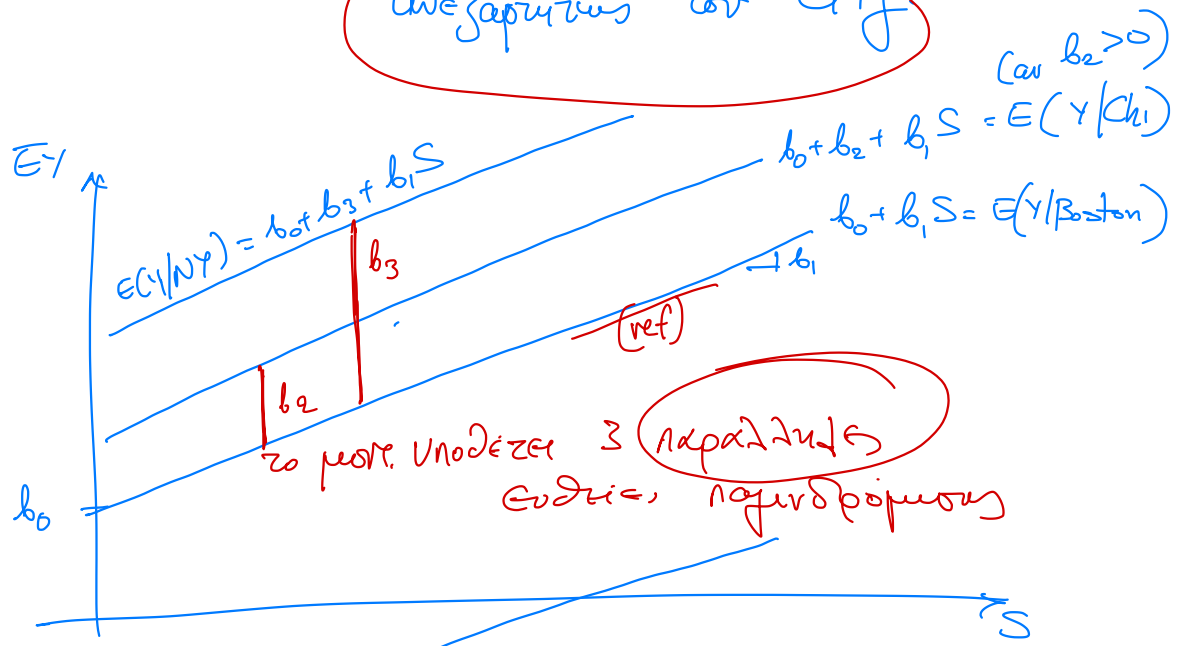
City	X_{CH}	X_{NY}	$Y = b_0 + b_1 S + b_2 X_{CH} + b_3 X_{NY}$
Boston	0	0	$EY = b_0 + b_1 S$
Chi	1	0	$EY = b_0 + b_2 + b_1 S$
NY	0	1	$EY = b_0 + b_3 + b_1 S$

$b_0 = X$

$b_2 = E(Y|Chi) - E(Y|Boston)$ ανεξαρτήτως του S

$b_3 = E(Y|NY) - E(Y|Boston)$ ανεξ. του S

b_1 = κλίση του S = μεταβολή του ΕΥ για αύξηση του S κατά 1, ανεξαρτήτως του City



(αν $b_2 > 0$)

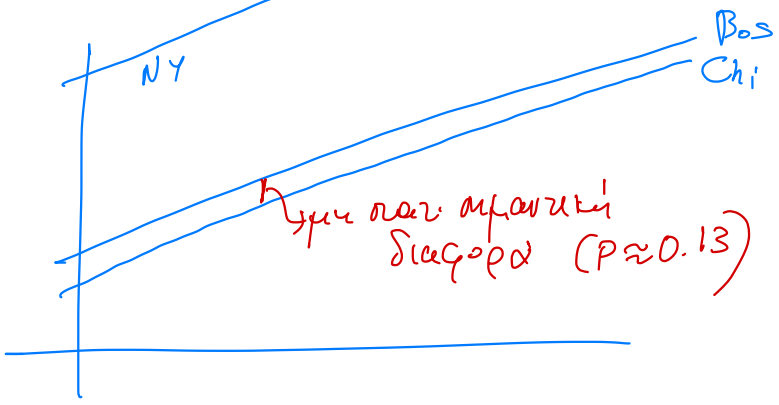
το μοντ. υποθέτει 3 παράλληλες ευθείες παλινδρόμησης

$\hat{b}_0 = 111$

$\hat{b}_1 = 0.12$

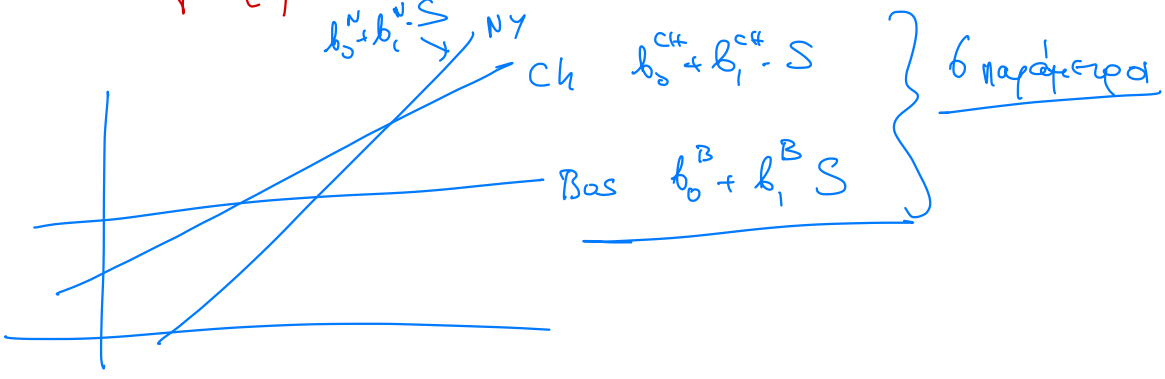
$\hat{b}_2 = -7.5$

$\hat{b}_3 = 61.5$



μικρή παρ. σημαντική διαφορά ($p \approx 0.13$)

Στη γενική περίπτωση έχουμε 3 ευθείες παλινδρόμησης Y ~ Size



Interaction Model

$$Y = b_0 + \underbrace{b_1 S}_{\text{effect size}} + \underbrace{b_2 X_{CH} + b_3 X_{NY}}_{\text{effect city}} + \underbrace{b_4 X_{CH} \cdot S + b_5 X_{NY} \cdot S}_{\text{interactions}}$$

(effect size + effect city = main)

	X_{CH}	X_{NY}	$Y = b_0 + b_1 S + b_2 X_{CH} + b_3 X_{NY} + b_4 X_{CH} S + b_5 X_{NY} S$
Bos	0	0	$Y = b_0 + b_1 S$
Chi	1	0	$Y = (b_0 + b_2) + (b_1 + b_4) \cdot S$
NY	0	1	$Y = (b_0 + b_3) + (b_1 + b_5) \cdot S$

b_4: Διαφορά ανάμεσα στη μεταβολή μέση ιχμής ανά μέτρο του CHI ως προς την αντίστοιχη " " ανά μέτρο στη Boston

---

Έλεγχος για σημαντικότητα αλληλεπιδράσεων

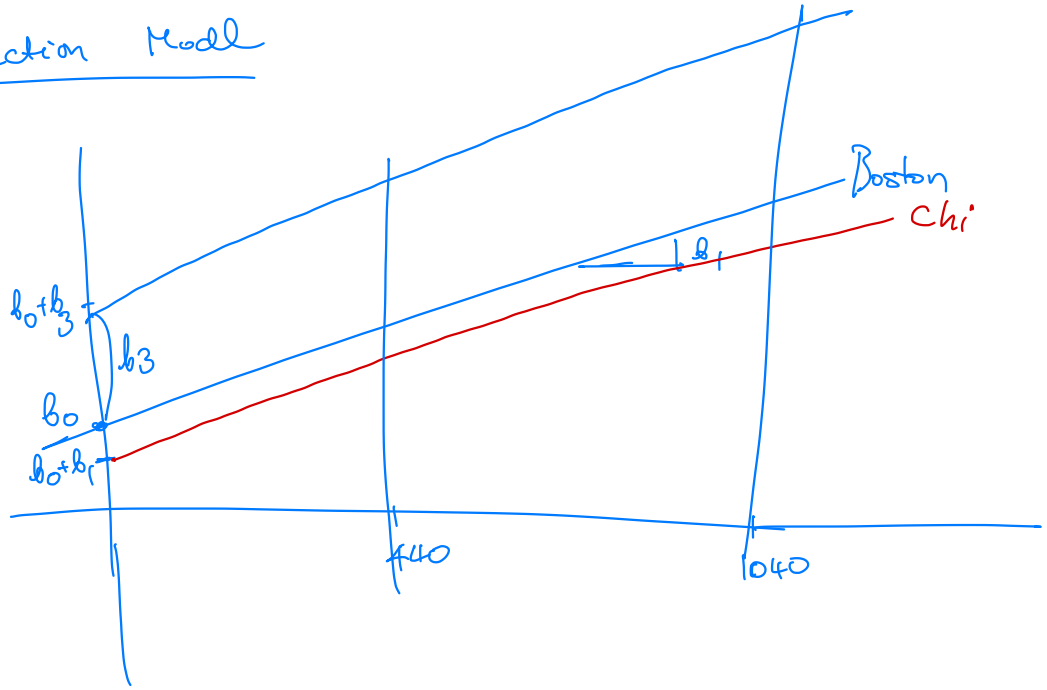
partial F-test: $H_0: b_4 = b_5 = 0$, H_1 τουλαχ. ένα $\neq 0$.

anova(main model, interaction model)

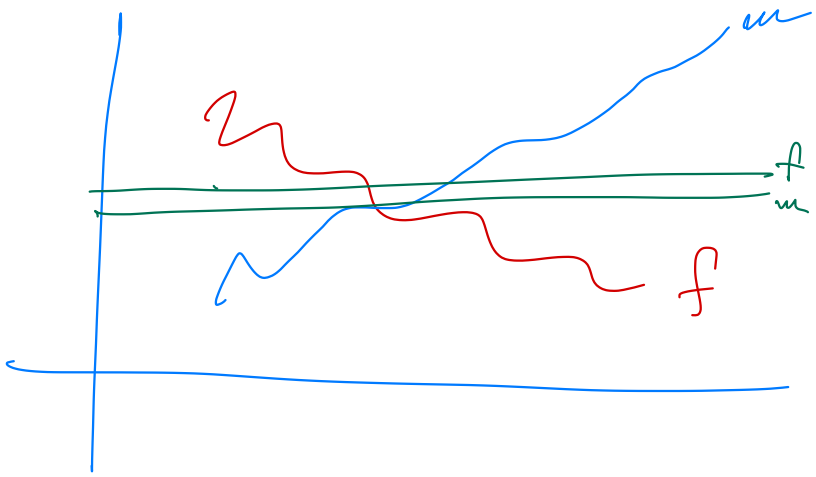
$p = 0.62$

Επαληθεύεται με plots

Interaction Model



Sauge dataset



$Y = b_0 + b_1 \text{age} + b_2 \text{sex}$

$Y = b_0 + b_1 \text{age} + b_2 X_M + b_3 X_M \cdot \text{age}$