

Παλιόδρομος - Ανάγνωση Διαστημάτων

16/2/2023

Υλικό

- 1) Σημειώσεις στο eclass
 - 2) " μάθημα να αναρτηθεί
 - 3) Βιβλιογραφία στο eclass
-

Software

Stata {Labs}
R (μάθημα)
οχι από Defete

Εργασία

5 Labs

Τελική Εξέταση

Παλινδρόμηση - Ανάλυση Διασποράς (Γραμμικά Μοντέλα)

Στατιστικά Μοντέλα - Μοντεφορνίον

Μοντέλα (Υπόδειγμα)

Αναπεράσταση πραγματικού φαινομένου

- Φυσικό / Βιολογικό

- Εικονικό (ΗΥ)

- Μαθηματικό

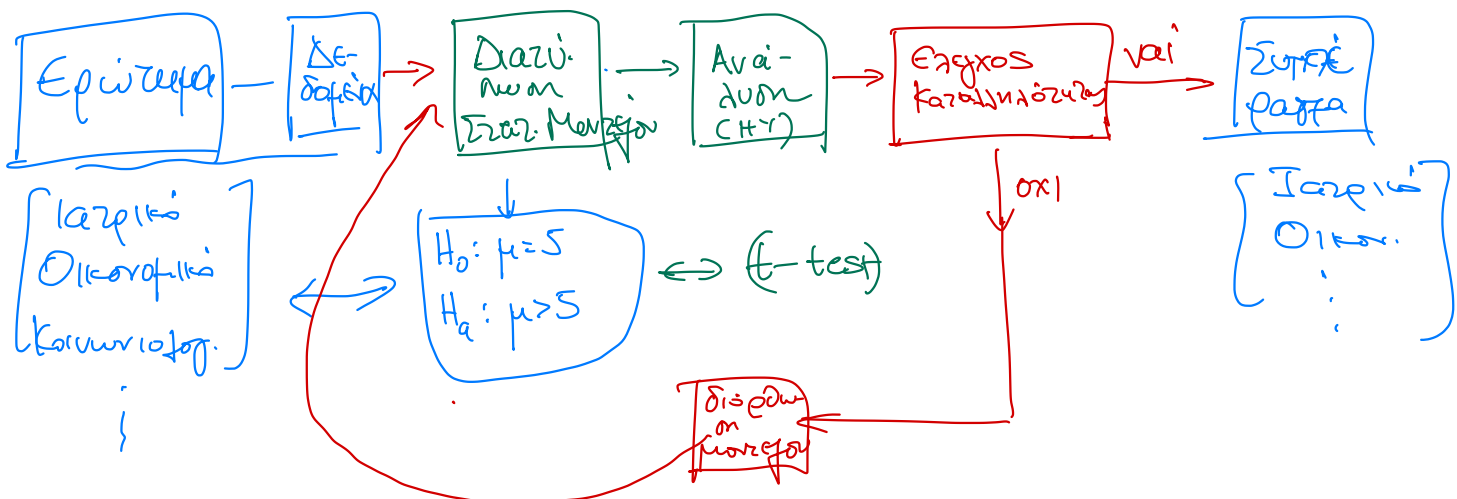
(μεσω μηθ. μεταβλητών & εξισώσεων)

- Στατιστικά Μοντέλα

Για φαινόμενα/πρόβλεψα με ετερο-
παρασάρματα (συντεταγ. από δεδομένα),
σφάλματα

Διαδικασία Μοντεφορνίον

GIGO (Garbage in
Garbage out)



Γραμμικά Μοντέλα / Παμνδρόμων - Ανάλυση Διασποράς.

Παράδειγμα dataset exercise

score : score φυσικής παράστασης (scale)

age : ηλικία ατόμων

sex : φύλο

extime : χρόνος ασκήσεως (min/week)

① Εξήγηση μεταβλητότητας μέσω φύλου

$$\mu_M = E(\text{score} | \text{Male}) \quad X \sim N(\mu_M, \sigma^2)$$

$$\mu_F = E(\text{score} | \text{Female}) \quad X \sim N(\mu_F, \sigma^2)$$

$$H_0: \mu_M = \mu_F \quad \text{vs} \quad H_1: \mu_M \neq \mu_F$$

$$p\text{-value} = 0.47$$

② Γενίκευση Έστω ότι έχουμε άλλη μεταβλητή

εργασία { χειρωνακική
πραγματ
ενδιάμεσο

$$\begin{aligned} X &\sim N(\mu_1, \sigma^2) \\ X &\sim N(\mu_2, \sigma^2) \\ X &\sim N(\mu_3, \sigma^2) \end{aligned}$$

$$H_0: \mu_1 = \mu_2 = \mu_3$$

$$H_1: \text{τουλάχιστον δύο διαφορές}$$

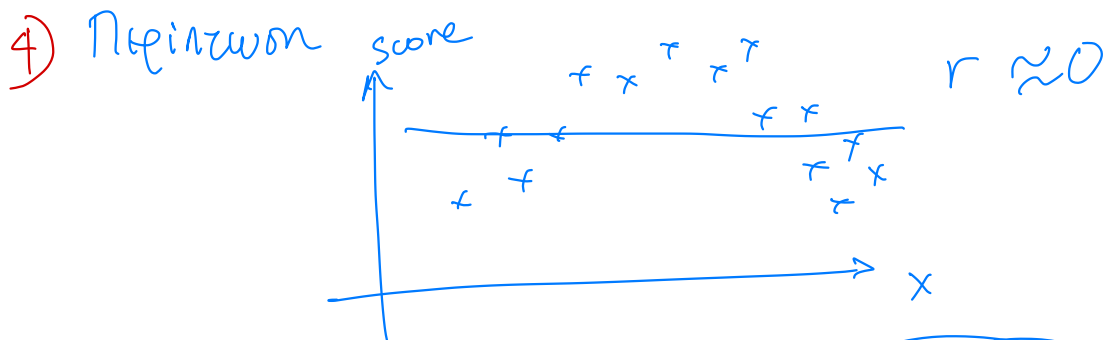
χ^2

Μέθοδος ANOVA (Ανάλυση Διασποράς)

3) Εξήγησε με 2. score μέσω ηλικίας

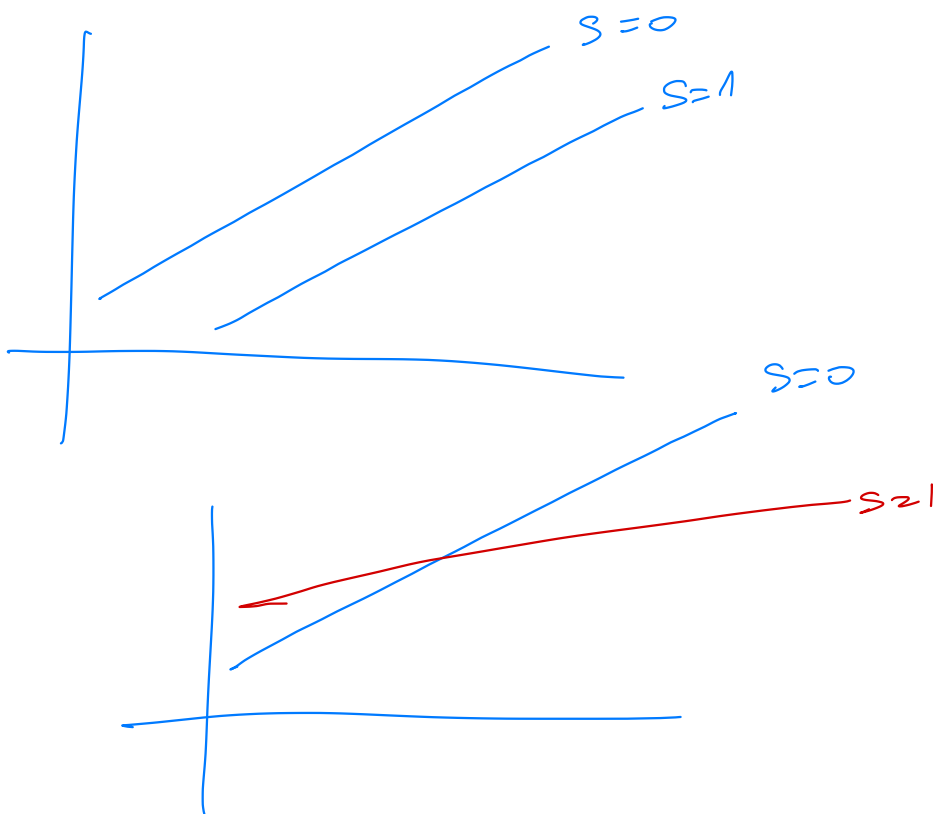
1) γραφικά: scatter plot.

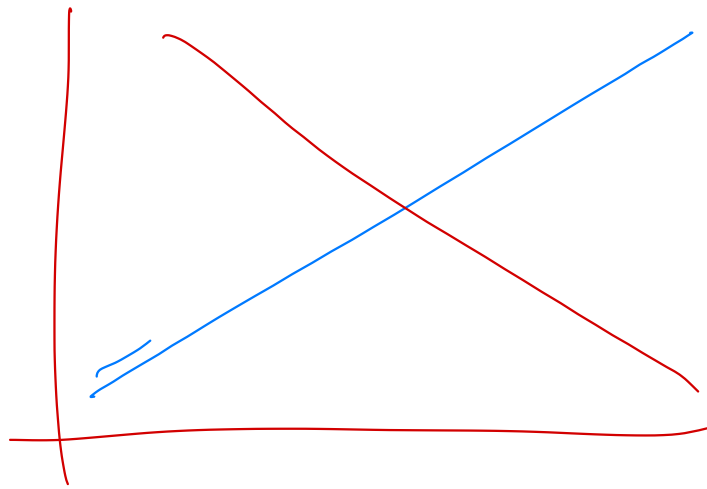
2) συντ. συσχέτισης $r(\text{score}, \text{age}) = \underline{-0.93}$



υπάρχει συσχέτιση εφ' όσον οχι γραμμική

5) Συσχέτιση score με age, extime





Ερωτήματα - Στόχοι της Ανάλυσης Παλινδρόμησης

Response variable (score) : έχει μεταβλητότητα στο δείγμα

1) Που αποδίδεται (πώς εξηγείται) η μεταβλητότητα; μέσω αλληλ μεταβλητών (παραγόντες)

2) Διερεύνηση συσχέτιση των response με αλληλ παραγόντες (γενικότερα από πραγματική συσχέτιση μεταξύ δύο μεταβ.

3) Παρατήρηση Διερεύνηση συσχέτισης
Συσχέτιση \neq Αιτιακή Σχέση
(correlation) (Causality)

Στόχοι Ανάλυσης Παλινδρόμησης

1) Ανίχνευση Συσχετίσεων } μιας μεταβλητής
Εξήγηση μεταβλητότητας } απόκλισης με
αλληλ παραγόντες

2) Πρόβλεψη { για πιθανό score ενός νέου (εκτός μελέτης) ατόμου με ορισμένους. Q_{obs} , age, extime

Machine Learning

