

# ΠΡΟΒΛΗΜΑ ΟΡΘΟΤΗΤΑΣ ΜΟΝΤΕΛΟΥ

- ΤΡΟΠΟΙ ΕΛΕΓΧΟΥ
- ΔΙΟΡΘΩΣΗ ΠΡΟΒΛΗΜΑΤΟΣ
- ΣΥΝΕΠΕΙΕΣ ΠΡΟΒΛΗΜΑΤΟΣ
- ΠΑΡΑΔΕΙΓΜΑΤΑ

# ΤΡΟΠΟΙ ΕΛΕΓΧΟΥ

- Γραφική παράσταση των υπολοίπων ως προς τις ανεξάρτητες μεταβλητές. Αν δεν παρατηρηθεί κάποια ιδιαίτερη μορφή και τα σημεία βρίσκονται τυχαία γύρω από το μηδέν το μοντέλο μπορεί να θεωρηθεί ορθό. Αν δούμε κάποια ιδιαίτερη γραφική παράσταση τότε η εξαρτημένη και η ανεξάρτητη μεταβλητή μπορεί να μην συνδέονται με μία γραμμική σχέση.

# ΤΡΟΠΟΙ ΕΛΕΓΧΟΥ

- Partial Regression Plot: Όταν το μοντέλο είναι ορθό, τα γραφήματα αυτά είναι γραμμικά. Αν όχι η εικόνα αυτών των γραφικών παραστάσεων ίσως μας βοηθήσει στο να αποφανθούμε για κάποιο μετασχηματισμό στις ανεξάρτητες μεταβλητές ή να συμπεριληφθεί στο μοντέλο κάποιος όρος

# ΣΥΝΕΠΕΙΕΣ

Οι συνέπειες ενός μη ορθού μοντέλου είναι: λάθος ερμηνεία των παραμέτρων του μοντέλου, λάθος προβλέψεις, λάθος εκτίμηση της κοινής διακύμανσης των σφαλμάτων. Από το τελευταίο προκύπτει η μη εγκυρότητα των όποιων διαστημάτων εμπιστοσύνης και ελέγχου υποθέσεων για τις παραμέτρους του μοντέλου.

# Τρόποι διόρθωσης του προβλήματος

- Κατάλληλος κλασικός μετασχηματισμός ή προσθήκη όρου.

1. Λογάριθμος  $Y_i = a X_i^\beta \varepsilon_i$

2. Λογάριθμος  $Y_i = a \exp(\beta X_i) \varepsilon_i$

3. Αντίστροφος  $Y_i = \frac{X_i}{a + \beta X_i + \varepsilon_i}$

$$Y^* = \ln \left( \frac{a}{Y} - 1 \right)$$

$$Y_i = \frac{a}{1 + \gamma \exp(-\beta X_i) \varepsilon_i}$$

# Box-Tidwell transformation

1. Προσαρμόζουμε το αρχικό μοντέλο
2. Έστω  $b_j$   
οι εκτιμητές των παραμέτρων.
3. Προσαρμόζουμε το μοντέλο που έχει ανεξάρτητες μεταβλητές τις :

$$X_1, \dots, X_k, Z_1, \dots, Z_k$$

όπου

$$Z_i = X_i \ln(X_i)$$

4.  $\gamma_j$   
οι εκτιμητές των παραμέτρων των νέων μεταβλητών

# Box-Tidwell transformation

5. Ο μετασχηματισμός που υποδεικνύεται είναι:

$$W_j = \begin{cases} X_j^{a_j}, & a_j \neq 0 \\ \ln(X_j), & a_j = 0 \end{cases}$$

όπου

$$a_j = \frac{\gamma_j}{b_j} + 1$$

# Box Tidwell transformation

- Κανονικά η διαδικασία επαναλαμβάνεται μέχρις ότου οι τελικές τιμές να είναι κοντά στη μονάδα.
- Συνήθως όμως αν επιτευχθεί η ορθότητα του μοντέλου σταματάμε.



# Παράδειγμα 1

- Στο αρχείο `trees.sav` καταγράφονται η διάμετρος, το ύψος και ο όγκος 31 δέντρων. Θέλουμε ένα μοντέλο πρόβλεψης του όγκου από τις άλλες μεταβλητές.

## Παράδειγμα 2

Chatterjee and Price (1980), p. 32

- Στο αρχείο `bacteria.sav` καταγράφεται ο αριθμός (σε χιλιάδες) των βακτηρίων που επέζησαν ( $nt$ ) όταν εκτέθηκαν σε ακτινοβολία για 1-15 χρονικές περιόδους των 6 λεπτών ( $t$ ). Από τη θεωρία είναι γνωστό ότι

$$nt = n_0 \exp(\beta t)$$

Επιβεβαιώνεται;