Εισαγωγή στην οικονομετρία (1404)

Διάλεξη 14

Άγγελος Αλεξόπουλος
Οικονομικό Πανεπιστήμιο Αθηνών -
Τμήμα Οικονομικής Επιστήμης
angelos@aueb.gr

Περίγραμμα Διάλεξης

- Περαιτέρω εξειδίκευση υποδείγματος
    - Ημιλογαριθμικός μετασχηματισμός
    - Λογαριθμικός μετασχηματισμός

    - Πρόβλεψη στο ημιλογαριθμικό & λογαριθμικό υπόδειγμα
    - Διάστημα εμπιστοσύνης στο ημιλογαριθμικό & λογαριθμικό υπόδειγμα

- Άσκηση

Μη γραμμικές συναρτήσεις - Εισαγωγή

- Μέχρι τώρα έχουμε υποθέσει ότι το υπόδειγμα είναι γραμμικό ως προς τις παραμέτρους (κλασικές υποθέσεις) και ως προς τις μεταβλητές
  - ωστόσο η γραμμική προσέγγιση ως προς τις μεταβλητές δεν είναι πάντοτε καλή
  - π.χ. διαφορετική κλίμακα μέτρησης μεταβλητών → δυσανάλογο μέγεθος διακυμάνσεων
- Εάν η σχέση μεταξύ της Y και της X είναι **μη γραμμική**:
  - Η επίδραση στην Y μιας αλλαγής στην X εξαρτάται από την τιμή της X - δηλαδή, η οριακή επίδραση της X δεν είναι σταθερή
  - Μια γραμμική παλινδρόμηση προσδιορίζεται λανθασμένα: λάθος εξειδίκευση υποδείγματος
  - Ο εκτιμητής της επίδρασης στην Y της X είναι μεροληπτικός
  - Λύση: μη γραμμική συνάρτηση παλινδρόμησης

Λογαριθμικοί μετασχηματισμοί

- Στην εμπειρική ανάλυση είναι σύνηθες η συναρτησιακή σχέση δύο μεταβλητών να μην είναι γραμμική
  - Λογάριθμοι –εξαρτημένη και/ή ανεξάρτητη μεταβλητή-, πολυωνυνική μορφή (π.χ. τετράγωνο ανεξάρτητης μεταβλητής), δυαδική μεταβλητή/ψευδομεταβλητή κτλ
  - Οι μετασχηματισμοί επηρεάζουν την ερμηνεία των συντελεστών

Λογάριθμος (φυσικός λογάριθμος, ln) – εξαρτημένη και/ή ανεξάρτητη μεταβλητή

- ✓Ο λογαριθμικός μετασχηματισμός μπορεί να εφαρμοστεί σε μεταβλητές που λαμβάνουν θετικές τιμές
- ✓Μειώνει την ασυμμετρία των υπό εξέταση μεταβλητών
- ✓Μειώνει τη μεταβλητότητα των υπό εξέταση μεταβλητών
  - ✓Ακραίες τιμές

- Ερμηνεία συντελεστών σε «ποσοστά»

Ιδιότητες λογαρίθμων

Ιδιότητες λογαρίθμων:

$$\ln(e) = 1$$
$$\ln(1) = 0$$
$$\ln(AB) = lnA + lnB$$
$$\ln\left(\frac{A}{B}\right) = lnA - lnB$$
$$\ln(A^k) = klnA$$

όπου A και B είναι θετικά και k είναι σταθερά

Γραμμικός-λογαριθμικός μετασχηματισμός

Το υπόδειγμα

$$Y_i = \beta_0 + \beta_1 lnX_i + u_i$$

καλείται γραμμικό-λογαριθμικό υπόδειγμα (lin-log) ή ημιλογαριθμικό

- Ερμηνεία β_1: μια αύξηση της μεταβλητής X κατά 1% οδηγεί σε μια $\frac{\beta_1}{100}$ (μονάδες μέτρησης) μεταβολή (αύξηση ή μείωση) στην μεταβλητή Y

$$\frac{dY}{dX} = \beta_1 \frac{1}{X}$$

- Η ελαστικότητα της Y ως προς X

$$\varepsilon = \frac{dY/Y}{dX/X} = \frac{dY}{dX} \cdot \frac{X}{Y} = \frac{\beta_1}{X} \cdot \frac{X}{Y} = \beta_1 \frac{1}{Y}$$

  - $\hat{\varepsilon} = \hat{\beta}_1 \frac{1}{Y}$

Λογαριθμικός-γραμμικός μετασχηματισμός

Το υπόδειγμα

$$lnY_i = \beta_0 + \beta_1 X_i + u_i$$

καλείται λογαριθμικό-γραμμικό υπόδειγμα (log-lin) ή ημιλογαριθμικό

- Ερμηνεία: μια αύξηση της μεταβλητής X κατά μονάδα οδηγεί σε μια $(100\beta_1)\%$ μεταβολή (αύξηση ή μείωση) στην μεταβλητή Y

Λύνουμε ως προς Y_i: $exp(lnY_i) = exp(\beta_0 + \beta_1 X_i + u_i) \Rightarrow Y_i = exp(\beta_0 + \beta_1 X_i + u_i)$

$$\frac{dY}{dX} = \beta_1\, exp(\beta_0 + \beta_1 X_i + u_i) = \beta_1 Y$$

- Η ελαστικότητα της Y ως προς X

$$\varepsilon = \frac{dY/Y}{dX/X} = \frac{dY}{dX} \cdot \frac{X}{Y} = \beta_1 Y \frac{X}{Y} = \beta_1 \mathrm{X}$$

  - $\hat{\varepsilon} = \hat{\beta}_1 \mathrm{X}$

Λογαριθμικός-λογαριθμικός μετασχηματισμός

Το υπόδειγμα

$$lnY_i = \beta_0 + \beta_1 lnX_i + u_i$$

καλείται λογαριθμικό-λογαριθμικό υπόδειγμα (log-log ή log-linear)

- Ερμηνεία: αναμενόμενη ποσοστιαία μεταβολή της Y όταν η X αυξηθεί κατά ένα ποσοστό

Λύνουμε ως προς Y_i: $Y_i = exp(\beta_0 + \beta_1 lnX_i + u_i)$

$$\frac{dY}{dX} = \frac{\beta_1}{X} exp(\beta_0 + \beta_1 lnX_i + u_i) = \frac{\beta_1}{X} Y$$

- Ο συντελεστής κλίσης β_1 μετράει την ελαστικότητα της Y ως προς X

$$\varepsilon = \frac{dY/Y}{dX/X} = \frac{dY}{dX} \cdot \frac{X}{Y} = \beta_1 \frac{Y}{X} \frac{X}{Y} = \beta_1$$

  - $\hat{\varepsilon} = \hat{\beta}_1$

Πρόβλεψη & Διάστημα εμπιστοσύνης– Λογαριθμικός γραμμικός μετασχηματισμός

- Πρόβλεψη σε μικρά δείγματα ($n < 30$)

$$\hat{Y}_n = \hat{Y}_0 = \exp\left(ln\hat{Y}_0\right) = exp\left(\hat{\beta}_0 + \hat{\beta}_1 X_0\right)$$

- Σε μεγάλα δείγματα «διορθωμένη πρόβλεψη»

$$\hat{Y}_c = \hat{Y}_0 = \exp\left(ln\hat{Y}_0\right) = exp\left(\hat{\beta}_0 + \hat{\beta}_1 X_0 + \frac{\hat{\sigma}^2}{2}\right) = \hat{Y}_n\, exp\left(\frac{\hat{\sigma}^2}{2}\right)$$

όπου $\hat{\sigma}^2$ ο εκτιμητής της διακύμανσης του διαταρακτικού όρου

- Διάστημα εμπιστοσύνης

$$\left[exp\left(\hat{Y}_n - t_c \cdot se(f)\right), exp\left(\hat{Y}_n + t_c \cdot se(f)\right)\right]$$

$se(f)$ το τυπικό σφάλμα της πρόβλεψης (διάλεξη 10 $\sqrt{Var\left(\hat{Y}_0 - Y_0\right)}$)

Πρόβλεψη & Διάστημα εμπιστοσύνης– Λογαριθμικός μετασχηματισμός

- Πρόβλεψη σε μικρά δείγματα ($n < 30$)

$$\hat{Y}_n = \hat{Y}_0 = \exp(ln\hat{Y}_0) = exp(\hat{\beta}_0 + \hat{\beta}_1 lnX_0)$$

- Σε μεγάλα δείγματα «διορθωμένη πρόβλεψη»

$$\hat{Y}_c = \hat{Y}_0 = \exp(ln\hat{Y}_0) = exp\left(\hat{\beta}_0 + \hat{\beta}_1 lnX_0 + \frac{\hat{\sigma}^2}{2}\right) = \hat{Y}_n\, exp\left(\frac{\hat{\sigma}^2}{2}\right)$$

όπου $\hat{\sigma}^2$ ο εκτιμητής της διακύμανσης του διαταρακτικού όρου

- Διάστημα εμπιστοσύνης

$$\left[exp(\hat{Y}_n - t_c \cdot se(f)), exp(\hat{Y}_n + t_c \cdot se(f))\right]$$

$se(f)$ το τυπικό σφάλμα της πρόβλεψης (διάλεξη 10 $\sqrt{Var(\hat{Y}_0 - Y_0)}$)

Σύνοψη

Υπόδειγμα	**Συνάρτηση**	**Κλίση** $\left(=\frac{dY}{dX}\right)$	**Ελαστικότητα** $\left(=\frac{dY}{dX}\cdot\frac{X}{Y}\right)$	**Ερμηνεία**
Γραμμικό	$Y_i = \beta_0 + \beta_1 X_i + u_i$	β_1	$\beta_1 \frac{X_i}{Y_i}$	$dY_i = \beta_1 dX_i$
Lin-Log	$Y_i = \beta_0 + \beta_1 lnX_i + u_i$	$\beta_1 \frac{1}{X_i}$	$\beta_1 \frac{1}{Y_i}$	$dY_i = (\beta_1/100)\ (\%dX_i)$
Log-Lin	$lnY_i = \beta_0 + \beta_1 X_i + u_i$	$\beta_1 Y_i$	$\beta_1 X_i$	$\%dY_i = (100\beta_1)dX_i$
Log-Log	$lnY_i = \beta_0 + \beta_1 lnX_i + u_i$	$\beta_1 \frac{Y_i}{X_i}$	β_1	$\%dY_i = \beta_1(\%dX_i)$

- Τα υποδείγματα είναι γραμμικά ως προς τις παραμέτρους β_0 και β_1
- Αν ισχύουν οι κλασικές υποθέσεις μπορούν να εκτιμηθούν με τη μέθοδο ΕΤ και οι εκτιμητές να είναι BLUE
- Οι έλεγχοι υποθέσεων και τα διαστήματα εμπιστοσύνης εφαρμόζονται και ερμηνεύονται «ως συνήθη».
- Διαφοροποίηση πρόβλεψης για μεγάλα δείγματα

Άσκηση (Hill et al.)

✓Έστω ότι έχετε ένα δείγμα 880 παρατηρήσεων για πωλήσεις σπιτιών στην πόλη Stockton στην Καλιφόρνια στα μέσα του 2005 (*price* τιμή πώλησης σε $) και διαθέτετε πληροφορίες για τις εξής μεταβλητές:

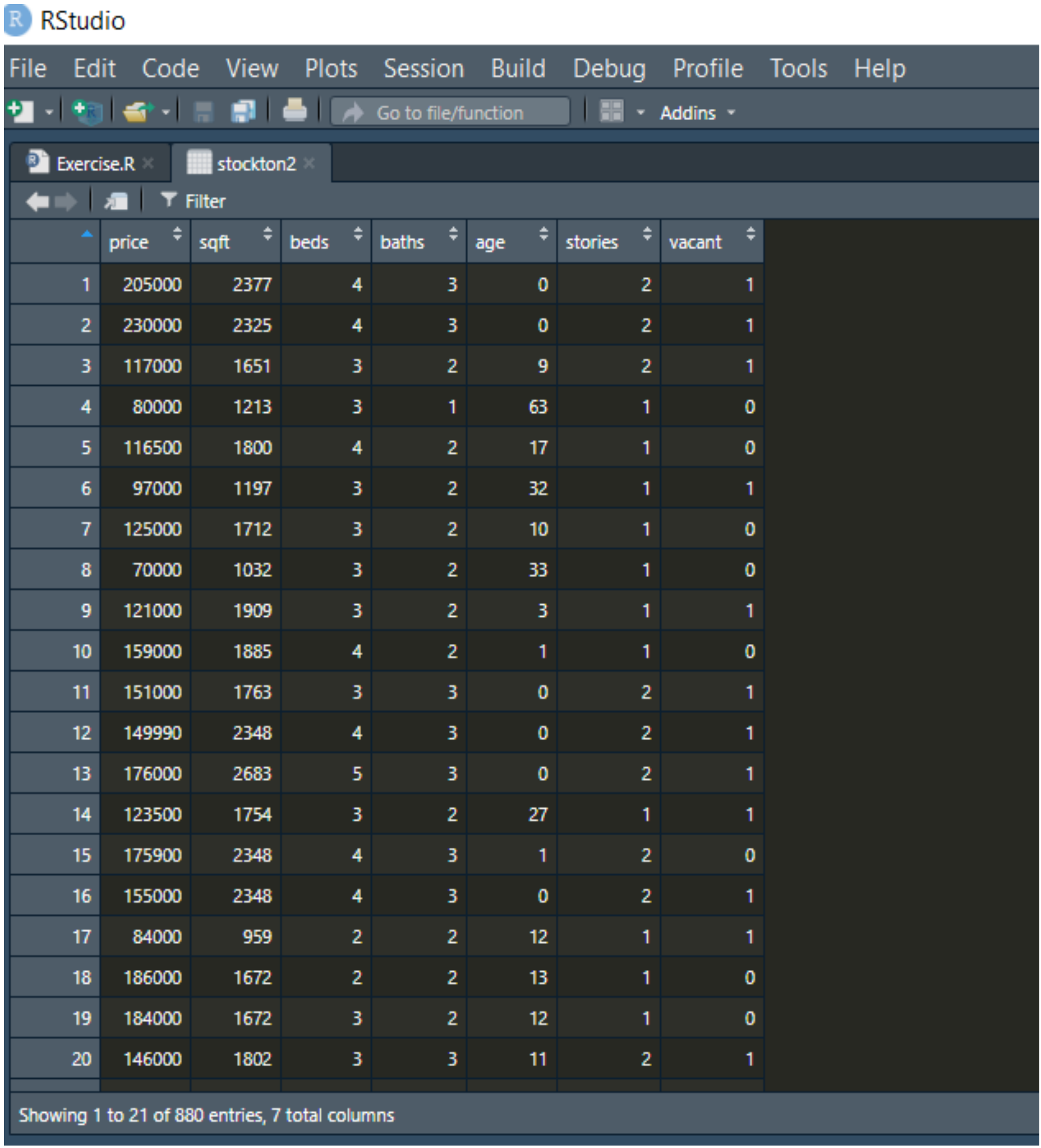
sqft εμβαδόν σε τετραγωνικά πόδια; *beds* αριθμός υπνοδωματίων;

baths αριθμός μπάνιων;

age η ηλικία του σπιτιού σε έτη; stories αριθμός ορόφων;

vacant νοικιασμένο ή μη

(αρχείο: stockton2.RDA)



	price	sqft	beds	baths	age	stories	vacant
1	205000	2377	4	3	0	2	1
2	230000	2325	4	3	0	2	1
3	117000	1651	3	2	9	2	1
4	80000	1213	3	1	63	1	0
5	116500	1800	4	2	17	1	0
6	97000	1197	3	2	32	1	1
7	125000	1712	3	2	10	1	0
8	70000	1032	3	2	33	1	0
9	121000	1909	3	2	3	1	1
10	159000	1885	4	2	1	1	0
11	151000	1763	3	3	0	2	1
12	149990	2348	4	3	0	2	1
13	176000	2683	5	3	0	2	1
14	123500	1754	3	2	27	1	1
15	175900	2348	4	3	1	2	0
16	155000	2348	4	3	0	2	1
17	84000	959	2	2	12	1	1
18	186000	1672	2	2	13	1	0
19	184000	1672	3	2	12	1	0
20	146000	1802	3	3	11	2	1

Showing 1 to 21 of 880 entries, 7 total columns

Ερώτημα i

i. Για το γραμμικό υπόδειγμα $price_i = \beta_0 + \beta_1 sqft_i + u_i$ η εκτίμηση ΕΤ έδωσε

$$\begin{array}{lccc} \widehat{price}_i = & -18385.65 & + \; 81.389 \, sqft_i & \\ (se) & (3256.424) & (1.918) & \end{array}, R^2 = 0.6721$$

με τους μέσους του δείγματος $\overline{sqrt} = 1611.968$ και $\overline{price} = 112810.8$

Ερμηνεύστε τους εκτιμημένους συντελεστές του υποδείγματος. Υπολογίστε την ελαστικότητα για το μέσο του δείγματος.

Λύση:

Για τον εκτιμητή του σταθερού όρου έχουμε ότι η τιμή πώλησης ενός σπιτιού με μηδενικό εμβαδό είναι -18385.65 (μη ρεαλιστικό)

Για τον εκτιμητή της κλίσης έχουμε ότι αν αυξηθεί το εμβαδό ενός σπιτιού κατά ένα τετραγωνικό πόδι, η τιμή του σπιτιού θα αυξηθεί κατά $81.389\$$

Ερώτημα i

(συνέχεια)

$$\widehat{price}_i = -18385.65 + 81.389\, sqft_i$$
$$(se) \quad (3256.424) \quad (1.918)$$

$\overline{sqft} = 1611.968$ και $\overline{price} = 112810.8$

- Η ελαστικότητα

$$\varepsilon = \frac{\%\Delta price}{\%\Delta sqft} = \frac{dprice/price}{dsqft/sqft} = \frac{dprice}{dsqft} \cdot \frac{sqft}{price} = \beta_1 \frac{sqft}{price}$$

Η μέση ελαστικότητας $\bar{\varepsilon} = \hat{\beta}_1 \frac{\overline{sqft}}{\overline{price}} = 81.389 * \left(\frac{1611.968}{112810.8}\right) = 1.163$

Στο μέσο, αν εμβαδό αυξηθεί κατά 1%, η τιμή πώλησης θα αυξηθεί κατά 1.163%

Ερώτημα ii

ii. Για το ημιλογαριθμικό (λογαριθμικό-γραμμικό) υπόδειγμα $\ln(price_i) = \beta_0 + \beta_1 sqft_i + u_i$ η εκτίμηση ΕΤ έδωσε

$$\underset{(se)}{\ln(\widehat{price_i})} = \underset{(0.0219)}{10.5938} + \underset{(0.000013)}{0.000596}\ sqft_i,\ R^2 = 0.7094$$

με τους μέσους του δείγματος $\overline{sqft} = 1611.968$ και $\overline{price} = 112810.8$

Ερμηνεύστε τους εκτιμημένους συντελεστές του υποδείγματος. Υπολογίστε την κλίση και την ελαστικότητα για το μέσο του δείγματος.

Λύση:

Για τον εκτιμητή του σταθερού όρου έχουμε ότι η τιμή πώλησης ενός σπιτιού με μηδενικό εμβαδό είναι 10.5938 (μη ρεαλιστικό)

Για τον εκτιμητή της κλίσης έχουμε ότι αν αυξηθεί το εμβαδό ενός σπιτιού κατά ένα τετραγωνικό πόδι, η τιμή του σπιτιού θα αυξηθεί κατά 0.06% $(0.000596 * 100 = 0.0596 \sim 0.06)$

Ερώτημα ii

(συνέχεια)

$$\ln(\widehat{price_i}) = 10.5938 + \ 0.000596\ sqft_i$$
$$(se) \qquad (0.0219) \quad (0.000013)$$

με τους μέσους του δείγματος $\overline{sqft} = 1611.968$ *και* $\overline{price} = 112810.8$

- Η *κλίση* υπολογίζεται ως εξής: $price_i = exp(\beta_0 + \beta_1 sqft_i + u_i)$

$$\frac{dprice}{dsqft} = \beta_1 price$$

και στο μέσο $\dfrac{dprice}{dsqft} = \hat{\beta}_1 \overline{price} = 0.000596 * 112810.8 = 67.23$

Στο μέσο, αν εμβαδό αυξηθεί *κατά κατά* ένα τετραγωνικό πόδι, η τιμή του σπιτιού θα αυξηθεί *κατά* 67.23$

Ερώτημα ii

(συνέχεια)

$$\ln(\widehat{price_i}) = 10.5938 + \ 0.000596 \ sqft_i$$
$$(se) \qquad (0.0219) \quad (0.000013)$$

με τους μέσους του δείγματος $\overline{sqft} = 1611.968$ και $\overline{price} = 112810.8$

- Η ελαστικότητα

$$\varepsilon = \frac{\%\Delta price}{\%\Delta sqft} = \frac{dprice/price}{dsqft/sqft} = \frac{dprice}{dsqft} \cdot \frac{sqft}{price} = \beta_1 price \cdot \frac{sqft}{price}$$
$$= \beta_1 sqft$$

Η μέση ελαστικότητας $\bar{\varepsilon} = \hat{\beta}_1 \ \overline{sqrt} = 0.000596 * 1611.968 = 0.9607$

Στο μέσο, αν εμβαδό αυξηθεί κατά 1%, η τιμή πώλησης θα αυξηθεί κατά ~1%

Ερώτημα iii

iii. Για το λογαριθμικό (λογαριθμικό-λογαριθμικό) υπόδειγμα $\ln(price_i) = \beta_0 + \beta_1 \ln(sqft_i) + u_i$ η εκτίμηση ΕΤ έδωσε

$$\begin{array}{cccc} \ln(\widehat{price_i}) = & 4.1707 & + & 1.0066 \ln(sqft_i) \\ (se) & (0.1655) & & (0.0225) \end{array}, R^2 = 0.6943$$

με τους μέσους του δείγματος $\overline{sqft} = 1611.968$ και $\overline{price} = 112810.8$ Ερμηνεύστε τους εκτιμημένους συντελεστές του υποδείγματος. Υπολογίστε την κλίση και την ελαστικότητα για το μέσο του δείγματος.

Λύση:

Για τον εκτιμητή του σταθερού όρου έχουμε ότι η τιμή πώλησης ενός σπιτιού με εμβαδό ενός τετραγωνικού ποδιού $(\ln(sqft_i) = 0 \rightarrow sqft_i = 1)$ είναι 4.1707 (μη ρεαλιστικό)

Για τον εκτιμητή της κλίσης έχουμε ότι αν αυξηθεί το εμβαδό ενός σπιτιού κατά 1%, η τιμή του σπιτιού θα αυξηθεί κατά 1% (1.0066%)

Ερώτημα iii

(συνέχεια)

$$\ln(\widehat{price_i}) = 4.1707 + 1.0066\ \ln(sqft_i)$$
$$(se) \qquad (0.1655) \quad (0.0225)$$

με τους μέσους του δείγματος $\overline{sqft} = 1611.968$ και $\overline{price} = 112810.8$

- Η κλίση υπολογίζεται ως εξής: $price_i = exp(\beta_0 + \beta_1 \ln(sqft_i) + u_i)$

$$\frac{dprice}{dsqft} = \beta_1 \frac{price}{sqft}$$

και στο μέσο $\frac{dprice}{dsqft} = \hat{\beta}_1 \frac{\overline{price}}{\overline{sqrt}} = 1.0066 * \frac{112810.8}{1611.968} = 70.44$

Στο μέσο, αν εμβαδό αυξηθεί κατά κατά ένα τετραγωνικό πόδι, η τιμή του σπιτιού θα αυξηθεί κατά 70.44$

Ερώτημα iii

(συνέχεια)

$$\ln(\widehat{price_i}) = 4.1707 + 1.0066 \ln(sqft_i)$$
$$(se) \qquad (0.1655) \quad (0.0225)$$

με τους μέσους του δείγματος $\overline{sqrt} = 1611.968$ και $\overline{price} = 112810.8$

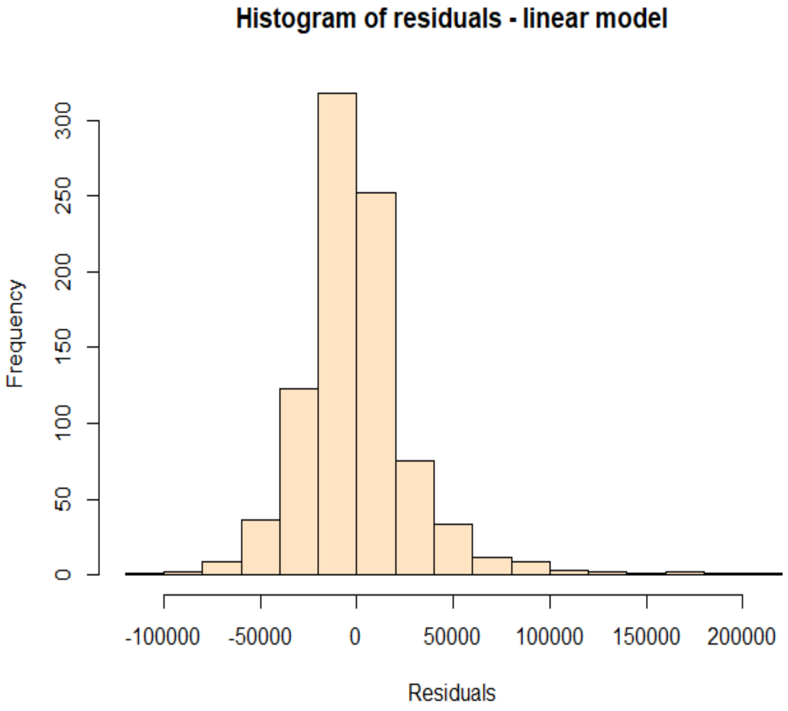
- Η ελαστικότητα

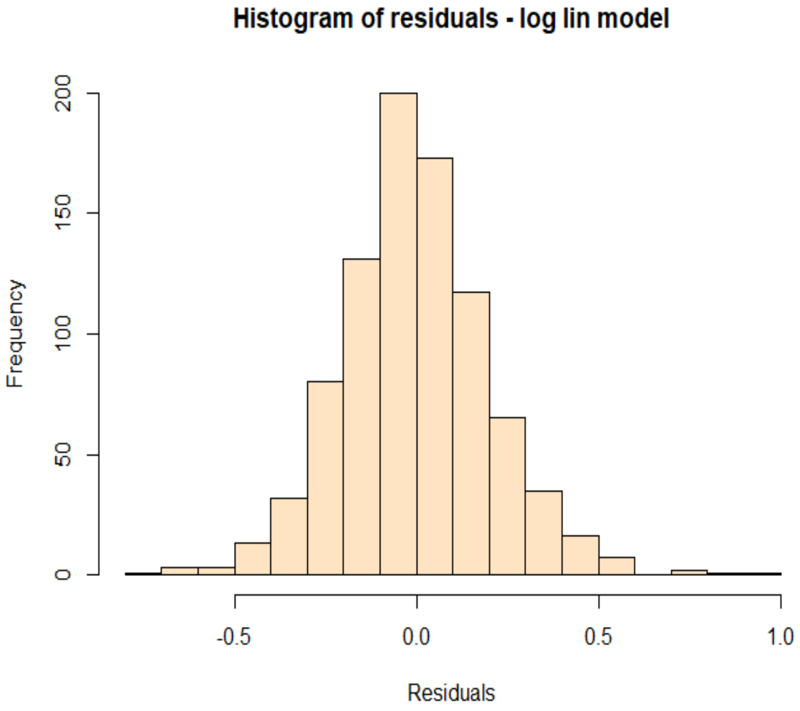
$$\varepsilon = \frac{\%\Delta price}{\%\Delta sqft} = \frac{dprice/price}{dsqft/sqft} = \frac{dprice}{dsqft} \cdot \frac{sqft}{price} = \beta_1 \frac{price}{sqrt} \frac{sqft}{price}$$
$$= \beta_1$$

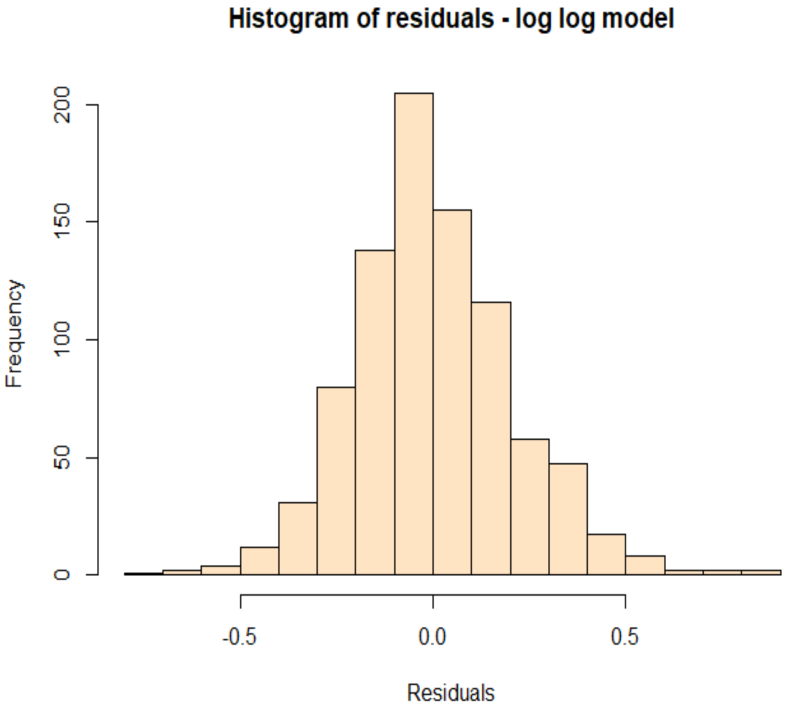
Σταθερή ελαστικότητα ($\hat{\beta}_1 = 1.0066$)

Ερώτημα iv

iv. Σας δίνονται τα ιστόγραμμα των καταλοίπων από τα εκτιμημένα υποδείγματα στα ερωτήματα i-iii. Η κατανομή των καταλοίπων συμβαδίζει με την υπόθεση της κανονικής κατανομής του διαταρακτικού όρου?

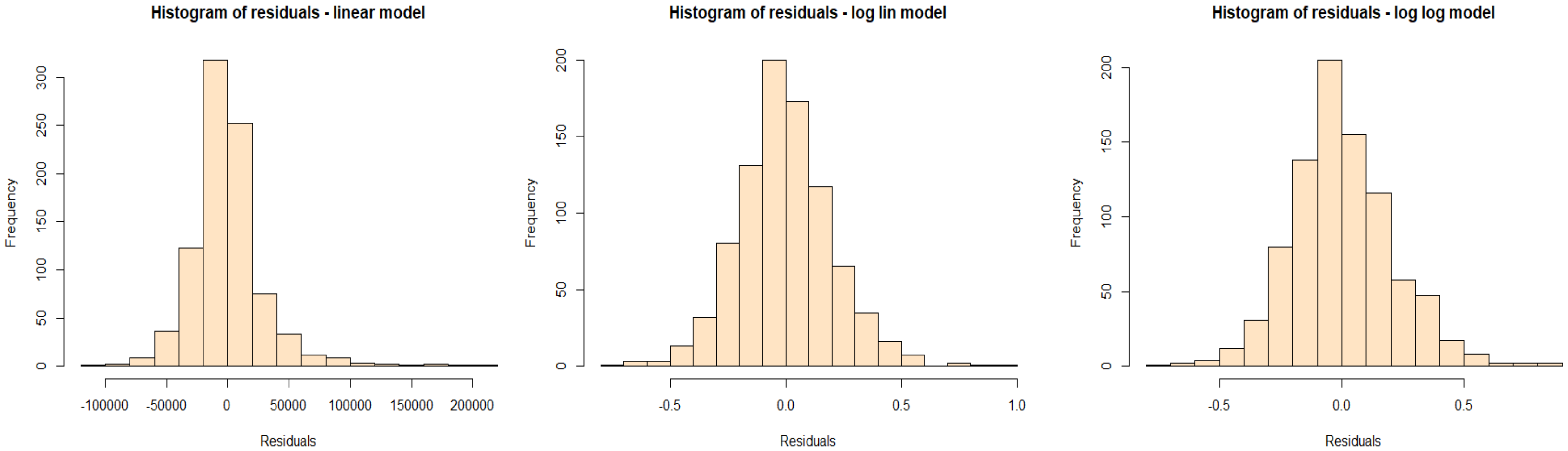






Ερώτημα iv

- Απόκλιση από την κανονική κατανομή
- Γραμμικό υπόδειγμα - δεξιά ασσυμετρία
- Ημιλογαριθμικό & λογαριθμικό υπόδειγμα – «πιο κοντά» στην κανονική κατανομή



Ερώτημα v

v. Για τα υποδείγματα στα ερωτήματα i-iii, ποια θα είναι η τιμή πώλησης όταν sqft=2700?

Σας δίνεται ότι για το ημιλογαριθμικό υπόδειγμα το τυπικό σφάλμα της παλινδρόμησης είναι 0.20303 και για το λογαριθμικό είναι 0.2083

Λύση:

Πρόβλεψη για το γραμμικό υπόδειγμα

$$\widehat{price}_i = -18385.65 + 81.389\, sqft_i$$
$$(se) \quad (3256.424) \quad (1.918)$$

για $sqft_0 = 2700$

$$\widehat{price}_0 = \hat{\beta}_0 + \hat{\beta}_1 sqft_0 = -18385.65 + 81.389 * 2700 = 201365$$

Ερώτημα v

(συνέχεια)

Σας δίνεται ότι για το ημιλογαριθμικό υπόδειγμα το τυπικό σφάλμα της παλινδρόμησης είναι 0.20303 και για το λογαριθμικό είναι 0.2083

- Πρόβλεψη για το ημιλογαριθμικό υπόδειγμα

$$\ln(\widehat{price_i}) = 10.5938 + 0.000596\ sqft_i$$
$$(se) \qquad (0.0219) \qquad (0.000013)$$

για $sqft_0 = 2700$

$$\widehat{price}_0 = \exp\left(\hat{\beta}_0 + \hat{\beta}_1 sqft_0 + \frac{\hat{\sigma}^2}{2}\right)$$
$$= \exp\left(10.5938 + 0.000596 * 2700 + \frac{0.20303^2}{2}\right) = 203516$$

Ερώτημα v

(συνέχεια)

Σας δίνεται ότι για το ημιλογαριθμικό υπόδειγμα το τυπικό σφάλμα της παλινδρόμησης είναι 0.20303 και για το λογαριθμικό είναι 0.2083

- Πρόβλεψη για το λογαριθμικό υπόδειγμα

$$\ln(\widehat{price_i}) = 4.1707 + 1.0066 \ln(sqft_i)$$
$$(se) \qquad (0.1655) \quad (0.0225)$$

αν $sqft_0 = 2700$

$$\widehat{price_0} = \exp(\widehat{price_0}) = exp\left(\hat{\beta}_0 + \hat{\beta}_1 \ln(sqft_0) + \frac{\hat{\sigma}^2}{2}\right)$$
$$= exp\left(4.1707 + 1.0066 \ln(2700) + \frac{0.2083^2}{2}\right)$$
$$= 188221$$

Ερώτημα vi

vi. Για τα υποδείγματα στα ερωτήματα i-iii κατασκευάστε το 95% διάστημα εμπιστοσύνης για την πρόβλεψη του προηγούμενου ερωτήματος.

Σας δίνεται ότι για το γραμμικό υπόδειγμα $\hat{\sigma} = 3519.2$, $\sum_{i=1}^{n}(X_i - \bar{X})^2 = 248768933.1$, για το ημιλογαριθμικό $\hat{\sigma} = 0.20303$ και για το λογαριθμικό είναι $\hat{\sigma} = 0.2083$ και $\sum_{i=1}^{n}(X_i - \bar{X})^2 = 85.34453$ και $\bar{X} = 7.3355$

Ερώτημα vi

Λύση:

- Διάστημα Εμπιστοσύνης για το γραμμικό υπόδειγμα

$$\left[\hat{Y}_0 - t_c \cdot se(f), \hat{Y}_0 + t_c \cdot se(f)\right]$$

$$= [201365 - 1.96267 * 30348.26, 201365 + 1.96267 * 30348.26]$$

$$= [141801,260928]$$

όπου $t_c = t_{a/2,n-2} = t_{0.025,878} = 1.96267$ *και*

$$se(f) = \hat{\sigma}\sqrt{\left[\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} + 1\right]} =$$

$$3519.2\sqrt{\left[\frac{1}{880} + \frac{(2700 - 1611.968)^2}{248768933.1} + 1\right]} = 30348.26$$

Ερώτημα vi

(συνέχεια)

- Διάστημα Εμπιστοσύνης για το ημιλογαριθμικό υπόδειγμα

$\left[exp\left(\hat{Y}_n - t_c \cdot se(f)\right), exp\left(\hat{Y}_n + t_c \cdot se(f)\right)\right] = [\exp(10.5938 + 0.000596 * 2700 - 1.96267 * 0.20363), \exp(10.5938 + 0.000596 * 2700 + 1.96267 * 0.20363)] = [133683{,}297315]$

όπου $t_c = t_{a/2,n-2} = t_{0.025,878} = 1.96267$ *και*

$se(f) = \hat{\sigma}\sqrt{\left[\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} + 1\right]} = 0.20303\sqrt{\left[\frac{1}{880} + \frac{(2700 - 1611.968)^2}{248768933.1} + 1\right]} = 0.20363$

Ερώτημα vi

(συνέχεια)

- Διάστημα Εμπιστοσύνης για το λογαριθμικό υπόδειγμα

$$\left[exp\left(\hat{Y}_n - t_c \cdot se(f)\right), exp\left(\hat{Y}_n + t_c \cdot se(f)\right)\right]$$
$$= [\exp(4.1707 + 1.0066\ln(2700) - 1.96267 * 0.20876), \exp(4.1707 + 1.0066\ln(2700) - 1.96267 * 0.20876)] = [122267, 277454]$$

όπου $t_c = t_{a/2,n-2} = t_{0.025,878} = 1.96267$ και

$$se(f) = \hat{\sigma}\sqrt{\left[\frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} + 1\right]}$$

$$= 0.2083\sqrt{\left[\frac{1}{880} + \frac{(\ln(2700) - 7.3355)^2}{85.34453} + 1\right]} = 0.20876$$

Ερώτημα vii

vii. Με βάση την ανάλυση σας ποια συναρτησιακή μορφή θα επιλέγατε για την σχέση μεταξύ της τιμής πώλησης και του εμβαδού? Εξηγείστε την επιλογή σας.

Λύση:

- Το γραμμικό υπόδειγμα δεν είναι μια καλή επιλογή καθώς τα κατάλοιπα εμφανίζουν ασσυμετρία προς τα δεξιά και επομένως δεν ακολουθούν την κανονική κατανομή
- Για τα άλλα 2 υποδείγματα τα κατάλοιπα εμφανίζουν παρόμοια συμπεριφορά (pattern), επίσης η ελαστικότητα της τιμής ως προς τις αλλαγές στο εμβαδόν είναι αρκετά πιστευτή, 1% αύξηση στο εμβαδό οδηγεί σε 1% της τιμής πώλησης.
- Το ημιλογαριθμικό υπόδειγμα ωστόσο έχει υψηλότερο R^2 και μικρότερο τυπικό σφάλμα της παλινδρόμησης/διαταρακτικού όρου και επομένως προτιμάτε από το λογαριθμικό αφού προσεγγίζει καλύτερα τα δεδομένα (ημιλογαριθμικό: $R^2 = 0.7094$, $\hat{\sigma} = 0.20303$; λογαριθμικό $R^2 = 0.6943$, $\hat{\sigma} = 0.2083$).

Περίληψη

- Ημιλογαριθμικός μετασχηματισμός
- Λογαριθμικός μετασχηματισμός

- Πρόβλεψη στο ημιλογαριθμικό & λογαριθμικό υπόδειγμα
- Διάστημα εμπιστοσύνης στο ημιλογαριθμικό & λογαριθμικό υπόδειγμα

Βιβλιογραφία

- ✓ Stock & Watson, κεφ. 8 (8.2)
- ✓ Gujarati & Porter, κεφ. 6
- ✓ Hill, Griffiths & Lim, Principles of Econometrics, 4th ed, κεφ. 4

Ευχαριστώ!

Ερωτήσεις?