

Εφαρμογές Η/Υ

Διδάσκων: Αναστάσιος Τασιόπουλος

Λογαριθμικά υποδείγματα: Log-level και Level-log

4.3 Λογαριθμικοί μετασχηματισμοί

4.3.1 Log-level μοντέλα

Στο log-level υπόδειγμα λογαριθμοποιείται μόνο η εξαρτημένη μεταβλητή. Στο παράδειγμα χρησιμοποιούμε τα δεδομένα **wage1** και εκτιμούμε το υπόδειγμα:

$$\log(wage) = \beta_0 + \beta_1 \cdot educ + \beta_2 \cdot tenure + e$$

- wage: μέσες ωριαίες αποδοχές 526 υπαλλήλων.
- educ: χρόνια σπουδών.
- tenure: χρόνια παραμονής του υπαλλήλου στην επιχείρηση.

Αποτελέσματα εκτίμησης

Μεταβλητή	Estimate	Std. Error	p-value
Intercept	0.404474	0.091696	1.25e-05
educ	0.086528	0.006991	< 2e-16
tenure	0.025814	0.002680	< 2e-16

R ²	Adj. R ²	F-statistic	Residual SE
0.3085	0.3059	116.7 (p < 2.2e-16)	0.4428

Συμπέρασμα. Οι συντελεστές των educ και tenure είναι στατιστικά σημαντικοί στο 5%, ενώ και το υπόδειγμα συνολικά είναι στατιστικά σημαντικό.

Ερμηνεία συντελεστών

- Στα log-level μοντέλα οι συντελεστές ερμηνεύονται ως προσεγγιστικές ποσοστιαίες μεταβολές της εξαρτημένης μεταβλητής.
- Αύξηση της educ κατά 1 έτος συνεπάγεται αναμενόμενη αύξηση της wage κατά περίπου $\beta_1 \cdot 100\% = 0.086528 \cdot 100\% = 8.65\%$.
- Αύξηση της tenure κατά 1 έτος συνεπάγεται αναμενόμενη αύξηση της wage κατά περίπου $\beta_2 \cdot 100\% = 0.025814 \cdot 100\% = 2.58\%$.

Σύντομη σύγκριση ερμηνείας

Μοντέλο	Ερμηνεία του β_1
$wage = \beta_0 + \beta_1 \cdot educ + \beta_2 \cdot tenure + e$ (level-level)	Μεταβολή της educ κατά 1 μονάδα μεταβάλλει την wage κατά β_1 μονάδες.
$\log(wage) = \beta_0 + \beta_1 \cdot educ + \beta_2 \cdot tenure + e$ (log-level)	Μεταβολή της educ κατά 1 μονάδα μεταβάλλει την wage περίπου κατά $\beta_1 \cdot 100\%$.

Προβλέψεις και διόρθωση μετασχηματισμού

Επειδή η εξαρτημένη μεταβλητή βρίσκεται σε λογαριθμική μορφή, η απλή εκθετοποίηση της πρόβλεψης υποεκτιμά συστηματικά τον αναμενόμενο μισθό στην αρχική κλίμακα. Για αυτό χρησιμοποιούμε διορθωτικό παράγοντα β^*_0 που εκτιμάται από βοηθητική παλινδρόμηση χωρίς σταθερά:

$$\hat{y}_{wage} = \beta^*_0 \cdot \exp(\hat{y}_{\log(wage)})$$

```
library(wooldridge)
ols1 = lm(log(wage) ~ educ + tenure, data = wage1)
ols_no_constant = lm(wage ~ exp(ols1$fit) - 1, data = wage1)
beta_star_0 = ols_no_constant$coef
```

Ο διορθωτικός συντελεστής είναι $\beta^*_0 \approx 1.123568$ (περίπου 1.124).

Πρόβλεψη για educ=6, tenure=15	fit	lwr	upr
Με διόρθωση	4.167696	3.739583	4.644820
Χωρίς διόρθωση	3.709342	3.328312	4.133993

Κύρια ιδέα. Στο log-level μοντέλο προβλέπουμε πρώτα το $\log(wage)$. Για να επιστρέψουμε σωστά στην κλίμακα του wage, δεν αρκεί το $\exp(\hat{y}_{\log})$ · χρειάζεται ο διορθωτικός παράγοντας β^*_0 .

4.3.2 Level-log μοντέλα

Στο level-log υπόδειγμα λογαριθμοποιείται τουλάχιστον μία ερμηνευτική μεταβλητή, ενώ η εξαρτημένη παραμένει στην αρχική της κλίμακα. Στο παράδειγμα εκτιμούμε:

$$wage = \beta_0 + \beta_1 \cdot \log(exper) + e$$

- wage: μέσες ωριαίες αποδοχές 526 υπαλλήλων.
- exper: συνολικά χρόνια εργασιακής εμπειρίας.

Αποτελέσματα εκτίμησης

Μεταβλητή	Estimate	Std. Error	p-value
Intercept	4.1202	0.3876	< 2e-16
log(exper)	0.7417	0.1479	7.29e-07

R ²	Adj. R ²	F-statistic	Residual SE
0.04579	0.04397	25.14 (p = 7.291e-07)	3.611

Ερμηνεία β_1 . Στα level-log μοντέλα μια αύξηση της ερμηνευτικής μεταβλητής κατά 1% μεταβάλλει την εξαρτημένη κατά περίπου $\beta_1/100$ μονάδες.

- Εδώ $\beta_1 = 0.7417$.
- Άρα αύξηση της exper κατά 1% αυξάνει την wage κατά περίπου $0.7417/100 = 0.00742$ δολάρια ανά ώρα.

Σύντομη σύγκριση ερμηνείας

Μοντέλο	Ερμηνεία του β_1
wage = $\beta_0 + \beta_1 \cdot \text{exper} + e$ (level-level)	Μεταβολή της exper κατά 1 μονάδα μεταβάλλει την wage κατά β_1 μονάδες.
wage = $\beta_0 + \beta_1 \cdot \log(\text{exper}) + e$ (level-log)	Μεταβολή της exper κατά 1% μεταβάλλει την wage κατά $\beta_1/100$ μονάδες.

Πρόβλεψη στο level-log υπόδειγμα

Επειδή η εξαρτημένη μεταβλητή δεν έχει λογαριθμοποιηθεί, δεν απαιτείται ειδική διόρθωση στις προβλέψεις. Χρησιμοποιούμε τον κλασικό τρόπο μέσω της predict().

```
library(wooldridge)
ols = lm(wage ~ log(exper), data = wage1)
predict(ols, list(exper = 20), interval = "confidence")
```

Πρόβλεψη για exper=20	fit	lwr	upr
Τιμή	6.342066	5.986828	6.697304

Τελικό συμπέρασμα. Ο αναμενόμενος μισθός ενός υπαλλήλου με 20 έτη συνολικής εργασιακής εμπειρίας εκτιμάται σε 6.342 δολάρια ανά ώρα.

Χρήσιμος κανόνας

Μορφή μοντέλου	Τι λογαριθμοποιείται	Πώς ερμηνεύεται ο συντελεστής
level-level	τίποτα	1 μονάδα στο X $\rightarrow$ β μονάδες στο Y
log-level	η εξαρτημένη Y	1 μονάδα στο X $\rightarrow$ περίπου $\beta \cdot 100\%$ στο Y
level-log	η ερμηνευτική X	1% στο X $\rightarrow$ περίπου $\beta/100$ μονάδες στο Y