Εφαρμοσμένη Οικονομετρία

Σπύρος Σκούρας
Ύδρας 28, 4ος όροφος
skouras@aueb.gr

Ώρες Γραφείου: Πέμπτη 5:30-6:30

Ύλη: Wooldridge, Introductory Econometrics, 2nd edition,

Χρονοσειρές: 10,11,12 (επιλογή από18)

Πάνελ: 13,14

Άλλα θέματα: (επιλογή από 15,16,17)

ΧΡΟΝΟΣΕΙΡΑ | ΣΤΟΧΑΣΤΙΚΗ ΑΚΟΛΟΥΘΙΑ:

Ακολουθία τυχαίων μεταβλητών δείκτης της οποίας είναι ο χρόνος, $\{y(t)\}_{t=1}^{T}$

$y(t)$ – τυχαία μεταβλητή y παρατηρούμενη την στιγμή t:

$$Y = \begin{bmatrix} y(1) \\ y(2) \\ . \\ y(t) \\ . \\ y(n) \end{bmatrix}$$

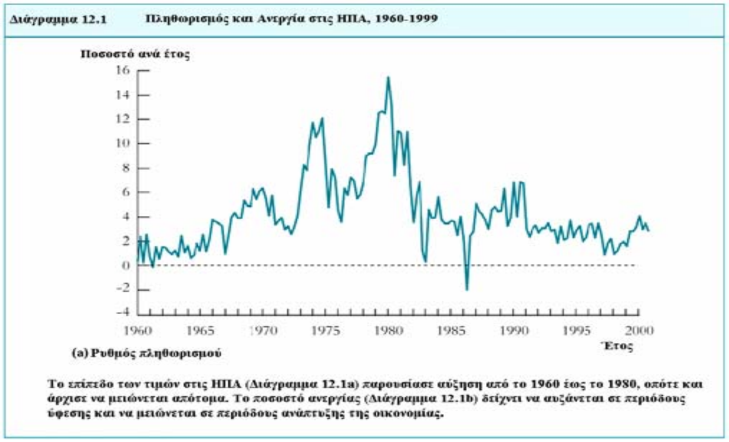
Στοιχεία χρονολογικών σειρών

Τα ***στοιχεία χρονολογικών σειρών*** συλλέγονται για την ίδια παρατηρήσιμη μονάδα, για πολλαπλές χρονικές περιόδους

- Συνολική Κατανάλωση και ΑΕΠ μιας χώρας (π.χ., τριμηνιαίες παρατηρήσεις για 20 χρόνια = 80 παρατηρήσεις).
- Συναλλαγματικές ισοτιμίες: Γεν/Δολάριο, Στερλίνα/Δολάριο και Ευρώ/Δολάριο: (ημερήσια στοιχεία για 1 χρόνο = 365 παρατηρήσεις).
- Κατά κεφαλήν κατανάλωση τσιγάρων σε μία πόλη.

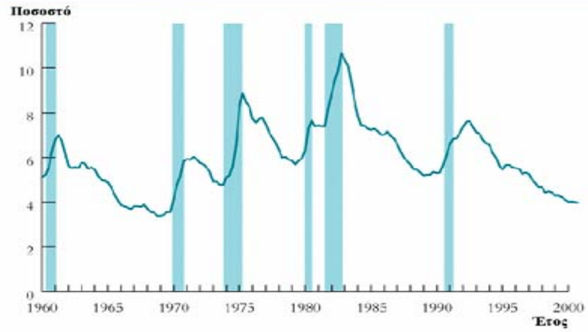
Διαγραμματική απεικόνιση χρονολογικών σειρών

Παράδειγμα 1: ο ρυθμός πληθωρισμού των ΗΠΑ



Παράδειγμα 2: το ποσοστό ανεργίας στις ΗΠΑ

Διάγραμμα 12.1 **Πληθωρισμός και Ανεργία στις ΗΠΑ, 1960-1999**



(b) Ποσοστό Ανεργίας

Το επίπεδο των τιμών στις ΗΠΑ (Διάγραμμα 12.1a) παρουσίασε αύξηση από το 1960 έως το 1980, οπότε και άρχισε να μειώνεται απότομα. Το ποσοστό ανεργίας (Διάγραμμα 12.1b) δείχνει να αυξάνεται σε περιόδους ύφεσης και να μειώνεται σε περιόδους ανάπτυξης της οικονομίας.

Γιατί χρησιμοποιούμε στοιχεία χρονολογικών σειρών;

- Για την ανάπτυξη υποδειγμάτων **πρόβλεψης**:
  - Ποιος αναμένεται να είναι ο ρυθμός πληθωρισμού τον επόμενο χρόνο;
- Για την εκτίμηση **δυναμικών αιτιωδών αποτελεσμάτων**:
  - Αν η Κεντρική Τράπεζα αυξήσει το επιτόκιο κατατεθειμένων διαθεσίμων σήμερα, πώς θα διαμορφωθούν τα επίπεδα πληθωρισμού και ανεργίας σε 3 μήνες από τώρα; σε 12 μήνες;
  - Ποιο είναι το διαχρονικό αποτέλεσμα στην κατανάλωση τσιγάρων από μια αύξηση της φορολογίας τσιγάρων;
- Άλλωστε, σε κάποιες περιπτώσεις δεν έχουμε άλλη επιλογή…
  - Τα επίπεδα πληθωρισμού και ανεργίας στις ΗΠΑ, μπορούν να παρατηρηθούν *μόνο* διαχρονικά.

Από τα δεδομένα χρονολογικών σειρών προκύπτουν νέα, τεχνικού χαρακτήρα ζητήματα:

- Χρονικές υστερήσεις.
- Διαχρονική συσχέτιση (σειριακή συσχέτιση ή αυτοσυσχέτιση).
- Υποδείγματα πρόβλεψης που δεν έχουν καμία αιτιώδη ερμηνεία (αποτελούν εξειδικευμένα εργαλεία για πρόβλεψη):
  - *Αυτοπαλίνδρομα υποδείγματα (σχήματα)* **(AR)**
  - *Αυτοπαλίνδρομα υποδείγματα κατανεμόμενων υστερήσεων* **(ADL)**
- Συνθήκες υπό τις οποίες μπορούν να εκτιμηθούν τα δυναμικά αποτελέσματα, καθώς και οι τρόποι με τους οποίους αυτά εκτιμώνται.
- Υπολογισμός των τυπικών σφαλμάτων, στην περίπτωση που αυτά εμφανίζουν αυτοσυσχέτιση.

Χρησιμοποίηση Υποδειγμάτων Παλινδρόμησης για τη Διενέργεια Προβλέψεων

- Η *πρόβλεψη* και η *εκτίμηση* των αιτιωδών αποτελεσμάτων είναι δύο τελείως διαφορετικές εργασίες.
- Όσον αφορά την πρόβλεψη,
  - ❖ο διορθωμένος συντελεστής προσδιορισμού $\bar{R}^2$ παίζει πολύ σημαντικό ρόλο!
  - ❖ η μεροληψία από παραλειπόμενες μεταβλητές δεν αποτελεί πρόβλημα!
  - ❖στα υποδείγματα πρόβλεψης δεν μας απασχολεί να ερμηνεύσουμε τους συντελεστές.
  - ❖η εξωτερική ισχύς κυριαρχεί: το εκτιμημένο με τα αρχικά στοιχεία υπόδειγμα πρέπει να ισχύει μέχρι το (τουλάχιστον εγγύς) μέλλον.

Στοιχεία Χρονολογικών Σειρών και Σειριακή Συσχέτιση: εισαγωγή

Θα πρέπει, πρώτα, να εισαγάγουμε κάποια ορολογία, αλλά και κάποιους συμβολισμούς.

Συμβολισμοί για τα στοιχεία χρονολογικών σειρών

•Y_t = η τιμή του Y την περίοδο t.

•Σύνολο στοιχείων: $Y_1,\ldots,Y_T$ = T παρατηρήσεις για την τ.μ. Y

• Εξετάζουμε μόνο:

- διαδοχικές,
- κατανεμημένες ανά ίσα χρονικά διαστήματα παρατηρήσεις,
- πχ κατά μήνα, από το 1960 έως το 1999, δίχως να παραλείπουμε μήνες (αλλιώς παλινδρόμηση περιπλέκεται).

ΥΠΟΔΕΙΓΜΑΤΑ ΓΙΑ ΧΡΟΝΟΣΕΙΡΕΣ

1. Στατικό υπόδειγμα: $y(t)=\beta(0)+\beta(1)z(t)+u(t),\ t=1,2,..n$

Χρήση όταν:

- πιστεύουμε ότι το $z(t)$ δεν επηρεάζει το $y(t+1)$
- μας ενδιαφέρει η σχέση $y(t)$ με $z(t)$ (και όχι άλλων μεταβλητών / υστερήσεων)

Παράδειγμα 1: Αν $u(t+1)=u(t) \Longrightarrow \Delta y(t)=\beta \Delta z(t)$

$(\Delta y(t)=y(t+1)-y(t))$

Ερμηνεία: «Ceteris Paribus» *οι μεταβολές στο y οφείλονται στις μεταβολές του z*

Παράδειγμα 2: Καμπύλη Phillips:

- y(t) – πληθωρισμός
- z(t) – ανεργία

Όμοιο με διαστρωματικό στατικό υπόδειγμα (cross-section)

2. Υπόδειγμα με πεπερασμένο αριθμό υστερήσεων (finite distributed lag):

- $y(t)=\beta(0)+\beta(1)z(t)+\beta(2)z(t-1)+\beta(3)z(t-2)$
~ (FDL 2ου βαθμού)
- $y(t)=\beta(0)+\beta(1)z(t)+\sum_{k=1}^{q}\beta(k+1)z(t-k)$
~ (FDL q βαθμού)

Ερμηνεία παραμέτρων $\Rightarrow$

Ερμηνεία παραμέτρων (q=2):

i. Συνέπειες προσωρινής μεταβολής z:

$$\left.\begin{array}{r} z(t)=c\forall t\neq t^* \\ z(t^*)=c+1 \\ u(t)=0\forall t \end{array}\right\} \Rightarrow \left\{\begin{array}{c} y(t^*)-y(t^*-1)=\beta(1) \\ y(t^*+1)-y(t^*-1)=\beta(2) \\ y(t^*+2)-y(t^*-1)=\beta(3) \end{array}\right\}$$

=> *β(k)* μετράει το *αναμενόμενο* αποτέλεσμα της μοναδιαίας προσωρινής μεταβολής του *z(t)* στο *y(t+k-1)*

ii: Συνέπειες μόνιμης μεταβολής z

$$\left.\begin{array}{r} z(t)=c\forall t<t^* \\ z(t^*)=c+1\forall t\geq t^* \\ u(t)=0\forall t \end{array}\right\} \Rightarrow \begin{array}{c} y(t^*+l)-y(t^*-1)=\beta(1)+\beta(2)+\beta(3), \\ l\geq 2 \end{array}$$

Άρα (β(1)+β(2)+β(3)) μετράει τον αναμενόμενο *μακροπρόθεσμο* αντίκτυπο στο $y(t)$ ως αποτέλεσμα μιας μοναδιαίας μόνιμης μεταβολής του $z(t)$

iii. Για $\beta(2) = \beta(3) = ... = \beta(q) = 0$ έχουμε το στατικό υπόδειγμα (υποπερίπτωση)

Χρήση όταν πιστεύουμε ότι το $z(t)$ επηρεάζει το $y(t+k)$

Παράδειγμα 1: $y(t)$: Βαθμός τεκνοποίησης

$z(t)$: επιδοτήσεις / φοροαπαλλαγές σε οικογένειες με παιδιά

Παράδειγμα 2: $y(t)$: ΑΕΠ

$z(t)$: Δαπάνες επενδύσεων

ΙΔΙΟΤΗΤΕΣ ΕΚΤΙΜΗΤΩΝ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ

Υπόθεση 1 (Υ1):	$Y=X\beta+u$	(γραμμικότητα)

$$\Leftrightarrow y(t)=\beta(0)+\beta(1)x(t,1)+\ldots+\beta(k)x(t,k)+u(t)\forall t$$

$$\text{π.χ. 1} \quad \left\{\begin{array}{l} x(t,1)=z(t) \\ x(t,2)=z(t-1) \\ x(t,3)=z(t-2) \end{array}\right\} \quad \text{(FDL(2))}$$

Χ – Πίνακας με *T* σειρές και *k* στήλες

Υπόθεση 2 (Υ2): *E(u(t)|X)=0* (αυστηρή εξωγένεια)

Ερμηνεία:

- $\left.\begin{array}{r}\Pr(u \le \Psi \mid X) = \Pr(u \le \Psi) \\ E(u) = 0\end{array}\right\} \Rightarrow \text{Υ2}$

- $\text{Υ2} \Rightarrow Cov(u(t), X) = 0 \forall t$

$\Leftrightarrow \; Cov(u(t), x(s, j)) = 0 \forall t, s, j$

$\Rightarrow Cov(u(t), x(t, j)) = 0 \forall t, j$ Σύγχρονη εξωγένεια

- Αληθοφάνεια Υ2 αμφισβητήσιμη λόγω:

1. Σφάλματα μετρήσεων του *X*
2. Παραληφθείσες μεταβλητές
3. Επιδράσεις του *y(t)* στο *x(t+λ)* (συνηθισμένες σε κοινωνικά φαινόμενα).

π.χ. 2: Αν *y(t)=β(0)+β(1)z(t)+β(2)z(t-1)+β(3)z(t-2)+u(t),*

E(u|X)=0

παραληφθείσα υστέρηση

και υποθέσουμε ότι:

y(t)=β(0)+β(1)z(t)+β(2)z(t-1)+u'(t),

τότε: *u'(t)=β(3)z(t-2)+u(t)*

=> *E(u'(t)|X)* ≠ *0 για x(t-1),x(t-2)*

π.χ. 3: *y(t):* φόνοι

z(t): αστυνομία

Υπόθεση 3 (Υ3): Όχι πλήρης συγγραμμικότητα

$\nexists j : x(t, j) = a \forall t$

$\nexists \beta^* : X(j) = X(-j)\beta^*,$

$X(-j) = [X(1), X(2), \ldots, X(j-1), X(j+1), \ldots, X(k)]$

$X(j) = [x(1, j), \ldots x(T, j)]'$

Θεώρημα 1 (Θ1) (Αμεροληψία Ε.Ε.Τ) :

Υ1, Υ2, Υ3 $\Rightarrow E(\hat{\beta}(j)) = \beta(j),\; j = 0,1,\ldots,k$

$\hat{\beta}$: εκτιμητής ελαχίστων τετραγώνων

Wooldridge, κεφ. 3 (απόδειξη-άσκηση)

Υπόθεση 4 (Υ4): Ομοσκεδαστικότητα:

$Var(u(t)|X)=\sigma^2 \forall t$

Ερμηνεία:

- $\Pr(u(t)\leq \Psi \mid X) = \Pr(u(t)\leq \Psi) \Rightarrow$ Υ4 (όπως και Υ2)
- Υ4 $\Rightarrow Var(y(t)\mid X) = Var(x(t)\beta \mid X) + \sigma^2$
- Ανεξάρτητες και εξαρτημένες μεταβλητές *μπορούν* να είναι ετερ/στικές

Ερμηνεία και αληθοφάνεια

- *X* πρέπει να συμπεριλαμβάνει *όλους* τους παράγοντες που επηρεάζουν τη διακύμανση του *Y (παρομοίως με εξωγένεια Υ2)*
- Διακύμανση *Y* δεν πρέπει να επηρεάζεται από επίπεδο *X*.

π.χ.

$$\left\{\begin{array}{l} Y: \text{επιτόκια} \\ X(1): \text{πληθωρισμός} \\ X(2): \text{έλλειμμα} \end{array}\right\}$$

=> μη ρεαλιστικό

(π.χ. γιατί $u(t)$ συμπεριλαμβάνει *κυβερνητική πολιτική* η οποία (1) έχει έντονο αντίκτυπο στην διακύμανση επιτοκίων, (2) επηρεάζεται από X)

Υπόθεση 5 (Υ5): Έλλειψη αυτοσυσχέτισης:

$Corr(u(t),u(s)|X)=0 \ \forall\ t\neq s$

Ερμηνεία/παρατηρήσεις:

- Υ5 $\Rightarrow Corr(u(t),u(s))=0$
- $y(t),z(t)$ είναι τυχαίο δείγμα (κλασσική υπόθεση σε υποδείγματα για διαστρωματικά στοιχεία) $\Rightarrow$ Υ5
- Δεν περιορίζει τη συμπεριφορά του X

Θεώρημα 2 (Θ2) (Διακυμάνσεις Ε.Ε.Τ) :

(Υ1-5)

$$\Rightarrow Var\left(\hat{\beta}(j)\,|\,X\right)=\sigma^2\left[\sum_{t=1}^{n}\left(x(t,j)-\bar{x}(j)\right)^2\left(1-R^2(j)\right)\right]^{-1}$$

$$R^2(j)=\frac{\sum_{t=1}^{n}\left(x(t,-j)\hat{\beta}(-j)-x(t,j)\right)^2}{\sum_{t=1}^{n}\left(x(t,j)-\bar{x}(j)\right)^2}$$

Άθροισμα τετραγώνων παλινδρόμησης *x(j)* στα άλλα *x* (*με σταθερό όρο*)

Άθροισμα τετραγώνων *x(j)*

Θεώρημα 3 (Θ3) (Αμερόληπτη Εκτίμηση Διακύμανσης Σφάλματος / Διαταρακτικού Όρου) :

$$(\text{Υ1-5}) \Rightarrow \hat{\sigma}^2=\frac{\sum_{t=1}^{n}\hat{u}_t^2}{n-k-1}$$

n-k-1: βαθμοί ελευθερίας (degrees of freedom)

Θεώρημα 4 (Θ4) (Gauss-Markov) :

(Υ1-5) ⇒ Ε.Ε.Τ. είναι {Άριστοι, Γραμμικοί, Αμερόληπτοι} δεδομένου του *X* (BLUE)

Υπόθεση 6 (Υ6): u(t)~N(0,σ²) (κανονικότητα)

Θεώρημα 5 (Θ5) (Κανονικές Κατανομές Ε.Ε.Τ.) :

$$(\text{Υ1-6}) \Rightarrow \quad \hat{\beta}\sim N\left(\beta,\sigma_{\beta}^2\,|\,X\right)$$

$$\frac{\hat{\beta}(j)-\beta}{\sqrt{\hat{\sigma}_{\beta}^2}}\sim t_{n-k-1}$$

$$\frac{(\text{ΑΤΠΠ}-\text{ΑΤΠ})/q}{\text{ΑΤΠ}/(n-k-1)}\sim F_{q,n-k-1}$$

- (*q*, *n-k-1*): βαθμοί ελευθερίας
- ΑΤΠΠ: Άθροισμα τετραγώνων περιορισμένης παλινδρόμησης
- *q*: αριθμός περιορισμών

Παράδειγμα 10.1: Καμπύλη Phillips

Υποθέτουμε Υ1-6 για y: πληθωρισμό, x: ανεργία

- ΗΠΑ ετήσια στοιχεία (1948-1996)

Με Ε.Ε.Τ καταλήγουμε στο υπόδειγμα:

$\widehat{\text{ΠΛΗΘ}}(t) = 1.42 + 0.468\text{ΑΝ}(t)$

(1.72) (0.289)

n = 49, R^2= 0.053, $\bar{R}^2$= 0.033

▪ Θέλουμε να εξετάσουμε αν υπάρχει αντίστροφη σχέση μεταξύ (y, x)

▪ Ο έλεγχος μπορεί να γραφεί ως: H_0: β(1)=0
H_1: β(1)<0

- Αλλά $\widehat{\beta(1)}>0$
- $t_{\widehat{\beta(1)}}$= 0.468/0.289 = 1.62, που με (n-k-1) = 47 β.ε. δίνει μόνο 11% πιθανότητα ότι β(1)=0

▪ Τα συμπεράσματα αυτά ισχύουν μόνο αν ισχύουν οι υποθέσεις Υ1-6

Παράδειγμα 10.2: Επιτόκια, Πληθωρισμός και Έλλειμμα

ι3: 3μηνιαία επιτόκια T-Bill

$\widehat{\iota 3}$ = 1.25 + 0.613ΠΛΗΘ + 0.7ΕΛΛΕΙΜΜΑ
(0.44) (0.076) (0.118)

n = 49, R^2= 0.697, $\bar{R}^2$= 0.683

- και τα δύο αυξάνουν τα επιτόκια
- στατιστικά σημαντικό
- οικονομικά σημαντικό
(1% ⇧ πληθωρισμός ⟹ 0.6% ⇧ επιτόκια)

Παράδειγμα 10.3

$$\log(\%\ \text{Απασχόληση}_t) = \underset{(0.77)}{-1.05} \underset{(0.65)}{-1.54}\log\left(\frac{\text{Μέσος Ελάχιστος Μισθός}}{\text{Μέσος Μισθός}}\times\%\text{Κάλυψης}\right)_t \underset{(.089)}{-0.12}\log(\text{ΑΕΠ}^{US}_t)$$

n = 38, R^2 = .661, $\bar{R}^2$ = .641, Πόρτο Ρίκο: 1950-1987

⇨ β_1 = ελαστικότητα
$t_{\hat{\beta}_1} = -2.37$

ΕΡΜΗΝΕΙΑ ΥΠΟΔΕΙΓΜΑΤΩΝ

- (Παράδειγμα 10.3): Λογαριθμικό

$$\log(y_t) = \beta_0 + \beta_1 \log(z_t) + u_t$$

$$\beta_1 = \frac{d\log(y_t)}{d\log(z_t)} = \frac{dy_t/y_t}{dz_t/z_t} \cong \frac{\%\text{ΑΛΛΑΓΗ } Y}{\%\text{ΑΛΛΑΓΗ } Z} \equiv \text{ ελαστικότητα}$$

Αν $Z = X_1/X_2$ Y - ανεργία

π.χ. X_1: μισθοί, X_2: τιμές, τότε Z = πραγματικοί μισθοί

$$\log(y_t) = \beta_0 + \beta_1\log(x_{1t}) + \beta_2\log(x_{2t}) + u_t$$

$\beta_2 = -\beta_1$ ⟹ αλλιώς **ονομαστικές** τιμές επηρεάζουν ανεργία

- <u>Λογαριθμικό με υστερήσεις:</u>

$$\log(y_t) = \beta_0 + \beta_1 \log(z_t) + \beta_2 \log(z_{t-1}) + \ldots + \beta_k \log(z_{t-k+1}) + u_t$$

$\beta_1 \equiv$ βραχυπρόθεσμη ελαστικότητα

$\sum_{i=1}^{k} \beta_i \equiv$ μακροπρόθεσμη ελαστικότητα

$\Rightarrow$ Αλλαγή k περιόδους μετά από μια μόνιμη αλλαγή

- <u>(10.4): Dummy Variables (Ψευδομεταβλητές)</u>

  - Χρησιμοποιούνται για να απομονώσουν «σημαντικές» περιόδους

$$y_t = \beta_0 + \beta_1 z_t + \beta_2 d_t + u_t$$

$$d_t = \begin{cases} 1, \text{ αν } t \in [t_1, t_2] \\ 0, \text{ διαφορετικά} \end{cases}$$

Παράδειγμα 10.4

$$\text{Γεννητικότητα}_t = \underset{(3.21)}{98.68} + \underset{(.030)}{.083}\text{Απαλλαγές}_t \underset{(7.46)}{-24.24}\text{ΔΠΠ}_t \underset{(4.08)}{-31.59}\text{Χάπι}_t$$

$n = 72$, $R^2 = .473$, $\bar{R}^2 = .450$: 1913-1984

$$\text{ΔΠΠ}_t = \begin{cases} 1 \text{ για } t \in [41,45] \\ 0 \text{ διαφορετικά} \end{cases}$$

$$\text{Χάπι}_t = \begin{cases} 1 \text{ για } t \geq 63 \\ 0 \text{ για } t \leq 63 \end{cases}$$

Παράδειγμα 10.4

$$\text{Γεννητικότητα}(t) = \underset{(3.28)}{95.87} + \underset{(.126)}{.073}\text{Απαλλαγές}(t) - \underset{(.1557)}{.0058}\text{Απ}(t-1) + \underset{(.126)}{.034}\text{Απ}(t-2) - \underset{(10.73)}{22.13}\text{ΔΠΠ}(t) - \underset{(3.98)}{31.3}\text{Χάπι}(t)$$

ΔΠΠ = 2ος Παγκόσμιος

$n = 70$, $R^2 = 0.499$, $\bar{R}^2 = 0.459$

- H_0: β(1) = β(2) = β(3) = 0: F-test: p =.012

  β(1), β(2), β(3) μοιάζουν ασήμαντες **?**

- $\widehat{\beta(1)} + \widehat{\beta(2)} + \widehat{\beta(3)} \cong .101$

Για να εκτιμήσουμε την τυπική απόκλιση του μακροπρόθεσμου αποτελέσματος (β(1)+β(2)+β(3)) παλινδρομούμε:

Γεννητικότητα(t) = β(0) + (β(1)+β(2)+β(3))Απ(t)

+ β(2)(Απ(t-1)-Απ(t))

+ β(3)(Απ(t-2)-Απ(t))

+ . . .

$\hat{\sigma}_{\hat{\beta}(1)+\hat{\beta}(2)+\hat{\beta}(3)} = .03 \Rightarrow t_{\beta(1)+\beta(2)+\beta(3)} = 3.37 \Rightarrow$ Μακροπρόθεσμο αποτέλεσμα σημαντικό

• <u>(10.5): Μελέτες Γεγονότων</u>

- Όπως οι D.V. αλλά δίνεται μεγαλύτερο βάρος στο τι γίνεται πριν και μετά από το γεγονός

Παράδειγμα 10.5

log(Εισαγωγές_Κίνα(t))= -17.8 + 3.12log(Παρ._Χημικών(t))
(21.05) (0.48)

+.196log(Παρ._Βενζίνης(t))+.983log($(t))
(.907) (.400)

+.060(Πριν_Κατάθεση(t)) -.032(Κατάθεση(t))
(.261) (.264)

-.566(Μετά(t))
(.286)

n = 131, R^2= 0.305, $\bar{R}^2$= 0.271 (Φεβ. ’78-Δεκ. ’88)

$$\text{Κατάθεση(t)} = \begin{cases} 1 \text{ για } t \in [\textbf{\underline{Οκτ. ’83}}, \text{Μαρ. ’84}] \\ 0 \end{cases}$$

$$\text{Πριν\_Κατάθεση(t)} = \begin{cases} 1 \text{ για } t \in [\text{Μαρ. ’83, Σεπ. ’83}] \\ 0 \end{cases}$$

$$\text{Μετά(t)} = \begin{cases} 1 \text{ για } t \in [\textbf{\underline{Οκτ. ’84}}, \text{Μαρ. ’85}] \\ 0 \end{cases}$$

• Κατάθεση(t) $\Longrightarrow$ Ασήμαντη

• Πριν_Κατάθεση(t) $\Longrightarrow$ Ασήμαντη

• Μετά(t) $\Longrightarrow$ Στατιστικά σημαντική.
Οικονομική σημασία:
Μεταβολή κατά 100[exp(-.566)-1]
$\cong$ -43.2%

• (10.6): Συνδυασμός και με ποσοτικές μεταβλητές

Παράδειγμα 10.6

$$\underset{}{Δημοκρατικό\ \%_t} = \underset{(0.12)}{.481} - \underset{(.0405)}{.0435}Εκλεγμένοι_t + \underset{(0.134)}{.0544}Εκλεγμένος_t$$

$$+\underset{(.0041)}{.0108}Εκλεγμένοι \times Καλά\_Νέα_t$$

$$-\underset{(.0033)}{.0077}Εκλεγμένοι \times Πλήθ_t$$

n = 20 [Τετραετίες 1966-92], $\bar{R}^2$= 0.663, R^2= 0.573

$$Εκλεγμένοι = \begin{cases} 1 \text{ αν Δημοκρατικοί} \\ -1 \text{ αν Ρεπουμπλικάνοι} \end{cases}$$

$$Εκλεγμένος = \begin{cases} 1 \text{ αν είναι ο Δ υποψήφιος} \\ -1 \text{ αν είναι ο Ρ υποψήφιος} \\ \text{Αλλιώς } 0 \end{cases}$$

Καλά_Νέα = # τριμήνων (από 15) με ανάπτυξη > 2.9%

Πληθωρισμός = Μέσος 15 τριμήνων

Προβλέψεις

Κλίντον – Ντόουλ, 1996:

Εκλεγμένοι = 1

Εκλεγμένος = 1

Καλά_Νέα = 3

Πληθ. = 3.019

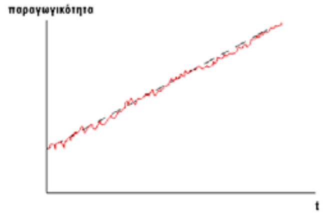
⇒ Δημ.% ≈ 50.11%

Αποτέλεσμα 54.65%

ΧΡΟΝΟΣΕΙΡΕΣ ΜΕ ΤΑΣΕΙΣ

Υποδείγματα για $\{y_t\}_{t=1}^{T}$:

1. Γραμμική Τάση: $y_t = a_0 + a_1 t + e_t, t = 1,2,....$



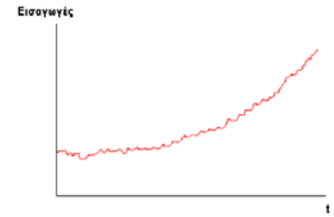
Ερμηνεία:

$a_1 = y_{t+1} - y_t$ αν $e_t = 0$

$a_1 > 0 \Rightarrow E(y_t) = a_0 + a_1 t$ $\Uparrow$ με το χρόνο

$a_1 < 0 \Rightarrow$ >> $\Downarrow$ με το χρόνο

2. Εκθετική τάση: $\log(y_t) = \beta_0 + \beta_1 t + e_t, t = 1,2,....$



Ερμηνεία:

$$\beta_1 \cong \Delta \log(y_t) \cong \frac{y_t - y_{t-1}}{y_t}$$

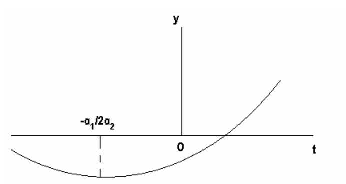
$\boxed{\text{Ρυθμός αύξησης}}$

3. Δευτεροβάθμια τάση:

$$y_t = a_0 + a_1 t + a_2 t^2 + e_t, t = 1,2,....$$

$$\frac{\Delta y_t}{\Delta t} = a_1 + 2a_2 t$$

Αν $a_1, a_2 > 0$



ΧΡΗΣΗ ΧΡΟΝΟΣΕΙΡΩΝ ΜΕ ΤΑΣΕΙΣ ΣΕ ΠΑΛΙΝΔΡΟΜΗΣΕΙΣ: ΑΠΟΤΕΛΕΣΜΑΤΑ ΠΑΡΑΛΕΙΨΗΣ

Παράδειγμα

- Πραγματικότητα: $y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t} + u_t$
  $x_{3t} = t$
- Υπόδειγμα: $y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t^{'}$

⇨ Κλασικές υποθέσεις παραβιάζονται (ενδογένεια, αυτοσυσχέτιση)

⇨ Εκτιμητές γίνονται μεροληπτικοί

⇨ Αν (και οι δύο) μεταβλητές *y* και *x* έχουν τάση, η μεροληψία θα είναι ιδιαίτερα έντονη

ΕΡΜΗΝΕΙΑ ΥΠΟΔΕΙΓΜΑΤΟΣ ΜΕ ΤΑΣΗ ΩΣ ΥΠΟΔΕΙΓΜΑ ΜΕΤΑΒΛΗΤΩΝ ΜΕ ΑΦΑΙΡΕΜΕΝΗ (ΑΠΟΜΟΝΩΜΕΝΗ) ΤΑΣΗ

Αφαίρεση τάσης: $\ddot{y}_t = y_t - \hat{a}_0 - \hat{a}_1 t$

$$\ddot{x}_t = x_t - \hat{a}_0^{'} - \hat{a}_1^{'} t$$

Μπορούμε να δείξουμε ότι η εκτίμηση οποιασδήποτε από τις δύο παλινδρομήσεις:

$$\ddot{y}_t = \beta_1 \ddot{x}_{1t} + \beta_2 \ddot{x}_{2t} + u_t$$

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 t + u_t$$

δίνει τα ίδια β.

⇒ *Η τάση πρέπει να συμπεριλαμβάνεται και όταν εμφανίζεται μόνο στο x_i (Αλλιώς μπορεί να μην αναγνωριστεί η σημασία του x_i)*

Παράδειγμα 10.7:

log(*Επενδύσεις_Σπίτια*) = -.550 + 1.241×log(*Τιμές_Σπίτια*)
(.043) (.382)

ΗΠΑ: 1947-’88, n = 42, R^2= 0.208, $\bar{R}^2$= 0.189

log(*Επενδύσεις_Σπίτια*) = .0081×*t*
(.0018)

Προβληματικά
(Αυτοσυσχέτιση)

log(*Τιμές_Σπίτια*) = .0044×*t*
(.0004)

⇒ **Υπάρχουν τάσεις**

log(*Επενδύσεις_Σπίτια*) = -.913 – .381×log(*Τιμές_Σπίτια*)
(.136) (.679)

+.0098×*t*
(.0035)

n = 42, R^2= .341

- Τιμές ασήμαντες
- Επενδύσεις ⇧ 1%/έτος

ΕΡΜΗΝΕΙΑ ΤΟΥ R^2 ΟΤΑΝ Η ΕΞΑΡΤΗΜΕΝΗ ΜΕΤΑΒΛΗΤΗ ΕΧΕΙ ΤΑΣΗ

- Η τάση συχνά εξηγεί μεγάλο μέρος των μεταβολών
- Για να απομονώσουμε τη σημασία άλλων μεταβλητών υπολογίζουμε το R^2 σε παλινδρόμηση με αφαιρεμένη τάση:

$$\ddot{y}_t = \beta_0 + \beta_1 x_t + \beta_2 t + u_t$$

log(*Επενδύσεις_Σπίτια*) = -.913 – .381×log(*Τιμές_Σπίτια*)
(.136) (.679)

+.0098×*t*
(.0035)

n = 42, R^2= .341

Αν όμως αφαιρέσουμε την τάση πριν την παλινδρόμηση

(10.10): log(*Επενδύσεις_Σπίτια με ΑΦΑΙΡΕΜΕΝΗ τάση*)
=*β0*+*β1*×log(*Τιμές_Σπίτια*) +.*β2*×*t*

⟹ ***R^2 = .08***

ΕΠΟΧΙΚΟΤΗΤΑ

- Συχνά οι εποχικότητες είναι ήδη αφαιρεμένες
- Αφαίρεση εποχικοτήτων:

π.χ. Μηνιαίες:

$$y_t = \beta_0 + \beta_1 \Phi\varepsilon\beta_t + \beta_2 \mathrm{M}\alpha\rho_t + \ldots + \beta_{11}\Delta\varepsilon\kappa_t + \ddot{y}_t \quad \textbf{(!)} \Leftrightarrow u_t$$

Χρησιμοποιούμε, $\ddot{y}_t$, αφού εκτιμήσουμε τα $\{\beta_0, \beta_1, \beta_2, .., \beta_{11}\}$

- Ίδια ζητήματα ερμηνείας με αφαίρεση τάσης

ΕΠΟΧΙΚΟΤΗΤΑ

- Συχνά οι εποχικότητες είναι ήδη αφαιρεμένες
- Αφαίρεση εποχικοτήτων:

π.χ. Μηνιαίες:

$$y_t = \beta_0 + \beta_1 \Phi\varepsilon\beta_t + \beta_2 \mathrm{M}\alpha\rho_t + \ldots + \beta_{11}\Delta\varepsilon\kappa_t + \ddot{y}_t \quad \textbf{(!)} \Leftrightarrow u_t$$

Χρησιμοποιούμε, $\ddot{y}_t$, αφού εκτιμήσουμε τα $\{\beta_0, \beta_1, \beta_2, .., \beta_{11}\}$

- Ίδια ζητήματα ερμηνείας με αφαίρεση τάσης

ΓΕΝΙΚΕΥΣΗ ΥΠΟΘΕΣΕΩΝ 1-6: ΜΕΛΕΤΗ ΑΣΥΜΠΤΩΤΙΚΩΝ ΙΔΙΟΤΗΤΩΝ ΕΕΤ

Βασικοί Ορισμοί:

(Αυστηρά) Στάσιμη Στοχαστική Ακολουθία/Χρονοσειρά:

Η Σ.Α. $\{y_t$: t = 1, 2, …} είναι *αυστηρά στάσιμη* αν

$\forall t_1, t_2, .. t_m$ η κατανομή $D_{y_{t_1}, y_{t_2}, \ldots y_{t_m}}$

είναι ίδια με την κατανομή $D_{y_{t_1+h}, y_{t_2+h}, \ldots y_{t_m+h}}$

- Ιδιότητα όλης της ακολουθίας και όχι μιας παρατήρησης
- Μπορεί να είναι δύσκολο να κρίνουμε αν μια Σ.Α. είναι στάσιμη
- Κάποιες χρονοσειρές (π.χ. με τάσεις) είναι φανερά <u>μη στάσιμες</u>

$\Longrightarrow$ Πιο ασθενής μορφή:

Στοχαστική Ακολουθία με Στάσιμη Συνδιακύμανση:

Η Σ.Α. $\{y_t$: $t = 1, 2, \ldots\}$ με πεπερασμένη διακύμανση, $E(y_t^2) < \infty \forall t$, έχει *στάσιμη συνδιακύμανση* αν

$$1) E(y_t) = k_1$$
$$2) \operatorname{var}(y_t) = k_2$$
$$3) \operatorname{cov}(y_t, y_{t+h}) = f(h) \forall (h, t)$$

Α.Σ. + $E(y_t^2) < \infty \Rightarrow$ Σ.Σ.	
Σ.Σ.	$\not\Rightarrow$ Α.Σ.

Στοχαστική Ακολουθία με Ασθενή Εξάρτηση:

Μια Σ.Α. έχει *Ασθενή Εξάρτηση* αν y_t και y_{t+h} είναι «σχεδόν ανεξάρτητες» καθώς το $h \to \infty$

Παράδειγμα ενός είδους Ασθενούς Εξάρτησης:

Μια Σ.Α. με στάσιμη συνδιακύμανση έχει *Ασυμπτωτική Έλλειψη Συσχέτισης* **(ΑΕΣ)** αν η συσχέτιση $Corr(y_t, y_{t+h})$ μικραίνει «αρκετά γρήγορα» καθώς $h \to \infty$:

«αρκετά γρήγορα», σημαίνει για παράδειγμα:

$$\exists \rho(h): Corr(y_t, y_{t+h}) \le \rho(h), 0 \le \rho(h) \le 1, \sum_{h=1}^{\infty} \rho(h) < \infty$$

$$\Rightarrow \lim_{h \to \infty} Corr(y_t, y_{t+h}) = 0 \quad \text{(ΑΕΣ)}$$

ΧΡΗΣΙΜΟΤΗΤΑ

$$\left\{ \begin{array}{c} \textbf{ΣΤΑΣΙΜΟΤΗΤΑΣ} \\ \textbf{ΑΣΘΕΝΟΥΣ ΕΞΑΡΤΗΣΗΣ} \end{array} \right\}$$

ΩΣ ΥΠΟΘΕΣΕΙΣ

Απαραίτητες για:
- Θεωρήματα Κεντρικού Ορίου
- Νόμους Μεγάλων Αριθμών
- Μαθησιμότητα σχέσεων από μια παρατήρηση χρονοσειράς

ΣΗΜΑΝΤΙΚΑ ΥΠΟΔΕΙΓΜΑΤΑ

1. Υπόδειγμα Κινητού Μέσου Πρώτου Βαθμού, MA(1):

$$y_t = e_t + a_1 e_{t-1}, t = 1,2,\ldots$$
$$e_t \sim D(0, \sigma^2)$$

Για $\alpha_1 = 0$ έχουμε το υπόδειγμα «Λευκού Θορύβου»

$$Corr(y_{t+1}, y_t) = a_1 / (1 + a_1^2) \quad \text{όχι } f(t) \quad \text{Στάσιμη Συνδ.}$$
$$Corr(y_{t+2}, y_t) = 0 \quad \text{Ασθενής Εξάρτηση}$$
$$D_{y_t} \text{ ίδια } \forall t \quad \text{Αυστηρή Στασιμότητα}$$

2. Αυτοπαλίνδρομο Υπόδειγμα Πρώτου Βαθμού, AR(1)

$$Y_t = \rho_1 Y_{t-1} + e_t, t = 1,2,\ldots$$
$$e_t \sim D(0, \sigma^2)$$

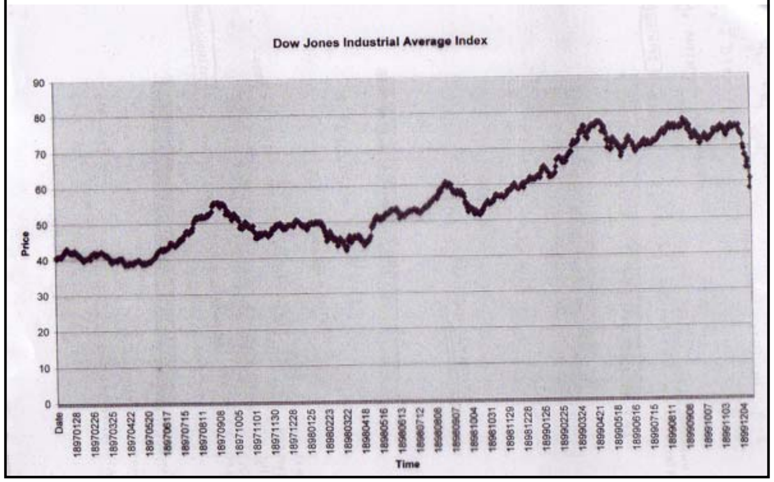
Μπορούμε να δείξουμε (Wooldridge, σελ.350-1):

$$Corr(y_{t+h}, y_t) = \rho^h{}_1$$

Αν $\rho_1 = 1$ ⟹ Υπόδειγμα Τυχαίας Διαδρομής

Αν $|\rho_1| < 1$ ⟹ Αυστηρά ασυσχέτιστο

Παράδειγμα χρονοσειράς που συχνά αντιμετωπίζεται σαν τυχαία διαδρομή (αντιπαράδειγμα στις υποθέσεις)

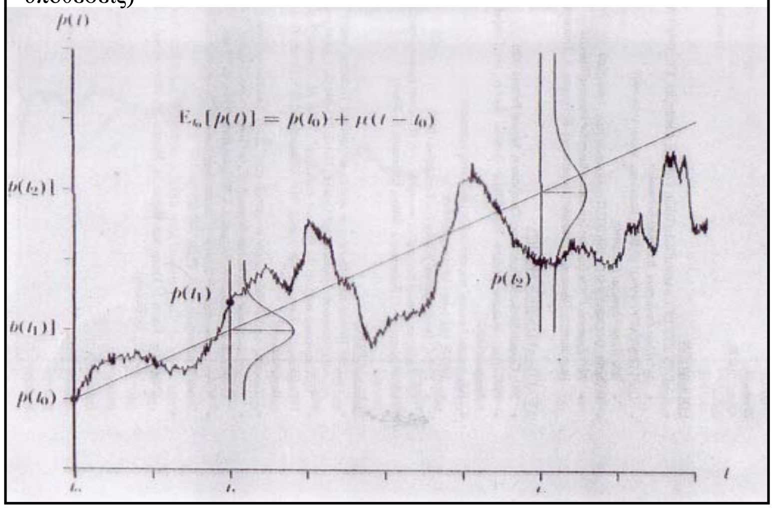


3. Σ.Α. με Τάση και Στασιμότητα και Ασθενή Εξάρτηση:

Σ.Α. της οποία αν της αφαιρέσουμε την τάση είναι στάσιμη και έχει ασθενή εξάρτηση.

π.χ.
$$Y_t = a_0 + a_1 t + e_t, t = 1,2,\ldots$$
$$e_t \sim D(0, \sigma^2)$$

Παράδειγμα τυχαίας διαδρομής με τάση (αντιπαράδειγμα στις υποθέσεις)



ΓΕΝΙΚΕΥΣΕΙΣ ΥΠΟΘΕΣΕΩΝ

(Υ1΄): Γραμμικότητα (Υ1) και Στασιμότητα και Ασθενή Εξάρτηση:

$$y_t = \beta_0 + \beta_1 x_{1t} + \ldots + \beta_k x_{kt} + u_t$$

- X μπορεί να συμπεριλαμβάνει υστερήσεις
- Εννοούμε ότι $\{(y_t, x_{1t}, \ldots, x_{kt})\}_{t=1}^{\infty}$ έχει ασθενή εξάρτηση

(Υ2΄): $E(u_t \mid X_t) = 0 \quad (\neq E(u_t \mid X))$

Σύγχρονη (ασθενής) εξωγένεια

- Πολύ πιο ασθενής υπόθεση

(Υ3΄) = (Υ3) : Όχι τέλεια συγγραμμικότητα

Θεώρημα: *Συνέπεια* ΕΕΤ

(Υ1΄-3΄) $\Rightarrow p\lim \hat{\beta}_j = \beta_j,\ j = 0,1,\ldots,k$

$$p\lim \hat{\beta}_j = \beta_j \Leftrightarrow \lim_{n\to\infty} \Pr\left(\left|\hat{\beta}_j - \beta_j\right| < \varepsilon\right) = 1$$

Παραδείγματα Υποδειγμάτων↔Συμβάντων

1. $y_t = \beta_0 + \beta_1 z_{1t} + \beta_2 z_{2t} + u_t$

   $E(u_t \mid z_{1t}, z_{2t}) = 0$ — Στατικό υπόδειγμα

   $z_{1t} = \delta_0 + \delta_1 y_{t-1} + \nu_t$ — Επιτρέπει *ανάδραση*

   Π.χ. $\left\{\begin{array}{l} z_{1t}: \text{νομισματική πολιτική} \\ y_t: \text{πληθωρισμός} \end{array}\right\}$

2. $y_t = a_0 + \delta_0 z_t + \delta_1 z_{t-1} + \delta_2 z_{t-2} + u_t$

   $E(u_t \mid z_t, z_{t-1}, z_{t-2}, z_{t-3}, \ldots) = 0$ — Υπόδειγμα πεπερασμένων υστερήσεων

$\Rightarrow$ Επιπλέον υστερήσεις άχρηστες
Γενικευμένες υποθέσεις ισχύουν

(Υ4΄): Ομοσκεδαστικότητα:

$$\operatorname{var}(u_t \mid X_t) = \sigma^2 \forall t$$

(Υ5΄): Όχι Αυτοσυσχέτιση:

$$E(u_t u_s \mid X_t, X_s) = 0$$

Θεώρημα: Ασυμπτωτική Κανονικότητα ΕΕΤ

Υ1΄-5΄ $\Rightarrow$ Ε.Ε.Τ έχουν κανονική κατανομή ασυμπτωτικά

$$\lim_{n\to\infty} \Pr\left(\hat{\beta}(n) < x\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{(x-\beta)^2}{\sigma^2 A^{-1}/n}\right)$$

$$p\lim(X'X/n)^{-1} = A^{-1}$$

- Οι στατιστικές *t* και *F* έχουν τις ίδιες κατανομές ασυμπτωτικά
- Η εκτιμήτρια διακύμανσης ΕΕΤ είναι ασυμπτωτικά αμερόληπτη

- Μπορούμε να δείξουμε ότι οι ΕΕΤ είναι Άριστοι ασυμπτωτικά
- Ανάλογα αποτελέσματα ισχύουν για Σ.Α. με τάσεις όταν συμπεριλάβουμε και το χρόνο ως μεταβλητή.

Παράδειγμα 11.4: *Αποτελεσματικότητα Αγορών*

$$\text{Αποδόσεις}_t = \underset{(.081)}{.180} + \underset{(.038)}{.059}\,\text{Αποδόσεις}_{t-1}$$

n = 689, R^2= .0035, Εβδομάδες (Τετάρτη) Ιαν.’76- Μαρ. ‘89

H_0: $\beta_1 = 0$ (Αποτελεσματικότητα)

$\Longrightarrow$ Δεν απορρίπτεται με βάση τη στατιστική *t*.

-AR(2)-

$$\text{Αποδόσεις}_t = \underset{(.081)}{0.186} + \underset{(.038)}{.060}\,\text{Αποδόσεις}_{t-1} - \underset{(.038)}{.038}\,\text{Αποδόσεις}_{t-2}$$

n = **688**, R^2= .0048

H_0: $\beta_1 = \beta_2 = 0$ (Αποτελεσματικότητα)

$$F(2,685) = \frac{(\text{ΑΤΠΠ} - \text{ΑΤΠ})/q}{\text{ΑΤΠ}/(n-k-1)} = \frac{1-R^2}{R^2}\frac{q}{n-k-1} = 1.65 \Rightarrow p \cong .193$$

Παράδειγμα 11.5 Καμπύλη Φίλιπς με προσδοκίες
Από π.10.1 έχουμε

ΠΛΗΘ(t) = 1.42 + 0.468ΑΝ(t)
(1.72) (0.289) n = 49, R^2= 0.053

ΗΠΑ ετήσια στοιχεία (1948-1996)

Αν έχουν σημασία οι προσδοκίες:

ΠΛΗΘ(t) – ΠροσδΠΛΗΘ(t) = β(ΑΝ(t) - μ)+e(t)
μ ~ «φυσική ανεργία»

ΠροσδΠΛΗΘ(t) = ΠΛΗΘ(t-1) (προσαρμόσιμες προσδοκίες)

ΠΛΗΘ(t) –ΠΛΗΘ(t-1) = 3.03 -.543ΑΝ(t)
(1.38) (.230) n=**48**, R=.088

μ=3.03/.543=5.58

Γενικευμένες υποθέσεις επιτρέπουν σχέση ΑΝ(t) με Δ ΠΛΗΘ(t+j) ή ΑΝ(t) με ΠΛΗΘ(j)

ΧΡΟΝΟΣΕΙΡΕΣ ΜΕ ΔΥΝΑΤΗ ΕΞΑΡΤΗΣΗ

$y_t = y_{t-1} + e_t, t = 1, 2, ...$ («Τυχαία Διαδρομή»)

$e_t \overset{iid}{\sim} D\left(0, \sigma^2\right)$

e_{t-1} ανεξάρτητο $\tau o \upsilon\ e_t \forall t$

- Υποπερίπτωση AR(1) με $\rho = 1$ και $\alpha_0 = 0$

Ιδιότητες Τυχαίας Διαδρομής:

- $\Rightarrow y_t = e_t + e_{t-1} + ... + e_1 + y_0$ (*Επανειλημμένη Αντικατάσταση*)

$\Rightarrow E\left(y_t\right) = E\left(\sum e_t\right) + E\left(y_0\right) = y_0 \forall t \geq 1$

$\neq f\left(t\right)$ (στάσιμη;)

- $\Rightarrow Var\left(y_t \mid y_0\right) = Var\left(e_t\right) + Var\left(e_{t-1}\right) + ... + Var\left(e_1\right)$

$= \sigma^2 t$

$= f\left(t\right)$ (άρα *μη στάσιμη*)

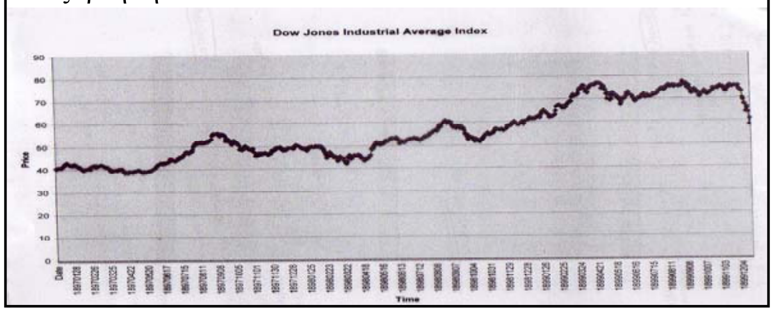
- $\Rightarrow y_{t+h} = e_{t+h} + e_{t+h-1} + ... + e_{t+1} + y_t$

$\Rightarrow E\left(y_{t+h} \mid y_t\right) = y_t \forall h \geq 1 \rightarrow$ Δυνατή Εξάρτηση

$\boxed{AR(1): E\left(y_{t+h} \mid y_t\right) = \rho_1^h y_t \left(\text{ασθενής εξάρτηση} : \left|\rho_1\right| < 1\right)}$

Παρατηρήσεις:

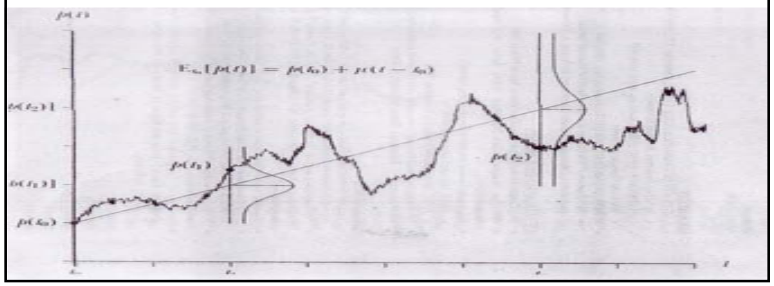
- Είναι δύσκολο να διακρίνουμε μια τυχαία διαδρομή με το μάτι (παράδειγμα DJIA)
- Μακροχρόνιος αντίκτυπος οικονομικής πολιτικής εξαρτάται από το αν η μεταβλητή έχει δυνατή εξάρτηση



Γενικεύσεις:

- e_t μπορεί να έχει ασθενή εξάρτηση
- Τυχαία διαδρομή με τάση (drift):

$$y_t = a_0 + y_{t-1} + e_t \Rightarrow y_t = a_0 t + e_t + e_{t-1} + \ldots + e_1 + y_0$$
$$\Rightarrow E(y_t) = a_0 t + y_0 \rightarrow a_0 > 0 \rightarrow \text{αύξηση}$$



Μετασχηματισμός Χρονοσειρών με Δυνατή Εξάρτηση:

- Διαφορισμός:

$$\Delta y_t = y_t - y_{t-1}, t = 2,3,\ldots$$
$$\gamma\iota\alpha \text{ Τ.Δ.}: \Delta y_t = e_t$$

- Λογαριθμικός Διαφορισμός:

$$\Delta \log(y_t) \cong \frac{y_t - y_{t-1}}{y_{t-1}}$$

(όπως με αποδόσεις στο π. 11.4)

Επιθυμητό αποτέλεσμα

- Αφαιρεί δυνατή εξάρτηση (αν y_t είναι I(1) – Σ.Α. με ολοκλήρωση 1ου βαθμού, το Δy_t είναι I(0)), π.χ. Τιμές - Αποδόσεις
- Αφαιρεί γραμμική τάση αν υπάρχει:

$$y_t = \gamma_0 + \gamma_1 t + \nu_t \Rightarrow \Delta y_t = \gamma_1 + \Delta \nu_t$$

- Συχνά η διαφορισμένη παλινδρόμηση μας δίνει την πληροφορία / παράμετρο που θέλουμε.

Χρήση αυτοσυσχέτισης για αναγνώριση I(1)

Παράδειγμα 11.6: Γεννητικότητα
Από π.10.4

Γεννητικότητα(t) = 95.87 +.073Απαλλαγές(t) – .0058Απ(t-1)
(3.28) (.126) (.1557)

+.034Απ(t-2) - 22.13ΔΠΠ(t) – 31.3Χάπι(t)
(.126) (10.73) (3.98)

Όμως: Γεννητικότητα(t)=.977Γεννητικότητα(t-1)
Απαλλαγές(t) =.964Απαλλαγές(t-1)

Δ Γεννητικότητα(t)= -.785 -.043Δ Απαλλαγές(t)
(.502) (.028)

n=71, R2=.032

Δ Γενν(t)= -.964 -.036Δ Απ(t) -.014Δ Απ(t-1) +**.110**Δ Απ(t-2)
(.468) (.027) (.028) **(.027)**

n=69,R2=.233

ΔΥΝΑΜΙΚΑ ΠΛΗΡΗ ΥΠΟΔΕΙΓΜΑΤΑ (Δ.Π.)

Το υπόδειγμα:

$$y_t = \beta_0 + \beta_1 x_{1t} + \ldots + \beta_k x_{kt} + u_t$$

x_t συμπεριλαμβάνει y_{t-j}, z_{t-l}

είναι Δ.Π. αν:

$$E(u_t \mid z_t, y_{t-1}, z_{1t-1}, \ldots) = 0$$

$$\Leftrightarrow E(y_t \mid z_t, y_{t-1}, z_{1t-1}, \ldots) = E(y_t \mid x_t, x_{t-1}, \ldots) = E(y_t \mid x_t)$$

Ερμηνεία έννοιας:

1. Επιπλέον υστερήσεις είναι άχρηστες
2. $$\text{Δ.Π.} \Rightarrow E(u_t \mid z_t, y_{t-1}, z_{1t-1}, \ldots) = 0$$
   $$\Rightarrow E(u_t \mid x_t, u_{t-1}, x_{t-1}, u_{t-2}, \ldots) = 0$$
   $$\Rightarrow E(u_t \mid x_t, x_s, u_s) = 0 \,\forall s < t$$
   $$\Rightarrow E(u_s E(u_t \mid x_t, x_s, u_s) \mid x_t, x_s) = 0 \,\forall s < t$$
   $$\Rightarrow E(E(u_t u_s \mid x_t, x_s, u_s) \mid x_t, x_s) = 0 \,\forall s < t$$
   $$\Rightarrow E(u_t u_s \mid x_t, x_s) = 0 \,\forall s < t$$
   $\Rightarrow$ Y5' ισχύει

   όμως Y5' δεν $\Rightarrow$ Δ.Π.

Παράδειγμα (από κεφάλαιο 12)

$$\left.\begin{aligned} y_t &= \beta_0 + \beta_1 y_{t-1} + u_t, |\beta_1| < 1 \\ u_t &= \rho u_{t-1} + e_t, |\rho| < 1 \end{aligned}\right\} \Rightarrow$$

1) Όχι σύγχρονη εξωγένεια:

$$Cov(y_{t-1}, u_t) = \rho Cov(y_{t-1}, u_{t-1}) \neq 0$$

2) Έλλειψη Δυναμικής Πληρότητας:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \rho(y_{t-1} - \beta_0 - \beta_1 y_{t-2}) + e_t \quad \sim\text{AR(2)}$$

$$E(y_t \mid y_{t-1}, y_{t-2}) \neq E(y_t \mid y_{t-1})\ ,$$

(προφανώς, αφού έχουμε δείξει ότι Δ.Π.→Y5΄)

Παράδειγμα 11.8: Γεννητικότητα

Από π.11.6

Δ Γενν(t)=	-.964	-.036Δ Απ(t)	-.014Δ Απ(t-1)	+.110Δ Απ(t-2)
	(.468)	(.027)	(.028)	(.027)

n=69,R2=.233

Όμως

Δ Γενν(t)= β0+β1ΔΑπ(t)+β2 ΔΑπ(t-1)+β3ΔΑπ(t-2)+.3ΔΓενν(t-1)

t-stat(H0: β4=0)=2.84 => π.11.6 **δυναμικά ελλιπή**

=> αυτοσυσχέτιση;

ικανοποιητικό υπόδειγμα;

Ερμηνεία Ομοσκεδαστικότητας σε Υποδείγματα Χρονοσειρών

Στο στατικό υπόδειγμα $y_t = \beta_0 + \beta_1 z_t + u_t$

- Y4΄→ $Var(u_t \mid z_t) = Var(y_t \mid z_t) = \sigma^2$

Επιτρέπεται $Var(y_t \mid y_{t-1}) = f(y_{t-1})$

Ενώ στο υπόδειγμα $y_t = \beta_0 + \beta_1 z_t + \beta_2 y_{t-1} + \beta_3 z_{t-1} + u_t$

- Y4΄→ $Var(u_t \mid z_t, y_{t-1}, z_{t-1}) = Var(y_t \mid z_t, y_{t-1}, z_{t-1}) = \sigma^2$

Άρα επιτρέπονται διαφορετικές σχέσεις για μέσο και διακύμανση:

$E(y_t \mid y_{t-1}) = f(y_{t-1})$

$Var(y_t \mid y_{t-1}) \neq f(y_{t-1})$

Αν συμπεριλαμβάνεται y_{t-1} αποκλείεται η «Δυναμική Ετεροσκεδαστικότητα»

ΚΕΦΑΛΑΙΟ 12

ΑΥΤΟΣΥΣΧΕΤΙΣΗ

και

ΕΤΕΡΟΣΚΕΔΑΣΤΙΚΟΤΗΤΑ

Κεφ. 10	**Κεφ. 11**	**Κεφ. 12**
Y1-5 «Κλασικές υποθέσεις» ***(μικρά δείγματα)***	Y1΄-5΄ «Ασθενείς εκδοχές κλασσικών υποθέσεων» *X(t)* αντί για *X* σε Y2,4,5	**- Y4’, Y5’**
• Γραμμικότητα: $\boxed{y_t = \beta_0 + \ldots + \beta_k x_{kt} + u_t}$ • Αυστηρή Εξωγένεια: $\boxed{E(u_t \mid X) = 0}$ • Ατελής Συγγραμικότητα: $\boxed{\nexists l : x_{lt} = a, \nexists \beta^* : x_l = x_{-l}\beta^*}$ • Ομοσκεδαστικότητα: $\boxed{Var(u_t \mid X) = \sigma^2}$ • Έλλειψη Αυτοσυσχ/σης: $\boxed{Corr(u_t, u_s \mid X) = 0}$	+ ***μεγάλα δείγματα*** Στασιμότητα Ασθ. εξάρτηση	

- Τι ξέρουμε ήδη;

⇨ Y1-3→ Αμεροληψία Ε.Ε.Τ (κφ. 10)

Y1΄-3΄→ Συνέπεια Ε.Ε.Τ (κφ. 11)

ΑΠΟΤΕΛΕΣΜΑΤΙΚΟΤΗΤΑ ΚΑΙ ΕΠΑΓΩΓΗ ΕΕΤ ΟΤΑΝ ΤΑ ΚΑΤΑΛΟΙΠΑ ΕΧΟΥΝ ΑΥΤΟΣΥΣΧΕΤΙΣΗ

Παράδειγμα:

$$y_t = \beta_0 + \beta_1 x_t + u_t$$
$$u_t = \rho u_{t-1} + e_t, t = 1,2,\ldots,n \quad AR(1)$$
$$|\rho| < 1$$
$$e_t \sim D(0, \sigma_e^2),$$
$$Corr(e_t, e_s) = 0$$

$$\Rightarrow E(u_t u_{t+j}) = \rho^j \sigma^2$$

Ιδιότητες ΕΕΤ, αν $\bar{x} = 0$ ($\Sigma(x_t) = 0$):

$$\hat{\beta}_1 = \frac{\sum(x_t(y_t - \bar{y}))}{\sum x_t^2} = \frac{\sum(x_t(\beta_0 + \beta_1 x_t + u_t - \bar{y}))}{\sum x_t^2}$$

$$\Rightarrow \hat{\beta}_1 = \frac{\sum(x_t(\beta_0 + \beta_1 x_t - \bar{y}))}{\sum x_t^2} + \frac{1}{\sum x_t^2}\sum(x_t u_t) = \beta_1 + \frac{1}{\sum x_t^2}\sum(x_t u_t)$$

$$Var(\hat{\beta}_1 \mid X) = \left(\frac{1}{\sum x_t^2}\right)^2 Var\left(\sum(x_t u_t)\right) =$$

$$= \left(\frac{1}{\sum x_t^2}\right)^2 \left(\sum x_t^2 Var(u_t) + 2\sum_{t=1}^{n-1}\sum_{j=1}^{n-t} x_t x_{t+j} E(u_t u_{t+j})\right) =$$

$$= \frac{\sigma^2}{\sum x_t^2} + 2\frac{\sigma^2}{\left(\sum x_t^2\right)^2}\sum_{t=1}^{n-1}\sum_{j=1}^{n-t}\rho^j x_t x_{t+j}$$

$$\rho = 0 \Rightarrow Var(\hat{\beta}_1 \mid X) = \frac{\sigma^2}{\sum x_t^2}$$ π.χ. Θ2.2

$$\hat{\sigma}_{\hat{\beta}_1} = \frac{\sum \hat{u}_t^2}{(n-k-1)\sum x_t^2}$$ Αμερόληπτος

$\rho > 0 \Rightarrow E(\hat{\sigma}_{\hat{\beta}_1}) < Var(\hat{\beta}_1 \mid X)$ ⇨ Μεροληπτικός

$\rho < 0 \Rightarrow E(\hat{\sigma}_{\hat{\beta}_1}) >< Var(\hat{\beta}_1 \mid X)$ ⇨ Μάλλον Μεροληπτικός

Οι στατιστικές *t, F, LM* χρησιμοποιούν τις εκτιμήσεις αυτές, άρα μεροληπτικές εκτιμήσεις προκαλούν και μη αναμενόμενες κατανομές στατιστικών

Ερμηνεία R^2 υπό Υ1'-3' (π.χ. με αυτοσυσχέτιση):

$$R^2 = 1 - \sigma_u^2 / \sigma_y^2 \sim R^2(\pi\lambda\eta\theta\upsilon\sigma\mu o\acute{\upsilon})$$

$$\hat{R}^2 = 1 - \hat{\sigma}_u^2 / \hat{\sigma}_y^2$$

- Υ1΄-3΄ $\overset{LLN}{\Rightarrow} \underset{n\to\infty}{p\lim}\hat{R}^2 = R^2$ (**συνέπεια**) *και με* αυτοσυσχέτιση
- Όμως $E(\hat{R}^2) \neq R^2$, (***μεροληψία***) υπό όποιες υποθέσεις

Στασιμότητα και ερμηνεία R^2

- Αν υπάρχουν τάσεις η εποχικότητες στις μεταβλητές η / και στο υπόδειγμα, το R^2 πρέπει να ερμηνευτεί ανάλογα
- Αν η εξαρτημένη μεταβλητή είναι I(1), τότε:

$$Var(y_t) = f(t), \frac{\partial f}{\partial t} > 0 \Rightarrow$$

οπότε το *Var*(*y*) δεν υπάρχει, οπότε και το R^2 που βασίζεται σε αυτό δεν έχει νόημα

Σημασία αυτοσυσχέτισης (Υ5) σε υποδείγματα με υστερήσεις εξαρτημένης μεταβλητής

Παράδειγμα

$E(y_t \mid y_{t-1}) = \beta_0 + \beta_1 y_{t-1}$ ~ Γραμμικότητα κ΄ Στασιμότητα

$|\beta_1| < 1$ ~ Ασθενής Εξάρτηση (Υ1΄)

$\Rightarrow y_t = \beta_0 + \beta_1 y_{t-1} + u_t$ ~ Ατελής Συγγραμμικότητα (Υ3΄)

$E(u_t \mid y_{t-1}) = 0$ ~ Σύγχρονη Εξωγένεια (Υ2΄)

Υ1΄-3΄→ Συνέπεια ΕΕΤ με υστερήσεις εξαρτημένης και με αυτοσυσχέτιση (χωρίς Υ5΄) ?

Μήπως όμως :

$\left.\begin{array}{l}\text{Υ1΄-3΄ +}\\ \text{Υστερήσεις}\\ \text{Εξαρτημενης}\\ \text{Μεταβλητής}\end{array}\right\} \rightarrow$ Υ5΄ ?

Δεν περιορίζεται από το υπόδειγμα

Όχι. *Στο παράδειγμα μας:*

$$Cov(u_t, u_{t-1}) = Cov(u_t, y_{t-1} - \beta_0 - \beta_1 y_{t-2}) = -\beta_1 Cov(u_t, y_{t-2})$$

$$E(u_t \mid y_{t-1}) = 0 \Leftrightarrow Cov(u_t, y_{t-1}) = 0 \not\Rightarrow Cov(u_t, y_{t-2}) = 0$$

$\left.\begin{array}{l}\text{Όχι Υ5΄}\\ \text{Υστερήσεις}\end{array}\right\} \xRightarrow[\text{υποδείγματα}]{\text{Σε κάποια}}$ Όχι σύγχρονη εξωγένεια (Υ2΄)

(Πρέπει να εξετάσουμε αυτοσυσχέτιση και υπόδειγμα μαζί)

Παράδειγμα

$$\left.\begin{array}{l} y_t = \beta_0 + \beta_1 y_{t-1} + u_t, |\beta_1| < 1 \\ u_t = \rho u_{t-1} + e_t, |\rho| < 1 \end{array}\right\} \Rightarrow$$

1) Όχι σύγχρονη εξωγένεια:
$Cov(y_{t-1}, u_t) = \rho Cov(y_{t-1}, u_{t-1}) \neq 0$

2) Έλλειψη Δυναμικής Πληρότητας:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \rho(y_{t-1} - \beta_0 - \beta_1 y_{t-2}) + e_t \quad \sim \text{AR(2)}$$

$E(y_t \mid y_{t-1}, y_{t-2}) \neq E(y_t \mid y_{t-1})$,
(προφανώς, αφού έχουμε δείξει ότι Δ.Π.→Υ5΄)

Έλεγχος για αυτοσυσχέτιση με τη στατιστική *t* (Ασυμπτωτικός)

Έστω $y_t = \beta_0 + \beta_1 x_{1t} + \ldots + \beta_k x_{kt} + u_t$, Γραμμ/τητα (Υ1)

$E(u_t \mid \underline{X}) = 0$ **Αυστηρή** Εξωγένεια (Υ2)
→ Όχι υστερήσεις της y_t

Πρόσθετες υποθέσεις ***για τα σφάλματα***

$u_t = \rho u_{t-1} + e_t, t = 1,2,\ldots,n$ AR(1) Σφάλματα

$|\rho| < 1$

$E(e_t \mid u_{t-1}, u_{t-2}, \ldots) = 0$ (ένα είδος) Εξωγένειας (Υ2)

$Var(e_t \mid u_{t-1}) = Var(e_t) = \sigma_e^2$ Ομοσκεδαστικότητα (Υ4)

$$\left\{\begin{array}{l} H_0\text{: Όχι αυτοσυσχέτιση} \leftrightarrow \rho = 0 \\ H_1\text{: Αυτοσυσχέτιση} \leftrightarrow \rho \neq 0 \text{ (ή } \rho > 0 \text{ όταν πιστεύουμε ότι } \rho \geq 0) \end{array}\right\}$$

Δυσκολία ελέγχου οφείλεται στο ότι τα σφάλματα δεν παρατηρούνται. Όμως:

Ασυμπτωτικά $\Leftrightarrow \left\{\begin{array}{l} H_0: \rho = 0 \\ H_1: \rho \neq 0 \end{array}\right\}$ στο υπόδειγμα: $\hat{u}_t = \rho \hat{u}_{t-1} + \zeta_t$

Οπότε μπορούμε να κάνουμε έναν ισοδύναμο έλεγχο σε υπόδειγμα για τα κατάλοιπα (τα οποία παρατηρούνται)

<u>Διαδικασία ελέγχου:</u>

i. Βρίσκουμε τα κατάλοιπα $\hat{u}_t$ από παλινδρόμηση:
$y_t = \beta_0 + \beta_1 x_{1t} + \ldots + \beta_k x_{kt} + u_t$

ii. Παλινδρομούμε: $\hat{u}_t = \rho \hat{u}_{t-1} + \zeta_t$, βρίσκοντας $\hat{\rho}, t_{\hat{\rho}}$

iii. Κάνουμε τον έλεγχο

Παρατηρήσεις:

• Οι πρόσθετες υποθέσεις είναι περιοριστικές

• Αν $Corr(u_t,u_{t-1}) = 0$ (αλλά π.χ. $Corr(u_t,u_{t-2}) \neq 0$), ο έλεγχος δεν θα παρατηρήσει υπαρκτή αυτοσυσχέτιση

• Ως ασυμπτωτικός έλεγχος, εφαρμογή για «μεγάλο *n*» σημαίνει ότι όποιες απορρίψεις μπορεί να μην έχουν οικονομική σημασία

Παράδειγμα 12.1

Στατική καμπύλη Phillips (1):

$$\widehat{\Pi\lambda\eta\theta}_t = 1.42 + .468\,A\nu\varepsilon\rho\gamma\acute{\iota}\alpha_t$$
$$(1.72) \quad (.289)$$

$n = 49$
$R^2 = .053$
ΗΠΑ: 1948-’96

Καμπύλη Phillips με προσδοκίες (2):

$$(\Pi\lambda\eta\theta_t - \Pi\lambda\eta\theta^e_t) = \beta_1(A\nu\varepsilon\rho\gamma\acute{\iota}\alpha_t - \mu_0) + e_t$$

$$\Pi\lambda\eta\theta^e_t = \Pi\lambda\eta\theta_{t-1}$$

$$\rightarrow \Delta\Pi\lambda\eta\theta_t = 3.03 - .543\,A\nu\varepsilon\rho\gamma\acute{\iota}\alpha_t$$
$$(1.38) \quad (.230)$$

$n = 48$
$R^2 = .108$

Για:

(1): $\hat{\rho} = .573, t_{\hat{\rho}} = 4.93, p = 0, (n = 48)$

$\rightarrow$ Υ5 *παραβιάζεται*

(2): $\hat{\rho} = -.036, t_{\hat{\rho}} = -.297, p = .775, (n = 47)$

$\rightarrow$ Υ5 *ok*

Έλεγχος Αυτοσυσχέτισης Durbin-Watson (για μικρά δείγματα)

Η στατιστική *DW* ορίζεται ως:

$$DW = \frac{\sum_{t=2}^{n}\left(\hat{u}_t - \hat{u}_{t-1}\right)^2}{\sum_{t=1}^{n}\hat{u}_t^2}$$

και έχει γνωστή κατανομή υπό ***Υ1-6*** (μειονέκτημα σε σχέση με έλεγχο *t*) ακόμα και σε *μικρά δείγματα* (πλεονέκτημα σε σχέση με έλεγχο *t παραδείγματος καμπύλης Φίλιπς*)

Μπορούμε να δείξουμε ότι:

1)

$$DW \approx 2\left(1-\hat{\rho}\right) \Rightarrow \begin{cases} \hat{\rho}=0 \Rightarrow DW \approx 2 \\ \hat{\rho}>0 \Rightarrow DW<2 \end{cases}$$

2) (Περίπλοκη) κατανομή στατιστικής DW εξαρτάται από:

N, k, ύπαρξη σταθερού όρου αλλά **και** X

3) Υπάρχουν όμως *διαστήματα* $[d_L, d_U]$ *ανεξάρτητα* του X, τέτοια ώστε:

$DW < d_L$ → Απορρίπτουμε H_0 σε α% επ.σ.σ.

$d_L \le DW \le d_U$ → Δεν ξέρουμε

$d_U < DW$ → Δεν απορρίπτουμε

Παράδειγμα 12.1 (επανάληψη αποτελεσμάτων)

Στατική καμπύλη Phillips (1):

$\widehat{\Pi\lambda\eta\theta}_t = 1.42 + .468 A\nu\varepsilon\rho\gamma\acute{\iota}\alpha_t$

(1.72) (.289)

$n = 49$
$R^2 = .053$
ΗΠΑ: 1948-’96

Καμπύλη Phillips με προσδοκίες (2):

$(\Pi\lambda\eta\theta_t - \Pi\lambda\eta\theta^e_t) = \beta_1(A\nu\varepsilon\rho\gamma\acute{\iota}\alpha_t - \mu_0) + e_t$

$\Pi\lambda\eta\theta^e_t = \Pi\lambda\eta\theta_{t-1}$

→ $\Delta\Pi\lambda\eta\theta_t = 3.03 - .543 A\nu\varepsilon\rho\gamma\acute{\iota}\alpha_t$

(1.38) (.230)

$n = 48$
$R^2 = .108$

Για:

(1): $\hat{\rho} = .573, t_{\hat{\rho}} = 4.93, p = 0, (n = 48)$

→ Y5 <u>*παραβιάζεται*</u>

Επίσης *DW=.80*

Από πίνακες ξέρουμε ότι αν *k=1, n=50* και *α=0.01*, τότε *d(L)=1.32 >DW =>* **απόρριψη**

(2): $\hat{\rho} = -.036, t_{\hat{\rho}} = -.297, p = .775, (n = 47)$

→ Y5 <u>*ok*</u>

Επίσης *DW=1.77*

Από πίνακες ξέρουμε ότι αν *k=1, n=50* και *α=0.01*, τότε *d(U)=1.59 < DW =>* ***<u>δεν απορρίπτεται</u>***

Έλεγχος Αυτοσυσχέτισης χωρίς ΑΕ (πχ *y(t-1) στο υπόδειγμα*)

i. Βρίσκουμε τα κατάλοιπα $\hat{u}_t$ από παλινδρόμηση:

$y_t = \beta' \underline{x}_t + u_t$ (με υστερήσεις όλων των μεταβλητών)

ii. Παλινδρομούμε: $\hat{u}_t = \gamma' x_t + \rho \hat{u}_{t-1} + \zeta_t$ βρίσκοντας $\hat{\rho}, t_{\hat{\rho}}$

- Η στατιστική $t_{\hat{\rho}}$ έχει ασυμπτωτικά κατανομή t ακόμη και όταν $Cov\left(x_{jt}, u_{t-1}\right) \neq 0$ (όχι ΑΕ-*Y2*)
- Συμπεριλαμβάνοντας $\gamma' x_t$ λαμβάνουμε υπ΄όψιν ότι $Cov\left(x_{jt}, u_{t-1}\right) \neq 0$ πράγμα που δε γίνεται με τον απλό έλεγχο βάση της στατιστικής t
- Πρέπει όμως να υποθέσουμε ότι $Var(u_t | x_t, u_{t-1})$=σ^2

Παράδειγμα 10.3

$\log(\%\, Απασχόληση_t) = -1.05$
(0.77)

$-1.54\log \left(\frac{Μεσος\ Ελάχιστος\ Μισθός}{Μέσος\ Μισθός} \times \%Κάλυψης \right)_t$
(0.65)

$-0.12\log(ΑΕΠ^{US}_t)$
$(.089)$

$n = 38$, $R^2 = .661$, $\bar{R}^2 = .641$, Πόρτο Ρίκο: 1950-1987

β_1 = ελαστικότητα

$t_{\hat{\beta}_1} = -2.37$

Παράδειγμα 10.9 (Π10.3 με τάση)

$\log(\%\, Απασχόληση_t) = -8.70$
(1.30)

$-1.69\log \left(\frac{Μεσος\ Ελάχιστος\ Μισθός}{Μέσος\ Μισθός} \times \%Κάλυψης \right)_t$
(0.044)

$-1.06\log(ΑΕΠ^{US}_t) - 0.032t$
$(.18)$ (0.005)

$n = 38$, $R^2 = .847$, $\bar{R}^2 = .834$, Πόρτο Ρίκο: 1950-1987

Επειδή $\log(ΑΕΠ^{US}_t) = 0.03t$ το ΑΕΠ γίνεται πολύ σ.σ. => σημασία έχει επίπεδο ΑΕΠ *σε σχέση με την τάση του*

Χωρίς να αλλάζουν άλλα συμπεράσματα

Π.χ. 12.2 (*10.3,10.9*)

$x_t \sim ΑΕΠ^{PR}_t$, $ΑΕΠ^{US}_t$, μεταβλητή ελάχιστου μισθού

$$\hat{u}_t = \beta_0 + \beta x_t + \ldots + .481\hat{u}_{t-1}$$
$$t_{\hat{\rho}} = 2.89$$
$$n = 37$$
$$p = .007$$

$\rightarrow$ Απορρίπτουμε H_0: $\rho = 0$

- Προηγούμενες επαγωγές **λάθος**
- Εκτιμήσεις συνεπείς

Έλεγχος Αυτοσυσχέτισης υψηλότερου βαθμού

Προσέγγιση: Γενίκευση ελέγχων πρώτου βαθμού

Υποθέσεις:

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \ldots \rho_q u_{t-q} + e_t, t = 1,2,\ldots,n$$
$$|\Sigma \rho_i| < 1$$
$$E(e_t \mid u_{t-1}, u_{t-2}, \ldots) = 0$$
$$Var(e_t \mid u_{t-1}) = Var(e_t) = \sigma_e^2$$
$$Var(u_t \mid x_t, u_{t-1}, \ldots, u_{t-q}) = \sigma_u^2$$

H_0: Όχι αυτοσυσχέτιση $\leftrightarrow \rho_1 = \rho_2 = \ldots = \rho_q = 0$
H_1: Αυτοσυσχέτιση $\leftrightarrow \rho_i \neq 0$ (για κάποιο i)

Διαδικασία ελέγχου

i. Βρίσκουμε τα κατάλοιπα $\hat{u}_t$ από παλινδρόμηση:

$y_t = \beta' \underline{x}_t + u_t$ (με επιθυμητές υστερήσεις μεταβλητών)

ii. Παλινδρομούμε:

$$\hat{u}_t = \gamma' x_t + \rho_1 \hat{u}_{t-1} + \rho_2 \hat{u}_{t-2} + \ldots + \rho_q \hat{u}_{t-q} + \zeta_t$$

Υπολογίζοντας **στατιστική** F για από κοινού σ.σ. ρ_i

- Η στατιστική θα έχει την επιθυμητή κατανομή *ασυμπτωτικά* ακόμη και χωρίς ΑΕ-*Υ2*
- Αν ισχύει ΑΕ-*Υ2* μπορούμε να μην συμπεριλάβουμε $\gamma' x_t$

Εναλλακτικά, χρησιμοποιούμε ότι ασυμπτωτικά:

$LM = (n-q) R^2_u \sim \chi^2_q$ **Έλεγχος Breusch - Godfrey**

Π.χ. 12.3 (10.5) – Έλεγχος για αυτοσυσχέτιση 3ου βαθμού

log(Εισαγωγές_Κίνα(t))= -17.8 + 3.12log(Παραγωγή(t))
(21.05) (0.48)
+.196log(Παρ._Βενζίνης(t))+.983log($(t))
(.907) (.400)
+.060(Πριν_Κατάθεση(t)) -.032(Κατάθεση(t))
(.261) (.264)
-.566(Μετά(t))
(.286)

n = 131,

R^2= 0.305,

$\bar{R}^2$= 0.271

(Φεβ. ’78-Δεκ. ’88)

Παλινδρομούμε:

$$\hat{u}_t = \gamma' x_t + \rho_1 \hat{u}_{t-1} + \rho_2 \hat{u}_{t-2} + \rho_3 \hat{u}_{t-3} + \zeta_t$$

F* = 5.12, β.ε. = *(3,118) => ***p=.0023*** => **Απορρίπτουμε H_0**

Διορθώσεις αυτοσυσχέτισης με Αυστηρή Εξωγένεια

- Αν στόχος είναι ένα ***δυναμικά πλήρες υπόδειγμα***, πρέπει να επαναπροσδιοριστεί συνολικά
- Αν στόχος είναι ***επαγωγή*** για παραμέτρους υποδείγματος με αυστηρή εξωγένεια, υπάρχουν πιο απλές λύσεις

BLUE Εκτιμήσεις με AR(1) σφάλματα

- Υ1-4
- Αντί για Υ5: $u_t = \rho u_{t-1} + e_t$

$\Rightarrow y_t - \rho\, y_{t-1} = (1-\rho)\beta_0 + \beta_1(x_t - \rho\, x_{t-1}) + e_t$

$\Rightarrow y^*_t = (1-\rho)\beta_0 + \beta_1 x^*_t + e_t$, ***t >1***

$y^*_t = y_t - \rho\, y_{t-1}$, $x^*_t = x_t - \rho\, x_{t-1}$, ημιδιαφορισμένες παρατηρήσεις (quasi-differenced data)

$y_1 = \beta_0 + \beta_1 x_1 + u_1$

$Cov(u_1, e_t) = 0$

Όμως $Var(u_1) = \sigma^2/(1-\rho^2) > \sigma^2 = Var(e_t)$

Οπότε ορίζουμε $y^*_1 = (1-\rho^2)^{0.5}\, y_1$,

$x^*_1 = (1-\rho^2)^{0.5}\, x_1$,

$u^*_1 = (1-\rho^2)^{0.5}\, u_1$,

Slide 116

SS1 ΒΛΘΕ ασψασ
Spyros Skouras, 6/13/2007

Χρησιμοποιώντας:

$y^*_t = (1-\rho)\beta_0 + \beta_1 x^*_t + e_t$, $t>=1$

Έχουμε:

- ΕΕΤ σε αυτή την παλινδρόμηση ⇔ Γενικευμένο Εκτιμητή Ελαχίστων Τετραγώνων στην αρχική (GLS)
- Υ1-5 για την νέα παλινδρόμηση
- Επαγωγή μπορεί να γίνει
- ΓΕΕΤ είναι BLUE (μετασχηματισμός διατηρεί γραμμικότητα)

Εφικτός ΓΕΕΤ

- Το ρ γενικά **δεν είναι γνωστό**
- Μπορεί όμως να αντικατασταθεί στον ΓΕΕΤ από μια συνεπή εκτίμηση

Διαδικασία Εφικτής ΓΕΕΤ

1) Παλινδρόμηση με ΕΕΤ για προσδιορισμό καταλοίπων
2) Παλινδρόμηση σε κατάλοιπα για εκτίμηση ρ
3) Εκτίμηση ΓΕΕΤ βασισμένη σε *εκτίμηση* για ρ

Ως αποτέλεσμα των σφαλμάτων εκτίμησης του ρ

- Ο Εφικτός ΓΕΕΤ μπορεί να είναι μεροληπτικός
- **Ασυμπτωτικά** επαγωγή γίνεται κανονικά
- Εφικτός ΓΕΕΤ **ασυμπτωτικά** αποτελεσματικότερος από OLS

Παραλλαγές διαδικασίας

- Επέκταση:

4) Επιστρέφουμε στο βήμα 2 όπου χρησιμοποιούμε τώρα τα κατάλοιπα από ΕΓΕΕΤ

Οι επιπτώσεις αυτής της επέκτασης δεν είναι γνωστές

- Cochrane-Orcutt (χωρίς την πρώτη παρατήρηση)
- Prais-Winsten (με)

Παράδειγμα 12.4 (*Παράδειγμα 10.5*)

log(Εισαγωγές_Κίνα(t))= -17.8 + 3.12log(Παραγωγή(t))
(21.05) (0.48)
-37.31 (23.33) **2.95 (0.65)**

(Φεβ. ’78-Δεκ. ’88)

n = 131,

R^2= 0.305,

+.196log(Παρ._Βενζίνης(t))+.983log($(t))
(.907) (.400)
1.05 (0.99) **1.14 (0.51)**

+.060(Πριν_Κατάθεση(t)) -.032(Κατάθεση(t))
(.261) (.264)
-.016 (.321) **-.033 (.323)**

-.565(Μετά(t))
(.286)
-.577 (.343)

CO: ρ^=.293 (.084)

n=130

R^2= 0.193 (?)

Παρόμοιες εκτιμήσεις

Ψηλότερες Τ.Α.

Σύγκριση ΕΕΤ και ΕΓΕΕΤ

- ΕΓΕΕΤ απαιτεί Υ1'-4' ***και*** ότι Cov(x_{t-1}+x_{t+1} , u_t)=0

Πχ ρ γνωστό, εκτίμηση Cochrane-Orcutt απαιτεί σύγχρονη εξωγένεια μετασχηματισμένων δεδομένων

Σύγχρονη εξωγένεια στην μετασχηματισνένη παλινδρ.

$\Rightarrow E[(x_t-\rho x_{t-1})(u_t-\rho u_{t-1})] = 0$

$\Rightarrow -\rho[E(x_{t-1}u_t)+E(x_t u_{t-1})] = 0$

$\Rightarrow E[(x_{t-1}+x_{t+1})\, u_t]=0$

Η επιπλέον υπόθεση μπορεί να παραβιάζεται

Οπότε και οι ΕΕΤ μπορεί να είναι προτιμητέες

Παράδειγμα 12.5 (*Παράδειγμα* 10.1): *Καμπύλη* Phillips

y: πληθωρισμός
x: ανεργία

- ΗΠΑ ετήσια στοιχεία (1948-1996)

Με Ε.Ε.Τ καταλήγουμε στο υπόδειγμα:

$$\widehat{\text{ΠΛΗΘ}}(t) = \underset{(1.72)}{1.42} + \underset{(0.289)}{0.468}\text{ΑΝ}(t)$$

CO **7.58** **-.665**
(2.38) **(.320)**

Αποτελέσματα πολύ πιο θεωρητικά βάσιμα

n = 49, R^2= 0.053, $\bar{R}^2$= 0.033

n = 48, R^2= 0.086, $\hat{\rho}$= .774 (.091)

Διαφόριση και Αυτοσυσχέτιση

→ Συχνά η διαφόριση μειώνει δραστικά την αυτοσυσχέτιση

Π.χ.

$$\left.\begin{aligned} y_t &= \beta_0 + \beta_1 x_t + u_t, t = 1,2,\ldots \\ u_t &= \rho u_{t-1} + e_t, \rho \approx 1 \end{aligned}\right\} \Rightarrow$$

$$\Rightarrow \Delta y_t = \beta_1 \Delta x_t + \Delta u_t$$

$$\Delta u_t = (\rho - 1)\Delta u_{t-1} + \Delta e_t, (\rho - 1) \approx 0$$

Π.χ. 12.6 (*10.2*)

$$3ΜΕπιτόκια_t = \underset{(.044)}{1.25} + \underset{(.076)}{.613}Πληθ_t + \underset{(.118)}{.700}Ελλειμμα_t$$

n = 49
R^2 =.697

Όμως: $\hat{u}_t = \underset{(.123)}{.530}\hat{u}_{t-1}$

Άρα παλινδρομούμε:

$$\Delta 3ΜΕπιτόκια_t = \beta_0 + \beta_1 \Delta Πληθ_t + \beta_2 \Delta Ελλειμμα_t$$

$\Delta\hat{u}_t = \underset{(.145)}{.068}\Delta\hat{u}_{t-1}$

... και συγχρόνως απαλλασσόμαστε από πιθανώς μη στάσιμες μεταβλητές

ΕΕΤ και επαγωγή εύρωστη ως προς αυτοσυσχέτιση **και ετεροσκεδαστικότητα**

Μπορεί να εφαρμοστεί όταν

- Παραβίαση αυστηρής εξωγένειας (πχ αυτοπαλίνδρομα)
- Παραβίαση ομοσκεδαστικότητας
- Άγνωστη μορφή αυτοσυσχέτισης (όχι AR(1))

Λογική:

Στο υπόδειγμα: $y_t = \beta_0 + \beta_1 x_{t1} + \ldots + \beta_k x_{tk} + u_t,\ t=1,\ldots,n$

Ορίζοντας: $x_{t1} = \delta_0 + \delta_2 x_{t2} + \ldots + \delta_k x_{tk} + \boldsymbol{r_t},\ t=1,\ldots,n,$
$E(r_t | x_{t2},\ldots,x_{tk})$

Προκύπτει ότι:

$$A\,\mathrm{var}\left(\hat{\beta}_1\right) = \left(\sum_{t=1}^{n} E(r_t^2)\right)^{-2} Var\left(\sum_{t=1}^{n} r_t u_t\right)$$

Οπότε «αρκεί» να εκτιμήσουμε Avar(β^1)

Διαδικασία

(i) Παλινδρόμηση και υπολογισμός

$$\hat{\sigma}_{\hat{\beta}_1}, \hat{\sigma}, \{u_t : t = 1,\ldots, n\}$$

(ii) Βοηθητική παλινδρόμηση και υπολογισμός

$$\{\hat{r}_t : t = 1,\ldots, n\}, \{\hat{a}_t = \hat{r}_t \hat{u}_t : t = 1,\ldots, n\}$$

(iii) Υπολογισμός

$$\hat{v} = \sum_{t=1}^{n} \hat{a}_t^2 + 2\sum_{h=1}^{g} [1 - h/(g+1)]\left(\sum_{t=h+1}^{n} \hat{a}_t \hat{a}_{t-h}\right)$$

Για κάποιο *g* το οποίο να αυξάνει με *n*, πχ 4(*n*/100)^(2/9)

(iv) Εκτίμηση Avar($\beta\hat{}_1$) εύρωστη προς αυτοσυσχέτιση

$$\left(\frac{\hat{\sigma}_{\hat{\beta}_1}}{\hat{\sigma}}\right)^2 \sqrt{\hat{v}}$$

(v) Μπορεί να χρησιμοποιηθεί σε στατιστικές t κλπ

Ιδιότητες

- Συνήθως εύρωστες τυπικές αποκλίσεις > κανονικές
- Προβληματική συμπεριφορά σε μικρά δείγματα (*n*<100)
- Απαιτείται επιλογή *g*
- ΕΕΤ είναι συχνά πολύ αναποτελεσματικοί όταν υπάρχει αυτοσυσχέτιση (υπάρχουν καλύτεροι εκτιμητές)

Π.χ. 12.7 (12.2,10.3,10.9)

$x_t \sim ΑΕΠ^{PR}_t$, $ΑΕΠ^{US}_t$, **μεταβλητή ελάχιστου μισθού**

$$\hat{u}_t = \beta_0 + \beta_1 x_{t1} + ... + .481\,\hat{u}_{t-1}$$
$$n = 37,\ p = .007$$
$$t_{\hat{\rho}} = 2.89$$
$$\hat{\beta}_1 = -.2123,\ \hat{\sigma}_{\hat{\beta}_1} = .0402,\ \hat{\sigma} = .0328$$
$$g = 2 \Rightarrow \hat{\nu} = 2$$
$$\hat{\sigma}_{\hat{\beta}_1}^{\ *} = .0426,\ t_{\hat{\beta}_1}^{\ *} = -4.98$$
$$\hat{\beta}_1^{CO} = -.1111,\ \hat{\sigma}_{\hat{\beta}_1}^{\ CO} = .0446$$

Παραβιάζεται ΑΕ ή αποτέλεσμα μικρού δείγματος;

ΕΤΕΡΟΣΚΕΔΑΣΤΙΚΟΤΗΤΑ ΣΕ ΠΑΛΙΝΔΡΟΜΗΣΕΙΣ ΧΡΟΝΟΣΕΙΡΩΝ
(Υ4΄: Var(u_t|x_t) = σ^2)

Στο υπόδειγμα: $y_t = \beta' x_t + u_t$ Υ1΄-3΄, Υ5΄ (E(u_t|x_t) = 0)

με $u_t^2 = \delta' x_t + \nu_t$

⇨Έλεγχος ομοσκεδαστικότητας (Υ4) *Breusch-Pagan* μπορεί να γίνει ως:

$$H_0 : \delta_1 = \delta_2 = ... = \delta_k = 0$$

βάσει $\hat{u}_t^2 = \hat{\delta}' x + \nu_t$ και χρησιμοποιώντας τη στατιστική *F* ασυμπτωτικά.

Παράδειγμα 12.8 (11.4): *Αποτελεσματικότητα Αγορών*

$$Αποδόσεις_t = .180 + .059 Αποδόσεις_{t-1}$$
$$(.081)\quad (.038)$$

n = 689, R^2= .0035, Εβδομάδες (Τετάρτη) Ιαν.'76- Μαρ. '89

H_0: β_1 = 0 (Αποτελεσματικότητα)

⟹ Δεν απορρίπτεται με βάση *t-stat*.

$$\hat{u}_t^2 = 4.66 - 1.104 αποδόσεις_{t-1} + \hat{\nu}_t$$
$$(0.43)\quad (0.201) \qquad\qquad n = 689,\ R^2=$$

Διακύμανση : εξαρτάται από παλαιότερες αποδόσεις
αυξάνει όταν αγορές πέφτουν

Δυναμική Ετεροσκεδαστικότητα

<u>Ορισμός:</u> $E\left(u_t^2 \mid u_{t-1}, u_{t-2}, ...\right) = f\left(u_{t-1}, u_{t-2}, ...\right)$

Παρατήρηση: Είναι συμβατή με τις υποθέσεις Gauss-Markov (Υ1΄-5΄) σε ***στατικά υποδείγματα***

π.χ. ***ARCH*** (συμβατό με Υ4΄):

$$E\left(u_t^2 \mid u_{t-1}, u_{t-2}, ..., x_t\right) = E\left(u_t^2 \mid u_{t-1}\right) = a_0 + a_1 u_{t-1}^2$$
$$\Leftrightarrow u_t^2 = a_0 + a_1 u_{t-1}^2 + \nu_t,\ E\left(\nu_t \mid u_{t-1}, u_{t-2}, ..., x_t\right) = 0$$

(ν_t, u_{t-1} έχουν εξάρτηση αφού πρέπει να ισχύει $\nu_t \geq -a_0 - a_1 u_{t-1}^2$)

Σημασία:

• Θεωρήματα (π.χ. Βέλτιστοι αμερόληπτοι γραμμικοί εκτιμητές) ισχύουν.

• Υπάρχουν όμως εκτιμήτριες περισσότερο αποδοτικές ασυμπτωτικά (μεροληπτικές αλλά συνεπείς)

• Συμπεριφορά διακύμανσης έχει οικονομική σημασία

*ARCH σε **δυναμικά υποδείγματα***

=> ασύμβατο με κλασική ομοσκεδαστικότητα (Υ4)

$$E(y_t \mid z_t, y_{t-1}, z_{t-1}, y_{t-1}, \ldots) = \beta_0 + \beta_1 z_t + \beta_2 y_{t-1} + \beta_3 z_{t-1}$$
$$Var(y_t \mid z_t, y_{t-1}, z_{t-1}, y_{t-1}, \ldots) = Var(u_t \mid z_t, y_{t-1}, z_{t-1}, y_{t-1}, \ldots)$$
$$= \alpha_0 + a_1 u_{t-1}^2$$

Σημασία:

• Δεν επηρεάζει αμεροληψία/συνέπεια

• Επηρεάζει συμπεριφορά:
- Στατιστικών ελέγχου
- Διακύμανσης Ε.Ε.Τ
- Εκτιμητών διακύμανσης

Π.χ. 12.9 (12.8, 11.4)

$$\hat{u}_t^2 = \underset{(.44)}{2.95} + \underset{(.036)}{.337}\hat{u}_{t-1}^2 + \hat{v}_t \qquad n = 688 \qquad R^2 = .114$$

$t_{\hat{\beta}_1} \approx 9 \Rightarrow ARCH(1)$ Καλύτερο υπόδειγμα για την διακύμανση από το πχ 12.8

$\hat{u}_t = .0014\hat{u}_{t-1}, \qquad t_{\hat{\rho}} = .038$

Συσχέτιση ≠ «Συσχέτιση Διακύμανσης»

Ετεροσκεδαστικότητα και αυτοσυσχέτιση μαζί

Λύσεις:

1) Μπορούμε να ελέγξουμε για ετεροσκεδαστικότητα σε ημιδιαφορισμένο υπόδειγμα όπου έχει αφαιρεθεί η αυτοσυσχέτιση (ΕΓΕΕΤ):

$y^*_t = (1-\rho)\beta_0 + \beta_1 x^*_t + e_t$, $\qquad t > 1$

$y^*_t = y_t - \rho\, y_{t-1}$, $x^*_t = x_t - \rho\, x_{t-1}$,

Και να χρησιμοποιήσουμε εκτιμήσεις για τυπικές αποκλίσεις παραμέτρων *εύρωστες προς ετεροσκεδαστικότητα*

2) Ανάπτυξη υποδείγματος για αυτοσυσχέτιση και ετεροσκεδαστικότητα

Παράδειγμα:

$$y_t = \beta_0 + \beta_1 x_{t1} + \dots + \beta_k x_{tk} + u_t$$
$$u_t = \sqrt{h_t} v_t$$
$$v_t = \rho v_{t-1} + e_t, |\rho| < 1$$
$$\mathrm{E}(e_t | \mathrm{X}) = 0, Var(e_t) = \sigma^2_e, \mathrm{E}(e_t e_{t-1}) = 0$$
$$h_t = f(x_t)$$
$$\Rightarrow Var(u_t) = \frac{\sigma^2_e}{1-\rho} h_t,$$
$$Var\left(\frac{u_t}{\sqrt{h_t}}\right) = Var(v_t) = \frac{\sigma^2_e}{1-\rho}$$

Οπότε η μετασχηματισμένη παλινδρόμηση:

$$\frac{y_t}{\sqrt{h_t}} = \frac{\beta_0}{\sqrt{h_t}} + \beta_1 \frac{x_{t1}}{\sqrt{h_t}} + \dots + \beta_k \frac{x_{tk}}{\sqrt{h_t}} + \frac{u_t}{\sqrt{h_t}}$$

Θα έχει ομοσκεδαστικά σφάλματα AR(1) αρκεί να ξέρουμε η να μπορούμε να εκτιμήσουμε *h* (την μορφή της ετεροσκεδαστικότητας)

Εφαρμόζουμε προηγούμενες λύσεις για αυτοσυσχέτιση

Εφικτός ΓΕΕΤ με ετεροσκεδασικότητα και αυτοσυσχέτιση

Διαδικασία Εφικτής ΓΕΕΤ

1) Παλινδρόμηση με ΕΕΤ για προσδιορισμό καταλοίπων

2) Παλινδρόμηση log(κατάλοιπα) σε x_t, x_{t-1}…

για εκτίμηση καταλοίπων *g*

3) Προσδιορισμός *f* ως *f*=exp(*g*)

4) Εκτίμηση μετασχηματισμένης παλινδρόμησης με ΕΓΕΕΤ υπο αυτοσυσχέτιση (Cochrane-Orcutt / Prais-Winsten)

Οι ιδιότητες του ΕΓΕΕΤ υπό αυτοσυσχέτιση ισχύουν

Κεφάλαιο18
Προχωρημένα θέματα στα υποδείγματα χρονοσειρών

ΟΛΟΚΛΗΡΩΣΗ-ΕΛΕΓΧΟΣ ΓΙΑ ΟΛΟΚΛΗΡΩΣΗ ΠΡΩΤΟΥ ΒΑΘΜΟΥ, I(1)

Στο υπόδειγμα αυτοπαλινδρόμησης 1ου βαθμού:

$y_t = a + \rho y_{t-1} + e_t, t = 1,2,....$

$E(e_t \mid y_{t-1}, y_{t-2},...., y_0) = 0$ ~ *Martingale difference sequence* (ως προς $\{y_{t-1}, y_{t-2},...\}$)

π.χ. $e_t \overset{iid}{\sim} D(0)$, ανεξάρτητο του y_0

I(1) ↔ $\rho = 1$

$\alpha = 0$ (Τυχαία Διαδρομή) $\alpha \neq 0$ (Τ.Δ. με τάση)

$H_0: \rho = 1$ $H_1: \rho < 1 \rightarrow$ I(0) (αν $|\rho| < 1$)

$\rho > 1$, μη ρεαλιστικό

Συνήθως περίπτωση $0 < \rho < 1$ έχει ενδιαφέρον (όχι $\rho < 1$)

Έλεγχος Dickey - Fuller

$\Rightarrow \Delta y_t = a + \theta y_{t-1} + e_t,$ $H_0: \theta = 0$, $H_1: \theta < 0$

Δυσκολία:

$t_{\hat{\theta}}$
- δεν έχει κατανομή t
- έχει ασυμπτωτική κατανομή *Dickey-Fuller*

Παράδειγμα 18.2

$\widehat{\Delta Επιτόκια3M_t} = .625 - .091 Επιτόκια3M_{t-1}$
(.261) (.037)

$n = 123$

$R^2 = .048$

$t_{\hat{\theta}} = -.091/.037 = -2.46 \rightarrow$ Κατανομή;

~ Από ασυμπτωτική κατανομή *Dickey-Fuller*

Επίπεδο Σημασίας (Ε.Σ.) %	1	2.5	5	10
Κριτική Τιμή	-3.43	-3.12	-2.86	-2.57

→ Δεν απορρίπτουμε $H_0 : \theta = 0$ σε Ε.Σ. 10%

$$\Updownarrow$$

$$\rho = 1$$

• Δεν απορρίπτουμε και $\rho = 0.9$

→ Συμπεριφορά επιτοκίων πολύ διαφορετική:

$$\rho = 0.9 : Corr(\text{ΕΠ}_{t+10}, \text{ΕΠ}_t) \approx 0.35$$
$$\rho = 1 : Corr(\text{ΕΠ}_{t+10}, \text{ΕΠ}_t) \approx 1$$

Επιπτώσεις Ολοκλήρωσης

→ *Παραβιάζεται αδύναμη εξάρτηση*

→ Στατιστική σημασία: *Ασυνέπεια Ε.Ε.Τ*

→ Οικονομική σημασία: Αντίκτυπος μεταβολών είναι μόνιμος

ΟΛΟΚΛΗΡΩΣΗ-ΕΛΕΓΧΟΣ ΓΙΑ ΟΛΟΚΛΗΡΩΣΗ ΠΡΩΤΟΥ ΒΑΘΜΟΥ, I(1)

<u>Στο υπόδειγμα αυτοπαλινδρόμησης βαθμού *k για Δy*:</u>

$$\Delta y_t = a_0 + \rho_1 y_{t-1} + \ldots + \rho_k \Delta y_{t-k} + e_t$$

$$H_0 : \mathrm{I}(1) \Leftrightarrow \rho_1 = 1 \Leftrightarrow \begin{cases} \Delta y_t = \beta_0 + \theta y_{t-1} + \gamma_1 \Delta y_{t-1} + \ldots + \gamma_p \Delta y_{t-p} \\ H_0 : \theta = 0 \end{cases}$$

<u>Έλεγχος *Augmented Dickey – Fuller (ADF)*:</u>

$$H_0 : \theta = 0$$
$$H_1 : \theta < 0$$

$t_{\hat{\theta}}$ ~ κατανομή *D-F*, *αρκεί να έχουμε αρκετές υστερήσεις*

$t_{\hat{\gamma}}$ ~ κατανομή *t* όπως συνήθως:

→ Μπορούμε να κάνουμε έλεγχο *F* για αρκετές υστερήσεις

→ Ώστε να αποφύγουμε υπερβολικά πολλές ή υπερβολικά λίγες υστερήσεις

Παράδειγμα 18.3

Παλινδρόμηση *ADF*:

$$\widehat{\Delta \Pi\lambda\eta\theta_t} = 1.36 - .310\Pi\lambda\eta\theta_{t-1} + .138\Delta\Pi\lambda\eta\theta_{t-1}$$
$$(.517)\ (.103)\qquad (.126)$$

1948-76

n=47

R^2=.172

$t_{\hat{\theta}} = -.310/.103 = -3.01 \Rightarrow$ Απόρριψη σε Ε.Σ. 5%

$t_{\hat{\gamma}} \approx 1.1 \Rightarrow$ Δεν απορρίπτουμε H_0: $\gamma = 0$

→ Προτιμούμε την παλινδρόμηση *D.F.*

$$\widehat{\Delta \Pi\lambda\eta\theta_t} = \ldots - .335\Pi\lambda\eta\theta_{t-1}$$

$t_{\hat{\theta}} = -3.13 \Rightarrow$ Απόρριψη ενισχύεται

Είναι αναμενόμενη η ενίσχυση;

<u>Σε αυτοπαλινδρόμηση με τάση:</u>

$$y_t = a_0 + \delta t + \rho y_{t-1} + e_t \Rightarrow \Delta y_t = a_0 + \delta t + \theta y_{t-1} + e_t$$

$$H_0 : \theta = 0 \Rightarrow$$
$$\Rightarrow \Delta y_t = a_0 + \delta t + e_t$$
$$\Rightarrow E(y_t) \sim 2^{o}\text{βάθμιο πολυώνυμο στο } t$$

(*Ασυνήθιστο*)

$$\Rightarrow H_0 : \begin{Bmatrix} \theta = 0 \\ \delta = 0 \end{Bmatrix}$$

•*Στατιστική F αυτού του ελέγχου έχει περίεργη ασυμπτωτική συμπεριφορά*

οπότε απλοποιούμε κάνοντας τον απλούστερο έλεγχο:

$$H_0 : \theta = 0$$

Όμως η χρονική τάση αλλάζει τις κριτικές τιμές (πιο δύσκολο να απορρίψουμε την *Ho*)

Ασυμπτωτική κατανομή $t_{\hat{\theta}}$ σε υπόδειγμα με τάση

Επίπεδο Σημασίας (Ε.Σ.) %	1	2.5	5	10
Κριτική Τιμή	-3.96	-3.66	-3.41	-3.12

Παράδειγμα 18.4

$$\widehat{gGDP_t} = 1.65 + .0059t - .210 logGDP_{t-1} + .264 gGDP_{t-2}$$
$$(.67) \quad (.0027) \quad (.087) \qquad (.165)$$

$n = 35$
$R^2 = .268$
Η.Π.Α: 1959-95

$$t_{\hat{\theta}} = -.210/.087 = -2.41$$

- Δεν απορρίπτουμε
- Όμως ρ^=0.79 << 1
- Πιθανώς λόγω μικρού δείγματος

Άλλες παρατηρήσεις

Η κατανομή της εκτίμησης για το δ είναι ιδιάζουσα

(Αν παραλείψουμε την τάση τα αποτελέσματα δεν αλλάζουν)

Διάφοροι άλλοι έλεγχοι πού μπορούν να γίνουν δίνουν παρόμοια αποτελέσματα

Συμπέρασμα

Δεν έχουμε αρκετές παρατηρήσεις για να έχουμε έντονη άποψη για την συμπεριφορά του GDP

ΦΑΙΝΟΜΕΝΙΚΗ ΣΥΣΧΕΤΙΣΗ ΚΑΙ ΠΑΛΙΝΔΡΟΜΗΣΗ

Είδη Φ.Σ.:

1. «Κλασική»:

$$\exists Z, X, Y : |Corr(X,Y)| >> 0$$
$$Corr(X - Z, Y - Z) \approx 0$$

(X,Y) σχετίζονται ***έμμεσα*** λόγω σχέσης με Z *(π.χ. τάση)*

2. Σε χρονοσειρές:

Αν X,Y~I(1) αλλά τις αντιμετωπίσουμε ως I(0), τότε *μοιάζουν* συσχετισμένες ακόμα και αν είναι **ανεξάρτητες**

Π.χ.

$$x_t = x_{t-1} + a_t, a_t \overset{iid}{\sim} D(0, \sigma_a^2)$$
$$y_t = y_{t-1} + e_t, e_t \overset{iid}{\sim} D'(0, \sigma_e^2)$$
$$x_0 = y_0 = 0$$
$$\{a_t\}, \{e_t\} \text{ ανεξάρτητες Σ.Α.}$$

Αν παλινδρομήσουμε $y_t = \beta_0 + \beta_1 x_t + u_t$

Παραβιάζεται η Στασιμότητα / Ασθ. Εξάρτηση

$\beta_1 = 0 \Rightarrow u_t = \sum_{k=1}^{t} e_k$ ~ *Τυχαία Διαδρομή (και β_0=0)*

*plimβ^* = ***0* ?**

1. Η στατιστική $t_{\hat{\beta}_1}$ δεν έχει αναμενόμενη κατανομή
2. Ακόμη και ασυμπτωτικά, $p\lim_{n\to\infty}\left(t_{\hat{\beta}_1}\right)=\infty$
3. Έλεγχος σημαντικότητας οδηγεί σε <u>λάθος συμπεράσματα</u>
4. Παρομοίως για R^2 : $p\lim_{n\to\infty} R^2 >> 1-\sigma_u^2 / \sigma_y^2$

(Άρα και για F, αφού: $F=\left(1-R^2\right)\frac{n-k-1}{q}$)

Ο ίδιος συλλογισμός ισχύει και για παλινδρομήσεις με πολλές μεταβλητές

ΣΥΝΟΛΟΚΛΗΡΩΣΗ

<u>Ορισμός:</u> $\{y_t\}_{t=1}^{\infty},\{x_t\}_{t=1}^{\infty} \sim I(1)$

$\exists\beta^* : y_t-\beta^* x_t \sim I(0)$

→ «Παράμετρος συνολοκλήρωσης»: β^*

<u>Οικονομική Ερμηνεία</u>

- $\left(y_t-\beta^* x_t\right)$ διακυμαίνεται γύρω από $E\left(y_t-\beta^* x_t\right)$
- $E\left(y_t-\beta^* x_t\right)$ είναι το «μακροπρόθεσμο σημείο ισορροπίας», π.χ. Διαφορά μηνιαίων/εξαμηνιαίων επιτοκίων.

Παράδειγμα 18.2 (ΣΥΝΕΧΕΙΑ)

$$\Delta\widehat{Επιτόκια3M_t} = .625 - .091 Επιτόκια3M_{t-1}$$
$$(.261)\quad(.037)$$

n = 123
R^2 =.048

$t_{\hat{\theta}} = -.091/.037 = -2.46$ → Ολοκλήρωση

Αντίστοιχα και για 6M_t

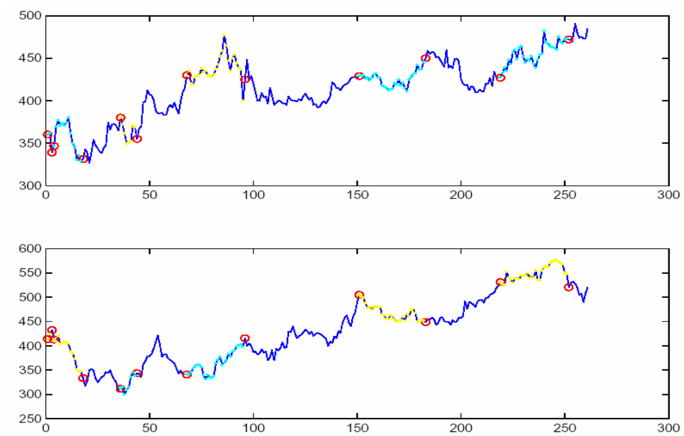
Όμως βρίσκουμε ότι:

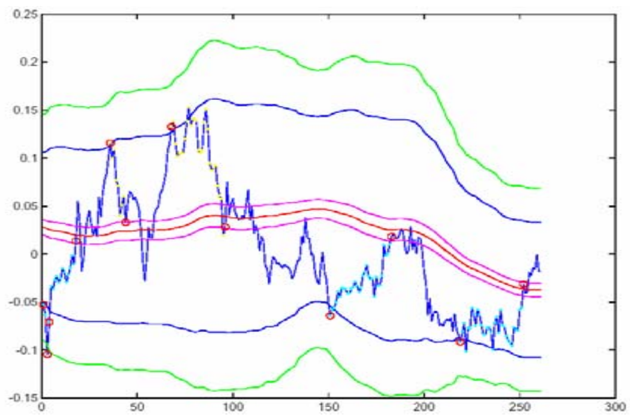
Δ(6M_t- 3M_t) =α -.67 (6M_t- 3M_t)

⇒t_stat=-7.71 (με κατανομή D-F)

⇒ Στασιμότητα I(0) – υπάρχει σημείο ισορροπίας (μακροπρόθεσμα)

Παράδειγμα: χρονοσειρές τιμών μετοχών και χαρτοφυλακίων





Ερμηνεία έννοιας «μακροπρόθεσμου σημείου ισορροπίας»

Ορίζουμε $\mu = E(6M\_t - 3M\_t)$

$\Rightarrow$ $6M\_t = 3M\_t + \mu + e\_t$

$E(e\_t)=0$

$e\_t \sim I(0)$

Τότε το ΜΣΙ επιτυγχάνεται όταν $e\_t = 0$

Όποιες αποκλίσεις αναμένεται να αντιστραφούν

Έλεγχοι Συνολοκλήρωσης

1. Με συγκεκριμένη παράμετρο (από οικονομική θεωρία):

$\Rightarrow$ Έλεγχος στασιμότητας $(y_t - \beta^* x_t)$ με $\left\{\begin{matrix} D\text{-}F \\ Augmented\ D.F. \end{matrix}\right\}$

2. Με άγνωστη παράμετρο:

i. Παλινδρομούμε $\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 x_t + \hat{u}_t$

Υπό γενικές συνθήκες $p\lim_{n\to\infty} \hat{\beta}_1 = \beta^*$

Πρόβλημα: Υπό *Ho* η παλινδρόμηση μας εμφανίζει φαινομενική συσχέτιση

=> Κλασσικές υποθέσεις δεν ισχύουν

Όμως έλεγχοι τύπου D-F μπορούν παρ’ όλα αυτά να γίνουν στα κατάλοιπα

ii. Κάνουμε έλεγχο τύπου *D-F* στα κατάλοιπα:

$$\Delta\hat{u}_t = a + \theta\hat{u}_{t-1} + e_t$$

Όμως $t_{\hat{\theta}} \nsim$ *Dickey - Fuller;*

$\boxed{t_{\hat{\theta}} \sim \textit{Engle/Granger} \text{ ασυμπτωτικά}}$

Απαιτεί μεγαλύτερο $|t_{\hat{\theta}}|$ για απόρριψη, δηλ.:

$|t_{\hat{\theta}}| <$ Κριτ.Τιμή $\Rightarrow \theta = 0 \Rightarrow \hat{u}_t \sim I(1) \Rightarrow$ Όχι Συνολ/ση

Επίπεδο Σημασίας (Ε.Σ.) %	1	2.5	5	10
Κριτική Τιμή	-3.90	-3.59	-3.34	-3.04

3. Με άγνωστη παράμετρο και πιθανή τάση:

i. Εκτιμούμε $\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 x_t + \hat{u}_t$

ii. Βρίσκουμε $t_{\hat{\theta}}$ στο $\Delta\hat{u}_t = a + \theta\hat{u}_{t-1} + e_t$

$\Rightarrow$ $t_{\hat{\theta}}$ έχει άλλη αλλά γνωστή κατανομή ασυμπτωτικά

Επίπεδο Σημασίας (Ε.Σ.) %	1	2.5	5	10
Κριτική Τιμή	-4.32	-4.03	-3.78	-3.50

Παράδειγμα 18.5 (11.8, 11.6,10.4)

Γεννητικότητα(t) = β_0 +.187Απαλλαγές(t) + β_2 t
R^2=.5 (.035)

ΔΓεννητικότητα(t) = β_0' - .043 ΔΑπαλλαγές(t)
R^2=.032 (.028)

Γιατί τόσο διαφορετικά αποτελέσματα;
ADF (1 υστέρηση και τάση) Γεννητικότητα => I(1)
ADF (1 υστέρηση και τάση) Απαλλαγές => I(1)

Μήπως υπάρχει σχέση συνολοκλήρωσης;

Στα κατάλοιπα από την πρώτη παλινδρόμηση
ADF με 1 υστέρηση => *t_stat* = -2.43 (ΚΤ –3.50 σε ΕΣ 10%)

$\Rightarrow$ Δεν μπορούμε να απορρίψουμε **έλλειψη συνολοκλήρωσης**

Εξήγηση προηγούμενων αποτελεσμάτων;
Σχέση επιπέδου: *φαινομενική συσχέτιση*
Σχέση διαφορών: χρειάζεται

Πως μπορούμε να έχουμε εκτιμητές με κανονικές κατανομές;

Νέα παλινδρόμηση με «περίπου» κλασσικές υποθέσεις

$$y_t = \beta_0 + \beta_1 x_t + \delta' z_t + u_t$$

$$y_t, x_t \text{ ειναι I(1)}$$

$$\mathrm{E}(u_t \Delta x_s) = 0$$

είδος εξωγένειας

Ικανοποιείται με εκτιμητή 'lead-lags', για τον οποίο:

z(t)=[Δx(t),Δx(t+1), Δx(t+2),…, Δx(t-1), Δx(t-2)…]'

<u>Παράδειγμα 18.6 (18.2, κλπ)</u>

Από προηγούμενη παλινδρόμηση

$\Delta Επιτόκια3M_t$ = .625 -.091$Επιτόκια3M_{t-1}$

(.261) (.037)

έχουμε απορρίψει έλλειψη ολοκλήρωσης.

Από την παλινδρόμηση:

Δ(6M_t- 3M_t) =α -.67 (6M_t- 3M_t)

έχουμε απορρίψει έλλειψη συνολοκλήρωσης (με βάση οικονομικό συλλογισμό – υπόθεση προσδοκιών)

$Επιτόκια6M_t$ = a +1.026$Επιτόκια3M_{t-1}$

(0.077)

(αλλά στατιστική t ίσως δεν ισχύει)

με lead-lag εκτιμητή (με δύο lead-lags) δίνει

$Επιτόκια6M_t$ = a +1.038$Επιτόκια3M_{t-1}$ +δ'z $_t$

(0.081)

z $_t$ = [$\Delta 3M_t$,$\Delta 3M_{t+1}$,$\Delta 3M_{t+2}$,$\Delta 3M_{t-1}$,$\Delta 3M_{t-2}$]'

(1.038-1)/.0081 = 4.69 => *στατιστικά* διαφορετικό από το 1 το οποίο θα περιμέναμε από θεωρία

Είναι όμως *οικονομικά* διαφορετικό;

Εφικτά υποδείγματα

$\{y_t\}_{t=1}^{\infty}, \{x_t\}_{t=1}^{\infty}$

Χωρίς Συνολοκλήρωση

$$\Delta y_t = \beta_0 + \beta_1 \Delta x_t \Rightarrow$$

$\Rightarrow$ Σχέση με αλλαγές εξαρτημένης, όχι επιπέδου

Με Συνολοκλήρωση

$$\Delta y_t = \beta_0 + \beta_1 \Delta x_t + \left(y_{t-1} - \beta^* x_{t-1}\right) \Rightarrow$$

$\Rightarrow$ Σχέση με *αλλαγές και επίπεδο εξαρτημένης* (ΥΔΛ)

Υποδείγματα Διόρθωσης Λαθών

$$\Delta y_t = a_0 + a_1 \Delta y_{t-1} + \gamma_0 \Delta x_t + \gamma_1 \Delta x_{t-1} + \delta\left(y_{t-1} - \beta x_{t-1}\right) + u_t$$

Ερμηνεία «όρου διόρθωσης λαθών»:

Εκφράζει πώς κινείται το y_t για να επανέλθει στη μακροχρόνια ισορροπία.

Π.χ. $\delta < 0, y_{t-1} > \beta x_{t-1} \Rightarrow y_{t-1}$ μεγαλύτερο από ότι προβλέπει η σχέση του με το x_{t-1} και θα μικρύνει

Δυσκολία:

Η παράμετρος β είναι άγνωστη

Παράδειγμα 18.7 (18.6, 18.2, κλπ)

Από προηγούμενη παλινδρόμηση

$\Delta Επιτόκια3M_t$ = .625 -.091$Επιτόκια3M_{t-1}$
(.261) (.037)

έχουμε απορρίψει έλλειψη ολοκλήρωσης.

Αν δεχτούμε υπόθεση προσδοκιών, τότε το κατάλληλο ΥΔΛ είναι:

$\Delta 6M_t = \alpha_0 + \gamma_0 \Delta 3M_{t-1} + \delta (6M_{t-1} - 3M_{t-2}) + u_t$

Εκτίμηση δίνει:

$\Delta 6M_t = .09 + 1.218\, \Delta 3M_{t-1} - .840\, (6M_{t-1} - 3M_{t-2})$

(.043) (.264) (.244)

n = 122 (τρίμηνα)

R^2=.790

δ^ δεν διαφέρει από 1 => όλη η διόρθωση γίνεται σε ένα τρίμηνο

Εκτίμηση β: «Διαδικασία 2 βημάτων *Engle-Granger*»

1. Εκτιμούμε $\hat{\beta}$ (πχ με κλασσική παλινδρόμηση)
2. Εκτιμούμε χρησιμοποιώντας $\left(y_{t-1} - \hat{\beta} x_{t-1}\right)$ ως μεταβλητή

⇨ Η ασυμπτωτική κατανομή των $\hat{a}, \hat{\gamma}$:

→ Δεν επηρεάζεται από το πρώτο βήμα

→ Είναι αποτελεσματική

Κεφάλαιο18
Προχωρημένα θέματα στα υποδείγματα χρονοσειρών

ΟΛΟΚΛΗΡΩΣΗ-ΕΛΕΓΧΟΣ ΓΙΑ ΟΛΟΚΛΗΡΩΣΗ ΠΡΩΤΟΥ ΒΑΘΜΟΥ, I(1)

<u>Στο υπόδειγμα αυτοπαλινδρόμησης 1ου βαθμού:</u>

$$y_t = a + \rho y_{t-1} + e_t, t = 1,2,....$$

$$E(e_t \mid y_{t-1}, y_{t-2}, ..., y_0) = 0$$ ~ *Martingale difference sequence* (ως προς $\{y_{t-1}, y_{t-2}, ...\}$)

π.χ. $e_t \overset{iid}{\sim} D(0)$, ανεξάρτητο του y_0

I(1) ↔ $\rho = 1$

$\alpha = 0$ (Τυχαία Διαδρομή) $\alpha \neq 0$ (Τ.Δ. με τάση)

$H_0 : \rho = 1$

($\rho > 1$, μη ρεαλιστικό)

$H_1 : \rho < 1 \rightarrow$ I(0) (αν $|\rho| < 1$)

(Συνήθως περίπτωση $0 < \rho < 1$ έχει ενδιαφέρον (όχι $\rho < 1$))

Έλεγχος Dickey - Fuller

$$\Rightarrow \Delta y_t = a + \theta y_{t-1} + e_t, \qquad H_0 : \theta = 0, \quad H_1 : \theta < 0$$

<u>Δυσκολία:</u>

$t_{\hat{\theta}}$
- δεν έχει κατανομή t
- έχει ασυμπτωτική κατανομή *Dickey-Fuller*

Παράδειγμα 18.2

$$\Delta\widehat{Επιτόκια3M_t} = .625 - .091 Επιτόκια3M_{t-1}$$
$$(.261)\ (.037)$$

$n = 123$
$R^2 = .048$

$t_{\hat{\theta}} = -.091/.037 = -2.46$ → Κατανομή;

~ Από ασυμπτωτική κατανομή *Dickey-Fuller*

Επίπεδο Σημασίας (Ε.Σ.) %	1	2.5	5	10
Κριτική Τιμή	-3.43	-3.12	-2.86	**-2.57**

→ Δεν απορρίπτουμε $H_0 : \theta = 0$ σε Ε.Σ. 10%

$$\Updownarrow$$

$$\rho = 1$$

• Δεν απορρίπτουμε και $\rho = 0.9$

→ Συμπεριφορά επιτοκίων πολύ διαφορετική:

$$\rho = 0.9 : Corr(\text{ΕΠ}_{t+10}, \text{ΕΠ}_t) \approx 0.35$$
$$\rho = 1 : Corr(\text{ΕΠ}_{t+10}, \text{ΕΠ}_t) \approx 1$$

Επιπτώσεις Ολοκλήρωσης

→ *Παραβιάζεται αδύναμη εξάρτηση*

→ Στατιστική σημασία: *Ασυνέπεια Ε.Ε.Τ*

→ Οικονομική σημασία: Αντίκτυπος μεταβολών είναι μόνιμος

ΟΛΟΚΛΗΡΩΣΗ-ΕΛΕΓΧΟΣ ΓΙΑ ΟΛΟΚΛΗΡΩΣΗ ΠΡΩΤΟΥ ΒΑΘΜΟΥ, Ι(1)

Στο υπόδειγμα αυτοπαλινδρόμησης βαθμού *k για Δy*:

$$\Delta y_t = a_0 + \rho_1 y_{t-1} + \ldots + \rho_k \Delta y_{t-k} + e_t$$

$$H_0 : \mathrm{I}(1) \Leftrightarrow \rho_1 = 1 \Leftrightarrow \begin{cases} \Delta y_t = \beta_0 + \theta y_{t-1} + \gamma_1 \Delta y_{t-1} + \ldots + \gamma_p \Delta y_{t-p} \\ H_0 : \theta = 0 \end{cases}$$

Έλεγχος *Augmented Dickey – Fuller (ADF)*:

$$H_0 : \theta = 0$$
$$H_1 : \theta < 0$$

$t_{\hat{\theta}}$ ~ κατανομή *D-F*, *αρκεί να έχουμε αρκετές υστερήσεις*

$t_{\hat{\gamma}}$ ~ κατανομή *t* όπως συνήθως:
→ Μπορούμε να κάνουμε έλεγχο *F* για αρκετές υστερήσεις

→ Ώστε να αποφύγουμε υπερβολικά πολλές ή υπερβολικά λίγες υστερήσεις

Παράδειγμα 18.3

Παλινδρόμηση *ADF*:

1948-76

$$\widehat{\Delta \Pi\lambda\eta\theta_t} = 1.36 - .310\Pi\lambda\eta\theta_{t-1} + .138\Delta\Pi\lambda\eta\theta_{t-1}$$
$$(.517) \quad (.103) \qquad (.126)$$

n=47

R^2=.172

$t_{\hat{\theta}} = -.310/.103 = -3.01 \Rightarrow$ Απόρριψη σε Ε.Σ. 5%

$t_{\hat{\gamma}} \approx 1.1 \Rightarrow$ Δεν απορρίπτουμε H_0: $\gamma = 0$

→ Προτιμούμε την παλινδρόμηση *D.F.*

$$\widehat{\Delta \Pi\lambda\eta\theta_t} = \ldots - .335\Pi\lambda\eta\theta_{t-1}$$

$t_{\hat{\theta}} = -3.13 \Rightarrow$ Απόρριψη ενισχύεται

Είναι αναμενόμενη η ενίσχυση;

Σε αυτοπαλινδρόμηση με τάση:

$$y_t = a_0 + \delta t + \rho y_{t-1} + e_t \Rightarrow \Delta y_t = a_0 + \delta t + \theta y_{t-1} + e_t$$

$$H_0 : \theta = 0 \Rightarrow$$
$$\Rightarrow \Delta y_t = a_0 + \delta t + e_t$$
$\Rightarrow E(y_t)$ ~ 2ºβάθμιο πολυώνυμο στο *t*
(*Ασυνήθιστο*)

$$\Rightarrow H_0 : \begin{Bmatrix} \theta = 0 \\ \delta = 0 \end{Bmatrix}$$

•*Στατιστική F αυτού του ελέγχου έχει περίεργη ασυμπτωτική συμπεριφορά* οπότε απλοποιούμε κάνοντας τον απλούστερο έλεγχο:

$$H_0 : \theta = 0$$

Όμως η χρονική τάση αλλάζει τις κριτικές τιμές (πιο δύσκολο να απορρίψουμε την *Ho*)

Ασυμπτωτική κατανομή $t_{\hat{\theta}}$ σε υπόδειγμα με τάση

Επίπεδο Σημασίας (Ε.Σ.) %	1	2.5	5	10
Κριτική Τιμή	-3.96	-3.66	-3.41	-3.12

Παράδειγμα 18.4

$$\widehat{gGDP}_t = 1.65 + .0059t - .210 logGDP_{t-1} + .264 gGDP_{t-2}$$
$$(.67) \quad (.0027) \quad (.087) \quad (.165)$$

$n = 35$
$R^2 = .268$
Η.Π.Α: 1959-95

$$t_{\hat{\theta}} = -.210/.087 = -2.41$$

- Δεν απορρίπτουμε
- Όμως ρ^=*0.79* << 1
- Πιθανώς λόγω μικρού δείγματος

Άλλες παρατηρήσεις

Η κατανομή της εκτίμησης για το δ είναι ιδιάζουσα

(Αν παραλείψουμε την τάση τα αποτελέσματα δεν αλλάζουν)

Διάφοροι άλλοι έλεγχοι πού μπορούν να γίνουν δίνουν παρόμοια αποτελέσματα

Συμπέρασμα

Δεν έχουμε αρκετές παρατηρήσεις για να έχουμε έντονη άποψη για την συμπεριφορά του GDP

ΦΑΙΝΟΜΕΝΙΚΗ ΣΥΣΧΕΤΙΣΗ ΚΑΙ ΠΑΛΙΝΔΡΟΜΗΣΗ

Είδη Φ.Σ.:

1. «Κλασική»:

$$\exists Z, X, Y : |Corr(X,Y)| >> 0$$
$$Corr(X - Z, Y - Z) \approx 0$$

(*X,Y*) σχετίζονται ***έμμεσα*** λόγω σχέσης με *Z (π.χ. τάση)*

2. Σε χρονοσειρές:

Αν *X,Y*~I(1) αλλά τις αντιμετωπίσουμε ως I(0), τότε *μοιάζουν* συσχετισμένες ακόμα και αν είναι **ανεξάρτητες**

Π.χ.

$$x_t = x_{t-1} + a_t, a_t \overset{iid}{\sim} D(0,\sigma_a^2)$$

$$y_t = y_{t-1} + e_t, e_t \overset{iid}{\sim} D'(0,\sigma_e^2)$$

$$x_0 = y_0 = 0$$

$\{a_t\},\{e_t\}$ ανεξάρτητες Σ.Α.

Αν παλινδρομήσουμε $y_t = \beta_0 + \beta_1 x_t + u_t$

Παραβιάζεται η Στασιμότητα / Ασθ. Εξάρτηση

$\beta_1 = 0 \Rightarrow u_t = \sum_{k=1}^{t} e_k \sim$ *Τυχαία Διαδρομή (και β_0=0)*

*plimβ^* = ***0*** **?**

1. Η στατιστική $t_{\hat{\beta}_1}$ δεν έχει αναμενόμενη κατανομή
2. Ακόμη και ασυμπτωτικά, $p\lim_{n\to\infty}\left(t_{\hat{\beta}_1}\right) = \infty$
3. Έλεγχος σημαντικότητας οδηγεί σε <u>λάθος συμπεράσματα</u>
4. Παρομοίως για R^2 : $p\lim_{n\to\infty} R^2 >> 1 - \sigma_u^2 / \sigma_y^2$

(Άρα και για *F*, αφού: $F = (1-R^2)\dfrac{n-k-1}{q}$)

Ο ίδιος συλλογισμός ισχύει και για παλινδρομήσεις με πολλές μεταβλητές

ΣΥΝΟΛΟΚΛΗΡΩΣΗ

<u>Ορισμός:</u> $\{y_t\}_{t=1}^{\infty}, \{x_t\}_{t=1}^{\infty} \sim I(1)$

$\exists \beta^* : y_t - \beta^* x_t \sim I(0)$

→ «Παράμετρος συνολοκλήρωσης»: β^*

<u>Οικονομική Ερμηνεία</u>

- $(y_t - \beta^* x_t)$ διακυμαίνεται γύρω από $E(y_t - \beta^* x_t)$
- $E(y_t - \beta^* x_t)$ είναι το «μακροπρόθεσμο σημείο ισορροπίας», π.χ. Διαφορά μηνιαίων/εξαμηνιαίων επιτοκίων.

Παράδειγμα 18.2 (ΣΥΝΕΧΕΙΑ)

$\widehat{\Delta Επιτόκια3M_t}$ = .625 -.091*Επιτόκια3M*$_{t-1}$
(.261) (.037)

n = 123
R^2 =.048

$t_{\hat{\theta}} = -.091/.037 = -2.46$ → Ολοκλήρωση

Αντίστοιχα και για 6M_t

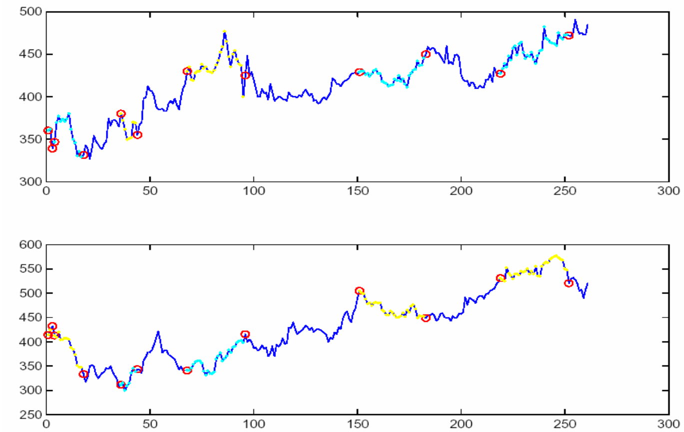
Όμως βρίσκουμε ότι:

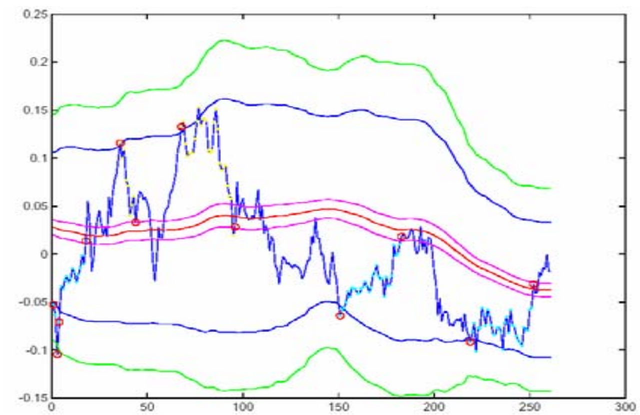
Δ(6M_t- 3M_t) =α -.67 (6M_t- 3M_t)

⇒t_stat=-7.71 (με κατανομή D-F)

⇒ Στασιμότητα I(0) – υπάρχει σημείο ισορροπίας (μακροπρόθεσμα)

Παράδειγμα: χρονοσειρές τιμών μετοχών και χαρτοφυλακίων





Ερμηνεία έννοιας «μακροπρόθεσμου σημείου ισορροπίας»

Ορίζουμε $\mu = E(6M\_t - 3M\_t)$

$\Rightarrow$ $6M\_t = 3M\_t + \mu + e\_t$

$E(e\_t)=0$

$e\_t \sim I(0)$

Τότε το ΜΣΙ επιτυγχάνεται όταν $e\_t = 0$

Όποιες αποκλίσεις αναμένεται να αντιστραφούν

Έλεγχοι Συνολοκλήρωσης

1. Με συγκεκριμένη παράμετρο (από οικονομική θεωρία):

$\Rightarrow$ Έλεγχος στασιμότητας $\left(y_t - \beta^* x_t\right)$ με $\left\{\begin{matrix} D\text{-}F \\ Augmented\ D.F. \end{matrix}\right\}$

2. <u>Με άγνωστη παράμετρο:</u>

i. Παλινδρομούμε $\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 x_t + \hat{u}_t$

Υπό γενικές συνθήκες $p\lim_{n\to\infty} \hat{\beta}_1 = \beta^*$

Πρόβλημα: Υπό *Ho* η παλινδρόμηση μας εμφανίζει φαινομενική συσχέτιση

=> Κλασσικές υποθέσεις δεν ισχύουν

Όμως έλεγχοι τύπου D-F μπορούν παρ' όλα αυτά να γίνουν στα κατάλοιπα

ii. Κάνουμε έλεγχο τύπου *D-F* στα κατάλοιπα:

$$\Delta\hat{u}_t = a + \theta\hat{u}_{t-1} + e_t$$

Όμως $t_{\hat{\theta}} \nsim$ *Dickey - Fuller;*

$\boxed{t_{\hat{\theta}} \sim \textit{Engle/Granger} \text{ ασυμπτωτικά}}$

Απαιτεί μεγαλύτερο $|t_{\hat{\theta}}|$ για απόρριψη, δηλ.:

$|t_{\hat{\theta}}| <$ Κριτ.Τιμή $\Rightarrow \theta = 0 \Rightarrow \hat{u}_t \sim I(1) \Rightarrow$ Όχι Συνολ/ση

Επίπεδο Σημασίας (Ε.Σ.) %	1	2.5	5	10
Κριτική Τιμή	-3.90	-3.59	-3.34	-3.04

3. <u>Με άγνωστη παράμετρο και πιθανή τάση:</u>

i. Εκτιμούμε $\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 x_t + \hat{u}_t$

ii. Βρίσκουμε $t_{\hat{\theta}}$ στο $\Delta\hat{u}_t = a + \theta\hat{u}_{t-1} + e_t$

$\Rightarrow$ $t_{\hat{\theta}}$ έχει <u>άλλη</u> αλλά γνωστή κατανομή ασυμπτωτικά

Επίπεδο Σημασίας (Ε.Σ.) %	1	2.5	5	10
Κριτική Τιμή	-4.32	-4.03	-3.78	-3.50

<u>Παράδειγμα 18.5 (11.8, 11.6,10.4)</u>

Γεννητικότητα(t) = β_0 +.187Απαλλαγές(t) + β_2 t
R^2=.5 (.035)

ΔΓεννητικότητα(t) = β_0' - .043 ΔΑπαλλαγές(t)
R^2=.032 (.028)

Γιατί τόσο διαφορετικά αποτελέσματα;
ADF (1 υστέρηση και τάση) Γεννητικότητα => I(1)
ADF (1 υστέρηση και τάση) Απαλλαγές => I(1)

Μήπως υπάρχει σχέση συνολοκλήρωσης;

Στα κατάλοιπα από την πρώτη παλινδρόμηση

ADF με 1 υστέρηση => *t_stat* = -2.43 (ΚΤ –3.50 σε ΕΣ 10%)

$\Rightarrow$ Δεν μπορούμε να απορρίψουμε **έλλειψη συνολοκλήρωσης**

Εξήγηση προηγούμενων αποτελεσμάτων;

Σχέση επιπέδου: *φαινομενική συσχέτιση*

Σχέση διαφορών: χρειάζεται

Πως μπορούμε να έχουμε εκτιμητές με κανονικές κατανομές;

Νέα παλινδρόμηση με «περίπου» κλασσικές υποθέσεις

$$y_t = \beta_0 + \beta_1 x_t + \delta' z_t + u_t$$

$$y_t, x_t \text{ ειναι I(1)}$$

$$\mathrm{E}(u_t \Delta x_s) = 0$$

είδος εξωγένειας

Ικανοποιείται με εκτιμητή ‘lead-lags’, για τον οποίο:

z(t)=[Δx(t),Δx(t+1), Δx(t+2),…, Δx(t-1), Δx(t-2)…]’

<u>Παράδειγμα 18.6 (18.2, κλπ)</u>

Από προηγούμενη παλινδρόμηση

$\Delta Επιτόκια3M_t$ = .625 -.091$Επιτόκια3M_{t-1}$
(.261) (.037)

έχουμε απορρίψει έλλειψη ολοκλήρωσης.

Από την παλινδρόμηση:

Δ(6M_t- 3M_t) =α -.67 (6M_t- 3M_t)

έχουμε απορρίψει έλλειψη συνολοκλήρωσης (με βάση οικονομικό συλλογισμό – υπόθεση προσδοκιών)

$Επιτόκια6M_t$ = a +1.026$Επιτόκια3M_{t-1}$
(0.077)

(αλλά στατιστική t ίσως δεν ισχύει)

με lead-lag εκτιμητή (με δύο lead-lags) δίνει

$Επιτόκια6M_t$ = a +1.038$Επιτόκια3M_{t-1}$ +δ’z $_t$
(0.081)

$z_t = [\Delta 3M_t, \Delta 3M_{t+1}, \Delta 3M_{t+2}, \Delta 3M_{t-1}, \Delta 3M_{t-2}]'$

(1.038-1)/.0081 = 4.69 => *στατιστικά* διαφορετικό από το 1 το οποίο θα περιμέναμε από θεωρία

Είναι όμως *οικονομικά* διαφορετικό;

Εφικτά υποδείγματα

$\{y_t\}_{t=1}^{\infty}, \{x_t\}_{t=1}^{\infty}$

Χωρίς Συνολοκλήρωση

$\Delta y_t = \beta_0 + \beta_1 \Delta x_t \Rightarrow$

$\Rightarrow$ Σχέση με αλλαγές εξαρτημένης, όχι επιπέδου

Με Συνολοκλήρωση

$\Delta y_t = \beta_0 + \beta_1 \Delta x_t + \left(y_{t-1} - \beta^* x_{t-1}\right) \Rightarrow$

$\Rightarrow$ Σχέση με *αλλαγές και επίπεδο εξαρτημένης* (ΥΔΛ)

Υποδείγματα Διόρθωσης Λαθών

$$\Delta y_t = a_0 + a_1 \Delta y_{t-1} + \gamma_0 \Delta x_t + \gamma_1 \Delta x_{t-1} + \delta\left(y_{t-1} - \beta x_{t-1}\right) + u_t$$

Ερμηνεία «όρου διόρθωσης λαθών»:

Εκφράζει πώς κινείται το y_t για να επανέλθει στη μακροχρόνια ισορροπία.

Π.χ. $\delta < 0, y_{t-1} > \beta x_{t-1} \Rightarrow y_{t-1}$ μεγαλύτερο από ότι προβλέπει η σχέση του με το x_{t-1} και θα μικρύνει

Δυσκολία:

Η παράμετρος β είναι άγνωστη

Παράδειγμα 18.7 (18.6, 18.2, κλπ)

Από προηγούμενη παλινδρόμηση

$\Delta Επιτόκια3M_t$ = .625 -.091$Επιτόκια3M_{t-1}$

(.261) (.037)

έχουμε απορρίψει έλλειψη ολοκλήρωσης.

Αν δεχτούμε υπόθεση προσδοκιών, τότε το κατάλληλο ΥΔΛ είναι:

$\Delta 6M_t = a_0 + \gamma_0 \Delta 3M_{t-1} + \delta\,(6M_{t-1} - 3M_{t-2}) + u_t$

Εκτίμηση δίνει:

$\Delta 6M_t = .09 + 1.218\, \Delta 3M_{t-1} - .840\, (6M_{t-1} - 3M_{t-2})$

(.043) (.264) (.244)

n = 122 (τρίμηνα)

R^2=.790

δ^ δεν διαφέρει από 1 => όλη η διόρθωση γίνεται σε ένα τρίμηνο

Εκτίμηση β: «Διαδικασία 2 βημάτων *Engle-Granger*»

1. Εκτιμούμε $\hat{\beta}$ (πχ με κλασσική παλινδρόμηση)
2. Εκτιμούμε χρησιμοποιώντας $\left(y_{t-1} - \hat{\beta} x_{t-1}\right)$ ως μεταβλητή

$\Rightarrow$ Η ασυμπτωτική κατανομή των $\hat{a}, \hat{\gamma}$:

→ Δεν επηρεάζεται από το πρώτο βήμα

→ Είναι αποτελεσματική

Οικονομετρική ανάλυση δεδομένων πάνελ

Wooldridge, Κεφ. 13

Είδη δεδομένων:

- Ομαδοποιημένα διαστρωματικά
- Πάνελ

Υπόδειγμα για **ομαδοποιημένα διαστρωματικά στοιχεία** από πολλά έτη

Γονιμότητα

72,74,…,84

Εξαρτημένη: Αριθμός παιδιών

Πως αλλάζει διαχρονικά (ελέγχοντας για παρατηρούμενους παράγοντες);

Πίνακας 13.1

Προσδιοριστικοί Παράγοντες της Γυναικείας Γονιμότητας

Εξαρτημένη μεταβλητή: *kids*		
Ανεξάρτητες μεταβλητές	**Συντελεστές**	**Τυπικά σφάλματα**
educ	−.128	.018
age	.532	.138
age^2	−.0058	.0016
black	1.076	.174
east	.217	.133
northcen	.363	.121
west	.198	.167
farm	−.053	.147
othrural	−.163	.175
town	.084	.124
smcity	.212	.160
y74	.268	.173
y76	−.097	.179
y78	−.069	.182
y80	−.071	.183
y82	−.522	.172
y84	−.545	.175
σταθερά	−7.742	3.052

$n = 1.129$
$R^2 = .1295$
$\bar{R}^2 = .1162$

Μεταβολές στην απόδοση της μόρφωσης και χάσμα μεταξύ φύλων

Εξαρτημένη: Ωριαίοι μισθοί

Δεδομένα: 78 και 85

δ1 μετράει την μεταβολή στον αντίκτυπο μόρφωσης

$$\begin{aligned}
\log(\widehat{wage}) &= \underset{(.093)}{.459} + \underset{(.124)}{.118}\, y85 + \underset{(.0067)}{.0747}\, educ + \underset{(.0094)}{.0185}\, y85 \cdot educ \\
&+ \underset{(.0036)}{.0296}\, exper - \underset{(.00008)}{.00040}\, exper^2 + \underset{(.030)}{.202}\, union \\
&- \underset{(.037)}{.317}\, female + \underset{(.051)}{.085}\, y85 \cdot female \\
n &= 1{,}084,\ R^2 = .426,\ \bar{R}^2 = .422.
\end{aligned}$$

Αποτεφρωτήρας απορριμάτων και τιμές κατοικιών

Εξαρτημένη: Πραγματικές τιμές

Δεδομένα: 78 (πριν φήμες), 81 (αρχή κατασκευής)

$$\widehat{rprice} = \underset{(3{,}093.0)}{101{,}307.5} - \underset{(5{,}827.71)}{30{,}688.27}\ nearinc$$
$$n = 142,\ R^2 = .165.$$

Δεδομένα: μόνο 78

$$\widehat{rprice} = \underset{(2{,}653.79)}{82{,}517.23} - \underset{(5{,}827.71)}{18{,}824.37}\ nearinc$$
$$n = 179,\ R^2 = .082.$$

Εξαρτημένη Τιμή: *rprice*

Ανεξάρτητη μεταβλητή	(1)	(2)	(3)
σταθερά	82,517.23 (2,726.91)	89,116.54 (2,406.05)	13,807.67 (11,166.59)
y81	18,790.29 (4,050.07)	21,321.04 (3,443.63)	13,928.48 (2,798.75)
nearinc	−18,824.37 (4,875.32)	9,397.94 (4,812.22)	3,780.34 (4,453.42)
y81 · nearinc	−11,863.90 (7,456.65)	−21,920.27 (6,359.75)	−14,177.93 (4,987.27)
Άλλες μεταβλητές	Καμία	*age*, *age*^{2}	Πλήρες σύνολο
Παρατηρήσεις *R*-τετράγωνο	321 .174	321 .414	321 .660

Πλήρες σύνολο συμπεριλαμβάνει απόσταση απο αυτοκινητόδρομο, εμβαδό κατοικίας, οικοπέδου, # δωματίων, # λουτρών

ΦΥΣΙΚΟ ΠΕΙΡΑΜΑ

Αντίκτυπος ορίου μέγιστης αποζημίωσης σε διάρκεια

$$\log(\widehat{durat}) = \underset{(0.031)}{1.126} + \underset{(.0447)}{.0077}\ afchnge + \underset{(.047)}{.256}\ highearn$$
$$+ \underset{(.069)}{.191}\ afchnge \cdot highearn$$
$$n = 5{,}626,\ R^2 = .021.$$

Φυσικό πείραμα – με μονάδα ελέγχου Α και αντιμετώπισης Β

$$y = \beta_0 + \delta_0 d2 + \beta_1 dB + \delta_1 d2 \cdot dB$$

Ψευδομεταβλητές αντιμετώπισης:

d2 - περιόδου, dB - μονάδας

$$\hat{\delta}_1 = (\bar{y}_{2,\mathrm{B}} - \bar{y}_{2,\mathrm{A}}) - (\bar{y}_{1,\mathrm{B}} - \bar{y}_{1,\mathrm{A}}).$$

Υποδείγματα πάνελ

$$\widehat{crmrte} = \underset{(20.76)}{128.38} - \underset{(3.42)}{4.16}\ unem$$
$$n = 46,\ R^2 = .033.$$

Απλή προσέγγιση

Εγκληματικότητα και ανεργία στις πόλεις

$$y_{it} = \beta_0 + \delta_0 d2_t + \beta_1 x_{it} + a_i + u_{it}, \quad t = 1, 2.$$

$$crmrte_{it} = \beta_0 + \delta_0 d87_t + \beta_1 unem_{it} + a_i + u_{it},$$

i, πόλεις (46)
t, έτη (2 – 1982 και 1987)
a_*i* απαρατήρητη / σταθερή επίδραση πόλης
δ_0 ψευδομεταβλητή έτους

$$\widehat{crmrte} = \underset{(12.74)}{93.42} + \underset{(7.98)}{7.94}\ d87 + \underset{(1.188)}{.427}\ unem$$
$$n = 92,\ R^2 = .012.$$

Προσέγγιση χρησιμοποιόντας παρατηρήσιμες μεταβλητές

και ομαδοποιημένα στοιχεία

Εκτίμηση παρατηρήσιμων παραμέτρων από εξίσωση πρώτων διαφορών:

$$y_{i2} = (\beta_0 + \delta_0) + \beta_1 x_{i2} + a_i + u_{i2} \quad (t = 2)$$

$$y_{i1} = \beta_0 + \beta_1 x_{i1} + a_i + u_{i1} \quad (t = 1).$$

Αν αφαιρέσουμε τη δεύτερη εξίσωση από την πρώτη, παίρνουμε:

$$(y_{i2} - y_{i1}) = \delta_0 + \beta_1(x_{i2} - x_{i1}) + (u_{i2} - u_{i1}),$$

ή

$$\Delta y_i = \delta_0 + \beta_1 \Delta x_i + \Delta u_i,$$

Υποθέσεις:

E(Δu|Δx)=0 [συνεπάγεται απο ΑΕ στο αρχικό υπόδειγμα]

Δx= δεν είναι σταθερό

$$\Delta \widehat{crmrte} = 15.40 + 2.22\ \Delta unem$$
$$(4.70) \quad (.88)$$
$$n = 46,\ R^2 = .127,$$

Εναλλακτική Ερμηνεία:

Αντίκτυπος μεταβολών

Απαιτεί:

δυσεύρετα δεδομένα

Σημαντική μεταβλητικότητα στο Δx

$$\Delta \log(wage_i) = \delta_0 + \beta_1 \Delta educ_i + \Delta u_i,$$

$$\Delta \widehat{slpnap} = -92.63 - .227\ \Delta totwrk - .024\ \Delta educ$$
$$(45.87)\ (.036) \qquad (48.759)$$
$$+ 104.21\ \Delta marr + 94.67\ \Delta yngkid + 87.58\ \Delta gdhlth$$
$$(92.86) \qquad (87.65) \qquad (76.60)$$
$$n = 239,\ R^2 = .150.$$

Ύπνος και εργασία:

Λεπτά ύπνου/εβδομάδα

Διαφορισμός αφαιρεί απαρατήρητα ατομικά χαρακτηρηστικά όπως επίπεδα ενέργειας

$$\Delta \log(\widehat{crime}) = .086 - .0040\ \Delta clrprc_{-1} - .0132\ \Delta clrprc_{-2}$$
$$(.064)\ (.0047) \qquad (.0052)$$
$$n = 53,\ R^2 = .193,\ \bar{R}^2 = .161.$$

Εγκληματικότητα και ποσοστά εκκαθάρισης

1972, 78

=> Εγκληματικότητα μπορεί να μειωθεί

Ανάλυση πολιτικής (αντίστοιχα με φυσικό πείραμα)

$$scrap_{it} = \beta_0 + \delta_0 y88_t + \beta_1 grant_{it} + a_i + u_{it}, \quad t = 1, 2,$$

Ποσοστό ελλατωματικών προϊόντων

1987,88

Ψευδομεταβλητή επιδότησης επαγγελματικής εκπαίδευσης

$$\Delta scrap_i = \delta_0 + \beta_1 \Delta grant_i + \Delta u_i.$$

$$\Delta \widehat{scrap} = -.564 - .739\ \Delta grant$$
$$(.405)\ (.683)$$
$$n = 54,\ R^2 = .022.$$

$$\Delta \log(\widehat{scrap}) = -.057 - .317\ \Delta grant$$
$$(.097)\ (.164)$$
$$n = 54,\ R^2 = .067.$$

Γενικότερα:

$$y_{it} = \beta_0 + \delta_0 d2_t + \beta_1 prog_{it} + a_i + u_{it}.$$

Ερμηνεία β_1:

- Η μεταβολή στη μέση τιμή του λόγο συμμετοχής στο πρόγραμμα
- Αν η συμμετοχή συμβαίνει μόνο στην 2η περίοδο => εκτιμητής διαφορών

$$\Delta \widehat{dthrte} = -.497 - .420\ \Delta open - .151\ \Delta admn$$
$$(.052)\ (.206)\qquad (.117)$$
$$n = 51,\ R^2 = .119.$$

Νομοθεσία και θανατηφόρα ατυχήματα

Θάνατοι / 100εκ μίλια οδήγησης

Ανοιχτές συσκευασίες

Ευκολότερες καταδίκες

50 πολιτείες + D.C., 1985; 1990

$$\log(uclms_{it}) = \theta_t + \beta_1 ez_{it} + a_i + u_{it},$$

Βιομηχανικές ζώνες και επιδόματα ανεργίας

22 πόλεις της Indiana 1980-8

Αριθμός αιτήσεων επιδομάτων

Ψευδομεταβλητή βιομηχανικής ζώνης

$$\Delta \log(uclms_{it}) = a_0 + a_1 d82_t + ... + a_7 d88_t + \beta_1 \Delta ez_{it} + \Delta u_{it}.$$

n=22*8=76

β1=.182 (.078)

Υποθέσεις:

- Αυστηρή εξωγένεια [στην αρχική σχέση]: Cov(x_itj,u_is)=0, all t,s,j
- Έλλειψη αυτοσυσχέτισης [στην σχέση διαφορών]: Cov(Δu_it,Δu_it-1)=0
- Ομοσκεδαστικότητα [στην σχέση διαφορών]

Εγκληματικότητα στην Βόρεια Καρολίνα

$$\begin{aligned}\Delta\log(\widehat{crmrte}) = {} & \underset{\substack{(.017)\\ {[.014]}}}{.008} - \underset{\substack{(.024)\\ {[.022]}}}{.100}\, d83 - \underset{\substack{(.024)\\ {[.020]}}}{.048}\, d84 - \underset{\substack{(.023)\\ {[.025]}}}{.005}\, d85 \\ & + \underset{\substack{(.024)\\ {[.021]}}}{.028}\, d86 + \underset{\substack{(.024)\\ {[.024]}}}{.041}\, d87 - \underset{\substack{(.030)\\ {[.056]}}}{.327}\, \Delta\log(prbarr) \\ & - \underset{\substack{(.018)\\ {[.039]}}}{.238}\, \Delta\log(prbconv) - \underset{\substack{(.026)\\ {[.045]}}}{.165}\, \Delta\log(prbpris) \\ & - \underset{\substack{(.022)\\ {[.025]}}}{.022}\, \Delta\log(avgsen) + \underset{\substack{(.027)\\ {[.101]}}}{.398}\, \Delta\log(polpc) \\ & n = 540,\ R^2 = .433,\ \bar{R}^2 = .422.\end{aligned}$$

90 κομητείες, 1981,82,…,87

Εγκλήματα/κάτοικος

Πιθανότητα σύλληψης, καταδίκης, φυλάκισης

Έτη φυλάκισης

Αστυνόμοι/κάτοικοι

Προηγμένες μέθοδοι Δεδομένων Πάνελ

Wooldridge, Κεφ 14

Εναλλακτικές μέθοδοι (από τον διαφορισμό) για εξάλειψη σταθερής επίδρασης

- Μετασχηματισμός σταθερών επιδράσεων
- Μετασχηματισμός τυχαίων επιδράσεων

Μετασχηματισμός σταθερών επιδράσεων

$$y_{it} = \beta_1 x_{it} + a_i + u_{it},\ t = 1, 2, ..., T.$$

Έχει διαχρονικό μέσο:

$$\bar{y}_i = \beta_1 \bar{x}_i + a_i + \bar{u}_i,$$

Οπότε αφαιρόντας:

$$\ddot{y}_{it} = \beta_1 \ddot{x}_{it} + \ddot{u}_{it},\ t = 1, 2, ..., T,$$

Και γενικότερα:

$$\ddot{y}_{it} = \beta_1 \ddot{x}_{it1} + \beta_2 \ddot{x}_{it2} + ... + \beta_k \ddot{x}_{itk} + \ddot{u}_{it},\ t = 1, 2, ..., T,$$

Εκτίμηση σταθερών επιδράσεων / “εντός”

Υποθέσεις

- E(u_it|x_jsk) για όλα τα j,s,k
- Επιτρέπεται Corr(a_i, x_isk)~=0
- *x_jsk* δεν μπορεί να είναι χρονικά σταθερό
- Ομοσκεδαστικότητα
- Έλλειψη αυτοσυσχέτισης

Ιδιότητες

- Συνηθισμένες OLS
- Βαθμοί ελευθερίας: *N(T-1)-k λόγω αφαίρεσης διαχρονικού μέσου*
- *R^2* μετράει εξηγησιμότητα διαχρονικών διακυμάνσεων εντός στρώματος

Επίδραση επαγγελματικής εκπαίδευσης σε ελαττώματα σε προϊόντα

Εξαρτημένη μεταβλητή: log(*scrap*)	
Ανεξάρτητες μεταβλητές	
d88	−.080 (.109)
d89	−.247 (.133)
grant	−.252 (.151)
$grant_{-1}$	−.422 (.210)
Παρατηρήσεις Βαθμοί ελευθερίας *R*-τετράγωνο	162 104 .201

54 εταιρείες

1987: 0 επιδοτήσεις

1988: 19 επιδοτήσεις

1999: 10 επιδοτήσεις

Εκτιμάται εντός

Παρουσιάζεται κλασσικά

Αν συμπεριλάβουμε τις μεταβλητές: log(sales), log(employ)

3 εταιρείες δεν έχουν τέτοια δεδομένα

5 παρατηρήσεις χάνονται γιατί λείπουν τα δεδομένα για κάποιες εταιρείες σε κάποια χρόνια

⇒ n=148, "μη ισορροπημένο πάνελ"

Επίδραση μεγαλώνει:

β_grant= -.297, t=-1,89

β_grant-1 = -.536 , t=-2,39

Εταιρείες με 1 παρατήρηση δεν έχουν επίδραση σε αποτελέσματα

Πιθανότητα να μην υπάρχει η παρατήρηση μπορεί να συσχετίζεται με επίδραση *a_i*

Απόδοση στην μόρφωση διαχρονικά

545 άνδρες

1980-7

log(educ_it)= β_0educ^2+β_1dmarried+β_2dsynd+β_3d81+…+β_9d87+ β_10*d81*educ+…+ β_16*d87*educ

β^_16=0.03 (t-stat = 2.48) => Απόδοση 3% μεγαλύτερη από το 1980

Απόδοση 1980 είναι μέρος του a_i

Ιδιαίτερος τρόπος χρήσης στατικής πληροφορίας για το στρώμα

Εκτίμηση σταθερών επιδράσεων [παραληφθείσες μεταβλητές]:

$$\hat{a}_i = \bar{y}_i - \hat{\beta}_1 \bar{x}_{i1} - \ldots - \hat{\beta}_k \bar{x}_{ik},\ i = 1, \ldots, N,$$

Αμερόληπτη, συνεπής *T*, όχι *N*

Ίδια αποτελέσματα με παλινδρόμηση ψευδομεταβλητών
- *d_i* για κάθε *i*
- Δύσκολο αν N μεγάλο, ανακριβές αν T μικρό
- R^2 τώρα συμπεριλαμβάνει ερμηνευτικότητα ψευδομεταβλητών

Σύγκριση εκτιμήσεων από μετασχηματισμούς σταθερών επιδράσεων και διαφορισμού

T=1: Δ Μη εφαρμόσιμος, ΣΕ όλες μεταβλητές 0

T=2: Ίδια αποτελέσματα

T>=3: Αμερόληπτοι και συνεπής (για N)

Αποτελεσματικότητα:

Εξαρτάται από την αυτοσυσέτιση του u_it [μη παρατηρήσιμο] σε σχέση με του Δu_it [παρατηρήσιμο]
- Καλό να χρησιμοποιούνται και οι δύο μετασχηματισμοί

Αν T>>0 χάρη σε ασυμπτωτική θεωρία χρονοσειρών:
- Δ προτιμότερος αν κάποιες μεταβλητές είναι I(1)
- ΣΕ προτιμότερες αν δεν ισχύει η ΑΕ

Μετασχηματισμός τυχαίων επιδράσεων

Αν $\text{Cov}(x_{itj}, a_i) = 0,\ t = 1, 2, \ldots, T.\ j = 1, 2, \ldots, k.$

δεν απομακρύνουμε απαρατήρητη σταθερή επίδραση στο υπόδειγμα:

$$y_{it} = \beta_0 + \beta_1 x_{it1} + \ldots + \beta_k x_{itk} + a_i + u_{it},$$

ή

$$y_{it} = \beta_0 + \beta_1 x_{it1} + \ldots + \beta_k x_{itk} + v_{it}. \quad v_{it} = a_i + u_{it}$$

Γιατί χάνουμε βαθμούς ελευθερίας (αποτελεσματικότητα)

Τότε όμως έχουμε αυτοσυσχέτιση:

$$\text{Corr}(v_{it}, v_{is}) = \sigma_a^2/(\sigma_a^2 + \sigma_u^2),\ t \neq s,$$

Αφαιρόντας την αυτοσυσχέτιση με διαδικασία αντίστοιχη με ημιδιαφορισμό:

$$y_{it} - \lambda \bar{y}_i = \beta_0(1 - \lambda) + \beta_1(x_{it1} - \lambda \bar{x}_{i1}) + \ldots + \beta_k(x_{itk} - \lambda \bar{x}_{ik}) + (v_{it} - \lambda \bar{v}_i),$$

$$\lambda = 1 - [\sigma_u^2/(\sigma_u^2 + T\sigma_a^2)]^{1/2},$$

Χρησιμοποιείται εκτίμηση για το λ

Επιτρέπονται μεταβλητές σταθερές στον χρόνο [αρκεί να είναι **ασυσχέτιστες** με ΣΕ]

Παρατηρόντας ότι:

$$v_{it} - \lambda \bar{v}_i = (1-\lambda)a_i + u_{it} - \lambda \bar{u}_i$$

Αν υπάρχει συσχέτιση μεταξύ a,x
Για λ->1 ασήμαντη [ΣΕ]
Για λ->0 σημαντική [ομαδοποιημένα στοιχεία]

Σύγκριση αποτελεσμάτων από μετασχηματισμούς ΤΕ και ΣΕ δίνει έλεγχο για έλλειψη συσχέτισης

Απόδοση στην μόρφωση διαχρονικά (ξανά)

545 άνδρες ΣΥΓΚΡΙΣΗ ΕΚΤΙΜΗΣΕΩΝ
1980-7
λ^=0.643

Εξαρτημένη μεταβλητή: log(wage)			
Ανεξάρτητες μεταβλητές	Ομαδοποιημένα ελάχιστα τετράγωνα	Τυχαίες επιδράσεις	Σταθερές επιδράσεις
educ	.091 (.005)	.092 (.011)	———
black	−.139 (.024)	−.139 (.048)	———
hispan	.016 (.021)	.022 (.043)	———
exper	.067 (.014)	.106 (.015)	———
exper^2	−.0024 (.0008)	−.0047 (.0007)	−.0052 (.0007)
married	.108 (.016)	.064 (.017)	.047 (.018)
union	.182 (.017)	.106 (.018)	.080 (.019)

Συμπεριλαμβάνονται και ψευδομεταβλητές έτους

Άλλες χρήσεις μετασχηματισμών πάνελ

Όταν σε συστάδες [οχι βασισμένες στον χρόνο] του δείγματος αναμένονται απαρατήρητες επιδράσεις:

π.χ. Εισόδημα κόρης σε οικογένειες με τουλάχιστον δύο κόρες

$$\log(incneeds_{fs}) = \beta_0 + \delta_0 sister2_s + \beta_1 teenbrth_{fs} + \beta_2 age_{fs} + \text{άλλοι παράγοντες} + a_f + u_{fs},$$

n=129, 1982, β^1=-.33, πολύ σ.σ.

$$\Delta\log(incneeds) = \delta_0 + \beta_1 \Delta teenbrth + \beta_2 \Delta age + \ldots + \Delta u.$$

n=129, 1982, β^1=-.08 , όχι σ.σ.

Π.χ.2:
149 ομοζυγωτικά δίδυμα

age, gender, race είναι ίδια οπότε παλινδρόμηση:

Δlog(earn)=b_0 Δeduc
β_0= 0.092, t=3.83
Μπορούν να χρησιμοποιηθούν και μετασχητατισμοί Τ.Ε.

Προηγμένες μέθοδοι Δεδομένων Πάνελ

Wooldridge, Κεφ 14

Εναλλακτικές μέθοδοι (από τον διαφορισμό) για εξάλειψη σταθερής επίδρασης

- Μετασχηματισμός σταθερών επιδράσεων
- Μετασχηματισμός τυχαίων επιδράσεων

Μετασχηματισμός σταθερών επιδράσεων

$$y_{it} = \beta_1 x_{it} + a_i + u_{it},\ t = 1, 2, ..., T.$$

Έχει διαχρονικό μέσο:

$$\bar{y}_i = \beta_1 \bar{x}_i + a_i + \bar{u}_i.$$

Οπότε αφαιρόντας:

$$\ddot{y}_{it} = \beta_1 \ddot{x}_{it} + \ddot{u}_{it},\ t = 1, 2, ..., T,$$

Και γενικότερα:

$$\ddot{y}_{it} = \beta_1 \ddot{x}_{it1} + \beta_2 \ddot{x}_{it2} + ... + \beta_k \ddot{x}_{itk} + \ddot{u}_{it},\ t = 1, 2, ..., T,$$

Εκτίμηση σταθερών επιδράσεων / “εντός”

Υποθέσεις

- E(*u_it*|*x_jsk*) για όλα τα j,s,k
- Επιτρέπεται Corr(*a_i, x_isk*)~=0
- *x_jsk* δεν μπορεί να είναι χρονικά σταθερό
- Ομοσκεδαστικότητα
- Έλλειψη αυτοσυσχέτισης

Ιδιότητες

- Συνηθισμένες OLS
- Βαθμοί ελευθερίας: *N(T-1)-k λόγω αφαίρεσης διαχρονικού μέσου*
- *R^2* μετράει εξηγησιμότητα διαχρονικών διακυμάνσεων εντός στρώματος

Επίδραση επαγγελματικής εκπαίδευσης σε ελαττώματα σε προϊόντα

Εξαρτημένη μεταβλητή: log(*scrap*)	
Ανεξάρτητες μεταβλητές	
d88	−.080 (.109)
d89	−.247 (.133)
grant	−.252 (.151)
$grant_{-1}$	−.422 (.210)
Παρατηρήσεις Βαθμοί ελευθερίας *R*-τετράγωνο	162 104 .201

54 εταιρείες

1987: 0 επιδοτήσεις

1988: 19 επιδοτήσεις

1999: 10 επιδοτήσεις

Εκτιμάται εντός

Παρουσιάζεται κλασσικά

Αν συμπεριλάβουμε τις μεταβλητές: log(sales), log(employ)

3 εταιρείες δεν έχουν τέτοια δεδομένα

5 παρατηρήσεις χάνονται γιατί λείπουν τα δεδομένα για κάποιες εταιρείες σε κάποια χρόνια

$\Rightarrow$ n=148, “μη ισορροπημένο πάνελ”

Επίδραση μεγαλώνει:

β_grant= -.297, t=-1,89

β_grant-1 = -.536 , t=-2,39

Εταιρείες με 1 παρατήρηση δεν έχουν επίδραση σε αποτελέσματα

Πιθανότητα να μην υπάρχει η παρατήρηση μπορεί να συσχετίζεται με επίδραση *a_i*

Απόδοση στην μόρφωση διαχρονικά

545 άνδρες

1980-7

log(educ_it)= β_0educ^2+β_1dmarried+β_2dsynd+β_3d81+…+β_9d87+ β_10*d81*educ+…+ β_16*d87*educ

β^_16=0.03 (t-stat = 2.48) => Απόδοση 3% μεγαλύτερη από το 1980

Απόδοση 1980 είναι μέρος του a_i

Ιδιαίτερος τρόπος χρήσης στατικής πληροφορίας για το στρώμα

Εκτίμηση σταθερών επιδράσεων [παραληφθείσες μεταβλητές]:

$$\hat{a}_i = \bar{y}_i - \hat{\beta}_1 \bar{x}_{i1} - \ldots - \hat{\beta}_k \bar{x}_{ik},\ i = 1, \ldots, N,$$

Αμερόληπτη, συνεπής *T*, όχι *N*

Ίδια αποτελέσματα με παλινδρόμηση ψευδομεταβλητών

- *d_i* για κάθε *i*
- Δύσκολο αν N μεγάλο, ανακριβές αν T μικρό
- R^2 τώρα συμπεριλαμβάνει ερμηνευτικότητα ψευδομεταβλητών

Σύγκριση εκτιμήσεων από μετασχηματισμούς σταθερών επιδράσεων και διαφορισμού

T=1: Δ Μη εφαρμόσιμος, ΣΕ όλες μεταβλητές 0

T=2: Ίδια αποτελέσματα

T>=3: Αμερόληπτοι και συνεπής (για N)
Αποτελεσματικότητα:

Εξαρτάται από την αυτοσυσέτιση του u_it [μη παρατηρήσιμο] σε σχέση με του Δu_it [παρατηρήσιμο]

- Καλό να χρησιμοποιούνται και οι δύο μετασχηματισμοί

Αν T>>0 χάρη σε ασυμπτωτική θεωρία χρονοσειρών:

- Δ προτιμότερος αν κάποιες μεταβλητές είναι I(1)

- ΣΕ προτιμότερες αν δεν ισχύει η ΑΕ

Μετασχηματισμός τυχαίων επιδράσεων

Αν $\text{Cov}(x_{itj}, a_i) = 0, \; t = 1, 2, ..., T. \; j = 1, 2, ..., k.$

δεν απομακρύνουμε απαρατήρητη σταθερή επίδραση στο υπόδειγμα:

$$y_{it} = \beta_0 + \beta_1 x_{it1} + ... + \beta_k x_{itk} + a_i + u_{it},$$

ή

$$y_{it} = \beta_0 + \beta_1 x_{it1} + ... + \beta_k x_{itk} + v_{it}. \qquad v_{it} = a_i + u_{it}$$

Γιατί χάνουμε βαθμούς ελευθερίας (αποτελεσματικότητα)

Τότε όμως έχουμε αυτοσυσχέτιση:

$$\text{Corr}(v_{it}, v_{is}) = \sigma_a^2/(\sigma_a^2 + \sigma_u^2), \; t \neq s,$$

Αφαιρόντας την αυτοσυσχέτιση με διαδικασία αντίστοιχη με ημιδιαφορισμό:

$$y_{it} - \lambda\bar{y}_i = \beta_0(1-\lambda) + \beta_1(x_{it1} - \lambda\bar{x}_{i1}) + ... + \beta_k(x_{itk} - \lambda\bar{x}_{ik}) + (v_{it} - \lambda\bar{v}_i),$$

$$\lambda = 1 - [\sigma_u^2/(\sigma_u^2 + T\sigma_a^2)]^{1/2},$$

Χρησιμοποιείται εκτίμηση για το λ

Επιτρέπονται μεταβλητές σταθερές στον χρόνο [αρκεί να είναι **ασυσχέτιστες** με ΣΕ]

Παρατηρόντας ότι:

$$v_{it} - \lambda\bar{v}_i = (1-\lambda)a_i + u_{it} - \lambda\bar{u}_i$$

Αν υπάρχει συσχέτιση μεταξύ a,x

Για λ->1 ασήμαντη [ΣΕ]

Για λ->0 σημαντική [ομαδοποιημένα στοιχεία]

Σύγκριση αποτελεσμάτων από μετασχηματισμούς ΤΕ και ΣΕ δίνει έλεγχο για έλλειψη συσχέτισης

Απόδοση στην μόρφωση διαχρονικά (ξανά)

545 άνδρες ΣΥΓΚΡΙΣΗ ΕΚΤΙΜΗΣΕΩΝ

1980-7

λ^=0.643

Εξαρτημένη μεταβλητή: log(*wage*)			
Ανεξάρτητες μεταβλητές	Ομαδοποιημένα ελάχιστα τετράγωνα	Τυχαίες επιδράσεις	Σταθερές επιδράσεις
educ	.091 (.005)	.092 (.011)	———
black	–.139 (.024)	–.139 (.048)	———
hispan	.016 (.021)	.022 (.043)	———
exper	.067 (.014)	.106 (.015)	———
$exper^2$	–.0024 (.0008)	–.0047 (.0007)	–.0052 (.0007)
married	.108 (.016)	.064 (.017)	.047 (.018)
union	.182 (.017)	.106 (.018)	.080 (.019)

Συμπεριλαμβάνονται και ψευδομεταβλητές έτους

΄Αλλες χρήσεις μετασχηματισμών πάνελ

Όταν σε συστάδες [οχι βασισμένες στον χρόνο] του δείγματος αναμένονται απαρατήρητες επιδράσεις:

π.χ. Εισόδημα κόρης σε οικογένειες με τουλάχιστον δύο κόρες

$$\log(incneeds_{fs}) = \beta_0 + \delta_0 sister2_s + \beta_1 teenbrth_{fs} + \beta_2 age_{fs} + \text{άλλοι παράγοντες} + a_f + u_{fs},$$

n=129, 1982, β^1=-.33, πολύ σ.σ.

$$\Delta\log(incneeds) = \delta_0 + \beta_1 \Delta teenbrth + \beta_2 \Delta age + \ldots + \Delta u.$$

n=129, 1982, β^1=-.08 , όχι σ.σ.

Π.χ.2:
149 ομοζυγωτικά δίδυμα

age, gender, race είναι ίδια οπότε παλινδρόμηση:

Δlog(earn)=b_0 Δeduc
β_0= 0.092, t=3.83
Μπορούν να χρησιμοποιηθούν και μετασχητατισμοί Τ.Ε.

Μοντέλα περιορισμένων εξαρτημένων μεταβλητών και διορθώσεις επιλογής δείγματος

Wooldridge, κεφ17

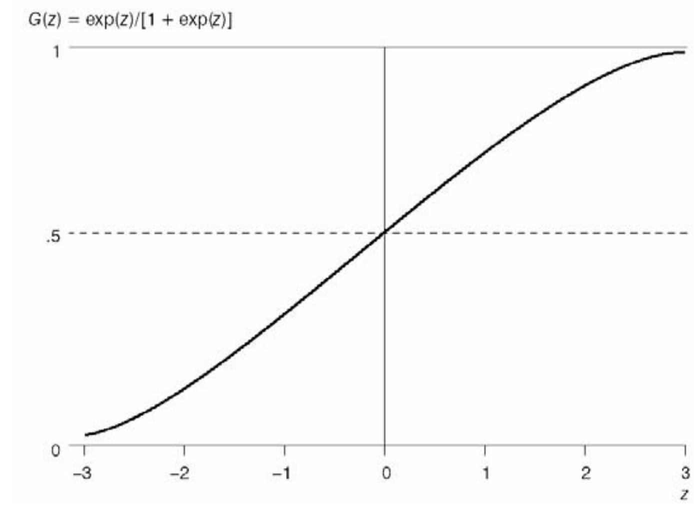
Υποδείγματα logit / probit

$$P(y=1|x) = P(y=1|x_1, x_2, ..., x_k),$$

$$P(y=1|x) = G(\beta_0 + ... + \beta_k x_k) = G(\beta_0 + x\beta),$$

Περιμπτώσεις:

$G(z) = z$	Γραμμικό υπόδειγμα πιθανότητας
$G(z) = \exp(z)/[1+\exp(z)] = \Lambda(z),$	logit
$G(z) = \Phi(z) \equiv \int_{-\infty}^{z} \varphi(v)dv,$ $\varphi(z) = (2\pi)^{-1/2}\exp(-z^2/2).$	probit



Ερμηνεία logit / probit

Υπόδειγμα μη παρατηρούμενης μεταβλητής

$$y^* = \beta_0 + x\beta + e,\ y = 1[y^* > 0],$$

E(e|x)=0

e~ L, N(0,1)

=>

$$P(y=1|x) = P(y^*>0|x) = P[e > -(\beta_0 + x\beta)|x]$$
$$= 1 - G[-(\beta_0 + x\beta)] = G(\beta_0 + x\beta),$$

Επίδραση ερμηνευτικών σε πιθανότητες

$$\frac{\partial p(x)}{\partial x_j} = g(\beta_0 + x\beta)\beta_j,\ \text{όπου}\ g(z) \equiv \frac{dG}{dz}(z)$$

g>0=> *β* καθορίζει πρόσημο επίδρασης
Σχετική σημασία επιδράσεων *β_k/β_i* ανεξάρτητη του *g*
Επίδραση μεγιστοποιείται για x ώστε *β_0* + x*β* =0
Probit *g*(0)=0.4
Logit *g*(0)= 0.25

Για διακρίτες ερμηνευτικές, επίδραση μεταβολής c -> c+1 είναι:

$$G[\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k(c_k + 1)]$$
$$- G(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k c_k).$$

Για συναρτήσεις ερμηνευτικών:

$$P(y = 1|z) = G(\beta_0 + \beta_1 z_1 + \beta_2 z_1^2 + \beta_3 \log(z_2) + \beta_4 z_3)$$

π.χ. $\partial P(y = 1|z)/\partial z_2 = g(\beta_0 + x\beta)(\beta_3/z_2)$

=> Αύξηση 1% z_2=> P(y=1) αλλάζει κατά % $g(\beta_0 + x\beta)(\beta_3/100)$

Εκτίμηση μέγιστης πιθανοφάνειας

$$f(y|x_i;\beta) = [G(x_i\beta)]^y[1 - G(x_i\beta)]^{1-y},\ y = 0, 1,$$

$$\ell_i(\beta) = y_i \log[G(x_i\beta)] + (1 - y_i)\log[1 - G(x_i\beta)]$$

$$: \mathcal{L}(\beta) = \sum_{i=1}^{n} \ell_i(\beta).$$

Μεγιστοποίηση υπολογιστικά περίπλοκη, όμως:
Συνεπής, ασυμπτωτικά κανονικές, ασ. αποτελεσματικές

Έλεγχος πολλαπλών υποθέσεων με στατιστική λόγου πιθανοφάνειας:

$$LR = 2(\mathcal{L}_{ur} - \mathcal{L}_r),$$

$$\mathcal{L}_{ur} \geq \mathcal{L}_r$$

$$LR \overset{a}{\sim} \chi_q^2$$

Μέτρηση ποιότητας προσαρμογής:

% σωστής πρόβλεψης για κάθε y

McFadden Pseudo-R^2 : $1 - \mathcal{L}_{ur}/\mathcal{L}_r$
Όπου περιορίζουμε β=0

Corr(y,y^)^2

Εκτίμηση επιδράσεων:

$$\Delta\hat{P}(y=1|x) \approx [g(\hat{\beta}_0 + x\hat{\beta})\hat{\beta}_j]\Delta x_j.$$

Σε δίαφορες τιμές του *x*:
Μέση τιμή
Τεταρτημόρια
Τιμές ψευδομεταβλητής

Παράδειγμα: Συμμετοχή έγγαμων γυναικών στο εργατικό δυναμικό

N=753
1975
Σy=428
Kids=#
nwifeinc – εισόδημα συζ $k

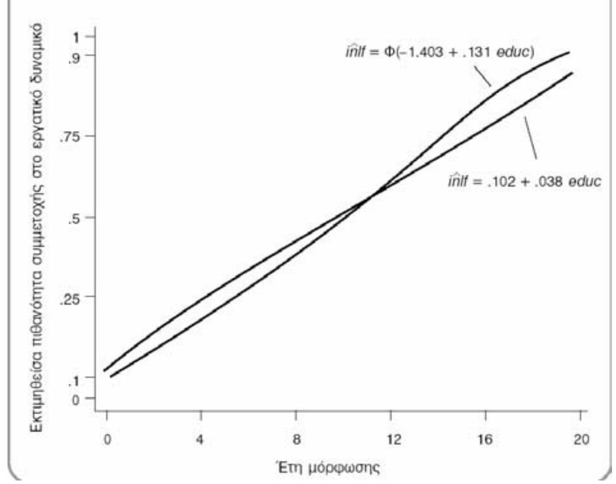
Εξαρτημένη μεταβλητή: *inlf*			
Ανεξάρτητες μεταβλητές	Γραμμικό μοντέλο πιθανότητας (μέθοδος ελαχίστων τετραγώνων)	Logit (Εκτίμηση μέγιστης πιθανοφάνειας)	Probit (Εκτίμηση μέγιστης πιθανοφάνειας)
nwifeinc	–.0034 (.0015)	–.021 (.008)	–.012 (.005)
educ	.038 (.007)	.221 (0.43)	.131 (.025)
exper	.039 (.006)	.206 (.032)	.123 (.019)
exper^2	–.00060 (.00018)	–.0032 (.0010)	–.0019 (.0006)
age	–.016 (.002)	–.088 (.015)	–.053 (.008)
kidslt6	–.262 (.032)	–1.443 (.204)	–.868 (.119)
kidsge6	.013 (.013)	.060 (.075)	.036 (.043)
σταθερά	.586 (.151)	.425 (.860)	.270 (.509)
Ποσοστό σωστής πρόβλεψης	73.4	73.6	73.4
Τιμή λογαριθμικής πιθανοφάνειας	———	–401.77	–401.309
Ψευδές *R*-τετράγωνο	.264	.220	.221

Σημαντικότερη διαφορά ΓΜΠ, logit, probit

Επίδραση στο ΓΜΠ είναι ανεξάρτητη του x
Σε logit, probit μείωνεται

π.χ.
Για μέσες τιμές του *x*:
Nwifeinc=20.13
Educ=12.3
Exper=10.6
Age=42.5
Επίδραση πρώτου μικρού παιδιόυ: -.334 (logit)
Επίδραση δεύτερου μικρού παιδιόυ: -.256 (logit)
Probit παρόμοιο
ΓΜΠ ενδιάμεσα (εκφράζοντας μέση επίδραση)

Εκτιμηθείσες πιθανότητες αντίδρασης ως προς τη μόρφωση για γραμμικά μοντέλα πιθανότητας και μοντέλα probit



Υπόδειγμα tobit

Όπως και στο log/prob-it:

$$y^* = \beta_0 + x\beta + u, \; u|x \sim \text{Normal}(0, \sigma^2) \qquad y = \max(0, y^*)$$

π.χ.

% εισοδήματος σε δαπάνη για οινοπνευματώδη

Ώρες εργασίας

Παρουσίες στα μαθήματα

Δωρεές

Πλεονεκτήματα σε σχέση με ΓΜΠ:

- Προβλέψεις πάντα θετικές
- Επίδραση στην εξαρτημένη εξαρτάται από επίπεδο ανεξάρτητης
- Πιθανή έλλειψη ομοσκεδαστικότητας
- Σίγουρη έλλειψη κανονικότητας

$$\mathrm{P}(y = 0|x) = \mathrm{P}(y^* < 0|x) = \mathrm{P}(u < -x\beta|x)$$
$$= \mathrm{P}(u/\sigma < -x\beta/\sigma|x) = \Phi(-x\beta/\sigma) = 1 - \Phi(-x\beta/\sigma) = 1 - \Phi(x\beta/\sigma)$$

Άρα για τυχαίο δείγμα

$$(2\pi\sigma^2)^{-1/2}\exp[-(y - x_i\beta)^2/(2\sigma^2)] = (1/\sigma)\varphi[(y - x_i\beta)/\sigma], \; y > 0$$

$$\mathrm{P}(y_i = 0|x_i) = 1 - \Phi(x_i\beta/\sigma),$$

Με συνάρτηση λογαριθμικής πιθανοφάνειας:

$$\ell_i(\beta, \sigma) = 1(y_i = 0)\log[1 - \varphi(x_i\beta/\sigma)] + 1(y_i > 0)\log\{(1/\sigma)\varphi[(y_i - x_i\beta)/\sigma]\}.$$

Έλεγχος λόγου πιθανοφάνειας παραμένει εφαρμόσιμος για πολλαπλούς περιορισμούς

Ερμηνεία εκτιμήσεων tobit:

β μετράνε επίδραση σε y* ενώ συνήθως το ενδιαφέρον εστιάζεται στο y:

πχ ευαισθησία εργασίας στην φορολογία

$$\mathrm{E}(y|x) = \mathrm{P}(y > 0|x) \cdot \mathrm{E}(y|y > 0, x) = \Phi(x\beta/\sigma) \cdot \mathrm{E}(y|y > 0, x)$$

$$\mathrm{E}(y|y > 0, x) = x\beta + \sigma\lambda(x\beta/\sigma).$$

$$\lambda(c) = \varphi(c)/\Phi(c) \quad \textbf{ανεστραμμένος λόγος Mills}$$

σλ>0 και συσχετισμένο με x

$$\mathrm{E}(y|x) = \Phi(x\beta/\sigma)[x\beta + \sigma\lambda(x\beta/\sigma)] = \Phi(x\beta/\sigma)x\beta + \sigma\varphi(x\beta/\sigma) \quad \mathbf{> 0}$$

Ερμηνεία εκτιμήσεων tobit (συνέχεια):

$$\partial\mathrm{E}(y|y > 0, x)/\partial x_j = \beta_j + \beta_j \cdot \frac{d\lambda}{dc}(x\beta/\sigma) \quad = \beta_j\{1 - \lambda(x\beta/\sigma)[x\beta/\sigma + \lambda(x\beta/\sigma)]\}$$

σλ>0 και συσχετισμένο με x

xβ+σλ >0

$$\frac{\partial\mathrm{E}(y|x)}{\partial x_j} = \frac{\partial\mathrm{P}(y > 0|x)}{\partial x_j} \cdot \mathrm{E}(y|y > 0, x) + \mathrm{P}(y > 0|x) \cdot \frac{\partial\mathrm{E}(y|y > 0, x)}{\partial x_j}$$

Επειδή $\mathrm{P}(y > 0|x) = \varphi(x\beta/\sigma)$,

$$\frac{\partial\mathrm{P}(y > 0|x)}{\partial x_j} = (\beta_j/\sigma)\varphi(x\beta/\sigma),$$

=> $$\frac{\partial\mathrm{E}(y|x)}{\partial x_j} = \beta_j\Phi(x\beta/\sigma)$$

Ερμηνεία εκτιμήσεων tobit (συνέχεια):

Ελαστικότητα: $$\frac{\partial \mathrm{E}(y|y>0,x)}{\partial x_1} \cdot \frac{x_1}{\mathrm{E}(y|y>0,x)}$$

Για διακριτές μεταβλητές παρόμοια λογική

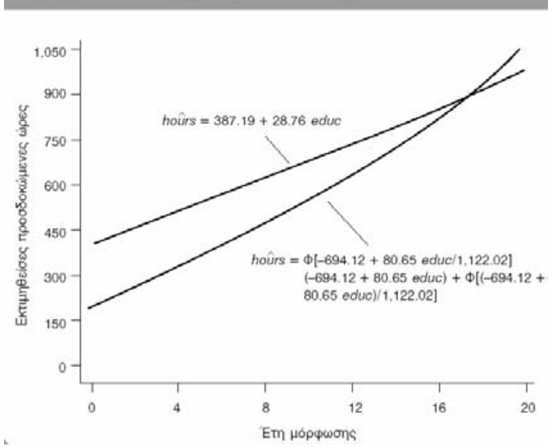
Πχ17.2 Ώρες εργασίας παντρεμένων γυναικών

n=753, 325 δούλεψαν 0 ώρες, 428 δούλεψαν από 12 ως 4,950 ώρες

Εξαρτημένη μεταβλητή: *hours*		
Ανεξάρτητες μεταβλητές	**Γραμμική (Ελαχίστων τετραγώνων)**	**Tobit (Εκτίμηση μέγιστης πιθανοφάνειας)**
nwifeinc	−3.45 (2.54)	−8.81 (4.46)
educ	28.76 (12.95)	80.65 (21.58)
exper	65.67 (9.96)	131.56 (17.28)
$exper^2$	−.700 (.325)	-1.86 (.54)
age	−30.51 (4.36)	−54.41 (7.42)
kidslt6	−442.09 (58.85)	-894.02 (111.88)
kidsge6	−32.78 (23.18)	-16.22 (38.64)
σταθερά	1,330.48 (270.78)	965.31 (446.44)
Τιμή λογαριθμικής πιθανοφάνειας	———	-3,819.09
R-τετράγωνο	.266	.274
$\hat{\sigma}$	750.18	1,122.02

Παράγοντας προσαρμογής μερικών επιδράσεων =**0.645**

Εκτιμηθείσες προσδοκώμενες τιμές της *hours* όσον αφορά τη μόρφωση για το γραμμικό μοντέλο και το μοντέλο Tobit



Ποιότητα προσδιορισμού υποδείγματος Tobit

Σε αντίθεση με ΓΜ αν δεν ισχύουν υποθέσεις τα μεγάλα δείγματα δεν εξασφαλίζουν καλές ιδιότητες εκτιμήσεων

Συχνό πρόβλημα προσδιορισμού:

Ίδιο πρόσημο σε:

$$\partial \mathrm{E}(y|y>0,x)/\partial x_j = \beta_j\{1-\lambda(x\beta/\sigma)[x\beta/\sigma+\lambda(x\beta/\sigma)]\}$$

$$\frac{\partial \mathrm{P}(y>0|x)}{\partial x_j} = (\beta_j/\sigma)\varphi(x\beta/\sigma).$$

πρόβλημα πχ αν Pr(ασφάλισης) αρνητικά με ηλικία αλλά ποσότητα ασφάλισης θετικά (επειδή αξία ζωής μεγαλύτερη)

Έλεγχος προσδιορισμού μέσω σύγκρισης με probit

Εξαρτημένη w=1[y>0]

Επειδή: $P(y_i = 0|x_i) = 1 - \Phi(x_i\beta/\sigma)$

w θα ακολουθεί probit με παραμέτρους β/σ

Οπότε αν αποκλίνουν στατιστικά σημαντικά εκτιμήσεις και πρόσημα τότε υπάρχει πρόβλημα προσδιορισμού

Υπόδειγμα poisson

Για εξαρτημένη διακριτή θετική μεταβλητή καταμέτρησης» y=0,1,2,3,…

πχ

Αριθμός παιδιών σε μια οικογένεια

Αριθμός συλλήψεων ανά έτος για ένα άτομο

Αριθμός αιτήσεων για πατέντα μιας επιχείρησης

$$E(y|x_1,x_2,...,x_k) = \exp(\beta_0 + \beta_1 x_1 + ... + \beta_k x_k)$$

=> $$\log[E(y|x_1,x_2,...,x_k)] = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k$$

=> $$\%\Delta E(y|x) \approx (100\beta_j)\Delta x_j$$

Κατανομή Poisson για μεταβλητή καταμέτρησης

$$P(y = h|x) = \exp[-\exp(x\beta)][\exp(x\beta)]^h/h!,\ h = 0, 1, ...,$$

Οπότε συνάρτηση λογαριθμικής πιθανοφάνειας για τυχαίο δείγμα:

$$\mathcal{L}(\beta) = \sum_{i=1}^{n} \ell_i(\beta) = \sum_{i=1}^{n} \{y_i x_i \beta - \exp(x_i\beta)\}$$

Τυπικές αποκλίσεις εκτιμήσεων μέγιστης πιθανοφάνειας, έλεγχοι περιορισμών λόγου πιθανοφάνειας κλπ υπολογίζονται από στατιστικά πακέτα

Μερική επίδραση:

$$\partial E(y|x_1,x_2,...,x_k)/x_j = \exp(\beta_0 + \beta_1 x_1 + ... + \beta_k x_k)\cdot\beta_j$$

Προσδιορισμός:

Συχνά υπερβολικά περιοριστικό ότι στην κατανομή Poisson:

Var(y|x)=E(y|x)

Όμως ακόμα και αν η κατανομή δεν είναι σωστή οι εκτιμητές παραμένουν συνεπείς και ασυμπτωτικά κανονικοί αλλά η τυπική τους απόκλιση απαιτεί ειδικό υπολογισμό

Υπολογισμός διευκολύνεται αν υποθέσουμε:

$$\mathrm{Var}(y|x) = \sigma^2 E(y|x)$$

Όπου σ>=<1

Πχ. 17.3 Αριθμός συλλήψεων ανδρών με προηγούμενες καταδίκες

n = 2725,

#(narr86=0)=1970,

#(narr86>5)=8

Εξαρτημένη μεταβλητή: *narr86*		
Ανεξάρτητες μεταβλητές	Γραμμική (μέθοδος ελαχίστων τετραγώνων)	Εκθετική (εκτίμηση οιονεί μέγιστης πιθανοφάνειας κατά Poisson)
pcnv	–.132 (.040)	–.402 (.085)
avgsen	–.011 (.12)	–.024 (.020)
tottime	.012 (.009)	.024 (.015)
ptime86	–.041 (.009)	–.099 (.021)
qemp86	–.051 (.014)	–.038 (.029)
inc86–442.09	–.0015 (.0003)	–.0081 (.0010)
black	–.327 (.045)	.661 (.074)
hispan	.194 (.040)	.500 (.074)
born60	–.022 (.033)	.051 (.064)
σταθερά	.577 (.038)	–.600 (.067)
Τιμή λογαριθμικής πιθανοφάνειας	———	–2,248.76
R-τετράγωνο	.073	.077
$\hat{\sigma}$	.829	1.232

Υπόδειγμα παλινδρόμησης με censoring

Παρομοίως με log/prob/tob-it:

$$y_i = \beta_0 + x_i\beta + u_i, u_i|x_i, c_i \sim \text{Normal}(0, \sigma^2) \qquad w_i = \min(y_i, c_i)$$

Όπου:

u ανεξάρτητο του c

c *μπορεί να διαφέρει* ανα άτομο

το ενδιαφέρον μας εστιάζεται στην *y* αλλά παρατηρούμε την *w*

πχ

Οικογενειακός πλούτος (σε ερωτηματολόγια με ερωτήσεις όπως >$500k)

Οριακή παραγωγικότητα όπως αντανακλάται μέσω μισθών με ελάχιστο όριο

Διάρκειες εώς γεγονός, πχ επόμενη σύλληψη

Όπως στο tobit ΕΕΤ μπορεί να οδηγήσει σε ασυνεπείς εκτιμήσεις

Εδώ λόγω τρόπου συλλογής δεδομένων όχι λόγω συμπεριφοράς ατόμων

Η πιθανότητα μια παρατήρηση να υποστεί censoring (περικομμένη τιμή) είναι:

$$P(w_i = c_i|x_i) = P(y_i \geq c_i|x_i) = P(u_i \geq c_i - x_i\beta) = 1 - \Phi[(c_i - x_i\beta)/\sigma]$$

Για τις υπόλοιπες η πυκνότητα του *w* είναι ίση με του *y* οπότε:

$$f(w|x_i, c_i) = 1 - \Phi[(c_i - x_i\beta)/\sigma], \; w = c_i$$

$$= (1/\sigma)\varphi[(w - x_i\beta)/\sigma], \; w < c_i.$$

Συνάρτηση λογαριθμικής πιθανοφάνειας προκύπτει άμεσα

Ερμηνεία παραμέτρων όπως και στο κλάσσικό υπόδειγμα - δεν χρειάζονται προσαρμογή όπως στο tobit

Πχ 17.4 Διάρκεια υποτροπής

durat: διάρκεια σε μήνες για έγκλειστους σε φυλακή Βόρειας Καρολίνας

n=1445

893 δεν επανασυνελήφθησαν [**62%**]

c= 70-81 ανάλογα με περίοδο παρακολούθησης του κάθε φυλακισμένου

priors = # καταδικών

tserved = # μηνών

felon = ψευδομεταβλητή εγκληματία

alcohol = ψευδομεταβλητή αλκοολικού

drugs = ψευδομεταβλητή ναρκομανή

Πίνακας 17.4

Εκτίμηση της εκ Νέου Σύλληψης με Βάση τη Περικομμένη Παλινδρόμη

Εξαρτημένη μεταβλητή: log(*durat*)		ΕΕΤ
Ανεξάρτητες μεταβλητές		
workprg	–.063 (.120)	
priors	–.137 (.021)	-.059 (0.009
tserved	–.019 (.003)	
felon	.444 (.145)	
alcohol	–.635 (.144)	-.262 (0.060
drugs	–.298 (.133)	
black	–.543 (.117)	
married	.341 (.140)	
educ	.023 (.025)	
age	.0039 (.0006)	
σταθερά	4.099 (.348)	
Τιμή λογαριθμικής πιθανοφάνειας	–1,597.06	
$\hat{\sigma}$	1.810	

ΕΕΤ δίνουν πολύ διαφορετικές εκτιμήσεις (υποεκτιμούν επιδράσεις)

Αν δεν ισχύει η κανονικότητα η προσέγγιση μας οδηγεί επίσης σε προβληματικές εκτιμήσεις

=> Σημαντικό να αποφεύγεται η περικοπή δεδομένων

Υπόδειγμα παλινδρόμησης με truncation

Παρομοίως με censoring:

$$y = \beta_0 + x\beta + u,\ u|x \sim \text{Normal}(0,\sigma^2)$$

Όπου δείγμα ***δεν είναι τυχαίο*** γιατί παρατηρούμε μόνο y_i<=c_i:

censoring μπορεί να είναι και μεγαλύτερο απλά δεν περικόβεται η *τιμή* του

Truncation περικόβεται η *παρατήρηση*

πχ

Δείγματα φτωχών οικογενειών

Κατανομή y_i δεδομένου ότι y_i<=c_i

$$g(y|x_i, c_i) = \frac{f(y|x_i\beta, \sigma^2)}{F(c_i|x_i\beta, \sigma^2)},\ y \le c_i,$$

- Μεγιστοποιόντας πιθανοφάνεια παίρνουμε συνεπείς και ασυμπτωτικά κανονικές εκτιμήσεις
- Απαιτεί κανονικότητα και ομοσκεδαστικότητα
- Αν περικόψουμε *παρατηρήσεις* σε δείγματα με περικομένες *τιμές* χάνουμε πληροφορία [υπήρχαν άτομα με υψηλές τιμές και συγκεκριμένα χαρακτηριστικά]
- ΕΕΤ σε δείγμα με περικομένες παρατηρήσεις υποεκτιμά επιδράσεις

Διορθώσεις επιλογής δείγματος

Γενίκευση παλινδρόμησης με truncation σε παλινδρομήσεις με *μη τυχαία δείγματα*

Πχ

Λόγω τρόπου συλλογής δεδομένων / μορφής ερωτηματολογίων

Ελλείψεις σε κάποια στοιχεία (ανεξάρτητη η εξαρτημένη)

Ελλείψεις σε y ως συνάρτηση άλλης μεταβλητής (incidental truncation)

π.χ. Εξίσωση προσφοράς μισθού που παρατηρείται μόνο για όσους συμβαίνει να εργάζονται

Σε μη ισοροπημένα πάνελ π.χ. όταν άτομα αποχορούν από το δείγμα

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + u,\ \mathrm{E}(u|x_1, x_2, \dots, x_k) = 0$$ ή $$y_i = x_i\beta + u_i$$

s_i μεταβλητή που καθορίζει άν το *i* συμπεριλαμβάνεται στο δείγμα {0,1}

$$s_i y_i = s_i x_i \beta + s_i u_i.$$

Συνεπής και αμερόληπτη εκτίμηση με ΕΕΤ απαιτεί:

$$\mathrm{E}(su) = 0$$

$$\mathrm{E}[(sx_j)(su)] = \mathrm{E}(sx_j u) = 0$$

Συμβατές περιπτώσεις για εξίσωση προσφοράς μισθού:

1. *s* συνάρτηση του *x* [εξωγενής επολογή δείγματος]

   πχ s=0 εξαρτάται από φύλο, εμπειρία κλπ

2. s ανεξάρτητο του *x,u*

3. s εξαρτάται από *x* και άλλους τυχαίους παράγοντες ανεξάρτητους του *u*

   πχ s=0 αν IQ<v όπου *v* ανεξάρτητο του *x*

Παράδειγμα ασυνέπειας:
Παλινδρόμηση με truncation: s_i=0 ⇔ y_i>c_i ⇔*u_i>c_i-x_iβ*=> συσχέτιση

Incidental truncation

π.χ. Εξίσωση προσφοράς εισοδήματος y=log(wage) όπου *s* καθορίζεται από το αν εργάζεται ο *i*

$$y = \boldsymbol{x}\boldsymbol{\beta} + u,\ \mathrm{E}(u|\boldsymbol{x}) = 0$$

$$s = 1[\boldsymbol{z}\boldsymbol{\gamma} + v \geq 0],$$

$$\mathrm{E}(u|\boldsymbol{x}, \boldsymbol{z}) = 0.$$

x **αυστηρό** υποσύνολο του *z*

v~N(0,1) και *u* ανεξάρτητο του *z*

=> $$\mathrm{E}(y|\boldsymbol{z}, v) = \boldsymbol{x}\boldsymbol{\beta} + \mathrm{E}(u|\boldsymbol{z}, v) = \boldsymbol{x}\boldsymbol{\beta} + \mathrm{E}(u|v),$$

=> $$\mathrm{E}(y|\boldsymbol{z}, v) = \boldsymbol{x}\boldsymbol{\beta} + \varrho v.$$

=> $$\mathrm{E}(y|\boldsymbol{z}, s) = \boldsymbol{x}\boldsymbol{\beta} + \varrho \mathrm{E}(v|\boldsymbol{z}, s)$$

=> $$\mathrm{E}(y|\boldsymbol{z}, s = 1) = \boldsymbol{x}\boldsymbol{\beta} + \varrho\lambda(\boldsymbol{z}\boldsymbol{\gamma})$$

Αν ρ~=0

Εκτίμηση Heckit:

1. Εκτιμούμε γ με probit χρησιμοποιόντας ότι $\mathrm{P}(s = 1|\boldsymbol{z}) = \Phi(\boldsymbol{z}\boldsymbol{\gamma})$ και δεδομένα για *s_i,z_i*
2. Υπολογίζουμε λ^_i=λ(z_iγ^)
3. Για παρατηρούμενα *i (s_i=1)* παλινρομούμε y σε x,λ^

Στατιστική t_λ δίνει έλεγχο για ρ=0 [τυχαίο δείγμα]
Τυπικά σφάλματα παλινδρόμησης απαιτούν (μικρή) διόρθωση

Παράδειγμα: Προσφορά μισθού παντρεμένων γυναικών

n=743

Σs_i=428 εργαζόμενες

Εξίσωση Προσφοράς Μισθού για τις Έγγαμες Γυναίκες

Εξαρτημένη μεταβλητή: log(*wage*)		
Ανεξάρτητες μεταβλητές	**Μέθοδος ελαχίστων τετραγώνων**	**Μέθοδος Heckit**
educ	.108 (.014)	.109 (.016)
exper	.042 (.012)	.044 (.016)
(Σχέση)	–.00081 (.00039)	–.00086 (.00044)
σταθερά	–.522 (.199)	–.578 (.307)
$\hat{\lambda}$	———	.032 (.134)
Μέγεθος δείγματος *R*-τετράγωνο	428 .157	428 .157

Παράδειγμα: Συμμετοχή έγγαμων γυναικών στο εργατικό δυναμικό

N=753

1975

Σy=428
Kids=#
nwifeinc – εισόδημα συζ $k

Εξαρτημένη μεταβλητή: *inlf*

Ανεξάρτητες μεταβλητές	Γραμμικό μοντέλο πιθανότητας (μέθοδος ελαχίστων τετραγώνων)	Logit (Εκτίμηση μέγιστης πιθανοφάνειας)	Probit (Εκτίμηση μέγιστης πιθανοφάνειας)
nwifeinc	–.0034 (.0015)	–.021 (.008)	–.012 (.005)
educ	.038 (.007)	.221 (0.43)	.131 (.025)
exper	.039 (.006)	.206 (.032)	.123 (.019)
exper^2	–.00060 (.00018)	–.0032 (.0010)	–.0019 (.0006)
age	–.016 (.002)	–.088 (.015)	–.053 (.008)
kidslt6	–.262 (.032)	–1.443 (.204)	–.868 (.119)
kidsge6	.013 (.013)	.060 (.075)	.036 (.043)
σταθερά	.586 (.151)	.425 (.860)	.270 (.509)
Ποσοστό σωστής πρόβλεψης Τιμή λογαριθμικής πιθανοφάνειας Ψευδές *R*-τετράγωνο	73.4 ——— .264	73.6 –401.77 .220	73.4 –401.309 .221