

ΔΜ με χρήση βοηθητικού χαρακτηριστικού

Περιγραφή:

Σε πληθυσμό μεγέθους N θεωρούμε τα χαρακτηριστικά

- U με παραμέτρους μ_U (μέση τιμή), τ_U (ολική τιμή), σ_{0U}^2 (διασπορά), και
- V με παραμέτρους μ_V (μέση τιμή), τ_V (ολική τιμή), σ_{0V}^2 (διασπορά)

Παίρνουμε δείγμα μεγέθους n με

- X_i η τιμή της i μονάδας του δείγματος για το χαρακτηριστικό U , και
- Y_i η τιμή της i μονάδας του δείγματος για το χαρακτηριστικό V

Ορίζουμε το λόγο (ratio)

$$r = \frac{\mu_U}{\mu_V} = \frac{\tau_U}{\tau_V}$$

ΔΜ με χρήση βοηθητικού χαρακτηριστικού

Παρατηρήσεις:

1. Ο λόγος r έχει ενδιαφέρον από μόνος του, π.χ.

μονάδα: οικογένεια

U : μηνιαίο εισόδημα οικογένειας

V : αριθμός των μελών της οικογένειας

τότε

$$r = \frac{\tau_U}{\tau_V} = \frac{\text{ολικό εισόδημα}}{\text{ολικός αριθμός μελών}} = \text{κατά κεφαλήν εισόδημα}$$

2. Ο λόγος r μπορεί να χρησιμοποιηθεί στην εκτίμηση των παραμέτρων του χαρακτηριστικού U

$$\mu_U = r \cdot \mu_V$$

$$\tau_U = r \cdot \tau_V$$

δηλαδή το V παίζει το ρόλο βοηθητικού χαρακτηριστικού για το οποίο όμως γνωρίζουμε το μ_V ή/και το τ_V

ΔΜ με χρήση βοηθητικού χαρακτηριστικού

Παρατηρήσεις (συνέχεια):

- π.χ. ένας έμπορος ενδιαφέρεται για τη χυμοποίηση ενός φορτίου πορτοκαλιών

μονάδα: πορτοκάλι

U : ποσότητα χυμού στο πορτοκάλι

ο έμπορος ενδιαφέρεται για το $\tau_U = N \cdot \mu_U$

- Για την εκτίμηση του μ_U (μέση ποσότητα χυμού ανά πορτοκάλι) μπορεί να πάρει ένα δείγμα πορτοκαλιών και να υπολογίσει τη μέση ποσότητα χυμού ανά πορτοκάλι στο δείγμα, $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$, συνεπώς

$$\hat{\tau}_U = N \cdot \bar{X} \tag{1}$$

- Για το N όμως, θα χρειαστεί να μετρήσει **ένα-ένα** τα πορτοκάλια του φορτίου; Προφανώς, μία τέτοια διαδικασία **δεν** είναι πρακτική!

ΔΜ με χρήση βοηθητικού χαρακτηριστικού

Παρατηρήσεις (συνέχεια):

2. Εναλλακτικά, κάνουμε χρήση βοηθητικού χαρακτηριστικού V . Ορίζουμε

V : βάρος του πορτοκαλιού

- μ_V : μέσο βάρος πορτοκαλιού
- τ_V : ολικό βάρος των πορτοκαλιών

Από το δείγμα πορτοκαλιών που έχει πάρει, ως εκτίμηση του μ_V μπορεί να υπολογίσει το μέσο βάρος ανά πορτοκάλι στο δείγμα, $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$

- Ισχύει

$$\tau_U = r \tau_V$$

όπου το τ_V όμως μπορεί να μετρηθεί εύκολα (γνωστό), για παράδειγμα με χρήση πλάστιγγας

- μπορούμε να εκτιμήσουμε το r με

$$\hat{r} = \frac{\bar{X}}{\bar{Y}}$$

οπότε μπορούμε να εκτιμήσουμε το τ_U με

$$\hat{\tau}_U = \hat{r} \tau_V = \frac{\bar{X}}{\bar{Y}} \tau_V$$

ΔΜ με χρήση βοηθητικού χαρακτηριστικού

Παρατηρήσεις (συνέχεια):

2. Σχόλιο: Είναι

$$\hat{\tau}_U = \frac{\bar{X}}{\bar{Y}} \tau_V = \frac{\tau_V}{\bar{Y}} \cdot \bar{X}$$

Η ποσότητα $\frac{\tau_V}{\bar{Y}}$ αποτελεί στην ουσία εκτίμηση του N στη σχέση (1), αφού

$$\hat{N} = \frac{\tau_V}{\bar{Y}} = \frac{\text{ολικό βάρος πορτοκαλιών του φορτίου}}{\text{εκτίμηση μέσου βάρους ανά πορτοκάλι}}$$

Εκτίμηση λόγου r σε ΑΤΔ

Εκτιμητής του $r = \frac{\mu_U}{\mu_V} = \frac{\tau_U}{\tau_V}$ είναι ο

$$\hat{r} = \frac{\bar{X}}{\bar{Y}} = \frac{\sum_i X_i}{\sum_i Y_i}$$

Πρόταση 1

Σε ΑΤΔ

1. Ο $\hat{r}$ δεν είναι ΑΕ του r .
2. Για $n \rightarrow N$ (n «μεγάλο») ισχύει $E\hat{r} \approx r$ (δηλ. προσεγγιστικά ΑΕ)
3. Για $n \rightarrow N$ (n «μεγάλο») ισχύει $\text{Var}(\hat{r}) \approx \frac{1}{\mu_V^2} (1 - f) \frac{\sigma_{0D}^2}{n}$, όπου

$$\sigma_{0D}^2 = \frac{1}{N-1} \sum_{j=1}^N (U_j - rV_j)^2$$

Απόδειξη.

Εκτιμητές λόγου των μ_U, τ_U σε ΑΤΔ

Εκτιμητής λόγου του μ_U σε ΑΤΔ

$$\hat{\mu}_U = \hat{r}\mu_V = \frac{\bar{X}}{\bar{Y}}\mu_V$$

Εκτιμητής λόγου του τ_U σε ΑΤΔ

$$\hat{\tau}_U = \hat{r}\tau_V = \frac{\bar{X}}{\bar{Y}}\tau_V$$

Παρατήρηση:

Ισχύει $\hat{\tau}_U = \hat{r}\tau_V = \hat{r}N\mu_V = N\hat{\mu}_U$

Εκτιμητές λόγου των μ_U, τ_U σε ΑΤΔ

Ιδιότητες των $\hat{\mu}_U, \hat{\tau}_U$:

1. Ο $\hat{\mu}_U$ δεν είναι ΑΕ του μ_U , και ο $\hat{\tau}_U$ δεν είναι ΑΕ του τ_U .
2. Για $n \rightarrow N$ είναι

$$E\hat{\mu}_U \approx \mu_U \quad \text{και} \quad E\hat{\tau}_U \approx \tau_U \quad (\text{δηλ. προσεγγιστικά ΑΕ})$$

3. Για $n \rightarrow N$ είναι

$$\text{Var}(\hat{\mu}_U) \approx (1-f) \frac{\sigma_{0D}^2}{n} \quad \text{και} \quad \text{Var}(\hat{\tau}_U) \approx N^2 \text{Var}(\hat{\mu}_U),$$

$$\text{όπου } \sigma_{0D}^2 = \frac{1}{N-1} \sum_{j=1}^N (U_j - rV_j)^2$$

4. Για $n \rightarrow N$ είναι

$$\hat{\text{Var}}(\hat{\mu}_U) \approx (1-f) \frac{\hat{\sigma}_{0D}^2}{n},$$

$$\text{όπου } \hat{\sigma}_{0D}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \hat{r}Y_i)^2 \text{ και}$$

$$\hat{\text{Var}}(\hat{\tau}_U) = \begin{cases} N^2 \hat{\text{Var}}(\hat{\mu}_U), & N \text{ γνωστό} \\ \left(\frac{\tau_U}{\bar{Y}}\right)^2 \hat{\text{Var}}(\hat{\mu}_U), & N \text{ άγνωστο} \end{cases}$$

Σύγκριση ΑΤΔ χωρίς βοηθητικό με ΑΤΔ με βοηθητικό

Για $(U_1, V_1), \dots, (U_N, V_N)$

- $\sigma_{0U}^2 = \frac{1}{N-1} \sum_{j=1}^N (U_j - \mu_U)^2$, η διασπορά του U
- $\sigma_{0V}^2 = \frac{1}{N-1} \sum_{j=1}^N (V_j - \mu_V)^2$, η διασπορά του V
- $\sigma_{0UV} = \frac{1}{N-1} \sum_{j=1}^N (U_j - \mu_U)(V_j - \mu_V)$, η συνδιασπορά των U, V
- $\rho = \frac{\sigma_{0UV}}{\sigma_{0U} \cdot \sigma_{0V}}$, ο συντελεστής συσχέτισης των U, V

Εκτίμηση του μ_U

χωρίς βοηθητικό	με βοηθητικό
-----------------	--------------

$\bar{X}$

με $\text{Var}(\bar{X}) = (1 - f) \frac{\sigma_0^2}{n}$

$\hat{\mu}_U = \hat{r} \mu_V$

με $\text{Var}(\hat{\mu}_U) \approx (1 - f) \frac{\sigma_{0D}^2}{n}$,
για $n \rightarrow N$

Σύγκριση ΑΤΔ χωρίς βοηθητικό με ΑΤΔ με βοηθητικό

Πρόταση 2

Για $n \rightarrow N$ είναι $\text{Var}(\hat{\mu}_U) < \text{Var}(\bar{X})$ (δηλαδή $\hat{\mu}_U$ πιο ακριβής από $\bar{X}$) εάν $r > 0$ και $\rho > \frac{1}{2} \frac{CV(V)}{CV(U)}$, όπου $CV(U) = \frac{\sigma_{0U}}{\mu_U}$, $CV(V) = \frac{\sigma_{0V}}{\mu_V}$ οι συντελεστές μεταβλητότητας των U, V , αντίστοιχα.

Απόδειξη.

Παρατηρήσεις:

- Γενικά μπορεί $\frac{1}{2} \frac{CV(V)}{CV(U)} > 1$ οπότε δεν ισχύει σε αυτή την περίπτωση η συνθήκη αφού $\rho \leq 1$
- Αν το V θεωρηθεί το U σε προγενέστερη χρονική περίοδο, τότε $\rho > 1/2$.
- Ανάλογο αποτέλεσμα ισχύει για την εκτίμηση του τ .

$$\hat{\tau}_U = \hat{r}\tau_V = N\hat{\mu}_V$$

$$\hat{\tau}_U = N\bar{X}$$

Εκτιμητές λόγου των μ_U, τ_U σε ΣΤΔ

Πληθυσμός: $(U_{\ell j}, V_{\ell j}), \ell = 1, \dots, L, j = 1, \dots, N_\ell$

Δείγμα: $(X_{\ell i}, Y_{\ell i}), \ell = 1, \dots, L, i = 1, \dots, n_\ell$

Υπάρχουν δύο κατηγορίες εκτιμητών λόγου στη ΣΤΔ

- χωριστοί εκτιμητές λόγου (separate ratio estimators)
- συνδυασμένοι εκτιμητές λόγου (combined ratio estimators)

Εκτιμητές λόγου των μ_U, τ_U σε ΣΤΔ

Α. Χωριστοί εκτιμητές λόγου:

Εκτίμηση του μ_U : $\mu_U = \sum_{\ell=1}^L W_{\ell} \mu_{U\ell}$

Εκτιμούμε το $\mu_{U\ell}$ **χωριστά** σε κάθε στρώμα χρησιμοποιώντας εκτιμητή λόγου σε ΑΤΔ, δηλαδή

$$\hat{\mu}_{U\ell} = \hat{r}_{\ell} \mu_{V\ell}, \quad \text{όπου } \hat{r}_{\ell} = \frac{\bar{X}_{\ell}}{\bar{Y}_{\ell}}$$

Ο χωριστός εκτιμητής λόγου του μ_U δίνεται από

$$\hat{\mu}_{U,st,s} = \sum_{\ell=1}^L W_{\ell} \hat{\mu}_{U\ell} = \sum_{\ell=1}^L W_{\ell} \hat{r}_{\ell} \mu_{V\ell} = \sum_{\ell=1}^L W_{\ell} \frac{\bar{X}_{\ell}}{\bar{Y}_{\ell}} \mu_{V\ell}$$

Εκτιμητές λόγου των μ_U, τ_U σε ΑΤΔ

Ιδιότητες του $\hat{\mu}_{U,st,s}$:

1. Ο $\hat{\mu}_{U,st,s}$ δεν είναι ΑΕ του μ_U .
2. Για $n_\ell \rightarrow N_\ell, \forall \ell$, είναι $E\hat{\mu}_{U\ell} \approx \mu_{U\ell}, \forall \ell$ (λόγω ΑΤΔ σε κάθε στρώμα), οπότε

$$E\hat{\mu}_{U,st,s} \approx \mu_U \quad (\text{δηλ. προσεγγιστικά ΑΕ})$$

3. Για $n_\ell \rightarrow N_\ell, \forall \ell$, είναι

$$\text{Var}(\hat{\mu}_{U,st,s}) = \sum_{\ell=1}^L W_\ell^2 \text{Var}(\hat{\mu}_{U\ell}) \approx \sum_{\ell=1}^L W_\ell^2 (1 - f_\ell) \frac{\sigma_{0D_\ell}^2}{n_\ell},$$

$$\text{όπου } \sigma_{0D_\ell}^2 = \frac{1}{N_\ell - 1} \sum_{j=1}^{N_\ell} (U_{\ell j} - r_\ell V_{\ell j})^2$$

4. Για $n_\ell \rightarrow N_\ell, \forall \ell$, είναι

$$\hat{\text{Var}}(\hat{\mu}_{U,st,s}) = \sum_{\ell=1}^L W_\ell^2 (1 - f_\ell) \frac{\hat{\sigma}_{0D_\ell}^2}{n_\ell},$$

$$\text{όπου } \hat{\sigma}_{0D_\ell}^2 = \frac{1}{n_\ell - 1} \sum_{i=1}^{n_\ell} (X_{\ell i} - \hat{r}_\ell Y_{\ell i})^2$$

Εκτιμητές λόγου των μ_U, τ_U σε ΣΔ

Εκτίμηση του τ_U : $\tau_U = N\mu_U$

Ο χωριστός εκτιμητής λόγου του τ_U δίνεται από

$$\hat{\tau}_{U,st,s} = N\hat{\mu}_{U,st,s} \quad (2)$$

$$\begin{aligned} &= N \sum_{\ell=1}^L W_{\ell} \hat{r}_{\ell} \mu_{V\ell} = \sum_{\ell=1}^L N_{\ell} \hat{r}_{\ell} \mu_{V\ell} = \sum_{\ell=1}^L \hat{r}_{\ell} \tau_{V\ell} \\ &= \sum_{\ell=1}^L \hat{\tau}_{U\ell} \end{aligned} \quad (3)$$

Εκτιμητές λόγου των μ_U, τ_U σε ΣΔ

Παρατηρήσεις:

1. Η σχέση (2) είναι εύχρηστη για θεωρητικές αποδείξεις.
2. Η σχέση (3) αποτελεί μία εναλλακτική μορφή του $\hat{\tau}_{U,st,s}$, η οποία μπορεί να προκύψει και με βάση τη σχέση

$$\tau_U = \sum_{\ell=1}^L \tau_{U\ell}$$

κάνοντας εκτίμηση των $\tau_{U\ell}$ με τους αντίστοιχους εκτιμητές λόγου (σε κάθε στρώμα χωριστά) $\hat{\tau}_{U\ell} = \hat{r}_{\ell} \tau_{V\ell}$.

Ιδιότητες του $\hat{\tau}_{U,st,s}$:

Προκύπτουν άμεσα από τη σχέση $\hat{\tau}_{U,st,s} = N \hat{\mu}_{U,st,s}$ και τις ιδιότητες του $\hat{\mu}_{U,st,s}$.

Εκτιμητές λόγου των μ_U, τ_U σε ΣΔ

Β. Συνδυασμένοι εκτιμητές λόγου:

Εκτίμηση του μ_U : $\mu_U = r\mu_V$ (απαιτεί γνώση του μ_V)

Κάνουμε εκτίμηση του λόγου r συνολικά (συνδυασμένα) από όλα τα στρώματα με

$$\hat{r} = \frac{\bar{X}_{st}}{\bar{Y}_{st}} = \frac{\sum_{\ell=1}^L W_{\ell} \bar{X}_{\ell}}{\sum_{\ell=1}^L W_{\ell} \bar{Y}_{\ell}}$$

Ο συνδυασμένος εκτιμητής λόγου του μ_U δίνεται από

$$\hat{\mu}_{U,st,c} = \hat{r}\mu_V = \frac{\bar{X}_{st}}{\bar{Y}_{st}}\mu_V$$

Εκτιμητές λόγου των μ_U, τ_U σε ΑΤΔ

Ιδιότητες του $\hat{\mu}_{U,st,c}$:

1. Ο $\hat{\mu}_{U,st,c}$ δεν είναι ΑΕ του μ_U .
2. Για $n \rightarrow N$ είναι

$$E\hat{\mu}_{U,st,c} = \mu_V E\left(\frac{\bar{X}_{st}}{\bar{Y}_{st}}\right) \approx \mu_V E\left(\frac{\bar{X}_{st}}{\mu_V}\right) = \mu_U \quad (\text{δηλ. προσεγγιστικά ΑΕ})$$

3. Για $n \rightarrow N$ είναι

$$\text{Var}(\hat{\mu}_{U,st,c}) = \text{Var}(\hat{r}\mu_V) \approx \text{Var}(\bar{X}_{st} - r\bar{Y}_{st}) = \sum_{\ell=1}^L W_{\ell}^2 (1 - f_{\ell}) \frac{\sigma_{0D_{\ell}^r}^2}{n_{\ell}},$$

$$\text{όπου } \sigma_{0D_{\ell}^r}^2 = \frac{1}{N_{\ell}-1} \sum_{j=1}^{N_{\ell}} (U_{\ell j} - rV_{\ell j} - (\mu_{U\ell} - r\mu_{V\ell}))^2$$

4. Για $n \rightarrow N$ είναι

$$\hat{\text{Var}}(\hat{\mu}_{U,st,c}) = \sum_{\ell=1}^L W_{\ell}^2 (1 - f_{\ell}) \frac{\hat{\sigma}_{0D_{\ell}^r}^2}{n_{\ell}},$$

$$\text{όπου } \hat{\sigma}_{0D_{\ell}^r}^2 = \frac{1}{n_{\ell}-1} \sum_{i=1}^{n_{\ell}} (X_{\ell i} - \hat{r}Y_{\ell i} - (\bar{X}_{\ell} - \hat{r}\bar{Y}_{\ell}))^2, \text{ με } \hat{r} = \frac{\bar{X}_{st}}{\bar{Y}_{st}}.$$

Εκτιμητές λόγου των μ_U, τ_U σε ΣΤΔ

Εκτίμηση του τ_U : $\tau_U = N\mu_U$

Ο συνδυασμένος εκτιμητής λόγου του τ_U δίνεται από

$$\hat{\tau}_{U,st,c} = N\hat{\mu}_{U,st,c} \quad (4)$$

$$\begin{aligned} &= N \frac{\bar{X}_{st}}{\bar{Y}_{st}} \mu_V = \frac{\bar{X}_{st}}{\bar{Y}_{st}} \tau_V \\ &= \hat{r} \tau_V \end{aligned} \quad (5)$$

Εκτιμητές λόγου των μ_U, τ_U σε ΣΔ

Παρατηρήσεις:

1. Η σχέση (4) είναι εύχρηστη για θεωρητικές αποδείξεις.
2. Η σχέση (5) αποτελεί μία εναλλακτική μορφή του $\hat{\tau}_{U,st,c}$, η οποία μπορεί να προκύψει και με βάση τη σχέση

$$\tau_U = r\tau_V$$

κάνοντας εκτίμηση του r συνολικά (συνδυασμένα) από όλα τα στρώματα

$$\text{με } \hat{r} = \frac{\bar{X}_{st}}{\bar{Y}_{st}}.$$

Ιδιότητες του $\hat{\tau}_{U,st,c}$:

Προκύπτουν άμεσα από τη σχέση $\hat{\tau}_{U,st,c} = N\hat{\mu}_{U,st,c}$ και τις ιδιότητες του $\hat{\mu}_{U,st,c}$.

Εκτιμητές παλινδρόμησης των μ_U, τ_U σε ΑΤΔ

Πληθυσμός: $(U_j, V_j), j = 1, \dots, N$

Δείγμα: $(X_i, Y_i), i = 1, \dots, n$

Γενική μορφή εκτιμητή παλινδρόμησης του μ_U

$$\hat{\mu}_{U,R} = \bar{X} + b (\mu_V - \bar{Y}), \quad \text{όπου } b \text{ σταθερά}$$

Παρατηρήσεις:

1. Αν $b = 0$ τότε $\hat{\mu}_{U,R} = \bar{X}$
2. Αν $b = \hat{r} = \frac{\bar{X}}{\bar{Y}}$ (καταχρηστικά διότι b σταθερά ανεξάρτητη του δείγματος) τότε $\hat{\mu}_{U,R} = \bar{X} + \hat{r}(\mu_V - \bar{Y}) = \hat{r} \mu_V$

Εκτιμητές παλινδρόμησης των μ_U, τ_U σε ΑΤΔ

Παρατηρήσεις (συνέχεια):

3. Ερμηνεία του $\hat{\mu}_{U,R}$:

Θεωρώντας γραμμική σχέση του V με το U

- Αν $U_j = bV_j, \forall j$, τότε $\mu_U = b\mu_V$,

άρα όταν χρησιμοποιούμε τον $\hat{\mu}_U = \hat{r}\mu_V$ έμμεσα υποθέτουμε τη σχέση $\mu_U = b\mu_V$

- Αν $U_j = bV_j + a, \forall j$, τότε $\mu_U = b\mu_V + a$,

άρα όταν χρησιμοποιούμε τον $\hat{\mu}_{U,R} = b\mu_V + (\bar{X} - b\bar{Y})$ έμμεσα υποθέτουμε τη σχέση $\mu_U = b\mu_V + a$.

Δηλαδή ο $\hat{\mu}_{U,R}$ «διορθώνει» με την ποσότητα $(\bar{X} - b\bar{Y})$ το λάθος που θα γινόταν αν σε αυτή την περίπτωση είχαμε χρησιμοποιήσει εκτιμητή λόγου $\hat{\mu}_U = b\mu_V$.

4. Ο $\hat{\mu}_{U,R}$ είναι ΑΕ του μ_U για κάθε τιμή της σταθεράς b , αφού

$$E(\hat{\mu}_{U,R}) = E\bar{X} + b\mu_V - bE\bar{Y} = \mu_U + b\mu_V - b\mu_V = \mu_U$$

Εκτιμητές παλινδρόμησης των μ_U, τ_U σε ΑΤΔ

Ερώτημα: Υπάρχει τιμή της σταθεράς b που δίνει καλύτερο (πιο ακριβή) εκτιμητή $\hat{\mu}_{U,R}$;

Πρόταση 3

Η διασπορά του $\hat{\mu}_{U,R}$, $\text{Var}(\hat{\mu}_{U,R})$, ελαχιστοποιείται ως προς b για

$$b_0 = \frac{\sigma_{0UV}}{\sigma_{0V}^2} = \rho \frac{\sigma_{0U}}{\sigma_{0V}}$$

και η ελάχιστη τιμή της είναι $(1 - f) \frac{\sigma_{0U}^2}{n} (1 - \rho^2)$.

Απόδειξη.

Παρατήρηση:

Επειδή $|\rho| \leq 1$, είναι $\text{Var}(\hat{\mu}_{U,R}) \leq (1 - f) \frac{\sigma_{0U}^2}{n} = \text{Var}(\bar{X})$.

Εκτιμητές παλινδρόμησης των μ_U, τ_U σε ΑΤΔ

Επειδή

$$b_0 = \frac{\sigma_{0UV}}{\sigma_{0V}^2} = \frac{\sum_{j=1}^N (U_j - \mu_U)(V_j - \mu_V)}{\sum_{j=1}^N (V_j - \mu_V)^2}$$

εξαρτάται από άγνωστες παραμέτρους, χρησιμοποιούμε τον εκτιμητή της

$$\hat{b}_0 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (Y_i - \bar{Y})^2}$$

Εκτιμητής παλινδρόμησης του μ_U σε ΑΤΔ

$$\hat{\mu}_{U,R} = \bar{X} + \hat{b}_0 (\mu_V - \bar{Y})$$

Εκτιμητές παλινδρόμησης των μ_U, τ_U σε ΑΤΔ

Ιδιότητες του $\hat{\mu}_{U,R}$:

1. Ο $\hat{\mu}_{U,R}$ δεν είναι ΑΕ του μ_U .
2. Για $n \rightarrow N$ είναι

$$E\hat{\mu}_{U,R} \approx \mu_U \quad (\text{δηλ. προσεγγιστικά ΑΕ})$$

αφού για $n \rightarrow N$ είναι $\hat{b}_0 \approx b_0$

3. Για $n \rightarrow N$ είναι

$$\text{Var}(\hat{\mu}_{U,R}) \leq \text{Var}(\bar{X}),$$

αφού για $n \rightarrow N$ είναι $\hat{b}_0 \approx b_0$ και εξ ορισμού
 $\text{Var}_{b_0}(\hat{\mu}_{U,R}) \leq \text{Var}_{b=0}(\hat{\mu}_{U,R}) = \text{Var}(\bar{X})$.

4. Για $n \rightarrow N$ είναι

$$\text{Var}(\hat{\mu}_{U,R}) \leq \text{Var}(\hat{\mu}_U), \quad \text{όπου } \hat{\mu}_U = \hat{r}\mu_V,$$

αφού $\text{Var}(\hat{\mu}_U) = \mu_V^2 \text{Var}(\hat{r}) = \mu_V^2 \text{Var}\left(\frac{\bar{X}}{\bar{Y}} - r\right) \approx \text{Var}(\bar{X} - r\bar{Y}) =$
 $\text{Var}(\bar{X} - r(\mu_V - \bar{Y})) \geq \text{Var}_{b_0}(\hat{\mu}_{U,R}) \approx \text{Var}_{\hat{b}_0}(\hat{\mu}_{U,R})$.

Εκτιμητές παλινδρόμησης των μ_U, τ_U σε ΑΤΔ

Εκτιμητής παλινδρόμησης του τ_U σε ΑΤΔ

$$\hat{\tau}_{U,R} = N \hat{\mu}_{U,R} \quad (6)$$

$$= N(\bar{X} + \hat{b}_0 (\mu_V - \bar{Y}))$$

$$= \hat{\tau}_U + \hat{b}_0 (\tau_V - \hat{\tau}_V) \quad (7)$$

Παρατηρήσεις:

1. Η σχέση (6) είναι εύχρηστη για θεωρητικές αποδείξεις.
2. Η σχέση (7) αποτελεί μία εναλλακτική μορφή του $\hat{\tau}_{U,R}$, η οποία μπορεί να προκύψει και με βάση τη γενική μορφή

$$\hat{\tau}_{U,R} = \hat{\tau}_U + b(\tau_V - \hat{\tau}_V), \quad b \text{ σταθερά}$$

όπου $\hat{b}_0$ είναι ο εκτιμητής του b_0 που ελαχιστοποιεί τη διασπορά του $\text{Var}(\hat{\tau}_{U,R})$ ως προς b .

Ιδιότητες του $\hat{\tau}_{U,R}$:

Προκύπτουν άμεσα από τη σχέση $\hat{\tau}_{U,R} = N \hat{\mu}_{U,R}$ και τις ιδιότητες του $\hat{\mu}_{U,R}$.

Εκτιμητές παλινδρόμησης των μ_U, τ_U σε ΣΤΔ

Πληθυσμός: $(U_{\ell j}, V_{\ell j}), \ell = 1, \dots, L, j = 1, \dots, N_\ell$

Δείγμα: $(X_{\ell i}, Y_{\ell i}), \ell = 1, \dots, L, i = 1, \dots, n_\ell$

Υπάρχουν δύο κατηγορίες εκτιμητών παλινδρόμησης στη ΣΤΔ

- χωριστοί εκτιμητές παλινδρόμησης (separate regression estimators)
- συνδυασμένοι εκτιμητές παλινδρόμησης (combined regression estimators)

Εκτιμητές παλινδρόμησης των μ_U, τ_U σε ΣΤΔ

Α. Χωριστοί εκτιμητές παλινδρόμησης:

Εκτίμηση του μ_U : $\mu_U = \sum_{\ell=1}^L W_{\ell} \mu_{U\ell}$

Εκτιμούμε το $\mu_{U\ell}$ **χωριστά** σε κάθε στρώμα χρησιμοποιώντας εκτιμητή παλινδρόμησης σε ΑΤΔ, δηλαδή

$$\hat{\mu}_{U\ell,R} = \bar{X}_{\ell} + \hat{b}_{0\ell} (\mu_{V\ell} - \bar{Y}_{\ell}), \quad \text{όπου} \quad \hat{b}_{0\ell} = \frac{\sum_{i=1}^{n_{\ell}} (X_{\ell i} - \bar{X}_{\ell})(Y_{\ell i} - \bar{Y}_{\ell})}{\sum_{i=1}^{n_{\ell}} (Y_{\ell i} - \bar{Y}_{\ell})^2}$$

Ο **χωριστός εκτιμητής παλινδρόμησης του μ_U** δίνεται από

$$\begin{aligned} \hat{\mu}_{U,st,R,s} &= \sum_{\ell=1}^L W_{\ell} \hat{\mu}_{U\ell,R} \\ &= \sum_{\ell=1}^L W_{\ell} (\bar{X}_{\ell} + \hat{b}_{0\ell} (\mu_{V\ell} - \bar{Y}_{\ell})) \end{aligned}$$

Εκτιμητές παλινδρόμησης των μ_U, τ_U σε ΣΤΔ

Ιδιότητες του $\hat{\mu}_{U,st,R,s}$:

Για $n_\ell \rightarrow N_\ell, \forall \ell$, είναι

1. $\hat{\mu}_{U,st,R,s} \approx \mu_U$ (δηλ. προσεγγιστικά ΑΕ)
2. $\text{Var}(\hat{\mu}_{U,st,R,s}) \approx \sum_{\ell=1}^L W_\ell^2 \text{Var}(\bar{X}_\ell)(1 - \rho_\ell^2)$
3. $\text{Var}(\hat{\mu}_{U,st,R,s}) \leq \text{Var}(\bar{X}_{st})$
4. $\text{Var}(\hat{\mu}_{U,st,R,s}) \leq \text{Var}(\hat{\mu}_{U,st,s})$

Εκτιμητές παλινδρόμησης των μ_U, τ_U σε ΣΤΔ

Εκτίμηση του τ_U : $\tau_U = N\mu_U$

Ο χωριστός εκτιμητής παλινδρόμησης του τ_U δίνεται από

$$\hat{\tau}_{U,st,R,s} = N\hat{\mu}_{U,st,R,s}$$

Ιδιότητες του $\hat{\tau}_{U,st,R,s}$:

Προκύπτουν άμεσα από τη σχέση $\hat{\tau}_{U,st,R,s} = N\hat{\mu}_{U,st,R,s}$ και τις ιδιότητες του $\hat{\mu}_{U,st,R,s}$.

Εκτιμητές παλινδρόμησης των μ_U, τ_U σε ΣΤΔ

Β. Συνδυασμένοι εκτιμητές παλινδρόμησης:

Εκτίμηση του μ_U : (απαιτεί γνώση του μ_V)

Γενική μορφή συνδυασμένου εκτιμητή παλινδρόμησης του μ_U

$$\hat{\mu}_{U,st,R,c} = \bar{X}_{st} + b (\mu_V - \bar{Y}_{st}), \quad \text{όπου } b \text{ σταθερά}$$

ο οποίος είναι ΑΕ του μ_U για κάθε τιμή της σταθεράς b .

Πρόταση 4

Η διασπορά του $\hat{\mu}_{U,st,R,c}$, $\text{Var}(\hat{\mu}_{U,st,R,c})$, ελαχιστοποιείται ως προς b για

$$b_0 = \frac{\sum_{\ell} W_{\ell}^2 (1 - f_{\ell}) \frac{1}{n_{\ell}} \sigma_{0UV\ell}}{\sum_{\ell} W_{\ell}^2 (1 - f_{\ell}) \frac{1}{n_{\ell}} \sigma_{0V\ell}^2}.$$

Απόδειξη.

Εκτιμητές παλινδρόμησης των μ_U, τ_U σε ΣΤΔ

Παρατήρηση:

Επειδή το b_0 εξαρτάται από άγνωστες παραμέτρους, χρησιμοποιούμε τον εκτιμητή του

$$\hat{b}_0 = \frac{\sum_{\ell} W_{\ell}^2 (1 - f_{\ell}) \frac{1}{n_{\ell}} \hat{\sigma}_{0UV\ell}}{\sum_{\ell} W_{\ell}^2 (1 - f_{\ell}) \frac{1}{n_{\ell}} \hat{\sigma}_{0V\ell}^2},$$

όπου

$$\hat{\sigma}_{0UV\ell} = \frac{1}{n_{\ell}-1} \sum_{i=1}^{n_{\ell}} (X_{\ell i} - \bar{X}_{\ell})(Y_{\ell i} - \bar{Y}_{\ell})$$

$$\hat{\sigma}_{0V\ell}^2 = \frac{1}{n_{\ell}-1} \sum_{i=1}^{n_{\ell}} (Y_{\ell i} - \bar{Y}_{\ell})^2$$

Ο συνδυασμένος εκτιμητής παλινδρόμησης του μ_U δίνεται από

$$\hat{\mu}_{U,st,R,c} = \bar{X}_{st} + \hat{b}_0 (\mu_V - \bar{Y}_{st})$$

Εκτιμητές παλινδρόμησης των μ_U, τ_U σε ΣΤΔ

Ιδιότητες του $\hat{\mu}_{U,st,R,c}$:

Για $n_\ell \rightarrow N_\ell, \forall \ell$, είναι

1. $E\hat{\mu}_{U,st,R,c} \approx \mu_U$ (δηλ. προσεγγιστικά ΑΕ)
2. $\text{Var}(\hat{\mu}_{U,st,R,c}) \leq \text{Var}(\bar{X}_{st})$
3. $\text{Var}(\hat{\mu}_{U,st,R,c}) \leq \text{Var}(\hat{\mu}_{U,st,c})$

Εκτίμηση του τ_U : $\tau_U = N\mu_U$

Ο συνδυασμένος εκτιμητής παλινδρόμησης του τ_U δίνεται από

$$\hat{\tau}_{U,st,R,c} = N\hat{\mu}_{U,st,R,c}$$

Ιδιότητες του $\hat{\tau}_{U,st,R,c}$:

Προκύπτουν άμεσα από τη σχέση $\hat{\tau}_{U,st,R,c} = N\hat{\mu}_{U,st,R,c}$ και τις ιδιότητες του $\hat{\mu}_{U,st,R,c}$.