Διάλεξη 5η 5/11/2019

Παλινδρόμηση και Ανάλυση Διακύμανσης

▶ ΑΣΚΗΣΕΙΣ που δώθηκαν στο μάθημα (φυλλάδιο)

ΑΣΚΗΣΗ 1

α.γ.π $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad , i=1,..,n$

με $\varepsilon_i \sim N(0,\sigma^2)$ και ασυσχ. ανά δύο.

i) ΕΕΤ(β_0, β_1) συμπίπτουν με τους ΕΜΠ. των β_0, β_1

ii) Να βρεθεί ο εκτιμητής μέγιστης πιθανοφάνειας (ΕΜΠ) της σ^2.

ΥΠΕΝΘΥΜΗΣΗ ΕΜΠ

Έστω $\underbrace{\text{Τυχαίο δείγμα}}_{}$ (n ανεξ. τ.μ με την ίδια κατανομή $f(w,\theta)$) $W_1, W_2, .., W_n$ από πληθυσμό με κατανομή $f(w,\theta)$ θ: άγνωστη παράμετρο. ||πχ|| $f(w,\theta) = \theta \cdot e^{-w\theta}$ $\theta>0$, $w>0$ Εκθετική.

Πρόβλημα Η εκτίμηση (προσέγγιση) της άγνωστης παραμέτρου θ.

↑

Μέθοδος μέγιστης Πιθανοφάνειας ⟶ δημιουργεί τους ΕΜΠ για την εκτίμηση της θ.

Πως υπολογίζω τον ΕΜΠ της θ.

Πιθανοφάνεια → η από κοινού κατανομή του τ.δ

=

Γινόμενο περιθωρίων κατανομών $\prod_{i=1}^{n} f_{W_i}(w_i;\theta)$ επειδή τα W_i, $i=1,..,n$ ανεξάρτητα.

| συμβολισμός

↓

$L(\theta)$ = από κοινού κατανομή των $W_1,..,W_n$ $= \prod_{i=1}^{n} f_{W_i}(w_i,\theta)$.

Ο ΕΜΠ είναι εκείνη η τιμή του θ. που μεγιστοποιεί την $L(\theta)$.

ή εκείνη η τιμή του θ που μεγιστοποιεί τον $\log L(\theta)$.

{επειδή log αύξουσα, όπου μεγιστοποιείται η $\log L(\theta)$, μεγιστοπ. και η $L(\theta)$}

Έστω $W_1,..,W_n$ από $f(w,\theta) = \theta \cdot e^{-w\theta}$

$$L(\theta) = \prod_{i=1}^{n} f(w_i;\theta) = \prod_{i=1}^{n} \theta \cdot e^{-\theta w_i} = \theta^n e^{-\theta \Sigma w_i}$$

ο ΕΜΠ της θ. μεγιστοποιεί την $L(\theta)$ ή $\log L(\theta) = n \log\theta - \theta \Sigma w_i$

$$\frac{\partial}{\partial\theta} \log L(\theta) = 0 \Rightarrow \frac{n}{\theta} - \Sigma w_i = 0 \Rightarrow \theta = \frac{1}{\frac{1}{n}\Sigma w_i} = \frac{1}{\bar{w}}$$

Άρα ο ΕΜΠ της θ είναι $\boxed{\hat{\theta} = \frac{1}{\bar{W}}}$.

ΛΥΣΗ

i) Οι ΕΕΤ προκύπταν από ελαχιστοποίηση ως προς β_0 και β_1.

$$S = \sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_i)^2 \quad (1)$$

ΕΜΠ των β_0, β_1

Από $\varepsilon_i \sim N(0, \sigma^2)$ και ε_i ασυσχ. ανά δύο προκύπτει ότι

$Y_i \sim N(\beta_0 + \beta_1 X_i, \sigma^2)$ και είναι ασυσχ ανά δύο

ή ανεξ αφού το ασυσχέτιστο στην κανονική σημαίνει ανεξαρτησία.

$$L(\beta_0, \beta_1) = \prod_{i=1}^{n} f_{Y_i}(y_i) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(y_i - \beta_0 - \beta_1 X_i)^2}$$

$$= \frac{1}{\sigma^2\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}\Sigma(y_i - \beta_0 - \beta_1 X_i)^2}$$

Για να βρώ ΕΜΠ των β_0, β_1 αρκεί να μεγιστ την $L(\beta_0, \beta_1)$

ή τον $\log L(\beta_0, \beta_1) = -n\log\sigma - \log\sqrt{2\pi} - \frac{1}{2\sigma^2}\underline{\Sigma(y_i - \beta_0 - \beta_1 X_i)^2}$

ή να μεγιστοποιήσω ως προς β_0, β_1

την $-\Sigma(y_i - \beta_0 - \beta_1 X_i)^2$

ή να ελαχιστοποιήσω ως προς β_0, β_1

την $\Sigma(y_i - \beta_0 - \beta_1 X_i)^2 \quad (2)$

Παρατηρώ ότι οι ΕΕΤ και οι ΕΜΠ προκύπτουν από ελαχιστοποίηση της ίδιας ποσότητας. Άρα ταυτίζονται.

Σπίτι

ii) Υπόδειξη

ΕΜΠ της σ^2

Φτάσαμε (πριν) $\log L(\beta_0, \beta_1, \sigma^2) = -\frac{n}{2}\log\sigma^2 - \log\sqrt{2\pi} - \frac{1}{2\sigma^2}\Sigma(y_i - \beta_0 - \beta_1 X_i)^2$

Ο ΕΜΠ της σ^2 θα προκύψει από την μεγιστοποίηση ως προς σ^2 της $\log(\beta_0, \beta_1, \sigma^2)$

δηλ από επίλυση ως προς σ^2, την $\frac{\partial}{\partial\sigma^2}\log L(\hat{\beta_0}, \hat{\beta_1}, \sigma^2) = 0$

....

Κ.Ο.Κ $\hat{\sigma}^2 = \frac{SSres}{n}$

ΘΕΩΡΙΑ

Μοντέλο πγπ :

$$\underset{\sim}{Y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon} \iff Y_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip} + \varepsilon_i, \quad i = 1, 2, \dots, n$$

ΕΕΤ: $\hat{\underset{\sim}{\beta}} = (X'X)^{-1} X' \underset{\sim}{Y}$

Εκτιμώμενο Μοντέλο πγπ : $\hat{\underset{\sim}{Y}} = X\hat{\underset{\sim}{\beta}}$

Υπόλοιπα : $\underset{\sim}{e} = \underset{\sim}{Y} - \hat{\underset{\sim}{Y}}$

Αποδείξαμε: $SS_{tot} = SS_{reg} + SS_{res}$

Αν $SS_{reg} >> SS_{res} \rightsquigarrow$ ΥΠΟΣΧΟΜΕΝΟ ΜΟΝΤΕΛΟ.

Ιδιότητες για $\hat{\underset{\sim}{\beta}}$.

Οι οποίες βασίζονται στις υποθέσεις για τα σφάλματα.

1) $E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$ — $E(\underset{\sim}{Y}) = X\underset{\sim}{\beta}$
2) $Var(\underset{\sim}{\varepsilon}) = \sigma^2 I_n \longrightarrow$ $Var(\underset{\sim}{Y}) = \sigma^2 I_n$
3) $\underset{\sim}{\varepsilon} \sim N_n(\underset{\sim}{0}, \sigma^2 I_n)$ — $\underset{\sim}{Y} \sim N_n(X\underset{\sim}{\beta}, \sigma^2 I_n)$

Αποδείξαμε κάποιες πρώτες ιδιότητες για $\hat{\underset{\sim}{\beta}}$.

Περαιτέρω Ιδιότητες των $\hat{\underset{\sim}{\beta}}$.

υπό την προϋπόθεση ότι ικανοποιούνται οι υποθέσεις για τα σφάλματα $\underset{\sim}{\varepsilon}$.

Ιδιότητα: Οι $\hat{\underset{\sim}{\beta}} \sim N_{p+1}(\underset{\sim}{\beta}, \sigma^2 (X'X)^{-1})$. (*)

Από ΘΠΣ (το έχω δει)

Πολυδιάστατη Κανονική Κατανομή.

Το τ.διάνυσμα $\underset{\sim}{W} = \begin{pmatrix} W_1 \\ \vdots \\ W_n \end{pmatrix}$ λέμε ότι ακολουθεί την n-διάστατη κανονική κατανομή με παραμέτρους $\underset{\sim}{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}$ και πίνακα Σ αν η πυκνότητα του $\underset{\sim}{W}$ είναι:

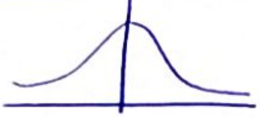
$$f_{\underset{\sim}{W}}(\underset{\sim}{w}, \mu, \Sigma) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \cdot e^{-\frac{1}{2}(\underset{\sim}{W} - \underset{\sim}{\mu})' \Sigma^{-1} (\underset{\sim}{W} - \underset{\sim}{\mu})}$$

και συμβολίζεται $\underset{\sim}{W} \sim N_n(\underset{\sim}{\mu}, \Sigma)$.

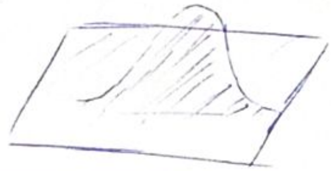
<u>ΠΑΡΑΤΗΡΗΣΗ</u>: Αναλογία με Κανονική

$$\frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

Μονοδιάστατη



Δισδιάστατη



Πιο βασικές ιδιότητες της $N_n(\underline{\mu}, \Sigma)$

① $\underline{\mu} = E(\underline{W})$

② $\Sigma = Var(\underline{W}) = \begin{bmatrix} Var(W_1) & & Cov(W_i,W_j) \\ & \ddots & \\ Cov(W_i,W_j) & & Var(W_n) \end{bmatrix}$

Είναι nxn συμμετρικός και θετικά ημιορισμένος.

③ Αν το $\underset{\sim}{W} \simeq \begin{pmatrix} W_1 \\ \vdots \\ W_n \end{pmatrix} \sim N_n\left(\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_1^2 & \cdots & Cov(W_i,W_j) \\ Cov(W_i,W_j) & \cdots & \sigma_n^2 \end{pmatrix} \right)$

τότε κάθε $W_i \sim N(\mu_i, \sigma_i^2 = Var(W_i))$ $i=1,\dots,n$

Το αντίστροφο δεν ισχύει, παρά μόνο αν οι W_i $i=1,\dots,n$ είναι ανεξάρτητες.

④ Αναλλοίωτο της $N_n(\underline{\mu}, \Sigma)$

α) Αν $\underset{\sim}{W} \sim N_n(\underline{\mu}, \Sigma)$ και A ένας pxn πίνακας

τότε $A\underset{\sim}{W} \sim N_p(A\underline{\mu}, A\Sigma A')$

β) Αν $\underset{\sim}{W} \sim N_n(\underline{\mu}, \Sigma)$ τότε $\underset{\sim}{\alpha}'\underset{\sim}{W} \sim N(\underset{\sim}{\alpha}'\underline{\mu}, \underset{\sim}{\alpha}'\Sigma\underset{\sim}{\alpha})$

⊛ <u>Αποδ</u>

Είναι $\hat{\underline{\beta}} = \underbrace{(X'X)^{-1}X'}_{(p+1)xn \text{ πίνακας}}Y$. Επειδή $\underline{Y} \sim N_n$ από την ιδιότητα 4α)

Το $\hat{\underline{\beta}} \sim N_{p+1}$

με $E(\hat{\underline{\beta}}) = \underline{\beta}$ και $Var(\hat{\underline{\beta}}) = \sigma^2 (X'X)^{-1}$ (βλ. προηγούμενο μάθημα).

Ιδιότητα: Οι ΕΕΤ $\hat{\beta}$ έχουν τη μικρότερη διακύμανση μεταξύ όλων των άλλων αμερόληπτων εκτιμητών των παραμέτρων β που είναι γραμμικές συνισεις των $Y_1, .., Y_n$ (ΘΕΩΡΗΜΑ Gass-Markov)

$\boxed{\hat{\beta} \text{ είναι ΑΟΕΔ}}$

Ιδιότητα: Υπό τις υποθέσεις για τα σφάλματα, τότε οι ΕΕΤ των β ταυτίζονται με τους ΕΜΠ των β

Περαιτέρω, ο ΕΜΠ της σ^2 είναι: $\frac{SSres}{n}$

Αποδ. Οι ΕΕΤ της β προκύπτουν από ελαχιστοποίηση ως προς β της $S = \sum_{i=1}^{n} \varepsilon_i^2 = \underset{\sim}{\varepsilon}' \underset{\sim}{\varepsilon} \overset{\pi\chi\pi}{=} (\underline{Y} - X\underline{\beta})'(Y - X\underline{\beta}) \quad (1)$

Εύρεση ΕΜΠ

Πιθανοφάνεια $= L(\beta)$ = Από κοινού κατανομή των $Y_1, .., Y_n$

$\underset{\text{σφαλμάτων}}{\overset{\text{υποθεση}}{=}} N_n(X\underline{\beta}, \sigma^2 I_n)$

(3)

$$L(\underline{\beta}) = \frac{1}{(2\pi)^{n/2} |\sigma^2 I_n|^{1/2}} e^{-\frac{1}{2}(\underline{y} - X\underline{\beta})'(\sigma^2 I_n)^{-1}(\underline{y} - X\underline{\beta})}$$

Ιδιότητα οριζουσών.

$$|\alpha A| = \alpha^n |A|$$

$$\frac{1}{(2\pi)^{n/2}[(\sigma^2)^n]^{1/2} |I_n|} e^{-\frac{1}{2}(y - X\underline{\beta})' \frac{1}{\sigma^2} I_n^{-1} (y - X\underline{\beta})}$$

$$= \frac{1}{(2\pi)^{n/2} \sigma^n} \cdot e^{-\frac{1}{2\sigma^2}(y - X\underline{\beta})'(y - X\underline{\beta})}$$

Οι ΕΜΠ των $\underline{\beta}$ προκύπτουν από μεγιστοποίηση της $L(\underline{\beta})$ ή της $\log L(\underline{\beta})$ ως προς $\underline{\beta}$. Δηλ από μεγιστ. ως προς $\underline{\beta}$ της $-\frac{n}{2}\log(2\pi) - n\log\sigma - \frac{1}{2\sigma^2}(y - X\underline{\beta})'(y - X\underline{\beta})$

ή από μεγιστ. ως προς $\underline{\beta}$ του $-(y - X\underline{\beta})'(y - X\underline{\beta})$

ή από ελαχιστ ως προς $\underline{\beta}$ του $(y - X\underline{\beta})'(y - X\underline{\beta}) \quad (2)$

Άρα από (1) και (2) $ΕΕΤ \equiv ΕΜΠ$.

ΕΜΠ της σ^2 Υπόδειξη: Ακολουθείται η ίδια πορεία με την αχπ (Ασκ 1)

ΠΑΡΑΤΗΡΗΣΗ (Ερμηνεία του εκτιμητή $\hat{\beta}_i$ $i=1,..,p$)

Ο ΕΕΤ $\hat{\beta}_i$ $i=1,2,..,p$ εκφράζει την μεταβολή της Y σε μοναδιαία μεταβολή της X_i για $i=1,..,p$ όταν οι άλλες ανεξάρτητες μεταβλητές παραμένουν σταθερές.

Ιδιότητα: $\frac{SSres}{\sigma^2} \sim X^2_{n-p-1}$

(Αποδ ξεφεύγει λιγάκι)

Ιδιότητα: Οι $\hat{\underline{\beta}}$ και MSres ανεξάρτητες τυχαίες ποσότητες

• t-ΤΕΣΤ για έλεγχο $H_0: \beta_i = \beta_i^*$, $i=1,..,p$

Στο μοντέλο της πγπ για τον έλεγχο της $H_0: \beta_i = \beta_i^*$ έναντι $H_a: \beta_i \neq \beta_i^*$ (β_i^* γνωστό) χρησιμοποιείται η στατιστική συν/ση:

$$t_i = \frac{\hat{\beta}_i - \beta_i^*}{\sqrt{MSres}\sqrt{C_{i+1,i+1}}}$$ με κατανομή t_{n-p-1} υπό H_0.

Και κ.π $|t_i| \geq t_{n-p-1, a/2}$ όπου $C_{i+1,i+1}$:
Το $(i+1)$-διαγώνιο στοιχείο του πίνακα $(X'X)^{-1}$ για $i=0,..,p$.

Πρακτική αξία του τεστ: ⓐ Δίνει την δυνατότητα να ελέγξω ότι η μεταβολή της Y για μοναδιαία μεταβολή της X_i έχει συγκεκριμένη τιμή.

ⓑ Αν $\beta_i^*=0$ δίνει τη δυνατότητα ελέγχου αν οι Y εξαρτώνται από την X_i.

$Y_i = \beta_0 + \beta_1 X_{i1} + .. + \beta_p X_{ip}$
Αν δεχτώ την $H_0: \beta_1 = 0$
ο όρος $\beta_1 X_{i1}$ εξαφανίζεται

Κατασκευή τεστ: Στηριζόμαστε στο εκτιμητή $\hat{\beta}_i$, της παραμέτρου β_i που εμφανίζεται στην Ho

Επειδή $\hat{\underline{\beta}} \sim N_{p+1}(\underline{\beta}, \sigma^2 (X'X)^{-1})$ και επειδή από ιδιότητα 3) της πολυδ. κανονικής

$$\hat{\beta}_i \sim N(\beta_i, \sigma^2 C_{i+1,i+1})$$

προφανώς $\dfrac{\hat{\beta}_i - \beta_i}{\sigma\sqrt{C_{i+1,i+1}}} \sim N(0,1)$ και έτσι $\dfrac{\hat{\beta}_i - \beta_i^*}{\sigma\sqrt{C_{i+1,i+1}}} \sim N(0,1)$ υπό την $H_0: \beta_i = \beta_i^*$

Αλλά $\dfrac{SS_{res}}{\sigma^2} \sim X^2_{n-p-1}$

Επομένως θεωρώ

$$\frac{\dfrac{\hat{\beta}_i - \beta_i^*}{\sigma\sqrt{C_{i+1,i+1}}}}{\sqrt{\dfrac{SS_{res}}{(n-p-1)\sigma^2}}} \underset{\text{αλγεβρικά}}{=} \frac{\hat{\beta}_i - \beta_i^*}{\sqrt{C_{i+1,i+1}}\sqrt{MS_{res}}} = t_i$$

Από άποψη κατανομής

υπό Ho $\dfrac{N(0,1)}{\sqrt{\dfrac{X^2_{n-p-1}}{n-p-1}}} \underset{MS_{res}}{\overset{\hat{\beta}_i \text{ ανεξ}}{=}} t_{n-p-1}$ υπό Ho

Κατασκευή κ.π

Μορφή της κ.π: Μεγάλες τιμές του t_i, $|t_i| \geq c$.

Υπολογισμός κρισ. σημείου c: **ΠΑΝΤΑ** από $\alpha = P(\text{Απορ } H_0 / H_0 \text{ αληθ}) = P(|t_i| \geq c / t_i \sim t_{n-p-1})$

$= P(|t_{n-p-1}| \geq c) = P(t_{n-p-1} \geq c \text{ ή } t_{n-p-1} \leq -c)$

συμ t_{n-p-1} $\quad 2P(t_{n-p-1} \geq c) \Rightarrow P(t_{n-p-1} \geq c) = \dfrac{\alpha}{2}$

$$\Downarrow$$

$$c = t_{n-p-1, \alpha/2}$$

F-ΤΕΣΤ για την ύπαρξη του μοντέλου πγπ

Για τον έλεγχο της $H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$

χρησιμοποιούμε την σ.σ: $F = \dfrac{MS_{reg}}{MS_{res}}$ με κατανομή $F_{p, n-p-1}$ υπό H_0

και Κ.Π $F \geq F_{p, n-p-1, \alpha}$.