Προσθήκη στο κεφάλαιο 20: Τετραγωνικές μορφές σε συμμετρικούς πίνακες

VAR1.

Πίνακες διακύμανσης (-συνδιακύμανσης)

* Έστω $z_1, \dots, z_k$ τυχαίες μεταβλητές. Ορίζω το $z = \begin{pmatrix} z_1 \\ \vdots \\ z_k \end{pmatrix} \in \mathbb{R}^k$, ένα διάνυσμα των k τυχαίων μεταβλητών $z_1, \dots, z_k$, από όποιο αυτές είναι οι συντεταγμένες.

* Αν $\mu_i := E(z_i)$ θέτουμε και $\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_k \end{pmatrix}$ και γράφουμε

$$E(z) := \begin{pmatrix} E(z_1) \\ \vdots \\ E(z_k) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_k \end{pmatrix} = \mu.$$

, Δηλαδή η E του τυχαίου διανύσματος ορίζεται να είναι το διάνυσμα των αναμενόμενων τιμών των z_i : $\mu_i = Ez_i$.

* Ορίζουμε επίσης τον $k \times k$ πίνακα διακύμανσης (-συνδιακύμανσης) της z, την $Var(z)$, που τη συμβολίζουμε και με Σ_z.

$$\Sigma_z = Var(z) := E\left[(z-\mu)(z-\mu)^T\right] =$$

$$= \begin{pmatrix} E(z_1-\mu_1)(z_1-\mu_1)^T & E(z_1-\mu_1)(z_2-\mu_2)^T & \dots \\ E(z_2-\mu_1)(z_1-\mu_1)^T & E(z_2-\mu_2)(z_2-\mu_2)^T & \dots \\ \vdots & \vdots & \end{pmatrix} = \begin{pmatrix} var(z_1) & Cov(z_1, z_2) & \\ Cov(z_1, z_2) & var(z_2) & \\ \vdots & & \ddots \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \\ \sigma_{k1} & & \ddots \quad \sigma_k^2 \end{pmatrix}$$

, όπου $\sigma_{ij} = Cov(z_i, z_j)$ και $\sigma_i^2 = Var(z_i)$.

VAR.2

* Πρόταση: Αν A $r\times k$ σταθερός πίνακας (δε συμπεριλαμβάνει τυχαίες μεταβλητές) και αν θέσουμε $Y := AZ$ τότε

SOS $\boxed{E(Y) = A\mu}$ και $\boxed{Var(Y) = \Sigma_Y = A\Sigma_Z A^T}$.

— Απόδ.: $E(Y) = E(A\cdot Z) = A \cdot E(Z) = A\cdot\mu$

$$Var(Y) = Var(AZ) = E\left[(AZ - E(AZ))\cdot(AZ - E(AZ))^T\right]$$

$$= E\left[(A\cdot(Z-\mu))(A\cdot(Z-\mu))^T\right] = E\left[A\,(Z-\mu)(Z-\mu)^T A^T\right]$$

$$= A\cdot E\left[(Z-\mu)(Z-\mu)^T\right]\cdot A^T = A\cdot\Sigma_Z\cdot A^T$$

Εφαρμογές – Παραδείγματα

* $Y = a_1 Z_1 + \dots + a_k Z_k$. Υπολογισμός $Var(Y)$:

$Y = a^T Z$, όπου $a = \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix}$. Τότε η παραπάνω πρόταση με $A = a^T$ δίνει: $\boxed{Var(Y) = a^T \Sigma_Z a}$.

---

* αν Z_i ανεξάρτητες & ισόνομες, $Var(Z_i) = \sigma^2$, $E(Z_i) = 0$

τότε αν $A_{r\times k}$ $Var(AZ) = \sigma^2 A\cdot I\cdot A^T = \sigma^2 A^T A$.

καθώς $\Sigma_Z = \sigma^2\cdot I = \begin{pmatrix} \sigma^2 & & 0 \\ & \ddots & \\ 0 & & \sigma^2 \end{pmatrix}$.

(Εκτός διαγωνίου $\sigma_{ij} = 0$ καθώς Z_i, Z_j ανεξάρτητα).

* Έχω $Z_1, \ldots, Z_k$ ανεξάρτητα $EZ_i = 0$, $\mathrm{var}(Z_i) = 1$. VAR3

Για δοσμένο Σ (συμμετρικό & με ιδιοτιμές ≥ 0) θέλω να βρώ A τέτοιο ώστε τα $Y = A \cdot Z$ να έχουν $\Sigma_Y = \mathrm{var}(Y) \overset{!}{=} \Sigma$!

(π.χ για να προσομοιώσω τυχαίες μεταβλητές με δοσμένο πίνακα διακύμανσης Σ).

Αν $\Sigma = Q \Lambda Q^T$ η φασματική διάσπαση του Σ

θέτω $A = \Sigma^{1/2} = Q \underbrace{\begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_k} \end{pmatrix}}_{\sqrt{\Lambda}} Q^T$. Τότε η πρόταση δίνει

$$\mathrm{var}(Y) = A \cdot \Sigma_Z \cdot A^T = A \cdot I \cdot A^T = A \cdot A^T = A \cdot A$$

$$= \left(Q \sqrt{\Lambda} Q^T\right) \cdot Q \sqrt{\Lambda} \left(Q^T\right) = Q \Lambda Q^T = \Sigma. \text{ ok}$$

Διοδιάστατη (και πολυδιάστατη) κανονική κατανομή.

Έστω $X_1 \sim N(\mu_1, 1)$ & $X_2 \sim N(\mu_2, 1)$ ανεξάρτητες.

Αν $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, τότε η από κοινού πυκνότητα της X, $f_X(x)$, (για $x \in \mathbb{R}^2$) δίνεται ως γινόμενο των πυκνοτήτων των X_1 & X_2: $f_{X_1}(x_1)$ και $f_{X_2}(x_2)$:

$$f_X(x) = f_X(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2)$$

$$= \frac{1}{2\pi} \cdot \exp\left(-\frac{1}{2}(x_1 - \mu_1)^2\right) \cdot \exp\left(-\frac{1}{2}(x_2 - \mu_2)^2\right)$$

$$= \frac{1}{2\pi} \exp\left(-\frac{1}{2}\left\{(x_1 - \mu_1)^2 + (x_2 - \mu_2)^2\right\}\right)$$

$$= \frac{1}{2\pi} \exp\left(-\frac{1}{2} \|x - \mu_X\|^2\right) \quad \text{όπου } x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \ \mu_X = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}.$$

Αν θέλω να μετασχηματίσω γραμμικά την $X=\begin{pmatrix}X_1\\X_2\end{pmatrix}$ | VAR4

έτσι ώστε να έχει ένα δοσμένο πίνακα διακύμανσης Σ (με αυστηρά θετικές ιδιοτιμές), όπως είδαμε στο προηγούμενο παράδειγμα πρέπει να πάρω $\boxed{Y=\Sigma^{1/2}X}$

Ποια θα είναι η πυκνότητα της Y?

Πρόταση: Αν $X\in\mathbb{R}^k$ έχει πυκνότητα $f_X(x)$, $x\in\mathbb{R}^k$, δηλ.

$$P(X\in A)=\int_A f_X(x)\,dx$$

και T μη-ιδιόμορφος $k\times k$ πίνακας. Τότε η $Y:=TX$ έχει πυκνότητα

$$f_Y(y)=f_X(T^{-1}y)\cdot\frac{1}{\det(T)},\ \text{δηλ}$$

$$P(Y\in B)=\int_B f_X(T^{-1}y)\frac{1}{\det(T)}\,dy,\quad y\in\mathbb{R}^k.$$

Εφαρμογή στην $Y=\Sigma^{1/2}X\in\mathbb{R}^2$.

$$f_Y(y)=\frac{1}{2\pi}\frac{1}{\det(\Sigma^{1/2})}\cdot\exp\left(-\frac{1}{2}\left\|\Sigma^{-1/2}y-\mu_X\right\|^2\right)$$

$$=\frac{1}{2\pi}\frac{1}{\sqrt{\det(\Sigma)}}\cdot\exp\left(-\frac{1}{2}\left\|\Sigma^{-1/2}y-\Sigma^{-1/2}\mu_Y\right\|^2\right)\quad \text{με } \mu_Y=\Sigma^{1/2}\mu_X$$

$$=\frac{1}{2\pi}\frac{1}{\sqrt{\det\Sigma}}\exp\left(-\frac{1}{2}(y-\mu_y)^T\Sigma^{-1/2\,T}\Sigma^{-1/2}(y-\mu_y)\right)$$

$$=\frac{1}{2\pi}\frac{1}{\sqrt{\det\Sigma}}\exp\left(-\frac{1}{2}(y-\mu_y)\Sigma^{-1}(y-\mu_y)\right).$$

Θα δείξε ότι Y έχει μία δισδιάστατη κανονική κατανομή με μέσο μ_Y και πίνακα διακύμανσης Σ_Y και θα συμβολίζουμε $Y \sim N(\mu_Y, \Sigma_Y)$ αν το Y προκύπτει

VAR5

ως $\boxed{Y = \Sigma_Y^{1/2} \cdot X}$ όπου $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ και $X_i \sim N(\mu_i, 1)$

ανεξάρτητες κανονικές κατανομές.

Τότε η Y θα έχει πυκνότητα

$$\boxed{f_Y(y) = \frac{1}{2\pi} \frac{1}{\sqrt{\det \Sigma}} \cdot \exp\left(-\frac{1}{2}(y-\mu_Y)^T \Sigma_Y^{-1} (y-\mu_Y)\right)}.$$

Άρα για να περιγράψω τις ισοϋψείς της $f_Y(y)$, δηλ. τον γεωμετρικό τόπο των σημείων $y \in \mathbb{R}^2$ με $f_Y(y) = c$ πρέπει να βρω τις ισοϋψείς της $(y-\mu_Y)^T \Sigma^{-1} (y-\mu_Y)$!

Μιά ακόμα εφαρμογή: ελλειψοειδή εμπιστοσύνης

VAR6

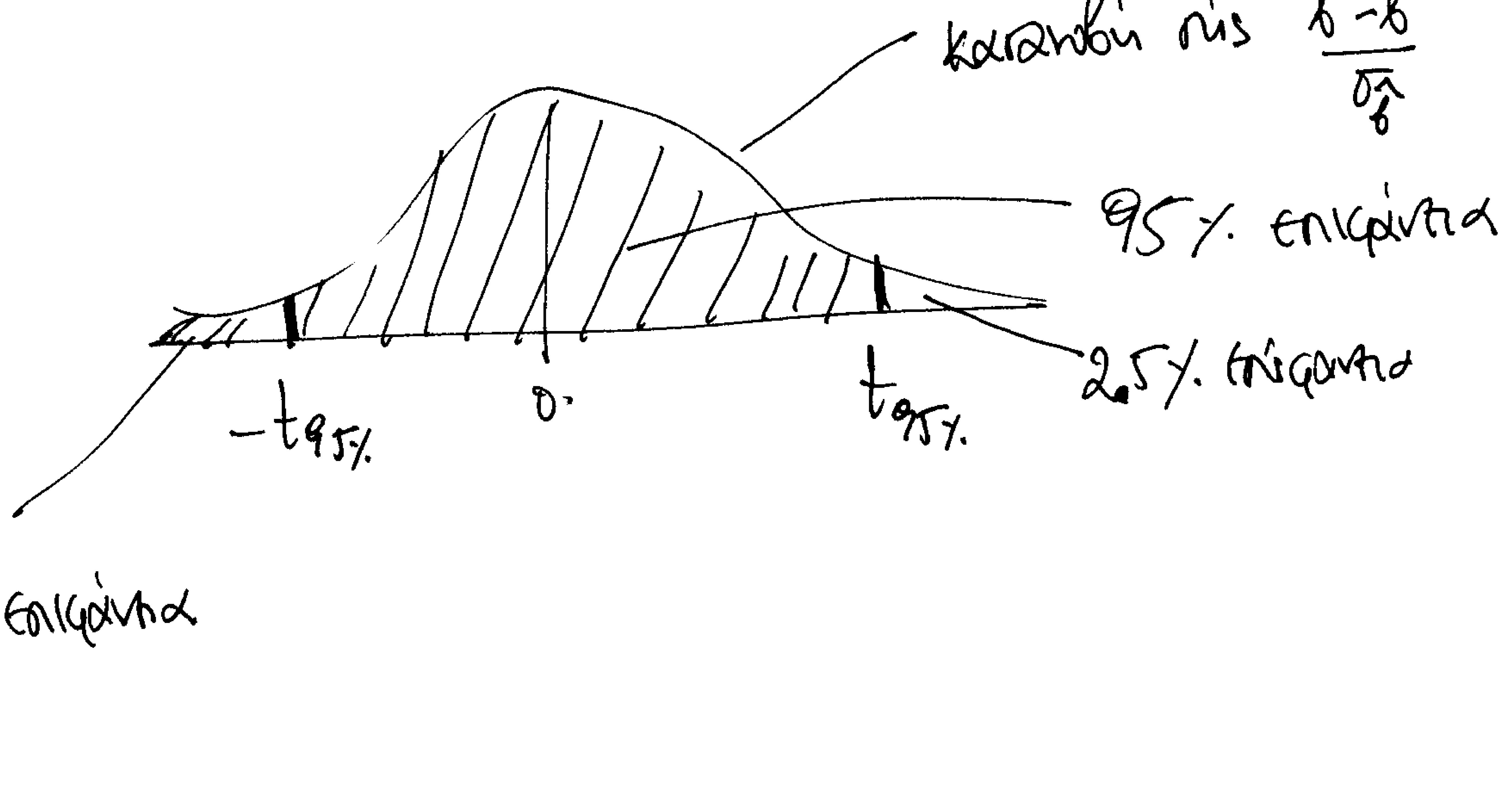
Εισαγωγική παρένθεση: διαστήματα εμπιστοσύνης.

— Έστω ότι εκτιμώ (στη βάση κάποιων παρατηρήσεων) μια παράμετρο β (π.χ τον μέσο μ) με μία εκτιμήτρια $\hat{\beta}$ (π.χ τον δειγματικό μέσο $\hat{\mu} = \overline{x}$).

— Ενδιαφέρομαι να κατασκευάσω "διαστήματα εμπιστοσύνης" (γύρω από το $\hat{\beta}$) για το β: διαστήματα μέσα στα οποία θα βρίσκεται με μεγάλη πιθανότητα η άγνωστη τιμή του β.

— στην Εκτιμητική θα γνωρίσετε περισσότερες λεπτομέρειες, αλλά μια τέτοια κατασκευή περνά από την μελέτη της κατανομής της τυχαίας μεταβλητής $(\hat{\beta}-\beta)/\sigma_{\hat{\beta}}$, όπου εδώ $\sigma_{\hat{\beta}}$ η τυπική απόκλιση του $\hat{\beta}$. Ας υποθέσουμε, ότι αυτή η κατανομή είναι γνωστή, και ότι επομένως μπορούμε να βρούμε μια "κρίσιμη τιμή" $t_{95\%}$, τέτοια ώστε

$$P\left(-t_{95\%} \leq \frac{\hat{\beta}-\beta}{\sigma_{\hat{\beta}}} \leq t_{95\%}\right) = 95\%$$



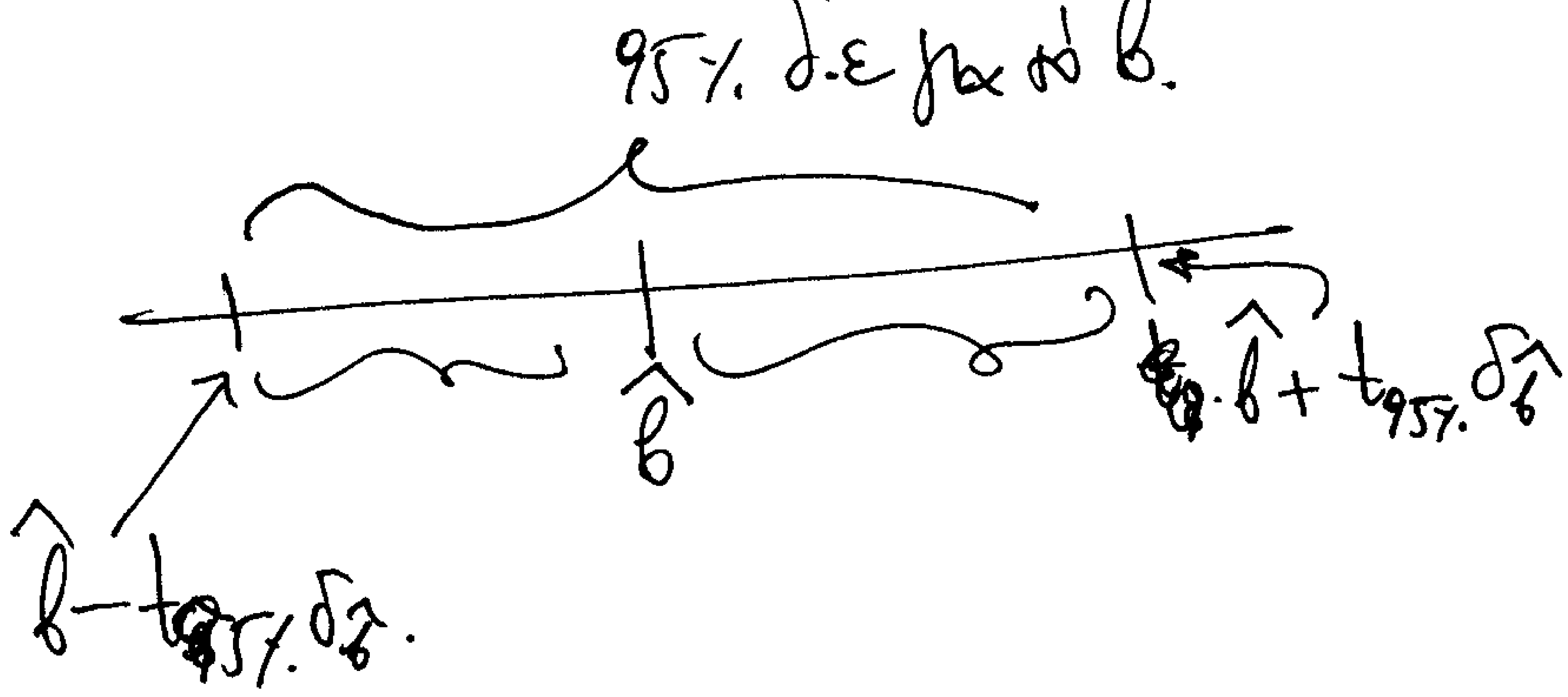
Τότε τα ~~διασ~~ διαστήματα εμπιστοσύνης κατασκευάζονται ως :

$$P\left(-t_{95\%}\,\hat{\sigma}_{\hat{b}} \leq \hat{b} - b \leq t_{95\%}\,\hat{\sigma}_{\hat{b}}\right) = 95\%$$

$$\Rightarrow P\left(-\hat{b} - t_{95\%}\,\hat{\sigma}_{\hat{b}} \leq -b \leq -\hat{b} + t_{95\%}\,\hat{\sigma}_{\hat{b}}\right) = 95\%$$

$$\Rightarrow P\left(\hat{b} - t_{95\%}\,\hat{\sigma}_{\hat{b}} \leq b \leq \hat{b} + t_{95\%}\,\hat{\sigma}_{\hat{b}}\right) = 95\%$$

- Άρα το 95% δ.ε για το b είναι : $\hat{b} \pm t_{95\%}\,\hat{\sigma}_{\hat{b}}$



- Τα ελλειψοειδή εμπιστοσύνης είναι μια γενίκευση των διαστημάτων εμπιστοσύνης όταν ενδιαφερόμαστε για δύο (ή περισσότερες) παραμέτρους b_1, b_2 ταυτόχρονα. Θέτω $b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$

~~και~~ - Έστω ότι έχω εκτιμήτριες $\hat{b}_1$ και $\hat{b}_2$ και ότι θέτω $\hat{b} = \begin{pmatrix} \hat{b}_1 \\ \hat{b}_2 \end{pmatrix}$. Ψάχνω μια περιοχή $\subset \mathbb{R}^2$ γύρω από το $\hat{b}$, τέτοια ώστε να ανήκει με μεγάλη πιθανότητα (π.χ 95%) σε αυτήν το b.

Για το σκοπό αυτό χρειάζομαι ακόμα ένα ορισμό και μια πρόταση: VAR8.

Ορισμός Αν $X_i \sim N(0,1)$, $i=1,\dots,n$ ανεξάρτητες τυχαίες μεταβλητές, τότε η κατανομή της $Z := X_1^2+\dots+X_n^2$ ονομάζεται χ-τετράγωνο με n βαθμούς ελευθερίας και συμβολίζεται με χ^2_n.

Πρόταση. Αν $(\hat{b}-b) \sim N(0,\Sigma)$ τότε $(\hat{b}-b)^T \Sigma^{-1}(\hat{b}-b) \sim \chi^2_2$ (με $b, \hat{b} \in \mathbb{R}^2$).

Απόδ.: Θέτω $Z = \Sigma^{-1/2}(\hat{b}-b)$ τότε η Z είναι κανονική με πίνακα συνδιακύμανσης $Var(Z) = \Sigma^{-1/2}\Sigma\Sigma^{-1/2} = I$

Επιπλέον:

$$(\hat{b}-b)^T \Sigma^{-1}(\hat{b}-b) = \left[\Sigma^{-1/2}(\hat{b}-b)\right]^T \cdot \left[\Sigma^{-1/2}(\hat{b}-b)\right] = Z^T Z = Z_1^2+Z_2^2$$

$$\sim \chi^2_2.$$

- Κατασκευή των ελλειψοειδών εμπιστοσύνης:

- Βρίσκουμε ~~...~~ το $\chi^2_{2,95\%}$, ~~...~~ το οποίο σημείο, δηλαδή τέτοιο ώστε

$$P(\text{μια } \chi^2_2\text{-κατανομή} \le \chi^2_{2,95\%}) = 95\%.$$

— Βρίσκουμε την ισοϋψή της $y^T\Sigma^{-1}y = \chi^2_{2,95\%}$

: ~~έλλ~~ μια ~~ελλειψ~~ έλλειψη με κέντρο το 0 και κορυφές πάνω στα ιδιοδιανύσματα του Σ^{-1}, q_1, q_2 και δη τις

$$\pm\sqrt{\frac{\chi^2_{2,95}}{\lambda_i}}\cdot q_i$$

, όπου λ_1, λ_2 οι ιδιοτιμές του Σ^{-1}.

— ~~Ερμηνεύεται~~ Το ελλειψοειδές (εδώ στον $\mathbb{R}^2$: έλλειψη) εμπιστοσύνης ~~θα~~ δηλ είναι εκείνα τα b για τα οποία

$$(\hat{b}-b)^T\Sigma^{-1}(\hat{b}-b) \le \chi^2_{2,95\%}$$

, δηλαδή το εσωτερικό της παραπάνω έλλειψης $\left(y^T\Sigma^{-1}y \le \chi^2_{2,95\%}\right)$, μετατοπισμένο κατά $\hat{b}$