Κεφάλαιο 10: Μέθοδος Μέγιστης Πιθανοφάνειας.

Εισαγωγή

Συνήθως υποθέτουμε κανονικότητα.

* Normal qq-plot
* test for normality.

$H_0: X \sim \text{Normal}$

$H_1: X$ όχι Normal

(1) **Bera - Jarque Test.**

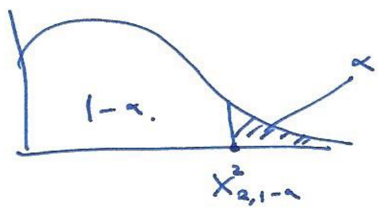
$$LM_{JB} = \frac{N}{6} \cdot S_k^2 + \frac{N}{6} \cdot (k_u - 3)^2$$

όπου

$$S_k = \frac{\frac{1}{N} \cdot \sum_{i=1}^{N} x_i^3}{\left(\frac{1}{N} \cdot \sum_{i=1}^{N} x_i^2\right)^{3/2}}, \quad k_u = \frac{\frac{1}{N} \cdot \sum_{i=1}^{N} x_i^4}{\left(\frac{1}{N} \cdot \sum_{i=1}^{N} x_i^2\right)^2}$$

$LM \sim \chi^2_2$ β.ε.

απορ. H_0 αν $LM > \chi^2_{2,1-\alpha}$.



(2) Μη παραμετρικοί έλεγχοι κανονικότητας

* Kolmogorov - Smirnov
* Shapiro-Wilk

10.1 Εκτίμηση Απλού Γραμμικού Μοντέλου με τη Μέθοδο Μέγιστης Πιθανοφάνειας.

$$y_i = \beta_1 + \beta_2 \cdot X_i + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2). \ \forall i$$

$$E(y_i | X_i) = E(\beta_1 + \beta_2 X_i + \varepsilon_i) = E(\beta_1) + E(\beta_2 X_i) + \underbrace{E(\varepsilon_i)}_{0}$$

$$= \beta_1 + \beta_2 \cdot X_i.$$

$$V(y_i | X_i) = V(\beta_1 + \beta_2 X_i + \varepsilon_i) = V(\varepsilon_i) = \sigma^2.$$

Άρα $y_i|X_i \sim N(\beta_1 + \beta_2 X_i, \sigma^2)$.

$$\text{Likelihood} = f(Y_1 = y_1, Y_2 = y_2, \ldots, Y_N = y_N) =$$

$$= f(y_1) \cdot f(y_2) \cdot \ldots \cdot f(y_N) = \prod_{i=1}^{N} f(y_i).$$

$$y_i \sim N(\beta_1 + \beta_2 X_i, \sigma^2) \Rightarrow f(y_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{ -\frac{1}{2\sigma^2} \cdot (y_i - \beta_1 - \beta_2 X_i)^2 \right\}$$

$$\text{Άρα} \quad Lik = \prod_{i=1}^{N} f(y_i) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{ -\frac{1}{2\sigma^2} \cdot (y_i - \beta_1 - \beta_2 X_i)^2 \right\} =$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{ -\frac{1}{2\sigma^2} \cdot (y_1 - \beta_1 - \beta_2 X_1)^2 \right\}$$

$$\cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{ -\frac{1}{2\sigma^2} \cdot (y_2 - \beta_1 - \beta_2 \cdot X_2)^2 \right\} \cdot$$

$$\cdot \ldots$$

$$\cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left\{ -\frac{1}{2\sigma^2} \cdot (y_N - \beta_1 - \beta_2 X_N)^2 \right\} =$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^N \cdot \exp\left\{ -\frac{1}{2\sigma^2} \cdot \sum_{i=1}^{N} (y_i - \beta_1 - \beta_2 X_i)^2 \right\}$$

$$= (2\pi\sigma^2)^{-\frac{N}{2}} \cdot \exp\left\{ -\frac{1}{2\sigma^2} \cdot \sum_{i=1}^{N} (y_i - \beta_1 - \beta_2 X_i)^2 \right\}.$$

$$\text{log-lik} = -\frac{N}{2} \cdot \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \cdot \sum_{i=1}^{N} (y_i - \beta_1 - \beta_2 X_i)^2 =$$

$$= -\frac{N}{2} \cdot \ln(2\pi) - \frac{N}{2} \cdot \ln(\sigma^2) - \frac{1}{2\sigma^2} \cdot \sum_{i=1}^{N} (y_i - \beta_1 - \beta_2 X_i)^2.$$

$$= -\frac{N}{2} \cdot \ln(2\pi) - \frac{N}{2} \cdot \ln(\sigma^2) - \frac{1}{2\sigma^2} \cdot \sum_{i=1}^{N} \varepsilon_i^2.$$

$$\frac{d\,\text{loglik}}{d\beta_1} = \cancel{-}\frac{1}{\cancel{2}\sigma^2}\cdot\sum_{i=1}^{N}\cancel{2}(y_i-\beta_1-\beta_2X_i)\cdot(\cancel{-1}) = \frac{1}{\sigma^2}\cdot\sum_{i=1}^{N}(y_i-\beta_1-\beta_2X_i)=0$$

$$\Rightarrow \sum_{i=1}^{N}\left[y_i-\beta_1-\beta_2X_i\right]=0. \quad (1)$$

$$\frac{d\,\text{loglik}}{d\beta_2} = \cancel{-}\frac{1}{\cancel{2}\sigma^2}\cdot\sum_{i=1}^{N}\cancel{2}\cdot(y_i-\beta_1-\beta_2X_i)\cdot(\cancel{-}X_i) = \frac{1}{\sigma^2}\cdot\sum_{i=1}^{N}(y_i-\beta_1-\beta_2X_i)(X_i)=0$$

$$\Rightarrow \sum_{i=1}^{N}\left[y_i-\beta_1-\beta_2X_i\right]\cdot X_i=0. \quad (2).$$

Σύστημα κανονικών εξισώσεων $\quad (1)=0$
$\qquad\qquad\qquad\qquad\qquad\quad (2)=0.$

$$\boxed{\hat\beta_1 = \bar y - \hat\beta_2\cdot\bar X}$$

$$\boxed{\hat\beta_2 = \frac{\sum(x_i-\bar x)(y_i-\bar y)}{\sum(x_i-\bar x)^2} = \frac{N\cdot\sum y_ix_i-\sum y_i\sum x_i}{N\cdot\sum x_i^2-\left(\sum x_i\right)^2}.}$$

$$\frac{d\,\text{loglik}}{d\sigma^2} = -\frac{N}{2}\cdot\frac{1}{\sigma^2}-\frac{1}{2}\cdot\sum(y_i-\beta_1-\beta_2X_i)^2\cdot\left(\frac{1}{\sigma^2}\right)' =$$

$$= -\frac{N}{2}\cdot\frac{1}{\sigma^2}-\frac{1}{2}\cdot\sum_{i=1}^{N}(y_i-\beta_1-\beta_2X_i)^2\cdot\left(-\frac{1}{\sigma^4}\right) =$$

$$\left(\frac{1}{\sigma^2}\right)' = \left((\sigma^2)^{-1}\right)' = -1\cdot(\sigma^2)^{-2}\cdot 1.$$

$$= -\frac{N}{2}\cdot\frac{1}{\sigma^2}+\frac{1}{2}\cdot\frac{1}{\sigma^4}\cdot\sum(y_i-\beta_1-\beta_2X_i)^2 =$$

$$= \frac{1}{2\sigma^2}\cdot\left[-N+\frac{1}{\sigma^2}\cdot\sum(y_i-\beta_1-\beta_2X_i)^2\right]=0$$

$$\Rightarrow -N+\frac{1}{\sigma^2}\cdot\sum(y_i-\beta_1-\beta_2X_i)^2=0 \Rightarrow$$

$$\Rightarrow \frac{\sum(y_i-\beta_1-\beta_2X_i)^2}{\sigma^2}=N \Rightarrow$$

$$\Rightarrow \boxed{\hat\sigma^2 = \frac{\sum(y_i-\beta_1-\beta_2X_i)^2}{N.}} = \boxed{\frac{\sum\left(y_i-\hat\beta_1-\hat\beta_2X_i\right)^2}{N} = \frac{\sum\hat\varepsilon_i^2}{N}}$$

10.2 ML Εκτιμητής παραμέτρων Πολλαπλού Γραμμικού Υποδείγματος.

$$y_i = \beta_1 + \beta_2 X_{i2} + \beta_3 X_{i3} + \dots + \beta_k X_{ik} + \varepsilon_i \quad , \quad \varepsilon_i \sim N(0, \sigma^2).$$

$$Y = X\beta + \varepsilon. \qquad \underset{(k+1)\times 1}{\theta} = (\beta_1, \beta_2, \dots, \beta_k, \sigma^2)'$$

$$\begin{aligned}
L_{ik} &= -\frac{N}{2}\cdot \ln(2\pi) - \frac{N}{2}\cdot \ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{N} \varepsilon_i^2 \\
&= -\frac{N}{2}\cdot \ln(2\pi) - \frac{N}{2}\cdot \ln(\sigma^2) - \frac{1}{2\sigma^2}\cdot \varepsilon'\varepsilon \qquad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{pmatrix}. \\
&= -\frac{N}{2}\cdot \ln(2\pi) - \frac{N}{2}\cdot \ln(\sigma^2) - \frac{1}{2\sigma^2}\cdot (Y - XB)'(Y - XB) \\
&= -\frac{N}{2}\cdot \ln(2\pi) - \frac{N}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2}\cdot \left(Y'Y - 2 B'X'y + \beta'X'XB\right).
\end{aligned}$$

$$\underset{(k+1)\times 1}{\frac{dL}{d\theta}} = \begin{bmatrix} \dfrac{dL}{dB} \\[2ex] \dfrac{dL}{d\sigma^2} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\sigma^2}\left[X'y - X'XB\right] = \dfrac{1}{\sigma^2}\cdot X'(y - XB) \\[2ex] -\dfrac{N}{2}\cdot\dfrac{1}{\sigma^2} + \dfrac{1}{2}\dfrac{1}{\sigma^4}\cdot (Y - XB)'(Y - XB) \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{1}{\sigma^2}\cdot (X'y - X'XB) \\[2ex] -\dfrac{N}{2}\cdot\dfrac{1}{\sigma^2} + \dfrac{1}{2}\dfrac{1}{\sigma^4}\cdot (y - XB)'(Y - XB) \end{bmatrix}\begin{matrix}(1)\\ \\ (2)\end{matrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \qquad \text{score vector.}$$

$$(1) = 0 \Rightarrow \frac{1}{\sigma^2}\cdot (X'y - X'XB) = 0 \Rightarrow X'y - X'XB = 0 \Rightarrow$$

$$\Rightarrow X'XB = X'y \Rightarrow \boxed{\hat{B} = (X'X)^{-1}\cdot X'y.}$$

$$(2)\ \frac{1}{2}\frac{1}{\sigma^2}\cdot \left(-N + \frac{1}{\sigma^2}\cdot(\varepsilon'\varepsilon)\right) = 0 \Rightarrow \frac{1}{\sigma^2}\varepsilon'\varepsilon = N \Rightarrow$$

Hessian Matrix: $H(\theta) = \frac{d}{d\theta}\left(\frac{dL}{d\theta}\right) = \frac{d^2 L}{d\theta' d\theta} = \frac{d}{d(\beta', \sigma^2)}\begin{pmatrix} \frac{dL}{d\beta} \\ \frac{dL}{d\sigma^2} \end{pmatrix} =$

$$= \begin{bmatrix} \underset{k \times k}{\frac{d^2 L}{d\beta' d\beta}} & \underset{1 \times k}{\frac{d^2 L}{d\sigma^2 d\beta}} \\ \underset{1 \times k}{\frac{d^2 L}{d\beta' d\sigma^2}} & \underset{1 \times 1}{\frac{d^2 L}{d\sigma^2 d\sigma^2}} \end{bmatrix}$$

$$\frac{d^2 L}{d\beta' d\beta} = \frac{d}{d\beta'}\left[\frac{1}{\sigma^2}\cdot(X'y - X'X\beta)\right] = -\frac{1}{\sigma^2}\cdot X'X.$$

$$\frac{d^2 L}{d\sigma^2 d\beta} = \frac{d}{d\sigma^2}\left[\frac{1}{\sigma^2}\cdot(X'y - X'X\beta)\right] = -\frac{1}{\sigma^4}\cdot X'(y - X\beta) = -\frac{1}{\sigma^4}\cdot X'\varepsilon.$$

$$\frac{d^2 L}{d\beta' d\sigma^2} = \frac{d}{d\beta'}\left[-\frac{N}{2}\cdot\frac{1}{\sigma^2} + \frac{1}{2}\frac{1}{\sigma^4}\cdot\underbrace{(y - X\beta)'(y - X\beta)}_{y'y - 2\beta'X'y + \beta'X'X\beta}\right] =$$

$$= \frac{1}{\cancel{2}}\cdot\frac{1}{\sigma^4}\cdot(-\cancel{2}X'y + \cancel{2}X'X\beta)' = -\frac{1}{\sigma^4}\cdot\left(X'\cdot(y - X\beta)\right)' =$$

$$= -\frac{1}{\sigma^4}\cdot(y - X\beta)'\cdot X = -\frac{1}{\sigma^4}\cdot\varepsilon' X.$$

$$\left(\frac{d^2 L}{d\sigma^2 d\beta}\right)' = \frac{d^2 L}{d\beta' d\sigma^2}.$$

$$\frac{d^2 L}{d\sigma^2 d\sigma^2} = \frac{d}{d\sigma^2}\left[-\frac{N}{2}\cdot\frac{1}{\sigma^2} + \frac{1}{2}\cdot\frac{1}{\sigma^4}\cdot(\varepsilon'\varepsilon)\right] = +\frac{N}{2}\cdot\frac{1}{\sigma^4} + \frac{1}{2}\cdot(-1)\cdot\frac{1}{(\sigma^2)^3}\cdot\varepsilon'\varepsilon =$$

$$= \frac{N}{2\sigma^4} - \frac{1}{\sigma^6}\cdot\varepsilon'\varepsilon.$$

$$H(\theta) = \begin{bmatrix} \dfrac{\partial^2 L}{\partial \beta' \partial \beta} & \dfrac{\partial^2 L}{\partial \sigma^2 \partial \beta} \\ \dfrac{\partial L^2}{\partial \beta' \partial \sigma^2} & \dfrac{\partial^2 L}{\partial \sigma^2 \partial \sigma^2} \end{bmatrix} = \begin{bmatrix} -\dfrac{1}{\sigma^2} \cdot X'X & -\dfrac{1}{\sigma^4} \cdot X'\varepsilon \\ -\dfrac{1}{\sigma^4} \cdot \varepsilon' X & \dfrac{N}{2\sigma^4} - \dfrac{1}{\sigma^6} \cdot \varepsilon'\varepsilon \end{bmatrix}.$$

Information Matrix.

$$I(\theta) = E\left[-H(\theta)\right] = E\left[-\frac{\partial^2 L}{\partial \theta' \partial \theta}\right] =$$

$$= E\left\{-\begin{bmatrix} -\dfrac{1}{\sigma^2} X'X & -\dfrac{1}{\sigma^4} X'\varepsilon \\ -\dfrac{1}{\sigma^4} \varepsilon' X & \dfrac{N}{2\sigma^4} - \dfrac{1}{\sigma^6} \varepsilon'\varepsilon \end{bmatrix}\right\} = \begin{bmatrix} \dfrac{1}{\sigma^2} \cdot X'X & 0 \\ 0 & -\dfrac{N}{2\sigma^4} + \dfrac{1}{\sigma^6} \cdot N \cdot \sigma^2 \end{bmatrix} =$$

$$\hat{\sigma}^2 = \frac{\varepsilon'\varepsilon}{N} \Rightarrow \varepsilon'\varepsilon = N \cdot \hat{\sigma}^2 \Rightarrow \boxed{E(\varepsilon'\varepsilon) = N \cdot \sigma^2.}$$

$$= \begin{bmatrix} \dfrac{1}{\sigma^2} \cdot X'X & 0 \\ 0 & \dfrac{1}{2} \cdot \dfrac{N}{\sigma^4} \end{bmatrix}.$$

$$\left(-\frac{N}{2\sigma^4} + \frac{N}{\sigma^4} = \frac{N}{\sigma^4}\left(1 - \frac{1}{2}\right) = \frac{1}{2} \cdot \frac{N}{\sigma^4}\right).$$

$$I^{-1}(\theta) = \begin{bmatrix} \sigma^2 \cdot (X'X)^{-1} & 0 \\ 0 & \dfrac{2\sigma^4}{N} \end{bmatrix}.$$

$$V(\hat{\beta}) = \sigma^2 (X'X)^{-1} \qquad V(\hat{\sigma}^2) = \frac{2\sigma^4}{N}.$$

$$V(\hat{\beta}) \geq \left[E\left(-\frac{\partial^2 L}{\partial \beta' \partial \beta}\right)\right]^{-1} = \left\{E\left(-H(\beta)\right)\right\}^{-1} = I(\beta)^{-1}.$$

Ιδιότητες

$$E\left[-H(\theta)\right] = E\left[-\frac{\partial^2 L}{\partial\theta'\partial\theta}\right] = E\left[\frac{\partial L}{\partial \theta}\cdot\left(\frac{\partial L}{\partial \theta}\right)'\right] = E\left(s(\theta)\cdot s'(\theta)\right)$$

$$s(\theta) = \frac{\partial L}{\partial \theta} = \begin{bmatrix} \frac{1}{\sigma^2}\cdot X'\varepsilon \\ -\frac{N}{2\sigma^2} + \frac{1}{2\sigma^4}\cdot\varepsilon'\varepsilon \end{bmatrix}.$$

$$\hat{\theta}_{ML} \sim N\left(\theta,\ Asy.Var(\hat{\theta}) = \left\{E\left(-H(\theta)\right)\right\}^{-1} = \begin{bmatrix} \frac{1}{\sigma^2}E(X'X) & 0 \\ 0' & \frac{N}{2\sigma^4} \end{bmatrix}^{-1}\right)$$

$$\boxed{\theta^{k+1} = \theta^{k} - \boxed{Q}^{-1}\cdot\frac{\partial L}{\partial \theta^{k}}}$$

approximation to ~~[illegible]~~ Hessian.

Εκτός από τον παραπάνω, ένας άλλος συνεπής εκτιμητής που χρησιμοποιείται συχνά στην πράξη για την εκτίμηση της ασυμπτωτικής διακύμανσης $Asy\mathrm{Var}(\hat{\boldsymbol{\theta}}_{ML})$ στηρίζεται στον ακόλουθο εκτιμητή της ασυμπτωτικής διακύμανσης του σκορ διανύσματος $Asy\mathrm{Var}\left(\mathrm{S}(\boldsymbol{\theta})\right)$:

$$\mathrm{S}(\hat{\boldsymbol{\theta}}_{ML})\mathrm{S}(\hat{\boldsymbol{\theta}}_{ML})' = \begin{bmatrix} \dfrac{\mathbf{X}'\hat{\boldsymbol{\varepsilon}}\hat{\boldsymbol{\varepsilon}}'\mathbf{X}}{\hat{\sigma}^4_{ML}} & \dfrac{1}{\hat{\sigma}^2_{ML}}\left(-\dfrac{\mathrm{N}}{2\hat{\sigma}^2_{ML}}\mathbf{X}'\hat{\boldsymbol{\varepsilon}} + \dfrac{1}{2\hat{\sigma}^4_{ML}}\mathbf{X}'\hat{\boldsymbol{\varepsilon}}\left(\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}\right)\right) \\ \dfrac{1}{\hat{\sigma}^2_{ML}}\left(-\dfrac{\mathrm{N}}{2\hat{\sigma}^2_{ML}}\hat{\boldsymbol{\varepsilon}}'\mathbf{X} + \dfrac{1}{2\hat{\sigma}^4_{ML}}\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}\hat{\boldsymbol{\varepsilon}}'\mathbf{X}\right) & \left(-\dfrac{\mathrm{N}}{2\hat{\sigma}^2_{ML}} + \dfrac{1}{2\hat{\sigma}^4_{ML}}\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}\right)\left(-\dfrac{\mathrm{N}}{2\hat{\sigma}^2_{ML}} + \dfrac{1}{2\hat{\sigma}^4_{ML}}\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}\right) \end{bmatrix}$$

Παίρνοντας όρια κατά πιθανότητα του εκτιμητή αυτού αποδεικνύεται εύκολα η συνέπεια του εκτιμητή αυτού ως εξής:

$$\mathrm{S}(\hat{\boldsymbol{\theta}}_{ML})\mathrm{S}(\hat{\boldsymbol{\theta}}_{ML})' \xrightarrow{p} \mathrm{E}\left[\mathrm{S}(\boldsymbol{\theta})\mathrm{S}(\boldsymbol{\theta})'\right] = Asy\mathrm{Var}\left(\mathrm{S}(\boldsymbol{\theta})\right).$$

Χρησιμοποιώντας κανόνες σύγκλισης κατά πιθανότητα (βλέπε Κεφάλαιο 6), το τελευταίο αποτέλεσμα συνεπάγεται ότι $\left[\mathrm{S}(\hat{\boldsymbol{\theta}}_{ML})\mathrm{S}(\hat{\boldsymbol{\theta}}_{ML})'\right]^{-1} \xrightarrow{p} \left[Asy\mathrm{Var}\left(\mathrm{S}(\boldsymbol{\theta})\right)\right]^{-1}$ και επομένως, λόγω της ισότητας της μήτρας πληροφοριών $\mathrm{E}\left[-\mathrm{H}(\boldsymbol{\theta})\right] = \mathrm{E}\left[\mathrm{S}(\boldsymbol{\theta})\mathrm{S}(\boldsymbol{\theta})'\right]$ η μήτρα διακύμανσης-συνδιακύμανσης $\left[\mathrm{S}(\hat{\boldsymbol{\theta}}_{ML})\mathrm{S}(\hat{\boldsymbol{\theta}}_{ML})'\right]^{-1}$ μπορεί επίσης να χρησιμοποιηθεί για την συνεπή εκτίμηση της ασυμπτωτικής διακύμανσης του εκτιμητή $\hat{\boldsymbol{\theta}}_{ML}$, $Asy\mathrm{Var}(\hat{\boldsymbol{\theta}}_{ML})$.

10.4 Τρεις εναλλακτικοί ισοδύναμοι ασυμπτωτικά έλεγχοι στατιστικών υποθέσεων (*)

Οι ασυμπτωτικές ιδιότητες του εκτιμητή Μεγίστης Πιθανοφάνειας $\hat{\boldsymbol{\theta}}_{ML}$ μπορούν να χρησιμοποιηθούν για να βρεθούν οι κατανομές τριών εναλλακτικών στατιστικών κριτηρίων. Αυτά εφαρμόζονται σε ελέγχους στατιστικών υποθέσεων για τους συντελεστές του γραμμικού υποδείγματος που παίρνουν την ακόλουθη γενική μορφή γραμμικών περιορισμών:[10]

$$\mathbf{R}\,\boldsymbol{\beta} = \mathbf{r}, \tag{7}$$

(JxK) (Kx1) = (r x1)

όπου J δηλώνει τον αριθμό των περιορισμών, **R** αποτελεί τη (JXK)-διάστασης μήτρα σχεδιασμού των περιορισμών και τέλος, **r** αποτελεί το (JX1)-διάστασης διάνυσμα των τιμών των περιορισμών.

Τα τρία κριτήρια που θα παρουσιάσουμε για τον έλεγχο των παραπάνω γραμμικών περιορισμών στηρίζονται στην ασυμπτωτική κατανομή του ML εκτιμητή $\hat{\boldsymbol{\beta}}_{ML}$ ή του διανύσματος των πρώτων παραγώγων $S(\boldsymbol{\theta})$. Αυτά είναι: το κριτήριο του λόγου πιθανοφανειών (likelihood ratio test, LR), το κριτήριο Wald (Wald test, W) και το κριτήριο του πολλαπλασιαστή Lagrange (lagrange multiplier, LM). Κάτω από τη μηδενική υπόθεση ότι ισχύουν από κοινού οι περιορισμοί που δίνονται από τη σχέση (7) και τα τρία αυτά κριτήρια ακολουθούν την χ^2-κατανομή με τους ίδιους βαθμούς ελευθερίας, οι οποίοι ισούνται με τον αριθμό των περιορισμών J. Το αποτέλεσμα αυτό ισχύει για μεγάλα δείγματα παρατηρήσεων. Επειδή και τα τρία αυτά κριτήρια ακολουθούν την ίδια κατανομή σε μεγάλα δείγματα παρατηρήσεων, αυτά θεωρούνται ως ασυμπτωτικά ισοδύναμα. Στη συνέχεια του κεφαλαίου παρουσιάζουμε καθένα από τα κριτήρια αυτά ξεχωριστά και τέλος, αποδεικνύουμε την αλγεβρική σχέση που τα συνδέει.

Το κριτήριο του λόγου πιθανοφανειών (LR)

Για το στατιστικό έλεγχο των γραμμικών περιορισμών της σχέσης (7), το κριτήριο του λόγου πιθανοφανειών συγκρίνει τη μέγιστη τιμή της λογαριθμική συνάρτησης πιθανοφάνειας $L=\ln\varphi(y_1, y_2, ..., y_N)$ που δίνεται από τη σχέση (3) όταν το γραμμικό υπόδειγμα εκτιμάται χωρίς περιορισμούς (unrestricted-UN) με εκείνη όταν εκτιμάται κάτω από περιορισμούς (restricted-RE). Η πρώτη από αυτές τις δύο τιμές της συνάρτησης $L=\ln\varphi(y_1, y_2, ...,y_N)$ υπολογίζεται αντικαθιστώντας το διάνυσμα τιμών του εκτιμητή $\hat{\boldsymbol{\theta}}_{ML}=(\hat{\boldsymbol{\beta}}_{ML}', \hat{\sigma}^2_{ML})'$ σε αυτή. Την τιμή αυτή της συνάρτησης $L=\ln\varphi(y_1, y_2, ...,y_N)$ θα τη

[10] Βλέπε Κεφάλαιο 4.

συμβολίζουμε στο εξής ως $L_{UN}(\hat{\boldsymbol{\theta}}_{ML})$. Με βάση τους παραπάνω ορισμούς, αυτή υπολογίζεται ως ακολούθως:

$$L_{UN}(\hat{\boldsymbol{\theta}}_{ML}) = -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln\hat{\sigma}^2_{ML} - \frac{\sum_{i=1}^{N}\hat{\varepsilon}_i^{\,2}}{2\hat{\sigma}^2_{ML}}$$

$$= -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln\left(\frac{\sum_{i=1}^{N}\hat{\varepsilon}_i^{\,2}}{N}\right) - \frac{\sum_{i=1}^{N}\hat{\varepsilon}_i^{\,2}}{2\left(\frac{\sum_{i=1}^{N}\hat{\varepsilon}_i^{\,2}}{N}\right)}$$

$$= -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln\left(\frac{\sum_{i=1}^{N}\hat{\varepsilon}_i^{\,2}}{N}\right) - \frac{N}{2} = -\frac{N}{2}\left[\ln(2\pi)+1\right] - \frac{N}{2}\ln\left(\frac{\sum_{i=1}^{N}\hat{\varepsilon}_i^{\,2}}{N}\right)$$

$$= -\frac{N}{2}\left[\ln(2\pi)+1\right] - \frac{N}{2}\ln\left(\sum_{i=1}^{N}\hat{\varepsilon}_i^{\,2}\right) + \frac{N}{2}\ln(N)$$

$$= -\frac{N}{2}\left[\ln(2\pi)+1-\ln(N)\right] - \frac{N}{2}\ln\left(\sum_{i=1}^{N}\hat{\varepsilon}_i^{\,2}\right) \qquad (8\alpha)$$

Η μέγιστη τιμή της λογαριθμικής συνάρτησης πιθανοφάνειας $L=\ln\varphi(y_1, y_2, ..., y_N)$ κάτω από περιορισμούς (restrictions-RE) βρίσκεται αν το γραμμικό υπόδειγμα εκτιμηθεί με τη μέθοδο ML κάτω με τους περιορισμούς $\mathbf{R}\boldsymbol{\beta}=\mathbf{r}$. Αντικαθιστώντας το διάνυσμα των τιμών του εκτιμητή αυτού στη συνάρτηση $L=\ln\varphi(y_1, y_2, ..., y_N)$ δίνει την τιμή αυτής κάτω από τους περιορισμούς. Τον εκτιμητή των παραμέτρων του γραμμικού υποδείγματος αυτού, θα τον συμβολίζουμε ως $\hat{\boldsymbol{\theta}}^*_{ML}=(\hat{\boldsymbol{\beta}}^{*\prime}_{ML}, \hat{\sigma}^{*2}_{ML})'$, ενώ τη μέγιστη τιμή της συνάρτησης $L=\ln\varphi(y_1,y_2, ...,y_N)$ που προκύπτει με βάση αυτόν θα τη συμβολίζουμε ως $L_{RE}(\hat{\boldsymbol{\theta}}^*_{ML})$.[11] Ανάλογα με τη σχέση 8(α), η τιμή της συνάρτησης αυτής υπολογίζεται ως εξής:

[11] Σημειώστε ότι μπορεί εύκολά να αποδειχθεί ότι ο ML εκτιμητής $\hat{\boldsymbol{\beta}}^*_{ML}$ είναι ίδιος με εκείνον των LS κάτω από τους περιορισμούς (βλέπε Κεφάλαιο 4). Αυτός δίνεται αναλυτικά ως εξής:

$\sum_{i=1}^{N} \hat{\varepsilon}_i^{*2}$

$$L_{RE}(\hat{\theta}^*_{ML}) = -\frac{N}{2}\left[\ln(2\pi)+1-\ln(N)\right] - \frac{N}{2}\ln\left(\sum_{i=1}^{N}\hat{\varepsilon}_i^{*2}\right), \quad \checkmark \qquad (8\beta)$$

όπου $\hat{\varepsilon}^*_i$ αποτελούν τα κατάλοιπα του γραμμικού υποδείγματος που υπολογίζονται με βάση τον εκτιμητή $\hat{\theta}^*_{ML} = (\hat{\beta}^{*\prime}_{ML}, \hat{\sigma}^{*2}_{ML})'$.

Έχοντας ορίσει τις μέγιστες τιμές της συνάρτησης πιθανοφάνειας (8α) και (8β), στη συνέχεια ορίζουμε το στατιστικό κριτήριο του λόγου των δύο πιθανοφανειών $L_{UR}(\hat{\theta}_{ML})$ και $L_{RE}(\hat{\theta}^*_{ML})$ ως εξής:

$H_0: RB = r$

$$LR = -2 \cdot \ln\left[\frac{Lik(\theta_{M_0})}{L(\theta_{M_1})}\right] \Rightarrow$$

$$\Rightarrow LR = -2\left[\ln L(\theta_{M_0}) - \ln L(\theta_{M_1})\right] \Rightarrow$$

$$\Rightarrow LR \equiv -2\left[L_{RE}(\hat{\theta}^*_{ML}) - L_{UR}(\hat{\theta}_{ML})\right] \qquad (9)$$

Το κριτήριο αυτό παίρνει πάντα θετικές τιμές καθώς στη βελτιστοποίηση μιας συνάρτησης ισχύει πάντα η ανισότητα $L_{UR}(\hat{\theta}_{ML}) > L_{RE}(\hat{\theta}^*_{ML})$, δηλ. το μέγιστο της συνάρτησης χωρίς περιορισμούς είναι μεγαλύτερο εκείνου όταν η συνάρτηση εκτιμάται με περιορισμούς. Αντικαθιστώντας τις σχέσεις (8α)-(8β) στη σχέση (9), το κριτήριο LR γράφεται πιο αναλυτικά ως εξής:

$$LR = -2\left[L_{RE}(\hat{\theta}^*_{ML}) - L_{UR}(\hat{\theta}_{ML})\right] = N\left[\ln\left(\sum_{i=1}^{N}\hat{\varepsilon}_i^{*2}\right) - \ln\left(\sum_{i=1}^{N}\hat{\varepsilon}_i^{2}\right)\right] \quad \checkmark$$

$$\hat{\beta}^*_{ML} = \hat{\beta}_{ML} + (X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]\left(r - R\hat{\beta}_{ML}\right).$$

Για να βρεθεί ο εκτιμητής αυτός θα πρέπει να μεγιστοποιήσουμε την ακόλουθη λογαριθμική συνάρτηση πιθανοφάνειας κάτω από τους γραμμικούς περιορισμούς $\mathbf{R}\,\boldsymbol{\beta} = \mathbf{r}$, δηλαδή

$$L = -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}(y'y - 2\beta'X'y + \beta'X'X\beta) - \lambda'(R\beta - r),$$

όπου λ αποτελεί το (JX1)-διάστασης διάνυσμα των πολλαπλασιαστών Lagrange. Από τις συνθήκες πρώτης τάξης του προβλήματος αυτού, μπορεί να αποδειχθεί ότι το διάνυσμα των πολλαπλασιαστών Lagrange δίνεται ως $\boldsymbol{\lambda} = [\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\boldsymbol{\beta} - \mathbf{r})$.

$$= N\ln\left(\frac{\sum_{i=1}^{N}\hat{\varepsilon}_i^{*2}}{\sum_{i=1}^{N}\hat{\varepsilon}_i^{2}}\right) = N\ln\left(\frac{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*}}{\hat{\varepsilon}^{\prime}\hat{\varepsilon}}\right) \qquad (10\alpha)$$

Όπως φαίνεται από την τελευταία σχέση, αν ισχύουν οι γραμμικοί περιορισμοί $\mathbf{R}\,\boldsymbol{\beta} = \mathbf{r}$, τότε η απόσταση μεταξύ των μέγιστων τιμών των λογαριθμικών συναρτήσεων $L_{UR}(\hat{\theta}_{ML})$ και $L_{RE}(\hat{\theta}^{*}_{ML})$ θα πρέπει να είναι πολύ κοντά στο μηδέν και όχι σημαντικά διάφορη από αυτό. Το ίδιο θα συμβαίνει και για τα αθροίσματα των καταλοίπων $\hat{\boldsymbol{\varepsilon}}^{*\prime}\hat{\boldsymbol{\varepsilon}}^{*}$ και $\hat{\boldsymbol{\varepsilon}}^{\prime}\hat{\boldsymbol{\varepsilon}}$ του γραμμικού υποδείγματος σε κάθε μια από τις δύο παραπάνω μεθόδους εκτίμησης της συνάρτησης πιθανοφάνειας, δηλ. χωρίς και με περιορισμούς, αντίστοιχα. Κάτω από την μηδενική υπόθεση ότι ισχύουν οι περιορισμοί $\mathbf{R}\,\boldsymbol{\beta} = \mathbf{r}$, αποδεικνύεται ότι η κατανομή του κριτηρίου LR καθώς $N \to \infty$ δίνεται ως εξής:

$$LR \sim \chi^2_{(J)},$$

όπου J δηλώνει τους βαθμούς ελευθερίας της χ^2-κατανομής. Οι βαθμοί αυτοί ισούνται με τον αριθμό των περιορισμών που ελέγχονται κάτω από τη μηδενική υπόθεση.

Εκτός από τον τύπο της σχέσης (10α), το κριτήριο του λόγου πιθανοφανειών LR μπορεί να παρουσιαστεί με δύο άλλους τύπους. Αυτοί μπορούν να βρεθούν από τη σχέση (10α) με κατάλληλους αλγεβρικούς χειρισμούς και δίνονται αντίστοιχα ως ακολούθως:

$$LR = N\ln\left(1+\frac{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*}}{\hat{\varepsilon}^{\prime}\hat{\varepsilon}}-1\right) = N\ln\left(1+\frac{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*}-\hat{\varepsilon}^{\prime}\hat{\varepsilon}}{\hat{\varepsilon}^{\prime}\hat{\varepsilon}}\right) \qquad (10\beta)$$

και

$$LR = N\ln\left(\frac{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*}}{\hat{\varepsilon}^{\prime}\hat{\varepsilon}}\right) = N\ln\left(\frac{1}{\dfrac{\hat{\varepsilon}^{\prime}\hat{\varepsilon}}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*}}}\right) = N\ln\left(\frac{1}{1+\dfrac{\hat{\varepsilon}^{\prime}\hat{\varepsilon}}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*}}-1}\right) = N\ln\left(\frac{1}{1+\dfrac{\hat{\varepsilon}^{\prime}\hat{\varepsilon}-\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*}}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*}}}\right) \qquad (10\gamma)$$

Οι σχέσεις (10β) και (10γ) του κριτηρίου LR θα χρησιμοποιηθούν για τη σύγκριση αυτού με τα άλλα δύο κριτήρια που χρησιμοποιούνται στον έλεγχο των περιορισμών $\mathbf{R}\boldsymbol{\beta}=\mathbf{r}$ που θα παρουσιάσουμε στη συνέχεια, δηλαδή τα κριτήρια W και LM.

Το κριτήριο Wald (W)

Το κριτήριο Wald (W) θεωρείται ως αντίστοιχο του κριτηρίου F, που παρουσιάστηκε στο Κεφάλαιο 4 για τον έλεγχο των περιορισμών $\mathbf{R}\boldsymbol{\beta}=\mathbf{r}$ για πεπερασμένα δείγματα παρατηρήσεων και κάτω από την υπόθεση ότι ο διαταρακτικός όρος κατανέμεται κανονικά.[12] Για τον έλεγχο των περιορισμών αυτών, το κριτήριο αυτό έχει το πλεονέκτημα ότι δεν απαιτεί την εκτίμηση του γραμμικού υποδείγματος κάτω από τους προς έλεγχο περιορισμούς. Αυτό στηρίζεται αποκλειστικά στις τιμές του ML εκτιμητή $\hat{\boldsymbol{\theta}}_{ML}$, χωρίς περιορισμούς.

Για την παρουσίαση του κριτηρίου W και την εύρεση της κατανομής του κάτω από τη μηδενική υπόθεση ότι ισχύουν οι περιορισμοί $\mathbf{R}\boldsymbol{\beta}=\mathbf{r}$, θεωρήστε την ασυμπωτική κατανομή του εκτιμητή $\hat{\boldsymbol{\beta}}_{ML}$ η οποία με βάση τα αποτελέσματα του Τμήματος 10.3 δίνεται ως

$$\hat{\boldsymbol{\beta}}_{ML} \sim \mathrm{N}\left[\boldsymbol{\beta}, Asy\mathrm{Var}(\hat{\boldsymbol{\beta}}_{ML})\right], \qquad \text{όπου} \qquad Asy\mathrm{Var}(\hat{\boldsymbol{\beta}}_{ML}) = \left[\frac{1}{\sigma^2}\mathrm{E}(\mathbf{X}'\mathbf{X})\right]^{-1},$$

καθώς η μήτρα διακύμανσης-συνδιακύμανσης $Asy\mathrm{Var}(\hat{\boldsymbol{\theta}}_{ML})$ είναι διαγώνια. Πολλαπλασιάζοντας την παραπάνω σχέση με τη μήτρα σχεδιασμού $\mathbf{R}$ και αφαιρώντας το διάνυσμα των περιορισμών $\mathbf{r}$ συνεπάγεται ότι

[12] Σημειώστε ότι το κριτήριο Wald το παρουσιάσαμε πρώτη φορά στο Κεφάλαιο 6, όπου το γράψαμε ως συνάρτηση του κριτηρίου F. Θυμηθείτε ότι το κριτήριο F και η κατανομή του σε πεπερασμένα δείγματα με βάση τον LS εκτιμητή $\hat{\boldsymbol{\beta}}$ δίνονται ως εξής: $\mathrm{F}=\dfrac{(\mathbf{R}\hat{\boldsymbol{\beta}}-\mathbf{r})'\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1}(\mathbf{R}\hat{\boldsymbol{\beta}}-\mathbf{r})}{\mathrm{J}\hat{\sigma}^2} \sim \mathrm{F}_{(\mathrm{J},\mathrm{N}-\mathrm{K})}$ (βλέπε Κεφάλαιο 4).

$H_0: R\beta = r$
$H_a: R\beta \neq r$

$$\mathbf{R}\hat{\boldsymbol{\beta}}_{\mathrm{ML}} - \mathbf{r} \sim \mathrm{N}\left[\mathbf{R}\boldsymbol{\beta} - \mathbf{r},\ \mathbf{R}Asy\mathrm{Var}(\hat{\boldsymbol{\beta}}_{\mathrm{ML}})\mathbf{R}'\right] \checkmark$$

$$\sim \mathrm{N}\left[\mathbf{0},\ \mathbf{R}Asy\mathrm{Var}(\hat{\boldsymbol{\beta}}_{\mathrm{ML}})\mathbf{R}'\right] \checkmark \qquad (11)$$

Το αποτέλεσμα της σχέσης (11) σημαίνει ότι η ακόλουθη τυποποιημένη τετραγωνική μορφή των αποστάσεων $\mathbf{R}\hat{\boldsymbol{\beta}}_{\mathrm{ML}} - \mathbf{r}$ υπολογισμένων με βάση τις τιμές του εκτιμητή $\hat{\boldsymbol{\beta}}_{\mathrm{ML}}$:

$$\boxed{(\mathbf{R}\hat{\boldsymbol{\beta}}_{\mathrm{ML}} - \mathbf{r})'\left[\mathbf{R}Asy\mathrm{Var}(\hat{\boldsymbol{\beta}}_{\mathrm{ML}})\mathbf{R}'\right]^{-1}(\mathbf{R}\hat{\boldsymbol{\beta}}_{\mathrm{ML}} - \mathbf{r})} \sim \chi^2_J \qquad (12)$$

ακολουθεί τη χ^2-κατανομή με J βαθμούς ελευθερίας, που ισούνται με τον αριθμό των περιορισμών. Αντικαθιστώντας τον εκτιμητή της διακύμανσης $Asy\mathrm{Var}(\hat{\boldsymbol{\beta}}_{\mathrm{ML}})$ που δίνεται ως $\hat{\sigma}^2_{\mathrm{ML}}(\mathbf{X}'\mathbf{X})^{-1}$ στη σχέση (12) δίνει τον ακόλουθο τύπο του κριτηρίου Wald για τον έλεγχο των περιορισμών $\mathbf{R}\ \boldsymbol{\beta} = \mathbf{r}$:

$Asy\mathrm{Var}(\hat{\beta}_{ML}) = \hat{\sigma}^2_{ML} \cdot (X'X)^{-1}$

$$\boxed{\begin{aligned} \mathrm{W} &= (\mathbf{R}\hat{\boldsymbol{\beta}}_{\mathrm{ML}} - \mathbf{r})'\left[\mathbf{R}\hat{\sigma}^2_{\mathrm{ML}}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1}(\mathbf{R}\hat{\boldsymbol{\beta}}_{\mathrm{ML}} - \mathbf{r}) \checkmark \\ &= \frac{(\mathbf{R}\hat{\boldsymbol{\beta}}_{\mathrm{ML}} - \mathbf{r})'\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1}(\mathbf{R}\hat{\boldsymbol{\beta}}_{\mathrm{ML}} - \mathbf{r})}{\hat{\sigma}^2_{\mathrm{ML}}} \checkmark \end{aligned}} \sim \chi^2_J \qquad (13)$$

Επειδή $\hat{\sigma}^2_{\mathrm{ML}}(\mathbf{X}'\mathbf{X})^{-1}$ αποτελεί συνεπή εκτιμητή της ασυμπτωτικής διακύμανσης $Asy\mathrm{Var}(\hat{\boldsymbol{\beta}}_{\mathrm{ML}})$, το κριτήριο W ακολουθεί την ίδια κατανομή με εκείνη της τετραγωνικής μορφής της σχέσης (12) για $\mathrm{N} \to \infty$. Δηλαδή, ισχύει το ακόλουθο αποτέλεσμα:

$$\mathrm{W} \sim \chi^2_{(J)}.$$

Αντικαθιστώντας το ακόλουθο αποτέλεσμα[13]

[13] Βλέπε Κεφάλαιο 4.

δὼ κεφ. 4

$$\hat{\boldsymbol{\varepsilon}}^{*\prime}\hat{\boldsymbol{\varepsilon}}^{*} - \hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}} = (\mathbf{R}\hat{\boldsymbol{\beta}}_{ML} - \mathbf{r})'\left[\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\right]^{-1}(\mathbf{R}\hat{\boldsymbol{\beta}}_{ML} - \mathbf{r}),$$

όπου $\hat{\boldsymbol{\varepsilon}}^{*}$ είναι το διάνυσμα των καταλοίπων του γραμμικού υποδείγματος που οι συντελεστές του εκτιμώνται με τη μέθοδο ML (ή LS) κάτω από τους περιορισμούς $\mathbf{R}\,\boldsymbol{\beta} = \mathbf{r}$, καθώς και τον εκτιμητή της διακύμανσης $\hat{\sigma}^2_{ML} = \dfrac{\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}}{N}$ στη σχέση (13) το κριτήριο W γράφεται ως εξής:[14]

$$W = \frac{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*} - \hat{\varepsilon}'\hat{\varepsilon}}{\hat{\sigma}^2_{ML}} \Rightarrow$$

$$\Rightarrow \boxed{W = \frac{\hat{\boldsymbol{\varepsilon}}^{*\prime}\hat{\boldsymbol{\varepsilon}}^{*} - \hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}}{\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}/N} = \frac{N(\hat{\boldsymbol{\varepsilon}}^{*\prime}\hat{\boldsymbol{\varepsilon}}^{*} - \hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}})}{\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}}.} \sim \chi^2_J . \qquad (14)$$

Ο τύπος αυτός του κριτηρίου W θα χρησιμοποιηθεί στη σύγκρισή του με τα κριτήριο LR και LM.

Το κριτήριο του πολλαπλασιαστή Lagrange (LM)

Σε αντίθεση με το κριτήριο W για τον έλεγχο των περιορισμών $\mathbf{R}\,\boldsymbol{\beta} = \mathbf{r}$, το κριτήριο LM βασίζεται στον εκτιμητή ML $\hat{\boldsymbol{\theta}}^{*}_{ML} = (\hat{\boldsymbol{\beta}}^{*\prime}_{ML}, \hat{\sigma}^{*2}_{ML})'$ κάτω από τους περιορισμούς αυτούς. Το κριτήριο αυτό στηρίζεται στην ακόλουθη αρχή. Αν οι περιορισμοί αυτοί ισχύουν, τότε η τυποποιημένη τετραγωνική μορφή των αποστάσεων του σκορ διανύσματος $S(\hat{\boldsymbol{\beta}}^{*}_{ML})$ που υπολογίζεται με βάση το ML εκτιμητή $\hat{\boldsymbol{\beta}}^{*}_{ML}$ δε θα πρέπει να διαφέρει στατιστικά από το μηδενικό διάνυσμα.[15]

---

[14] Σημειώστε ότι η αντίστοιχη μορφή του κριτηρίου F είναι $F = \dfrac{(\hat{\boldsymbol{\varepsilon}}^{*\prime}\hat{\boldsymbol{\varepsilon}}^{*} - \hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}})/J}{\hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}}/(N-K)}$.

[15] Θυμηθείτε ότι το σκορ διάνυσμα υπολογισμένο με βάση τον ML εκτιμητή $\hat{\boldsymbol{\theta}}_{ML}$ χωρίς περιορισμούς ισούται πάντα με το μηδενικό διάνυσμα εξ ορισμού.

Για την παρουσίαση του κριτηρίου LM και την εύρεση της ασυμπτωτικής του κατανομής, θεωρήστε την κατανομή του σκορ διανύσματος για το διάνυσμα $\boldsymbol{\theta}=(\boldsymbol{\beta}',\sigma^2)'$ [16]

$$\boxed{\mathrm{S}(\boldsymbol{\theta}) \sim \mathrm{N}\left\{\mathbf{0}, Asy\mathrm{Var}\left[\mathrm{S}(\boldsymbol{\theta})\right]\right\},}$$

όπου $Asy\mathrm{Var}\left[\mathrm{S}(\boldsymbol{\theta})\right]=\mathrm{E}\left[\mathrm{S}(\boldsymbol{\theta})\mathrm{S}(\boldsymbol{\theta})'\right]$. Επειδή η μήτρα $\mathrm{E}\left[\mathrm{S}(\boldsymbol{\theta})\mathrm{S}(\boldsymbol{\theta})'\right]$ είναι διαγώνια και ισούται με την $\mathrm{E}\left[-\mathrm{H}(\boldsymbol{\theta})\right]$ λόγω της ισότητας της μήτρας πληροφοριών (βλ. Τμήμα 10.3), η ασυμπτωτική κατανομή του $\mathrm{S}(\boldsymbol{\theta})$ συνεπάγεται ότι αυτή του σκορ διανύσματος για το υποδιάνυσμα του $\boldsymbol{\theta}$ που περιλαμβάνει τους συντελεστές του υποδείγματος $\boldsymbol{\beta}$ δίνεται ως ως εξής:

$$\mathrm{S}(\boldsymbol{\beta}) \sim \mathrm{N}\left\{\mathbf{0}, Asy\mathrm{Var}\left[\mathrm{S}(\boldsymbol{\beta})\right]\right\} \tag{15}$$

όπου $Asy\mathrm{Var}\left(\mathrm{S}(\boldsymbol{\beta})\right)=\dfrac{1}{\sigma^2}\mathrm{E}(\mathbf{X}'\mathbf{X})$. Με βάση τη σχέση (15) αποδεικνύεται ότι η ακόλουθη τετραγωνική μορφή του σκορ διανύσματος:

$$\boxed{\mathrm{S}(\hat{\boldsymbol{\beta}}^*_{ML})'\left[Asy\mathrm{Var}\left(\mathrm{S}(\boldsymbol{\beta})\right)\right]^{-1}\mathrm{S}(\hat{\boldsymbol{\beta}}^*_{ML})} \sim \chi^2_J \tag{16}$$

κάτω από τους περιορισμούς $\mathbf{R}\,\boldsymbol{\beta}=\mathbf{r}$ ακολουθεί τη χ^2-κατανομή με J βαθμούς ελευθερίας. Η σχέση (16) αποτελεί μια πιο γενική μορφή του κριτηρίου LM. Αντικαθιστώτας σε αυτή την αναλυτική σχέση του σκορ διανύσματος για το γραμμικό υπόδειγμα $\boxed{\mathrm{S}(\hat{\boldsymbol{\beta}}^*_{ML})=\dfrac{1}{\hat{\sigma}^{*\,2}_{ML}}(\mathbf{X}'\hat{\boldsymbol{\varepsilon}}^*)}$, που εκτιμάται κάτω από τους περιορισμούς, καθώς και το συνεπή εκτιμητή της $\boxed{Asy\mathrm{Var}\left(\mathrm{S}(\boldsymbol{\beta})\right) \text{ που δίνεται ως } \dfrac{1}{\hat{\sigma}^{*\,2}_{ML}}(\mathbf{X}'\mathbf{X})}$ δίνει τον ακόλουθο τύπο του LM κριτηρίου

[16] Βλέπε Παράρτημα του κεφαλαίου για την απόδειξη τη σχέσης αυτής.

$LM = S(\hat{\beta}^*_{ML})' \cdot [AsyVar(S(\beta))]^{-1} \cdot S(\hat{\beta}^*_{ML}) \Rightarrow$

$LM = \frac{1}{\hat{\sigma}^{*2}_{ML}}(X'\hat{\varepsilon})' \cdot \left[\frac{1}{\hat{\sigma}^{*2}_{ML}} \cdot X'X\right]^{-1} \cdot \frac{1}{\hat{\sigma}^{*2}_{ML}} \cdot X'\hat{\varepsilon}^* \Rightarrow$

$$\Rightarrow \boxed{LM \equiv \frac{1}{\hat{\sigma}^{*\,2}_{ML}}(\mathbf{X}'\hat{\varepsilon}^*)'(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\hat{\varepsilon}^*),} \sim \chi^2_J \qquad (17)$$

που αναφέρεται στο γραμμικό υπόδειγμα. Επειδή $\frac{1}{\hat{\sigma}^{*\,2}_{ML}}(\mathbf{X}'\mathbf{X})$ αποτελεί συνεπή εκτιμητή της διακύμανσης του σκορ διανύσματος $S(\boldsymbol{\beta})$ κάτω από την μηδενική υπόθεση ότι ισχύουν οι περιορισμοί αποδεικνύεται ότι το κριτήριο LM ακολουθεί την ίδια κατανομή με την τετραγωνική μορφή του σκορ διανύσματος που δίνεται από τη σχέση (16). Δηλαδή, ισχύει το αποτέλεσμα

$$LM \dot{\sim} \chi^2_{(J)}.$$

Εκτός από τον τύπο που δίνεται από τη σχέση (17), ένας άλλος τύπος του LM κριτηρίου που συναντάται συχνά στη πράξη για τον έλεγχο των περιορισμών $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$ είναι ο ακόλουθος:

$$\boxed{LM = N \cdot R^2,} \qquad (18\alpha)$$

όπου R^2 αποτελεί το συντελεστή προσδιορισμού της βοηθητικής παλινδρόμησης των τιμών του διανύσματος των καταλοίπων $\hat{\varepsilon}^*$ πάνω στις K-ανεξάρτητες μεταβλητές του υποδείγματος.[17] Σε μορφή γραμμικής άλγεβρας, αυτή γράφεται ως

$$\boxed{\hat{\varepsilon}^* = \mathbf{X}\gamma + \nu.}$$

Για την απόδειξη της σχέσης (18α), γράψτε το συντελεστή R^2 της παραπάνω βοηθητικής παλινδρόμησης ως εξής:

$$R^2 = \frac{ESS}{TSS} = \frac{\hat{\gamma}'X'X\hat{\gamma}}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^*} = \frac{\hat{\varepsilon}^{*\prime}X(X'X)^{-1}(X'X)(X'X)^{-1}X'\hat{\varepsilon}^*}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^*} = \frac{\hat{\varepsilon}^{*\prime}X(X'X)^{-1}X'\hat{\varepsilon}^*}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^*},$$

καθώς $\hat{\gamma} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\hat{\varepsilon}^*$. Αντικαθιστώντας την τελευταία σχέση στη σχέση (17) και χρησιμοποιώντας τον ορισμό της διακύμανσης $\hat{\sigma}^{*\,2}_{ML} = \frac{1}{N}\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^*$ συνεπάγεται

$$(17) \Rightarrow LM = \frac{1}{\hat{\sigma}^{*2}_{ML}} \cdot \hat{\varepsilon}^{*\prime} \cdot X \cdot (X'X)^{-1} \cdot X'\hat{\varepsilon}^* \cdot \frac{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^*}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^*} = \frac{1}{\hat{\sigma}^{*2}_{ML}} \cdot R^2 \cdot \hat{\varepsilon}^{*\prime}\hat{\varepsilon}^* =$$

$$\boxed{LM = \frac{1}{\hat{\sigma}^{*\,2}_{ML}} R^2 \hat{\varepsilon}^{*\prime}\hat{\varepsilon}^* = N\,R^2,}$$

$$= N \cdot R^2.$$

$$= \frac{1}{\frac{1}{N} \cdot \cancel{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^*}} \cdot R^2 \cdot \cancel{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^*} = \underline{\underline{N \cdot R^2}}.\ \checkmark.$$

που αποδεικνύει τη σχέση (18α).

Ο τύπος του κριτηρίου LM που δίνεται από τη σχέση (18α) διευκολύνει σημαντικά την εφαρμογή αυτού στον έλεγχο υποθέσεων, καθώς απαιτεί την απλή εκτίμηση του συντελεστή προσδιορισμού R^2 της βοηθητικής παλινδρόμησης του διανύσματος των καταλοίπων $\hat{\varepsilon}^*$ του υποδείγματος που εκτιμάται κάτω από τους περιορισμούς στις K ανεξάρτητες μεταβλητές του υποδείγματος. Η ερμηνεία που δίνεται στη σχέση αυτή του κριτηρίου LM είναι επίσης πολύ απλή. Αν ισχύουν οι περιορισμοί $\mathbf{R}\,\boldsymbol{\beta} = \mathbf{r}$, τότε η τιμή του συντελεστή R^2 της παραπάνω βοηθητικής παλινδρόμησης θα πρέπει να είναι πολύ κοντά στο μηδέν και όχι στατιστικά διάφορη από την τιμή αυτή. Αυτό συμβαίνει γιατί στην περίπτωση αυτή οι τιμές του διανύσματος των καταλοίπων $\hat{\varepsilon}^*$ θα πρέπει να είναι πολύ κοντά σε εκείνες του διανύσματος χωρίς περιορισμούς, δηλαδή του $\hat{\varepsilon}$. Μάλιστα και τα δύο αυτά διανύσματα τιμών των καταλοίπων θα πρέπει να είναι ορθογώνια των ανεξάρτητων μεταβλητών του υποδείγματος, διαφορετικά δε θα ισχύουν οι γραμμικοί περιορισμοί $\mathbf{R}\,\boldsymbol{\beta} = \mathbf{r}$.

Τέλος για τη σύγκριση του κριτηρίου LM με τα άλλα δύο κριτήρια τα LR και W, που αναφέρθηκαν προηγουμένως, στη συνέχεια παρουσιάζουμε έναν άλλο αναλυτικό τύπο αυτού που δίνεται ως εξής:

$$LM = N\frac{(\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^* - \hat{\varepsilon}'\hat{\varepsilon})}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^*} \qquad (18\beta)$$

[17] Σημειώστε ότι την μορφή αυτή του LM κριτηρίου τη συναντήσαμε σε πολλούς ασυμπτωτικούς ελέγχους που παρουσιάσαμε στα Κεφάλαια 7 και 8 για τη διάγνωση ύπαρξης αυτοσυσχέτισης ή ετεροσκεδαστικότητας στο διαταρακτικό όρο ε_i.

Ο τύπος αυτός αποδεικνύεται στο παράρτημα του κεφαλαίου.

Σχέση μεταξύ των τριών κριτηρίων LR, W και LM

Χρησιμοποιώντας τις αναλυτικές σχέσεις των τριών κριτηρίων LR, W και LM για τον έλεγχο των περιορισμών $\mathbf{R}\,\boldsymbol{\beta} = \mathbf{r}$ που παρουσιάστηκαν προηγουμένως, στο τμήμα αυτό αποδεικνύουμε μια γνωστή ανισότητα που ισχύει μεταξύ των τιμών τους. Αυτή δίνεται ως εξής:

$$\boxed{\mathrm{W} \geq \mathrm{LR} \geq \mathrm{LM}} \qquad (19)$$

Αν και τα τρία αυτά κριτήρια ακολουθούν την ίδια κατανομή και είναι ασυμπτωτικά ισοδύναμα, η ανισότητα αυτή δείχνει ότι οι τιμές τους θα διαφέρουν σε πεπερασμένα δείγματα παρατηρήσεων. Αυτό θα έχει ως αποτέλεσμα η δύναμη του κριτηρίου ή το μέγεθος (σφάλμα τύπου Ι) των ελέγχων που διεξάγονται με βάση τα κριτήρια αυτά να διαφέρει σημαντικά σε μικρά δείγματα παρατηρήσεων. Πιο συγκεκριμένα, αν μας ενδιαφέρει η δύναμη του κριτηρίου ενός ελέγχου η παραπάνω ανισότητα σημαίνει ότι το κριτήριο W έχει υψηλότερη δύναμη του κριτηρίου να απορρίπτει τη μηδενική υπόθεση όταν αυτή δεν είναι αληθινή σε σχέση με τα άλλα δύο κριτήρια. Αυτό συμβαίνει γιατί οι τιμές του κριτηρίου αυτού είναι μεγαλύτερες ή ίσες των άλλων δύο και βρίσκονται πιο δεξιά στο δεξί άκρο της χ^2-κατανομής, που αποτελεί την περιοχή απόρριψης. Το ίδιο θα συμβαίνει και για το κριτήριο LR σε σχέση με το LM.

Για την απόδειξη της ανισότητας (19), θα χρησιμοποιήσουμε τους δύο πρώτους όρους του αναπτύγματος κατά Taylor της συνάρτησης ln(1+z), για κάποια μεταβλητή z. Σύμφωνα με το ανάπτυγμα αυτό, η συνάρτηση ln(1+z) μπορεί να προσεγγιθεί ως

$$\ln(1+z) = z - \frac{1}{2}z^2 \qquad (20)$$

Θέτοντας στο παραπάνω ανάπτυγμα $z = \dfrac{(\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*} - \hat{\varepsilon}'\hat{\varepsilon})}{\hat{\varepsilon}'\hat{\varepsilon}}$, η σχέση (10β) του κριτηρίου LR συνεπάγεται

$$LR = N \ln\left[1+\frac{(\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*} - \hat{\varepsilon}'\hat{\varepsilon})}{\hat{\varepsilon}'\hat{\varepsilon}}\right] = \frac{N(\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*} - \hat{\varepsilon}'\hat{\varepsilon})}{\hat{\varepsilon}'\hat{\varepsilon}} - \frac{N}{2}\left[\frac{(\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*} - \hat{\varepsilon}'\hat{\varepsilon})}{\hat{\varepsilon}'\hat{\varepsilon}}\right]^2 = W - \frac{N}{2}\left[\frac{(\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*} - \hat{\varepsilon}'\hat{\varepsilon})}{\hat{\varepsilon}'\hat{\varepsilon}}\right]^2$$

Επειδή $\left[\dfrac{(\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*} - \hat{\varepsilon}'\hat{\varepsilon})}{\hat{\varepsilon}'\hat{\varepsilon}}\right]^2 \geq 0$, η τελευταία σχέση συνεπάγεται ότι $W \geq LR$. Για να αποδείξουμε το δεύτερο σκέλος της ανισότητας (19) (δηλαδή ότι $LR \geq LM$), θα χρησιμοποιήσουμε το ανάπτυγμα (19) θεωρώντας ότι $z = \dfrac{(\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*} - \hat{\varepsilon}'\hat{\varepsilon})}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*}}$. Τότε, η σχέση (8c) για το κριτήριο LR συνεπάγεται

$$LR = N \ln\left[\frac{1}{1+\dfrac{(\hat{\varepsilon}'\hat{\varepsilon} - \hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*})}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*}}}\right] = -N \ln\left[1+\frac{(\hat{\varepsilon}'\hat{\varepsilon} - \hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*})}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*}}\right] = -N\ln\left[1-\frac{(\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*} - \hat{\varepsilon}'\hat{\varepsilon})}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*}}\right]$$

$$= N\left[\frac{(\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*} - \hat{\varepsilon}'\hat{\varepsilon})}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*}}\right] + \frac{N}{2}\left[\frac{(\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*} - \hat{\varepsilon}'\hat{\varepsilon})}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*}}\right]^2 = LM + \frac{N}{2}\left[\frac{(\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*} - \hat{\varepsilon}'\hat{\varepsilon})}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*}}\right]^2.$$

Από την τελευταία σχέση συνεπάγεται ότι $LR \geq LM$, καθώς ισχύει $\left[\dfrac{(\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*} - \hat{\varepsilon}'\hat{\varepsilon})}{\hat{\varepsilon}^{*\prime}\hat{\varepsilon}^{*}}\right]^2 \geq 0.$