Section 6 :
Tests d'hypothéses linéaires (bis)

STAT-F-406

Master en sciences mathématiques, Master en statistique

Davy Paindaveine

Université libre de Bruxelles

2023–2024

Contenu du chapitre

Tests d'hypothèses linéaires sur β

Dans le modèle semi-fort homoscédastique, reconsidérons le problème de test (au niveau $\alpha \in]0,1[$)

$$\begin{cases} \mathcal{H}_0 : R\beta = r \\ \mathcal{H}_1 : R\beta \neq r, \end{cases}$$

où la matrice R de rang maximal p et le p-vecteur r sont fixés.

On a vu que, sous $\mathcal{H}_0$,

$$W \stackrel{\text{def}}{=} \frac{n}{\hat{\sigma}^2}(R\hat{\beta} - r)'(RQ^{-1}R')^{-1}(R\hat{\beta} - r) \xrightarrow{\mathcal{D}} \chi^2_p$$

(on écrira W plutôt que W_{asympt} ici). Le test associé rejette donc $\mathcal{H}_0$ au niveau asymptotique α si et seulement si $W > \chi^2_{p,1-\alpha}$.

Estimateur des moindres carrés contraint

Considérons l'espace paramétrique

$$\Theta = \{(\beta, \sigma^2) \in \mathbb{R}^k \times \mathbb{R}_0^+\}$$

et son équivalent contraint

$$\Theta_0 = \{(\beta, \sigma^2) \in \Theta : R\beta = r\},$$

qui sont de dimension $k+1$ et $k+1-p$, respectivement.

L'estimateur des moindres carrés contraint

$$\tilde{\beta} \stackrel{\text{def}}{=} \arg\min_{\beta : R\beta = r} \sum_{i=1}^{n} (Y_i - \beta' X_i)^2$$

va jouer un rôle important dans cette section.

Estimateur des moindres carrés contraint

On peut obtenir une expression explicite de $\tilde{\beta}$, en déterminant les points critiques de la Lagrangienne

$$L(\beta, \lambda) = \sum_{i=1}^{n} (Y_i - \beta' X_i)^2 + \lambda'(R\beta - r),$$

où $\lambda \in \mathbb{R}^p$. Puisque

$$\begin{aligned}
\nabla_\beta L(\beta, \lambda) &= -2 \sum_{i=1}^{n} (Y_i - \beta' X_i) X_i + R'\lambda \\
&= -2 \sum_{i=1}^{n} X_i Y_i + 2 \sum_{i=1}^{n} X_i X_i' \beta + R'\lambda \\
&= -2n\Big((\tfrac{1}{n} \textstyle\sum_{i=1}^{n} X_i Y_i) - Q\beta - \tfrac{1}{2n} R'\lambda \Big) \\
&= -2nQ\big(\hat{\beta} - \beta - \tfrac{1}{2n} Q^{-1} R'\lambda\big),
\end{aligned}$$

Estimateur des moindres carrés contraint

on doit avoir

$$\tilde{\beta} = \hat{\beta} - \tfrac{1}{2n} Q^{-1} R' \lambda.$$

En imposant la condition $0 = \nabla_\lambda L(\beta, \lambda) = R\beta - r$, on obtient

$$R\tilde{\beta} = R\hat{\beta} - \tfrac{1}{2n} R Q^{-1} R' \lambda = r,$$

donc

$$\tfrac{1}{2n}\lambda = (RQ^{-1}R')^{-1}(R\hat{\beta} - r).$$

Ceci livre donc l'unique solution

$$\tilde{\beta} = \hat{\beta} - Q^{-1}R'(RQ^{-1}R')^{-1}(R\hat{\beta} - r),$$

qui doit être l'unique minimum global (pourquoi?)

Estimateur des moindres carrés contraint

De même qu'à l'estimateur $\hat{\beta}$ est associé l'estimateur de la variance

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{\beta}' X_i)^2,$$

à l'estimateur contraint $\tilde{\beta}$ est associé l'estimateur contraint de la variance

$$\tilde{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \tilde{\beta}' X_i)^2.$$

Notons que $\tilde{\sigma}^2 \geq \hat{\sigma}^2$ p.s.

Sur le test de Wald

Une autre manière de montrer que $\tilde{\sigma}^2 \geq \hat{\sigma}^2$ est d'écrire

$$
\begin{aligned}
n(\tilde{\sigma}^2 - \hat{\sigma}^2) &= \textstyle\sum_{i=1}^n \{(Y_i - \tilde{\beta}' X_i)^2 - (Y_i - \hat{\beta}' X_i)^2\} \\
&= \textstyle\sum_{i=1}^n (\hat{\beta} - \tilde{\beta})' X_i (2Y_i - (\tilde{\beta} + \hat{\beta})' X_i) \\
&= (\hat{\beta} - \tilde{\beta})' \textstyle\sum_{i=1}^n X_i (2Y_i - X_i' (\tilde{\beta} + \hat{\beta})) \\
&= n(\hat{\beta} - \tilde{\beta})' (2Q\hat{\beta} - Q(\tilde{\beta} + \hat{\beta})) \\
&= n(\hat{\beta} - \tilde{\beta})' Q (\hat{\beta} - \tilde{\beta})
\end{aligned}
$$

(pourquoi?) Ceci montre aussi que les valeurs minimales contraintes et non contraintes sont proches ssi les minimas correspondants le sont.

Sur le test de Wald

Puisque $\tilde{\beta} = \hat{\beta} - Q^{-1}R'(RQ^{-1}R')^{-1}(R\hat{\beta} - r)$, on a de plus

$$\begin{aligned}
&n(\hat{\beta} - \tilde{\beta})'Q(\hat{\beta} - \tilde{\beta}) \\
&\quad = n\{Q^{-1}R'(RQ^{-1}R')^{-1}(R\hat{\beta} - r)\}'Q\{Q^{-1}R'(RQ^{-1}R')^{-1}(R\hat{\beta} - r)\} \\
&\quad = n(R\hat{\beta} - r)'(RQ^{-1}R')^{-1}RQ^{-1}QQ^{-1}R'(RQ^{-1}R')^{-1}(R\hat{\beta} - r) \\
&\quad = n(R\hat{\beta} - r)'(RQ^{-1}R')^{-1}(R\hat{\beta} - r) = \hat{\sigma}^2 W,
\end{aligned}$$

de sorte que

$$W = \frac{n}{\hat{\sigma}^2}(\hat{\beta} - \tilde{\beta})'Q(\hat{\beta} - \tilde{\beta}) = \frac{n(\tilde{\sigma}^2 - \hat{\sigma}^2)}{\hat{\sigma}^2},$$

ce qui donne deux nouvelles interprétations à la statistique de Wald W.

Test du rapport de vraisemblance gaussien

Plaçons-nous dans le cadre du modèle fort avec normalité.

Le test du rapport de vraisemblance pour

$$\begin{cases} \mathcal{H}_0 : \theta \in \Theta_0 = \{(\beta, \sigma^2) \in \mathbb{R}^k \times \mathbb{R}_0^+ : R\beta = r\} \\ \mathcal{H}_1 : \theta \in \Theta \setminus \Theta_0 = \{(\beta, \sigma^2) \in \mathbb{R}^k \times \mathbb{R}_0^+ : R\beta \neq r\} \end{cases}$$

rejette $\mathcal{H}_0$ au niveau asymptotique α si et seulement si

$$-2 \ln \Lambda = -2 \ln \frac{L_{\tilde{\theta}}^{(n)}}{L_{\hat{\theta}}^{(n)}} > \chi^2_{s,1-\alpha},$$

où $s = (k+1) - (k+1-p) = p$ et où $\tilde{\theta}$ et $\hat{\theta}$ sont respectivement les MLEs gaussiens contraints et non contraints de θ.

Test du rapport de vraisemblance gaussien

Rappelons que la vraisemblance gaussienne est donnée par

$$L_\theta = L_{\beta,\sigma^2} = \big(\tfrac{1}{\sqrt{2\pi\sigma^2}}\big)^n \exp\big(-\tfrac{1}{2\sigma^2}\textstyle\sum_{i=1}^n (Y_i-\beta' X_i)^2\big)\big(\prod_{i=1}^n f^{X_i}(X_i)\big).$$

On vérifiera aisément (comment?) que $\tilde\theta = (\tilde\beta, \tilde\sigma^2)$ et $\hat\theta = (\hat\beta, \hat\sigma^2)$, où $\tilde\beta$, $\tilde\sigma^2$, $\hat\beta$ et $\hat\sigma^2$ sont les estimateurs considérés aux pages précédentes.

On a donc

$$\begin{aligned}
-2\ln\Lambda &= -2\ln \frac{\big(\tfrac{1}{\sqrt{2\pi\tilde\sigma^2}}\big)^n \exp\big(-\tfrac{1}{2\tilde\sigma^2}\sum_{i=1}^n (Y_i-\tilde\beta' X_i)^2\big)\big(\prod_{i=1}^n f^{X_i}(X_i)\big)}{\big(\tfrac{1}{\sqrt{2\pi\hat\sigma^2}}\big)^n \exp\big(-\tfrac{1}{2\hat\sigma^2}\sum_{i=1}^n (Y_i-\hat\beta' X_i)^2\big)\big(\prod_{i=1}^n f^{X_i}(X_i)\big)} \\
&= -2\ln \frac{\big(\tfrac{1}{\sqrt{\tilde\sigma^2}}\big)^n \exp\big(-\tfrac{n}{2}\big)}{\big(\tfrac{1}{\sqrt{\hat\sigma^2}}\big)^n \exp\big(-\tfrac{n}{2}\big)} = -2\ln\left(\frac{\tilde\sigma^2}{\hat\sigma^2}\right)^{-n/2} = n\ln\frac{\tilde\sigma^2}{\hat\sigma^2}.
\end{aligned}$$

Test du rapport de vraisemblance gaussien

Le test du rapport de vraisemblance gaussien rejette donc $\mathcal{H}_0$ au niveau asymptotique α si et seulement si

$$-2 \ln \Lambda = n \ln \frac{\tilde{\sigma}^2}{\hat{\sigma}^2} > \chi^2_{p,1-\alpha}.$$

Par le théorème de Wilks, ce test est asymptotiquement de niveau α dans le modèle fort avec normalité.

Qu'en est-il dans le modèle semi-fort homoscédastique?

Test des multiplicateurs de Lagrange

Pour un problème de test générique du type

$$\begin{cases} \mathcal{H}_0 : \theta \in \Theta_0 \\ \mathcal{H}_1 : \theta \in \Theta \setminus \Theta_0, \end{cases}$$

le test des multiplicateurs de Lagrange rejette $\mathcal{H}_0$ pour de grandes valeurs de la statistique

$$\|(\nabla_\theta \ln L_\theta)|_{\theta=\tilde{\theta}}\|_\Sigma^2,$$

où $\tilde{\theta}$ est le MLE contraint de θ et où $\|z\|_\Sigma^2 = z'\Sigma^{-1}z$ est fondée sur une matrice de variance-covariance Σ livrant une loi asymptotique standard.

Test des multiplicateurs de Lagrange gaussien

Plaçons-nous de nouveau dans le cadre du modèle fort avec normalité.

Puisque

$$L_\theta = L_{\beta,\sigma^2} = \big(\tfrac{1}{\sqrt{2\pi\sigma^2}}\big)^n \exp\big(- \tfrac{1}{2\sigma^2} \textstyle\sum_{i=1}^n (Y_i - \beta' X_i)^2\big)\big(\prod_{i=1}^n f^{X_i}(X_i)\big),$$

on a alors

$$\begin{aligned} \nabla_\beta \ln L_\theta &= \nabla_\beta \big(- \tfrac{1}{2\sigma^2} \textstyle\sum_{i=1}^n (Y_i - \beta' X_i)^2\big) \\ &= \tfrac{1}{\sigma^2} \textstyle\sum_{i=1}^n (Y_i - \beta' X_i) X_i. \end{aligned}$$

Ici, la statistique que nous utiliserons est alors de la forme

$$\|(\nabla_\beta \ln L_\theta)|_{\theta=\tilde{\theta}}\|_\Sigma^2$$

plutôt que $\|(\nabla_\theta \ln L_\theta)|_{\theta=\tilde{\theta}}\|_\Sigma^2$ (pourquoi ne pas inclure $(\nabla_{\sigma^2} \ln L_\theta)|_{\theta=\tilde{\theta}}$?)

Quelle matrice Σ utiliser?

Dans le modèle fort avec normalité,

$$\frac{1}{\sqrt{n}} \nabla_\beta \ln L_\theta = \frac{1}{\sqrt{n}\sigma^2} \sum_{i=1}^n (Y_i - \beta' X_i) X_i \xrightarrow{\mathcal{D}} \mathcal{N}_k(0, V),$$

où

$$V = \mathrm{Var}\left[\tfrac{1}{\sigma^2} \varepsilon X\right] = \tfrac{1}{\sigma^4} \mathrm{E}[\varepsilon^2 X X'] = \tfrac{1}{\sigma^4} \mathrm{E}[\mathrm{E}[\varepsilon^2 | X] X X'] = \tfrac{1}{\sigma^2} \mathrm{E}[X X'].$$

Ceci suggère de considérer

$$\left(\tfrac{1}{\sqrt{n}} \nabla_\beta \ln L_\theta\right)' \left(\tfrac{1}{\sigma^2} \mathrm{E}[X X']\right)^{-1} \left(\tfrac{1}{\sqrt{n}} \nabla_\beta \ln L_\theta\right) = \|\nabla_\beta \ln L_\theta\|^2_{\Sigma = nV},$$

qui suit asymptotiquement une loi standard (la loi χ^2_k).

Test des multiplicateurs de Lagrange gaussien

La statistique de test naturelle qui en résulte est donc

$$
\begin{aligned}
&\Big(\tfrac{1}{\sqrt{n}}(\nabla_\beta \ln L_\theta)|_{\theta=\tilde\theta}\Big)'\Big(\tfrac{1}{\tilde\sigma^2}Q\Big)^{-1}\Big(\tfrac{1}{\sqrt{n}}(\nabla_\beta \ln L_\theta)|_{\theta=\tilde\theta}\Big)\\
&= \Big(\tfrac{1}{\sqrt{n}\tilde\sigma^2}\textstyle\sum_{i=1}^n (Y_i-\tilde\beta' X_i)X_i\Big)'\Big(\tfrac{1}{\tilde\sigma^2}Q\Big)^{-1}\Big(\tfrac{1}{\sqrt{n}\tilde\sigma^2}\textstyle\sum_{i=1}^n (Y_i-\tilde\beta' X_i)X_i\Big)\\
&= \tfrac{n}{\tilde\sigma^2}\Big(\tfrac{1}{n}\textstyle\sum_{i=1}^n (Y_i-\tilde\beta' X_i)X_i\Big)'Q^{-1}\Big(\tfrac{1}{n}\textstyle\sum_{i=1}^n (Y_i-\tilde\beta' X_i)X_i\Big)\\
&= \tfrac{n}{\tilde\sigma^2}(Q\hat\beta - Q\tilde\beta)'Q^{-1}(Q\hat\beta - Q\tilde\beta) = \tfrac{n}{\tilde\sigma^2}(\hat\beta-\tilde\beta)'Q(\hat\beta-\tilde\beta) = \tfrac{n(\tilde\sigma^2-\hat\sigma^2)}{\tilde\sigma^2}.
\end{aligned}
$$

Par le lemme de Slutzky, on a que, sous $\mathcal{H}_0$,

$$
\frac{n(\tilde\sigma^2-\hat\sigma^2)}{\tilde\sigma^2} = \frac{\hat\sigma^2}{\tilde\sigma^2}\times\frac{n(\tilde\sigma^2-\hat\sigma^2)}{\hat\sigma^2} \xrightarrow{\mathcal{D}} \chi^2_p,
$$

et ce même dans le modèle semi-fort homoscédastique! Le test qui en résulte rejette donc $\mathcal{H}_0$ au niveau asymptotique α ssi $\frac{n(\tilde\sigma^2-\hat\sigma^2)}{\tilde\sigma^2} > \chi^2_{p,1-\alpha}$.

Relation entre les trois tests

Les trois tests rejettent $\mathcal{H}_0$ au niveau asymptotique α lorsque

$$LM = \frac{n(\tilde{\sigma}^2 - \hat{\sigma}^2)}{\tilde{\sigma}^2} = n(1 - (\tfrac{\tilde{\sigma}^2}{\hat{\sigma}^2})^{-1}) > \chi^2_{p,1-\alpha},$$

$$LR = n \ln \tfrac{\tilde{\sigma}^2}{\hat{\sigma}^2} > \chi^2_{p,1-\alpha},$$

$$W = \frac{n(\tilde{\sigma}^2 - \hat{\sigma}^2)}{\hat{\sigma}^2} = n(\tfrac{\tilde{\sigma}^2}{\hat{\sigma}^2} - 1) > \chi^2_{p,1-\alpha}.$$

Puisque $1 - x^{-1} \leq \ln x \leq x - 1$ pour tout $x \geq 1$, on a

$$LM \leq LR \leq W \quad \text{p.s.}$$

Par conséquent, le test LR est également valide dans le modèle semi-fort homoscédastique!

Les courbes de puissances sont aussi dans l'ordre ci-dessus.