Test shody dvou regresních přímek

David Hampel

Ústav statistiky a operačního výzkumu,
Mendelova univerzita v Brně

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Kurz pokročilých statistických metod
Global Change Research Centre AS CR, 5.–7. 8. 2015

Tato akce se koná v rámci projektu: Vybudování vědeckého týmu environmentální metabolomiky a ekofyziologie a jeho zapojení do mezinárodních sítí (ENVIMET; r.č. **CZ.1.07/2.3.00/20.0246**) realizovaného v rámci Operačního programu Vzdělávání pro konkurenceschopnost

Obsah

Motivace

- Otázka, zda mají dvě regresní funkce stejné všechny nebo alespoň některé parametry je relativně častá.
- Jednoduché přístupy jako t-test nelze použít.
- Testování je možné pomocí obecného přístupu, tzv. obecné lineární hypotézy, pomocí níž lze testovat shodu i více než dvou regresních funkcí.

Maticový zápis modelu vícenásobné regrese

Model vícenásobné regrese můžeme vyjádřit jako

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

kde:

$\mathbf{Y}$ – náhodný vektor pozorování závisle proměnné;

$\mathbf{X}$ – matice pozorování nezávisle proměnných;

$\boldsymbol{\beta}$ – vektor neznámých parametrů;

$\boldsymbol{\epsilon}$ – vektor náhodných chyb.

Maticový zápis modelu vícenásobné regrese

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{p1} \\ 1 & X_{12} & X_{22} & \cdots & X_{p2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{pn} \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

Obecná lineární hypotéza

- Obecnou lineární hypotézu zapisujeme symbolicky jako

$$\mathrm{H}_0 : \quad \mathbf{A}\boldsymbol{\beta} = \mathbf{a}, \tag{1}$$

kde $\mathbf{A}$ je libovolná reálná matice typu $m \times k$ a hodnosti $m \leq k$.

- Dále, $\mathbf{a}$ je vektor typu $m \times 1$ pro který má rovnice (1) řešení.
- Testovací statistika pro obecnou lineární hypotézu je ve tvaru

$$F = \frac{1}{ms^2}(\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{a})'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1}(\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{a}), \tag{2}$$

kde s^2 je odhad reziduálního rozptylu.

- Za platnosti H_0 má statistika F Fisherovo-Snedecorovo rozdělení se stupni volnosti m a $n - k$.
- Hypotézu H_0 zamítáme, pokud $F > F_{1-\alpha}(m; n-k)$.

Test shody parametrů dvou přímek

Rovnice první přímky:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, \quad i = 1, \ldots, n_1$$

Rovnice druhé přímky:

$$Y_i^* = \beta_0^* + \beta_1^* X_i^* + \epsilon_i^*, \quad i = 1, \ldots, n_2$$

Maticový zápis modelů

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_{n_1} \end{bmatrix} = \begin{bmatrix} 1 & X_1 \\ \vdots & \vdots \\ 1 & X_{n_1} \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{n_1} \end{bmatrix}$$

$$\begin{bmatrix} Y_1^* \\ \vdots \\ Y_{n_2}^* \end{bmatrix} = \begin{bmatrix} 1 & X_1^* \\ \vdots & \vdots \\ 1 & X_{n_2}^* \end{bmatrix} \cdot \begin{bmatrix} \beta_0^* \\ \beta_1^* \end{bmatrix} + \begin{bmatrix} \epsilon_1^* \\ \vdots \\ \epsilon_{n_2}^* \end{bmatrix}$$

Maticový zápis společného modelu

$$
\begin{bmatrix} Y_1 \\ \vdots \\ Y_{n_1} \\ Y_1^* \\ \vdots \\ Y_{n_2}^* \end{bmatrix} = \begin{bmatrix} 1 & X_1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{n_1} & 0 & 0 \\ 0 & 0 & 1 & X_1^* \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & X_{n_2}^* \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_0^* \\ \beta_1^* \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{n_1} \\ \epsilon_1^* \\ \vdots \\ \epsilon_{n_2}^* \end{bmatrix}
$$

Možnosti testování

- Test shody konstant:

$$H_0: \quad \beta_0 = \beta_0^*$$

$$\beta_0 - \beta_0^* = 0$$

Maticový zápis:

$$\begin{bmatrix} 1 & 0 & -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_0^* \\ \beta_1^* \end{bmatrix} = 0$$

$$\mathbf{a} = 0, \quad \mathbf{A} = \begin{bmatrix} 1 & 0 & -1 & 0 \end{bmatrix}$$

Možnosti testování

- Test shody směrnic:

$$H_0: \quad \beta_1 = \beta_1^*$$

$$\beta_1 - \beta_1^* = 0$$

Maticový zápis:

$$\begin{bmatrix} 0 & 1 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_0^* \\ \beta_1^* \end{bmatrix} = 0$$

$$\mathbf{a} = 0, \quad \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & -1 \end{bmatrix}$$

Možnosti testování

- Test shody konstant i směrnic:

$$H_0: \quad \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} \beta_0^* \\ \beta_1^* \end{bmatrix}$$

$$\begin{bmatrix} \beta_0 - \beta_0^* \\ \beta_1 - \beta_1^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Maticový zápis:

$$\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_0^* \\ \beta_1^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\mathbf{a} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}$$

Příklad

- Data: `dve_primky_dohromady.sta`
- Odhady parametrů:
  Kosnt_upper: $\hat{\beta}_0 = 1,12$
  Narea_upper: $\hat{\beta}_1 = 1,89$
  Konst_lower: $\hat{\beta}_0^* = -8,49$
  Narea_lower: $\hat{\beta}_1^* = 6,10$
- Matice $(\mathbf{X}'\mathbf{X})^{-1}$:

$$\begin{bmatrix} 6,58 & -1,52 & 0,00 & 0,00 \\ -1,52 & 0,36 & 0,00 & 0,00 \\ 0,00 & 0,00 & 3,31 & -1,18 \\ 0,00 & 0,00 & -1,18 & 0,43 \end{bmatrix}$$

- $s^2 = 4,15$

Příklad

- Testovací statistika:

$$F = \frac{1}{ms^2}(\mathbf{A}\hat{\beta} - \mathbf{a})'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1}(\mathbf{A}\hat{\beta} - \mathbf{a})$$

- Test shody směrnic:

$$\mathbf{a} = 0, \quad \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & -1 \end{bmatrix}$$

- Počet řádků matice **X**: $n = 26$
- Počet řádků matice **A**: $m = 1$
- Počet sloupců matice **A**: $k = 4$

Příklad

$$\mathbf{A}\hat{\beta} - \mathbf{a} = \begin{bmatrix} 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1{,}12 \\ 1{,}89 \\ -8{,}49 \\ 6{,}10 \end{bmatrix} - 0$$

$$= 1{,}89 - 6{,}10 = -4{,}21$$

Příklad

$$\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}' =$$

$$\begin{bmatrix} 0 & 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 6{,}58 & -1{,}52 & 0{,}00 & 0{,}00 \\ -1{,}52 & 0{,}36 & 0{,}00 & 0{,}00 \\ 0{,}00 & 0{,}00 & 3{,}31 & -1{,}18 \\ 0{,}00 & 0{,}00 & -1{,}18 & 0{,}43 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix} =$$

$$\begin{bmatrix} -1{,}52 & 0{,}36 & 1{,}18 & -0{,}43 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ -1 \end{bmatrix} =$$

$$0{,}36 + 0{,}43 = 0{,}79$$

Příklad

$$[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1} = 0{,}79^{-1} = \frac{1}{0{,}79} = 1{,}27$$

Příklad

$$F = \frac{1}{ms^2}(\mathbf{A}\hat{\beta} - \mathbf{a})'[\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}']^{-1}(\mathbf{A}\hat{\beta} - \mathbf{a}) =$$

$$\frac{1}{1 \cdot 4{,}15}(-4{,}21) \cdot 1{,}27 \cdot (-4{,}21) = 5{,}42$$

$$F_{1-\alpha}(m; n-k) = F_{0{,}95}(1; 26-4) = F_{0{,}95}(1; 22) = 4{,}30$$

$$5{,}42 = F > F_{1-\alpha}(m; n-k) = 4{,}30$$

$\Rightarrow$ H_0 o shodě směrnic zamítáme