Cvičení z NSTP022
13. týden cvičení (20. – 24. 5. 2013)

Neymanova-Pearsonova věta, Regrese

1. Nechť $X_1, \ldots, X_n$ je náhodný výběr z alternativního rozdělení s parametrem $p \in (0,1)$. Najděte pomocí Neymanovy-Pearsonovy věty kritický obor pro test hypotézy $H_0 : p = p_0$ proti alternativě $H_1 : p = p_1 > p_0$, kde $p_0, p_1 \in (0,1)$ jsou známé konstanty.

2. Nechť $X_1, \ldots, X_n$ je náhodný výběr z $\mathsf{N}(0, \sigma^2)$. Najděte pomocí Neymanovy-Pearsonovy věty kritický obor pro test hypotézy $H_0 : \sigma = \sigma_0$ proti alternativě $H_1 : \sigma = \sigma_1 < \sigma_0$, kde $\sigma_0, \sigma_1 > 0$ jsou známé konstanty.

3. Nechť naše pozorování $Y_1, \ldots, Y_n$ splňují model
$$Y_i = \beta_1 + \beta_2\, x_i + e_i, \qquad i = 1\ldots, n,$$
kde $x_1, x_2, \ldots, x_n$ jsou známé konstanty, $e_1, \ldots, e_n$ jsou nezávislé stejně rozdělené náhodné veličiny s nulovou střední hodnotou a konečným rozptylem σ^2. Uvažujme následující odhady neznámých parametrů β_1 a β_2:
$$\hat\beta_2 = \frac{\sum_{i=1}^n (x_i - \bar x_n) Y_i}{\sum_{i=1}^n (x_i - \bar x_n)^2}, \qquad \hat\beta_1 = \bar Y_n - \hat\beta_2\, \bar x_n,$$
kde $\bar x_n = \frac{1}{n}\sum_{i=1}^n x_i$. Ukažte, že odhady $\hat\beta_1$ a $\hat\beta_2$ jsou nestranné a spočtěte rozptyl odhadu $\hat\beta_2$.

4. Nechť $Y_1, Y_2, \ldots, Y_n$ jsou nezávislé náhodné veličiny a Y_i má normální rozdělení $\mathsf{N}(\beta_1 + \beta_2\, x_i, \sigma^2)$, kde $x_1, x_2, \ldots, x_n$ jsou známé konstanty a také σ^2 je známé. Odvoďte maximálně věrohodné odhady parametrů β_1 a β_2.

Opakování z přednášky

Neymanova-Pearsonova věta. Nechť $\mathbf{x} = (x_1, \ldots, x_n)$ a $p_0(\mathbf{x})$, $p_1(\mathbf{x})$ jsou nezáporné měřitelné funkce na $(\mathbb{R}^n, \mathcal{B}^n)$ takové, že
$$\int_{\mathbb{R}^n} p_0(\mathbf{x})\, d\nu_n(\mathbf{x}) = \int_{\mathbb{R}^n} p_1(\mathbf{x})\, d\nu_n(\mathbf{x}) = 1, \tag{13.1}$$
kde ν_n je σ-konečná míra na $(\mathbb{R}^n, \mathcal{B}^n)$. Nechť pro dané $\alpha \in (0,1)$ existuje takové kladné číslo c, že pro množinu
$$W^* = \{\mathbf{x} \in \mathbb{R}^n : p_1(\mathbf{x}) \geq c\, p_0(\mathbf{x})\}$$
platí
$$\int_{W^*} p_0(\mathbf{x})\, d\nu_n(\mathbf{x}) = \alpha. \tag{13.2}$$
Pak pro libovolnou množinu $W \in \mathcal{B}^n$ splňující
$$\int_{W} p_0(\mathbf{x})\, d\nu_n(\mathbf{x}) \leq \alpha \tag{13.3}$$
platí
$$\int_{W^*} p_1(\mathbf{x})\, d\nu_n(\mathbf{x}) \geq \int_{W} p_1(\mathbf{x})\, d\nu_n(\mathbf{x}). \tag{13.4}$$

Aplikace v testování hypotéz. Nechť $X_1, \ldots, X_n$ je náhodný výběr z rozdělení s hustotou $f(x, \theta)$ vzhledem k nějaké σ-konečné míře ν. Chceme testovat hypotézu $H_0 : \theta = \theta_0$ proti alternativě $H_1 : \theta = \theta_1$, kde $\theta_1 \neq \theta_0$. Potom test s kritickým oborem
$$W^* = \left\{ (x_1, \ldots, x_n) : \prod_{i=1}^n f(x_i, \theta_1) \geq c \prod_{i=1}^n f(x_i, \theta_0) \right\}, \tag{13.5}$$
kde
$$\int_{W^*} \prod_{i=1}^n f(x_i, \theta_0)\, d\nu(x_1) \ldots d\nu(x_n) = \alpha \tag{13.6}$$
má nejmenší pravděpodobnost chyby 2. druhu mezi všemi testy na hladině α.

Výsledky

1. $\sum_{i=1}^{n} X_i \geq K$
2. $\sum_{i=1}^{n} X_i^2 \leq K$
3. $\mathsf{var}\,(\hat{\beta}_2) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \bar{x}_n)^2}$.
4. $\hat{\beta}_2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x}_n) Y_i}{\sum_{i=1}^{n}(x_i - \bar{x}_n)^2}, \qquad \hat{\beta}_1 = \bar{Y}_n - \hat{\beta}_2\, \bar{x}_n.$