

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

**Neymanova-Pearsonova věta, Regrese**

1. Nechť  $X_1, \dots, X_n$  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  $X_1, \dots, X_n$  je náhodný výběr z  $N(0, \sigma^2)$ . 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  $Y_1, \dots, Y_n$  splňují model

$$Y_i = \beta_1 + \beta_2 x_i + e_i, \quad i = 1, \dots, n,$$

kde  $x_1, x_2, \dots, x_n$  jsou známé konstanty,  $e_1, \dots, e_n$  jsou nezávislé stejně rozdělené náhodné veličiny s nulovou střední hodnotou a konečným rozptylem  $\sigma^2$ . Uvažujme následující odhady neznámých parametrů  $\beta_1$  a  $\beta_2$ :

$$\hat{\beta}_2 = \frac{\sum_{i=1}^n (x_i - \bar{x}_n) Y_i}{\sum_{i=1}^n (x_i - \bar{x}_n)^2}, \quad \hat{\beta}_1 = \bar{Y}_n - \hat{\beta}_2 \bar{x}_n,$$

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 spočítejte rozptyl odhadu  $\hat{\beta}_2$ .

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

**Opakování z přednášky**

**Neymanova-Pearsonova věta.** Nechť  $\mathbf{x} = (x_1, \dots, x_n)$  a  $p_0(\mathbf{x}), p_1(\mathbf{x})$  jsou nezáporné měřitelné funkce na  $(\mathbb{R}^n, \mathcal{B}^n)$  takové, že

$$\int_{\mathbb{R}^n} p_0(\mathbf{x}) d\nu_n(\mathbf{x}) = \int_{\mathbb{R}^n} p_1(\mathbf{x}) d\nu_n(\mathbf{x}) = 1, \quad (13.1)$$

kde  $\nu_n$  je  $\sigma$ -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

$$\int_{W^*} p_0(\mathbf{x}) d\nu_n(\mathbf{x}) = \alpha. \quad (13.2)$$

Pak pro libovolnou množinu  $W \in \mathcal{B}^n$  splňující

$$\int_W p_0(\mathbf{x}) d\nu_n(\mathbf{x}) \leq \alpha \quad (13.3)$$

platí

$$\int_{W^*} p_1(\mathbf{x}) d\nu_n(\mathbf{x}) \geq \int_W p_1(\mathbf{x}) d\nu_n(\mathbf{x}). \quad (13.4)$$

**Aplikace v testování hypotéz.** Nechť  $X_1, \dots, X_n$  je náhodný výběr z rozdělení s hustotou  $f(x, \theta)$  vzhledem k nějaké  $\sigma$ -konečné míře  $\nu$ . 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

$$W^* = \left\{ (x_1, \dots, x_n) : \prod_{i=1}^n f(x_i, \theta_1) \geq c \prod_{i=1}^n f(x_i, \theta_0) \right\}, \quad (13.5)$$

kde

$$\int_{W^*} \prod_{i=1}^n f(x_i, \theta_0) d\nu(x_1) \dots d\nu(x_n) = \alpha \quad (13.6)$$

má nejmenší pravděpodobnost chyby 2. druhu mezi všemi testy na hladině  $\alpha$ .

## Výsledky

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