3 Postačující statistika, monotónní poměr věrohodnosti a rovnoměrně nejsilnější test

Postačující statistika

- **Def. 1.** Libovolná funkce $T(.) : \mathcal{Y} \to \mathbb{R}^r$ pro nějaké $r \in \mathbb{N}^+$ náhodného výběru $(X_1, \ldots, X_n)^T$, kde $T = T(X_1, \ldots, X_n)$ nezávisí na parametru θ se nazývá **statistika**. Hodnota $t = T(x_1, \ldots, x_n)$ příslušící realizacím $x_1, \ldots, x_n$ se nazývá *realizace statistiky*.
- **Def. 2.** Statistika $T(\mathbf{X})$ se nazývá *postačující*, pokud podmíněné rozdělení $\mathbf{X}$ při dané $T(\mathbf{X})$ nezávisí na θ.
- Předcházející definice říká, že pokud bude $T(\mathbf{X})$ postačující statistika pro θ, potom každá inference o θ závisí na $\mathbf{X}$ pouze přes hodnotu $T(\mathbf{X})$, tj., pokud $\mathbf{x}$ a $\mathbf{y}$ jsou dvě realizace a platí, že $T(\mathbf{x}) = T(\mathbf{y})$, potom inference o θ bude stejná nezávisle na tom, zda jsme pozorovali $\mathbf{X} = \mathbf{x}$ nebo $\mathbf{Y} = \mathbf{y}$.
- **Věta. 1.** Je-li $f_1(\mathbf{x}|\theta)$ sdružená hustota $\mathbf{X}$ a $f_2(t|\theta)$ je hustota $T(\mathbf{X})$, potom $T(\mathbf{X})$ je postačující statistika pro θ, pokud pro každé $\mathbf{x}$ je podíl
$$\frac{f_1(\mathbf{x}|\theta)}{f_2(t|\theta)}$$
konstanta nezávislá na θ.

Příklad 3.1. Postačující statistika binomického rozdělení Nechť X_i, $i = 1, 2, \ldots, N$ jsou iid Bernoulliho pokusy a $X = \sum_{i=1}^N X_i$. Potom $X \sim Bin(N, p)$. Ukažte, že $T(\mathbf{X}) = \sum_{i=1}^N X_i$ je postačující statistika pro p.

Poznámka: Zadáním úkolu je projít si odvození (případně si odvození samostatně provést a následně jej s uvedeným textem zkontrolovat). Řešení příkladu do R-Skriptu uveďte proto pouze v následující formě:

```
#============================================
#            P R I K L A D      3.1
#============================================

# Nastudovano.
```

Příklad 3.2. Postačující statistika normálního rozdělení Nechť $X_i \sim N(\mu, \sigma^2)$, kde $i = 1, 2, \ldots, N$ jsou iid proměnné a σ^2 poznáme. Ukažte, že $T(\mathbf{X}) = \sum_{i=1}^N X_i/N = \bar{X}$ je postačující statistika pro μ.

Poznámka: Zadáním úkolu je projít si odvození (případně si odvození samostatně provést a následně jej s uvedeným textem zkontrolovat). Řešení příkladu do R-Skriptu uveďte proto pouze v následující formě:

```
#============================================
#            P R I K L A D      3.2
#============================================

# Nastudovano.
```

Monotónní poměr věrohodnosti

Chceme otestovat nějakou $H_0 \rightarrow$ hledáme vhodný test. Při testování H_0 mohou nastat čtyři situace:

1. H_0 zamítáme, když H_0 neplatí. (Správně)
2. H_0 zamítáme, když H_0 platí. (Chyba I. druhu; CHPD)
3. H_0 nezamítáme, když H_0 neplatí. (Chyba II. druhu; CHDD)
4. H_0 nezamítáme, když H_0 platí. (Správně)

Statistický test volíme vždy tak, abychom obě chyby minimalizovali. Nicméně jsme limitováni vzájemným propojením obou chyb:

- S klesající $\Pr(CHPD)$ roste $\Pr(CHDD)$.
- S klesající $\Pr(CHDD)$ roste $\Pr(CHPD)$.

Pravděpodobnost chyby I. druhu umíme shora omezit pomocí hladiny významnosti α. Volbu vhodného testu potom provádíme pomocí Neyman-Pearsonova přístupu:

- Hladinu významnosti α předem volíme, a následně hledáme takový test, který má při dané alternativě největší sílu β^* (tj. má nejmenší $\Pr(CHDD)$).

Mějme tedy test $\mathcal{T}$, hladinu významnosti α, $\alpha \geq \Pr(CHPD)$ a sílu $\beta^*(\theta, \mathcal{T})$ stanovenou za platnosti alternativní hypotézy. Bohužel obecně platí, že pokud θ_1 a θ_2 jsou dvě různé hodnoty $\theta \in \Theta_1$, potom test, pro který je síla $\beta^*(\theta_1, \mathcal{T})$ maximalizovaná, může být jiný než test, pro který je maximalizovaná $\beta^*(\theta_2, \mathcal{T})$. Obecně neexistuje jeden test, který by maximalozoval $\beta^*(\theta, \mathcal{T})$ pro všechna $\theta \in \Theta_1$. Pro některé typy hypotéz ale takový test existuje a nazývá se *rovnoměrně nejsilnější test* (značíme jej $\mathcal{T}_{UMP}$; (*uniformly most powerful test*)).

Def.3. (Rovnoměrně nejsilnější test) Testujeme $H_0 : \theta \in \Theta_0$ oproti $H_1 : \theta \in \Theta_1$ testem $\mathcal{T}_{UMP}$ na hladině významnosti α. Nechť existuje taková pozorovaná hladina významnosti $\alpha(\mathcal{T}_{UMP})$, pro kterou je $\alpha(\mathcal{T}_{UMP}) \leq \alpha$. Test $\mathcal{T}_{UMP}$ se nazývá rovnoměrně nejsilnější test, pokud

$$\beta^*(\theta|\mathcal{T}) \leq \beta^*(\theta|\mathcal{T}_{UMP}), \forall \theta \in \Theta_1.$$

Def.4. (Monotónní poměr věrohodnosti) Testujeme $H_0 : \theta \in \Theta_0$ oproti $H_1 : \theta \in \Theta_1$ testem $\mathcal{T}$ na hladině významnosti α. Nechť pro $\theta_1 \in \Theta$ a $\theta_2 \in \Theta$ platí $\theta_1 < \theta_2$. Pokud je poměr věrohodností $L(\theta_2|\mathbf{x})/L(\theta_1|\mathbf{x})$ monotónní (rostoucí / klesající) funkcí $\mathcal{T}(\mathbf{X})$, potom $\mathcal{T} = \mathcal{T}_{UMP.}$

Pozn: Monotónní poměr věrohodnosti existuje pro jednostranné alternativní hypotézy (ne však vždy). Pokud pro vybranou alternativní hypotézu monotónní poměr věrohodnosti existuje, potom pro tuto hypotézu existuje také UMP test. Pro oboustranné hypotézy UMP testy neexistují.

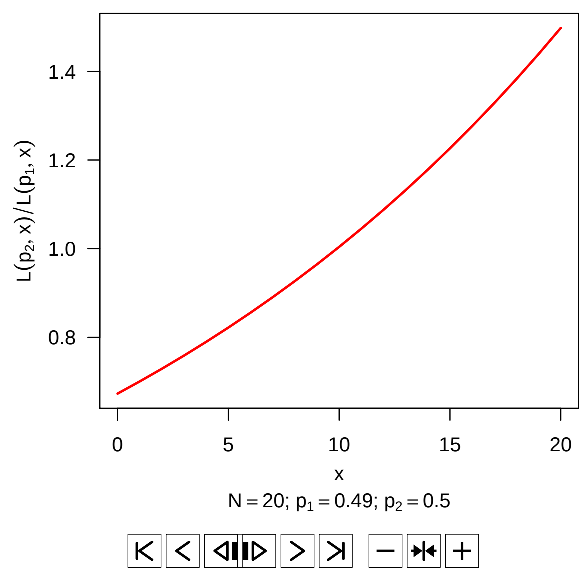
Příklad 3.3. Monotónní poměr věrohodnosti pro parametr p binomického rozdělení
Nechť $X \sim Bin(N, p)$, kde $\theta = p$. Funkce věrohodnosti je definovaná jako

$$L(\theta|x) = p^x(1-p)^{N-x},$$

kde $x = \sum_{i=1}^{N} x_i$.

Ukažte, že pro nějaké dvě hodnoty p_1, p_2, kde $0 < p_1 < p_2 < 1$ je poměr věrohodnosti $\frac{L(p_2|\mathbf{x})}{L(p_1|\mathbf{x})}$ monotónní (rostoucí) funkcí statistiky $T(\mathbf{X}) = X = \sum_{i=1}^{N} X_i$. Zvolte $N = 20$, $p_1 = 0.4$, $p_2 = 0.5$.

Vytvořte animaci zobrazující, že poměr věrohodnosti $\frac{L(p_2|\mathbf{x})}{L(p_1|\mathbf{x})}$ je monotónní funkcí statistiky $T(\mathbf{X}) = X = \sum_{i=1}^{N} X_i$ pro pevně zvolené $N = 20$, $p_2 = 0.5$, $p_1 \in \{0.49, 0.47, \ldots, 0.03, 0.01\}$.



Obrázek 1: Monotónní poměr věrohodnosti pro parametr p binomického rozdělení

Příklad 3.4. Rovnoměrně nejsilnější test pro parametr p binomické rozdělení
Na základě závěru příkladu 3 určete, zda existuje rovnoměrně nejsilnější test pro otestování

1. $H_{02} : p \leq p_0$ oproti $H_{12} : p > p_0$,
2. $H_{03} : p \geq p_0$ oproti $H_{13} : p < p_0$,
3. $H_{01} : p = p_0$ oproti $H_{11} : p \neq p_0$.

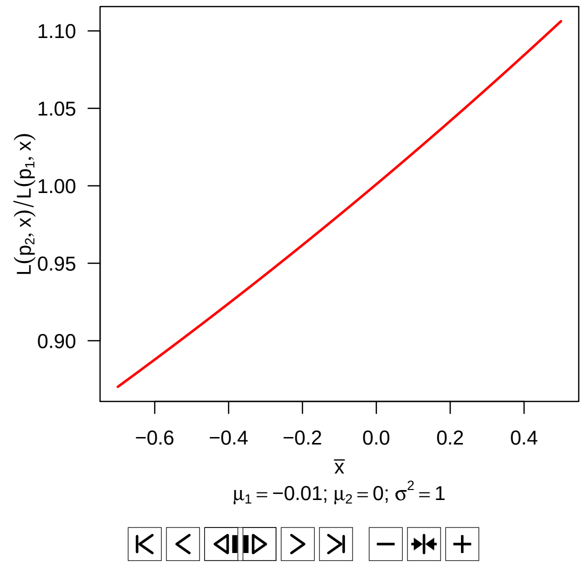
Pokud ano, uveďte příklad testovací statistiky rovnoměrně nejsilnějšího testu. Zdůvodněte, proč je právě tato statistika testovací statistikou rovnoměrně nejsilnějšího testu.

Příklad 3.5. Monotónní poměr věrohodnosti pro parametr μ normálního rozdělení
Nechť $X \sim N(\mu, \sigma^2)$, $\sigma^2 > 0$ je **známý rozptyl** a $\boldsymbol{\theta} = (\mu, \sigma^2)^T$. Funkce věrohodnosti je definovaná jako

$$L(\theta|x) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2}.$$

Ukažte, že pro nějaké dvě hodnoty μ_1, μ_2, kde $\mu_1 < \mu_2$ je poměr věrohodnosti $\frac{L((\mu_2,\sigma^2)^T|\mathbf{x})}{L((\mu_1,\sigma^2)^T|\mathbf{x})}$ monotónní funkcí statistiky $T(\mathbf{X}) = \bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$. Zvolte $\mu_1 = -0.01$, $\mu_2 = 0$, $\sigma^2 = 1$, $n = 20$.

Vytvořte animaci zobrazující, že poměr věrohodnosti $\frac{L((\mu_2,\sigma^2)^T|\mathbf{x})}{L((\mu_1,\sigma^2)^T|\mathbf{x})}$ je monotónní (rostoucí) funkcí statistiky $T(\mathbf{X}) = \bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ pro pevně zvolené $\mu_2 = 0$, $\sigma^2 = 1$, a $\mu_1 =\in \{-0.01, -0.03, \dots, -0.49\}$.



Obrázek 2: Monotónní poměr věrohodnosti pro parametr μ normálního rozdělení při známém rozptylu σ^2

Příklad 3.6. Rovnoměrně nejsilnější test pro parametr μ normálního rozdělení
Na základě závěru příkladu 3.5 určete, zda existuje rovnoměrně nejsilnější test pro otestování

1. $H_{02}: \mu \leq \mu_0$ oproti $H_{12}: \mu > \mu_0$,
2. $H_{03}: \mu \geq \mu_0$ oproti $H_{13}: \mu < \mu_0$,
3. $H_{01}: \mu = \mu_0$ oproti $H_{11}: \mu \neq \mu_0$.

za předpokladu, že $X \sim N(\mu, \sigma^2)$, kde σ^2 známe. Pokud ano, uveďte příklad testovací statistiky rovnoměrně nejsilnějšího testu. Zdůvodněte, proč je právě tato statistika testovací statistikou rovnoměrně nejsilnějšího testu.