MATEMATICKO-FYZIKÁLNÍ FAKULTA
Univerzita Karlova

Martin Farda

Waldův sekvenční test

3/2021

Předpoklady

- $X_1, \ldots, X_n$ iid náhodné veličiny.
- hustota $f(x, \theta)$ vzhledem k nějaké σ-konečné míře μ.
- $\theta \in (\underline{\theta}, \overline{\theta}), \quad \underline{\theta} < \overline{\theta}$.
- Chceme provést test hypotézy $H_0 : \theta = \theta_0$ proti alternativě $H_1 : \theta = \theta_1$, kde $\theta_0 \neq \theta_1 \in (\underline{\theta}, \overline{\theta})$.

Značení

- Označme α, β pravděpodobnosti chyb prvního a druhého druhu.
- Definujme $Z_i = Z_i(\theta_0, \theta_1) = \log \frac{f(X_i,\theta_1)}{f(X_i,\theta_0)}$, jsou-li hustoty nenulové, $Z_i = 1$, jsou-li obě nulové, $Z_i = -\infty$, je-li nulová pouze hustota v čitateli a $Z_i = \infty$, je-li nulová pouze hustota ve jmenovateli.
- $Q_n = Q_n(\theta_0, \theta_1) = \sum_{i=1}^{n} Z_i$
- $a = \log A, \quad b = \log B.$
- Operační charakteristiku testu $S(b, a)$ označme $L_S(\theta)$. Operační charakteristika je pravděpodobnost přijetí hypotézy H_0 testem $S(b, a)$ při skutečné hodnotě parametru θ.

Waldův test

- Srovnáváme $\prod_{i=1}^{n} \frac{f(X_i,\theta_1)}{f(X_i,\theta_0)}$ s A, B.
- $Q_n(\theta_0, \theta_1) \leq b$, pak přijmeme H_0.
- $Q_n(\theta_0, \theta_1) \geq a$, pak přijmeme H_1.
- V případě přijetí H_0, nebo H_1 zastavíme náhodný výběr. Jinak pokračujeme v náhodném výběru.
- $b < a$ třeba volit tak, aby odpovídali pravděpodobnosti chyb.

Waldův test

- Test daný tímto rozhodovacím pravidlem nazveme Waldův sekvenční test (nebo sekvenční test poměru pravděpodobností). Označme jej $S = S(b, a)$.
- Nerovnosti určující rozhodovací pravidlo nazveme kritické nerovnosti.
- Rozsah náhodného výběru je náhodná veličina $N = min\{n; Q_n(\theta_0, \theta_1) \notin (b, a)\}$.

Aproximace b,a

- Spočíst za daných α, β hodnoty a, b je obtížné. Používá se jejich aproximace.

Lemma

Pro Waldův sekvenční test $S(b, a)$ platí $a \leq \log \frac{1-\beta}{\alpha}$ a $b \geq \log \frac{\beta}{1-\alpha}$.

- Pro $\alpha^{'}, \beta^{'}$ chyby prvního a druhého druhu testu $S(\log \frac{\beta}{1-\alpha}, \log \frac{1-\beta}{\alpha})$ platí
  $\alpha^{'} < \frac{\alpha}{1-\beta}, \quad \beta^{'} < \frac{\beta}{1-\alpha}, \quad \alpha^{'} + \beta^{'} \leq \alpha + \beta$.
- b, a aproximujeme hodnotami $b^* = \log \frac{\beta}{1-\alpha}$ a $a^* = \log \frac{1-\beta}{\alpha}$.
- Wald ukázal, že použitím testu $S(\log \frac{\beta}{1-\alpha}, \log \frac{1-\beta}{\alpha})$ místo $S(b, a)$, s danými pravděpodobnostmi chyb 1. a 2. druhu se zvýší střední rozsah výběru.

Test skončí s pravděpodobností 1

Věta. Test skončí s pravděpodobností 1

Nechť $P(\frac{f(X_1,\theta_1)}{f(X_1,\theta_0)} > 1) > 0$ a $P(\frac{f(X_1,\theta_1)}{f(X_1,\theta_0)} < 1) > 0, \quad \theta \in (\underline{\theta}, \overline{\theta})$.
Nechť $-\infty < b < 0 < a < \infty$.
Pak test $S(b,a)$ skončí s pravděpodobností 1. Dále $E[N^k] < \infty \ \forall k > 0$. A $\exists t_0(\theta) > 0$ takové, že $E[exp(tN)] < \infty \quad \forall t \leq t_0(\theta)$

Věta

Nechť je splněn předpoklad předchozí věty. Nechť $E[exp(tZ_1)] < \infty \quad \forall t \quad \forall\theta \in (\underline{\theta}, \overline{\theta})$. Nechť $b < 0 < a$.
Pak $\exists h(\theta)$ reálná na $(\underline{\theta}, \overline{\theta})$, která je rovna nule pouze v bodě θ^* splňujícím $E_{\theta^*}[Z_1] = 0$ a splňující $E[exp(h(\theta)Q_N)] = 1 \quad \forall\theta \in (\underline{\theta}, \overline{\theta})$.

- $E[exp(h(\theta)Q_N)]$ můžeme rozepsat pomocí podmíněné střední hodnoty, kdy budeme podmiňovat tím, že je přijata hypotéza nebo alternativa.
- Aproximujeme podmíněné střední hodnoty pomocí hodnot a, b.

$$E[exp(h(\theta)Q_N)|H_0 \quad \text{přijata}] \tilde{=} exp(h(\theta)b)),$$
$$E[exp(h(\theta)Q_N)|H_1 \quad \text{přijata}] \tilde{=} exp(h(\theta)a))$$

Aproximace operační charakteristiky

- Získáváme aproximaci pro operační charakteristiku

$$\hat{L}_S(\theta) = \frac{1 - exp(h(\theta)a))}{exp(h(\theta)b) - exp(h(\theta)a)}, \quad E_\theta[Z_1] \neq 0,$$
$$\hat{L}_S(\theta^*) = \frac{a}{a - b}, \quad E_{\theta^*}[Z_1] = 0.$$

- $h(\theta)$ spočteme jako nenulové řešení rovnice $E[exp(h(\theta)Z_1)] = 1$.

Aproximace průměrného rozsahu výběru

- průměrný rozsah náhodného výběru při testu S označíme $EN = E_S(N) = E_S(N;\theta)$.
- Nechť platí předpoklad věty, že test skončí s pravděpodobností 1 a nechť $0 < E|Z_1| < \infty$.
- Pak $EN < \infty$ a $EN = \frac{EQ_N}{EZ_1}$.
- Rozepíšeme pomocí podmíněných středních hodnbot a aproximujeme podmíněné střední hodnoty pomocí hodnot a, b podobně jako při aproximaci operační charakteristiky.

$$E[Q_N|H_0 \quad \text{přijata}] \doteq b,$$
$$E[Q_N|H_1 \quad \text{přijata}] \doteq a.$$

Aproximace průměrného rozsahu výběru

- Dostáváme aproximaci pro průměrný rozsah náhodného výběru.

$$\hat{\hat{E}}N = \frac{\hat{L}_S(\theta)b + (1 - \hat{L}_S(\theta))a}{EZ_1}, \quad EZ_1 \neq 0,$$

$$\hat{\hat{E}}N = \frac{-ab}{E_{\theta^*}Z_1^2}, \quad E_{\theta^*}Z_1 = 0$$

Aproximace průměrného rozsahu výběru

- A speciálně pro θ_0, θ_1 dostáváme po aproximaci b, a následující:

$$\hat{\hat{E}}_{\theta_0} N = \frac{(1-\alpha)\log\frac{\beta}{1-\alpha} + \alpha\log\frac{1-\beta}{\alpha}}{E_{\theta_0} Z_1},$$

$$\hat{\hat{E}}_{\theta_1} N = \frac{\beta\log\frac{\beta}{1-\alpha} + (1-\beta)\log\frac{1-\beta}{\alpha}}{E_{\theta_1} Z_1}.$$

Příklad

- Nechť máme $X_1, X_2, \ldots$ posloupnost iid s $Alt(p), 0 \leq p \leq 1$.
- Chceme testovat $H_0 : p = p_0$ proti $H_1 : p = p_1, \quad p_0 < p_1$.
- Pro $0 < p_0 < p_1 < 1$ máme

  $$Q_n = \log\left(\left(\frac{p_1(1-p_0)}{p_0(1-p_1}\right)^{\sum_1^n X_i} \frac{1-p_1}{1-p_0}^n\right).$$
- Postup řešení na tabuli