

Statistika

Testování hypotéz – statistická indukce – Úvod do problému

Roman Biskup

(zapálený) statistik ve výslužbě, aktuálně analytik v praxi ;-)
roman.biskup(at)email.cz

21. února 2012



Obsah

Nulová a alternativní hypotéza

Chyby testování

Průkaznost hypotéz

Logické dělení testů



Myšlenka testování hypotéz

- ▶ = způsob ověřování domněnek o statistickém souboru:
 1. typu (zákona) rozdělení;
 2. parametrech rozdělení a charakteristik souboru: $\mu, \sigma^2, \sigma, \lambda, \tilde{x}, \dots$;
 3. shodě dvou a více rozdělení;
 4. shodě parametrů rozdělení a charakteristik dvou a více souboru.
- ▶ Myšlenka testování hypotéz
 - ▶ Dvojice vzájemně se vylučujících se hypotéz
 - H_0 Nulová hypotéza; hypotéza, která má být zamítnuta, matematické vyjádření hypotézy obsahuje „=“, „nic se nestalo“, zamítnutí má vážnější důsledky
 - H_A Alternativní hypotéza (H_1); hypotéza, která má být prokázána, matematické vyjádření hypotézy neobsahuje „=“, „něco se stalo“, zamítnutí nemá tak vážné důsledky.
 - ▶ za platnosti H_0 se pomocí matematické statistiky odvodí tzv. **testové kritérium** T (testová statistika) – to je (přesně nebo asymptoticky) náhodnou veličinou se známým rozdělením,
 - ▶ z možných hodnot, kterých může testové kritérium T nabývat, část svědčí pro nulovou hypotézu (**obor nezamítnutí** – K^c) a část pro její zamítnutí (**kritický obor** – K),
 - ▶ body oddělující obor nezamítnutí a kritický obor se nazývají **kritické hodnoty** a jsou jimi kvantily příslušných rozdělení.



Možné chyby testování

Výsledek testování	Skutečnost	
	H_0	H_A
H_0	$H_0 H_0$	$H_0 H_A$
H_A	$H_A H_0$	$H_A H_A$

Výsledek testování	Skutečnost	
	H_0	H_A
H_0	$T \in K^c H_0$	$T \in K^c H_A$
H_A	$T \in K H_0$	$T \in K H_A$

- ▶ správný výsledek ($H_0|H_0$, $H_A|H_A$)
- ▶ nesprávný výsledek
 - ▶ chyba prvního druhu ($H_A|H_0$)
 - ▶ nulová hypotéza byla zamítnutá ačkoliv byla správná (alternativní hypotéza neplatí)
 - ▶ chyba druhého druhu ($H_0|H_A$)
 - ▶ nulová hypotéza nebyla zamítnutá ačkoliv byla chybná (platí alternativní hypotéza)



Pravděpodobnosti chyb testování

Výsledek testování	Skutečnost	
	H_0	H_A
H_0	$P(H_0 H_0)$	$P(H_0 H_A)$
H_A	$P(H_A H_0)$	$P(H_A H_A)$

- ▶ pravděpodobnosti nesprávných výsledků
 - ▶ pravděpodobnost chyby prvního druhu – $P(H_A|H_0) = \alpha$ – tzv. **hladina významnosti**
 - ▶ pravděpodobnost chyby druhého druhu – $P(H_0|H_A) = \beta$
- ▶ správný výsledek
 - ▶ pravděpodobnost, že nenastane chyba prvního druhu – $P(H_0|H_0) = 1 - \alpha$ – tzv. **spolehlivost testu**
 - ▶ pravděpodobnost, že nenastane chyba druhého druhu – $P(H_A|H_A) = 1 - \beta$ – tzv. **síla testu**



Požadavky na testy

- ▶ Požadavky na testy:
 - ▶ maximalizovat spolehlivost a sílu testu, tj.
 - ▶ minimalizovat chybu prvního a druhého druhu.
- ▶ Pod kontrolou (prostřednictvím hladiny významnosti) držíme chybu prvního druhu.



Průkaznost hypotéz

Metafora o hypotézách před soudem

- ▶ Nulová hypotéza před soudem (alternativní samozřejmě také)
 - H_0 nulová hypotéza jde za mříže (alternativní je nevinná),
 - H_A nulová hypotéza je nevinná (alternativní je odsouzena).
- ▶ Práce obhájce nulové hypotézy
 - ▶ shromažďuje důkazy o nevině nulové hypotézy a
 - ▶ důkazy o vině alternativní hypotézy.
- ▶ Práce žalobce nulové hypotézy
 - ▶ shromažďuje důkazy o vině nulové hypotézy a
 - ▶ důkazy o nevině alternativní hypotézy.
- ▶ Práce soudce
 - ▶ dá pouze na obhájce: odsoudit alternativní hypotézu, pokud tak neučiní, nemůže prokázat vinu nulové
 - ▶ dá pouze na žalobce: odsoudit nulovou hypotézu, pokud tak neučiní, nemůže prokázat vinu alternativní
- ▶ Testování hypotéz (testové kritérium)= Soudce, jenž se rozhoduje pouze na základě práce obhájce (Vychází z nevinu nulové hypotézy).



Průkaznost hypotéz

Reálný dopad

- ▶ $T \in K^c$
 - ▶ Nulová hypotéza se **nezamítá** ... – nulovou hypotézu nelze prokázat (potvrdit, přijmou, ...), to znamená, že s $(1 - \alpha) \cdot 100\%$ může, ale nemusí platit.
 - ▶ **Nepodařilo** se prokázat alternativní hypotézu ...
 - ▶ Vše s předem zvolenou chybou α .
 - ▶ Říká se: Nepodařilo se prokázat alternativní hypotézu na hladině významnosti α , nebo
 - ▶ nelze zamítnout nulovou hypotézu na hladině významnosti α .
- ▶ $T \in K$
 - ▶ Nulová hypotéza se **zamítá**, to znamená, že s $(1 - \alpha) \cdot 100\%$ neplatí.
 - ▶ Podařilo se prokázat alternativní hypotézu ...
 - ▶ Vše s předem zvolenou chybou α .
 - ▶ Říká se: Alternativní hypotéza byla prokázána na hladině významnosti α , nebo
 - ▶ nulovou hypotézu byla zamítnuta na hladině významnosti α .



Logické dělení testů

- ▶ Dle objektu testování:
 - ▶ parametrické testy,
 - ▶ neparametrické testy.
- ▶ Dle počtu uvažovaných souborů (k):
 - ▶ jeden,
 - ▶ dva,
 - ▶ více než dva.
- ▶ Dle „vztahu“ mezi soubory (pro dva a více):
 - ▶ **nezávislé soubory** – statistické jednotky napříč výběrovými soubory nemají vzájemnou souvislost (provázanost)
 - ▶ **závislé soubory** – statistické jednotky napříč výběrovými soubory mají vzájemnou souvislost (provázanost)
 - opakované pozorování, společný vlastník, stejné podmínky měření;
 - pozorování napříč skupinami tvoří k -tice (pro dva výběry páry),
- ▶ Dle vyrovnanosti počtu pozorování ve souborech (pro dva a více):
 - ▶ **vyvážený pokusný plán** – $n_1 = n_2 = \dots = n_k$,
 - ▶ **nevyvážený pokusný plán**



Výběrové statistiky (kvantitativní znak) – značení I

► Jeden soubor:

- Výběr ze základního souboru (populace): $X_1, X_2, \dots,$
- realizace náhodného výběru z homogenní populace:

$x_1, x_2, \dots, x_n \dots \dots \dots n, \bar{x}, s^2$

► Dva soubory:

- Výběr ze dvou základních souborů (populací), které jsou stejně rozdělené $X_{11}, X_{12}, \dots$ a $X_{21}, X_{22}, \dots$ a liší se v jednom (alternativním) faktoru
- realizace dvou náhodných výběrů z homogenní populace:

$x_{11}, x_{12}, \dots, x_{1n_1} \dots \dots \dots n_1, \bar{x}_1, s_1^2$
 $x_{21}, x_{22}, \dots, x_{2n_2} \dots \dots \dots n_2, \bar{x}_2, s_2^2$

► Výběr z více než dvou základních souborů (populací):

- Více než dva základní soubory (populace), které jsou stejně rozdělené $Y_{11}, Y_{12}, \dots$ a $Y_{21}, Y_{22}, \dots, \dots Y_{k1}, Y_{k2}, \dots$ a liší se v jednom (kategorionálním) faktoru
- realizace k náhodných výběrů z homogenní populace:



Výběrové statistiky (kvantitativní znak) – značení II

i	n_i	výběr	$\bar{y}_i$	s_i^2
1	n_1	$y_{11}, y_{12}, \dots, y_{1n_1}$	$\bar{y}_1$	s_1^2
2	n_2	$y_{21}, y_{22}, \dots, y_{2n_2}$	$\bar{y}_2$	s_2^2
$\vdots$				
k	n_k	$y_{k1}, y_{k2}, \dots, y_{kn_k}$	$\bar{y}_k$	s_k^2
$\sum$	n	–	$\bar{y}_\bullet$	s_c^2

$$\blacktriangleright \bar{y}_\bullet = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij}, \quad s_c^2 = \frac{1}{n-1} \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_\bullet)^2$$

$$\blacktriangleright \bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}, \quad s_i^2 = \frac{1}{n_i-1} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2, \text{ pro } i = 1, \dots, k$$

$$\blacktriangleright \bar{y}_\bullet = \frac{1}{n} \sum_{i=1}^k \bar{y}_i n_i \quad s_c^2 \neq \frac{1}{n-k} \sum_{i=1}^k (n_i-1) s_i^2 = s_r^2 ?$$

