Klasická úloha testování hypotéz

V ní je postavena preferovaná hypotéza proti alternativě. Konstrukce testu (návrh, odvození, …) je dána tím, že výstupem musí být buď výrok o potvrzení hypotézy nebo její vyloučení ve prospěch alternativy. K dispozici jsou „vstupní“ data tvořící náhodný výběr. Hodnocením kvality testu jsou „dosažená“ pravděpodobnost chyby prvního druhu (zamítnutí hypotéza, pokud platí) a pravděpodobnost chyby druhého druhu (přijetí hypotézy, pokud neplatí).
Základem užití takového (klasického) testování je:

1. Hypotéza a alternativa se vzájemně vylučují (disjunkce).
2. Úloha je dichotomií, tj. test musí rozhodnout. Test končí buď přijetím hypotézy nebo (vylučovací = XOR) jejím zamítnutím. Nic jiného není možné.
3. Návrh klasického testu je dán
   a. ZADÁNÍM MEZNÍ PRAVDĚPODOBNOSTI CHYBY PRVNÍHO DRUHU.
   b. Hledáním „testovacího postupu“ který za garance (její horní meze) pravděpodobnosti chyby prvního druhu (v nějakém smyslu) minimalizuje pravděpodobnost chyby druhého druhu.

Většina obtíží konstrukce testů za klasických podmínek je v tom, že od testu požadujeme aby vždycky rozhodl (ve výše uvedeném smyslu volby mezi hypotézou a alternativou). Pokud připustíme, že ne nad každými daty test musí rozhodnout, dostáváme se k

úloze testování s neurčitostí testového výstupu.

Zde opět (v základním tvaru) pracujeme s preferovanou hypotézou a proti ní postavenou alternativou. Opět se jedná o dichotomii MOŽNÁ je jen hypotéza nebo alternativa. Hypotéza a alternativa se opět vzájemně vylučují. Ale výstupem testu může být jen jediný výrok z následující trojice:

- Přijetí hypotézy.
- Zamítnutí hypotézy ve prospěch alternativy.
- NEROZHODNUTÍ. Tj. výstupem testu je konstatování, že na základě disponibilních dat a při použití daného testového postupu nelze rozhodnout.

Pokud přijmeme možnost výroku o NEROZHODNUTÍ máme k dispozici elegantní (a hlavně jednoduché) odhady mezí pro testové kritérium ve tvaru „poměru věrohodností“:
Jsou zadány pravděpodobnosti chyb prvního a druhého druhu:

$$P(\text{zamítnutí hypotézy} \mid \text{hypotáza platí}) = \alpha$$
$$P(\text{přijetí hypotézy} \mid \text{hypotáza neplatí}) = \beta$$ [1]

Hledáme dvě čísla $A\ a\ B$ k nim máme testové kritérium ve tvaru:

- $\prod_{i=1}^{n} \frac{f_A(x_i)}{f_H(x_i)} \leq B$ pak přijmeme hypotézu H,
- $\prod_{i=1}^{n} \frac{f_A(x_i)}{f_H(x_i)} \geq A$ pak zamítneme hypotézu H ve prospěch alternativy A a
- $B < \prod_{i=1}^{n} \frac{f_A(x_i)}{f_H(x_i)} < A$ zde nerozhodneme. [2]

Výhodou je, že u takto formulované úlohy máme k dispozici efektivní odhady čísel $A\ a\ B$: $B^* = \frac{\beta}{1-\alpha}$ a $A^* = \frac{1-\beta}{\alpha}$. U takových testů je základním problémem dokázat že se odhady $A^*\ a\ B^*$ čísel $A\ a\ B$ „příliš neliší“ od ideálních $A\ a\ B$. Vyjádření a důkaz lze nalézt např. v C. Radhakrishna Rao: Lineární metody

---

[1] Pozor na častý rozdíl značení pravděpodobnosti chyby druhého druhu u klasických testů, zde β tam $1-\beta$, v některých textech.

[2] U klasické (a základní) úlohy testování hypotéz hledáme JEDINÉ číslo C (pro uvedené vlastnosti pravděpodobností chyb prvního a druhého druhu), pokud $\prod_{i=1}^{n} \frac{f_A(x_i)}{f_H(x_i)} \geq C$ zamítáme hypotézu H ve prospěch alternativy A a pokud je $\prod_{i=1}^{n} \frac{f_A(x_i)}{f_H(x_i)} < C$ nezamítáme hypotézu H.

statistické indukce a jejich aplikace, ACADEMIA, Praha 1978, str. 519-520[3]. Dalším problémem u takových „testování s neurčitostí výstupu" je konstruovat testy tak, aby „nepříliš často" vedly na NEROZHODNUTÍ. Obecnější úlohou je nalezení jiných tvarů testových kritérií než odvozených z „poměru pravděpodobností" $\prod_{i=1}^{n}\frac{f_A(x_i)}{f_H(x_i)}$ a k nim jejich rozhodovacích mezí (s nějakými předem danými vlastnostmi).
Testy s neurčitostí se často používají u tzv. „přejímek měřením" nebo „přejímek srovnáním" [4]. V dost případech lze situaci „nerozhodnutí" interpretovat jako nedostatek počtu měření nebo pozorování = pro rozhodnutí je třeba „vstupní náhodný výběr" doplnit. To vede na

sekvenční testy (Waldovy)

V tomto přístupu opětovně hledáme dvě čísla $A\ a\ B$ k nim máme testové kritérium ve tvaru:

- $\prod_{i=1}^{n}\frac{f_A(x_i)}{f_H(x_i)} \le B$ pak přijmeme hypotézu H,
- $\prod_{i=1}^{n}\frac{f_A(x_i)}{f_H(x_i)} \ge A$ pak zamítneme hypotézu H ve prospěch alternativy A a
- $B < \prod_{i=1}^{n}\frac{f_A(x_i)}{f_H(x_i)} < A$ zde test indikuje potřebu **doplnit měření nebo pozorování**.

Výhodou je, že u i takto formulované úlohy máme k dispozici (stejné) efektivní odhady čísel $A\ a\ B$: $B^* = \frac{\beta}{1-\alpha}$ a $A^* = \frac{1-\beta}{\alpha}$. U takových testů je opětovně základním problémem dokázat že se odhady $A^*\ a\ B^*$ čísel $A\ a\ B$ „příliš neliší" od ideálních $A\ a\ B$. U těchto testů dále potřebujme „zajistit" aby testování neprobíhalo trvalou indikací potřeby doplnit další a další měření nebo pozorování. Tj. aby takové testování (v nějakém smyslu) končilo rozhodnutím. K tomu je důležité číslo

$$N = min\left\{n; \prod_{i=1}^{n}\frac{f_A(x_i)}{f_H(x_i)} \notin (B, A)\right\}, \text{ reálně } N^* = min\left\{n; \prod_{i=1}^{n}\frac{f_A(x_i)}{f_H(x_i)} \notin (B^*, A^*)\right\}$$

tj. nejmenší počet kroků za který testování skončí. Je zřejmé, že takový počet kroků je náhodnou veličinou. Pokud je podmínka konce testu vyjádřena $E(N) < +\infty$ a jako testové kritérium volíme výše uvedený poměr „pravděpodobností" mluvíme o sekvenčních Waldových testech hypotéz. Příslušné (základní) důkazy lze nalézt (stejně jako v předchozím typu úlohy) např. v C. Radhakrishna Rao: Lineární metody statistické indukce a jejich aplikace, ACADEMIA, Praha 1978, str. 519-523 (jedná se o stejnou konstrukci jako u předchozí modifikace úlohy doplněné o důkaz splnění požadavku $E(N) < +\infty$).

V teorii sekvenčních testů jsou rozebírány následující problémy:

- Srovnání sekvenčních testů s klasickými.
- Jinak vyjádřené podmínky skončení testu.
- Odhady počtů potřebných měření nebo pozorování (velikosti vstupního náhodného výběru) pro data (a dvojici hypotéza – alternativa) konkrétních vlastností.
- Jiné tvary testového kritéria než zde uvedeného (tj. $\prod_{i=1}^{n}\frac{f_A(x_i)}{f_H(x_i)}$ nebo z něj odvozeného) = ne-waldovské testy.

[3] Zde a i jinde jsou tvrzení dokazována pro sekvenční testy, viz dále.
[4] Jedná se o postupy kdy se z dodávky „vybere" náhodně několik kusů, ty se zkontrolují a na základě výsledku jejich kontroly se usoudí o kvalitě celé dodávky. Používají se hlavně tam, kde se kontrolou kontrolované předměty stanou dále nepoužitelné (např. v lékařských dodávkách se při kontrole naruší sterilita nebo ve vojenských se kontrolované náboje zničí výstřelem v „cejchované – měřicí" hlavni = destruktivní kontrolní postupy) nebo je vlastní (plná) kontrola velmi drahá nebo … . Obvykle je první test testem s neurčitostí, pokud rozhodne přejímka končí, pokud nerozhodne následuje nějaký klasický nebo sekvenční test = více výběrové testy.