TESTOVÁNÍ HYPOTÉZ V MARKETINGU

Obecný postup testování

1. Formulace hypotézy,
2. Výpočet testového kritéria
3. Nalezení kritické hodnoty pro zvolenou hladinu významnosti (vymezení oboru přijetí a kritického oboru)
4. Rozhodnutí o přijetí nebo zamítnutí hypotézy

Rozdělení statistických hypotéz

Statistické hypotézy tvoří jen část vědeckých (nebo alespoň odborných) hypotéz. Týkají se náhodných veličin a rozdělujeme je do dvou velkých tříd na:

- Parametrické hypotézy: vztahují se k jednomu parametru nebo několika parametrům daného pravděpodobnostního rozdělení náhodné veličiny (neboli znaku populace).
- Neparametrické hypotézy: netýkají se parametrů rozdělení náhodné veličiny, nýbrž jiných statistických vlastností, např. tvaru pravděpodobnostního rozdělení této veličiny (může nás třeba zajímat, zda lze chování náhodné veličiny modelovat binomickým rozdělením, normálním rozdělením apod.).

Testované hypotézy

Zopakujme, že v každém statistickém testu vystupují proti sobě dvě hypotézy:

- testovaná hypotéza (testované tvrzení), kterou nazýváme *nulová hypotéza* a značíme H_0,
- a *alternativní hypotéza*, značena H_1.

H_1 je obvykle logickou negací hypotézy H_0.

Testování hypotézy

Při testování hypotézy máme k dispozici především výsledek náhodného výběru. Na základě výběrových dat máme nyní rozhodnout, zda testovanou hypotézu přijmout nebo zamítnout.

Další postup:

- Počítáme tzv. *testové kritérium T* jakožto funkci dat, která jsme získali náhodným výběrem, a dále vymezujeme na reálné ose podmnožinu zvanou *kritický obor*.
  - Padne-li hodnota *testového kritéria T* pro získaný vzorek dat do kritického oboru, zamítneme testovanou, tj. nulovou hypotézu.
  - Naopak, padne-li hodnota testového kritéria mimo kritický obor, testovanou hypotézu přijímáme.
- Kritický obor vymezuje obvykle reálné číslo zvané *kritická hodnota K* (kritický obor lze vymezit také kvantilem). Tato kritická hodnota se hledá buďto ve statistických tabulkách nebo se počítá s využitím vhodného softwaru (často např. Excelu).

Poznámka

- Přijetí statistické hypotézy není matematickým důkazem platnosti testovaného tvrzení!
- Testování hypotéz nemusí vždy vést ke správným rozhodnutím, což je přirozené, neboť jde o náhodný proces využívající omezené informace náhodného výběru.
- Nejistota závěru statistického testu souvisí mimo jiné s tzv. hladinou významnosti , kterou si statistik stanovuje při provádění statistického testu hypotézy.
- Zdůrazněme opět, že základem statistického testování je náhodnost výběru, která souvisí s tím, zda vybraná data byla vybrána nezávisle na sobě. Tuto skutečnost lze rovněž testovat,

Obecný postup testování

-

1. Formulace nulové hypotézy H_0 a alternativní hypotézy H_1,

2. Výpočet testového kritéria T,

3. Nalezení kritické hodnoty K pro zvolenou hladinu významnosti α (vymezení kritického oboru C),

4. Porovnání K a T, resp. rozhodnutí, zda $T \in C$, a dle toho přijetí nebo zamítnutí H_0.

Závěr testu

- Je-li $T \in C$, H_0 se zamítá. Je-li $T \notin C$, H_0 se přijímá.

Věrohodnost testu

- Protože rozhodnutí přijmout nebo zamítnout hypotézu závisí na omezené informaci obsažené ve vzorku dat, můžeme se při testování dopustit chyb dvojího druhu:
  - Zamítneme nulovou hypotézu, která ve skutečnosti platí. Tím se dopustíme *chyby prvního druhu*. Pravděpodobnost chyby prvního druhu se značí α a nazývá se *hladina významnosti*.
  - Přijmeme nulovou hypotézu, která ve skutečnosti neplatí. Dopustíme se tak *chyby druhého druhu*. Pravděpodobnost chyby druhého druhu se značí β. Pravděpodobnost 1 – β se nazývá síla testu. Je to pravděpodobnost, že test povede k oprávněnému zamítnutí testované nulové hypotézy.
- Hladina významnosti α se při testu volí obvykle 0,05, 0,01 nebo 0,1.

Alternativa – testování pomocí počítače

- Kromě hladiny významnosti se využívají při testování hypotéz také tzv. **p-hodnoty**. Často jsou součástí výstupů matematických počítačových programů.
- P-hodnota vyjadřuje pravděpodobnost obdržení nebo překročení spočteného testového kritéria.
- Pokud je p-hodnota menší než stanovená hladina významnosti, příp. rovna této hladině, nulová hypotéza se zamítá. V opačném případě se nulová hypotéza přijímá.
- $p \leq \alpha \ \rightarrow \ H_1$
- $p > \alpha \ \rightarrow \ H_0$

Základní statistické testy

(A) Jednovýběrový t-test.

(B) Dvouvýběrový t-test s rovností rozptylů.

(C) Dvouvýběrový t-test s nerovností rozptylů.

(D) Dvouvýběrový párový t-test.

(E) Dvouvýběrový F-test pro rovnost rozptylů.

(A) Test předpokladu o střední hodnotě základního souboru (Jednovýběrový t-test střední hodnoty)

- Nechť $X = (X_1,\dots, X_n)$ je náhodný výběr z rozdělení $N(\mu, \sigma^2)$, přitom σ^2 není známo.

Dále:

- $\bar{X}$ = výběrový průměr spočtený z dat $X_1, X_2, \dots, X_n$
- S = výběrová směrodatná odchylka spočtená z dat $X_1, X_2, \dots, X_n$
- μ_0 = předpoklad o μ zadaný statistikem
- n = rozsah výběru.

Postup testování

1: Testuje se
- nulová hypotéza H_0: $\mu = \mu_0$
- alternativní hypotéza H_1: $\mu \neq \mu_0$

2. Testové kritérium T má tvar

$$T = \frac{\bar{X} - \mu_0}{S} \cdot \sqrt{n},$$

3. Kritickou hodnotou K je zde kritická hodnota Studentova rozdělení s n-1 stupni volnosti na hladině α, která se značí $t_{n-1}(\alpha)$.

4. Je-li $|T| \geq t_{n-1}(\alpha)$ zamítá se H_0 a přijímá se H_1, jinak se přijímá H_0.

Kritická hodnota - poznámky

- Pro kritickou hodnotu $t_{n-1}(\alpha)$ platí $P(|T| \geq t_{n-1}(\alpha))$=α.
- Kritická hodnota – ve statistických tabulkách
- Jiná možnost – v **Excelu** pomocí funkce TINV(α ; *n*-1).

Příklad

Automat na plnění litrových lahví je podle výrobce seřízen tak, že střední hodnota objemu naplněných lahví je 1000 ml. Kontrola jakosti 25 naplněných lahví ukázala, že průměrný objem náplně byl 998 ml se směrodatnou odchylkou 5 ml. Je automat seřízený správně? Testujte na hladině významnosti 0,05, předpokládejte normální rozdělení základního souboru.

Řešení příkladu

1. Testuje se
   - nulová hypotéza H_0: $\mu = 1000$
   - alternativní hypotéza H_1: $\mu \neq 1000$
2. Testové kritérium:

$$|T| = \frac{|\bar{x} - \mu_0| . \sqrt{n}}{s} = \frac{|998 - 1000| . \sqrt{25}}{5} = 2$$

3. Kritická hodnota $t_{24}(\alpha = 0{,}05) = 2{,}06$

4. Závěr: Na hladině významnosti 0,05 nemůžeme zamítnout nulovou hypotézu a tedy automat je seřízený správně.

(B) Test významnosti rozdílu mezi dvěma populačními průměry (Dvouvýběrový t-test s rovností rozptylů)

- Uvažujme dva náhodné výběry *Y* a *Z* z normálního rozdělení.
- Předpokládejme, že rozptyly obou souborů jsou stejné; v praxi si to pro konkrétní hodnoty můžete ověřit *F*-testem.
- Výběrové charakteristiky proměnné *Y* jsou $\overline{x_1}$, s_1^2, n_1
- Výběrové charakteristiky proměnné *Z* jsou $\overline{x_2}$, s_2^2, n_2
- Rozsahy výběrů nemusí být stejné.

Testování - Dvouvýběrový t-test s rovností rozptylů

Postup testování:

1. Stanovení hypotézy: $H_0 : \mu_1 - \mu_2 = \Delta$, $H_1 : \mu_1 - \mu_2 \neq \Delta$.
2. Testové kritérium: $T = \dfrac{(\bar{x}_1 - \bar{x}_2 - \Delta)}{\sqrt{(n_1 - 1).s_1^2 + (n_2 - 1).s_2^2}} . \sqrt{\dfrac{n_1.n_2.(n_1 + n_2 - 2)}{n_1 + n_2}}$.
3. Pro $|T|$ je obor přijetí: $\langle 0 , t_{n_1+n_2-2}(\alpha) \rangle$, kritický obor: $(t_{n_1+n_2-2}(\alpha) , +\infty)$, kde $t_{n_1+n_2-2}(\alpha)$ je kritická hodnota
4. Závěr.

Test významnosti rozdílu mezi dvěma populačními průměry

- Speciální případ testu když $\Delta = 0$.

Příklad

- Ve sportovním areálu jsou dva okruhy Y a Z, které vypadají stejně dlouhé. Závodník běžel šestkrát okruhem Y a pětkrát okruhem Z, naměřené časy v sekundách jsou v tabulce. Zjistěte, jestli je na hladině významnosti 0,05 čas oběhu okruhu Y stejný jako čas oběhu okruhu Z.

Y	*Z*
61	52
54	57
55	49
60	49
54	51
58	

Řešení příkladu

1. Stanovení hypotézy: $H_0 : \mu_1 - \mu_2 = 0$, $H_1 : \mu_1 - \mu_2 \neq 0$.
2. Pro výpočet testového kritéria musíte nejdříve dopočítat výběrové průměry a rozptyly pro oba soubory hodnot. Potom má testové kritérium hodnotu:

$$|T| = \frac{|\bar{x}_1 - \bar{x}_2 - \Delta|}{\sqrt{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}} \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}} =$$

$$= \frac{|57 - 51{,}6 - 0|}{\sqrt{(6-1) \cdot 9{,}6 + (5-1) \cdot 10{,}8}} \cdot \sqrt{\frac{6 \cdot 5 \cdot (5 + 6 - 2)}{5 + 6}} = 2{,}801.$$

3. Kritickou hodnotu $t_9(0{,}05)$ vypočítáte pomocí funkce TINV. Potom je obor přijetí: $\langle 0\,,\, 2{,}262 \rangle$, kritický obor: $(2{,}262\,,\, +\infty)$.
4. Závěr: Testové kritérium spadá do kritického oboru, a tedy hypotézu o stejné délce běhu oběma okruhy zamítneme na hladině významnosti 0,05.

(C) Test významnosti rozdílu mezi dvěma populačními průměry (Dvouvýběrový t-test s nerovností rozptylů)

- Uvažujme dva náhodné výběry *Y* a *Z* z normálního rozdělení.
- Předpokládejme, že rozptyly obou souborů jsou stejné; v praxi si to pro konkrétní hodnoty můžete ověřit *F*-testem.
- Výběrové charakteristiky proměnné *Y* jsou $\overline{x_1}$, s_1^2, n_1
- Výběrové charakteristiky proměnné *Z* jsou $\overline{x_2}$, s_2^2, n_2
- Rozsahy výběrů nemusí být stejné.

Testování – Dvouvýběrový t-test s nerovností rozptylů

1. Stanovení hypotézy: $H_0 : \mu_1 - \mu_2 = \Delta$, $H_1 : \mu_1 - \mu_2 \neq \Delta$
2. Testové kritérium: $T = \dfrac{(\bar{x}_1 - \bar{x}_2 - \Delta)}{\sqrt{\dfrac{s_1^2}{n_1 - 1} + \dfrac{s_2^2}{n_2 - 1}}}$.
3. Pro $|T|$ je přibližná kritická hodnota: $T_{krit.} = \dfrac{\dfrac{s_1^2}{n_1 - 1} t_{n_1 - 1}(\alpha) + \dfrac{s_2^2}{n_2 - 1} t_{n_2 - 1}(\alpha)}{\dfrac{s_1^2}{n_1 - 1} + \dfrac{s_2^2}{n_2 - 1}}$, obor přijetí: $\langle 0 , T_{krit.} \rangle$, kritický obor: $(T_{krit.} , + \infty)$.
4. Závěr.

Test významnosti rozdílu mezi dvěma populačními průměry

- Speciální případ testu když $\Delta = 0$.

Postup v EXCELu

- V Excelu pro dvouvýběrový *t*-test s nerovností rozptylů můžete využít analytický nástroj Dvouvýběrový t-test s nerovností rozptylů a funkci TTEST, když indikátor typ nastavíte na 3.
- Tento nástroj však vypočítá testové kritérium poněkud jinak, a to jako veličinu s *t*-rozdělením

$$T = \frac{(\bar{x}_1 - \bar{x}_2 - \Delta)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

- a s počtem stupňů volnosti

$$df = Round\left(\frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}} \right).$$

Poznámka

- Kritická hodnota, jejíž vzorec jsme vám ukázali, i výpočet pomocí programu Excel je jen aproximací. Tento test je přibližný a ztrácí hodně informací z naměřených hodnot.

(D) Dvouvýběrový párový t-test

- Vhodný test když výběry nejsou nezávislé.
- Uvažujme náhodný výběr z dvourozměrného rozdělení. Naměřené hodnoty jsou: $(y_1, z_1), (y_2, z_2), \dots, (y_n, z_n)$, kde μ_1 je střední hodnota veličiny *Y* a μ_2 je střední hodnota veličiny *Z* .
- Označme $x_1 = y_1 - z_1, x_2 = y_2 - z_2, \dots, x_n = y_n - z_n$. Potom $x_1, x_2, \dots, , x_n$ jsou prvky normální veličiny *X* s průměrem $\bar{x}$ a rozptylem s^2 .
- Párový test zkoumá, jestli je rozdíl středných
- hodnot veličin *Y* a *Z* roven hodnotě Δ.

Postup testování

1. Stanovení hypotézy: $H_0 : \mu_1 - \mu_2 = \Delta$, $H_1 : \mu_1 - \mu_2 \neq \Delta$.
2. Testové kritérium: $T = \dfrac{(\bar{x} - \Delta)\sqrt{n}}{s}$.
3. Pro $|T|$ je obor přijetí: $\langle 0 , t_{n-1}(\alpha) \rangle$, kritický obor: $(t_{n-1}(\alpha) , +\infty)$, kde t_{n-1} je kritická hodnota Studentova rozdělení.
4. Závěr.

Příklad

- V následující tabulce jsou data *Y* a *Z*, která určují přesnost zásahu šipkou 6 hráčů pravou a levou rukou. Zjistěte, jestli na hladině významnosti 0,05 je přesnost zásahu oběma rukama stejná, tedy že rozdíl v přesnosti je 0.

Y	*Z*
2,8	2,5
2,6	2,1
3,2	3
1,9	2,1
2,2	2,4
2,6	2,4

Řešení

1. Stanovení hypotézy: $H_0 : \mu_1 - \mu_2 = 0$, $H_1 : \mu_1 - \mu_2 \neq 0$.
2. Abyste mohli vypočítat testové kritérium, potřebujete znát hodnoty veličiny X.
   Z nich pak vypočítáte $\overline{x} = 0{,}1333$ a $s = 0{,}2805$. Potom testové kritérium:

   $$|T| = \frac{|\overline{x} - \Delta| . \sqrt{n}}{s} = \frac{|0{,}1333 - 0| . \sqrt{6}}{0{,}2805} = 1{,}164 .$$
3. Obor přijetí: $\langle 0 \,;\, 2{,}571 \rangle$, kritický obor: $(2{,}571 \,;\, +\infty)$
4. Závěr: Testové kritérium spadá do oboru přijetí, a tedy nelze na základě získaných dat zamítnout hypotézu, že přesnost zásahu šipkou levou a pravou rukou je stejná.

(E) Dvouvýběrový F – test pro rozptyly

- Uvažujme dva náhodné výběry *Y* a *Z* z normálního rozdělení.
- Výběrové charakteristiky proměnné *Y* jsou $\overline{x_1}$, s_1^2, n_1
- Výběrové charakteristiky proměnné *Z* jsou $\overline{x_2}$, s_2^2, n_2
- Rozsahy výběrů nemusí být stejné
- Předpokládejme, že $s_1^2 \geq s_2^2$

Postup testování

1. Stanovení hypotézy: $H_0 : \sigma_1^2 = \sigma_2^2$, $H_1 : \sigma_1^2 \neq \sigma_2^2$
2. Testové kritérium: $T = \dfrac{s_1^2}{s_2^2}$ kde větší z rozptylů je v čitateli.
3. Kritický obor: $\left(F_{(n_1-1),\,(n_2-1)}(\alpha), +\infty\right)$, kde $F_{(n_1-1),\,(n_2-1)}(\alpha)$ je kritická hodnota F- rozdělení.
4. Závěr.

Příklad

- Ve sportovním areálu jsou dva okruhy Y a Z, které vypadají stejně dlouhé. Závodník běžel šestkrát okruhem Y a pětkrát okruhem Z, naměřené časy v sekundách jsou v tabulce. Zjistěte, jestli je na hladině významnosti 0,05 rozptyl času oběhu okruhu Y stejný jako rozptyl času oběhu okruhu Z.

Y	*Z*
61	52
54	57
55	49
60	49
54	51
58	

Řešení příkladu

1. $H_0 : \sigma_1^2 = \sigma_2^2$, $H_1 : \sigma_1^2 \neq \sigma_2^2$
2. Pomocí funkce VAR.VÝBĚR vypočítáte výběrový rozptyl z hodnot ve sloupcích *Y* a *X*. Pro výpočet testového kritéria uspořádáte rozptyly tak, aby vyšší rozptyl byl v čitateli: $T = \frac{s_2^2}{s_1^2} = \frac{10{,}8}{9{,}6} = 1{,}125$.
3. Pro nalezení kritického oboru musíte určit kritickou hodnotu *F*-rozdělení, a to například pomocí funkce FINV takto:
   FINV(0,05;4;5)=5,19.
   Pozor, počet stupňů volnosti musí odpovídat pořadí rozptylů v čitateli a pak ve jmenovateli. $C = (5{,}19\ , +\infty)$
4. Závěr: testové kritérium nespadá do kritického oboru, přijímáme nulovou hypotézu, rozptyly časů jsou pro oba okruhy stejné.

Děkuji za pozornost.