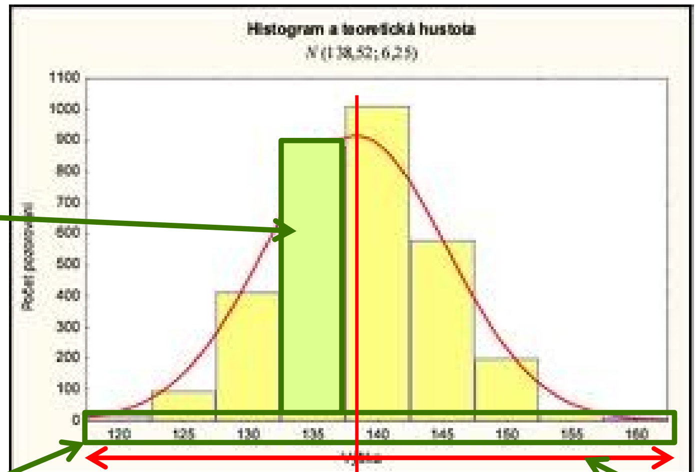
VI. ŠTATISTICKÉ TESTOVANIE

Opakovanie – normálne rozloženie



$\sim N(\mu,\sigma^2)$

stredná hodnota rozptyl

relatívna četnosť =

četnostná hustota

triediace intervaly

rozptyl

priemer=medián

interval $\mu \pm 3\sigma$ obsahuje 99,7 % populace

Normalita dát

Pretože má normalita rozdelenia dobré vlastnosti, vďaka ktorým môžeme o skúmanom súbore vyvodzovať závery, snažíme sa pri skúmaných javoch overiť, či nepochádzajú z tohto rozloženia. Pokiaľ by sme použili štatistický test predpokladajúci normalitu dát na dáta, ktoré majú iné rozloženie, došli by sme ku skresleným výsledkom. Normalitu môžeme testovať rôznym spôsobom.

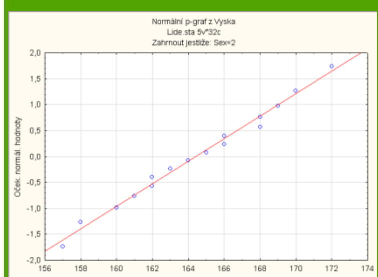
Shapiro-Wilksov test

Neparametrický test použiteľný aj pre malé rozsahy s dobrou silou testu. Pokiaľ máme teda malé N (N < 50), použijeme tento test. U väčších rozsahov si môžeme test zvoliť. Pri porovnaní testov normality vidíme, že pre veľké N vychádzajú veľmi podobne. V prípade malého rozsahu výberu (malého N), pri rôznych záveroch testov, sa budeme riadiť radšej Shapiro-Wilksovým testom.

Kolmogorov-Smirnovov test

Dokáže dobre nájsť odľahlé hodnoty, ale počíta skôr so symetriou dát než priamo s normalitou. Mal by byť počítaný len v prípade, že poznáme teoretickú smerodajnú odchýlku. V praxi túto hodnotu väčšinou nepoznáme, preto používame jeho modifikáciu – Lilieforsov test.

Normál – pravdepodobnostný graf (N-P plot)



Zobrazuje odchýlky hodnôt od pravdepodobnostnej priamky. Pomocou vizuálnej metódy je možné poznať, kedy dáta vyzerajú, že pochádzajú z normálneho rozloženia.

Test dobrej zhody

Test dáva dobré výsledky, ale je náročný na N, teda množstvo dát, aby bolo možné vytvoriť dostatočný počet tried hodnôt.

V prípade veľkého počtu pozorovaní N>30 a zamietnutí hypotézy o normalite dát (graf rozloženia sa veľmi nevychyľuje) nebude chybou analyzovať dáta ako normálne rozložené.

Parametrické vs. neparametrické testy

PARAMETRICKÉ TESTY

- Majú predpoklady o rozložení vstupujúcich dát (napr. normálne rozloženie)
- Pri rovnakom N a dodržaní predpokladov majú vyššiu silu testu než testy neparametrické
- Pokiaľ nie sú dodržané predpoklady parametrických testov, potom ich sila testu prudko klesá a výsledok testu môže byť chybný a nezmyselný

NEPARAMETRICKÉ TESTY

- Nemajú predpoklady o rozložení vstupujúcich dát, je ich teda možné použiť aj pri asymetrickom rozložení, odľahlých hodnotách, či nedetekovateľnom rozložení
- Znížená sila týchto testov je spôsobená redukciou informačnej

  hodnoty pôvodných dát, kedy neparametrické testy nevyužívajú pôvodné hodnoty, ale najčastejšie len ich poradie

One-sample vs. two sample testy

JEDNO-VÝBEROVÉ TESTY (one-sample)

- Porovnávajú jedu vzorku (one sample, jednovýberové testy) s referenčnou hodnotou (poprípade so štatistickým parametrom cieľovej populácie)
- V teste je teda porovnávané rozloženie hodnôt (vzorka) s jediným číslom (referenčná hodnota, hodnota cieľovej populácie)
- Otázka položená v teste môže byť vztiahnutá k priemeru, rozptylu, podielu hodnôt aj ďalším štatistickým parametrom popisujúcim vzorku

DVOJ-VÝBEROVÉ TESTY (two-sample)

- Porovnávajú navzájom dva vzorky (two sample, dvojvýberové testy)
- V teste sú porovnávané dve rozloženia hodnôt
- Otázka položená v teste môže byť opäť vztiahnutá k priemeru, rozptylu, podielu hodnôt aj ďalším štatistickým parametrom popisujúcim vzorku
- Okrem testov pre dve skupiny hodnôt existujú samozrejme

  i testy pre viac skupín dát

Nepárový vs. párový design

NEPÁROVÝ DESIGN

- Skupiny porovnávaných dát sú na sebe úplne nezávislé (tiež nezávislý, independent design), napr. Ľudia z rôznych zemí, nezávislé skupiny pacientov s odlišnou liečbou atď.
- Pri výpočte je nevyhnutné brať v úvahu charakteristiky oboch skupín dát

PÁROVÝ DESIGN

- Medzi objektami v porovnávaných skupinách existuje väzba, daná napr. človekom pred a po operácii, reakcie rovnakého kmeňa krýs atď.
- Väzba môže byť buď priamo daná alebo len predpokladaná (v tom prípade je nutné ju overiť)
- Test je v podstate uskutočňovaný na diferenciach skupín, nie na ich pôvodných dátach

Typ srovnání	Parametrický test	Neparametrický test
2 skupiny dat nepárově:	Nepárový t-test	Mann Whitney test
2 skupiny dat párově:	Párový t-test	Wilcoxon test, znaménkový test
Více skupin nepárově:	ANOVA	Kruskal- Wallis test
Korelace:	Pearsonův koeficient	Spearmanův koeficient

ÚVOD DO TESTOVANIA HYPOTÉZ

Častou úlohou štatistika je na základe dát overiť predpoklady o parametroch, alebo type rozloženia, z ktorého pochádza náhodný výber. Taký predpoklad sa nazýva NULOVÁ HYPOTÉZA (označenie H_0). H_0 vyjadruje nejaký teoretický predpoklad, často skeptického rázu a užívateľ ju musí stanoviť dopredu, bez prihliadnutia k dátovému súboru. Proti H_0 staviame ALTERNATÍVNU HYPOTÉZU (označenie H_A), ktorá nám hovorí, čo platí, v prípade, že neplatí H_0. Alternatívna hypotéza je formulovaná tak, aby mohla platiť len jedna z týchto dvoch hypotéz. Pravdivosť H_A by znamenala objavenie nejakých nových skutočností, alebo zásadnejšiu zmenu v doterajších predstavách. Napríklad výskumník by chcel na základe dát preveriť tézu (nový objav), že pasívne fajčenie škodí zdraviu.

Testovaním hypotéz sa myslí rozhodovací postup, ktorý je založený na danom náhodnom výbere a pomocou ktorého rozhodneme o zamietnutí, či nezamietnutí H_0.

OBOJ-PRAVO-ĽAVO STRANNÁ ALTERNATÍVA

X1, ..., Xn . . . náhodný výber
h(ζ) . . . parametrická funkcia

L(ζ) . . . rozloženie náhodného výberu
c . . . číselná konštanta

OBOJSTRANNÁ ALTERANTÍVA

H_0: h(ζ) = c . . . Jednoduchá nulová hypotéza

H_A: h(ζ) ≠ c . . . Zložená obojstranná alternatívna hypotéza

PRAVOSTRANNÁ ALTERNATÍVA

H_0: h(ζ) ≤ c . . . Zložená ľavostranná nulová hypotéza

H_A: h(ζ) > c . . . Zložená pravostranná alternatívna hypotéza

ĽAVOSTRANNÁ ALTERNATÍVA

H_0: h(ζ) ≥ c . . . Zložená pravostranná nulová hypotéza

H_A: h(ζ) < c . . . Zložená ľavostranná alternatívna hypotéza

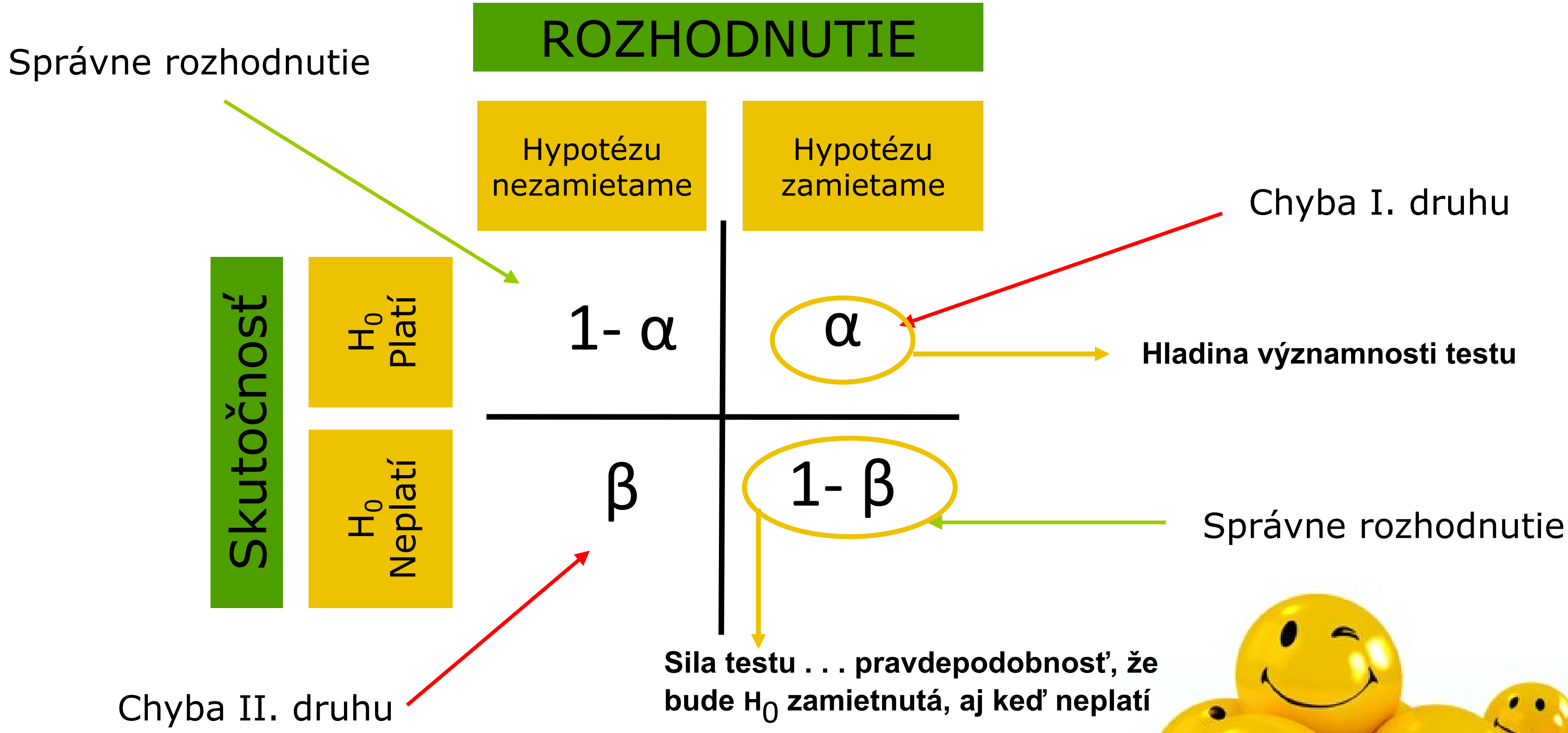
Pomocou testovania H_0 proti H_A
zamietneme, alebo nezamietneme NULOVÚ HYPOTÉZU

Možné chyby při testování hypotéz

Pri testovaní hypotéz sa môžeme dopustiť jednej z dvoch chýb:

CHYBA 1.DRUHU . . . H_0 zamietneme, aj keď v skutočnosti platí

CHYBA 2.DRUHU . . . H_0 nezamietneme, aj keď v skutočnosti neplatí



skutočnosť	**rozhodnutie**	
	zdravý	**nemocný**
som zdravý	**zdravý a neliečený**	**zdravý a liečený**
som nemocný	**nemocný a neliečený**	**nemocný a liečený**

ŠTATISTICKÉ TESTOVANIE

Štatistické testovanie odpovedá na otázku: „Je pozorovaný rozdiel náhodný?“

Odpoveď na otázku získame pomocou štatistického modelu

– pomocou TESTOVEJ ŠTATISTIKY

$$\textbf{Testová štatistika} = \frac{\textbf{Pozorovaná hodnota – Očakávaná hodnota}}{\textbf{Variabilita dát}} * \sqrt{\textbf{Veľkosť vzorku}}$$

KRITICKÝ OBOR

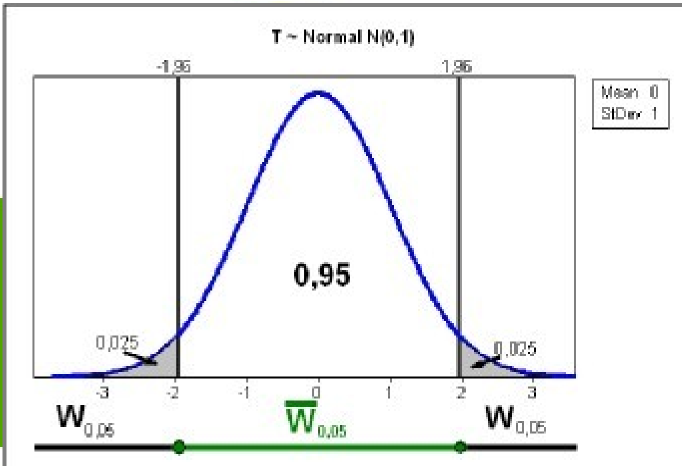
Množina všetkých hodnôt, ktorých môže testové kritérium nadobudnúť, sa rozpadá na obor nezamietnutia $\mathbf{H}_0$ a obor zamietnutia $\mathbf{H}_0$ (nazýva sa tiež kritický obor). Tieto dva obory sú oddelené kritickými hodnotami (pre danú hladinu významnosti α sa dá nájsť v štatistických tabuľkách).

Ak číselná realizácia testovej štatistiky (t_0) padne do kritického oboru,

potom $\mathbf{H}_0$ zamietame na hladine významnosti α a znamená to skutočné vyv

Ak t_0 padne do oboru nezamietnutia, potom ide jedine o mlčanie,

ktoré platnosť nulovej hypotézy len pripúšťa.



INTERVALY SPOĽAHLIVOSTI

p HODNOTA

p-hodnota

Významnosť hypotézy hodnotíme podľa získanej tzv. p-hodnoty, ktorá vyjadruje **pravdepodobnosť, s akou číselné realizácie výberu podporujú H_0, ak je pravdivá.**

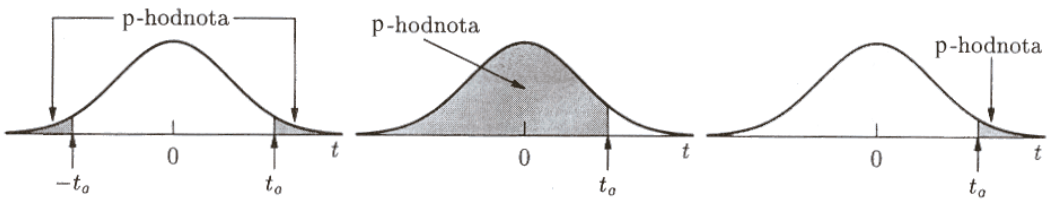
P-hodnotu porovnáme s α (hladina významnosti, stanovujeme ju na 0,05, tzn., že pripúšťame 5% chybu testu, teda, že zamietneme H_0, aj keď v skutočnosti platí).

P-hodnotu získame pri testovaní hypotéz v štatistickom software.

Ak je p-hodnota **≤ α, potom H_0 zamietame** na hladine významnosti α a prijímame H_A.

Ak je p-hodnota **> α, potom H_0 nezamietame** na hladine významnosti α.

P-hodnota vyjadruje pravdepodobnosť za platnosti H_0, s ktorou by sme získali rovnakú alebo extrémnejšiu hodnotu testovej štatistiky.



Doporučený postup pri testovaní hypotéz

1. Stanovíme $\mathbf{H_0}$ a $\mathbf{H_A}$. Pritom je vhodné zvoliť ako $\mathbf{H_A}$ ten predpoklad, ktorého prijatie znamená závažné opatrenie a malo by k nemu dôjsť iba s malým rizikom omylu.
2. Zvolíme hladinu významnosti α. Spravidla volíme $\alpha = 0{,}05$, menej často 0,1 nebo 0,01.
3. Nájdeme vhodné testové kritérium a na základe zistených dát vypočítame jeho realizáciu.
4.

a) Ak testujeme pomocou kritického oboru, potom ho stanovíme. Ak realizácia t_0 padla do kritického oboru, $\mathbf{H_0}$ zamietame na hladine významnosti α a prijímame $\mathbf{H_A}$. V opačnom prípade $\mathbf{H_0}$ nezamietame na hladine významnosti α.

b) Ak testujeme pomocou p-hodnoty, vypočítame ju a porovnáme ju s hladinou významnosti α. Ak $p \le \alpha$, potom $\mathbf{H_0}$ zamietame na hladine významnosti α a prijímame $\mathbf{H_A}$. Ak je $p > \alpha$, potom $\mathbf{H_0}$ nezamietame na hladine významnosti α.

5. Na základe rozhodnutí, ktoré sme učinili o $\mathbf{H_0}$, uskutočníme nejaké konkrétne opatrenia.