

TESTOVÁNÍ STATISTICKÝCH HYPOTÉZ

1

Vytvořeno s podporou projektu Průřezová inovace studijních programů Lesnické a dřevařské fakulty MENDELU v Brně (LDF) s ohledem na disciplíny společného základu (reg. č. CZ.1.07/2.2.00/28.0021) za přispění finančních prostředků EU a státního rozpočtu České republiky.

TESTOVÁNÍ STATISTICKÝCH HYPOTÉZ –ZÁKLADNÍ POJMY

teorie testů viz **teorie text I, str. 102 – 111**
testy pro 1 a 2 výběry **teorie text I, str. 112 – 134**

Statistická hypotéza je určitá domněnka (předpoklad) o jakékoli vlastnosti **ZÁKLADNÍHO SOUBORU**. O ZS „nic“ nevíme „jistě, stoprocentně“, pokud ho neznáme celý (což je naprosto výjimečný případ).

Pokud řekneme „ZS má normální rozdělení“, je to jen naše domněnka, protože neznáme všechny prvky ZS. Proto **každé tvrzení (hypotézu) o ZS musíme testovat**. Teprve na základě testu, který vychází z naměřených (výběrových) dat můžeme rozhodnout, zda tuto domněnku (hypotézu) můžeme považovat za platnou nebo za neplatnou.

Test statistické hypotézy je pravidlo (kritérium), které na základě **dat náhodného výběru** objektivně doporučuje rozhodnutí, má-li být **ověřovaná hypotéza zamítnuta či nikoliv**.

2

DRUHY HYPOTÉZ

Pro každý test musíme formulovat nulovou a alternativní hypotézu:

Testovaná hypotéza se nazývá **nulová hypotéza (H_0)**. Předpokládáme, že platí, pokud nemáme k dispozici **dostatečný statistický důkaz** její neplatnosti.

Pokud zamítneme platnost nulové hypotézy, předpokládáme, že platí **alternativní hypotéza (H_1)**.

H_0 a H_1 musí být formulovány tak, aby nebyla možná žádná „třetí možnost“. Buď platí H_0 nebo H_1 , nic dalšího není možné.

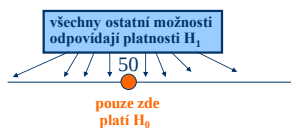
3

DRUHY HYPOTÉZ

Hypotézy se mohou formulovat jako **oboustranné** nebo jako **jednostranné**.

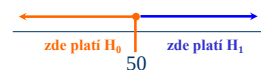
Oboustranná hypotéza:

$$H_0: \mu = 50 \quad H_1: \mu \neq 50$$



Jednostranná hypotéza:

$$H_0: \mu \leq 50 \quad H_1: \mu > 50$$



$$(H_0: \mu \geq 50 \quad H_1: \mu < 50)$$

4

TESTOVÉ KRITÉRIUM

Testy statistických hypotéz jsou obecně založeny na **testovém kritériu** (náhodná veličina, jejíž rozdělení je známo **pro případ platnosti i neplatnosti nulové hypotézy**).



5

TESTOVÉ KRITÉRIUM

Interval možných hodnot testového kritéria je rozdělen na dvě části – **interval nezamítnutí (přijetí)** a **interval zamítnutí (nepřijetí)** testované hypotézy.

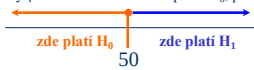
Interval zamítnutí je ta část všech možných hodnot testového kritéria, kde je vysoce nepravděpodobné, že by testové kritérium mohlo nabýt tyto hodnoty (**za předpokladu, že platí nulová hypotéza**).

Hranice intervalů tvoří **kritický bod (kritická hodnota)**, což je kvantil (hodnota) testového kritéria určený na základě hladiny významnosti α a (zpravidla) velikosti výběru n .

6

OBECNÝ PRINCIP
STATISTICKÉHO TESTU

Budeme testovat hypotézy $H_0: \mu \leq 50$ oproti $H_1: \mu > 50$.
Za ZS zde považujeme „nekonečně“ velký základní soubor (ZS), jehož střední hodnota μ je nejvýše rovna hodnotě 50. Jedná se tedy o jednostrannou hypotézu.
Pro všechny hodnoty μ menší nebo rovno 50 platí H_0 , pro ostatní hodnoty H_1 .

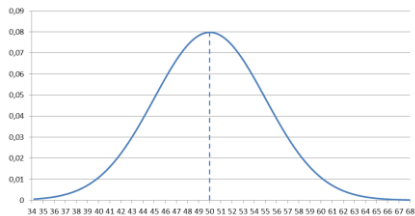


Problémem je, že my neznáme (a nikdy znát nebudeme) celý ZS. K dispozici máme pouze jeho podmnožinu – výběrový soubor (VS). Otázka, kterou řešíme, tedy zní:
Můžeme na základě dat z VS s dostatečnou pravděpodobností zamítnout hypotézu, že ZS (z něhož pochází náš VS) má střední hodnotu menší nebo rovnou 50?

7

OBECNÝ PRINCIP
STATISTICKÉHO TESTU

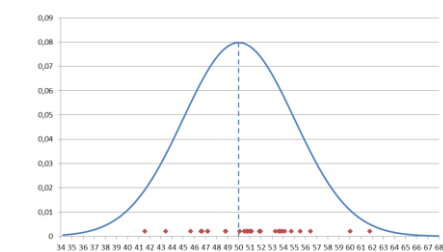
Nejprve si musíme umět představit „situaci“, tj. jak by vypadala data, pokud by platila nulová hypotéza. Tento případ pro střední hodnotu 50 ukazuje obrázek. Jedná se o normální rozdělení se střední hodnotou 50. Na ose X jsou hodnoty souboru (číselná osa), na ose Y je frekvenční funkce (tj. pravděpodobnosti výskytu jednotlivých hodnot).



8

OBECNÝ PRINCIP
STATISTICKÉHO TESTU

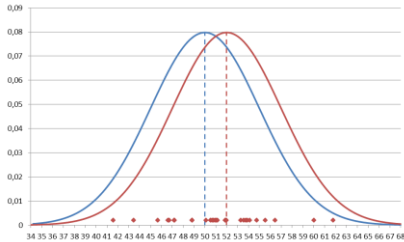
Poté naměříme data – získáme VS (data označena červenými kosočtverci). Vidíme, že střední hodnota dat bude vychýlena poněkud doprava - výpočtem zjistíme, že výběrový průměr je přibližně 52.



9

OBECNÝ PRINCIP
STATISTICKÉHO TESTU

Znamená to, že náš výběrový soubor má rozdělení (vzhledem k dostatečnému rozsahu výběru – 30 hodnot - předpokládáme normální rozdělení) se střední hodnotou 52 (červená křivka).



10

OBECNÝ PRINCIP
STATISTICKÉHO TESTU

Otázka nyní zní:

Je odchylka teoretického (modrého) a experimentálního (červeného) rozdělení tak velká, že nulovou hypotézu mohu s dostatečnou spolehlivostí zamítnout (potom bychom zamítli nulovou hypotézu a „věřili“ v platnost alternativní hypotézy)

nebo

je odchylka mezi teoretickým a experimentálním rozdělením natolik malá (pouze náhodně daná našim konkrétním výběrem), že nemůžeme s dostatečnou pravděpodobností vyloučit, že kdybychom měřili velký („nekonečný“) počet hodnot, získali bychom rozdělení odpovídající teoretickému (potom bychom nemohli nulovou hypotézu zamítnout)?

11

OBECNÝ PRINCIP
STATISTICKÉHO TESTU

Z této otázky vyplývá, že podstatou testu je rozhodnutí, zda s „dostatečnou“ spolehlivostí můžeme zamítnout tvrzení obsažené v nulové hypotéze. Onu „dostatečnou“ spolehlivost si určujeme sami, nejobvyklejší hodnota je 95 %. Znamená to, že nulovou hypotézu zamítneme pouze tehdy, pokud data VS umožňují toto zamítnutí učinit s pravděpodobností nejméně 95 %.

Musíme se tedy stanovit jednoznačnou hranici, která nám určí, v jakém intervalu možných hodnot testované proměnné se přikloníme k zamítnutí nulové hypotézy a v jakém intervalu k tomu ještě nebudeme mít dostatečně silný „statistický důkaz“ (dostatečně silnou spolehlivost).

12

OBECNÝ PRINCIP
STATISTICKÉHO TESTU

Touto hranicí je kritická hodnota. Určuje se pomocí hodnoty α (u statistických testů se nazývá **chyba prvního druhu** – podrobněji viz dále). Pokud si nastavíme např. α na hodnotu 0,05 („5%“), potom říkáme, že nulovou hypotézu zamítneme pouze v tom případě, kdy budeme mít minimální pravděpodobnost zamítnutí nulové hypotézy $1 - 0,05 = 0,95$ („95%“).

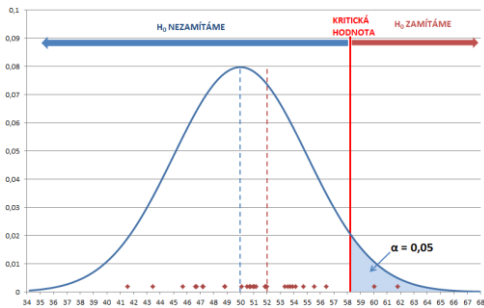
Kritickou hodnotu (KH) stanovíme tak, že musíme **určit pro teoretické rozdělení hodnotu (kvantil) pro pravděpodobnost α** (obvykle 0,05 - v případě oboustranného testu by tato pravděpodobnost byla dělená na poloviny, stejně jako u intervalů spolehlivosti).

KH určíme např. v Excelu podle funkce **Q.INV**, kde na místo čtverečku vložíme název použitého rozdělení (např. **T.INV** pro t-rozdělení, **CHISQ.INV** pro chi-kvadrát rozdělení, **NORM.INV** pro normální rozdělení apod.). Do funkce zadáme jeho parametry, tj. požadovanou pravděpodobnost (**α pro jednostranný test nebo $\alpha/2$ pro oboustranný test**) a zpravidla **počet stupňů volnosti** (obvykle **$n-1$**). Situaci ukazuje následující obrázek.

V tomto příkladu je KH 58,38.

13

OBECNÝ PRINCIP
STATISTICKÉHO TESTU



14

OBECNÝ PRINCIP
STATISTICKÉHO TESTU

S KH srovnáváme hodnotu vypočítanou přímo z dat VS – **testové kritérium (TK)**. Tedy srovnáváme kvantil stanovený pro teoretické rozdělení (nulovou hypotézu) – KH - s kvantilem stanoveným z VS (měřených dat), tedy pro alternativní hypotézu.

Pokud vypočítané TK bude vlevo od KH (**$TK < KH$**), nebudeme mít dostatečnou pravděpodobnost pro zamítnutí nulové hypotézy (protože KH určuje min. možnou pravděpodobnost zamítnutí nulové hypotézy), tedy **nulovou hypotézu nezamítneme**.

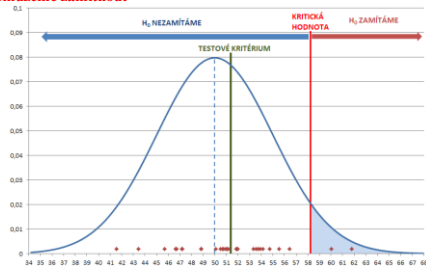
Pokud vypočítané testové kritérium bude vpravo od KH (**$TK > KH$**), budeme mít dostatečnou pravděpodobnost pro zamítnutí nulové hypotézy, neboť potom víme, že nulovou hypotézu zamítáme nejméně s $(1 - \alpha)\%$ pravděpodobností (obvykle $1 - 0,05 = 0,95$ – „95% pravděpodobnost“). Pokud TK padne do tohoto intervalu, je velmi málo pravděpodobné, že by data z VS mohla pocházet z teoretického rozdělení (v našem příkladu z normálního rozdělení se střední hodnotou 50). **Nulovou hypotézu tedy zamítneme**.

$TK < KH \Rightarrow$ nezamítáme nulovou hypotézu
 $TK > KH \Rightarrow$ zamítáme nulovou hypotézu

15

OBECNÝ PRINCIP
STATISTICKÉHO TESTU

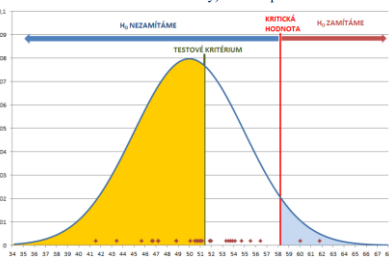
TK se stanovuje podle vzorce specifického pro daný test. V našem případě má TK hodnotu 51,37. TK je menší než KH (51,4 < 58,4), nulovou hypotézu tedy nemůžeme zamítnout



16

OBECNÝ PRINCIP
STATISTICKÉHO TESTU

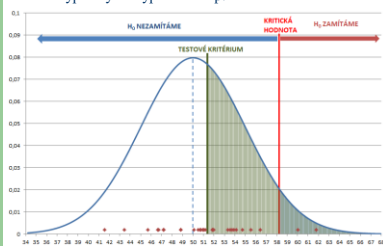
TK také vymezuje skutečnou maximální pravděpodobnost, s jakou můžeme nulovou hypotézu zamítnout. Je to plocha vlevo od TK pod teoretickým rozdělením (tj. distribuční funkce této náhodné veličiny) – žlutá plocha na obrázku.



17

OBECNÝ PRINCIP
STATISTICKÉHO TESTU

Tato pravděpodobnost se prakticky obvykle určuje tak, že se stanoví tzv. p-hodnota (podrobněji viz dále), což je zelená plocha na obrázku a pravděpodobnost zamítnutí nulové hypotézy se vypočítá $1 - p$.

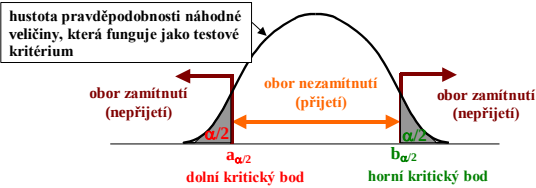


18

V tomto příkladu je p-hodnota asi 0,40, tedy maximální pravděpodobnost zamítnutí nulové hypotézy je $1 - 0,4 = 0,6$ („60%“). Nedosahuje tedy požadované hranice pravděpodobnosti zamítnutí 0,95 („95%“), což je další „důkaz“ toho, že nemůžeme zamítnout nulovou hypotézu.

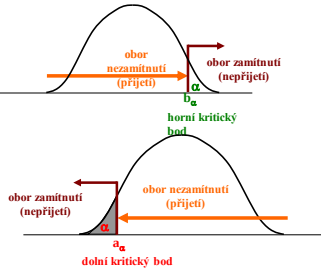
TESTOVÉ KRITÉRIUM
PRO OBOUSTRANNÝ TEST

U **oboustranného testu** (který je nejčastější) používáme **dvě kritické hodnoty**, které určujeme na základě **pravděpodobnosti $\alpha/2$** .



19

TESTOVÉ KRITÉRIUM
PRO JEDNOSTRANNÝ TEST



U **jednostranného testu** (který používáme podobně jako jednostranné intervaly spolehlivosti pouze ve speciálních případech, např. při testování, zda nějaká hodnota přesáhla v ZS zadanou hranici, normu, apod.) používáme **jednu kritickou hodnotu**, kterou určujeme na základě **pravděpodobnosti α** .

20

OBECNÝ POSTUP TESTU

- ✦ formulace nulové hypotézy (H_0) a alternativní hypotézy (H_1).
- ✦ volba chyby I. druhu α (a toho vyplývající požadavku na zamítnutí H_0 s minimální pravděpodobností $1 - \alpha$).
- ✦ volba druhu testu a testového kritéria
- ✦ určení kritického oboru (oboru nepřijetí) testového kritéria na základě jejího rozdělení pravděpodobnosti a hladiny významnosti
- ✦ rozhodnutí o výsledku testu, tj. zda
 - ♦ zamítnout nulovou hypotézu (jestliže vypočítaná hodnota testového kritéria padne do oboru nepřijetí),
 - ♦ nezamítnout nulovou hypotézu (jestliže vypočítaná hodnota testového kritéria padne do oboru přijetí).

Říkáme, že výsledek testu je **významný (signifikantní) na hladině významnosti α** , pokud jeho rozhodnutí vede k zamítnutí nulové hypotézy s využitím zvolené hodnoty α .

Pokud H_0 nezamítneme, neznamená to, že určitě („stoprocentně“) platí, znamená to pouze, že nemáme dostatečný **statistický důkaz** její neplatnosti.

21

POSTUP TESTOVÁNÍ NA PŘÍKLADU

Byla prověřována správnost měření výškoměru. Byla 15 x změřena výška, jejíž hodnota je přesně známa ($\mu_0 = 20$ m). Z výsledků měření se získal průměr měřených výšek $= 19,2$ m se směrodatnou odchylkou $S = 1,1$ m. Stanovte, zda-li výškoměr měří správně.

1 formulace H_0 a H_1 :

H_0 : Výškoměr měří správně.

H_1 : $\mu \neq \mu_0$ (výběrový soubor o parametrech $\bar{x}$ a S^2 pochází ze základního souboru s rozdělením $N(\mu, \sigma^2)$, přičemž parametr μ se rovná normované hodnotě μ_0).

22

POSTUP TESTOVÁNÍ NA PŘÍKLADU

H_1 : Výškoměr **ne**měří správně.

H_1 : $\mu \neq \mu_0$ (výběrový soubor o parametrech $\bar{x}$ a S^2 pochází ze základního souboru s rozdělením $N(\mu, \sigma^2)$, přičemž parametr μ se **nerovná** normované hodnotě μ_0).

2 stanovení hodnoty α :

chybu I. druhu stanovíme na 0.05, tj. pro zamítnutí nulové hypotézy potřebujeme pravděpodobnost nejméně $1 - \alpha = 0.95$.

23

POSTUP TESTOVÁNÍ NA PŘÍKLADU

3 výběr typu testu (testového kritéria)

vzhledem k charakteru úlohy dané formulací hypotéz vybereme **test střední hodnoty pro jeden výběr**, který testuje námi zvolené hypotézy.

Testové kritérium je

$$t = (\bar{x} - \mu_0) \cdot \frac{\sqrt{n}}{S} \qquad t = (19,2 - 20) \cdot \frac{\sqrt{15}}{1,1} = -2,82$$

Pro tento test jako testové kritérium **používáme t-rozdělení** a jeho **konkrétní vypočítaná hodnota** pro naše zadání je **- 2,82**. Tuto hodnotu budeme porovnávat k kritickým mezím.

24

POSTUP TESTOVÁNÍ NA PŘÍKLADU

4 stanovení kritické hodnoty

vzhledem k tomu, že pracujeme (podle formulace hypotéz a typu úlohy) s **oboustranným testem**, musíme zvolit **2 kritické hodnoty** – dolní a horní. Studentovo t-rozdělení je symetrické kolem 0 – stačí nám tedy určit jednu hodnotu a použít ji s kladným i záporným znaménkem.

Využijeme funkce Excelu **=T.INV (0,975;14)**, kde 0,975 je hodnota $1-\alpha/2 = 1 - 0,05/2 = 0,975$. Hodnota 14 je počet stupňů volnosti $(n-1)=15-1$. **Výsledná hodnota je 2,145**. Tím jsou určeny **kritické meze -2,145 a +2,145**.

POSTUP TESTOVÁNÍ NA PŘÍKLADU

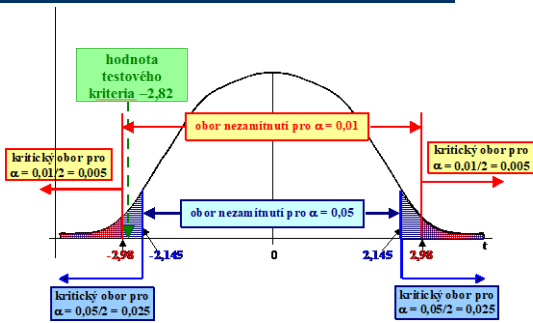
5 rozhodnutí o výsledku testu

kritická hodnota **-2,145** je menší než vypočítaná hodnota testového kritéria **-2,82**, znamená to, že **-2,82** padlo do oboru nepřijetí – tedy **nulovou hypotézu zamítáme**.

Závěr testu:

Výškoměr neměří správně. Odchylka naměřených hodnot od správné hodnoty je statisticky významná – tj. s pravděpodobností nejméně 95% se přikláname k možnosti, že i kdybychom měli velký (teoreticky „nekonečný“) počet výšek, střední hodnota by se nerovнала přesně 20 m.

POSTUP TESTOVÁNÍ NA PŘÍKLADU



POSTUP TESTOVÁNÍ NA PŘÍKLADU

Z předcházejícího obrázku vyplývá, že H_0 můžeme zamítnout s pravděpodobností vyšší než 95 % - testové kritérium (zelená šipka) je mimo „modrý 95% interval, ale nemůžeme ji zamítnout s pravděpodobností 99% - testové kritérium je uvnitř „červeného“ intervalu, který byl sestrojen pro $\alpha = 0,01$. Zatím tedy můžeme konstatovat, že H_0 můžeme zamítnout s pravděpodobností mezi 95 – 99%.

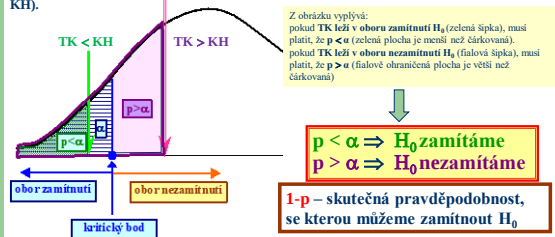
Výpočet skutečné pravděpodobnosti zamítnutí je spojen s výpočtem tzv. **p-hodnoty** (viz následující snímek) – tato hodnota je velmi důležitá, protože **poskytuje druhý způsob, jak určit výsledek statistického testu**. Navíc je tento způsob prezentace výsledku testu daleko častější než srovnání kritické hodnoty a testového kritéria. V podstatě všechny statistické programy uvádí jako jediný výsledek testu právě p-hodnotu, podle které musíte umět rozhodnout o výsledku testu.

28

p-HODNOTA

p-hodnota je pravděpodobnost, že získáme stejné nebo extrémnější testové kritérium než je vypočítané, **za předpokladu, že ve skutečnosti platí nulová hypotéza**.

p-hodnota je tedy pravděpodobnost od testového kritéria (TK) směrem „ven“ z rozdělení – na obrázku zelená ploška (pro TK nižší než KH) a ploška fialově orámovaná (pro TK vyšší než KH).



29

ROZHODNUTÍ O VÝSLEDKU TESTU SHRNUTÍ

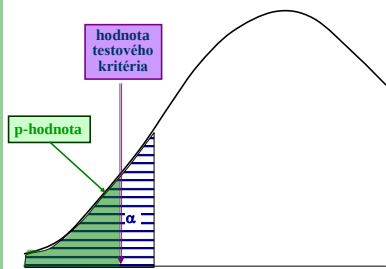
Máme tedy k dispozici dva způsoby, jak rozhodnout o výsledku testu:

- $TK < KH \Rightarrow$ nezamítáme nulovou hypotézu
- $TK > KH \Rightarrow$ zamítáme nulovou hypotézu
- $p > \alpha \Rightarrow$ nezamítáme nulovou hypotézu
- $p < \alpha \Rightarrow$ zamítáme nulovou hypotézu

Pamatujte, že vždy můžete srovnávat mezi sebou pouze vzájemně srovnatelné hodnoty, tj. kvantily (hodnoty) rozdělení (TK a KH), resp. pravděpodobnosti (p a α). Samozřejmě pro stejná data a stejný test oba způsoby vyhodnocení musí poskytnout stejnou interpretaci.

30

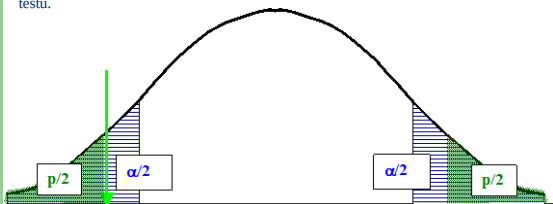
p – HODNOTA (jednostranný test)



31

p-HODNOTA (oboustranný test)

U oboustranného testu se p-hodnota stejně jako α „dělí“ na poloviny na obě strany testu.



32

p-HODNOTA

Podle p - hodnoty můžeme pouze zamítnout nebo nezamítnout H_0 . **Nemůžeme podle ní „srovnávat“ různé testy a jejich výsledky.** Není totiž možné obecně říci, že pokud dosáhneme v testu hodnotu $p = 0,005$, je H_0 tohoto testu zamítnuta „spolehlivěji“ než u testu, kde $p = 0,01$.

- Důvodem je to, že p-hodnota závisí na:
- ✦ velikosti efektu - rozdílu mezi nulovou a alternativní hypotézou
 - ✦ velikosti výběru

Příklady uvádí následující snímky.

33

p-HODNOTA

Tento příklad ukazuje vliv velikosti výběru na výsledek testu (na hodnotu p), přičemž velikost efektu je stále stejná ($100-99 = 1$). Pokud máme malý výběr ($n = 10$), velikost efektu je považována za statisticky neprůkaznou (znamená to, že na základě 10 dat VS nemůžeme zamítnout H_0 , tj. tvrzení, že střední hodnota ZS se rovná 100). Ovšem pro velký výběr ($n = 1000$) je H_0 zamítnuta.

Vyplyvá z toho, že při použití hodně velkého výběru i velmi malá velikost efektu (tj. velmi malý rozdíl mezi „naměřenou“ a „teoretickou“ hodnotou) je prokázána jako statisticky významná.

Střední hodnota pro H_1	Střední hodnota pro H_0	n	t	t_{krit}	p
99	100	10	0,300	2,262	0,771
99	100	100	0,995	1,984	0,322
99	100	1000	3,161	1,962	0,002

34

p-HODNOTA

Tento příklad ukazuje, že i při stejné velikosti výběru (zde $n = 20$) závisí výsledek testu na velikosti efektu. Pokud jsou H_0 a H_1 od sebe „dostatečně“ vzdáleny, i malý výběr vede ke statistickému zamítnutí H_0 .

Střední hodnota pro H_1	Střední hodnota pro H_0	n	t	t_{krit}	p
99	100	20	0,436	2,093	0,668
97	100	20	1,308	2,093	0,207
95	100	20	2,179	2,093	0,042
90	100	20	4,359	2,093	0,000

Musíme si ale položit otázku:
jaká velikost efektu je pro řešení konkrétního praktického problému prakticky (reálně) významná?

35

STATISTICKÁ VÝZNAMNOST A PRAKTICKÁ DŮLEŽITOST VELIKOSTI EFEKTU

Je nutné rozlišovat mezi statistickou významností (danou p - hodnotou) a reálnou (ekologickou, ekonomickou, biologickou,...) důležitostí studované velikosti efektu !!

Praktická (reálná) důležitost rozdílu $H_0 - H_1$	Statistická významnost (podle p)	
	nevýznamná	významná
nedůležitý	n v pořádku	n příliš velké
důležitý	n příliš malé	n v pořádku

36

INTERVAL SPOLEHLIVOSTI EFEKTU A JEHO SPOJITOST S VÝSLEDKEM TESTU

Rozdíl mezi statistickou významností a praktickou důležitostí testované veličnosti efektu je možné názorně vyložit na intervalu spolehlivosti (IS) efektu. Je nutné si uvědomit, že nás zajímá, **zda je možné věřit tomu, že velikost efektu je prokazatelná v základním souboru**. IS nám se stanovenou pravděpodobností **pokryje interval, ve kterém se vyskytuje studovaný efekt v základním souboru**. Vztahy mezi statistickou a reálnou významností ukazuje obrázek na následujícím snímku.

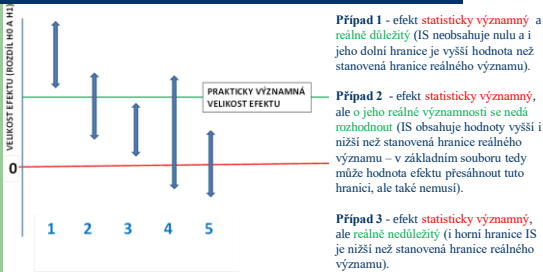
Na obrázku je uvedeno 5 možností vztahu mezi statistickou a reálnou významností.

Za **statisticky významný efekt** považujeme ten, jehož **IS neobsahuje nulu** (znamená to, že ani v základním souboru nemůžeme předpokládat, že by studovaný rozdíl mezi nulovou a alternativní hypotézou neexistoval).

Za **reálné (prakticky) důležitý** ten efekt, jehož **celý IS převyšuje hranici prakticky důležité velikosti efektu** (zde čárkovaná čára – určeno rozhodnutím výzkumníka na základě předchozích zkušeností, údajů z literatury apod. – pro konkrétní řešený problém) - znamená to, **prokázání daného efektu testem jako významného bude mít i praktickou důležitost pro řešení daného problému**.

37

STATISTICKÁ vs. REÁLNÁ VÝZNAMNOST



38

VÝZNAM STATISTICKÉ VÝZNAMNOSTI A REÁLNÉ DŮLEŽITOSTI VELIKOSTI EFEKTU

Z předchozích poznatků plyne, že **cílem statistického testu je „odhalit“ (tj. prokázat jako statisticky významný) takový rozdíl mezi nulovou a alternativní hypotézou, který má praktickou důležitost pro určitý řešený problém**. Přitom bychom chtěli tohoto cíle dosáhnout **při co nejmenším možném počtu měřených hodnot**.

Tento cíl je možné splnit pomocí tzv. **analýzy síly testu**. Tato analýza odpovídá na dvě otázky:

- jakou **minimální velikost výběru** potřebuji, aby použitý test prokázal jako významnou právě takovou velikost efektu, kterou považuji za prakticky důležitou?
- pokud mám již naměřená data, nakolik se mohu „spolehnout“ na rozhodnutí testu, tj. **jak je výsledek testu spolehlivý?**

Analýza síly testu vychází z pojmů chyb I. a II. druhu, které budou vysvětleny na následujících snímcích.

39

CHYBY TESTU

Při hodnocení výsledku testu musíme mít stále na paměti, že ani **test nerozhodně „stoprocentně“, „zcela jistě“**, protože nemáme úplnou informaci o ZS. Pamatujte tedy, že i každý výsledek statistického testu musíme brát „s rezervou“, tj. s vědomím, že rozhodnutí doporučené testem může být „chybné“, což znamená že **doporučení testu zamítnout nebo nezamítnout hypotézu nemusí být nutně v souladu s realitou** (test totiž vychází z dat VS a ta mohou být „náhodou“ tak zvláštní, že to vede test k rozhodnutí, které ve skutečnosti v ZS neplatí).

Test tedy může doporučit zamítnutí hypotézy, zatímco tato hypotéza v reálu platí nebo naopak může doporučit nezamítnutí hypotézy, která ve skutečnosti neplatí.

Naši výhodou je to, že **pravděpodobnosti těchto možných „chybných“ rozhodnutí testu si můžeme nastavit** a tím určit, nakolik bude náš test „spolehlivý“, tj. s jakou pravděpodobností mohou jeho doporučení „věřit“.

Zdůrazňuji, že **„chybou“ testu se zde rozumí pouze možnost, že test může doporučit rozhodnutí, které je v rozporu s realitou** (nikoli tedy chyba způsobená chybným výpočtem nebo chybným číslem, apod.).

40

CHYBA I. A II. DRUHU, SÍLA TESTU

Chyba I. druhu (α) – pravděpodobnost, že test **zamítne hypotézu, která ve skutečnosti platí**.

Chyba II. druhu (β) – pravděpodobnost, že test **NEzamítne hypotézu, která ve skutečnosti NEplatí**.

Síla testu ($1 - \beta$) – pravděpodobnost, že test **správně zamítne hypotézu, která ve skutečnosti neplatí**.

41

CHYBA I. A II. DRUHU

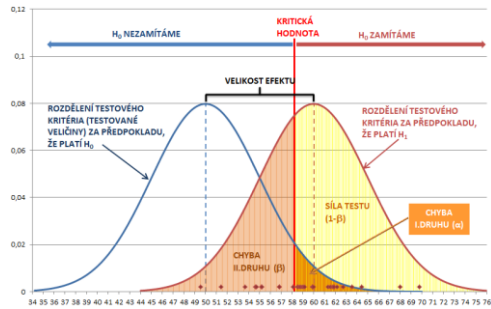
		H ₀ VE SKUTEČNOSTI	
		platí	neplatí
H ₀ PODLE ROZHODNUTÍ TESTU	platí	činíme správné rozhodnutí	dopouštíme se chyby II. druhu (s pravděp. β)
	neplatí	dopouštíme se chyby I. druhu (s pravděp. α)	činíme správné rozhodnutí (s pravděp. $S = 1 - \beta$)

42

CHYBA I. A II. DRUHU

Graficky si jednotlivé pravděpodobnosti spojené s chybami testů ukážeme na následujícím obrázku:
Předpokládáme, že pracujeme s jednostranným testem (kvůli snazšímu grafickému znázornění) s hypotézami: $H_0: \mu \leq 50$ oproti $H_1: \mu > 50$
Pokud platí, data jsou rozdělena podle „modrého“ rozdělení $N(50,5^2)$. Neměřili jsme data výběrového souboru (červené body), jehož rozdělení je znázorněno „červeně“ – $N(60,5^2)$ – to je rozdělení pro případ platnosti H_1 .
Chyba I. druhu – nastane, pokud test chybně zamítne H_0 . V tom případě v reálu platí H_0 (tedy modré rozdělení), ale TK „vyjde“ vyšší než KH – pravděpodobnost této chyby je tedy plocha pod modrým rozdělením za KH – **tmavě žlutá čárkovaná plocha**.
Chyba II. druhu – nastane, pokud test chybně nezamítne H_0 . V tom případě v reálu platí H_1 (tedy červené rozdělení), ale TK „vyjde“ menší než KH – pravděpodobnost této chyby je tedy plocha pod červeným rozdělením před KH – **oranžová plocha**.
Síla testu – je vyjádřena pravděpodobností, že test zamítne ve skutečnosti neplatnou hypotézu H_0 . Znamená to, že v reálu platí H_1 (červené rozdělení) a test správně zamítne H_0 , tedy platí že $TK > KH$, tedy síla testu je vyjádřena plochou pod červeným rozdělením za KH – **světle žlutá plocha**

CHYBA I. A II. DRUHU

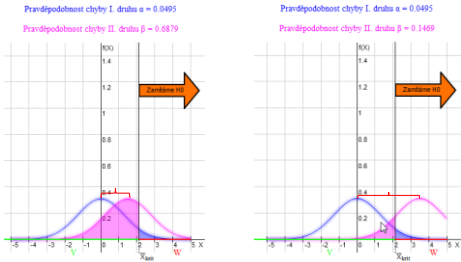


FAKTORY OVLIVŇUJÍCÍ SÍLU TESTU

- Síla testu závisí na následujících faktorech:
- ✦ odchylka mezi hodnotou testovaného parametru v nulové hypotéze a skutečnou hodnotou parametru – **velikost efektu**
 - ✦ **variabilita** (směrodatné odchylce nebo rozptylu) **základního souboru**
 - ✦ **velikost výběrového souboru**
 - ✦ **na hladina významnosti α**
 - ✦ **typ testu**

FAKTORY OVLIVŇUJÍCÍ SÍLU TESTU

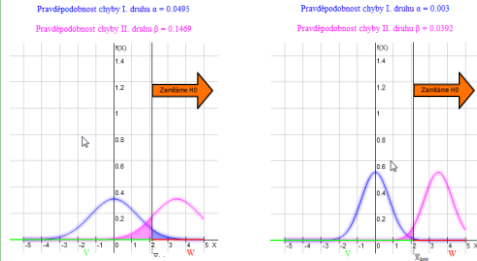
Obrázky ukazují vliv velikosti efektu (červené sorky) na sílu testu. Pokud jsou rozdělení reprezentující H_0 a H_1 blízko sebe (velikost efektu je malá), rozdělení se více „překrývají“ a velikost chyby II. druhu (fialová plocha) je větší – tedy síla testu (bílá plocha pod fialovou křivkou) menší. **Se zvětšující se velikostí efektu se zvětšuje i síla testu.**



46

FAKTORY OVLIVŇUJÍCÍ SÍLU TESTU

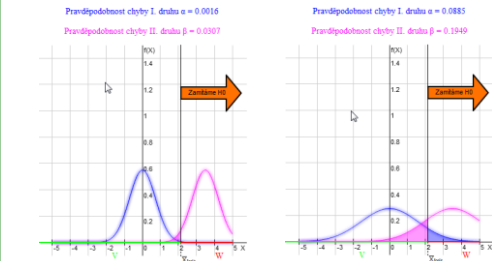
Obrázky ukazují vliv změny variability dat na sílu testu. **Variabilita dat a síla testu jsou nepřímo úměrné, tedy se zmenšující se variabilitou se zvyšuje síla testu.** Rozdělení vlevo má vyšší variabilitu dat (rozdělení jsou „širší“, „roztaženější“) a je vidět výrazně vyšší chyba II. druhu (fialová plocha). Rozdělení vpravo má nižší variabilitu (je „stříhlejší“) a chyba II. druhu je menší, čímž se zvyšuje síla testu (bílá plocha pod fialovou křivkou).



47

FAKTORY OVLIVŇUJÍCÍ SÍLU TESTU

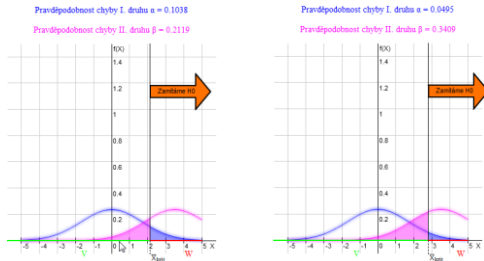
Obrázky ukazují vliv změny velikosti výběru dat (n) na sílu testu. **Velikost výběru a síla testu jsou přímo úměrné, tedy se zvětšující se velikostí výběru se zvyšuje síla testu.** Rozdělení vlevo má vyšší n (rozdělení jsou „stříhlejší“, protože s vyšším n souvisí spolehlivější odhad střední hodnoty, prostě máme víc informací o ZS) a chyba II. druhu je menší, čímž se zvyšuje síla testu (bílá plocha pod fialovou křivkou). Rozdělení vpravo má nižší n (je „širší“) a je vidět výrazně vyšší chyba II. druhu.



48

FAKTORY OVLIVŇUJÍCÍ SÍLU TESTU

Obrázek ukazuje vliv změny chyby I. druhu (α) na sílu testu. Změna hodnoty α znamená vlastně změnu kritické hodnoty. **Hodnota α a síla testu jsou přímo úměrné, se zvyšující se hodnotou chyby I. druhu se zvyšuje i síla testu** (protože klesá chyba II. druhu β). Na obrázku vlevo je α přibližně 0,10 (KH je asi 2,05) a síla testu je asi 79% (1-0,21). Na pravém obrázku je α přibližně 0,05 (KH je asi 2,9) a síla testu je asi 66% (1-0,34).



49

ÚČEL ANALÝZY SÍLY TESTU

před provedením pokusu (apriorní analýza)	
zjišťujeme	známe (zadáváme)
potřebnou velikost výběru	- chybu I. druhu (alfa) - požadovanou sílu testu (1 - beta) - velikost efektu, kterou potřebujeme detekovat

Pokud předpokládáme, že naměřená data budeme vyhodnocovat statistickým testem, měla by samotnému měření předcházet analýza síly testu, která odpoví na otázku: Kolik hodnot musím minimálně naměřit, aby test detekoval jako významnou právě takovou velikost efektu, kterou považují za prakticky důležitou a přitom zachoval akceptovatelné velikosti chyb I. i II. druhu (a tedy přijatelnou sílu testu)?

50

ÚČEL ANALÝZY SÍLY TESTU

po provedení pokusu (aposteriorní analýza)	
zjišťujeme	známe (zadáváme)
skutečnou sílu testu	- chybu I. druhu (alfa) - velikost výběru (N) - velikost efektu

Po provedení pokusu (měření) je analýza síly testu důležitá především tehdy, **pokud nezamítneme H_0** . Pokud totiž H_0 zamítneme, víme, že tak činíme nejméně s pravděpodobností $1 - \alpha$ (obvykle 0,95 – 95% a víme, že max. pravděpodobnost chyby I. druhu – tedy že H_0 byla zamítnuta chybně – je α). **Pokud ale H_0 nezamítneme, nevíme o velikosti chyby II. druhu** – tedy že H_0 byla chybně nezamítnuta, „přijata“ – **nic konkrétního**. Měli bychom ji tedy vypočítat a stanovit skutečnou sílu testu – tím zjistíme, nakolik je závěr o nezamítnutí H_0 oprávněný.

Blokové schéma postupu při provedení experimentu s využitím analýzy síly testu je zde: http://user.mendelu.cz/drapela/Statisticke_metody/Prezentace/zakladni_postup_analyzy_sily_testu.doc

51

ANALÝZA SÍLY TESTU - PŘÍKLAD

V úpravně vody byly sledovány hodnoty obsahu chloru v pitné vodě po zavedení nové čistící metody vody.
Dle normy je povolený obsah chloru ve vodě 0,3 mg.l⁻¹. Určete, zda skutečný obsah chloru ve vodě odpovídá normě. Dále potřebujeme zjistit, kolik vzorků je nutné odebrat, aby možná chyba testu, která by způsobila vážné následky, nepřesáhla 5%.

Předběžně bylo odebráno 23 vzorků (obsah Cl ve vodě vmg.l⁻¹):

0.10	0.15	0.25	0.15	0.30	0.25	0.25	0.30	0.35
0.55	0.70	0.70	0.25	0.20	0.15	0.65	0.55	0.50
0.30	0.35	0.30	0.25	0.80				

52

ANALÝZA SÍLY TESTU - PŘÍKLAD

Zadání je typická úloha na **jednostranný test**, protože nás zajímá pouze překročení normy (pokud bude voda obsahovat méně chlóru, a přitom bude z ostatních hledisek vyhovovat kvalitě pitné vody, bude to jen dobře). Formulace hypotéz bude tedy následující:

H₀: obsah Cl ≤ 0,3 mg.l⁻¹ H₁: obsah Cl > 0,3 mg.l⁻¹

53

ANALÝZA SÍLY TESTU - PŘÍKLAD

Nyní musíme analyzovat, co by obě chyby testu znamenaly z praktického hlediska.

Chyba I. druhu znamená zamítnutí ve skutečnosti platné nulové hypotézy. V tomto případě platí, že pokud je nulová hypotéza platná, je voda z hlediska obsahu chloru v pořádku. **Když test chybně tuto hypotézu zamítne, znamenalo by to, že si mylně myslíme, že ve vodě je nepovolený obsah chloru**, což by zřejmě znamenalo další úpravy zaváděné čistící metody, další laboratorní zkoušky, apod., **Tento závěr by tedy znamenal vyšší náklady vodárny, ale neznamenal by zdravotní ohrožení obyvatel.**

Pokud by došlo k **chybě II. druhu**, znamenalo by to, že vodě je vyšší než povolený obsah chloru a přitom test by „doporučil“ nulovou hypotézu o nezávadnosti vody nezamítnout. Zde by zřejmě **následky mohly být vážnější**, protože zvýšený obsah chloru ve vodě by mohl citlivějším osobám způsobit i **zdravotní obtíže**, ostatním by přinejmenším vadil např. zvýšený zápach chloru z vody.

54

ANALÝZA SÍLY TESTU - PŘÍKLAD

Pokud bychom použili k testování již odebraných 23 vzorků, použili bychom klasický **jednovýběrový t-test** (porovnání střední hodnoty naměřených dat s konstantou - normou). Testové kritérium se vypočítá podle vzorce

t = |μ - c| · √n / S

kde je c daná konstanta - hodnota normy, S je bodový odhad směrodatné odchylky, n je velikost výběru. Statistická analýza potvrdila, že výběr pochází z normálního rozdělení a jeho statistické charakteristiky jsou:

Aritmetický průměr	0.363
Median	0.300
Směrodatná odchylka	0.202
Počet hodnot	23

55

ANALÝZA SÍLY TESTU - PŘÍKLAD

Testové kritérium vyjde **1.494** a **kritická hodnota** (kvantil t-rozdělení pro α = 0,05 a 22 stupňů volnosti) je **1.717** (např. v Excelu =T.INV(0,95;22)). Hodnota 0,95 je spanovena jako 1 - α = 1 - 0,05 (pro jednostranný test počítáme s hodnotou α). Vzhledem k tomu, že **platí TK < KH**, test **neumožnil** zamítnout nulovou hypotézu, tedy **měli** bychom se přiklonit k hypotéze, že voda je z hlediska obsahu chloru v pořádku.

Jak již bylo uvedeno, v případě **nezamítnutí** nulové hypotézy je velmi důležité zjistit velikost chyby II. druhu a tím i spolehlivost závěru testu.

Vypočet síly testu a potřebnou velikost výběru vypočítáme pomocí jednoduché tabulky v Excelu, kde jsou také uvedeny všechny potřebné vzorce - zde: http://user.mendelu.cz/drapela/Statistické_metody/Data_do_cviceni/Velikost_vyberu.xls pro příklad v prezentaci použijte list „velikost výběru - I V test“

Různé řešené příklady na velikost výběru, minimální detekovatelný rozdíl hypotéz (velikost efektu) a sílu testu, které si můžete přepočítat pomocí excelovského souboru uvedeného výše, jsou k prostudování zde: http://user.mendelu.cz/drapela/Statistické_metody/Data_do_cviceni/Velikost_vyberu.doc

56

ANALÝZA SÍLY TESTU - PŘÍKLAD

Směrodatná odchylka (S)	0.202
Velikost efektu (δ)	0.063
Chyba I.druhu (α)	0.1
Požadovaná chyba II.druhu (β)	0.05
Velikost výběru (n)	23
n vypočítané	121.25 122
detekovatelný rozdíl	0.1447
t-beta	-0.221
beta skut.	0.5866
síla testu	0.4134

Vzhledem k tomu, **nulová hypotéza nebyla zamítnuta**, je prvním krokem je výpočet skutečné síly testu – tedy **aposteriorní analýza síly testu**

Směrodatnou odchylku převezmeme z výsledků výběru, velikost efektu je rozdíl mezi normou (0,3) a vypočítaným výběrovým průměrem (0,363), tedy 0,063. Chybu I. druhu máme nastavenou na obvyklé hodnotě 0,05, vzhledem k jednostrannému testu a konstrukci t-rozdělení v Excelu nastavíme hodnotu α na 0,10 (zadááme dvojnásobek požadované hodnoty) – v příkladu je použita starší verze funkce TINV, nikoli novější T.INV

Výsledná **síla testu je pouze 0,41**, tj. **asi 41 %**, tedy **chyba II. druhu je asi 59 %**.

57

ANALÝZA SÍLY TESTU - PŘÍKLAD

Jinými slovy, **test udělá chybu II. druhu** při dané velikosti výběru **téměř v 60%** (tj. nesprávně nezamítne nulovou hypotézu o přijatelném obsahu chloru ve vodě). S takovým testem nemůžeme být jistě spokojeni. Při stanovených parametrech testu by test téměř vždy „ohlásil“ bezproblémový obsah chloru, zatímco ve skutečnosti by to nebylo pravda.

Musíme tedy **určit takovou velikost výběru**, která by **zabezpečila rozhodnutí testu s dostatečnou jistotou**. V našem případě nás zajímá **hlavně chyba II. druhu** (zdůvodnění viz výše), kterou **nastavíme na 5%, tj. sílu testu na 95%**. Dále nastavíme **prakticky důležitou velikost efektu na 0,07 mg.l⁻¹**. Tím předpokládáme (např. na základě předběžných studií, údajů z literatury, apod.), že z hlediska praktických důsledků (např. možného poškození zdraví odběratelů vody) je důležitý rozdíl 0,07 mg.l⁻¹. Na základě takto stanovené velikosti výběru test jako statisticky významné prokáže zvýšení obsahu chloru ve vodě na 0,37 mg.l⁻¹ a více.

Chceme tedy **stanovit takovou velikost výběru**, která zabezpečí, **že tento rozdíl bude testem označen jako statisticky významný** (tedy existující i v základním souboru) **za předpokladu dodržení velikosti chyb I i II. druhu na hodnotě 0,05**. Tyto podmínky jsou poměrně velmi přísné, předpokládají udržení úrovně síly testu na 95 %, přičemž za obecně akceptovatelnou sílu testu (kdy následky chyby II. druhu nepředstavují hlavní riziko) se považuje asi 80%).

58

ANALÝZA SÍLY TESTU - PŘÍKLAD

Vzhledem k tomu, že **výpočet velikosti výběru (n)** je závislý na hodnotách t-rozdělení (které zase závisí na počtu stupňů volnosti, tedy **n-1**), výpočet **probíhá iterativně**, tj. v několika „kolech“ a zastaví se tehdy, až se výsledky ve dvou za sebou následujících výpočtech nezmění. Pokud zadáme výchozí hodnoty, tj. většinu ponecháme stejných, jen velikost efektu upravíme na 0,07, chyba I. druhu zůstane na 0,1, chyba II. druhu na 0,05), získáme v prvním kroku tyto výsledky:

Směrodatná odchylka (S)	0.202
Velikost efektu (δ)	0.07
Chyba I.druhu (α)	0.1
Poladovaná chyba II.druhu (β)	0.05
Velikost výběru (n)	23
n vypočítané	98.216 99
detekovatelný rozdíl	0.1447
t-beta	0.055
beta skot.	0.8218
sila testu	0.493

Vycházíme z původní velikosti výběru (23 hodnot), nově vypočítaná velikost výběru je 99 a touto hodnotou zaměníme původní velikost výběru - tj. číslo 23 přepíšeme hodnotou 99 a kurzorem klepneme do jakékoliv prázdné buňky tabulky. Tím se provede další „kolo“ výpočtu

59

ANALÝZA SÍLY TESTU - PŘÍKLAD

Výsledkem je velikost výběru 92. Znovu tímto číslem zaměníme hodnotu 99 a spustíme další kolo výpočtu. Vidíme, že výsledek se již nezměnil – **vypočítaná velikost výběru je tedy 92 měření**.

Směrodatná odchylka (S)	0.202
Velikost efektu (δ)	0.07
Chyba I.druhu (α)	0.1
Poladovaná chyba II.druhu (β)	0.05
Velikost výběru (n)	99
n vypočítané	91.848 92
detekovatelný rozdíl	0.0974
t-beta	1.7874
beta skot.	0.9385
sila testu	0.9615

Směrodatná odchylka (S)	0.202
Velikost efektu (δ)	0.07
Chyba I.druhu (α)	0.1
Poladovaná chyba II.druhu (β)	0.05
Velikost výběru (n)	92
n vypočítané	91.983 92
detekovatelný rozdíl	0.097
t-beta	1.6621
beta skot.	0.95
sila testu	0.95

60

ANALÝZA SÍLY TESTU - PŘÍKLAD

Tato velikost výběru je značně vysoká – prakticky čtyřnásobná oproti původnímu počtu měření – to by bylo zřejmě nepřijatelné z časového i finančního hlediska. Proto je nutné se pokusit přijatelným způsobem změnit podmínky testu (tj. tak, aby test měl stále akceptovatelnou spolehlivost).

Vzhledem k tomu, že jsme si zdůvodnili, že chyba I. druhu nemá takovou závažnost, můžeme její velikost zvýšit na 0.15 (tj. připouštíme, že v 15% případů je možné, že test nesprávně zamítne nulovou hypotézu a bude nám mylně „vnucovat“, že voda obsahuje příliš vysoký obsah chloru). Prepíšeme hodnotu α na 0.15 (v našem případě na 0.30, protože musíme vzhledem k jednostrannému testu zadávat dvojnásobek chyby I. druhu) a vyjde potřebná velikost výběru. Přes 2 iterace dojdeme k výsledku 62 měření.

Pokud by i tento výsledek byl příliš vysoký, museli bychom „sáhnout“ i na chybu II. druhu. Zkusíme ji nastavit na 0.10 (tedy sílu testu snížíme na 90 %), což je stále velmi „přijatelná“ hodnota, která zaručuje, pouze v jednom z deseti případů bude nesprávně „přijata“ hypotéza o správném obsahu chloru ve vodě. V tomto případě získáme jako potřebnou velikost výběru 46 měření.

Tato řešení z Excelu jsou na následujícím snímku.

61

ANALÝZA SÍLY TESTU - PŘÍKLAD

Směrodatná odchylka (S)	0.202	Směrodatná odchylka (S)	0.202
Velikost efektu (δ)	0.07	Velikost efektu (δ)	0.07
Chyba I.druhu (α)	0.3	Chyba I.druhu (α)	0.3
Požadovaná chyba II.druhu (β)	0.05	Požadovaná chyba II.druhu (β)	0.1
Velikost výběru (n)	62	Velikost výběru (n)	46
n vypočítané	61.407	n vypočítané	45.955
detekovatelný rozdíl	0.0697	detekovatelný rozdíl	0.07
t-beta	1.6833	t-beta	1.3018
beta skut.	0.0487	beta skut.	0.0998
sila testu	0.9513	sila testu	0.9002

62

ANALÝZA SÍLY TESTU - PŘÍKLAD

pokud budeme tento výsledek považovat za přijatelný, můžeme jej interpretovat takto:

pokud provedeme 46 měření obsahu chloru ve vodě, tato velikost výběru zaručí, že test detekuje hodnotu 0,07 mg·l⁻¹ jako statisticky významný rozdíl střední hodnoty měřených dat od normy při zachování chyby I. druhu 0,15 a síle testu 90 %. Připouštíme tedy 15 % riziko, že test nesprávně zamítne nulovou hypotézu a 10 % riziko, že test nesprávně „přijme“ nulovou hypotézu.

63

TYPY TESTŮ

- Parametrické testy:
- testují **parametr základního souboru** (např. střední hodnotu, rozptyl, ...),
 - mají **vyšší sílu testu** (tedy schopnost správně zamítnout ve skutečnosti neplatnou nulovou hypotézu) než testy neparametrické, jsou tedy „přísnější na data“,
 - vyžadují **splnění určitých podmínek** (obvykle normální rozdělení, nepřítomnost extrémů, dostatečný rozsah výběru)
 - pokud jsou splněny podmínky, **používají se prioritně**

64

TYPY TESTŮ

- Neparametrické testy:
- testujeme **jinou hypotézu o rozdělení základního souboru než je hypotéza o jeho parametru**
 - mají **nižší sílu testu** (tedy schopnost správně zamítnout ve skutečnosti neplatnou nulovou hypotézu) než testy parametrické, mají vyšší tendenci „nezamítnout“ nulovou hypotézu (v hraničních případech – kdy je testové kritérium velmi blízké kritické hodnotě – mohou vést k nezamítnutí nulové hypotézy, zatímco parametrický test pro stejná data nulovou hypotézu zamítne)
 - **pro stejnou sílu testu je nutná větší velikost výběru** než u parametrických testů
 - **širší použití než parametrické** (lze testovat většinou i ZS hodnot **slovních** znaků, především **ordinálních**, tj. rozlišujících dle relace (např. pořadové testy), některé dokonce i pro hodnoty **nominálních** znaků, tj. zařazujících jen do skupin (např. znaménkový test)
 - **nezávislé na rozdělení a na přítomnosti extrémních hodnot, vhodné pro malé výběry**
 - všechny obvyklé parametrické testy mají své neparametrické „obdoby“

65

TYPY TESTŮ

- Pro jeden výběr**
- používá se pro **porovnávání parametru odhadnutého z měřených dat** (např. střední hodnotu, rozptyl,...) – P – **se zadanou konstantou** (např. předem známou hodnotou, teoretickou hodnotou, normou, ...) – K
- Pro dva výběry**
- používá se obvykle pro **porovnání parametrů odhadnutých ze dvou výběrů**, nejobvyklejší úloha je **testování shody těchto parametrů** (např. střední tloušťky dvou porostů se rovnají) – **dvouvýběrový test pro nezávislé výběry** (oba výběry, ze kterých se odhadovaly parametry, jsou vzájemně nezávislé, např. dva rozdílné porosty),
 - zvláštním případem je **porovnání dvou měření prováděných na stejných jedincích** (např. dvěma metodami, dvěma přístroji, před a po provedení určitého zásahu, ve dvou časových okamžicích, – obvykle se opět testuje shoda měření na určitém jedinci) – **dvouvýběrový test pro závislé výběry (párový test)** – měřené hodnoty už nejsou nezávislé, ale „závisí“ na měřeném jedinci.

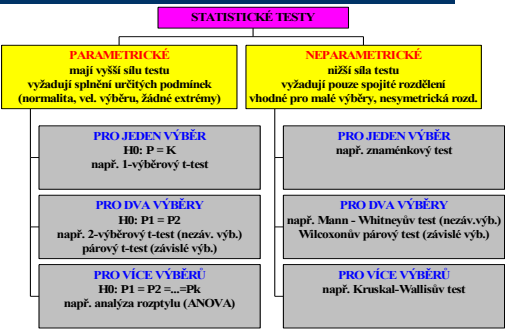
66

TYPY TESTŮ

Pro více výběrů totéž jako pro dva výběry, ale **testujeme simultánní hypotézu, že se zároveň rovnají střední hodnoty odhadnuté z několika** (tři a více) **výběrů**. Používá se speciální metodika umožňující testování simultánní hypotézy – blíže viz **analýza rozptylu (ANOVA)**.

67

TYPY TESTŮ



68

TYPY TESTŮ

Podrobněji o jednotlivých testech pro střední hodnoty, rozptyly, četnosti apod. pro 1 a 2 výběry viz **teorie text I, str. 111 – 134**
http://user.mendelu.cz/drapela/Statisticke_metody/teorie%20text%20I.pdf

dále „Návody k použití statistických programů“ (provedení v příslušném programu a interpretace výsledků)

Excel
http://user.mendelu.cz/drapela/Statisticke_metody/Navody_k_pouziti_statistickych_programu/Excel/T-test.EXE
http://user.mendelu.cz/drapela/Statisticke_metody/Navody_k_pouziti_statistickych_programu/Excel/Parovy_T-test.EXE
http://user.mendelu.cz/drapela/Statisticke_metody/Navody_k_pouziti_statistickych_programu/Excel/Kolmogorov_Smirnov_uv_test_1_vyber.EXE
http://user.mendelu.cz/drapela/Statisticke_metody/Navody_k_pouziti_statistickych_programu/Excel/Kolmogorov_Smirnov_uv_test_2_vybery.EXE

Statistica
http://user.mendelu.cz/drapela/Statisticke_metody/Navody_k_pouziti_statistickych_programu/Statistica/Stat-1-vyberovy_T-test.exe
http://user.mendelu.cz/drapela/Statisticke_metody/Navody_k_pouziti_statistickych_programu/Statistica/Stat-2V_t-test.exe
http://user.mendelu.cz/drapela/Statisticke_metody/Navody_k_pouziti_statistickych_programu/Statistica/Stat-parovy_T-test.exe
http://user.mendelu.cz/drapela/Statisticke_metody/Navody_k_pouziti_statistickych_programu/Statistica/Stat-Mann-Whitneyuv_Utest-neparam.exe

69
