

Přednáška 6

Pojmy z oblasti statistického testování

Typy testů

Normalita dat a její význam pro testování

Parametrické vs. neparametrické testy

- **Parametrické testy**

- Mají předpoklady o rozložení vstupujících dat (např. normální rozložení)
- Při stejném N a dodržení předpokladů mají vyšší sílu testu než testy neparametrické
- **Pokud nejsou dodrženy předpoklady parametrických testů, potom jejich síla testu prudce klesá a výsledek testu může být zcela chybný a nesmyslný**

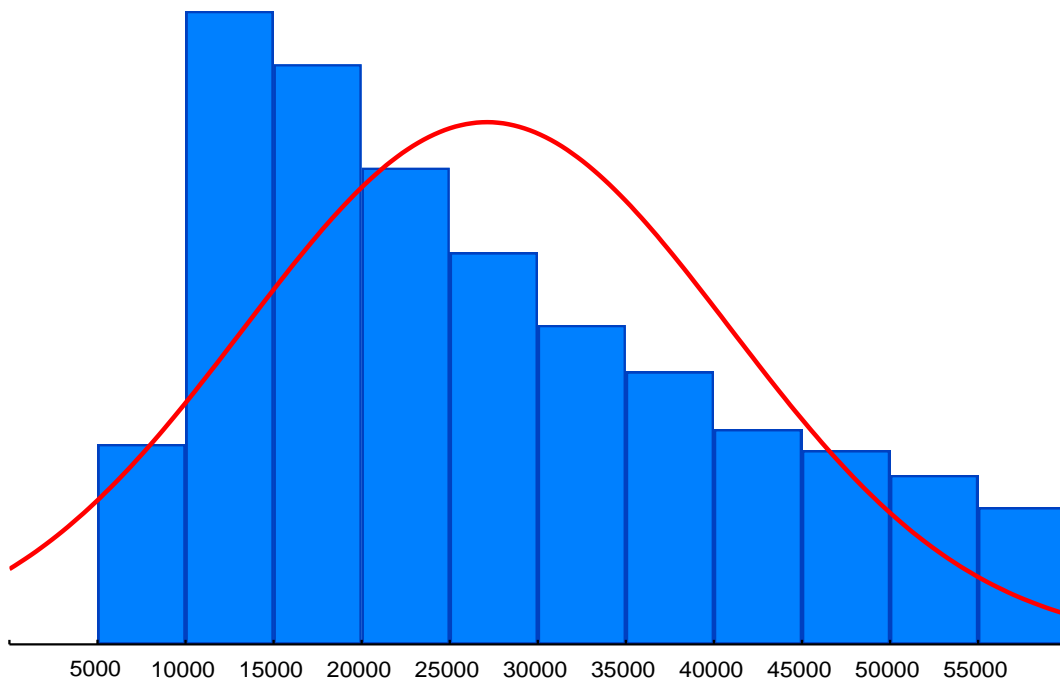
- **Neparametrické testy**

- Nemají předpoklady o rozložení vstupujících dat, lze je tedy použít i při asymetrickém rozložení, odlehlých hodnotách, či nedetekovatelném rozložení
- Snížená síla těchto testů je způsobena redukcí informační hodnoty původních dat, kdy neparametrické testy nevyužívají původní hodnoty, ale nejčastěji pouze jejich pořadí

Problémy parametrických a neparametrických testů

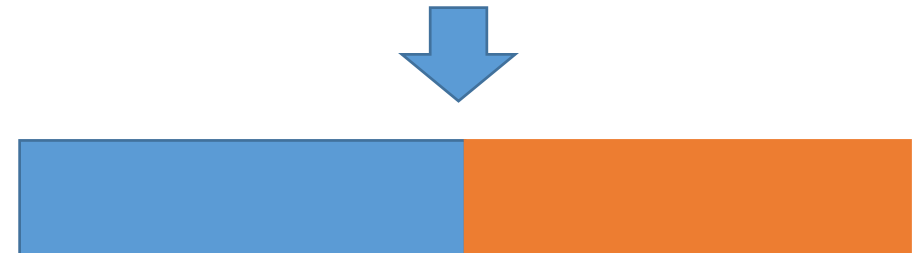
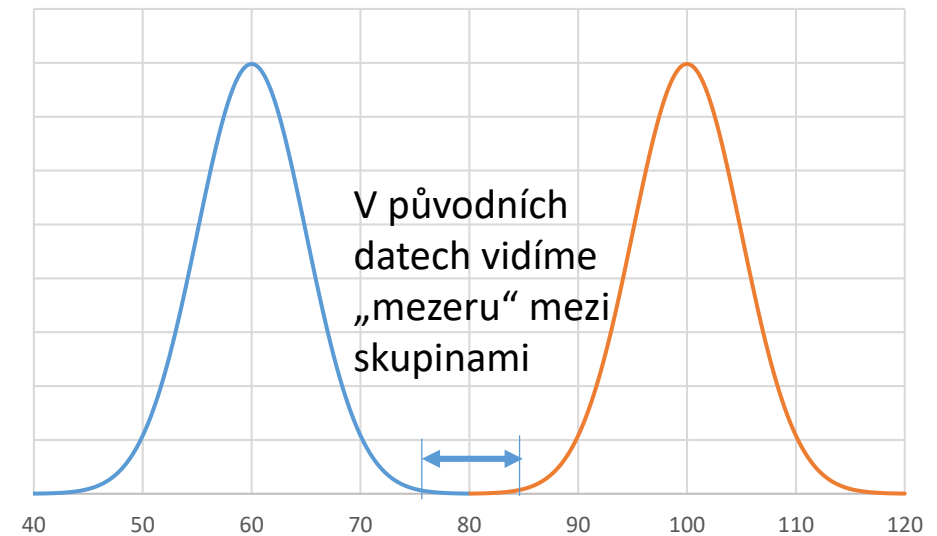
Parametrické testy

- Reálná data neodpovídají modelovému rozdělení



Neparametrické testy

- Díky převodu dat na pořadí ztrácíme část informace



Po převodu na pořadí o tuto informaci přicházíme.

Jednovýběrové a dvouvýběrové testy

- **Jednovýběrové testy (one-sample)**

- Srovnávají jeden vzorek (one sample, jednovýběrové testy) s referenční hodnotou (popřípadě se statistickým parametrem cílové populace)
- V testu je tedy srovnáváno rozložení hodnot (vzorek) s jediným číslem (referenční hodnota, hodnota cílové populace)
- Otázka položená v testu může být vztažena k průměru, rozptylu, podílu hodnot i dalším statistickým parametrům popisujícím vzorek

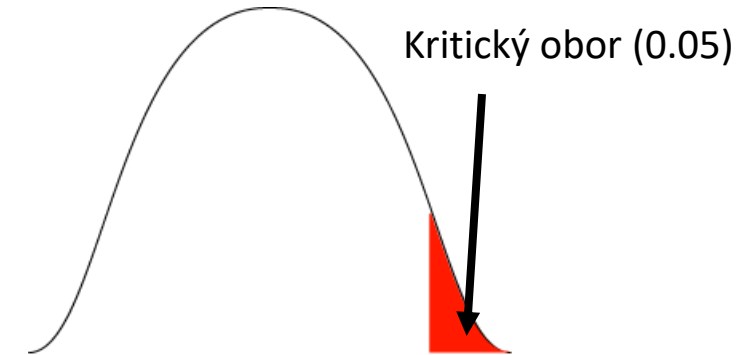
- **Dvouvýběrové testy (two-sample)**

- Srovnávají navzájem dva vzorky (two sample, dvouvýběrové testy)
- V testu jsou srovnávány dvě rozložení hodnot
- Otázka položená v testu může být opět vztažena k průměru, rozptylu, podílu hodnot i dalším statistickým parametrům popisujícím vzorek
- Kromě testů pro dvě skupiny hodnot existují samozřejmě i testy pro více skupin dat

Jednostranné a oboustranné hypotézy

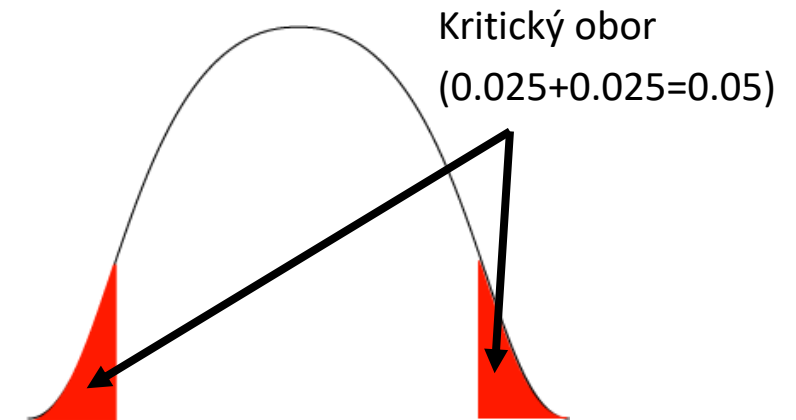
- **Jednostranné testy (one-tailed)**

- Hypotéza testu je postavena asymetricky, tedy ptáme se na větší než/ menší než
- Test může mít pouze dvojí výstup – jedna z hodnot je větší (menší) než druhá a všechny ostatní případy
- Pouze v případě jasné hypotézy – jinak může být napadnuto za účelovost (při vhodném výběru směru testování snaží potvrzení významnosti)



- **Oboustranné testy (two-tailed)**

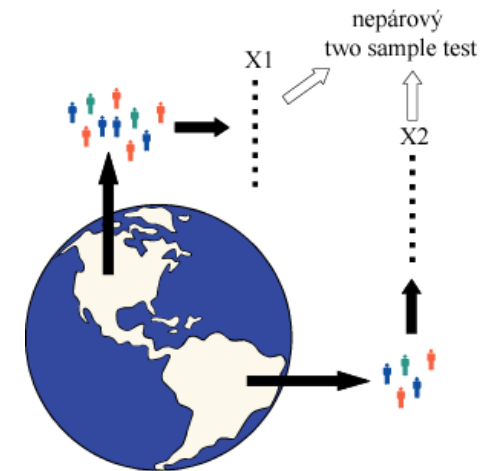
- Hypotéza testu se ptá na otázku rovná se/nerovná se
- Test může mít trojí výstup – menší - rovná se – větší než
- Situace nerovná se je tedy souhrnem dvou možných výstupů testu (menší+větší)
- Významově neutrální



Nepárový vs. párový design

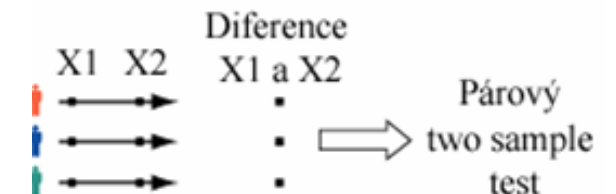
• Nepárový design

- Skupiny srovnávaných dat jsou na sobě zcela nezávislé (též nezávislý, independent design), např. lidé z různých zemí, nezávislé skupiny pacientů s odlišnou léčbou atd.
- Při výpočtu je nezbytné brát v úvahu charakteristiky obou skupin dat



• Párový design

- Mezi objekty v srovnávaných skupinách existuje vazba, daná např. člověkem před a po operaci, reakce stejného kmene krys atd.
- Vazba může být buď přímo dána nebo pouze předpokládána (v tom případě je nutné ji ověřit)
- Test je v podstatě prováděn na diferencích skupin, nikoliv na jejich původních datech



Důležité poznámky k testování hypotéz

- **Nezamítnutí nulové hypotézy neznamena automaticky její přijetí!** Může se jednat o situaci, kdy pro zamítnutí nulové hypotézy nemáme dostatečné množství informace.
- **Dosažená hladina významnosti testu** (ať už 5 %, 1 % nebo 10 %) **nesmí být slepě brána jako hranice pro existenci / neexistenci testovaného efektu.**
- **Malá p-hodnota nemusí znamenat velký efekt.** Hodnota testové statistiky a p-hodnota mohou být ovlivněny velkou velikostí vzorku a malou variabilitou pozorovaných dat.
- **Na výsledky testování musí být nahlíženo kriticky** – jedná se o závěr založený „pouze“ na jednom výběrovém souboru.
- **Statistická významnost** indikuje, že pozorovaný rozdíl není náhodný, ale nemusí znamenat, že je významný i ve skutečnosti. Důležitá je i **praktická (klinická) významnost.**

Statistické testy a normalita

- Normalita dat je jedním z předpokladů tzv. parametrických testů (testů založených na předpokladu nějakého rozložení) – např. t-testy
- Obecně lze říci, že každá statistická metoda, v jejímž algoritmu je obsažen výpočet průměru nebo směrodatné odchylky má předpoklad normálního rozložení
- Pokud data nejsou normální, neodpovídají ani modelovému rozložení, které je použito pro výpočet (t-rozložení) a test tak může lhát
- Řešením je tedy:
 - Transformace dat za účelem dosažení normality jejich rozložení
 - Neparametrické testy – tyto testy nemají předpoklady o rozložení dat (nebo jen minimální)

Typ srovnání	Parametrický test	Neparametrický test
2 skupiny dat nepárově:	Nepárový t-test	Mannův-Whitneyho test
2 skupiny dat párově:	Párový t-test	Wilcoxonův test, znaménkový test
Více skupin nepárově:	ANOVA (analýza rozptylu)	Kruskalův- Wallisův test
Korelace:	Pearsonův koeficient	Spearmanův koeficient

Testy normality

- Testy normality pracují s nulovou hypotézou, že není rozdíl mezi zpracovávaným rozložením a normálním rozložením. Vždy je ovšem dobré prohlédnout si i histogram, protože některé odchylky od normality, např. bimodalitu některé testy neodhalí.

Chí-kvadrát test dobré shody

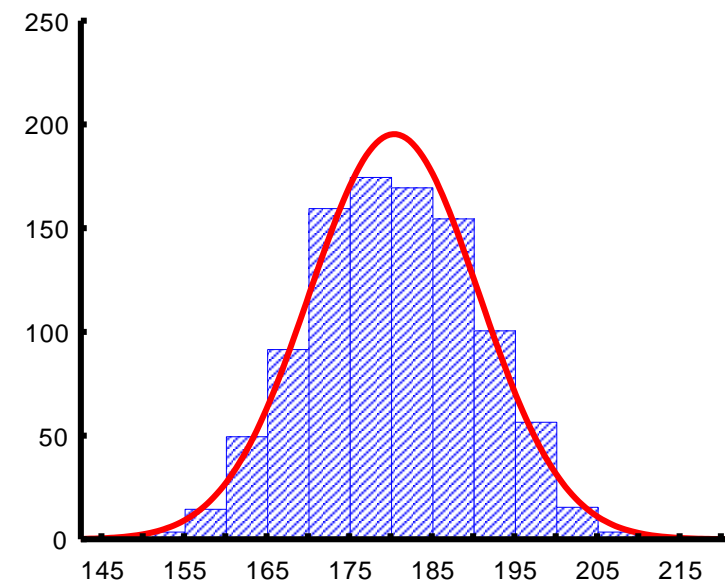
- V testu dobré shody jsou data rozdělena do kategorií (obdobně jako při tvorbě histogramu), tyto intervaly jsou normalizovány (převedeny na normální rozložení) a podle obecných vzorců normálního rozložení jsou k nim dopočítány očekávané hodnoty v intervalech, pokud by rozložení bylo normální. Pozorované normalizované četnosti jsou poté srovnány s očekávanými četnostmi pomocí χ^2 testu dobré shody. Test dává dobré výsledky, ale je náročný na n , tedy množství dat, aby bylo možné vytvořit dostatečný počet tříd hodnot.

Kolmogorovův - Smirnovův test

- Tento test je často používán, dokáže dobře najít odlehlé hodnoty, ale počítá spíše se symetrií hodnot než přímo s normalitou. Jde o neparametrický test pro srovnání rozdílu dvou rozložením. Je založen na zjištění rozdílu mezi reálným kumulativním rozložením (vzorek) a teoretickým kumulativním rozložením. Měl by být počítán pouze v případě, že známe průměr a směrodatnou odchylku hypotetického rozložení, pokud tyto hodnoty neznáme, měla by být použita jeho modifikace – Lilieforsův test.

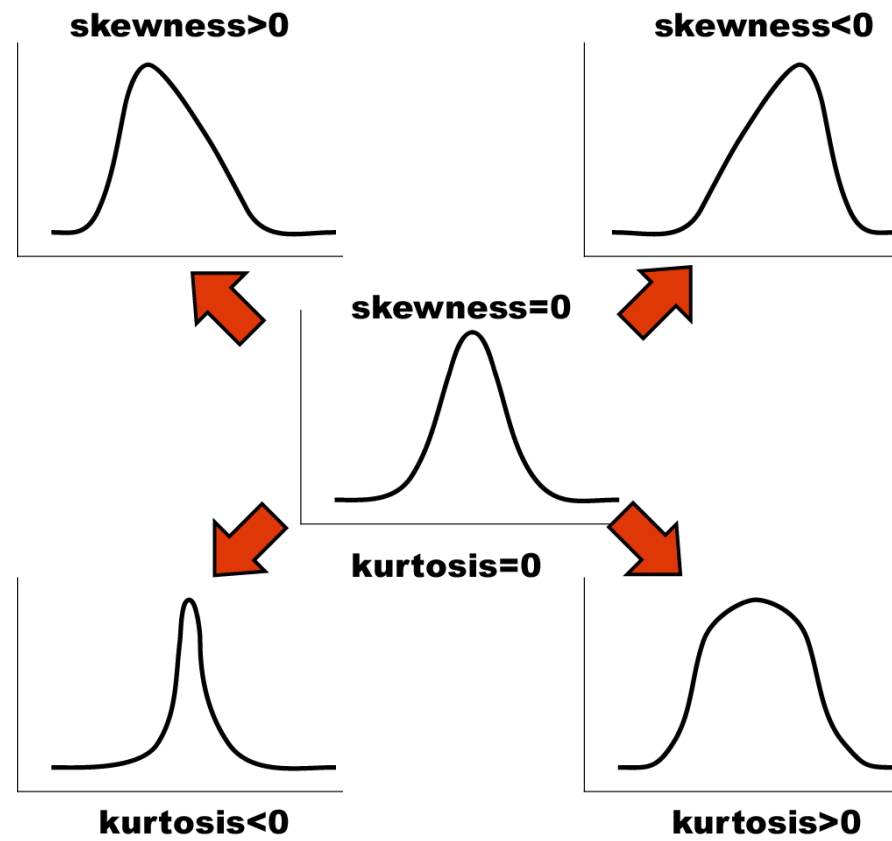
Shapirův-Wilkův test

- Jde o neparametrický test použitelný i při velmi malých n (10) s dobrou silou testu, zvláště ve srovnání s alternativními typy testů, je zaměřen na testování symetrie.

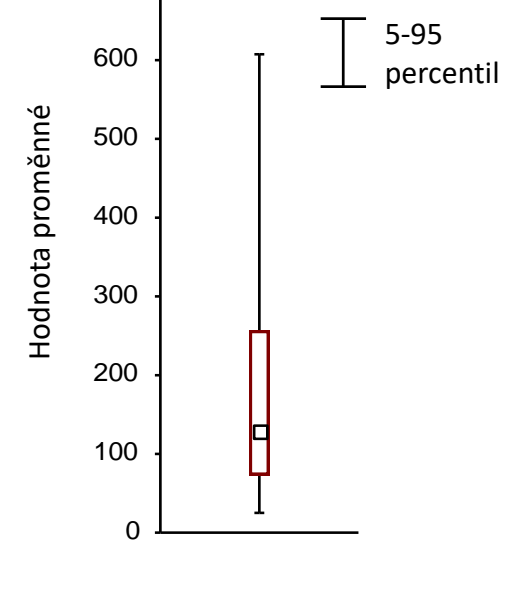
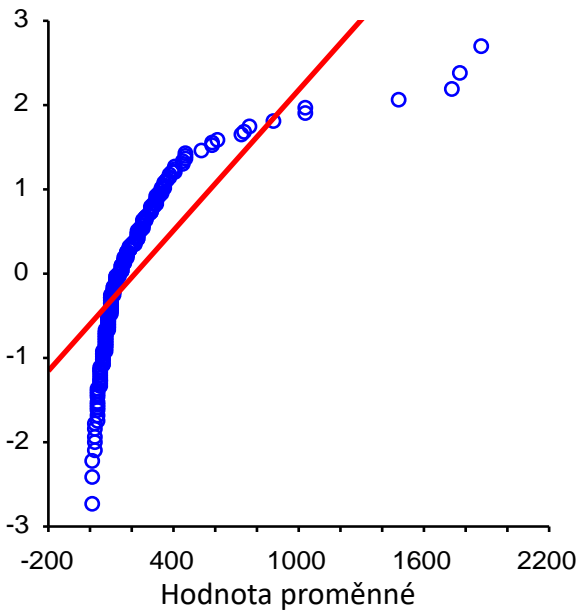
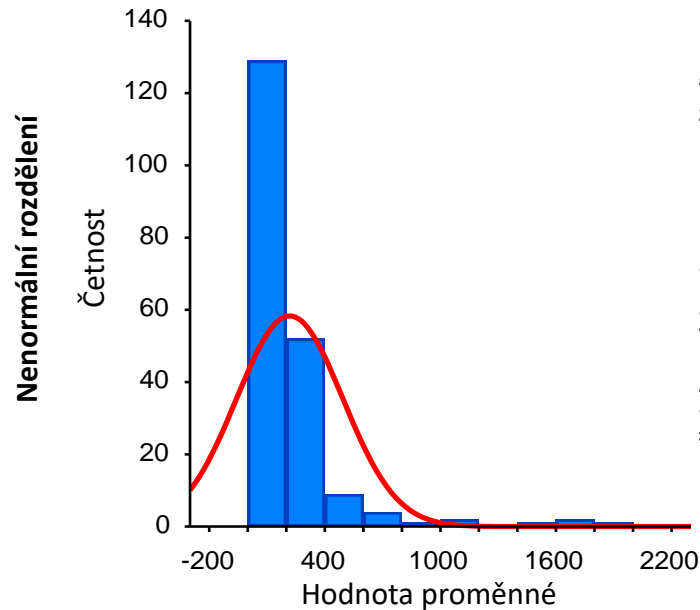
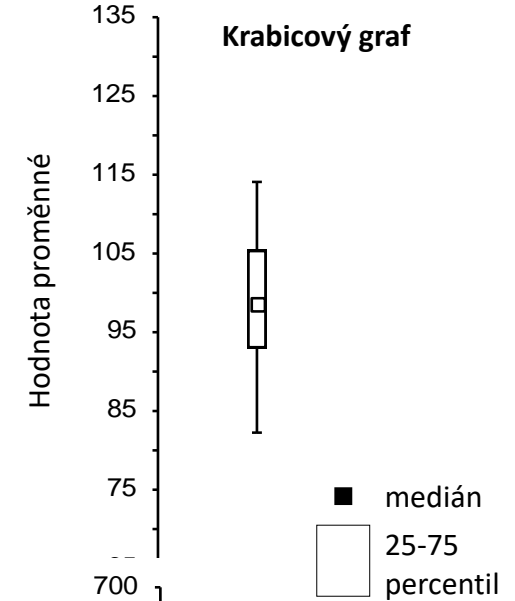
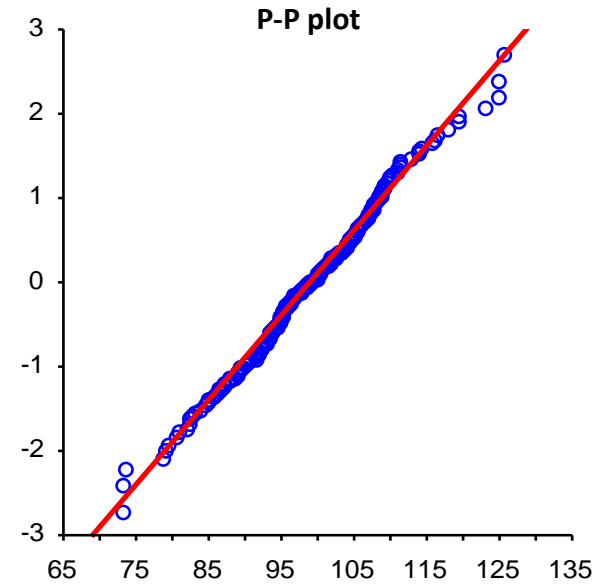
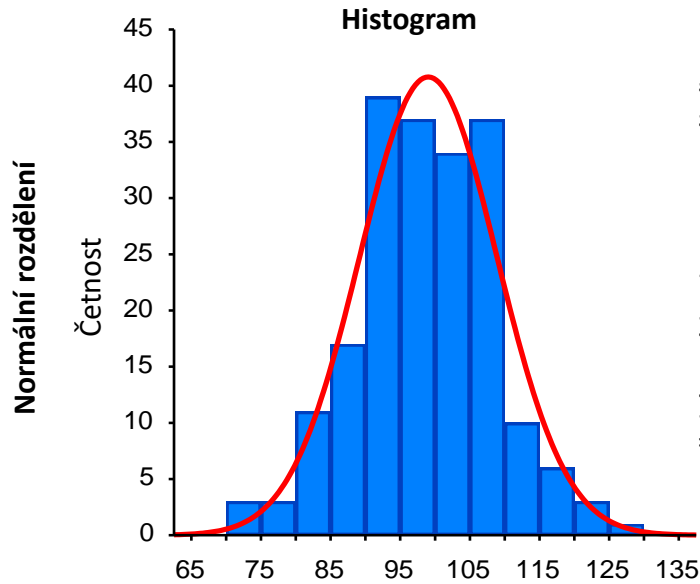


Šikmost a špičatost jako testy normality

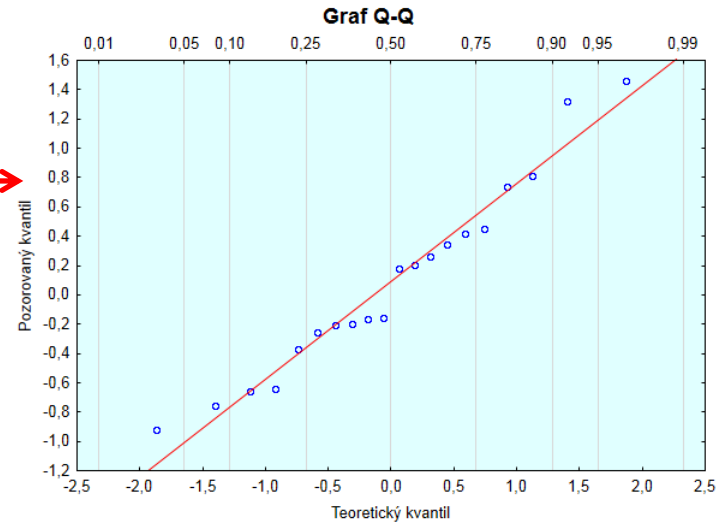
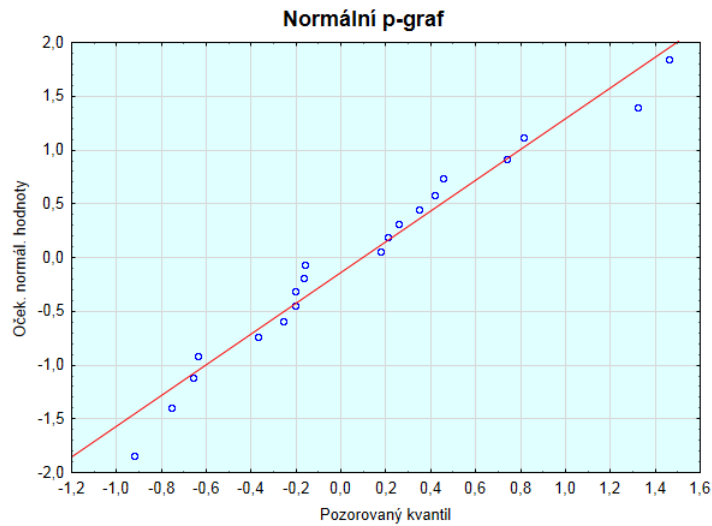
- Parametry normálního rozdělení, skewness a kurtosis mohou být využity pro testování normality, ale pouze pro velké vzorky (šikmost – 100, špičatost – 500).



Vizuální hodnocení normality I

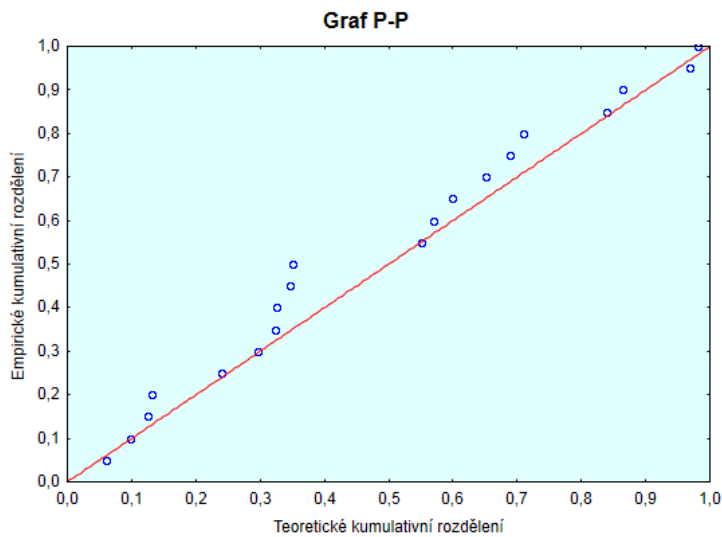


Vizuální hodnocení normality II



???

- Pouze výměna os
- Znárodněn pozorovaný a teoretický kvantil



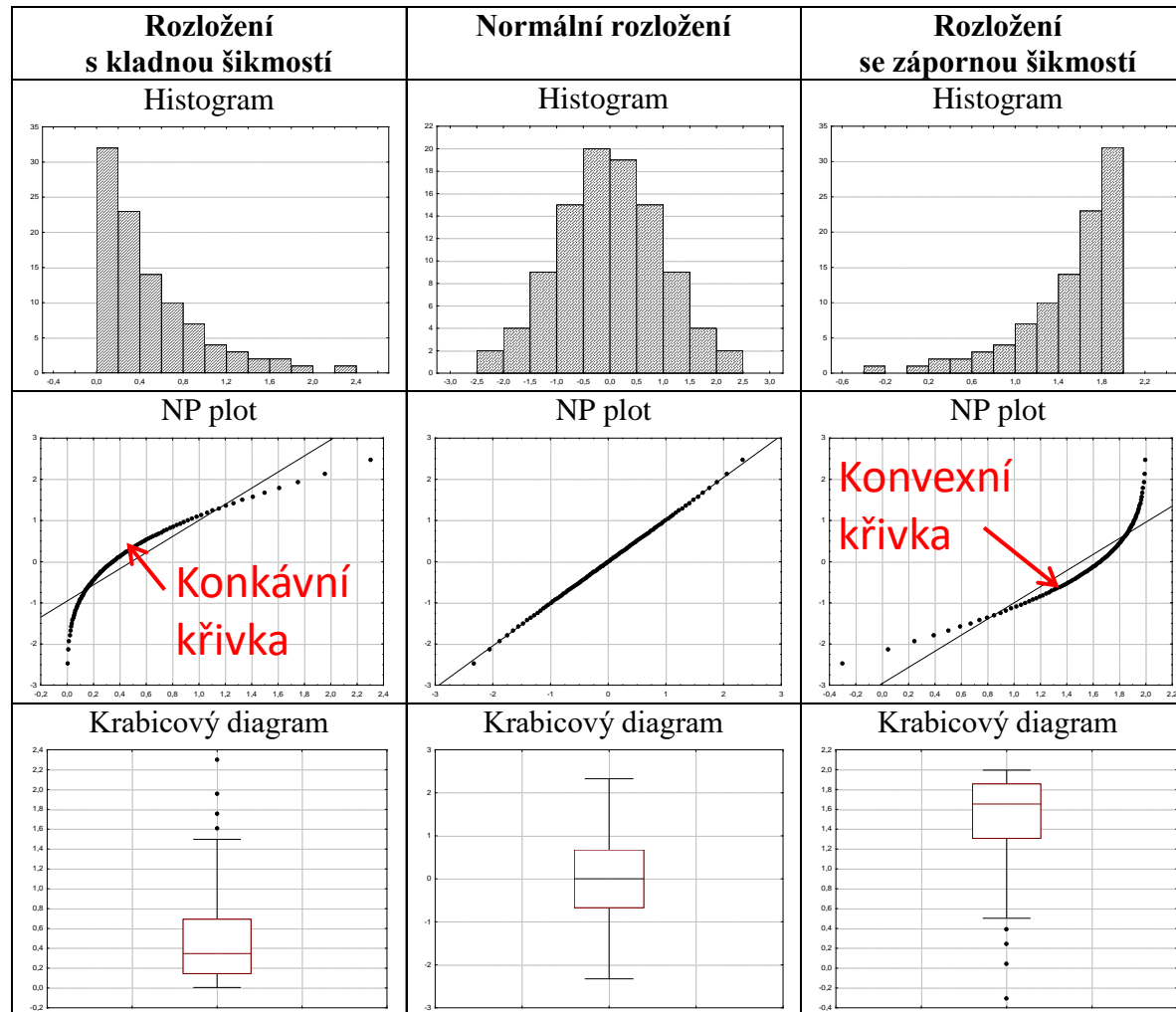
- Vykresleno kumulativní rozdělení

PAMATUJ:

**Pocházejí-li data z normálního rozložení,
pak body budou ležet okolo přímky**



Vizuální hodnocení normality III



Výukové materiály: Výpočetní statistika,
RNDr. Marie Budíková, Dr., 2011

Parametrické jednovýběrové statistické testy

Jednovýběrový t-test

Jednovýběrový test rozptylu

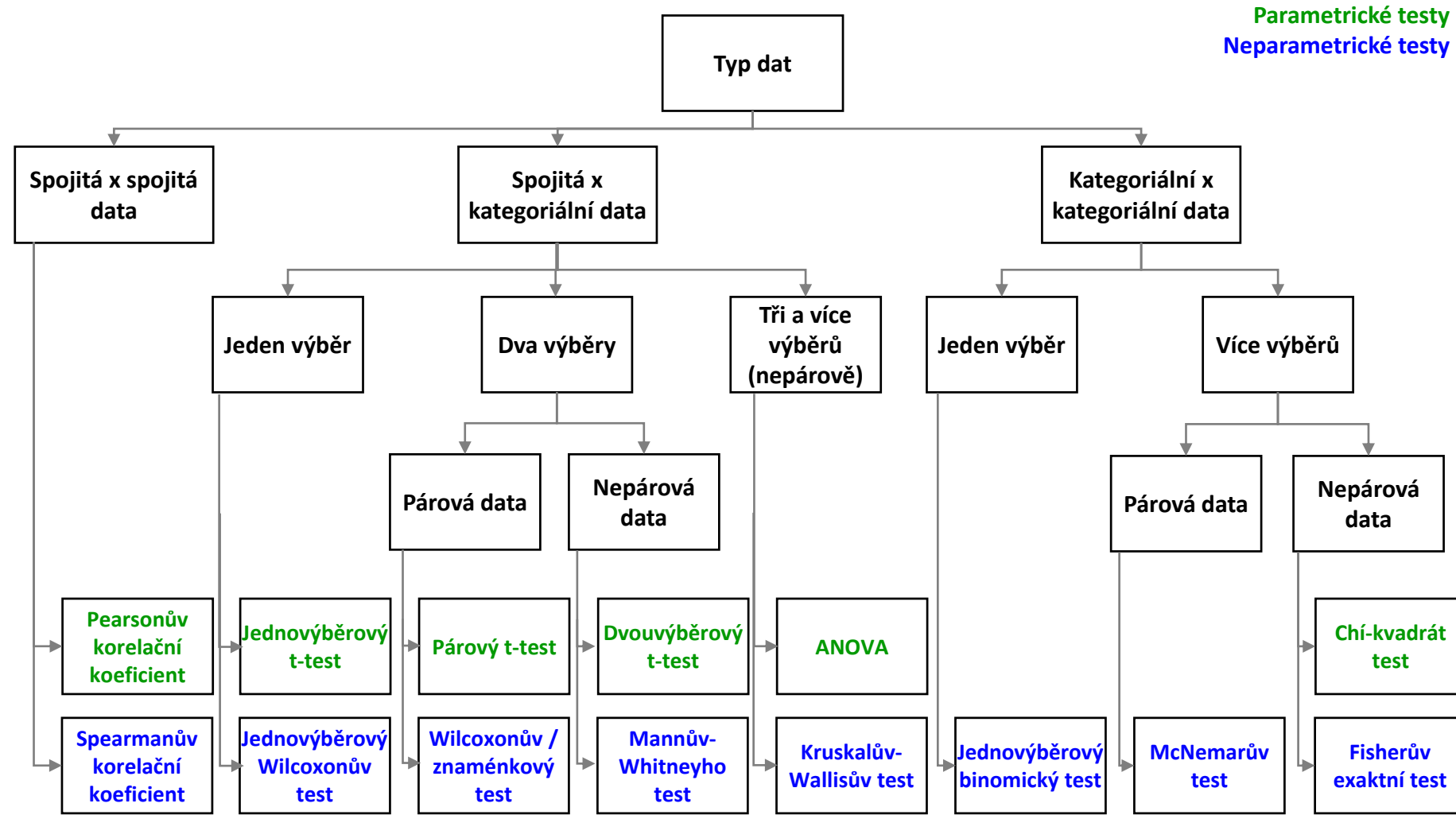
Anotace

- Jednovýběrové statistické testy srovnávají některou popisnou statistiku vzorku (průměr, směrodatnou odchylku) s jediným číslem, jehož význam je ze statistické hlediska hodnota cílové populace
- Z hlediska statistické teorie jde o ověření, zda daný vzorek pochází z testované cílové populace.

Shrnutí statistických testů

Typ srovnání	Nulová hypotéza	Parametrický test	Neparametrický test
1 výběr dat vs. referenční hodnota	Střední hodnota je rovna zvolené referenční hodnotě.	jednovýběrový t-test / z-test	Jednovýběrový Wilcoxonův test
2 nezávislé skupiny dat (test shody středních hodnot)	Střední hodnoty se mezi skupinami neliší.	nepárový t-test	Mannův-Whitneyho test
2 nezávislé skupin dat (test shody rozptylů = homoskedasticity)	Rozptyl obou skupin je shodný.	F-test	Levenův test
2 párově závislé výběry dat	Rozdíl (diference) párových hodnot je nulový.	párový t-test	Wilcoxonův test; znaménkový test
Shoda rozdělení výběru s teoretickým rozdělením	Rozdělení dat odpovídá teoretickému (vybranému) rozdělení.	test dobré shody (χ^2 test)	Shapirův-Wilkův test; Kolmogorovův-Smirnovův test; Lilieforsův test
3 a více skupin nepárově (test shody středních hodnot)	Střední hodnoty se mezi skupinami neliší.	ANOVA	Kruskalův-Wallisův test
Korelace	Neexistuje vztah mezi hodnotami dvou výběrů.	Pearsonův korelační koeficient	Spearmanův korelační koeficient

Základní rozhodování o výběru statistických testů



Parametrické jednovýběrové testy

- Předpoklad: **normalita dat**
- **Jednovýběrový z-test** (porovnání základního a výběrového souboru, známe střední hodnotu a rozptyl základního souboru)
- **Studentův jednovýběrový t-test** (testování rozdílů dvou středních hodnot) - (porovnání základního a výběrového souboru, známe střední hodnotu ale neznáme rozptyl základního souboru; nahrazujeme jej výběrovým rozptylem našich dat)
- **Chi-kvadrát test** (testování rozdílu cílová vs. výběrová populace)

Jednovýběrový z a t test

- V případě jednovýběrových testů jde o srovnání výběru dat (tedy one sample) s cílovou populací. Pro parametrické testy musí mít datový soubor normální rozložení.
- Rozdíl mezi jednovýběrovým z-testem a t-testem spočívá ve znalosti rozptylu základního souboru (z-test) nebo jeho nahrazení výběrovým rozptylem našich dat (t-test)

z-test:

$$z = \frac{\bar{x} - \mu}{\rho} \sqrt{N}$$

H_0	H_A	Testová statistika	Kritická hodnota
$\bar{x} \leq \mu$	$\bar{x} > \mu$	z / t	$z > z_{1-\alpha} / t > t_{1-\alpha}^{N-1}$

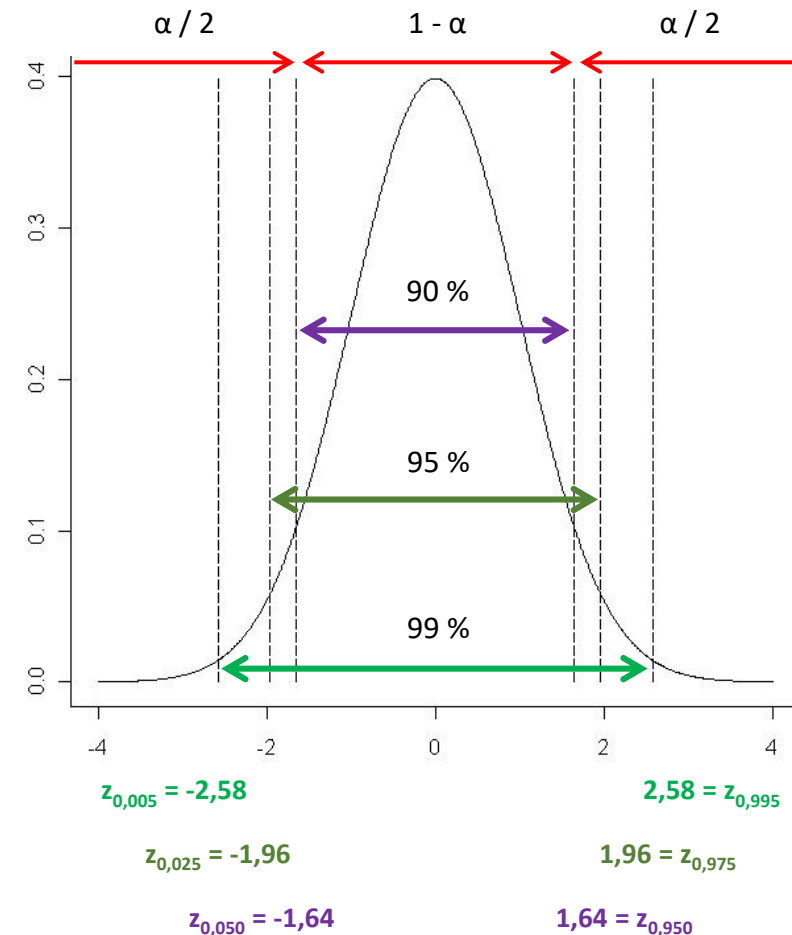
t-test:

$$t = \frac{\bar{x} - \mu}{s} \sqrt{N}$$

$\bar{x} \geq \mu$	$\bar{x} < \mu$	z / t	$z < z_{\alpha} / t < t_{\alpha}^{N-1}$
$\bar{x} = \mu$	$\bar{x} \neq \mu$	z / t	$ z > z_{1-\alpha/2} / t > t_{1-\alpha/2}^{N-1}$

Příklad: z-test pro jeden výběr 1

- Při populačním epidemiologickém průzkumu se zjistilo, že průměrný objem prostaty u mužů je 32,73 ml (SD = 18,12 ml).
- Na hladině významnosti testu $\alpha = 0,05$ chceme ověřit, jestli se muži nad 70 let liší od celé populace.
- Máme náhodný výběr o velikosti $n = 100$ a výběrový průměr 36,60 ml.
- Chceme ověřit platnost:
 - $H_0: \mu = 32,73$
 - $H_A: \mu \neq 32,73$



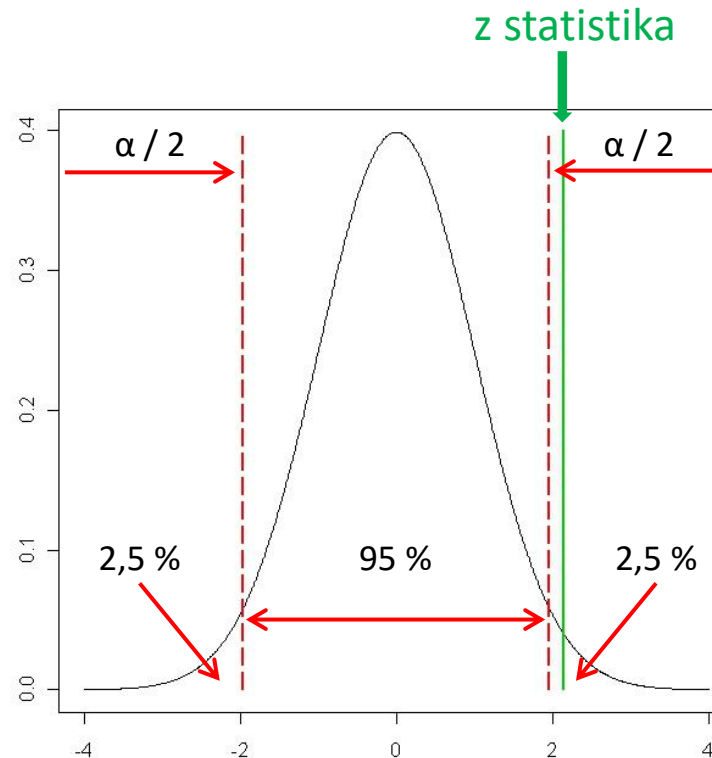
Příklad: z-test pro jeden výběr 2

- Hodnota testové statistiky:
$$z = \frac{\bar{x} - \mu}{\rho} \sqrt{N} = \frac{36,60 - 32,73}{18,12} \sqrt{100} = 2,14$$

- Můžeme zamítnout nulovou hypotézu na hladině významnosti testu $\alpha = 0,05$ nebo ne?

$$z = 2,14 > 1,96 \quad (z_{1-\alpha/2} = z_{0,975})$$

- Nulovou hypotézu o rovnosti objemu prostaty u mužů nad 70 let populační hodnotě 32,73 ml zamítáme na hladině významnosti $\alpha = 0,05$, protože výsledná hodnota z statistiky je větší než kritická hodnota (příslušný kvantil) rozdělení $N(0,1)$.



Příklad: t-test pro jeden výběr



- Určitá linka autobusové městské dopravy má v době dopravní špičky průměrnou rychlost 8 km/hod. Uvažovalo se o tom, zda změna trasy by vedla ke změně průměrné rychlosti. Nová trasa byla proto projeta v deseti náhodně vybraných dnech a byly zjištěny tyto průměrné rychlosti: 8,4; 7,9; 9,0; 7,8; 8,0; 7,8; 8,5; 8,2; 8,2; 9,3. Rozhodněte, zda změna trasy vede ke změně průměrné rychlosti. Předpokládáme normální rozdělení a $\alpha=0,05$.

- Postup:

1. Na hladině významnosti 0,05 testujeme hypotézu $H_0: \mu = 8$, proti $H_A: \mu \neq 8$

2. Vypočteme aritmetický průměr a rozptyl výběrového souboru.

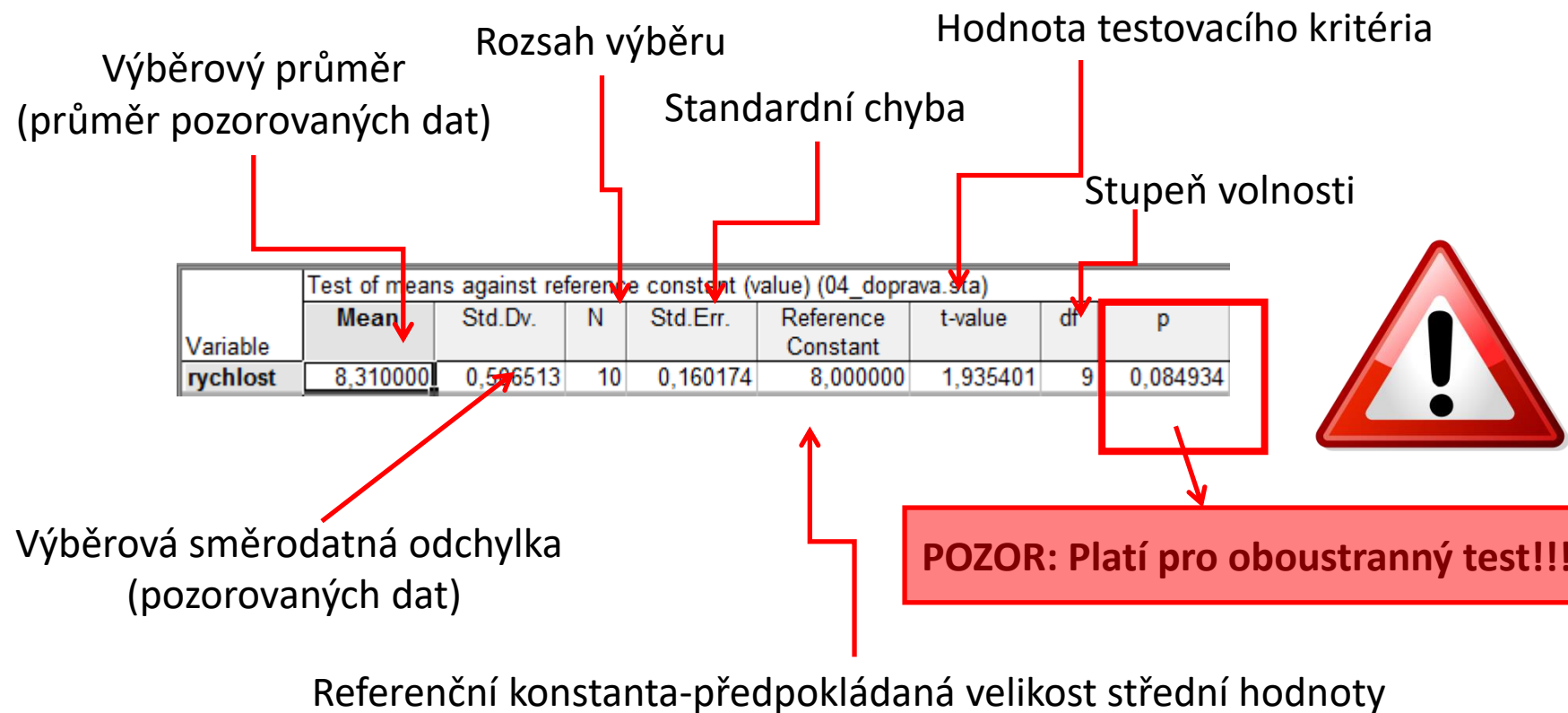
3. Vypočteme testovou statistiku t:
$$t = \frac{\bar{x} - \mu}{s} \sqrt{N} = \frac{8,310 - 8}{0,507} \sqrt{10} = 1,934$$

4. Vypočtené t porovnáme s kritickou hodnotou:

$$t_{1-\alpha/2}^{N-1} = t_{0,975}^9 = 2,262$$

5. Je-li $|t| \leq t_{1-\alpha/2}^{N-1}$ -> statisticky nevýznamný rozdíl testovaných parametrů při zvolené α ; **nulovou hypotézu nezamítáme**, na hladině významnosti $\alpha=0,05$ se nepodařilo prokázat, že by změna trasy měla za následek změnu průměrné rychlosti.

Typické výstupy SW (Statistica, v jiných obdobné)



Příklad k řešení: t-test 1

Data - koncentrace antibiotika v cílovém orgánu

- Při 1000 měřeních antibiotika byla zjištěna v cílovém orgánu průměrná koncentrace 202,5 jednotek a směrodatná odchylka 44 jednotek.
- Požadovaná koncentrace antibiotika je 200 jednotek.

Výzkumné otázky

1. Je daný rozdíl 2,5 významný vzhledem k variabilitě znaku na hladině významnosti 5%?
2. Jaká je skutečná hladina významnosti?

Příklad k řešení: t-test 1

Data - koncentrace antibiotika v cílovém orgánu

- Při 1000 měřeních antibiotika byla zjištěna v cílovém orgánu průměrná koncentrace 202,5 jednotek a směrodatná odchylka 44 jednotek.
- Požadovaná koncentrace antibiotika je 200 jednotek.

Výzkumné otázky

1. Je daný rozdíl 2,5 významný vzhledem k variabilitě znaku na hladině významnosti 5%?
2. Jaká je skutečná hladina významnosti?

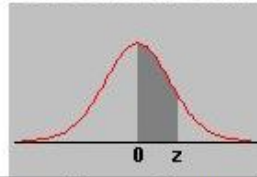
$$t = \frac{\bar{x} - \mu}{s} \sqrt{N} = \frac{202,5 - 200}{44} \sqrt{1000} = 1,797$$

Příklad k řešení: t-test 1

$$t = \frac{\bar{x} - \mu}{s} \sqrt{N} = \frac{202,5 - 200}{44} \sqrt{1000} = 1,797 \sim 1,8$$

$$t_{1-\alpha/2}^{N-1} = t_{0,975}^{999} = 1,960$$

Area between 0 and z



	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706

Výzkumné otázky

1. Je daný rozdíl 2,5 významný vzhledem k variabilitě znaku na hladině významnosti 5%?
 - **Nulovou hypotézu nezamítáme**
2. Jaká je skutečná hladina významnosti?
 - **$p = 2 * (1 - 0,4641) = 0,072$**

Příklad k řešení: t-test 2

Data - aktivita enzymu v buňkách

- Při zjišťování aktivity enzymu v buňkách na vzorku 25 měření byl zjištěn průměr 3,5 jednotek a směrodatná odchylka 1.

Výzkumné otázky

1. otázka zní, zda se naměřené hodnoty našeho vzorku liší od výsledků dřívější rozsáhlé studie zaměřené na celou cílovou populaci, kde byla zjištěna průměrná aktivita 2,5 jednotky?
2. otázka – jakou minimální odchylku X od jiné hodnoty bychom zachytili při daných hodnotách?
3. za předpokladu, že z praktického hlediska je významná odchylka již 0,2 jednotky, jaký minimální počet měření musíme provést, abychom ji byli schopni prokázat ?

Příklad k řešení: t-test 2

Data - aktivita enzymu v buňkách

- Při zjišťování aktivity enzymu v buňkách na vzorku 25 měření byl zjištěn průměr 3,5 jednotek a směrodatná odchylka 1.

Výzkumné otázky

1. otázka zní, zda se naměřené hodnoty našeho vzorku liší od výsledků dřívější rozsáhlé studie zaměřené na celou cílovou populaci, kde byla zjištěna průměrná aktivita 2,5 jednotky?

$$t = \frac{\bar{x} - \mu}{s} \sqrt{n} = \frac{3,5 - 2,5}{1} \sqrt{25} = 5$$

$$t_{0,975}^{24} = 2,064 \Rightarrow t > t_{1-\alpha/2}^{24} \Rightarrow H_0 \text{ zamítnuta při } \alpha \leq 0,05$$

Příklad k řešení: t-test 2

Data - aktivita enzymu v buňkách

- Při zjišťování aktivity enzymu v buňkách na vzorku 25 měření byl zjištěn průměr 3,5 jednotek a směrodatná odchylka 1.

Výzkumné otázky

2. otázka – jakou minimální odchylku X od jiné hodnoty bychom zachytili při daných hodnotách?

$$t = \frac{\bar{x} - \mu}{s} \sqrt{n} = \frac{d}{s} \sqrt{n} \rightarrow d = \frac{t_{1-\alpha/2}^v}{\sqrt{n}} s \rightarrow d = \frac{2,064}{5} 1$$

Příklad k řešení: t-test 2

Data - aktivita enzymu v buňkách

- Při zjišťování aktivity enzymu v buňkách na vzorku 25 měření byl zjištěn průměr 3,5 jednotek a směrodatná odchylka 1.

Výzkumné otázky

3. za předpokladu, že z praktického hlediska je významná odchylka již 0,2 jednotky, jaký minimální počet měření musíme provést, abychom ji byli schopni prokázat ?

$$t = \frac{\bar{x} - \mu}{s} \sqrt{n} = \frac{d}{s} \sqrt{n} \rightarrow n = \left(\frac{t_{1-\alpha/2}^v}{d} s \right)^2$$

Jednovýběrový test pro rozptyl

- V případě one sample testů jde o srovnání výběru dat (tedy one sample) s cílovou populací. Pro parametrické testy musí mít datový soubor normální rozložení.

Chi-kvadrát test:

$$\chi^2 = \frac{(N - 1)s^2}{\sigma^2}$$

H_0	H_A	Testová statistika	Kritická hodnota
$s^2 \leq \sigma^2$	$s^2 > \sigma^2$	χ^2	$\chi^2 > \chi_{1-\alpha}^{2(N-1)}$
$s^2 \geq \sigma^2$	$s^2 < \sigma^2$	χ^2	$\chi^2 < \chi_{\alpha}^{2(N-1)}$
$s^2 = \sigma^2$	$s^2 \neq \sigma^2$	χ^2	$\chi^2 > \chi_{1-\alpha/2}^{2(N-1)}$ nebo $\chi^2 < \chi_{\alpha/2}^{2(N-1)}$

Neparametrické jednovýběrové statistické testy

Jednovýběrový t-test

Jednovýběrový test rozptylu

Parametrické vs. neparametrické testy

Parametrické testy

- Mají předpoklady o rozložení vstupujících dat (např. normální rozložení)
- Při stejném N a dodržení předpokladů mají vyšší sílu testu než testy neparametrické
- **Pokud nejsou dodrženy předpoklady parametrických testů, potom jejich síla testu prudce klesá a výsledek testu může být zcela chybný a nesmyslný**



Neparametrické testy

- Vyžadují méně předpokladů o rozložení vstupujících dat, lze je tedy použít i při asymetrickém rozložení, odlehlých hodnotách, či nedetekovatelném rozložení
- Snížená síla těchto testů je způsobena redukcí informační hodnoty původních dat, kdy neparametrické testy nevyužívají původní hodnoty, ale nejčastěji pouze jejich pořadí
- Souvisí s malou velikostí souboru (nejsme schopni normalitu dat ověřit)

Proč nemusí parametrický a neparametrický test vyjít stejně?

Jednovýběrový Wilcoxonův test

- Předpokladem je symetrické rozdělení dat kolem mediánu.
- Testuje, zda je **medián** jednoho výběru roven hodnotě c (v případě párového designu je $x_{0.5}$ reprezentováno mediánem rozdílu hodnot)

$$H_0: x_{0.5} = c \text{ proti } H_1: x_{0.5} \neq c.$$

Postup:

1. Spočítáme rozdíly hodnot výběru s testovanou hodnotou mediánu.
2. Absolutní hodnoty rozdílů uspořádáme vzestupně a přiřadíme jim pořadí.
3. Spočítáme statistiky S_w^+ a S_w^- , které odpovídají **součtu pořadí kladných (S_w^+) a záporných rozdílů (S_w^-)**. Jako finální hodnotu testové statistiky bereme minimum z S_w^+ a S_w^- . Nulovou hypotézu zamítáme, pokud hodnota testové statistiky menší nebo rovna tabelované kritické hodnotě (při dané hladině významnosti a počtu nenulových rozdílů).

nebo

3. Pro $N > 30$ lze využít asymptotické normality statistiky S_w^+

$$E(S_w^+) = \frac{n(n+1)}{4}$$

$$D(S_w^+) = \frac{n(n+1)(2n+1)}{24}$$

$$Z = \frac{S_w^+ - E(S_w^+)}{\sqrt{D(S_w^+)}} \approx N(0,1)$$

- Pokud $|Z| \geq u_{1-\alpha/2}$ zamítáme nulovou hypotézu, že medián výběru je roven hodnotě c .

Jednovýběrový znaménkový test

- Lze použít v situaci, kdy není splněn předpoklad symetrie rozdělení kolem mediánu.
- Testuje, zda je medián jednoho výběru roven hodnotě c (v případě párového designu je $x_{0,5}$ reprezentováno mediánem rozdílu hodnot)

$$H_0: x_{0,5}=c \text{ proti } H_1: x_{0,5} \neq c.$$

Postup:

1. Spočítáme rozdíly hodnot výběru s testovanou hodnotou mediánu.
 2. Spočítáme statistiku S_z^+ , která odpovídá počtu kladných rozdílů → **test nevyužívá hodnot pořadí původních dat ale pouze informaci, zda se hodnota realizuje nad nebo pod mediánem** → dochází ke snížení síly testu
 3. Nulovou hypotézu zamítáme, pokud statistika S_z^+ realizuje v kritickém oboru hodnot $W=(0,k_1) \cup (k_2,n)$, kde n odpovídá počtu nenulový rozdílů a hodnoty k_1 a k_2 lze dohledat v matematických tabulkách.
- **nebo**
3. Pro $N > 20$ lze využít asymptotické normality statistiky S_z^+ .

$$E(S_z^+) = \frac{n}{2} \quad D(S_z^+) = \frac{n}{4} \quad Z = \frac{S_z^+ - E(S_z^+)}{\sqrt{D(S_z^+)}} \approx N(0,1)$$

Pokud $|Z| \geq u_{1-\alpha/2}$ zamítáme nulovou hypotézu, že medián výběru je roven hodnotě c .

Příklad: jednovýběrový test


- U 15 náhodně vybraných pacientů byla vyhodnocena doba, kterou museli strávit v čekárně, než byli sestrou pozváni do ordinace. Na 5% hladině významnosti testujte nulovou hypotézu, že medián čekací doby je roven půl hodině.



Příklad: jednovýběrový test – Wilcoxonův test

- U 15 náhodně vybraných pacientů byla vyhodnocena doba, kterou museli strávit v čekárně, než byli sestrou pozváni do ordinace. Na 5% hladině významnosti testujte nulovou hypotézu, že medián čekací doby je roven půl hodině.

Pacient č.	čekací doba (min)	medián	rozdíl	rozdíl	pořadí
1	1	30	-29	29	15
2	45	30	15	15	10
3	25	30	-5	5	3.5
4	15	30	-15	15	10
5	34	30	4	4	2
6	19	30	-11	11	8
7	31	30	1	1	1
8	25	30	-5	5	3.5
9	8	30	-22	22	14
10	12	30	-18	18	12
11	20	30	-10	10	6
12	15	30	-15	15	10
13	40	30	10	10	6
14	20	30	-10	10	6
15	10	30	-20	20	13


$$S_w^+ = 19$$

$$S_w^- = 101$$

$$\min(S_w^+, S_w^-) = 19$$

Kritická hodnota $w_{15}(0,05) = 25$

Hodnota testové statiky je menší než kritická hodnota → **zamítáme H_0**

Příklad: jednovýběrový test – Znaménkový test

- U 15 náhodně vybraných pacientů byla vyhodnocena doba, kterou museli strávit v čekárně, než byli sestrou pozváni do ordinace. Na 5% hladině významnosti testujte nulovou hypotézu, že medián čekací doby je roven půl hodině.

Pacient č.	čekací doba (min)	medián	rozdíl	Větší než medián?
1	1	30	-29	Ne
2	45	30	15	Ano
3	25	30	-5	Ne
4	15	30	-15	Ne
5	34	30	4	Ano
6	19	30	-11	Ne
7	31	30	1	Ano
8	25	30	-5	Ne
9	8	30	-22	Ne
10	12	30	-18	Ne
11	20	30	-10	Ne
12	15	30	-15	Ne
13	40	30	10	Ano
14	20	30	-10	Ne
15	10	30	-20	Ne



$$S_z^+ = 4$$

Kritický obor: $W = (0, 3) \cup (12, 15)$

Hodnota statistiky se realizuje mimo kritický obor hodnot → **nezamítáme H_0**

Příklad: Řešení v softwaru

1) Výstup Wilcoxonova testu

Testová statistika: $\min(S_w^+, S_w^-)$

Wilcoxon Matched Pairs Test (v cekarne.sta)				
Marked tests are significant at p < .05000				
Pair of Variables	Valid N	T	Z	p-value
doba & median	15	19.00000	2.328644	0.019879

Statistika a p-hodnota pro asymptotickou variantu testu (používat pouze pro $N > 30$)

Počet nenulových rozdílů

2) Výstup znaménkového testu

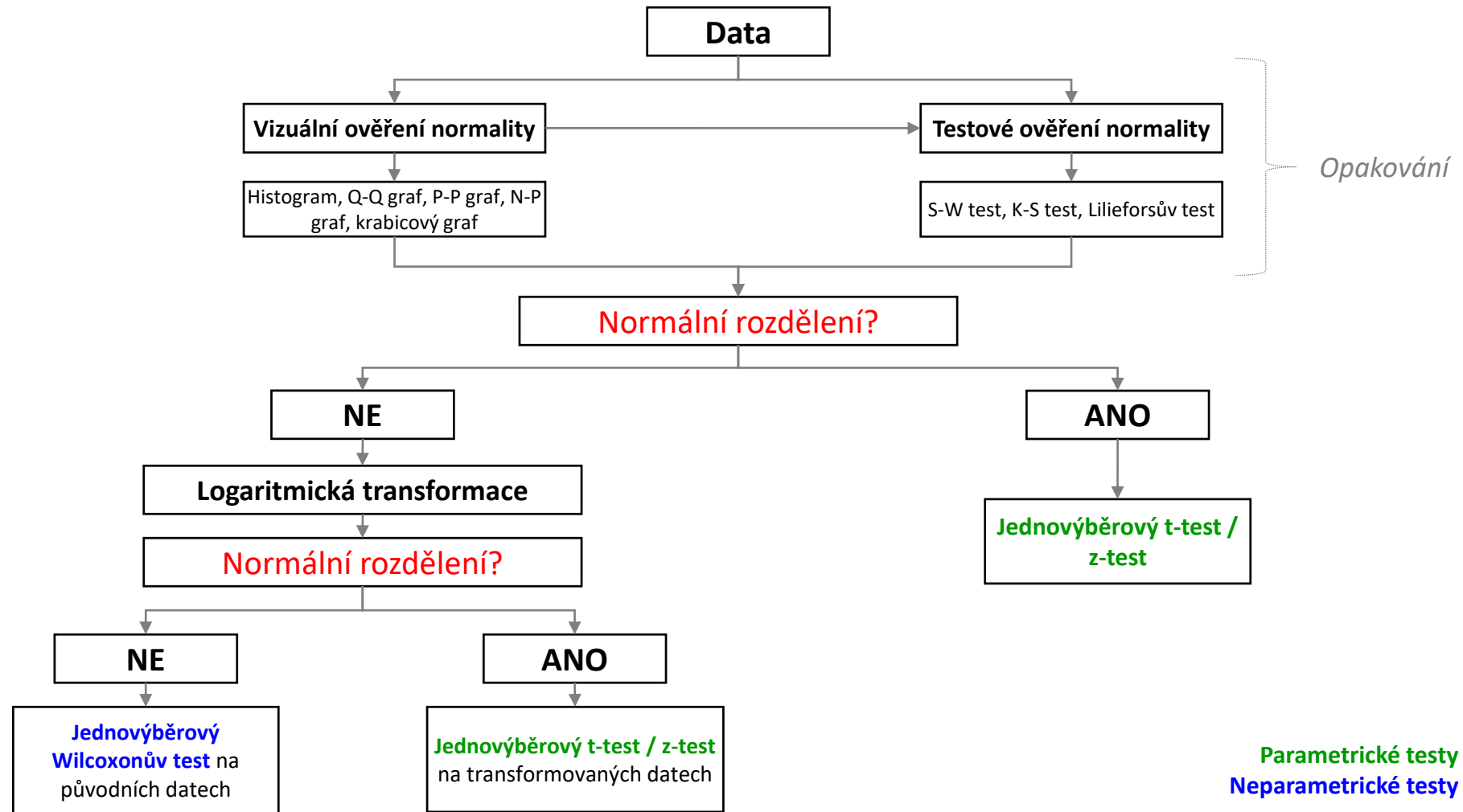
Podíl hodnot menších než testovaný medián

Sign Test (v cekarne.sta)				
Marked tests are significant at p < .05000				
Pair of Variables	No. of Non-ties	Percent v < V	Z	p-value
doba & median	15	73.33333	1.549193	0.121335

Statistika a p-hodnota pro asymptotickou variantu testu (používat pouze pro $N > 20$)

Počet nenulových rozdílů

Schéma při testování pomocí jednovýběrových testů



Parametrické dvouvýběrové statistické testy

Dvouvýběrový nepárový t-test

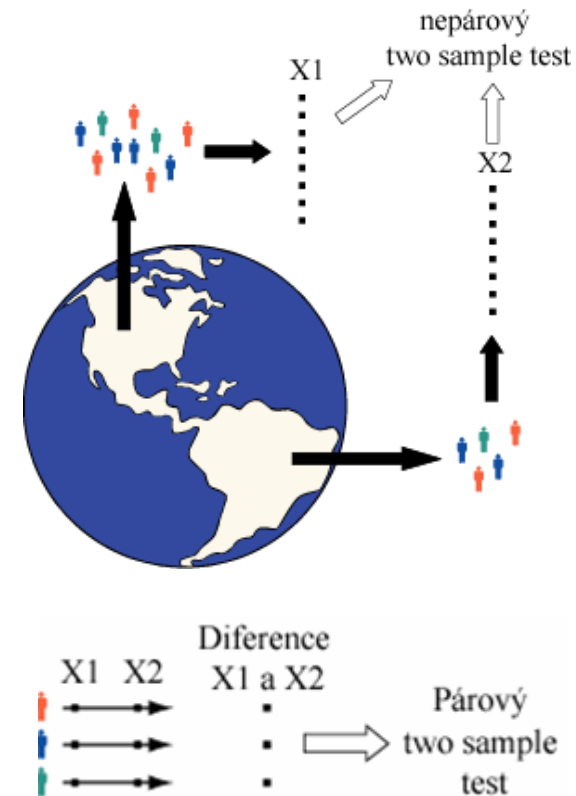
Dvouvýběrový párový t-test

Anotace

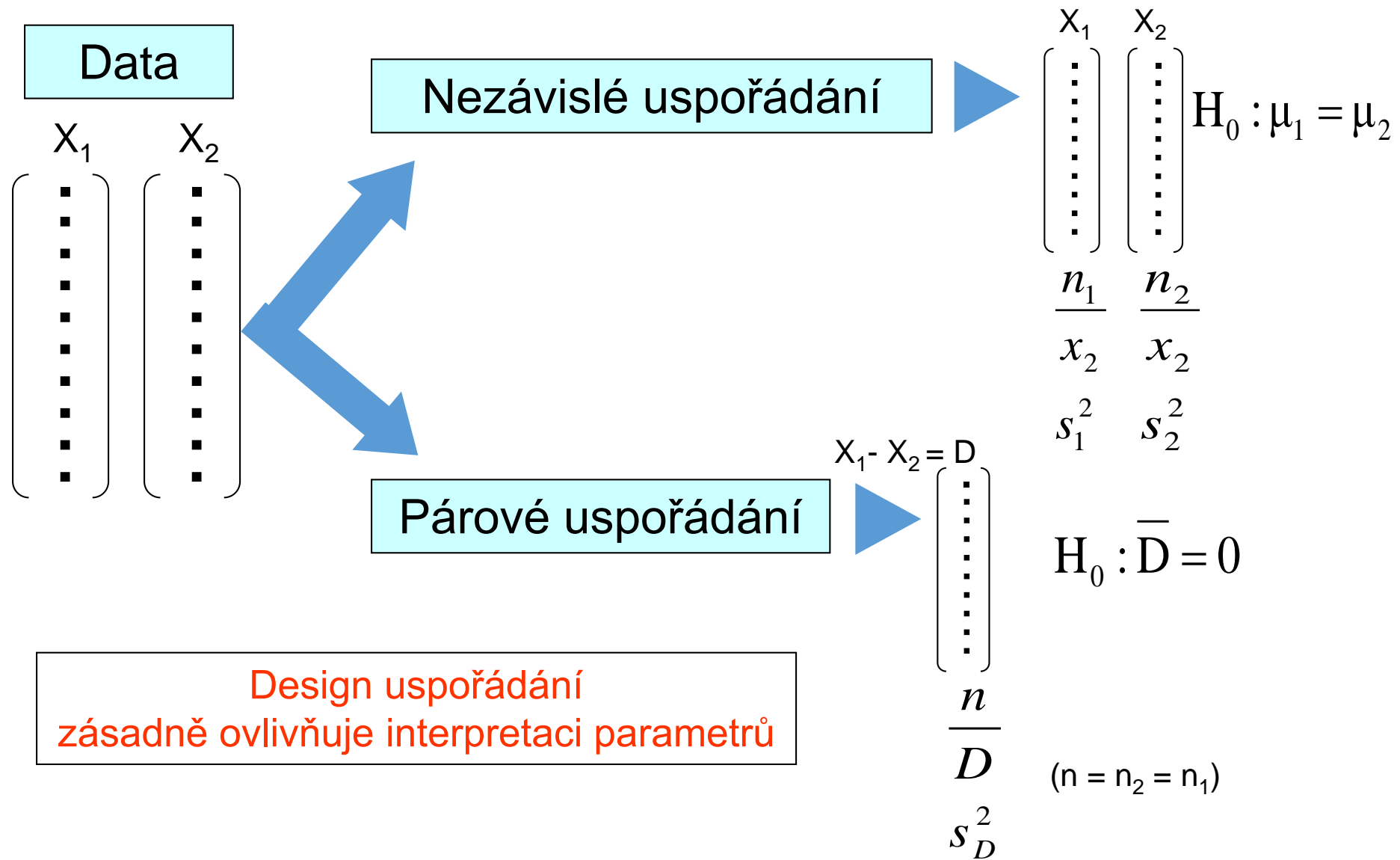
- Jedním z nejčastějších úkolů statistické analýzy dat je srovnání spojitých dat ve dvou skupinách pacientů.
- Na výběr je celá škála testů, výběr konkrétního testu se pak odvíjí od toho, zda je o srovnání párové nebo nepárové a zda je vhodné použít test parametrický (má předpoklady o rozložení dat) nebo neparametrický (nemá předpoklady o rozložení dat, nicméně má nižší vypovídací sílu).
- Nejznámějšími testy z této skupiny jsou tzv. t-testy používané pro srovnání průměrů dvou skupin hodnot

Dvouvýběrové testy: párové a nepárové I

- Při použití two sample testů srovnáváme spolu dvě rozložení. Jejich základním dělením je podle designu experimentu na testy párové a nepárové.
- Základním testem pro srovnání dvou nezávislých rozložení spojitých čísel je nepárový two-sample t-test
- Základním testem pro srovnání dvou závislých rozložení spojitých čísel je párový two-sample t-test

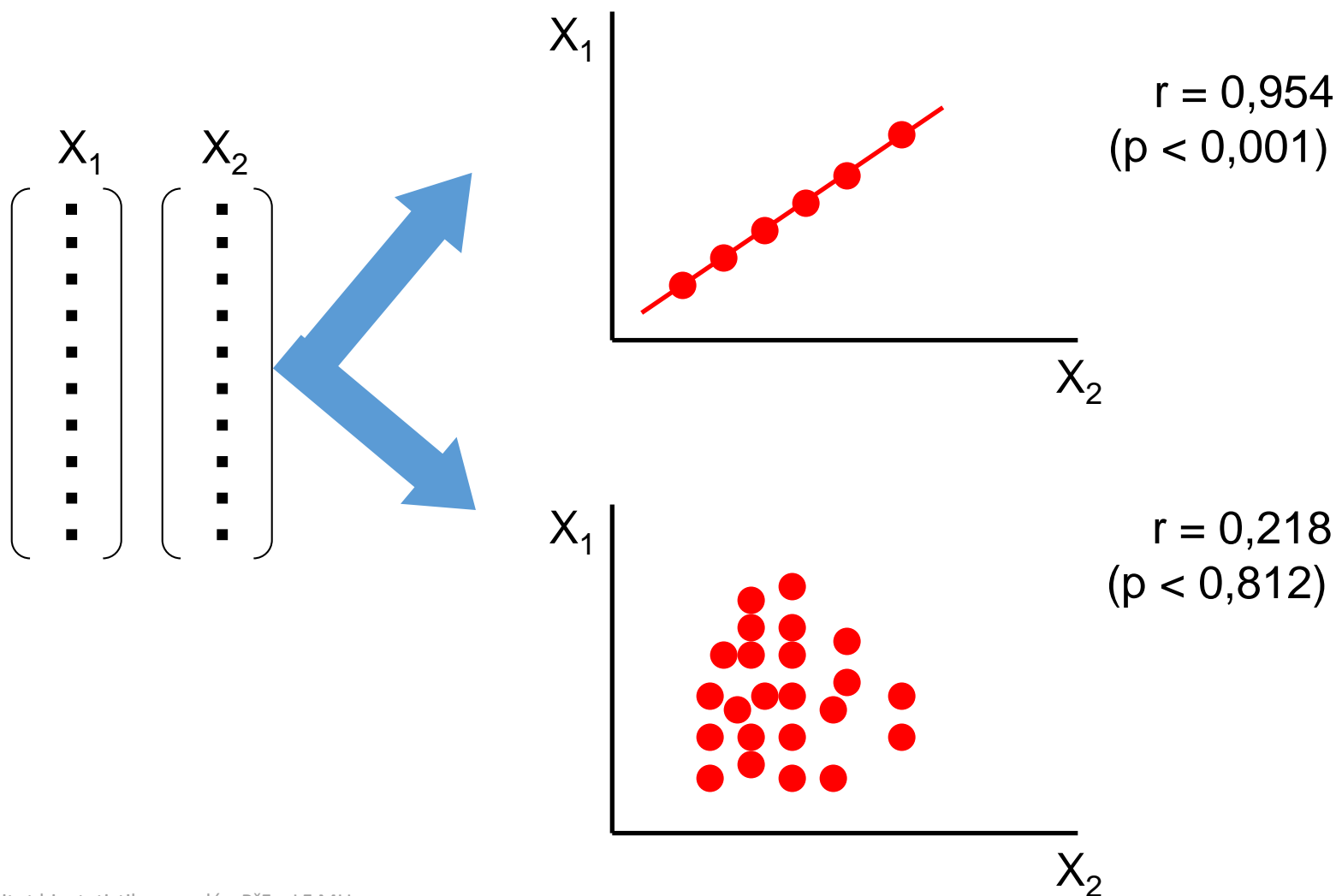


Dvouvýběrové testy: párové a nepárové II



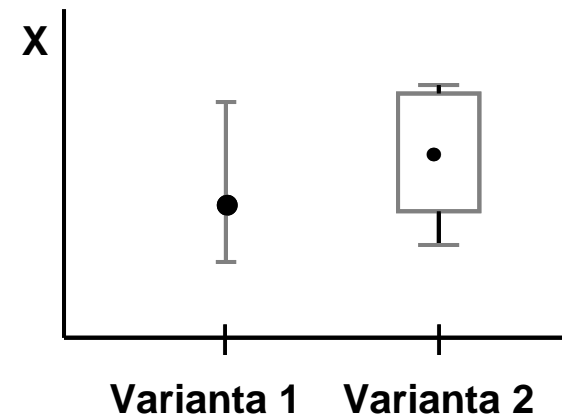
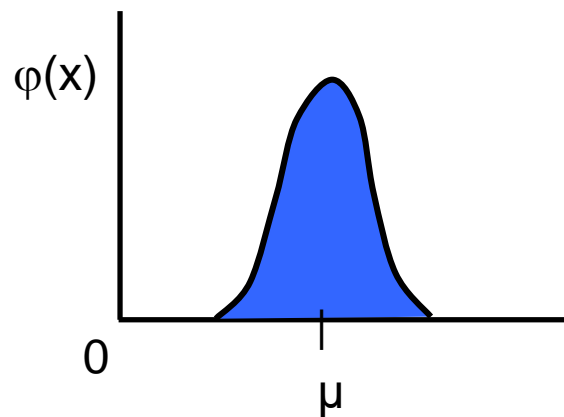
Dvouvýběrové testy: párové a nepárové III

- Identifikace párovitosti (Korelace, Kovariance)



Předpoklady nepárového dvouvýběrového t-testu

- Náhodný výběr subjektů jednotlivých skupin z jejich cílových populací
- Nezávislost obou srovnávaných vzorků
- Přibližně normální rozložení proměnné ve vzorcích, drobné odchylky od normality ovšem nejsou kritické, test je robustní proti drobným odchylkám od tohoto předpokladu, normalita může být testována testy normality
- Rozptyl v obou vzorcích by měl být přibližně shodný (homoscedastic). Tento předpoklad je testován několika možnými testy – Levenův test nebo F-test.
- Vždy je vhodné prohlédnout histogramy proměnné v jednotlivých vzorcích pro okometrické srovnání a ověření předpokladů normality a homogenity rozptylu – nenahradí statistické testy, ale poskytne prvotní představu.



Nepárový dvouvýběrový t-test – výpočet I

- nulová hypotéza: průměry obou skupin jsou shodné, alternativní hypotéza je, že nejsou shodné, two tailed test
- prohlédnout průběh dat, průměr, medián apod. pro zjištění odchylek od normality a nehomogenita rozptylu, provést F –test
- F-test pro srovnání dvou výběrových rozptylů
 - Používá se pro srovnání rozptylu dvou skupin hodnot, často za účelem ověření homogenity rozptylu těchto skupin dat.
- V případě ověření homogenity je testována hypotéza shody rozptylů (two tailed); v případě shodných rozptylů je vše v pořádku a je možné pokračovat ve výpočtu t-testu, v opačném případě není vhodné test počítat.

H_0	H_A	Testová statistika
$\sigma_1^2 \leq \sigma_2^2$	$\sigma_1^2 > \sigma_2^2$	$F = \frac{s_1^2}{s_2^2}$
$\sigma_1^2 \geq \sigma_2^2$	$\sigma_1^2 < \sigma_2^2$	$F = \frac{s_2^2}{s_1^2}$
$\sigma_1^2 = \sigma_2^2$	$\sigma_1^2 \neq \sigma_2^2$	$F = \frac{\max(s_1^2; s_2^2)}{\min(s_1^2; s_2^2)}$

Nepárový dvouvýběrový t-test – výpočet II

- Výpočet testové statistiky (stupně volnosti jsou $\nu = n_1 + n_2 - 2$):

$$t = \frac{\text{Rozdíl} - \text{průměr}}{SE(\text{rozdíl průměrů})} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$
 vážený odhad rozptylu

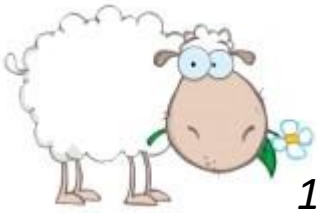
- výsledné t srovnáme s tabulární hodnotou t pro dané stupně volnosti a α (obvykle $\alpha=0,05$)
- Lze spočítat interval spolehlivosti pro rozdíl průměrů (např. 95%), počet stupňů volnosti a s^2 odpovídají předchozím vzorcům

$$(\bar{x}_1 - \bar{x}_2) \pm t_{0,975} SE(\bar{x}_1 - \bar{x}_2) = (\bar{x}_1 - \bar{x}_2) \pm t_{0,975} \sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

Nepárový dvouvýběrový t-test – výpočet shrnutí

- Nulová hypotéza: průměry obou skupin jsou shodné
 - Alternativní hypotéza je, že nejsou shodné.
- Prohlédnout průběh dat, průměr, medián apod.
 - Ověřit normalitu dat (např. Shapiro-Wilk test)
 - Ověřit homogenitu rozptylů (F-test)
 - V případě ověření homogenity je testována hypotéza shody rozptylů; v případě shodných rozptylů je vše v pořádku a je možné pokračovat ve výpočtu t-testu, v opačném případě není vhodné test počítat.
- Vypočítat hodnotu testové statistiky a p-hodnotu. Když je vypočítaná p-hodnota menší než 0,05, zamítáme nulovou hypotézu.

Přednáška 7



Příklad : Nepárový dvouvýběrový t-test

1. skupina, N=30

- Průměrná hmotnost ovcí v čase páření byla srovnávána pro kontrolní skupinu a skupinu krmenou zvýšenou dávkou potravy. Kontrolní skupina obsahuje 30 ovcí, skupina se zvýšeným příjmem potravy pak 24 ovcí.
- Vlastní experiment byl prováděn tak, že na začátku máme 54 ovcí (ideálně stejného plemene, stejně staré atd.), které náhodně rozdělíme do dvou skupin (náhodné rozdělování objektů do pokusných skupin je objektem celého specializovaného odvětví statistiky nazývaného randomizace). Poté co experiment proběhne, musíme nejprve ověřit teoretický předpoklad pro využití nepárového t-testu. Pro obě skupiny jsou vykresleny grafy (můžeme též spočítat základní popisnou statistiku), na kterých můžeme posoudit normalitu a homogenitu rozptylu, kromě okometrického pohledu můžeme pro ověření normality použít testy normality, pro ověření homogenity rozptylu pak F-test.
- Pokud platí všechny předpoklady dvouvýběrového nepárového t-testu, můžeme spočítat testovou statistiku, výsledné t je 2,43 s 52 stupni volnosti, podle tabulek je $t_{0,975(52)} = 2,01$, tedy $|t| > t_{0,975(52)}$ a nulovou hypotézu můžeme zamítnout, skutečná pravděpodobnost je pak 0,018. Rozdíl mezi skupinami je 1,59 kg ve prospěch skupiny se zvýšeným příjmem.

$$t = \frac{\text{Rozdíl .prumeru}}{SE(\text{rozdil .prumeru})} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \quad v = n_1 + n_2 - 2$$

- Pro rozdíl mezi oběma soubory jsou spočítány 95% intervaly spolehlivosti jako $1,59 \pm 2,01 * (0,655)$ kg, což odpovídá rozsahu 0,28 až 2,91 kg. To, že interval spolehlivosti nezahrnuje 0 je dalším potvrzením, že mezi skupinami je významný rozdíl – jde o další způsob testování významnosti rozdílů mezi skupinami dat – nulovou hypotézu o tom, že rozdíl průměrů dvou skupin dat je roven nějaké hodnotě zamítáme v případě, kdy 95% interval spolehlivosti rozdílu nezahrnuje tuto hodnotu (v tomto případě 0).

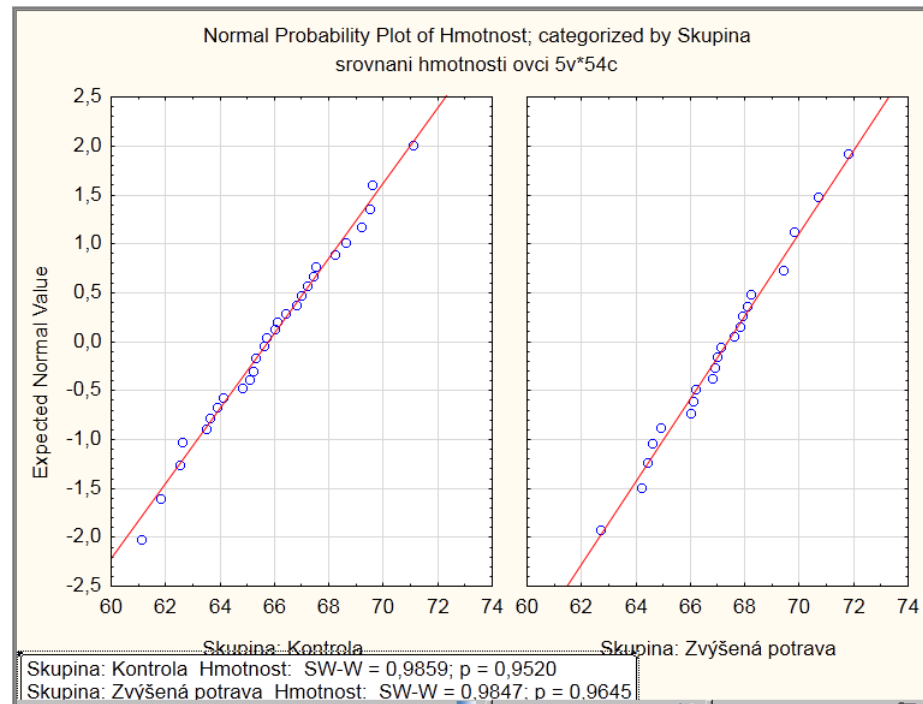
$$(\bar{x}_1 - \bar{x}_2) \pm t_{0,975} SE(\bar{x}_1 - \bar{x}_2) = (\bar{x}_1 - \bar{x}_2) \pm t_{0,975} \sqrt{s^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$



2. skupina, N=24

Příklad : Nepárový dvouvýběrový t-test

- Nejprve ověřte normalitu hmotnosti jednak ve skupině kontroly a ve skupině se zvýšenou potravou



- V obou případech se tečky odchylují od přímky jenom málo a p-hodnoty S-W testu převyšují 0,05. Předpoklad o normálním rozložení dat v obou skupinách je oprávněný.

Příklad : Nepárový dvouvýběrový t-test

•POZOR: Výstupní tabulku vyhodnocujeme zezadu!!!

Výběrový průměr u 1. skupiny

Výběrový průměr u 2. skupiny

Výběrová směrodatná odchylka u 2. skupiny

Rozsah výběru 1. skupiny

Rozsah výběru 2. skupiny

T-tests; Grouping: Skupina (srovnani hmotnosti ovcí)											
Group 1: Kontrola											
Group 2: Zvýšená potrava											
Variable	Mean Kontrola	Mean Zvýšená potrava	t-value	df	p	Valid N Kontrola	Valid N Zvýšená potrava	Std.Dev. Kontrola	Std.Dev. Zvýšená potrava	F-ratio Variances	p Variances
Hmotnost	65,77333	67,36667	-2,43226	52	0,018483	30	24	2,497162	2,252470	1,229066	0,617383

Hodnota testové statistiky
(pro test shody středních hodnot)

Počet stupňů volnosti

Testová statistika pro test shody rozptylů
(F-test)

**Tyto sloupce lze interpretovat pouze
pokud rozdíl mezi rozptyly byl neprůkazný !!!**

Párový dvouvýběrový t-test

- Skupiny dat jsou spojeny přes objekt měření, příkladem může být měření parametrů pacienta před léčbou a po léčbě (nemusí jít přímo o stejný objekt, dalším příkladem mohou být např. krysy ze stejné linie).
- Oba soubory musí mít shodný počet hodnot, protože všechna měření v jednom souboru musí být spárována s měřením v druhém souboru. Při vlastním výpočtu se potom počítá se změnou hodnot (diferencí) subjektů v obou souborech.
- V případě, že se nejedná o měření na témže subjektu je vhodné si před párovým testem ověřit si, zda existuje vazba mezi oběma skupinami – vynesení do grafu, korelace.

Existuje několik možných designů experimentu, stručně lze sumarizovat:

- pokus je párový a jako párový se projeví
- párové provedení pokusu – párově se neprojeví
 - možná párovost není
 - špatně provedený pokus – malé n, velká variabilita, špatný výběr jedinců
- čekali jsme nezávislé a jsou
- čekali jsem nezávislé a nejsou
 - vazba
 - náhoda

Párový dvouvýběrový t-test

- Tento test nemá žádné předpoklady o rozložení vstupních dat, protože je počítán až na základě jejich diferencí.
- Tyto diference by měly být normálně rozloženy a otázkou v párovém t-testu je, zda se průměrná hodnota diferencí rovná nějakému číslu, typicky jde o srovnání s nulou jako důkaz neexistence změny mezi oběma spárovanými skupinami.
- V podstatě jde o one sample t-test, kde místo rozdílu průměru vzorku a cílové populace je uveden průměr diferencí a srovnávané číslo (0 v případě otázky, zda není rozdíl mezi vzorky).

- Pro srovnání s 0 (testovou statistikou je t rozložení):
$$t = \frac{\bar{D}}{s} \sqrt{n} \quad \nu = n - 1$$

- Někdy je obtížné rozhodnout, zda jde nebo nejde o párové uspořádání, párový test by měl být použit pouze v případě, že můžeme potvrdit vazbu (korelace, vnesení do grafu), jedním z důvodů proč toto ověřovat je fakt, že v případě párového t-testu není nutné brát ohled na variabilitu původních dvou souborů, tento předpoklad však platí pouze v případě vazby mezi proměnnými. Výpočet obou typů testů se vlastně liší v použité s, jednou jde o s diferencí, v druhém případě o složený odhad rozptylu obou souborů.
- Zda je párové uspořádání efektivnější lze určit na základě:
 - Síly vazby
 - Je-li s_D výrazně menší než $s_{x_1-x_2}$

- Závislost je možné rozepsat pomocí vzorce:
$$s_D^2 \cong \sigma_{x_1}^2 + \sigma_{x_2}^2 - 2Cov(x_1; x_2)$$

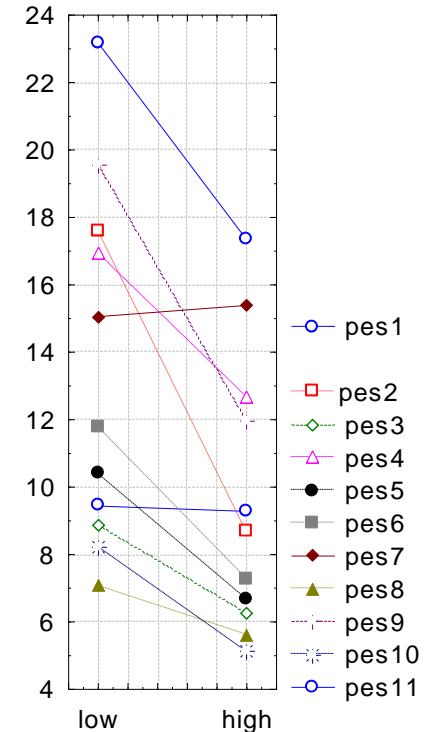
- v případě $Cov=0$, tedy v případě neexistence vazby pak s_D^2 odpovídá součtu původních rozptylů, tedy přibližně $S_{x_1-x_2}$.

Příklad 1: Párový dvouvýběrový test

- Byl prováděn pokus s dietou 11 diabetických psů, každý pes byl vystaven dvěma dietám s odlišným typem sacharidů (snadno vstřebatelné X pozvolna se rozkládající na glukózu), hodnoty krevní glukózy v průběhu jednotlivých diet mají být srovnány pro zjištění vlivu diety na hladinu krevní glukózy. Protože každý pes absolvoval obě diety, jde o párové uspořádání, kdy výsledky hodnoty v obou pokusech jsou spojeny přes pokusné zvíře.
- Nulová hypotéza zní, že skutečný průměrný rozdíl mezi oběma dietami je 0, alternativní hypotéza zní, že to není 0.
 - Pro každého psa je spočítán rozdíl mezi jeho hladinou glukózy při obou dietách a měly by být ověřeny předpoklady pro one sample t-test – tedy alespoň přibližně normální rozložení.
 - Je spočítána testová charakteristika, výpočet vlastně probíhá jako one-sample t-test, kde je zjišťována významnost průměru diferencí obou souborů jako rozdíl mezi touto hodnotou a nulou (nula je hodnota, kterou by průměrná diference měla nabývat, pokud platí nulová hypotéza). $T=4.37$ s 10 stupni volnosti, skutečná hodnota $p=0,0014$ a tedy na hladině $p=0,05$ můžeme nulovou hypotézu zamítnout

$$t = \frac{\text{rozdíl}_\text{průměru}_\text{vzorku}_\text{a}_\text{populace}}{SE(\text{průměru})} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{\bar{x} - \mu}{s} \sqrt{n}$$

- Závěrem můžeme říci, že nulová hypotéza neexistence rozdílu mezi oběma dietami byla zamítnuta, což znamená, že high-fibre dieta má významný vliv na snížení hladiny krevní glukózy.



Příklad 2: Párový dvouvýběrový test

- Byl prováděn pokus s dietou u 18 diabetických krys, každá krysa byla vystavena dvěma dietám (jedné nové speciální a jedné kontrolní dietě). Protože každá krysa absolvovala obě diety, jde o párové uspořádání, kdy hodnoty v obou pokusech jsou spojeny přes pokusné zvíře. Zjistěte, zda testovaná dieta způsobí změnu hmotnosti u krys (zda se liší hmotnost krys po nové speciální a po kontrolní dietě).
- 1. Nulová hypotéza zní, že skutečný průměrný rozdíl v hmotnosti krys po speciální a kontrolní dietě je nulový (speciální dieta nevedla ke změně hmotnosti ve srovnání s kontrolní dietou), alternativní hypotéza zní, že rozdíl hmotností je odlišný od nuly (speciální dieta vedla ke změně hmotnosti ve srovnání s kontrolní dietou).
- 2. Pro každou krysu je spočítán rozdíl hmotností naměřených po obou dietách a měly by být ověřeny předpoklady pro jednovýběrový t-test – alespoň přibližně normální rozložení diferencí.
- 3. Je spočítána testová statistika, výpočet vlastně probíhá jako jednovýběrový t-test, kde je zjišťována významnost průměru diferencí obou souborů jako rozdíl mezi touto hodnotou a nulou (0 je hodnota, kterou by průměrná diference měla nabývat, pokud platí nulová hypotéza). $T = -1,72$ s 17 stupni volnosti, skutečná p-hodnota = 0,102 a tedy na hladině významnosti $\alpha = 0,05$ nemůžeme nulovou hypotézu zamítnout.

$$t = \frac{\text{rozdíl}_\text{průměru}_\text{vzorku}_\text{a}_\text{populace}}{SE(\text{průměru})} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{\bar{x} - \mu}{s} \sqrt{n}$$

- 4. Závěrem můžeme říci, že nulová hypotéza neexistence rozdílu vlivu na snížení váhy mezi oběma dietami nebyla zamítnuta.



Příklad: Párový dvouvýběrový test

Výběrový průměr Výběrová směrodatná odchylka

Počet pozorování

T-test for Dependent Samples (efektivita diety pro krysy)								
Marked differences are significant at $p < ,05000$								
Variable	Mean	Std.Dv.	N	Diff.	Std.Dv. Diff.	t	df	p
testovaná dieta	186,0556	59,52011						
kontrolní dieta	191,7222	69,65022	18	-5,66667	13,91994	-1,72714	17	0,102266

Hodnota testovacího kritéria

Průměrná hodnota diferencí

Výběrová směrodatná odchylka diferencí

Neparametrické dvouvýběrové statistické testy

Nepárový Mannův-Whitneyův test

Párový Wilcoxonův a znaménkový test

Mannův-Whitneyův U test

- Neparametrická alternativa dvouvýběrového t-testu.
- Počítá s pořadím dat v souborech namísto s originálními daty.
- Předpoklad: rozdělení pravděpodobnosti veličiny ve skupinách se může lišit pouze posunutím.

Postup:

1. Stanovíme nulovou a alternativní hypotézu ($F(x)$ =distribuční funkce):

$$H_0: F(x_1)=F(x_2)$$

$$H_1: F(x_1)\neq F(x_2).$$

2. Čísla obou souborů jsou sloučena a je určeno jejich pořadí v tomto sloučeném souboru.
3. Pro oba výběry zvlášť je spočítán součet pořadí (T_1 a T_2).
4. Ze součtů pořadí ve skupinách je určena finální hodnota testové statistiky U .

$$U_1 = n_1 n_2 + \frac{n_1 - (n_1 + 1)}{2} - T_1$$

$$U_2 = n_1 n_2 + \frac{n_2 - (n_2 + 1)}{2} - T_2$$

$$U = \min(U_1, U_2)$$

5. Hodnotu testové statistiky U porovnáme s kritickou hodnotou testu, pokud je tato hodnota menší než kritická hodnota testu, zamítáme nulovou hypotézu shody distribučních funkcí obou skupin.

Mannův-Whitneyův U test – asymptotická varianta

5. Pro velká n_1 a n_2 (>30) lze využít asymptotické normality statistiky U.

$$E(U) = \frac{n_1 n_2}{2} \quad D(U) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

6. Pro testování lze využít Z-statistiky:

$$Z = \frac{U - E(U)}{\sqrt{D(U)}} \approx N(0,1)$$

7. Pokud $|Z| \geq u_{1-\alpha/2}$ zamítáme nulovou hypotézu o shodnosti distribučních funkcí

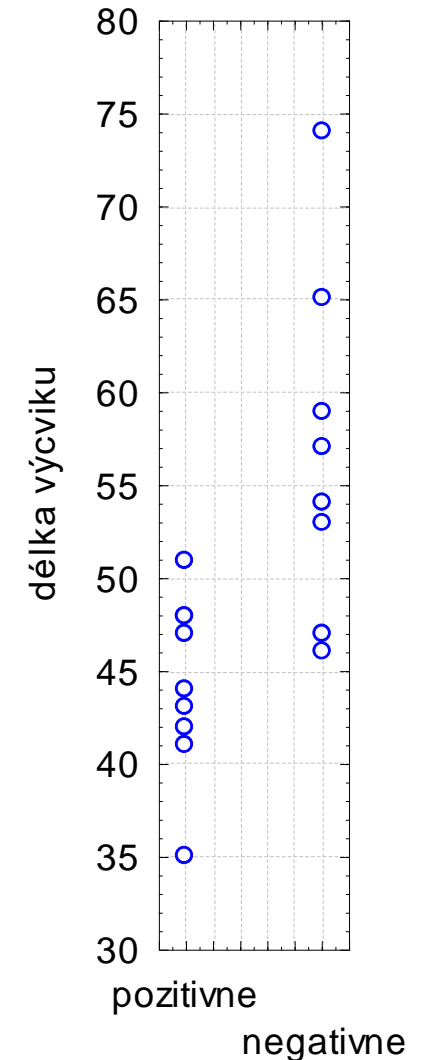
Mannův-Whitneyův U test

- Stejně jako řada jiných neparametrických testů počítá i tento test s pořadím dat v souborech namísto s originálními daty. Jde o neparametrickou obdobu nepárového t-testu a z těchto neparametrických testů má nejvyšší sílu testu (95% párového t-testu).
- V případě Mann-Whitney testu jsou nejprve čísla obou souborů sloučena a je vytvořeno jejich pořadí v tomto sloučeném souboru, pak jsou hodnoty vráceny do původních souborů a nadále se pracuje již jen s jejich pořadím.
- Pro oba soubory je tedy vytvořen součet pořadí a menší z obou součtů je porovnán s kritickou hodnotou testu, pokud je tato hodnota menší než kritická hodnota testu, zamítáme nulovou hypotézu shody distribučních funkcí obou skupin.

X1	X2	ALL	Rank ALL	X1 rank	X2 rank
27	25	25	5	6	5
35	29	29	7,5	11	7,5
38	31	31	9	13	9
37	23	23	4	12	4
39	18	18	2	14	2
29	17	17	1	7,5	1
41	32	32	10	15	10
	19	19	3		3
		27	6		
		35	11		
		38	13		
		37	12		
		39	14		
		29	7,5		
		41	15		

Příklad: Mannův-Whitneyův U test

- 17 štěňat bylo trénováno v chození na záchod metodou pozitivní motivace (pochvala, když jde na záchod venku) nebo negativní motivace (trest, když jde na záchod doma). Jako parametr bylo měřeno, za kolik dní je štěně vycvičeno.
- Nulová hypotéza je, že není rozdíl v metodách tréninku, tedy, že oběma metodami je štěně vycvičeno za stejnou dobu.
- Po srovnání rozložení + kvůli nízkému počtu hodnot je vhodné použít neparametrický test.
- Je vytvořeno pořadí hodnot v kompletním souboru.
- Hodnota testové statistiky je určena ze součtu pořadí hodnot v jednotlivých skupinách.
- **Jak dopadne testování?**



Příklad: Řešení v softwaru

Součet pořadí T_1

Součet pořadí T_2

Hodnota Z statistiky

Mann-Whitney U Test (Spreadsheet15)										
By variable skupina										
Marked tests are significant at p <,05000										
variable	Rank Sum Group 1	Rank Sum Group 2	U	Z	p-value	Z adjusted	p-value	Valid N Group 1	Valid N Group 2	2*1sided exact p
delka	49,50000	103,5000	135,0000	-2,11695	0,034265	-2,11955	0,034045	8	9	0,027396

Hodnota testové statistiky

Asymptotická p-hodnota

Přesná p-hodnota

(použít, jestliže rozsah výběru je menší než 30)

Párový Wilcoxonův a znaménkový test

- Vycházíme z rozdílů párových hodnot a přecházíme na design jednovýběrových testů
- Testuje, zda je **medián diferencí (D)** párových hodnot roven hodnotě c
 $H_0: D_{0.5}=c$ proti $H_1: D_{0.5} \neq c$.

Wilcoxonův párový test

1. Spočítáme rozdíly **diferencí** výběru s testovanou hodnotou mediánu $= c$.
2. Absolutní hodnoty rozdílů uspořádáme vzestupně a přiřadíme jim pořadí.
3. Spočítáme statistiky S_w^+ a S_w^- , které odpovídají součtu pořadí kladných (S_w^+) a záporných rozdílů (S_w^-). Jako finální hodnotu testové statistiky bereme minimum z S_w^+ a S_w^- . Nulovou hypotézu zamítáme, pokud hodnota testové statistiky menší nebo rovna tabelované kritické hodnotě (při dané hladině významnosti a počtu nenulových rozdílů).

Znaménkový párový test

1. Spočítáme rozdíly **diferencí** výběru s testovanou hodnotou mediánu $= c$.
2. Spočítáme statistiku S_z^+ , která odpovídá počtu kladných rozdílů \rightarrow test nevyužívá hodnot pořadí původních dat ale pouze informaci, zda se hodnota realizuje nad nebo pod mediánem \rightarrow dochází ke snížení síly testu
3. Nulovou hypotézu zamítáme, pokud statistika S_z^+ realizuje v kritickém oboru hodnot $W=(0,k_1)U(k_2,n)$, kde n odpovídá počtu nenulových rozdílů a hodnoty k_1 a k_2 lze dohledat v matematických tabulkách.

Příklad 2: Párový dvouvýběrový test

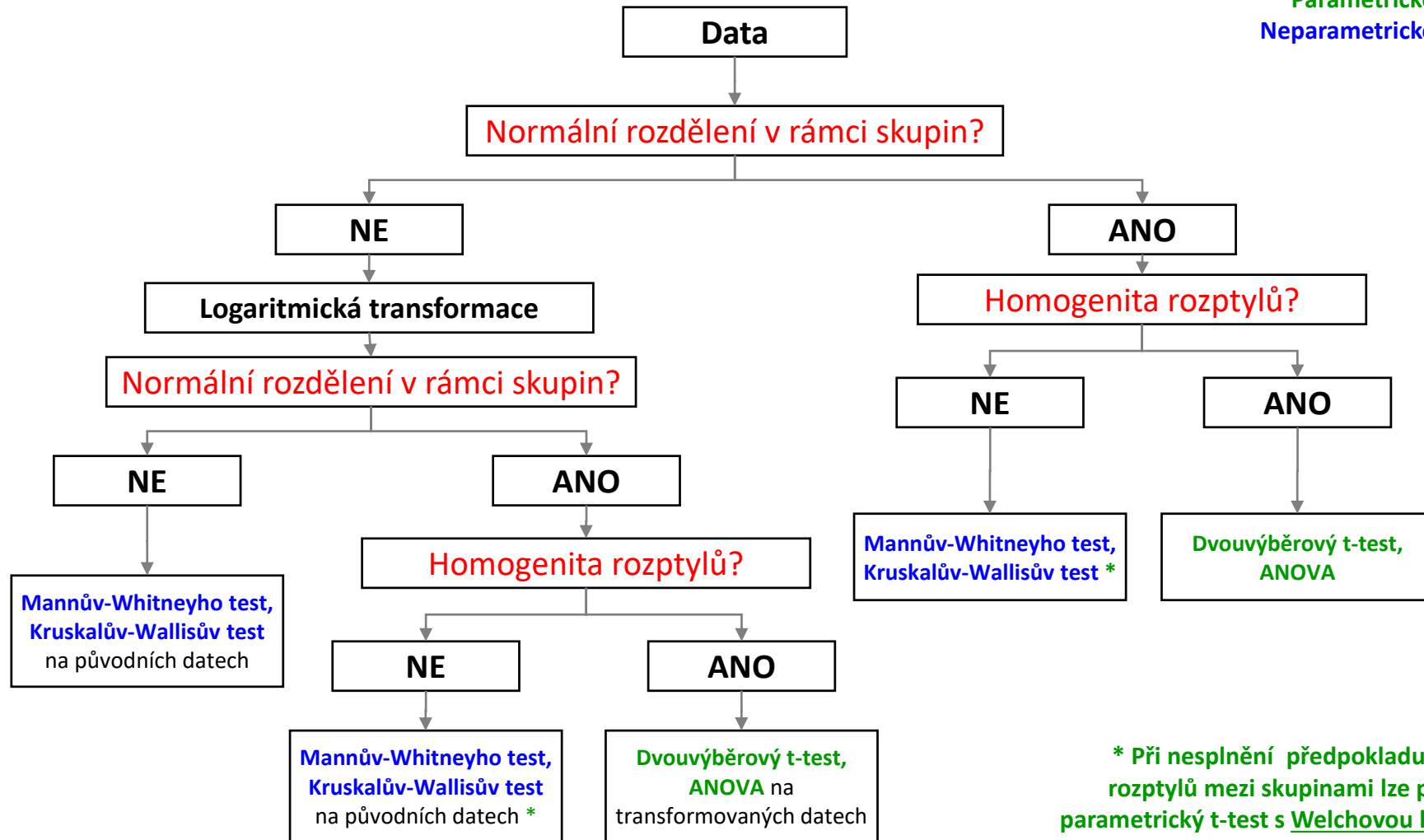
Byla testována nová dieta pro laboratorní krysy, při pokusu byl zjišťován její vliv na hmotnost v různých liniích kryš, bylo proto zvoleno párové uspořádání kdy krysy v obou dietách jsou spojeny přes svoji linii, tj. na začátku byly dvojice kryš stejné linie, jedna z nich byla náhodně přiřazena k dietě, druhá z dvojice pak do druhé diety.

1. nulová hypotéza je, že váha kryš není ovlivněna použitou dietou, alternativní, že ovlivnění dietou existuje
2. spočítáme diference – tyto diference jsou nenormální a proto je vhodné využít neparametrický test
3. Spočítáme sumu pořadí kladných a záporných diferencí, zde je menší suma záporných diferencí – 31
4. výsledkem výpočtu je $p > 0,05$ a tedy nemáme dostatečné důkazy pro zamítnutí nulové hypotézy, nelze říci, že by nová dieta byla efektivnější než stará
5. pro doplnění výsledků je vhodné zjistit také skutečnou velikost rozdílu hmotností ve skupinách, např. ve formě mediánu



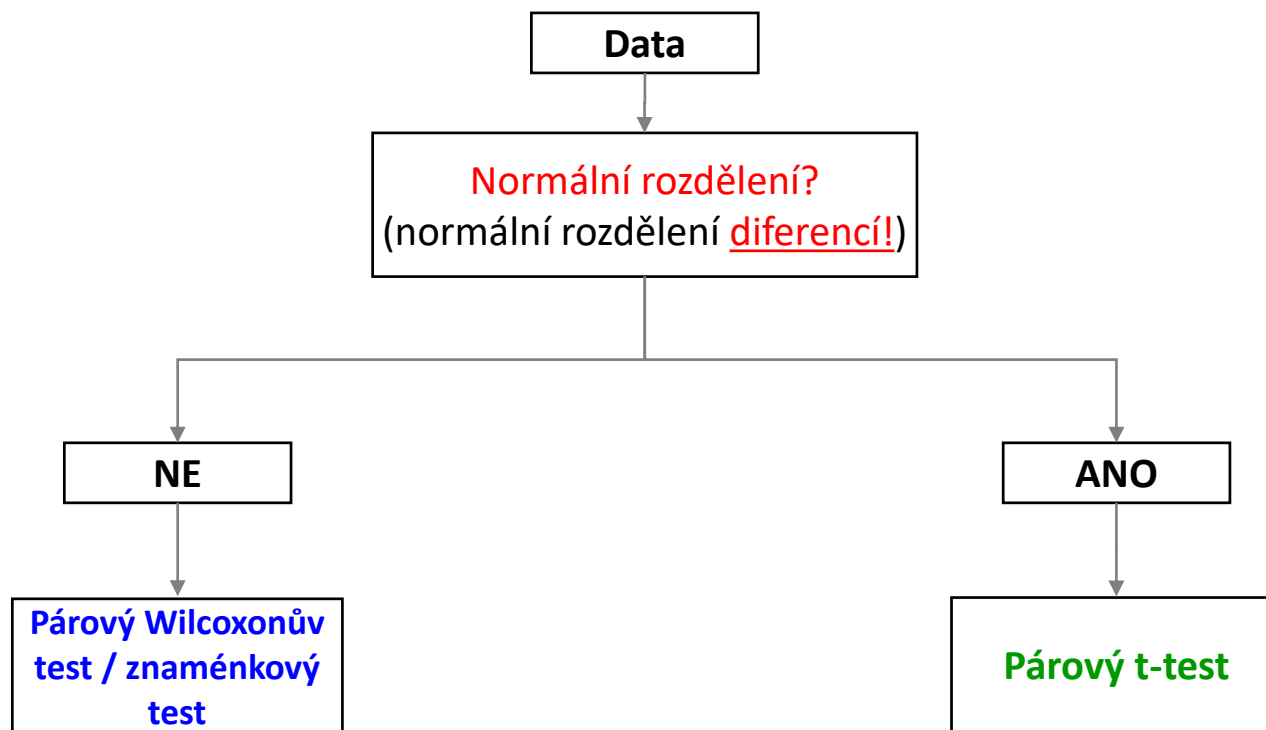
Schéma při testování 2 a více skupin

Parametrické testy
Neparametrické testy



* Při nesplnění předpokladu shody rozptylů mezi skupinami lze použít i parametrický t-test s Welchovou korekcí

Schéma při testování pomocí párových testů



Parametrické testy
Neparametrické testy

Binomické rozdělení

Popis binomického rozložení

Testování hypotéz binomicky rozložených dat

Anotace

- Kromě spojitých dat se setkáváme také s daty kategoriálními, jejichž nejjednodušším případem jsou data binární.
- Binární data jsou popsána binomickým rozdělením, od chování binomického rozdělení je odvozena popisná statistika binárních dat (procento výskytu jevu), její interval spolehlivosti a binomické testy pro srovnání procentuálního výskytů jevů v různých skupinách.

Alternativní rozdělení

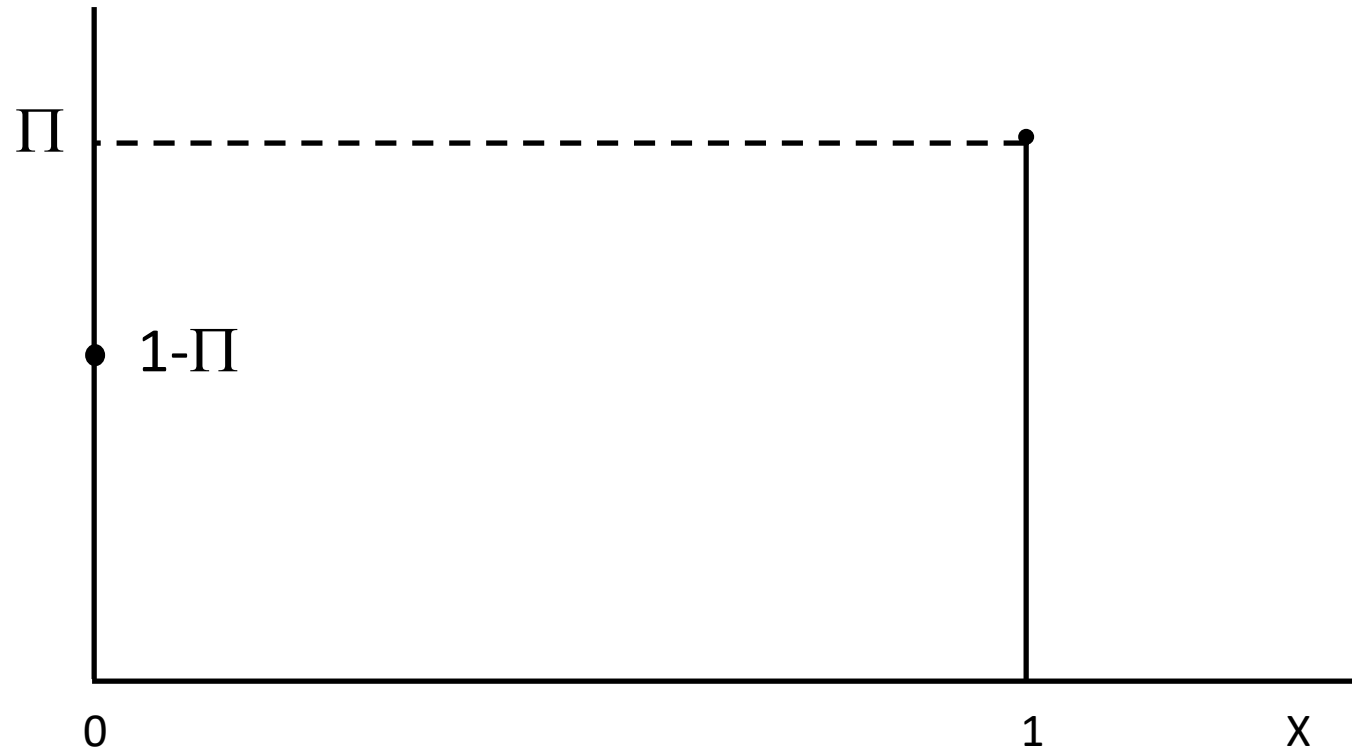
- Nastane jedna ze dvou možných varianta

$$\Pi(x) = \Pi \text{ pro } X = 1$$

$$\Pi(x) = 1 - \Pi \text{ pro } X = 0$$

$$\Pi(x) = 0 \text{ jinak}$$

} $X = 1 \dots\dots\text{jev}$



Binomické rozdělení

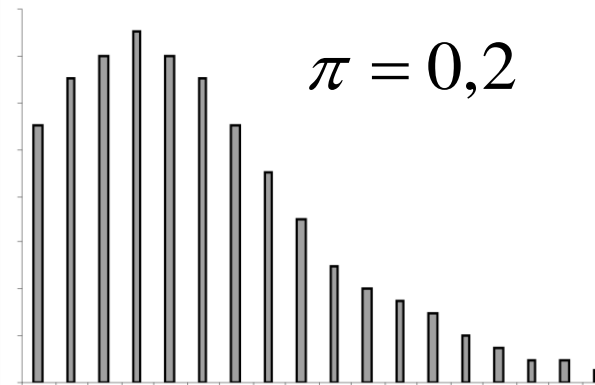
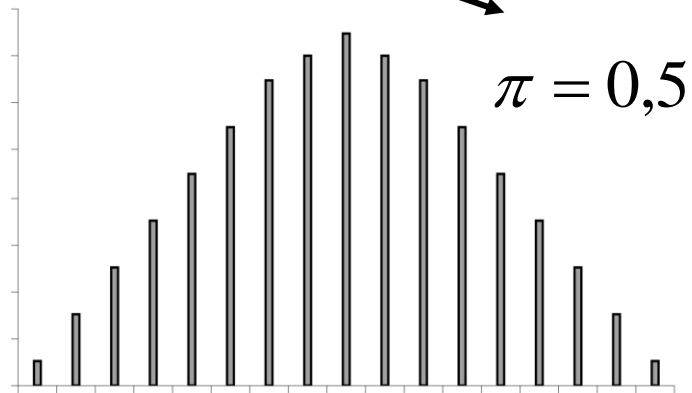
X celkový počet nastání jevu v n nezávislých pokusech

$$E(x) = n \cdot \Pi$$

$$D(x) = n \cdot \Pi (1 - \Pi)$$

$$\Pi \sim p$$

jediný parametr distribuce
určuje tvar distribuce



Binomické rozdělení jako model pro zkoumání výskytu sledovaného jevu

n počet nezávislých opakování (dotazů)

X počet lidí s jistým symptomem

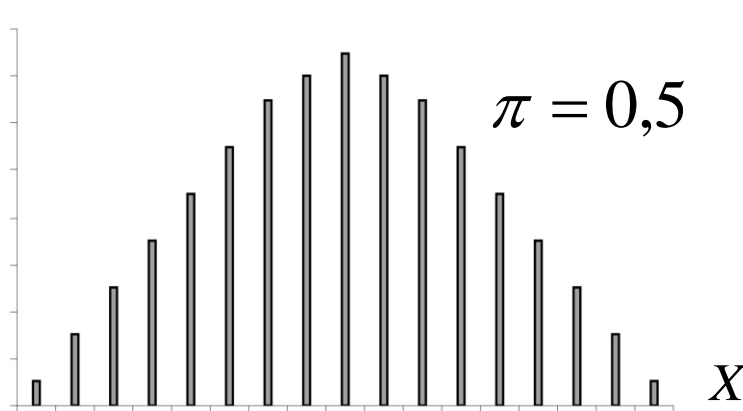
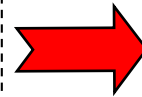
r znamená celkový počet nastání jevu v n nezávislých experimentech

$r : 0 \dots n$

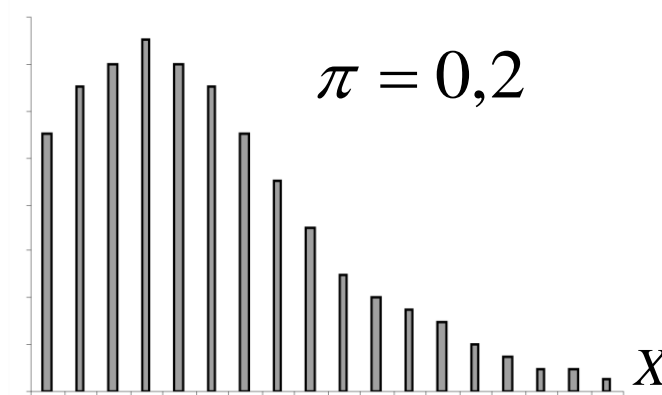
$p \sim \pi$.. jediný parametr binomického rozložení

p relativní četnost nastání jevu

p určuje tvar distribuce



$$p = \frac{r}{n}$$



Binomická proměnná X

Binomické rozdělení jako model

Jev: narození chlapce $\pi = 0,5$

n : rodina s 5 dětmi

r : 0,1,2,3,4,5 chlapců

$$P(r) = \binom{n}{r} \cdot p^r \cdot (1-p)^{(n-r)} = \frac{n!}{r!(n-r)!} \cdot p^r \cdot q^{(n-r)}$$

$$r = 0: \frac{5!}{(0! 5!)} \cdot (0,5)^0 \cdot (0,5)^5 = 0,031$$

$$r = 1: \frac{5!}{(1! 4!)} \cdot (0,5)^1 \cdot (0,5)^4 = 0,15625$$

$$r = 2: P(r) = 0,3125$$

$$r = 3: P(r) = 0,3125$$

$$r = 4: P(r) = 0,15625$$

$$r = 5: P(r) = 0,031$$

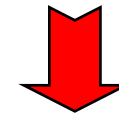
X : Binomická proměnná

Střed rozložení:

Rozptyl: $E(x) = n \cdot p$

$$D(x) = n \cdot p \cdot (1 - p)$$

Příklad: $n = 100$ respondentů
 $r = 20$ má symptom



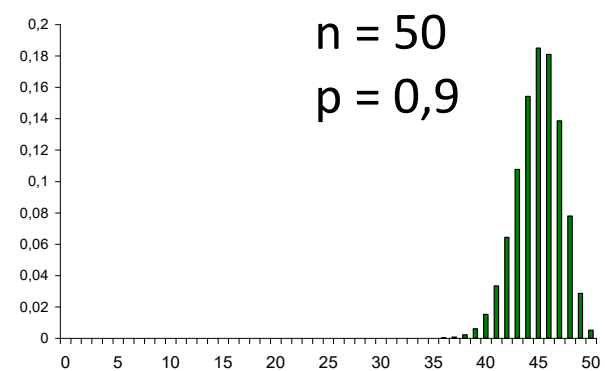
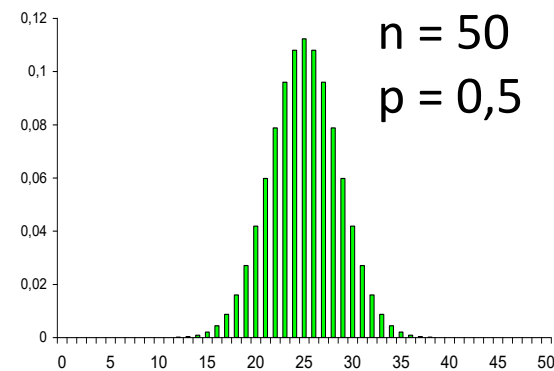
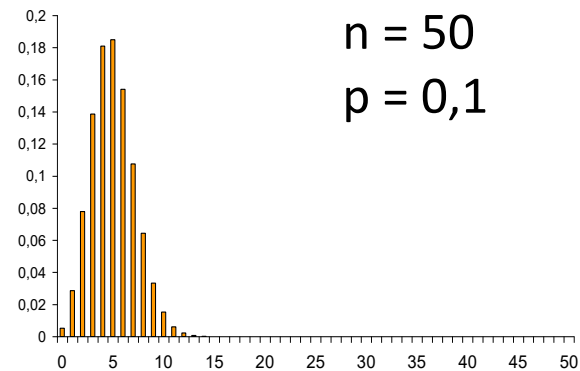
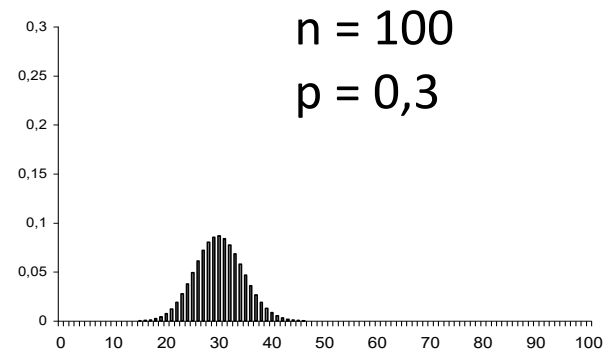
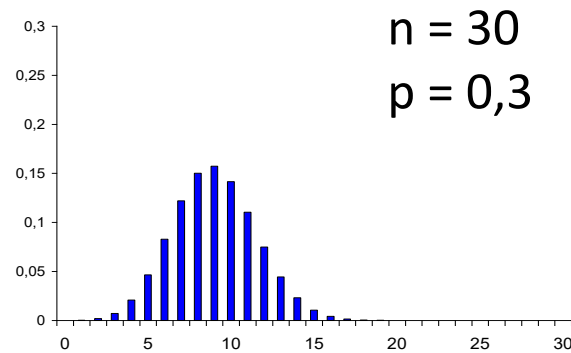
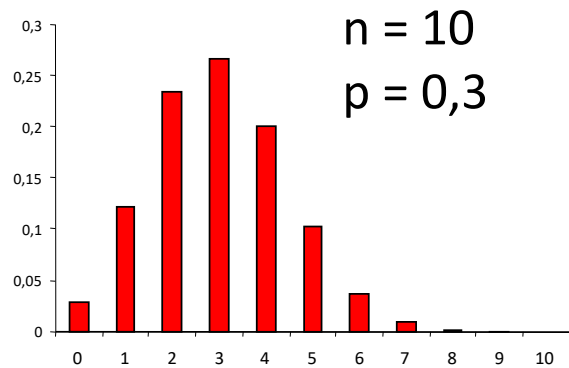
$$E(x) = n \cdot p = 20$$

je střed rozložení
a nejpravděpodobnější
hodnota

Binomické rozdělení jako model

$$P(x = r) = \frac{n!}{r!(n-r)!} \cdot p^r \cdot q^{(n-r)}$$

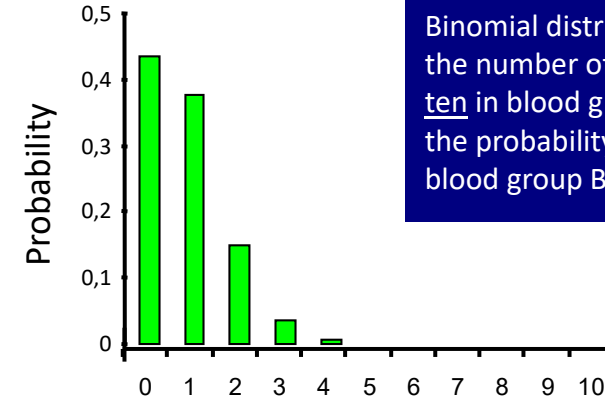
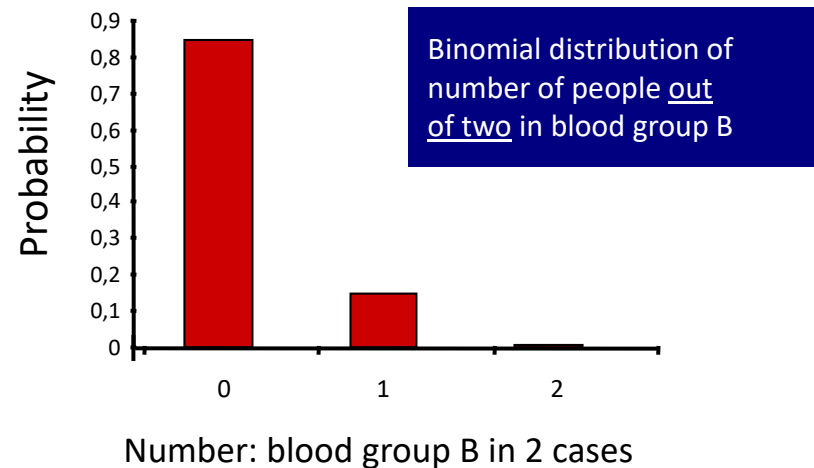
$$q = 1 - p$$



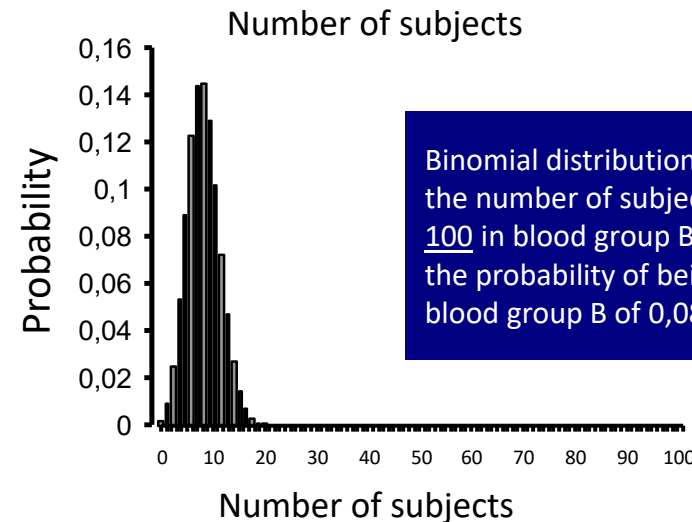
Aplikace binomického rozdělení

- Výskyt krevní skupiny B v určité populaci: $p = 0,08$

		Number in blood group B	Probability
B	B	2	0,0064
not B	B	1	0,0736
B	not B	1	0,0736
not B	not B	0	0,8464



Binomial distribution showing the number of subjects out of ten in blood group B based on the probability of being in blood group B of 0,08.



Binomial distribution showing the number of subjects out of 100 in blood group B based on the probability of being in blood group B of 0,08.

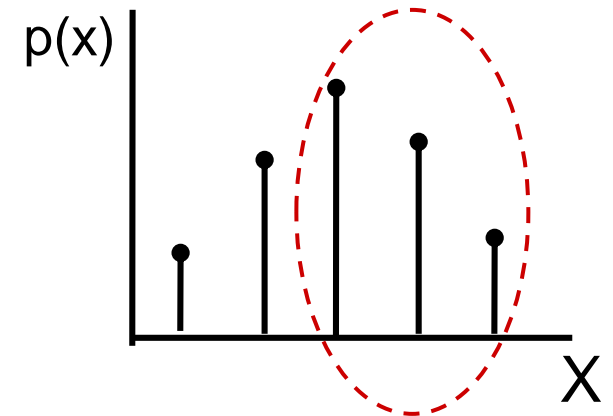
Aplikace binomického rozdělení

- Populace: 60% jedinců má zvýšenou hladinu cholesterolu; výběr: 5 lidí
- Kolik lidí má ve výběru vyšší hladinu cholesterolu ?
 - $n \cdot p = 5 \cdot 0,6 = 3$ lidé $\sim E(x)$
- Jaká je P, že právě 3 lidé budou mít vyšší hladinu cholesterolu ? \sim Tzn. Výběr přesně odpovídá dané populaci ?

$$P_{(3)} = \frac{5!}{3!(5-3)!} \cdot (0,6)^3 \cdot (0,4)^2 = 0,346 = 35\%$$

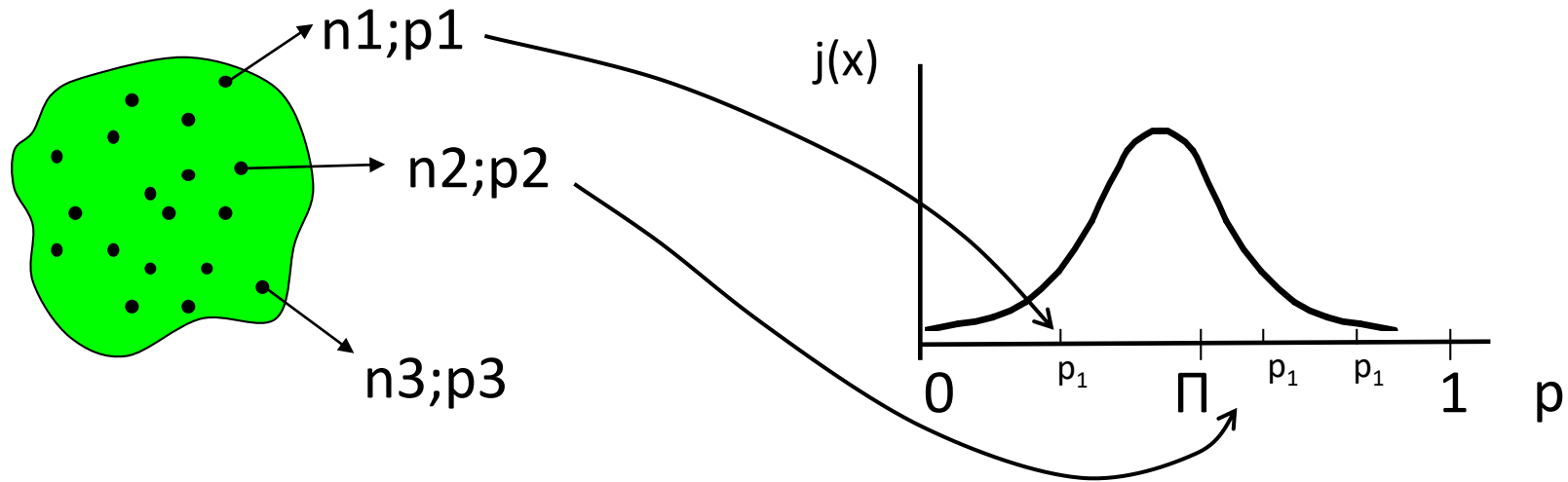
- Jaká je P, že většina jedinců (tedy minimálně 3) má vyšší hladinu cholesterolu ? \sim Tzn. výběr alespoň obecně odpovídá zkoumané populaci ?

- $P(X > 3) = P(3) + P(4) + P(5) = 0,346 + 0,259 + 0,078 = 68\%$

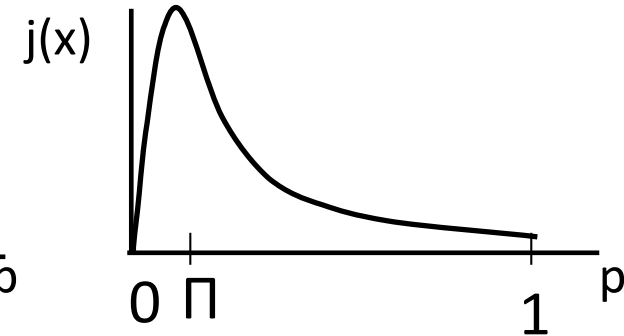
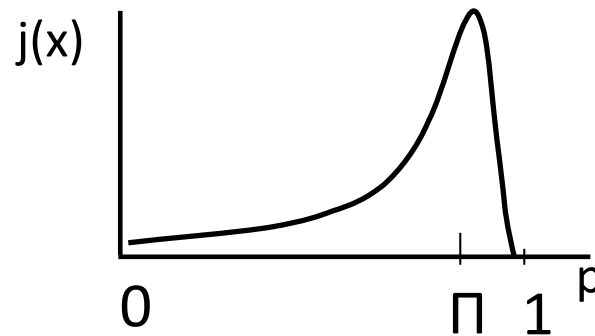


Odhad parametru Π binomického rozdělení

- Při vícenásobném odhadu se parametr Π chová jako normálně rozložen



U malých nebo velkých hodnot p (Π) je však předpoklad normality omezen



Odhad parametru Π binomického rozložení

$$\pi \approx \hat{p}; \quad \hat{p} = r/n$$

1) Bodový

$$\hat{p}; \quad s_{\hat{p}}^2 = \frac{\hat{p}(1-\hat{p})}{n-1}$$

2) Intervalový – aproximace

$$\hat{p} - Z_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n-1}} \leq \pi \leq \hat{p} + Z_{1-\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n-1}}$$

$$\pi: \hat{p} \pm Z_{1-\alpha/2} \cdot \sqrt{\frac{p(1-p)}{n-1}}$$

Odhad parametru π binomického rozložení: příklad I

X: % jedinců s daným znakem

n = 100 jedinců

r = 60; $\hat{p} = 0,6$

$s_{\hat{p}} = 0,049$

Interval spolehlivosti : 95 %

$Z_{0,975} = 1,96$

$$0,6 - 1,96 \cdot 0,049 \leq \pi \leq 0,6 + 1,96 \cdot 0,049$$

$$0,504 \leq \pi \leq 0,697$$



$$P(0,504 \leq \pi \leq 0,697) \geq 0,95$$

Odhad parametru Π binomického rozložení

- Intervalový odhad bez aproximací na normální rozložení

$$L_1 = \frac{r}{r + (n - r + 1) \cdot F_{\alpha/2}^{(v_1; v_2)}}$$



spodní limit intervalu

$$v_1 = 2(n - r + 1); \quad v_2 = 2r$$

$$L_2 = \frac{(r + 1) \cdot F_{\alpha/2}^{(v'_1; v'_2)}}{n - r + (r + 1) \cdot F_{\alpha/2}^{(v'_1; v'_2)}}$$



horní limit intervalu

$$v'_1 = 2(r + 1) = v_2 + 2$$

$$v'_2 = 2(n - r) = v_1 - 2$$

$$P(L_1 \leq \pi \leq L_2) \geq 1 - \alpha$$

Odhad parametru Π binomického rozložení: příklad II

Náhodný vzorek $n = 200$ jedinců.

Zjištěno pouze $r = 4$ jedinci bez určitého znaku.

$$\hat{p} = \frac{4}{200} = \underline{\underline{0,02}}$$

95% interval spolehlivosti = ?

Spodní hranice

$$v_1 = 2(n - r + 1) = 2(200 - 4 + 1) = 394$$

$$v_2 = 2r = 2 \cdot 4 = 8$$

$$F_{1-\alpha/2}^{(394;8)} = \underline{\underline{3,67}}$$

$$L_1 = \frac{4}{4 + (200 - 4 + 1) \cdot 3,67} = \underline{\underline{0,0055}}$$

Horní hranice

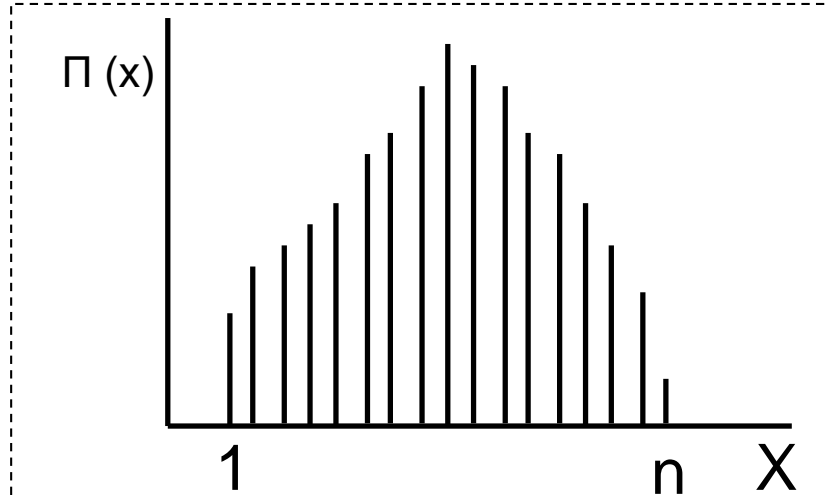
$$v'_1 = 2(r + 1) = 10$$

$$v'_2 = 2(n - r) = 2(200 - 4) = 392$$

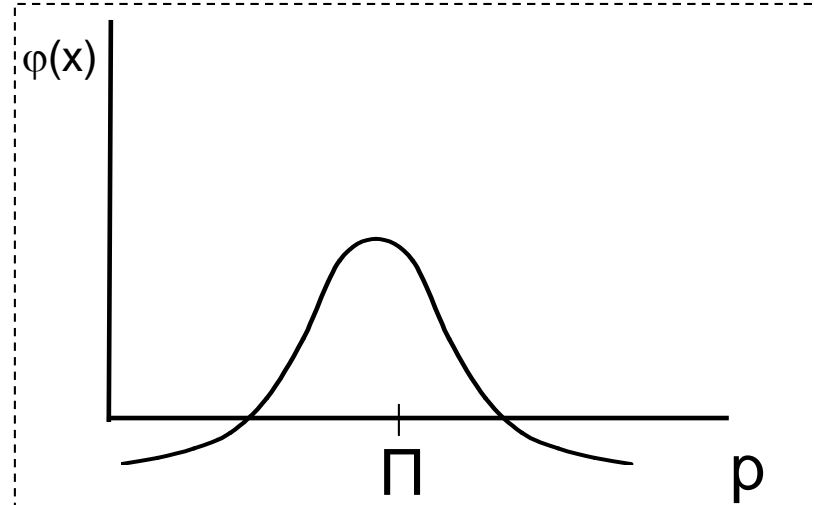
$$F_{1-\alpha/2}^{(10;392)} = \underline{\underline{2,08}}$$

$$L_2 = \frac{(4 + 1) \cdot 2,08}{200 - 4 + (4 + 1) \cdot 2,08} = \underline{\underline{0,051}}$$

Binomické rozložení v datech: vizualizace

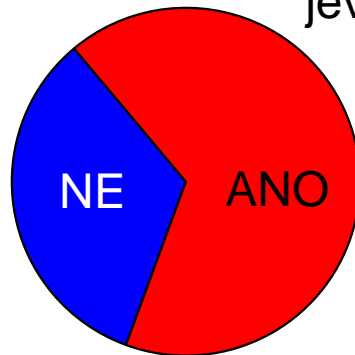


Pravděpodobnost výskytu hodnot X

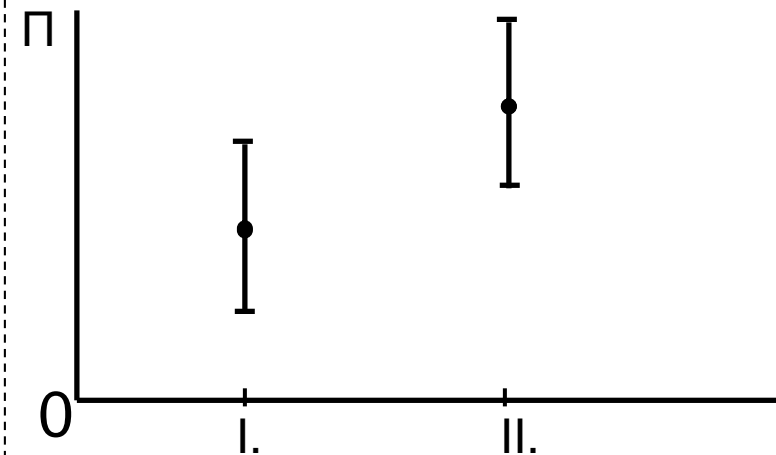


Modelové rozložení odhadovaného parametru

n opakování jev ANO
jev NE



Binární podstata původních hodnot



Interval spolehlivosti pro Π

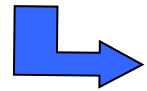
Statistické testování binomických dat

I.

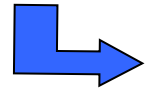
Liší se odhad \underline{p} od předpokládané hodnoty P ?

II.

Liší se dva nebo více odhadů \underline{p} ?



- závislé odhady -



- nezávislé odhady -

III.

Je výskyt kategorií dvou jevů nezávislý ?

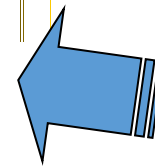
IV.

Hodnocení relativního rizika z výskytu určitého jevu v rámci skupiny lidí

Jednovýběrový binomický test

H_0	H_A	Testová statistika	Interval spolehlivosti
$p \leq \Pi$	$p > \Pi$	z	$z > z_{1-\alpha}$
$p \geq \Pi$	$p < \Pi$	z	$z < z_{\alpha}$
$p = \Pi$	$p \neq \Pi$	z	$\frac{1}{2}z\frac{1}{2} > z_{1-\alpha/2}$

$$Z = \frac{n \cdot \hat{p} - n \cdot \pi}{\sqrt{n \cdot \hat{p}(1 - \hat{p})}} \cong \frac{|n \cdot \hat{p} - n \cdot \pi| - 0,5}{\sqrt{n \cdot \hat{p}(1 - \hat{p})}}$$



Korekce na
kontinuitu

H_0	H_A	Testová statistika	Interval spolehlivosti
$p \leq \Pi$	$p > \Pi$	$L_1 = \frac{(r + 1)F_{\alpha, v_1', v_2'}}{n - r + (r + 1)F_{\alpha, v_1', v_2'}}$	$p = r / n > L_1$
$p \geq \Pi$	$p < \Pi$	$L_2 = \frac{r}{r + (n - r + 1)F_{\alpha, v_1', v_2'}}$	$p < L_2$
$p = \Pi$	$p \neq \Pi$	$L_1; L_2 (F_{\alpha/2}; F_{1-\alpha/2})$	$p < L_2 \vee p > L_1$

Test π ? p : Příklad 1

- Stromy s pozměněným tvarem koruny
 - $n = 9\ 000$ jedinců
 - $r = 2\ 250$ změněných jedinců
- Jak je pravděpodobná změna u až $1/3$ jedinců?

$$Z = \frac{n \cdot p - n \cdot \pi}{\sqrt{p(1-p) \cdot n}} = \frac{2250 - 3000}{\sqrt{0,25 \cdot 0,75 \cdot 9000}} = \underline{\underline{-18,26}}$$

$$\alpha = 5\ %; \quad Z_{1-\alpha/2} = 1,96; \quad Z_{1-\alpha} = 1,645$$

$Z > Z_{1-\alpha/2}$ zamítáme $H_0: p < 0,01$

- 95 % Interval spolehlivosti ... p : (0,241; 0,258)

Test $\pi ? p$: Příklad 2

- Pravděpodobnost narození chlapce je asi $1/2$.
- Máte zhodnotit výsledky průzkumu populace, která žije v silně poškozeném životním prostředí.
- Průzkum se týká 1000 náhodně vybraných rodin a zjištěný podíl narozených chlapců je 0.41.

Jaké jsou vaše závěry o této populaci (zda se rodí stejný podíl chlapců jako v běžné populaci?)

Jak se váš odhad zpřesní, když použijete vzorek $n = 10\,000$ rodin při zachování odhadu $p = 0.41$?

Test π ? p : Příklad 2

- Použijeme jednovýběrový binomický test s nulovou hypotézou $H_0: p=\pi$, hladina významnosti $\alpha=0,05$
- Testová statistika: _____ a příslušný kvantil: _____
- Protože _____ nulovou hypotézu ???
- Interval spolehlivosti: _____
- Pokud použijeme $n=10\ 000$, bude int. spolehlivosti ???:

Test π ? p : Příklad 2

- Použijeme jednovýběrový binomický test s nulovou hypotézou $H_0: p=\pi$, hladina významnosti $\alpha=0,05$
- Testová statistika: $Z = \frac{n \cdot \hat{p} - n \cdot \pi}{\sqrt{n \cdot \hat{p}(1 - \hat{p})}} = \frac{1000 \cdot 0,41 - 1000 \cdot 0,5}{\sqrt{1000 \cdot 0,41 \cdot 0,59}} = -5,79$ a příslušný kvantil: $Z_{1-\frac{\alpha}{2}} = Z_{0,975} = 1,96$
- Protože $|Z| > Z_{0,975}$ nulovou hypotézu zamítáme. Chlapci se ve zkoumavé populaci nerodí s pravděpodobností 0,5.
- Interval spolehlivosti: $\pi: \hat{p} \pm Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{p(1-p)}{n-1}} = 0,4 \pm Z_{0,975} \cdot 0,046 = 0,41 \pm 1,96 \cdot 0,016 = 0,41 \pm 0,03$
- Pokud použijeme $n=10\ 000$, bude int. spolehlivosti užší:

$$\pi: \hat{p} \pm Z_{1-\frac{\alpha}{2}} \cdot \sqrt{\frac{p(1-p)}{n-1}} = 0,41 \pm 1,96 \cdot 0,005 = 0,41 \pm 0,01$$

Test $\pi ? p$: Příklad 3

- Příklad testu bez aproximace na normální rozložení

12 jedinců bylo zkoumáno pro výskyt určitého znaku,
10 jedinců znak nemělo

? Jak hodně se tento výsledek liší od výsledku 6 - 6: tedy od situace, kdy polovina jedinců znak má?

a) Využití distribuční funkce

r	0	1	2	3	4	5	6	7	8	9	10	11	12
P(r)	0,0002 4	0,0029 3	0,0161 1	0,0537 1	0,1208 5	0,1933 5	0,2255 9	0,1933 6	0,1208 5	0,0537 1	0,0161 1	0,0029 3	0,0002 4

$$P(r \geq 10) = 0,01611 + 0,00393 + 0,00024 = 0,01928$$

$H_0: p = 0,5$ je tedy značně nepravděpodobná

b) Pozorované $\hat{p} = \frac{10}{12} = 0,833$ překročilo horní limit 95 % intervalu
spolehlivosti pro p :

$$p = 0,5: L_2 = \frac{(6+1) \cdot 2,64}{12 - 6 + (6+1) \cdot 2,64} = \underline{\underline{0,755}}$$

Dvouvýběrový binomický test ($p_1 \neq p_2$)

$$Z = \frac{|\hat{p}_1 - \hat{p}_2|}{\sqrt{\frac{\bar{p}(1-\bar{p})}{n_1} + \frac{\bar{p}(1-\bar{p})}{n_2}}}$$

$$\bar{p} = \frac{n_1 \cdot \bar{p}_1 + n_2 \cdot \bar{p}_2}{n_1 + n_2}$$

$$(\hat{p}_1 - \hat{p}_2) \pm Z_{(1-\alpha/2)} \cdot \sqrt{\frac{\bar{p}(1-\bar{p})}{n_1} + \frac{\bar{p}(1-\bar{p})}{n_2}}$$

Dvouvýběrový binomický test ($p_1 \neq p_2$)

- Tento příklad je ukázkou testování rozdílů mezi dvěma binomickými populacemi (tedy srovnání dvou odhadů parametru p).
- Celkem 49 pokusných myší bylo použito k testování toxického preparátu během dvouměsíční kultivace. Následující tabulka obsahuje původní data zároveň s testem nulové hypotézy: Podíl přežívajících jedinců je u zasažené populace stejný.

	Alive	Dead	Total	Proportion alive	Proportion dead
Treated	15	9	24	$\hat{p}_1 = 0,625$	$\hat{q}_1 = 0,375$
Not Treated	10	15	25	$\hat{p}_2 = 0,400$	$\hat{q}_2 = 0,600$
Total	25	24	49	$\hat{p} = 0,510$	$\hat{q} = 0,490$

$$Z = \frac{0,625 - 0,400}{\sqrt{\frac{(0,510)(0,490)}{24} + \frac{(0,510)(0,490)}{25}}} = \frac{0,225}{\sqrt{0,010413 + 0,009996}} = 1,573$$

$$Z_{0,05(2)} = t_{0,05(2)} = 1,96$$

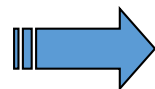


Nezamítáme $H_0: 0,10 < P < 0,20$

S korekcí na kontinuitu:

$$Z = \frac{\frac{15 - 0,5}{24} - \frac{10 + 0,5}{25}}{0,143} = \frac{0,604 - 0,420}{0,143} = 1,287$$

$$Z_{0,05(2)} = t_{0,05(2)} = 1,96$$



Nezamítáme $H_0: 0,10 < P < 0,20$

Přednáška 8

Kontingenční tabulky

Test dobré shody

Fisherův přesný test

McNemar test

Odds ratio a relativní riziko

Anotace

- Analýza kontingenčních tabulek umožňuje analyzovat vazbu mezi dvěma kategoriálními proměnnými.
- Základním způsobem testování je tzv. chi-square test, který srovnává pozorované četnosti kombinací kategorií oproti očekávaným četnostem, které vychází z teoretické situace, kdy je vztah mezi proměnnými náhodný.
- Test dobré shody je využíván také pro srovnání pozorovaných četností proti očekávaným četnostem daným určitým pravidlem (typickým příkladem je Hardy-Weinbergova rovnováha v genetice)
- Specifickým typem výstupů odvozených z kontingenčních tabulek jsou tzv. odds ratio a relativní rizika, využívaná často v medicíně pro identifikaci a popis rizikových skupin pacientů.

Co je kontingenční tabulka ?

- Frekvenční sumarizace dvou kategoriálních proměnných (binárních, nominálních nebo ordinálních proměnných).
- Obecně: R x C kontingenční tabulka (R – počet kategorií jedné proměnné, C – počet kategorií druhé proměnné).
- Speciální případ: 2 x 2 tabulka = čtyřpolní tabulka.
- Kontingenční tabulky: absolutních četností, celkových procent, řádkových/sloupcových četností
- Příklad: Sumarizace vyšetřených osob podle pohlaví a výsledku diagnostického testu.

Pohlaví	Výsledek vyšetření		Celkem
	Nemocný	Zdravý	
Muž	45	11	56
Žena	25	6	31
Celkem	70	17	87

Ukázka kontingenční tabulky

- Vztah pohlaví a výskytu onemocnění (pozor na hodnocení nesmyslného vztahu)

	Nemocný	Zdravý	Celkem
Muž	a	b	a + b
Žena	c	d	c + d
Celkem	a + c	b + d	a + b + c + d = N

Marginální absolutní četnost

Celkový počet hodnot

Simultánní absolutní četnost



	Nemocný	Zdravý	Celkem
Muž	45	11	56
Žena	25	6	31
Celkem	70	17	87



Jsou více nemocní muži nebo ženy?

Test dobré shody - základní teorie

Testová statistika:

$$\chi^2 = \sum \frac{\left[\begin{array}{c} \text{pozorovaná} - \text{očekávaná} \\ \text{četnost} \quad \text{četnost} \end{array} \right]^2}{\text{očekávaná četnost}}$$

$$\chi^2 = \underbrace{\frac{\left[\begin{array}{c} \text{pozorovaná} - \text{očekávaná} \\ \text{četnost} \quad \text{četnost} \end{array} \right]^2}{\text{očekávaná četnost}}}_{\text{1. jev}} + \underbrace{\frac{\left[\begin{array}{c} \text{pozorovaná} - \text{očekávaná} \\ \text{četnost} \quad \text{četnost} \end{array} \right]^2}{\text{očekávaná četnost}}}_{\text{2. jev}} + \dots$$

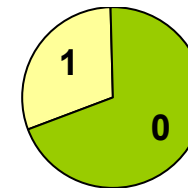
$$\chi^2 > \chi^2_{(1-\alpha)} \text{ (s.v.)} \quad \dots \text{ zamítáme } H_0$$

1 - hladina významnosti

stupně volnosti

Test dobré shody: příklad

Binomické jevy (1/0)



$$\chi^2_{(1)} = \frac{\left[\frac{\text{pozorovaná četnost} - \text{očekávaná četnost}}{\text{očekávaná četnost}} \right]^2}{\text{I. jev 1}} + \frac{\left[\frac{\text{pozorovaná četnost} - \text{očekávaná četnost}}{\text{očekávaná četnost}} \right]^2}{\text{II. jev 2}}$$

Příklad



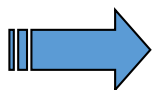
10 000 lidí hází mincí → rub: 4 000 případů (R)
→ líc: 6 000 případů (L)



Lze výsledek považovat za statisticky významně odlišný (nebo neodlišný) od očekávaného poměru R : L = 1 : 1 (tzn. že je výsledek hodů mincí náhodný)?

$$\chi^2 = \frac{(4000 - 5000)^2}{5000} + \frac{(6000 - 5000)^2}{5000} = 400$$

Tabulková hodnota: $\chi^2_{(0,95)} (v = k - 1 = 1) = \underline{3,84}$ ($0,95 = 1 - \alpha$)

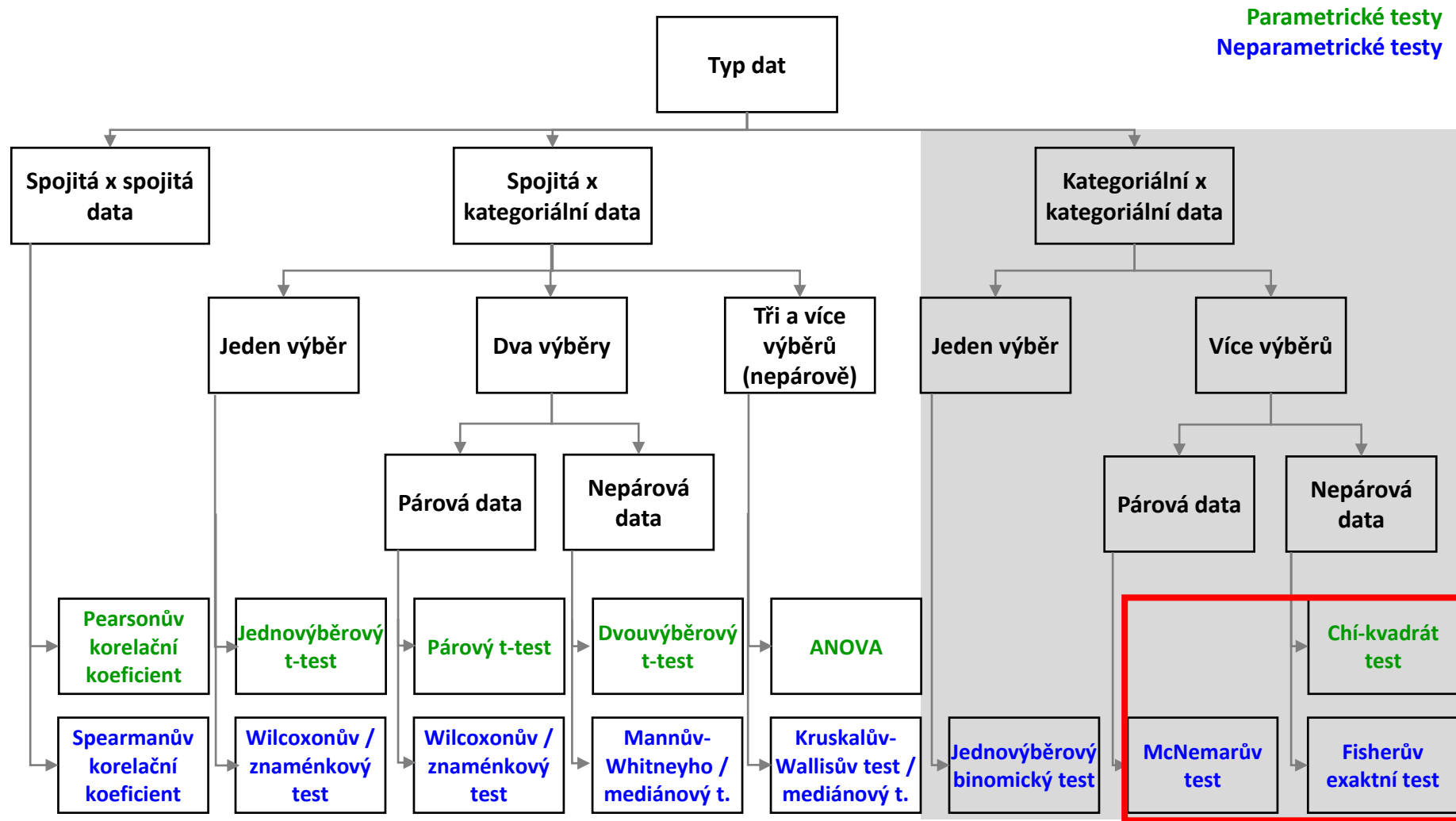


Rozdíl je vysoce statisticky významný ($p < 0,001$)

Kontingenční tabulka - hypotézy

- **NEZÁVISLOST** (Pearsonův chí-kvadrát test, Fisherův exaktní test)
 - Jeden výběr, 2 charakteristiky – obdoba nepárového uspořádání
 - Např.: existence vztahu mezi barvou očí a známkou z biostatistiky u studentů
- **SHODA STRUKTURY** (Pearsonův chí-kvadrát test, Fisherův exaktní test)
 - Tzv. test homogeneity
 - Více výběrů, jedna charakteristika – obdoba nepárového uspořádání
 - Např.: věková struktura pacientů s diabetem v K nemocnicích (tj. K výběrů)
- **SYMETRIE** (McNemarův test)
 - Jeden výběr, opakovaně jedna charakteristika – obdoba párového uspořádání
 - Např.: posouzení stavu stromů ve dvou sezónách

Základní rozhodování o výběru statistických testů - analýza kontingenčních tabulek



Kontingenční tabulka - obecně

- Máme dvě nominální veličiny, X (má r variant) a Y (má s variant)
- Kontingenční tabulka typu $r \times s$

$x_{[j]}$ \ $y_{[k]}$	$y_{[1]}$	$y_{[s]}$	$n_{j.}$
$x_{[1]}$	n_{11}	n_{1s}	$n_{1.}$
.
.
$x_{[r]}$	n_{r1}	n_{rs}	$n_{r.}$
$n_{.k}$	$n_{.1}$.	.	$n_{.s}$	n

Marginální absolutní četnost

Marginální absolutní četnost

Simultánní absolutní četnost

- Označení:
 - n_{jk} - simultánní absolutní četnost,
 - $n_{j.}$ - marginální absolutní četnost

Kontingenční tabulky H0 :Nezávislost dvou jevů A a B

**Kontingenční
tabulka
2 x 2**

↘ ↓ B \ A →	+	-	Podíl (+)
+	a	b	$\frac{a}{(a+b)}$ p₁
-	c	d	$\frac{c}{(c+d)}$ p₂
Podíl (+)	$\frac{a}{(a+c)}$	$\frac{b}{(b+d)}$	

$$N = a + b + c + d$$

$$P(B^+) = \frac{(a+b)}{N}$$

$$P(B^-) = \frac{(c+d)}{N}$$

Očekávané četnosti:

$$F_{(A)} = \frac{(a+b)(a+c)}{N}$$

$$F_{(C)} = \frac{(a+c)(d+c)}{N}$$

$$\chi^2_{\nu=1} = \sum_{i=1}^4 \frac{(f_i - F_i)^2}{F_i}$$

$$F_{(B)} = \frac{(a+b)(b+d)}{N}$$

$$F_{(D)} = \frac{(b+d)(c+d)}{N}$$

$$\nu = 1 = (r-1) * (c-1)$$

$$P_{(A)}; P_{(B)}$$

$$\chi^2_c = \sum \sum \frac{(|f_{ij} - F_{ij}| - 0,5)^2}{F_{ij}}$$

Očekávané četnosti v kontingenční tabulce

- Očekávané četnosti pro výpočet testu dobré shody v kontingenční tabulce odpovídají tabulce, která nemá žádný vztah mezi řádky a sloupce (náhodný vztah řádků a sloupců)

$$\chi^2_{(1)} = \frac{\left[\begin{array}{cc} \text{pozorovaná} & \text{očekávaná} \\ \text{četnost} & \text{četnost} \end{array} - \right]^2}{\text{očekávaná četnost}}$$

Počítáno pro každou buňku tabulky

	☠	😊
A	10	0
B	0	10

Pozorovaná tabulka

	☠	😊
A	5	5
B	5	5

Očekávaná tabulka

Testování nezávislosti – Pearsonův chí-kvadrát test

- Souvisí spolu výskyt dvou nominálních znaků měřených na jediném výběru?
- Příklad: Barva očí (modrá, zelená, hnědá) a barva vlasů (hnědá, černá, blond) u vybraných 30 studentů jsou nezávislé.
- **Nulová hypotéza**: Znaky X a Y jsou nezávislé náhodné veličiny.
- **Alternativní hypotéza**: Znaky X a Y jsou závislé náhodné veličiny.
- Test: **Pearsonův chí-kvadrát**

$$K = \sum_{j=1}^r \sum_{k=1}^s \frac{(n_{jk} - e_{jk})^2}{e_{jk}} \stackrel{\text{H}_0 \text{ platí}}{\approx} \chi^2((r-1)(s-1))$$

- Očekávané (teoretické) četnosti e_{jk} : $e_{jk} = \frac{n_{j.} \cdot n_{.k}}{n}$
- H_0 zamítáme na hladině významnosti α , pokud $K \geq \chi_{1-\alpha}^2((r-1)(s-1))$

- **Předpoklady testu ?**

Testování nezávislosti – Pearsonův chí-kvadrát test

Předpoklady Pearsonova chí-kvadrát testu:

- **Jednotlivá pozorování** shrnutá v kontingenční tabulce jsou **nezávislá**, tj. každý prvek patří jen do jedné buňky kont. tabulky, nemůže zároveň patřit do dvou.
- **Podmínky dobré aproximace:** Očekávané (teoretické) četnosti jsou aspoň v 80 % případů větší nebo rovné 5 a ve 100 % případů nesmí být pod 2 (pokud není tento předpoklad splněn, je vhodné sloučit kategorie s nízkými četnostmi).

- **Měření síly závislosti:**

$$V = \sqrt{\frac{K}{n(m-1)}}, \text{ kde } m = \min\{r, s\}, V \text{ je z intervalu } (0,1)$$

Cramérův koeficient:

Význam hodnot: 0-0,1....zanedbatelná závislost

0,1-0,3...slabá závislost

0,3-0,7...střední závislost

0,7-1 silná závislost

Kontingenční tabulky: příklad

gen \ †	Ano	Ne	Σ
Ano	20	82	102
Ne	10	54	64
Σ	30	136	166

$$F_A = 102 * 30 / 166 = 18,43$$

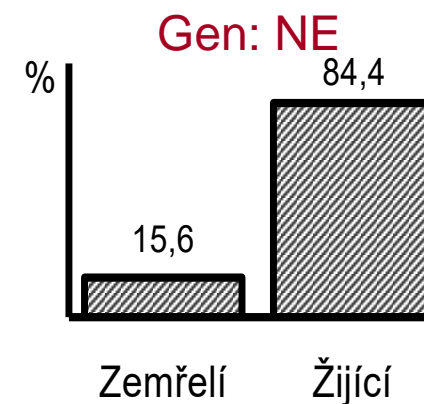
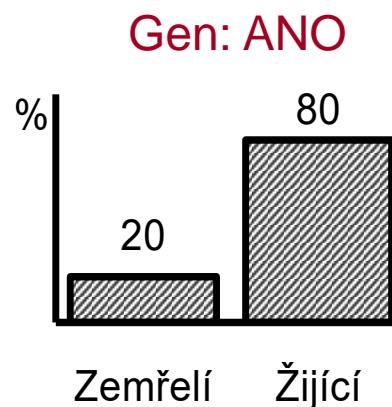
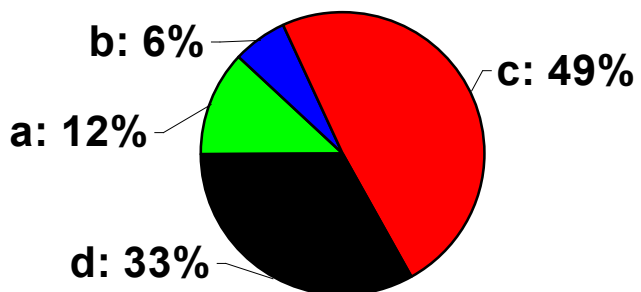
$$F_B = 102 * 136 / 166 = 83,57$$

$$F_C = 11,57$$

$$F_D = 52,43$$

$$\chi^2_{(1)} = \frac{(20 - 18,43)^2}{18,43} + \frac{(82 - 83,57)^2}{83,57} + \frac{(10 - 11,57)^2}{11,57} + \frac{(54 - 52,43)^2}{52,43} = 0,423 \quad 0,423 < \chi^2_{0,95}^{(1)} = 3,84$$

Kontingenční tabulka v obrázku



Výstup řešení v SW

Tab.1: Pozorované četnosti

Summary Frequency Table (07 prikklad_z_prednasky_K
Marked cells have counts > 10
(Marginal summaries are not marked)

Gen	Stav_pacienta úmrtí	Stav_pacienta žijící	Row Totals
přítomen	20	82	102
nepřítomen	10	54	64
All Grps	30	136	166

Tab. 2: Očekávané četnosti

Summary Table: Expected Frequencies (07 prikklad_z_pre
Marked cells have counts > 10
Pearson Chi-square: ,421322, df=1, p=,516278

Gen	Stav_pacienta úmrtí	Stav_pacienta žijící	Row Totals
přítomen	18,43373	83,5663	102,0000
nepřítomen	11,56627	52,4337	64,0000
All Grps	30,00000	136,0000	166,0000



Jsou splněny podmínky dobré aproximace?

Tab. 3: Paersonův chí-kvadrát

Hodnota testové statistiky Počet stupňů volnosti p- hodnota

Statistic	Chi-square	df	p
Pearson Chi-square	,4213223	df=1	p=,51628
M-L Chi-square	,4277117	df=1	p=,51311
Phi for 2 x 2 tables	,0503794		
Tetrachoric correlation	,0949754		
Contingency coefficient	,0503156		

R x C kontingenční tabulka

Výběr: N lidí ze sociologického průzkumu (delikventi)

Jev A: Původ z rozvrácených rodin

Jev B: Stupeň zločinnosti I < II < III < IV

A \ B	I.	II.	III.	IV.	Σ
ANO	a	b	c	d	číslo 1
NE	e	f	g	h	
Σ	číslo2				

Stupně volnosti:

$$(R-1) * (C-1) = 1 * 3 = 3$$

$$F_a = \frac{\text{číslo 1} \cdot \text{číslo 2}}{N}$$

Tabulky: $\chi^2_{(1-\alpha)}^{(v)}$

Očekávané četnosti:

$$p_a = \frac{a}{a + e}$$

$$p_b = \frac{b}{b + f}$$

$$p_c = \frac{c}{c + g}$$

$$p_d = \frac{d}{d + h}$$

Rekódování kategoriálních proměnných na binární

Původní NYHA	Dummies				Vzhledem k referenci		
	NYHA I	NYHA II	NYHA III	NYHA IV	NYHA II ref	NYHA III ref	NYHA IV ref
I	1	0	0	0	0	0	0
I	1	0	0	0	0	0	0
I	1	0	0	0	0	0	0
II	0	1	0	0	1		
II	0	1	0	0	1		
III	0	0	0	0		1	
III	0	0	0	0		1	
IV	0	0	1	1			1
IV	0	0	1	1			1

Rekódování kategoriálních proměnných na binární

- Kategoriální a ordinální data mohou do analýzy vstupovat jako binární proměnné
- Kategoriální data (nelze seřadit) -> dummies
- Ordinální data (lze seřadit)
 - Dummies
 - Definice referenční kategorie (obvykle kategorie s nejnižším rizikem pro hodnocený endpoint)
- Příklad: The New York Heart Association (NYHA) Functional Classification

Původní	Dummies				Vzhledem k referenci		
	NYHA	NYHA I	NYHA II	NYHA III	NYHA IV	NYHA II ref	NYHA III ref
I	1	0	0	0	0	0	0
I	1	0	0	0	0	0	0
I	1	0	0	0	0	0	0
II	0	1	0	0	1		
II	0	1	0	0	1		
III	0	0	0	0		1	
III	0	0	0	0		1	
IV	0	0	1	1			1
IV	0	0	1	1			1

Test dobré shody: příklad I



Ověřte na datech z pokusu se 100 květinami určitého druhu, že barva květů se geneticky štěpí v poměru žlutá : červená = 3 : 1.



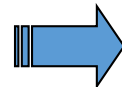
H_0 : Pozorovaná frekvence pro jednotlivé barvy květů jsou vzorkem populace mající poměr mezi žlutými a červenými květy 3 : 1.

Součet frekvencí u obou barev květů (f_i) se rovná 100 a pozorované frekvence u kategorií barvy budou srovnány s očekávanými frekvencemi (uvedeny v závorkách):

	Kategorie barvy		
	Žlutá	Červená	n
$f_{\text{poz.}}$	84	16	100
$f_{\text{oček.}}$	75	25	

$$\chi^2 = \sum \frac{(f_{\text{poz.}} - f_{\text{oč.}})^2}{f_{\text{oč.}}} = \frac{(84 - 75)^2}{75} + \frac{(16 - 25)^2}{25} = 4,320$$

St. volnosti = $n = k - 1 = 1$



Zamítáme hypotézu shody srovnávaných četností

Při testování H_0 jsme použili matematický zápis ($0,025 < P < 0,05$). Z tabulek χ^2 rozložení vidíme, že pravděpodobnost překročení hranice 2,706 je 0,1 (10 %), což může být stručně zapsáno jako $P(\chi^2 \geq 2,706) = 0,10$.

Dále lze zjistit pro $P(\chi^2 \geq 3,841) = 0,05$. V řešené úloze jsme dospěli k hodnotě testové statistiky $\chi^2 = 4,320$. Pro tento případ lze tedy psát $0,025 < P(\chi^2 \geq 4,320) < 0,05$; a jednodušeji $0,025 < P < 0,05$. Jde v podstatě o přibližné určení hranic chyby 1. druhu.

Test dobré shody: příklad II

Tento příklad je rozšířením problému z příkladu 1 na srovnání pozorovaných a očekávaných frekvencí pro více kategorií sledovaného znaku:

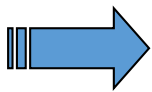


Celkem bylo zkoumáno 250 semen určitého druhu rostliny a roztríděno do následujících kategorií: žluté/hladké; žluté/vrásčité; zelené/hladké; zelené/vrásčité. Předpokládaný poměr výskytu těchto kategorií v populaci je 9 : 3 : 3 : 1. Následující tabulka obsahuje původní data z pozorování a dále postup při testování H_0 .

	žluté/hladké	žluté/vrásčité	zelené/hladké	zelené/vrásčité	n
$f_{\text{poz.}}$	152	39	53	6	250
$f_{\text{oček.}}$	140,6250	46,8750	46,8750	15,6250	

$$v = k - 1 = 3$$

$$\chi^2 = \frac{11,3750^2}{140,6250} + \frac{7,8750^2}{46,8750} + \frac{6,1250^2}{46,8750} + \frac{9,6250^2}{15,6250} = 8,972$$



Zamítáme hypotézu shody pozorovaných četností s očekávanými

Test dobré shody: příklad III

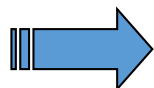
Složitější příklady řešené srovnáváním frekvencí je možné rozdělit na testování dílčích hypotéz:

- ✓ Předpokládejme, že chceme pro data z předchozí úlohy testovat hypotézu existence štěpného poměru 9 : 3 : 3 pro první tři kategorie semen:

	žluté/hladké	žluté/vrásčité	zelené/hladké	n
$f_{\text{poz.}}$	152	39	53	244
$f_{\text{oček.}}$	146,400	48,800	48,800	

$n = k - 1 = 2$

$$\chi^2 = \frac{5,600^2}{146,40} + \frac{9,800^2}{48,80} + \frac{4,200^2}{48,80} = 2,544$$



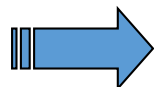
Nezamítáme hypotézu shody pozorovaných četností s očekávanými.

- ✓ Nyní otestujeme hypotézu štěpného poměru kategorií zelené/vrásčité:ostatní typy = 1:15

	zelené/vrásčité	ostatní	n
$f_{\text{poz.}}$	6	244	25
$f_{\text{oček.}}$	15,625	234,375	

$n = k - 1 = 1$

$$\chi^2 = \frac{9,625^2}{15,625} + \frac{9,625^2}{234,375} = 6,324$$



Zamítáme hypotézu shody pozorovaných četností s očekávanými.

Testování homogenity (shody struktury)

- Motivace: Zajímá nás výskyt nominálního znaku u r nezávislých výběrů z r různých populací.
- Příklad: Je zájem o sport stejný u děvčat jako u chlapců?
- Nulová hypotéza: pravděpodobnostní rozdělení kategoriální proměnné je stejné v různých populací
- Test: **Pearsonův chí-kvadrát**

		Dívky	Chlapci	
Zájem o sport	Ano	a	b	$a+b$
	Ne	c	d	$c+d$
		$a+c$	$b+d$	n

Některé marginální četnosti (buď sloupcové nebo řádkové) jsou předem pevně stanoveny

Test homogenity binomických rozložení



Jev: Úmrtnost na leukemii

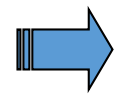
Předpoklad: $\Pi = 0,6$

Absolutní četnost jevu označena r_i

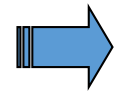
$$\bar{p} = \frac{\sum p_i}{S}$$

Sledovalo s autorů z s zemí:

Autor	n_i	r_i	p_i
1			
2			
⋮			
⋮			
⋮			
s	$\sum n_i = N$		



Test homogenity binomických rozložení



Po možném sloučení s výběrů

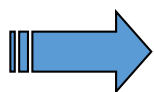
Test shody reálného r ($\sum r_i$) a $n \cdot \Pi$

$$\chi^2_{S-1} = \frac{(\sum r_i p_i - \bar{p} \sum r_i)}{\bar{p} (1 - \bar{p})}$$

$$\chi^2_{(1)} = \frac{\left(\left| \sum r_i - N \cdot \Pi \right| - \frac{1}{2} \right)^2}{N \cdot \Pi \cdot (1 - \Pi)}$$

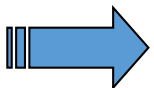
Test homogenity binomických četností: příklad

Pomocí χ^2 rozložení lze rovněž posuzovat homogenitu většího množství nezávislých pokusů testujících tutéž hypotézu.



Bylo provedeno 6 nezávislých výběrů z populace mladých mužů, kteří v dětství onemocněli těžkým zánětem mozkových blan.

H_0 : V této populaci se vyskytují praváci a leváci v poměru 1 : 1.



Nalezněte v literatuře příslušné vztahy pro testování homogenity všech šesti výběrových populací a na základě výsledků tohoto testu rozhodněte o dalším postupu.

Následující tabulka obsahuje původní data a výsledek testování (v závorkách jsou uvedeny očekávané četnosti):

Vzorek	Praváci	Leváci	n	χ^2	St. volnosti
1	3 (7)	11 (7)	14	4,5714	1
2	4 (8)	12 (8)	16	4,000	1
3	15 (10)	5 (10)	20	5,000	1
4	14 (9)	14 (9)	18	5,5556	1
5	13 (8,5)	4 (8,5)	17	4,7647	1
6	17 (11)	5 (11)	22	6,5455	1

$$\chi^2_{heterogenita} = 30,2$$

$$v = s - 1 = 5$$

$$P < 0,001$$

Jednoduchým testováním lze zjistit, že všechny testy pro jednotlivé výběry jsou významné, což znamená, že ani v jednom případě nebyla potvrzena shoda očekávaných a pozorovaných četností. Test homogenity štěpného poměru v zkoumaných populacích rovněž vedl k zamítnutí možnosti sloučit jednotlivé výběry a posuzovat je jako celek (kromě testovaného poměru 1 : 1 neexistuje tedy v datech žádný jiný jednotný štěpný poměr mezi oběma vlastnostmi).

V případě, že by tento test neprokázal odchylky mezi jednotlivými výběrovými populacemi, bylo by možné jednotlivé odběry sloučit a posuzovat jako homogenní vzorek.

χ^2 test - příklad frakcionace složitější kontingenční tabulky I

Cílem rozsáhlejšího průzkumu populace bylo prozkoumat vztah mezi dvěma typy chorob a krevními skupinami u lidí. Konkrétní data jsou uvedena v tabulce:

Krevní skupina	Žaludeční vředy	Rakovina žaludku	Kontrola	Celkem
0	983	383	2892	4258
A	679	416	2625	3720
B	134	84	570	788
Celkem	1796	883	6087	8766

Vypočítejte testovou charakteristiku pro tuto kontingenční tabulku a otestujte nulovou hypotézu nezávislosti jevů ($\chi^2 = 40,54$; 4 st. volnosti)

χ^2 test - příklad frakcionace složitější kontingenční tabulky II

K podrobnějšímu průzkumu složitějších tabulek výrazně napomáhá prepis původní tabulky do podoby procentického zastoupení kategorií:

Krevní skupina	Žaludeční vředy	Rakovina žaludku	Kontrola
0	983	383	2892
A	679	416	2625
B	134	84	570
Celkem	1796	883	6087

Z této tabulky je patrné:

- 1.** Jsou jenom malé rozdíly v distribuci krevních skupin u kontroly a u skupiny nemocných rakovinou žaludku.
- 2.** Pacienti s vředy mají mnohem častěji krevní skupinu 0.

Na základě těchto poznatků je možné sestavit menší kontingenční tabulku, která otestuje hypotézu o shodné distribuci krevních skupin pro nemocné rakovinou a pro zdravé lidi.

Sestavte tuto tabulku a otestujte nulovou hypotézu.

($\chi^2 = 5,64$ (2 st. v.), P je přibližně rovna 0,06)

χ^2 test - příklad frakcionace složitější kontingenční tabulky III

- Z tohoto dílčího testu vyplývá možnost sloučení skupiny nemocných rakovinou a zdravých lidí neboť se vzhledem k distribuci krevních skupin chovají jako homogenní populace.
- Dalším logickým krokem v podrobné analýze je testování shody relativních četností výskytu krevních skupin A a B mezi kombinovaným vzorkem (sloučená skupina s rakovinou a kontrola) a mezi vzorkem lidí nemocných žaludečními vředy - tzn. nyní neuvažujeme krevní skupinu 0. Výsledkem tohoto testu je $\chi^2 = 0,68$ (1 st. vol.); $P > 0,7$. Vzorky pro krevní skupiny A a B lze tedy sloučit do směsného vzorku A + B.
- Nyní otestujeme shodu relativních četností výskytu skupiny 0 oproti A + B, a to mezi kombinovanou populací (kontrola + nemocní rakovinou) a mezi vzorkem nemocných vředařů ($c^2 = 34,29$; 1 st. vol.).
- Lze tedy shrnout, že vysoká hodnota původního c^2 se 4 st. volnosti byla způsobena zvýšenou četností lidí s krevní skupinou 0 mezi nemocnými žaludečními vředy.

χ^2 test - příklad frakcionace složitější kontingenční tabulky IV

Průběh hodnocení lze shrnout do tabulky:

Srovnání	St. volnosti	χ^2
0, A, B skupina u pacientů s rakovinou (r) x kontrola (k)	2	5,64
A, B skupina u pacientů s vředy x kombinovaný vzorek (r + k)	1	0,68
0, A, B skupina u pacientů s s vředy x kombinovaný vzorek (r + k)	1	34,29
Celkem	4	40,61

Celkový součet testových statistik χ^2 (40,61) odpovídá přibližně původní hodnotě χ^2 (40,54). Což platí i o stupních volnosti (4). Tato skutečnost potvrzuje, že jsme detailním rozbořem vyčerpali informační obsah původní kontingenční tabulky a kromě popsané závislosti (zvýšený výskyt krevní skupiny 0 u lidí s žaludečními vředy) jsou jednotlivé kategorie zkoumaných jevů zcela nezávislé.

Kontingenční tabulka 2 x 2: Řešení při nedostatečné velikosti vzorku

Yates' corection

Fisher's exact test



H_0 : Nezávislost jevů

Test analyzuje všechny možné 2 x 2 tabulky, které dávají stejnou sumu řádků a sloupců jako tabulka zdrojová.

Algoritmus každé tabulce přiřazuje pravděpodobnost, že taková situace nastane, je-li H_0 pravdivá.

Spectacle wearing among juvenile delinquents and non-delinquents who failed a vision test (Weindling et al., 1986)

		Juvenile delinquents	Non- delinquents	Total
Spectacle wearers	Yes	1	5	6
	No	8	2	10
	Total	9	7	16

Kontingenční tabulka 2 x 2: Řešení při nedostatečné velikosti vzorku

Všechny možné varianty tabulky s danou sumou řádků a sloupců

(I)	0	6	(V)	4	2
	9	1		5	5
(II)	1	5	(VI)	5	1
	8	2		4	6
(III)	2	4	(VII)	6	0
	7	3		3	7
(IV)	3	3			
	6	4			

Pravděpodobnost náhodného vzniku variant tabulky

	a	b	c	d	P
(I)	0	6	9	1	0,00087
(II)	1	5	8	2	0,02360
(III)	2	4	7	3	0,15734
(IV)	3	3	6	4	0,36713
(V)	4	2	5	5	0,33042
(VI)	5	1	4	6	0,11014
(VII)	6	0	3	7	0,01049
Total					0,99999

Fisherův exaktní test

- Využití ve čtyřpolní tabulce (v současnosti i větší díky vyššímu výkonu počítačů) s nízkými četnostmi, které znemožňují použití Pearsonova chí-kvadrát testu.
- Patří mezi neparametrické testy pracující s daty na nominální škále, v nejjednodušší podobě ve dvou třídách: pozitivní/negativní, úspěch/neúspěch apod.
- Nulová hypotéza předpokládá rovnoměrné zastoupení sledovaného znaku u dvou nezávislých souborů.
- Slovo exaktní (přímý) znamená, že se přímo vypočítává pravděpodobnost odmítnutí, resp. platnosti nulové hypotézy.

Fisherův exaktní test

- Výpočet „přesné“ p-hodnoty, která zde hraje roli testové statistiky:
 - spočítá se parciální pravděpodobnost čtyřpolní tabulky p_1 :

Sledovaný jev	Skupina		Celkem
	Experimentální	Kontrolní	
Ano	a	b	$a + b$
Ne	c	d	$c + d$
Celkem	$a + c$	$b + d$	n

$$p_1 = \frac{(a+b)! * (c+d)! * (a+c)! * (b+d)!}{N! * a! * b! * c! * d!}$$

- Spočítá se pa všech možných tabulek při zachování marginálních četností (řádkové a sloupcové součty) a výsledná p-hodnota je součtem pa menších nebo stejných jako p_1 , která přísluší pozorované tabulce.

Test hypotézy o symetrii (McNemarův test pro čtyřpolní tabulku)

- Motivace: Na osobách sledujeme binární proměnnou před pokusem a po něm, cílem je zjistit, zda došlo ke změně v rozdělení této proměnné.
- **Analýza párových dichotomických proměnných**

Četnostní tabulka

		po		n _{j.}
		+	-	
před	+	a	b	a+b
	-	c	d	c+d
		n _k	a+c b+d	n

Tabulka teoretických pravděpodobností

		po		
		+	-	
před	+	p_{11}	p_{12}	$p_{1.}$
	-	p_{21}	p_{22}	$p_{2.}$
		$p_{.1}$	$p_{.2}$	

- Nulová hypotéza: $p_{ij} = p_{ji}$, pokus nemá vliv na výskyt daného znaku

- Testová statistika:
$$\chi^2 = \frac{(|b - c| - 1)^2}{b + c}$$
 pokud je větší než kritická hodnota χ^2 rozdělení o jednom stupni volnosti (vhodné pro počty údajů $b+c > 8$), pak nulovou hypotézu zamítáme

McNemarův test: příklad I

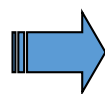
- Zjistěte, zda úspěch našich sportovců na Olympiádě nebo ve Světovém poháru vede ke změně postojů žáků ke sportování.
- Nulová hypotéza: Počet žáků, kteří změní svůj postoj pozitivním směrem, je pouze náhodně odlišný od počtu žáků, kteří změní svůj postoj negativním směrem.

		Postoj po výuce		
		+	-	
Postoj před výukou	+	5	3	8
	-	16	2	18
		21	5	26

$$\chi^2 = \frac{(|3 - 16| - 1)^2}{3 + 16} = 7,58$$

Stupně volnosti

Tabulky: $\chi^2_{1-\alpha} (v = k(k-1)/2 = 1) = 3,84$



H₀ zamítnuta

- Závěr: Úspěch našich sportovců má pozitivní vliv na postoj žáků vzhledem k provozování sportu.

McNemarův test: příklad II

Příklad: Srovnání 2 metod stanovení antigenu v krvi (antigen vždy přítomen)



H_0 : metoda 1 = metoda 2

Metoda 1	Metoda 2	Frekvence
úspěch	úspěch	202
úspěch	neúspěch	60
neúspěch	úspěch	42
neúspěch	neúspěch	10

} $\Sigma = 102$

$$\chi_{(c)}^2 = \frac{(|60 - 42| - 1)^2}{102} = 2,83$$

Tabulky: $\chi_{1-\alpha}^2 (v=1) = 3,84$



H_0 nezamítnuta

Aplikace analýzy 2 x 2 tabulky pro hodnocení rizika

I. Prospektivní studie - odhad relativního rizika

Jedinci jsou sledováni prospektivně, zda se vyskytne nějaká vlastnost.
VÝBĚR JE DÁN SLOUPCEM

OBEČNĚ

		Skupina	
		1	2
Znak	ANO	a	b
	NE	c	d

Riziko: $\frac{a}{(a+c)}$ $\frac{b}{(b+d)}$

$$RR = \frac{\frac{a}{(a+c)}}{\frac{b}{(b+d)}}$$

PŘÍKLAD

		Retardace plodu	
		Symetrická	Asymetrická
Agpar skore > 7	ANO	2	33
	NE	14	58

$$RR = \frac{2/16}{33/91} = 0,345$$

2/16=0,13 33/91=0,36

Riziko u "symetrické skupiny" je asi 35 % rizika u asymetrické skupiny

$$SE(\ln RR) = \sqrt{\frac{1}{a} - \frac{1}{a+c} + \frac{1}{b} - \frac{1}{b+d}}$$

$$IS: \ln RR - Z_{1-\alpha/2} \cdot SE(\ln RR)$$

$$\ln RR + Z_{1-\alpha/2} \cdot SE(\ln RR)$$

H₀: RR = 1

Aplikace analýzy 2 x 2 tabulky pro hodnocení rizika

II. Retrospektivní studie - "ODDS RATIO"

Zcela zásadně odlišný přístup od retrospektivní studie
VÝBĚR JE DÁN VLASTNOSTÍ - ŘÁDKEM
 Není tedy možné analyzovat relativní riziko, protože přípravou řádků můžeme měnit velikost kontrol.

OBEČNĚ

		Skupina	Skupina
		1	2
Znak	ANO	a	b
	NE	c	d

odds a/c b/d

$$\text{Odds ratio} : \frac{a/c}{b/d}$$

$$SE(\ln OR) = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

PŘÍKLAD

		Vady chrupu	
		ANO	NE
Plavání týdně	< 6h	32	118
	≥ 6h	17	127

$$OR = (32/17) / (118/127) = 2,026$$

$$\ln(OR) = 0,706$$

$$SE(\ln(OR)) = 0,326$$

Relative risk vs. Odds ratio ?

Relative risk
(relativní riziko)



Odds ratio
(poměr šancí)

- Smysl RR a OR
- Výpočet
- Srovnatelnost
- Interpretace
- Výhody a nevýhody

- Aplikace v klinickém hodnocení

Smysl RR a OR

- Popis vlivu faktoru (léčba, klinický parametr) na výskyt události (úmrtí, progresse aj.)

Relative risk
(relativní riziko)



Odds ratio
(poměr šancí)

- ☑ Snadná přirozená interpretace rizik vyjádřených jako procento událostí

ALE

- ☑ Matematická omezení pro některé aplikace

- ☑ Pouze málo lidí má přirozenou schopnost interpretovat OR

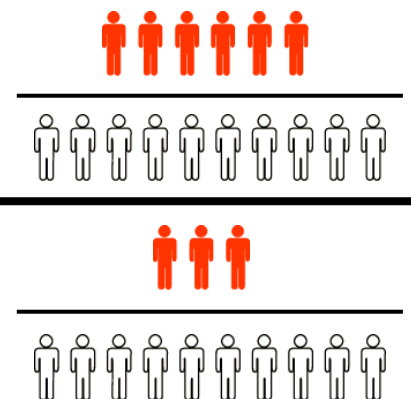
ALE

- ☑ OR v řadě aplikací výhodnější matematické vlastnosti



Výpočet

 event  bez eventu

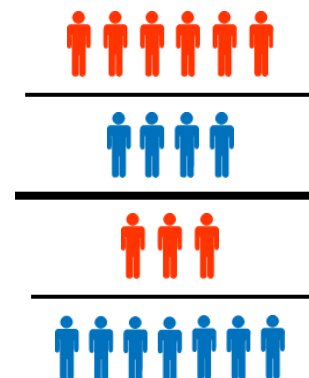
Relative risk
(relativní riziko)


$$RR = \frac{\frac{6}{10}}{\frac{3}{10}} = 2$$

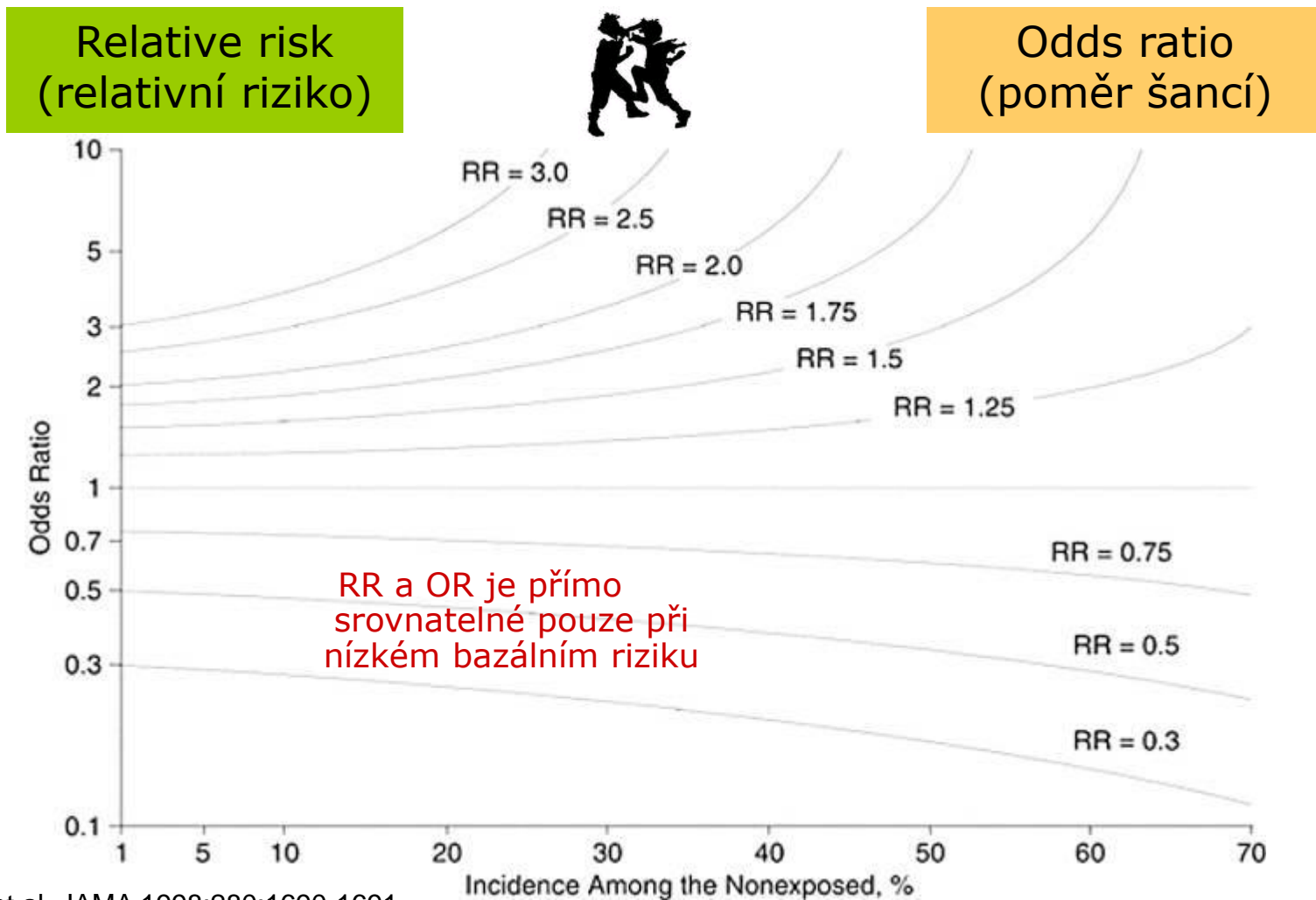


A B
 

Odds ratio
(poměr šancí)


$$OR = \frac{\frac{6}{4}}{\frac{3}{7}} = \frac{6}{3} = 2$$

Vztah mezi RR a OR



Zhang, J. et al. JAMA 1998;280:1690-1691.

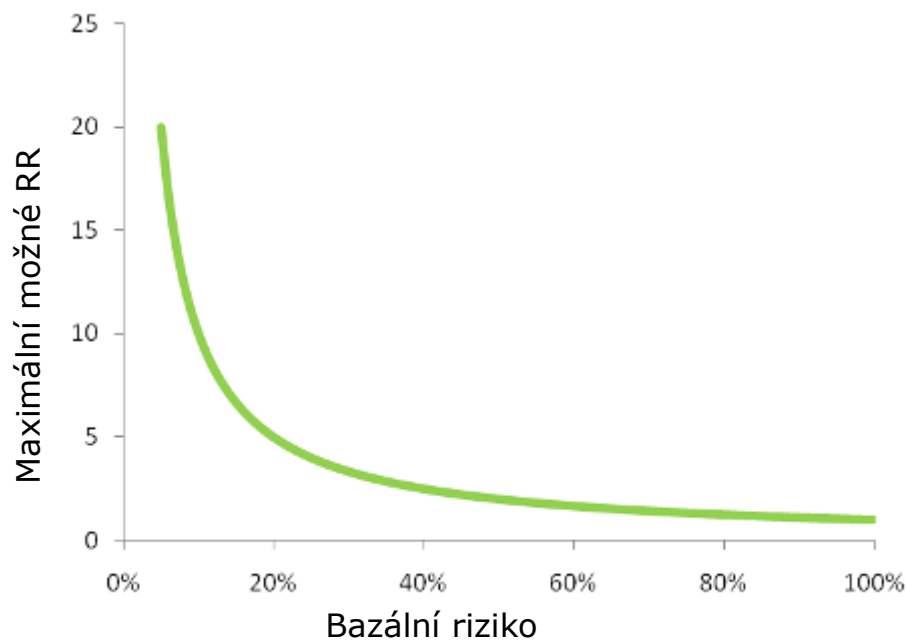
Srovnatelnost RR a OR I: maximum

Relative risk
(relativní riziko)



Odds ratio
(poměr šancí)

- RR mění své maximum podle bazálního rizika



- ✓ RR ve studiích s různým bazálním rizikem jsou nesrovnatelná !!!!

- ✓ Odds ratio má vždy rozsah od 0 do nekonečna
- ✓ Velikost OR není závislá na velikosti bazálního rizika

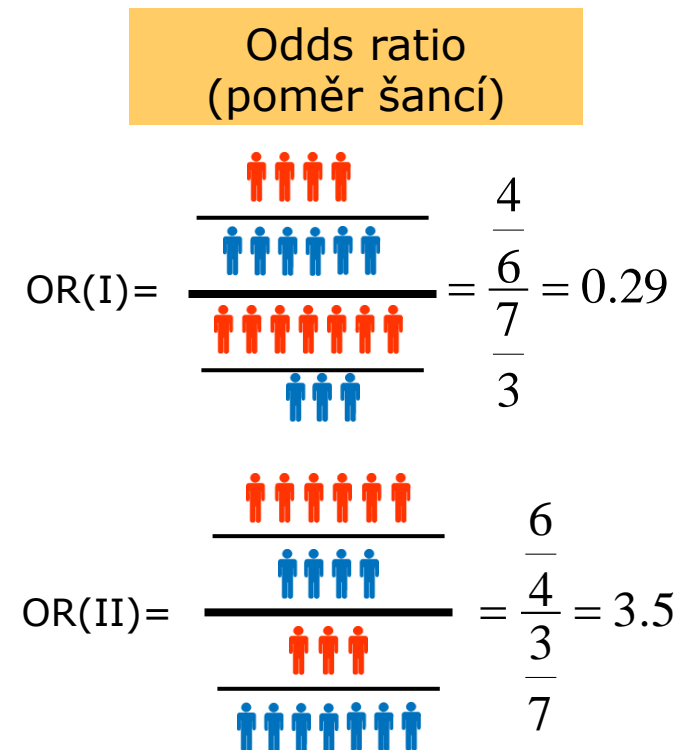
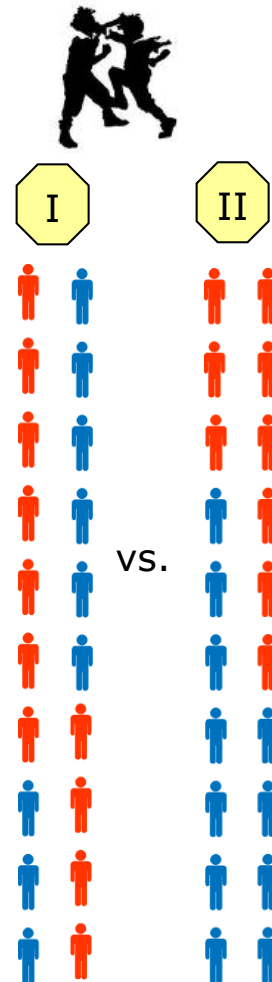
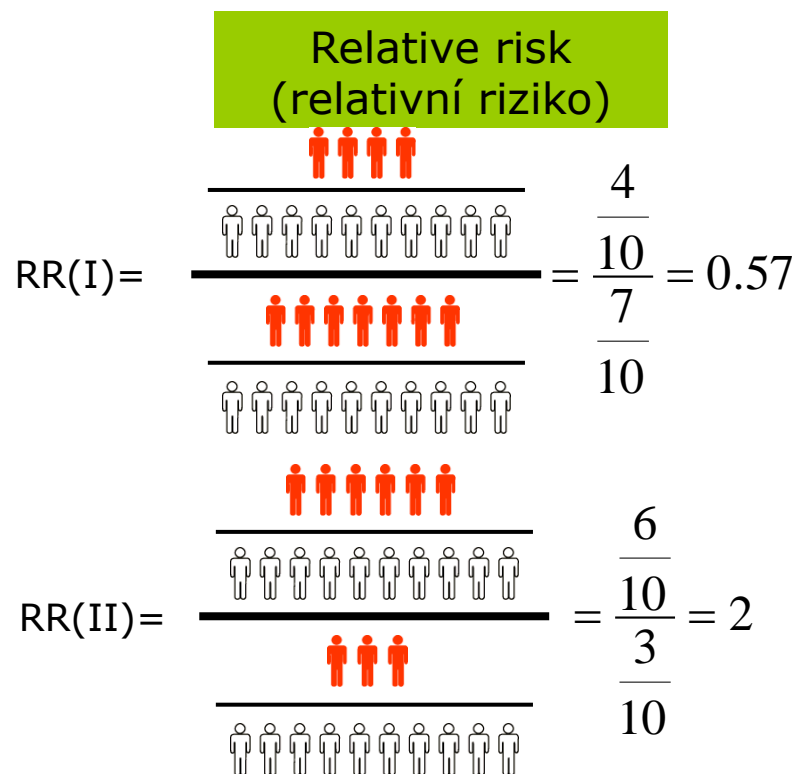
- ✓ OR lze použít pro srovnání studií s různým bazálním rizikem !!!!

- ✓ Výhodné pro metaanalýzu

Srovnatelnost RR a OR I: symetrie

- Existuje mezi RR a O rozdíl v případě

výměny definice eventu a non-eventu?

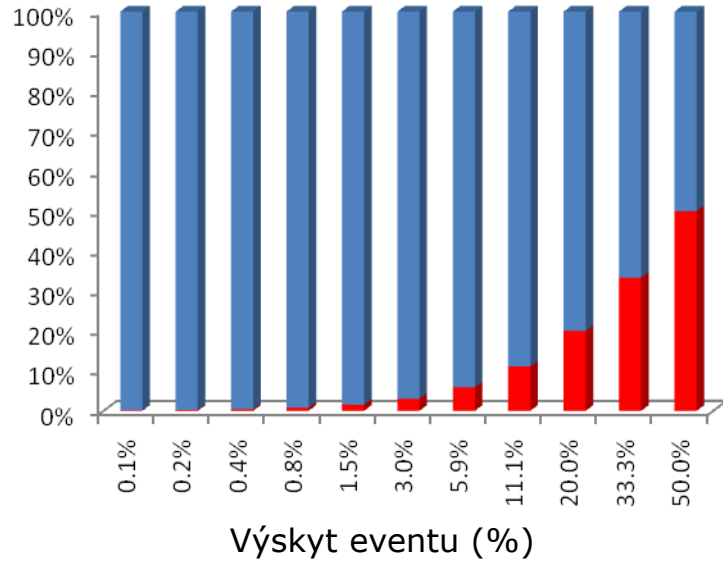


$$RR(I) \neq \frac{1}{RR(II)}$$

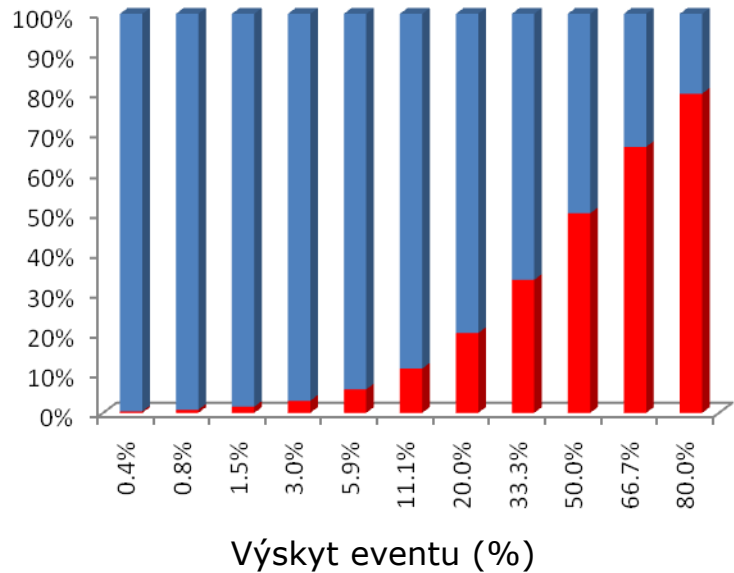
$$OR(I) = \frac{1}{OR(II)}$$

RR a OR ve studiích s různou mírou bazálního rizika

Control

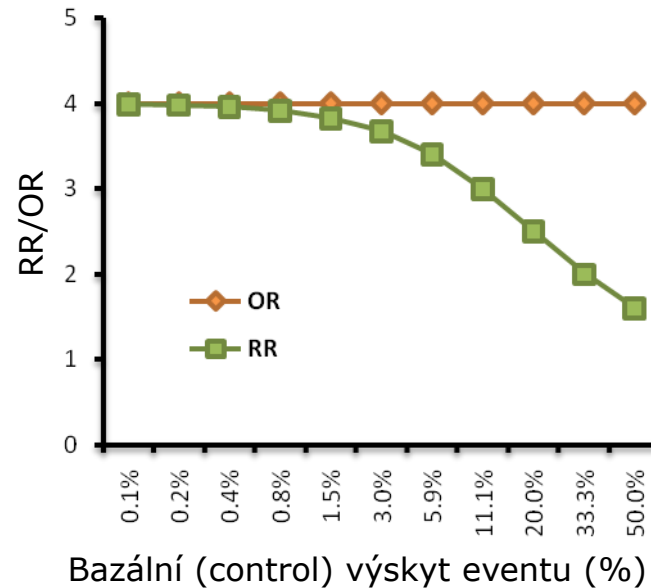


Case



Odds ratio

Ve skupině „Case“ připadá na jednoho pacienta bez eventu 4x tolik pacientů s eventem než ve skupině „Control“



Relative risk

Pacient ve skupině „Case“ má x-krát zvýšenou pravděpodobnost výskytu eventů než pacient ve skupině „Control“. X-krát závisí na bazálním výskytu eventů.

RR a OR v prospektivních a retrospektivních studiích

Prospektivní studie

- ☑ Sledování výskytu eventů a následná analýza jeho příčin
- ☑ Převážně kohortní studie



- ☑ Bazální výskyt eventů je dán vlastnostmi kohorty pacientů
- ☑ Bezproblémové využití RR



Relative risk
(relativní riziko)

Retrospektivní studie

- ☑ **Zpětné sledování příčin eventů**
- ☑ **Převážně case-control studie**
- ☑ **Výběrem pacientů ovlivňujeme bazální výskyt eventů**



- ☑ RR nelze použít – ovlivněno bazálním výskytem eventů
- ☑ Využití OR – není ovlivněno designem studie



Odds ratio
(poměr šancí)

Relative risk vs. Odds ratio: shrnutí

Relative risk
(relativní riziko)



Odds ratio
(poměr šancí)

- ☑ **Intuitivně snadno interpretovatelné**
- ☑ **Pro prospektivní studie**
- ☑ **Maximum se liší podle bazální hodnoty výskytu eventů**

- ☑ Retrospektivní studie
- ☑ Aplikace v metaanalýze
- ☑ Standardní výstup logistické regrese
- ☑ Rozsah vždy 0 až nekonečno, není ovlivněno bazálním výskytem eventů
- ☑ Obtížnější interpretace

Přednáška 9

Poissonovo rozdělení

Popis rozložení a jeho využití

Anotace

- Poissonovo rozdělení se používá pro popis četnosti výskytu jevu na experimentální jednotku, příkladem může být počet mutací bakterií na Petriho misku nebo počet srdečních poruch na jednotku času

Poissonovo rozdělení

Celkový počet jevů v n nezávislých pokusech

$$\left. \begin{array}{l} E(x) = n p \\ D(x) = n p \end{array} \right\} E(x) = D(x)$$

$$P(r) = \frac{e^{-\mu} \cdot \mu^r}{r!} = e^{-\lambda} \cdot \frac{\lambda^r}{r!}$$

$\mu = \lambda =$ průměrný počet jevů z n pokusů

$$\hookrightarrow P(X = 0) = e^{-\mu}$$

$$\hookrightarrow P(X = 1) = e^{-\mu} \cdot \mu^1$$

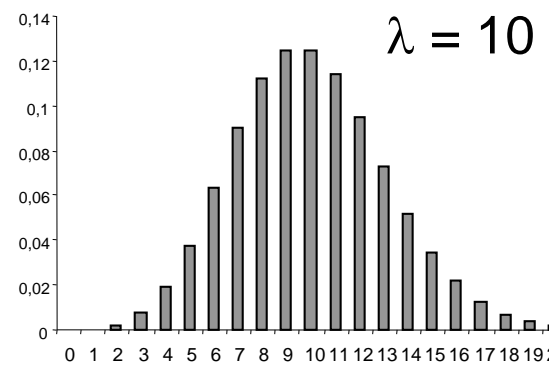
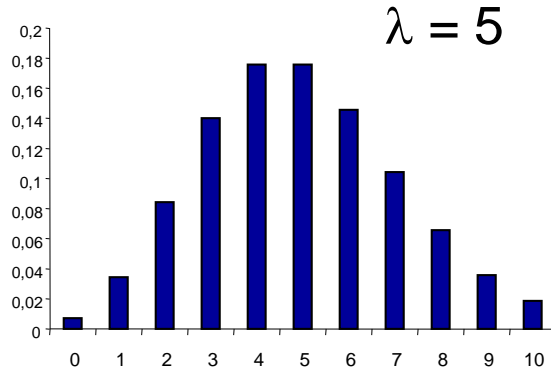
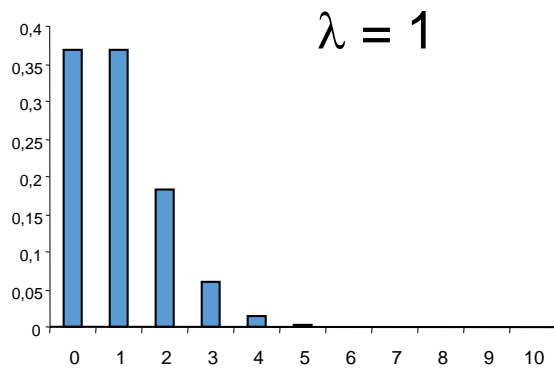
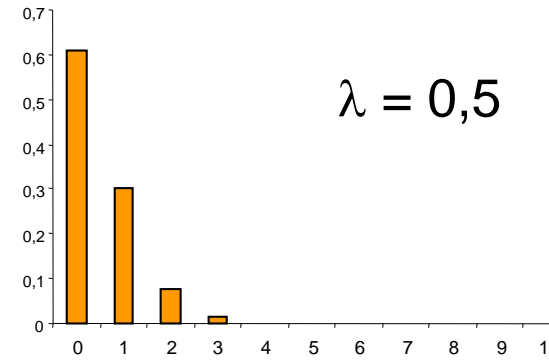
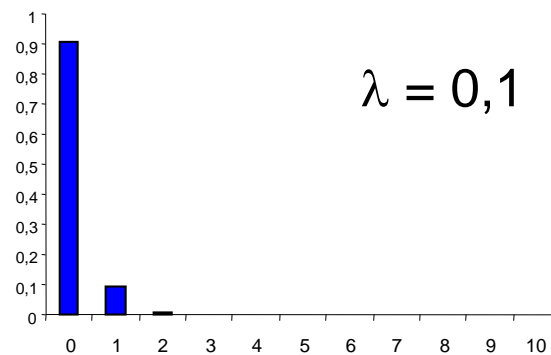
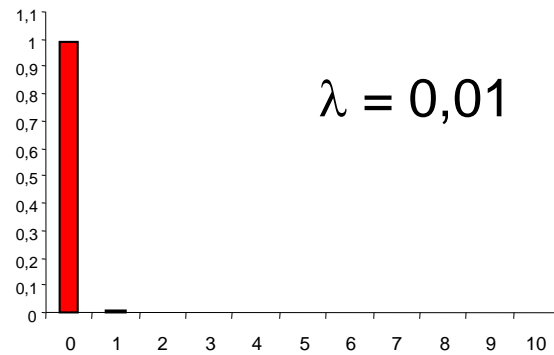
$$\hookrightarrow P(X = 2) = \frac{e^{-\mu} \cdot \mu^2}{2}$$

$$\hookrightarrow P(X = 3) = \frac{e^{-\mu} \cdot \mu^3}{(3)(2)}$$

$$\hookrightarrow P(X = 4) = \frac{e^{-\mu} \cdot \mu^4}{(4)(3)(2)}$$

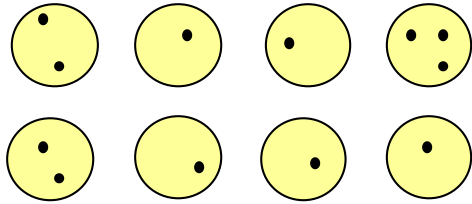
Poissonovo rozdělení jako model

$$P(x = r) = e^{-\lambda} \cdot \frac{\lambda^r}{r!}$$

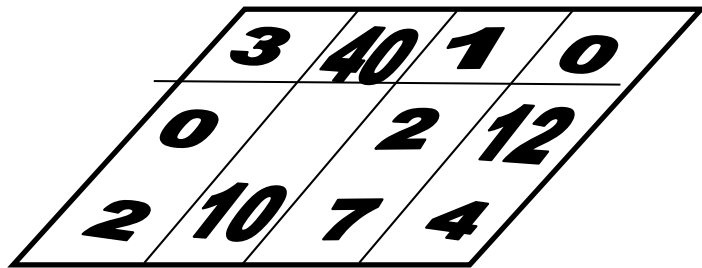


Poissonovo rozdělení v přírodě existuje

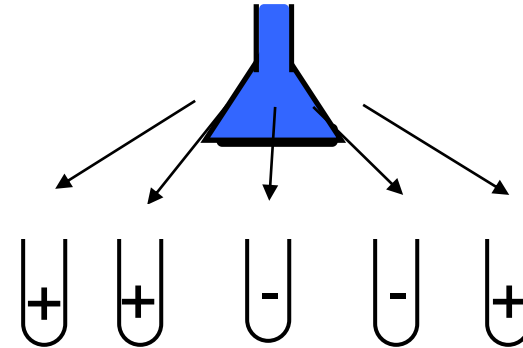
Mutace bakterií na inkubačních miskách



Výskyt jevu v prostoru
(počet žížal na určité plochu pole)



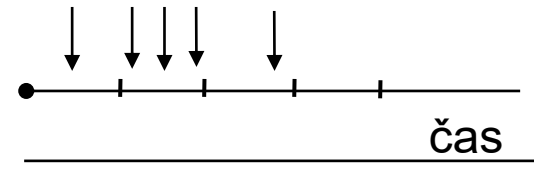
Orientační stanovení jevu
(při produkci plynu bakteriemi)



The most probable number
technique

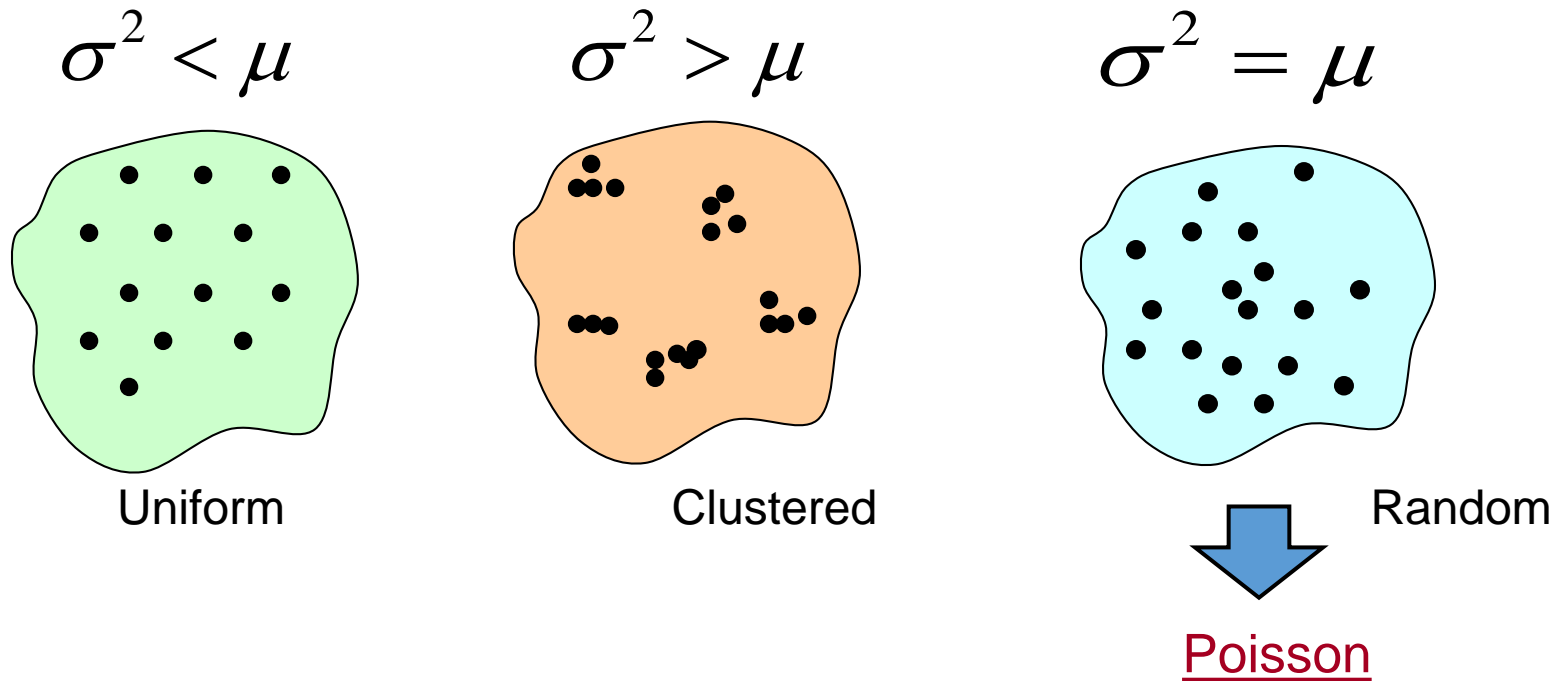
Výskyt jevu v čase

(srdeční arytmie v určitých časových intervalech)



Poissonovo rozdělení jako model pro náhodný výskyt jevů

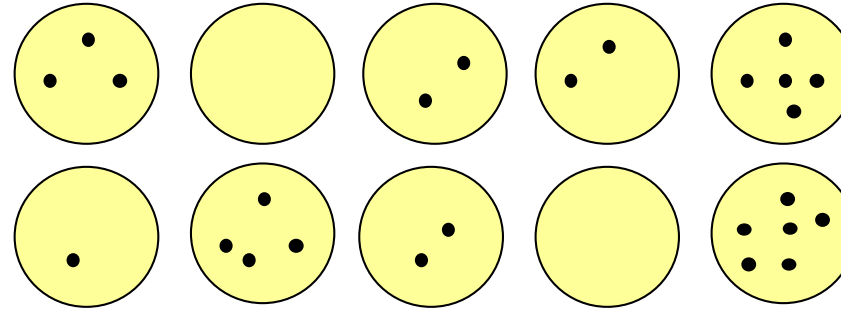
Předpoklad: náhodná distribuce jevu mezi studovanými objekty
(příp. v čase, v prostoru).



Pokud je λ spíše větší ($\sim 5 - 10$), pak Poisson odpovídá spíše binomickému až normálnímu rozložení.

Formální prezentace Poissonova rozložení

Př: pokus.....10 000 bakterií na misce
n = 10 misek
Jev: mutace (r=25)
 λprůměrný počet mutantů na
jednu misku



$$r = 25$$

$$\bar{x} \approx \lambda = 25/10 = 2,5$$

95 % IS:

$$\bar{x} - Z_{1-\alpha/2} \cdot \sqrt{\frac{\bar{x}}{n}} \leq \lambda \leq \bar{x} + Z_{1-\alpha/2} \cdot \sqrt{\frac{\bar{x}}{n}}$$

$$2,5 - 1,96 \cdot \sqrt{0,25} \leq \lambda \leq 2,5 + 1,96 \cdot \sqrt{0,25}$$

$$1,52 \leq \lambda \leq 3,48$$

Poissonova náhodná proměnná

- Při měření počtu krvinek změněných určitou chorobou (relativně vzácné) je pozorován zředěný vzorek krve pod mikroskopem v komůrce rozdělené na stejně velká pole. Sledovaná veličina, udávající počet krvinek v i -tém poli může být považována za rozdělenou podle Poissonova rozložení:
- $n = 169$ = počet nezávislých pozorování proměnné
- $r = 10$ = počet pozorovaných krvinek
- Jaká je hodnota parametru λ Poissonova rozložení a jaká je jeho interpretace ?
- Jaký je interval 95% spolehlivosti pro parametr λ
- Pokud bychom sledovali celkový počet červených krvinek (opět v $n = 169$ nezávislých políčkách), bylo by i tuto proměnnou možno považovat za rozdělenou podle Poissonova rozložení ? Uvažujte celkový počet pozorovaných krvinek jako 2013.

Výpočet intervalu spolehlivosti pro λ (bez aproximace na normální rozložení)

Spodní hranice IS

$$L_1 = \frac{\chi_{1-\alpha/2}^2 (f_1=2r)}{2}$$

Horní hranice IS

$$L_2 = \frac{\chi_{\alpha/2}^2 (f_2=f_1+2)}{2}$$

Poissonova náhodná proměnná

Konstantní zářič: $n = 2608$ časových intervalů (každý 7,5 s)

i : počet částic v intervalu (x)

s_i : pozorovaná četnost intervalů s i částicemi

$$P(x = i) = \frac{\lambda^i \cdot e^{-\lambda}}{i!} \sim p_i$$

Poissonova proměnná:

* Výborný model pro experimenty, v nichž je během časového průběhu zjišťován počet výskytu určitého jevu

i	Počet intervalů s právě i zaznamenanými částicemi s_i	teoretické četnosti np_i	$\frac{(s_i - np_i)^2}{np_i}$
0	57	54,399	0,1244
1	203	210,523	0,2688
2	383	407,361	1,4568
3	525	525,496	0,0005
4	532	508,418	1,0938
5	408	393,515	0,5332
6	273	253,817	1,4498
7	139	140,325	0,0125
8	45	67,882	7,7132
9	27	29,189	0,1642
10	10	17,075 (= $P\{\xi \geq 10\}$)	0,0677
11	4		
12	2		
13	0		
	$n = 2608$	2608,00	12,8849

Poissonovo rozdělení: jednovýběrový test

$$P_{(r)} = \frac{(e^{-\lambda} \cdot \lambda^r)}{r!}$$

Př: Počet hnízd křepelek na dané ploše

$$\left. \begin{array}{l} n = 8\,000 \quad \text{"pod lokality"} \\ r = 28 \end{array} \right\} \hat{p} = 0,0035$$

Nechť je srovnávací soubor
(předchozí průzkum)

$$p_o = 0,0020$$

$$\underline{p_o \cdot 8\,000 = 16 = \mu = \lambda}$$

$$\underline{H_o : p \leq p_o \sim \mu \leq 16 \quad ?}$$

1) Vztít data jako pocházející z populace:

$$P(r = 28) = \frac{e^{-16} \cdot 16^{28}}{28!} = \underline{0,00192}$$

$$2) \left. \begin{array}{l} P(r \geq 28) = ? \\ [0,00411] \end{array} \right\} < 0,05 \Rightarrow \underline{H_o \text{ zamítnuta}}$$



$r = 28$ je příliš velké pro populaci s p_o



$\underline{p > p_o}$, aby $r = 28$ bylo
pravděpodobnější

Analýza rozptylu

Parametrická analýza rozptylu

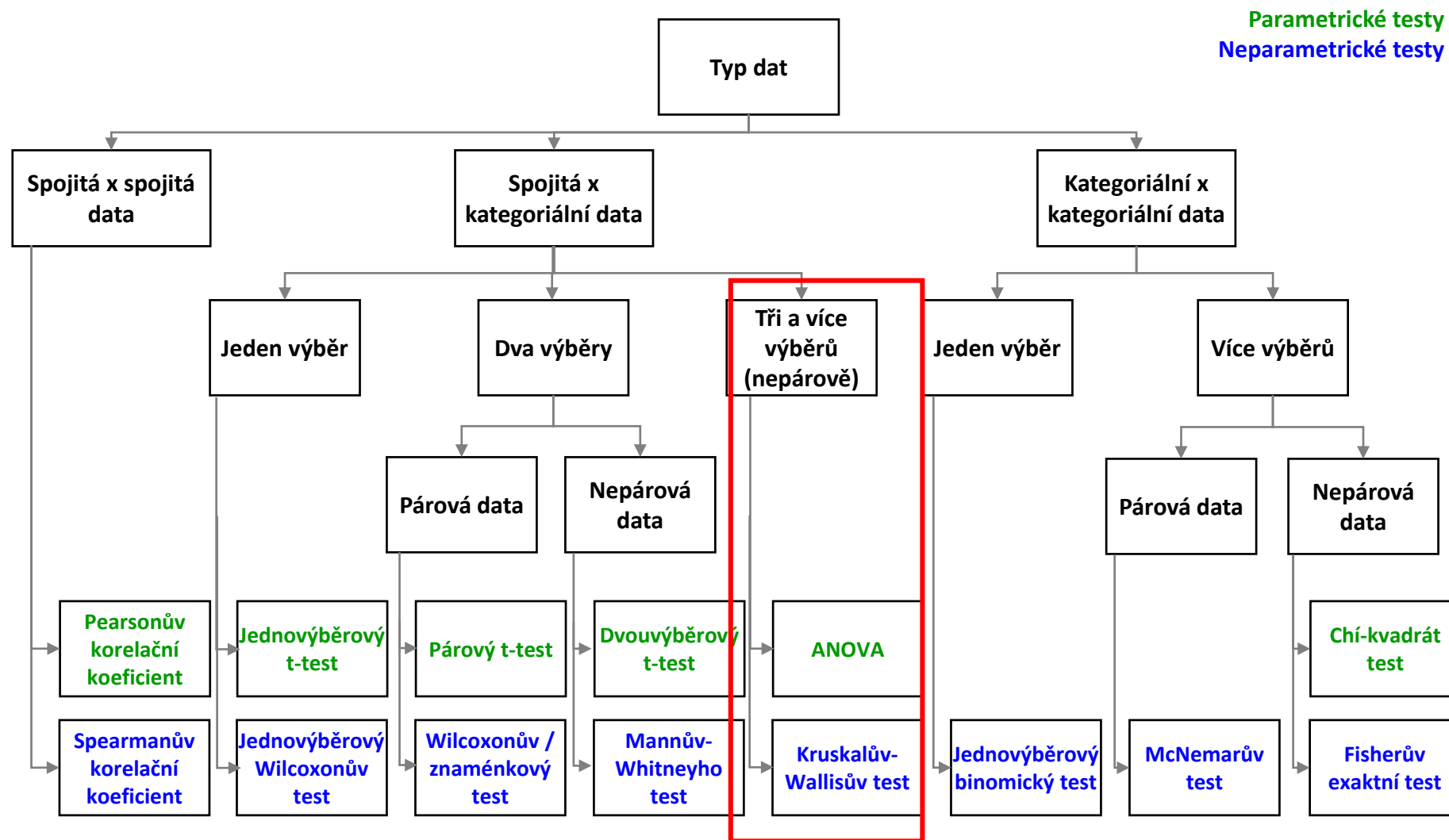
Post hoc testy

Kruskal-Wallis test

Anotace

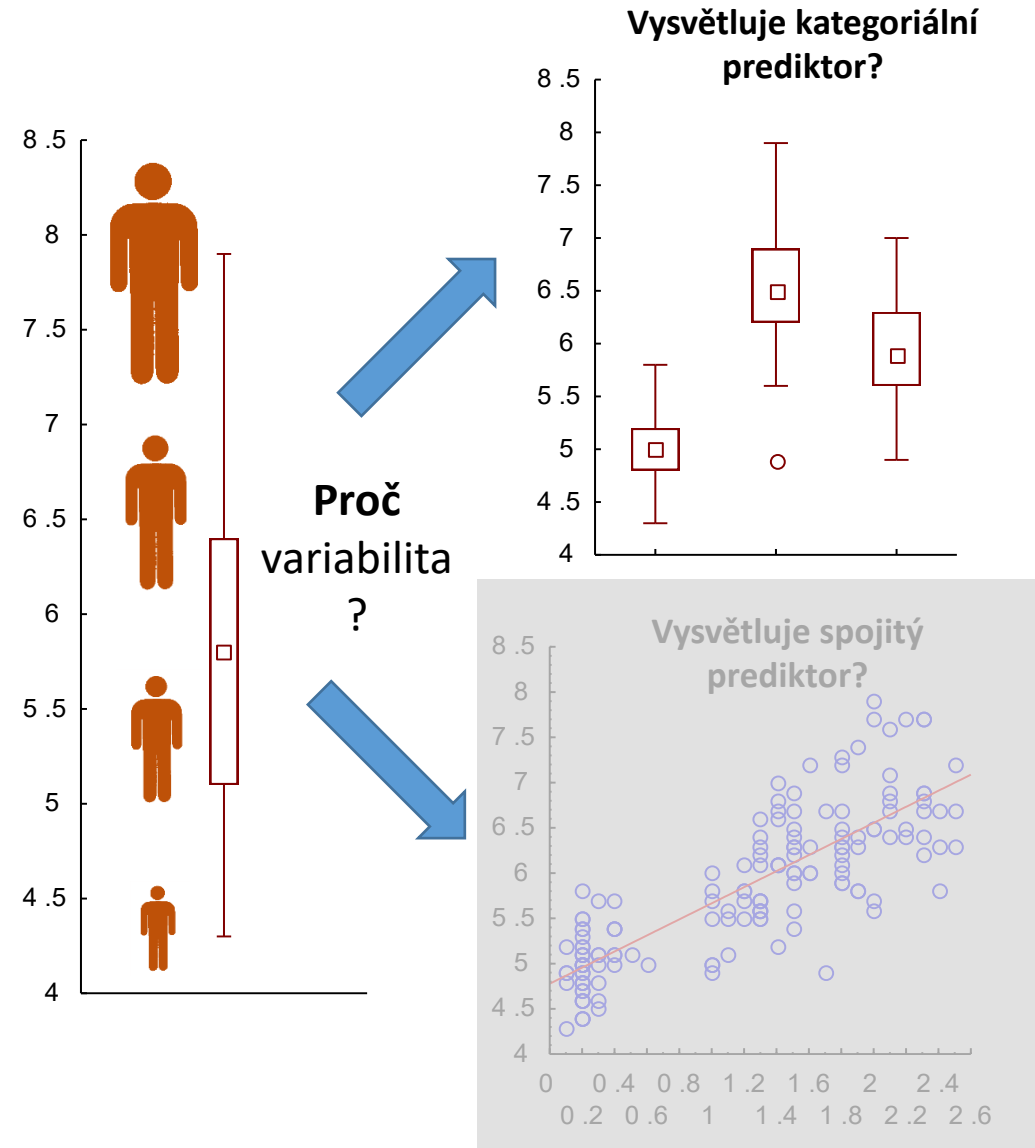
- Analýza rozptylu je základním nástrojem pro analýzu rozdílů mezi průměry v několika skupinách objektů.
- Základní myšlenka, na níž je ANOVA založena, je rozdělení celkové variability v datech (neznámé, dané pouze náhodným rozložením) na část systematickou (spjatou s kategoriemi pacientů, vysvětlená variabilita) a část náhodnou. Pokud systematická, tedy nenáhodná a vysvětlitelná část variability převažujeme, považujeme daný kategoriální faktor za významný pro vysvětlení variability dat.
- Analýza rozptylu vyhodnocuje pouze celkový vliv faktoru na variabilitu, v případě analýzy jednotlivých kategorií je třeba využít tzv. post-hoc testy

Základní rozhodování o výběru statistických testů



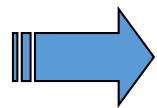
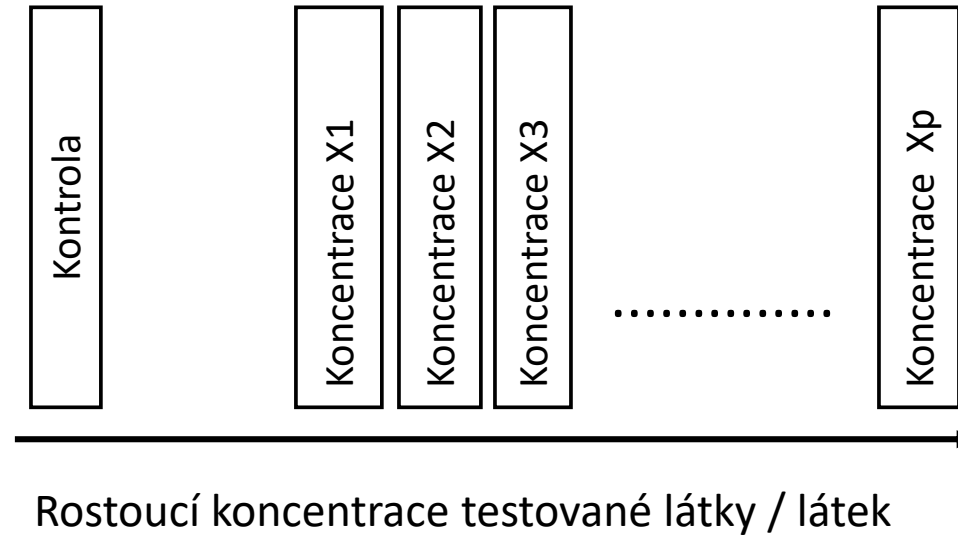
Cíl stochastického modelování

- Obecným cílem je snaha vysvětlit variabilitu predikované proměnné (endpoint, Y) pomocí prediktorů (vysvětlující proměnná, faktor, X)
- Jak predikovaná proměnná, tak prediktor mohou být různého typu
 - Binární
 - Kategoriální
 - Ordinální
 - Spojitá
- Cenzorovaná (-> analýza přežití)
- Kombinace datového typu predikované proměnné a prediktoru určuje použitou metodu analýzy

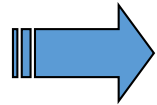


Analýza rozptylu - ANOVA

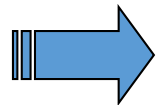
Základní technika
sloužící
k posouzení rozdílů
mezi více úrovněmi
pokusného zásahu



Celkově významné změny v reakci biologického systému



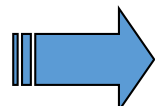
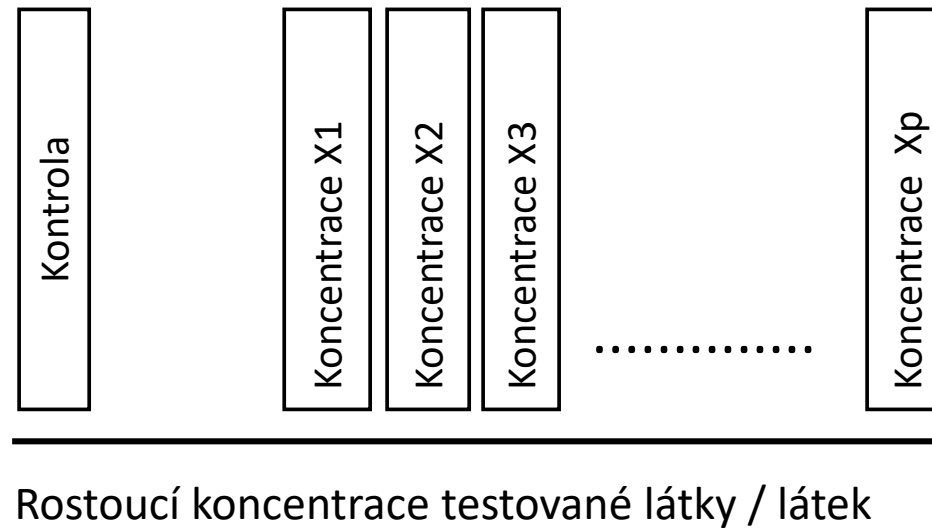
Vzájemné rozdíly účinku jednotlivých dávek



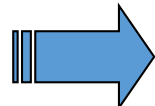
Rozdíly účinku dávek od kontroly

Analýza rozptylu - ANOVA

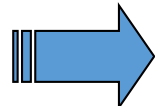
Významné kroky
analýzy, vedoucí k
efektivnímu srovnání
variant



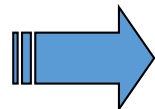
Splnění předpokladů analýzy
Transformace dat



Relevantnost kontroly
(vliv vlastní aplikace látek)



Vhodnost modelu ANOVA pro účely testu



Vlastní srovnání variant
Minimalizace chyb při ověřování hypotéz

Analýza rozptylu - ANOVA

SPLNĚNÍ PŘEDPOKLADŮ ANOVA JE NEZBYTNOU PODMÍNKOU
POUŽITÍ TÉTO TECHNIKY

ANOVA
= parametrická
analýza dat

1. Předpoklad nezávislosti
opakování experimentu

2. Homogenita rozptylu
v rámci pokusných
variant

Normalita rozložení
3. v rámci pokusných variant

ALTERNATIVOU JSOU NEPARAMETRICKÉ METODY

ANOVA – předpoklady

- Symetrické rozložení hodnot a normalita odchylek od hodnoceného modelu ANOVA. Velkou část dat lze adekvátně normalizovat použitím logaritmické transformace. Předpoklad lognormální transformace může pochopitelně být teoreticky vyloučen u mnoha datových souborů obsahujících diskrétní parametry, kde je indikována vhodnost jiného typu transformace. U asymetricky rozložených a u diskrétních dat je nutné využít neparametrické alternativy analýzy rozptylu.
- Homogenita rozptylu je nutným předpokladem pro smysluplnost vzájemných srovnání pokusných variant. U testů toxicity by splnění tohoto předpokladu mělo být ověřováno (Bartlettův test), neboť vážné rozdíly (až řádové) v jednotkách testovaného parametru mohou nastat v důsledku inhibice dávkami látky. Nehomogenita rozptylu je často ve vztahu k nenormalitě (asymetrii) dat a lze ji odstranit vhodnou normalizující transformací.
- Statistická nezávislost reziduí vyhodnocovaného modelu ANOVA. Pokud odhad a posouzení korelačních vztahů mezi pokusnými variantami není přímo předmětem výzkumu, lze jejich vliv na vyhodnocení odstranit znáhodněním dat v rámci pokusných variant - tedy změnou pořadí v náhodné. Rozsah vlivu těchto autokorelačních vztahů musí být ovšem primárně omezen správností experimentálního uspořádání.
- Aditivita jako předpoklad týkající se složitějších experimentálních uspořádání. Exaktní otestování aditivity více pokusných faktorů je procedura poměrně náročná na experimentální design vyvážený co do počtu opakování. Je rovněž obtížné testovat interakci na nestandardních datech, neboť případná transformace může změnit charakter odchylek původních dat od hodnoceného modelu ANOVA.

Omezení aplikace ANOVA lze řešit

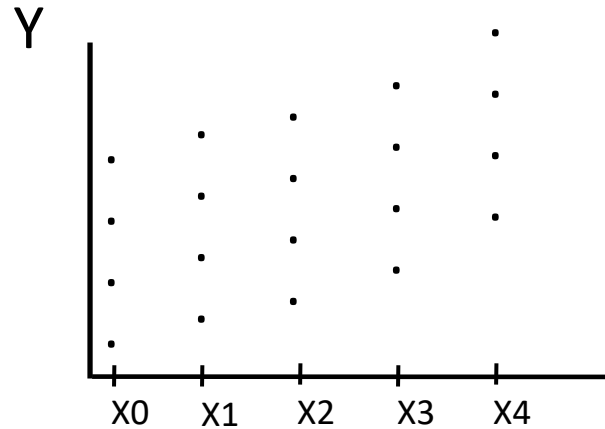
- **Chybějící data.** Vážným problémem jsou chybějící údaje o celé skupině kombinací testovaných látek, například u faktoriálních pokusů, kdy je znemožněno hodnocení experimentu jako celku.
- **Různé počty opakování.** Jde o typický jev pro experimentální datové soubory. Při různých počtech opakování v experimentálních variantách jsou testy ANOVA citlivější na nenormalitu dat. Pokud jsou počty opakování zcela odlišné (až na řádové rozdíly), je nutno použít neparametrické techniky nebo analýzu rozptylu nevyvážených pokusů.
- **Odlehlé hodnoty.** Ojedinelé odlehlé hodnoty musí být před parametrickou analýzou rozptylu vyloučeny.
- **Nedostatek nezávislosti mezi rezidui modelu.** Jde o závažný nedostatek, zkreslující výsledek F-testu. Velmi často je tato skutečnost důsledkem špatného provedení nebo naplánování experimentu.
- **Nehomogenita rozptylu.** Velmi častý nedostatek experimentálních dat, často související s nenormalitou rozložení nebo s odlehlými hodnotami.
- **Nenormalita dat.** I v tomto případě lze situaci upravit vyloučením odlehlých hodnot nebo normalizující transformací.
- **Nedítivita kombinovaného vlivu více pokusných zásahů.** Tuto situaci lze testovat jednak speciálními testy aditivity nebo přímo F testem kontrolujícím významnost vlivu interakce pokusných zásahů. Při významné interakci je nutné prozkoumat především její charakter ve vhodném experimentálním uspořádání.

Modely analýzy rozptylu

Model I. Pevný model

X_0	X_1	X_2	X_3	X_4
.
.
.
.
.
.
.
.
.

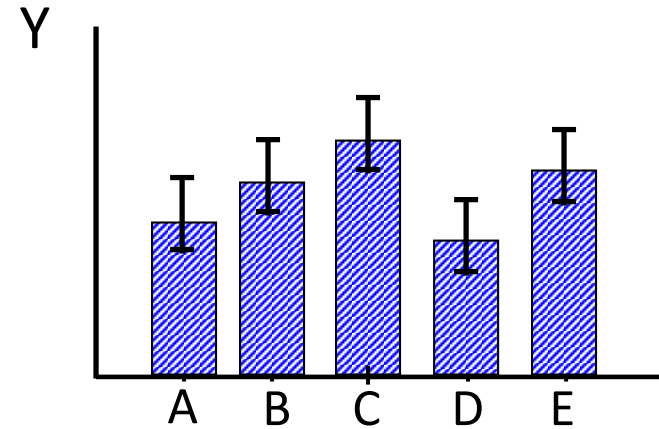
$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$



Model II. Náhodný model

A	B	C	D	E
.
.
.
.
.
.
.
.
.
.

$$y_{ij} = \mu + A_i + \varepsilon_{ij}$$



Princip ANOVA

- Základním principem ANOVY je porovnání rozptylu připadajícího na:
 - Rozdělení dat do skupin (tzv. effect, variance between groups)
 - Variabilitu objektů uvnitř skupin (tzv. error, variance within groups), předpokládá se, že jde o náhodnou variabilitu (=error)

1. Variabilita mezi skupinami

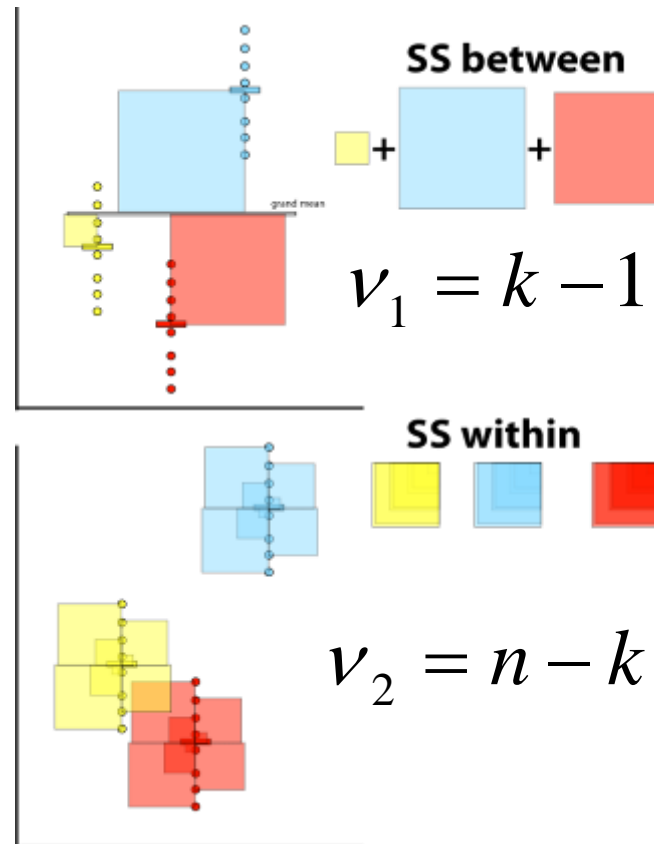
Rozptyl je počítán pro celkový průměr (tzv. grand mean) a průměry v jednotlivých skupinách dat

Stupně volnosti jsou odvozeny od počtu skupin (= počet skupin -1)

2. Variabilita uvnitř skupin

Rozptyl je počítán pro průměry jednotlivých skupin a objekty uvnitř příslušných, celková variabilita je pak sečtena pro všechny skupiny

Stupně volnosti jsou odvozeny od počtu hodnot (= počet hodnot - počet skupin)



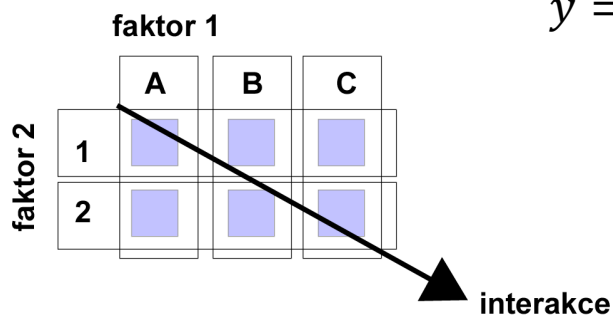
$$F = \frac{\text{between_groups}}{\text{within_groups}}$$

Výsledný poměr (F) porovnáme s tabulkami F rozložení pro v_1 a v_2 stupňů volnosti

SS=sum of squares

Design modelu

- Design modelu znamená jaké proměnné a v jakých kombinacích budou vysvětlovat hodnocenou proměnnou
- Obecně je vhodné ať již expertně nebo jako výsledek předběžné analýzy vytvořit a ověřit hypotézy o vzájemných vztazích proměnných a podle těchto předběžných výsledků vytvářet finální model
- Tvorba designu modelu úzce souvisí s pojmy:
 - Analýza pouze hlavních efektů proměnných
 - Analýza interakcí mezi proměnnými a složitost interakcí
- Design modelu lze vyjádřit graficky nebo v rovnici nebo pomocí maticového zápisu

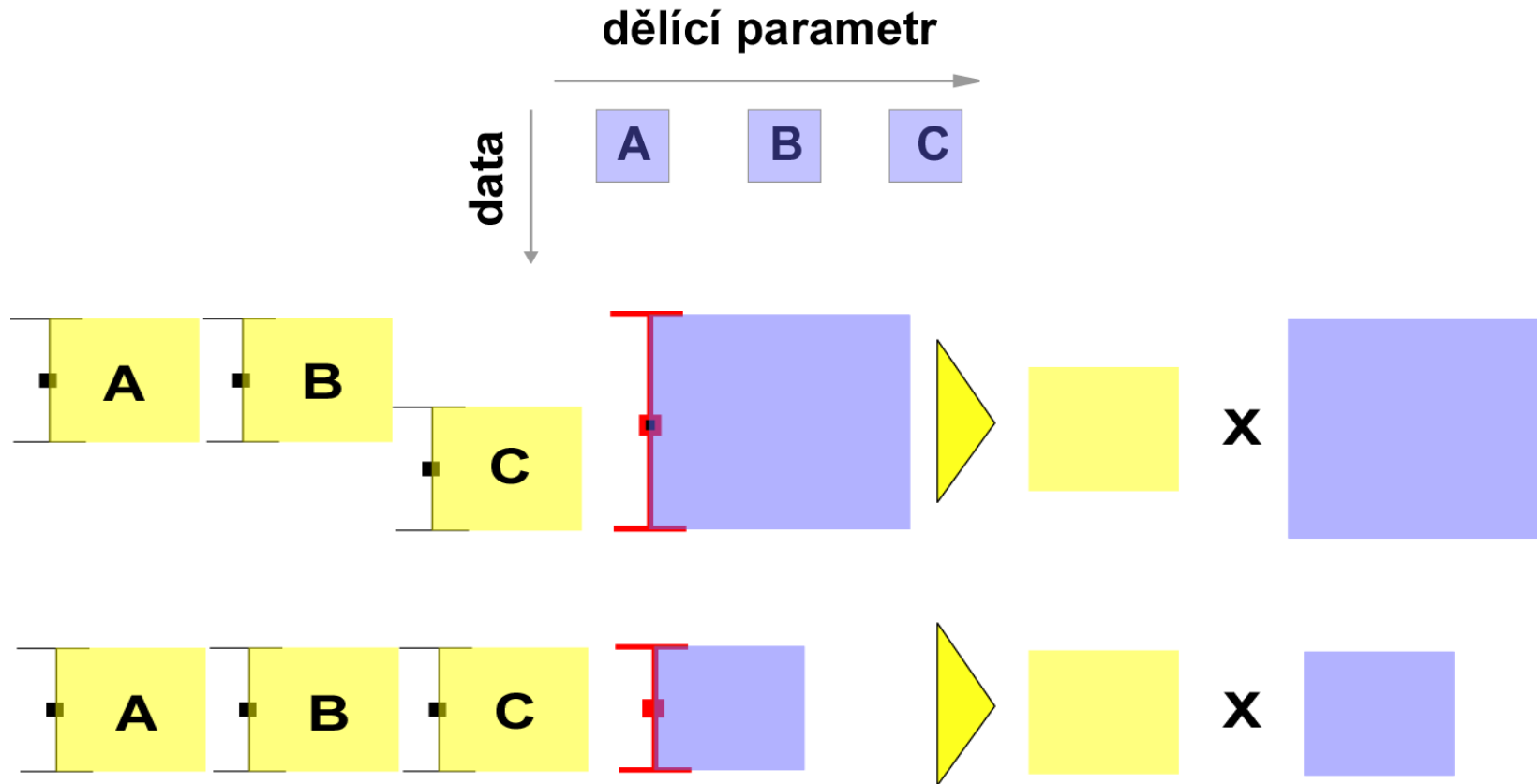


$$y = hmotnost * 1.5 + věk * 3.6 + hmotnost * věk * 1.8 + 9$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

Jednoduchý ANOVA design

- Nejjednodušším případem ANOVA designu je rozdělení na skupiny podle jednoho parametru



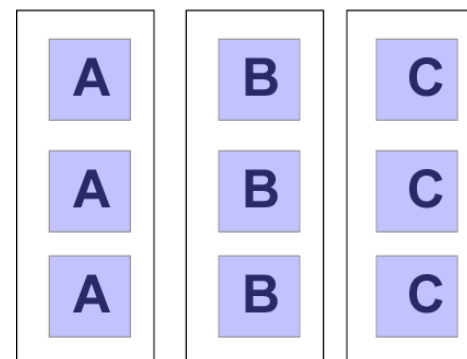
Nested ANOVA

- Rozdělení skupin na náhodné podskupiny (např. opakování experimentu)
- Cílem je zjistit, zda data v jedné skupině nejsou pouhou náhodou
- Nejprve je testována shoda podskupin v hlavních skupinách,
 - pokud jsou shodné, je vše v pořádku
 - pokud nejsou, stále lze zjišťovat, zda se variabilita uvnitř hlavních skupin liší od celkové variability

jednoduchá ANOVA

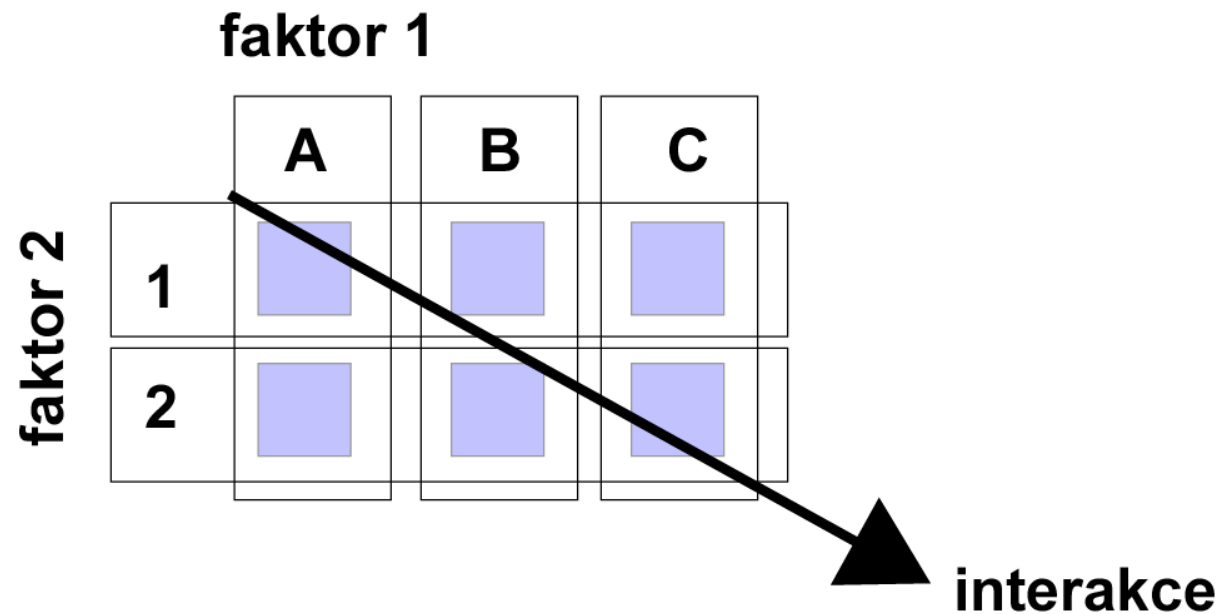


nested ANOVA



Two way ANOVA

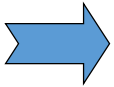
- Pro rozdělení do kategorií je zde více parametrů
- Na rozdíl od nested ANOVY nejde o náhodná opakování experimentu, ale o řízené zásahy (např.vliv pH a koncentrace O₂)
- Kromě vlivu hlavních faktorů se uplatňuje i jejich interakce



ANOVA – základní výstup

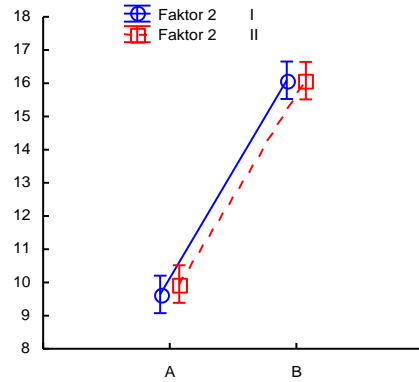
- Základním výstupem analýzy rozptylu je Tabulka ANOVA - frakcionace komponent rozptylu

Zdroj rozptylu	St. v.	SS	MS	F
Pok. zásah (mezi skupinami)	a - 1	SS_B	$SS_B/(a - 1)$	MS_B/MS_E
Uvnitř skupin	N - a	SS_E	$SS_E/(N - a)$	
Celkem	N - 1	SS_T		

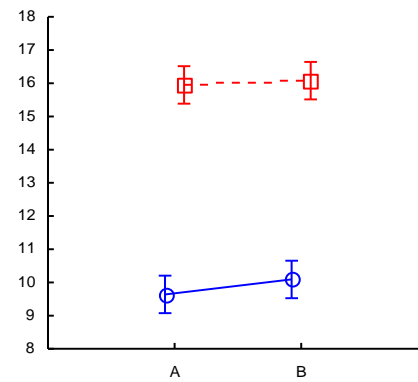
SS_B/SS_T  Kvantifikovaný podíl rozdílu mezi pokusnými zásahy na celkovém rozptylu

MS_B/MS_T  Statistická významnost rozdílu

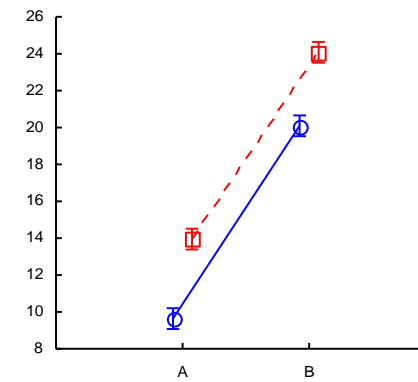
Hlavní efekty a interakce



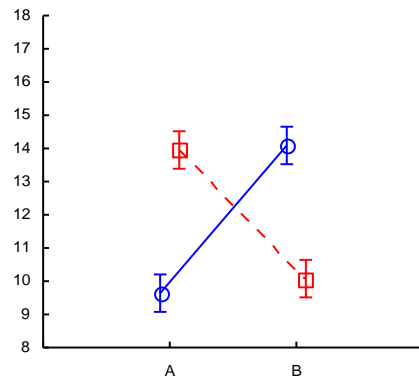
	SS	D.f.	MS	F	p
Intercept	33487	1	33487	8165.3	0.000
Faktor 1	1978	1	1978	482.2	0.000
Faktor 2	1	1	1	0.3	0.602
F1*F2	1	1	1	0.3	0.570
Error	804	196	4		



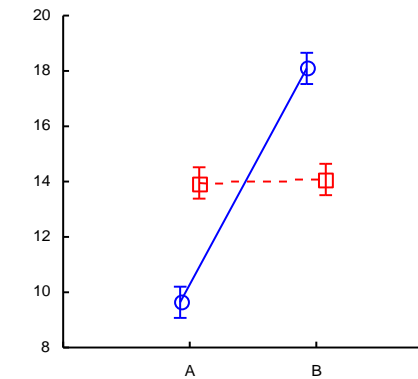
	SS	D.f.	MS	F	p
Intercept	33487	1	33487	8165.3	0.000
Faktor 1	4	1	4	1.0	0.314
Faktor 2	1891	1	1891	461.1	0.000
F1*F2	1	1	1	0.3	0.570
Error	804	196	4		



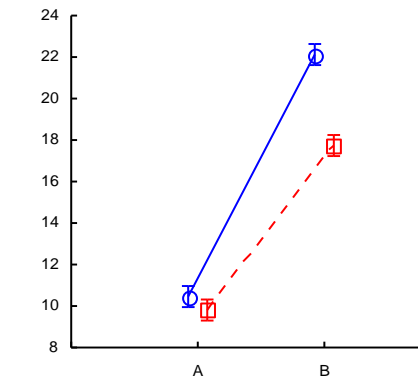
	SS	D.f.	MS	F	p
Intercept	57391	1	57391	13993	0.000
Faktor 1	5293	1	5293	1290.7	0.000
Faktor 2	861	1	861	209.9	0.000
F1*F2	1	1	1	0.3	0.570
Error	804	196	4		



	SS	D.f.	MS	F	p
Intercept	28511	1	28511	6952.0	0.000
Faktor 1	4	1	4	1.0	0.314
Faktor 2	1	1	1	0.3	0.602
F1*F2	867	1	867	211.3	0.000
Error	804	196	4		

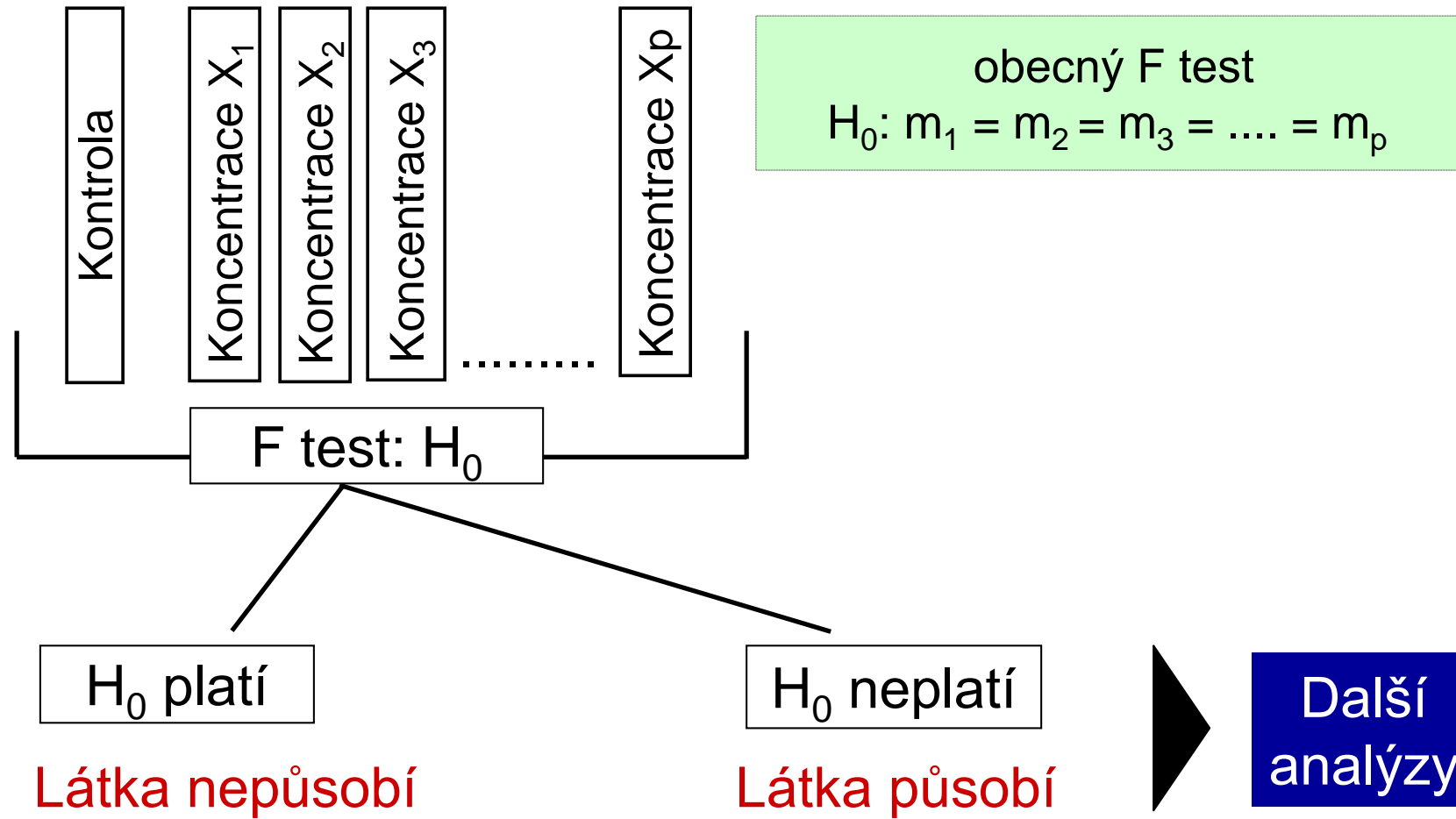


	SS	D.f.	MS	F	p
Intercept	38863	1	38863	9476.2	0.000
Faktor 1	920	1	920	224.3	0.000
Faktor 2	1	1	1	0.3	0.602
F1*F2	867	1	867	211.3	0.000
Error	804	196	4		



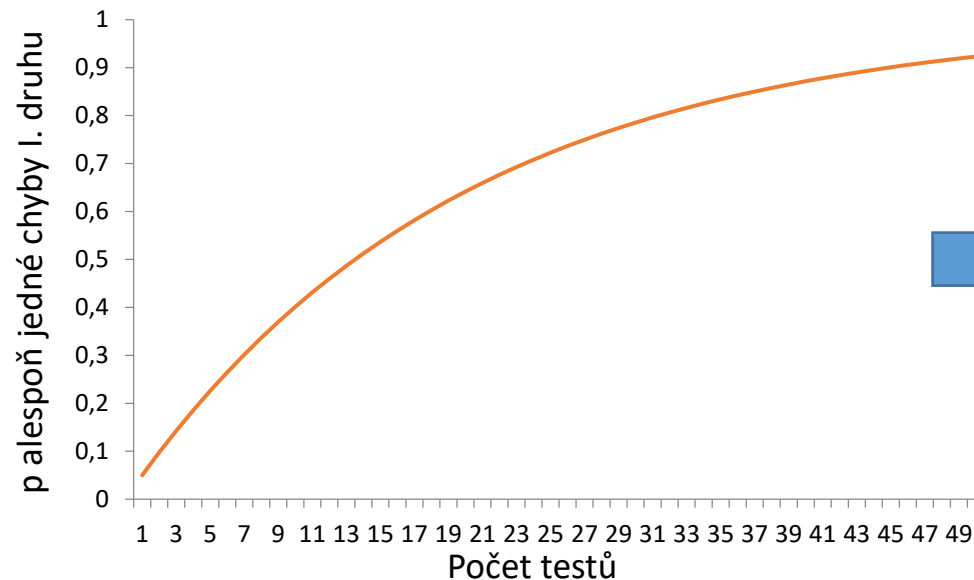
	SS	D.f.	MS	F	p
Intercept	45203	1	45203	13596	0.000
Faktor 1	4799	1	4799	1443.4	0.000
Faktor 2	316	1	316	95.0	0.000
F1*F2	175	1	175	52.5	0.000
Error	652	196	3		

Analýza rozptylu - obecný F test



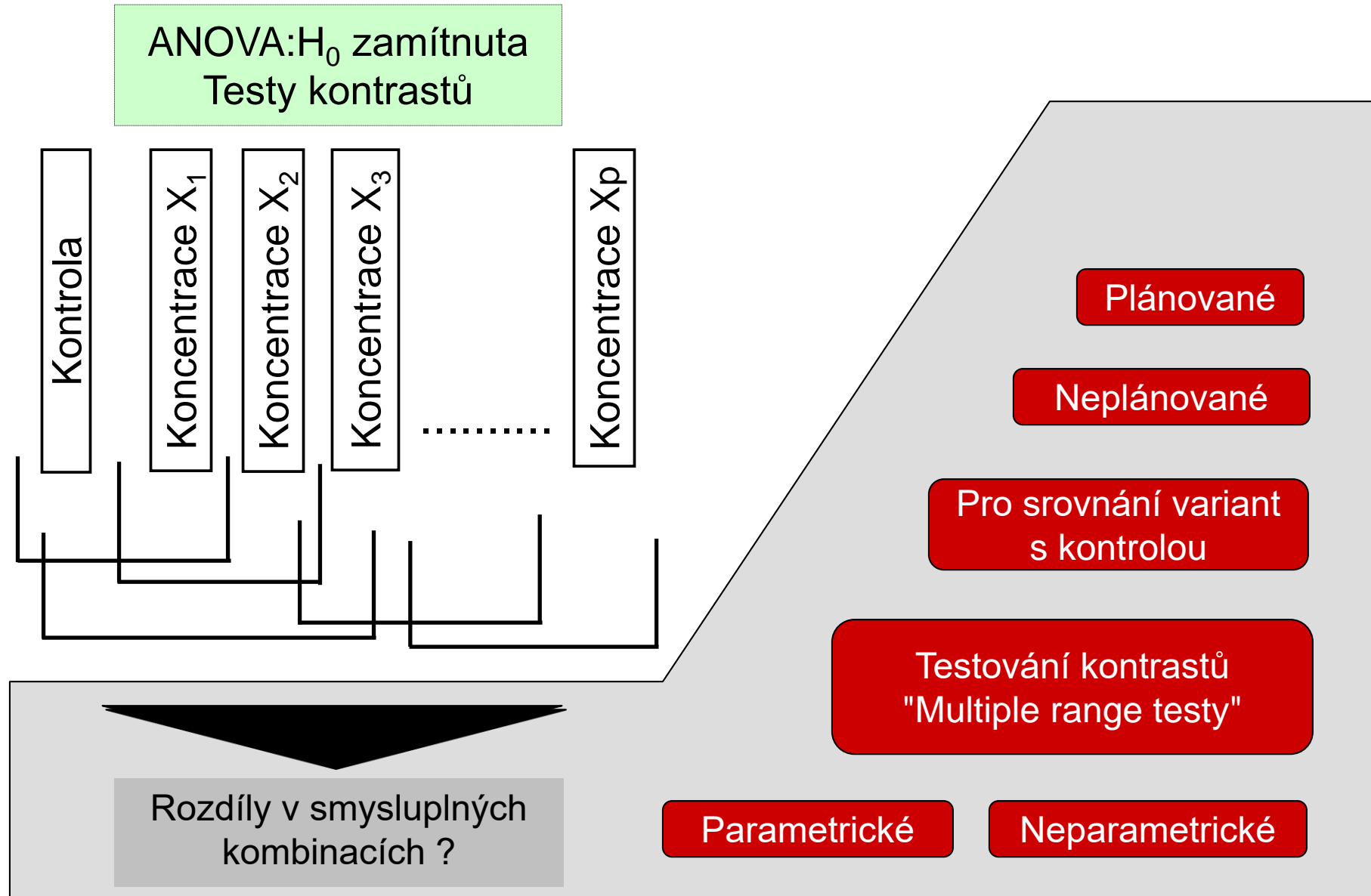
Testování dílčích hypotéz

- V řadě analýz je třeba pracovat se vzájemným testováním více skupin objektů stylem každý s každým
- Obecný postup analýzy je
 - Testování celkové významnosti – všechny skupiny navzájem (ENG: among groups)
 - Pokud je zjištěna celková významnost pokračuje testování analýzou již konkrétních kombinací dvojic skupin (ENG: between)
- Problémem je vliv mnohonásobného testování na statistickou významnost testů:
 - Každý jeden test má $\alpha=0.05$ (chyba I. druhu)
 - Při mnohonásobném testování stoupá pravděpodobnost, že alespoň u jednoho testu dojde k chybnému zamítnutí nulové hypotézy (tedy k chybě I. druhu)

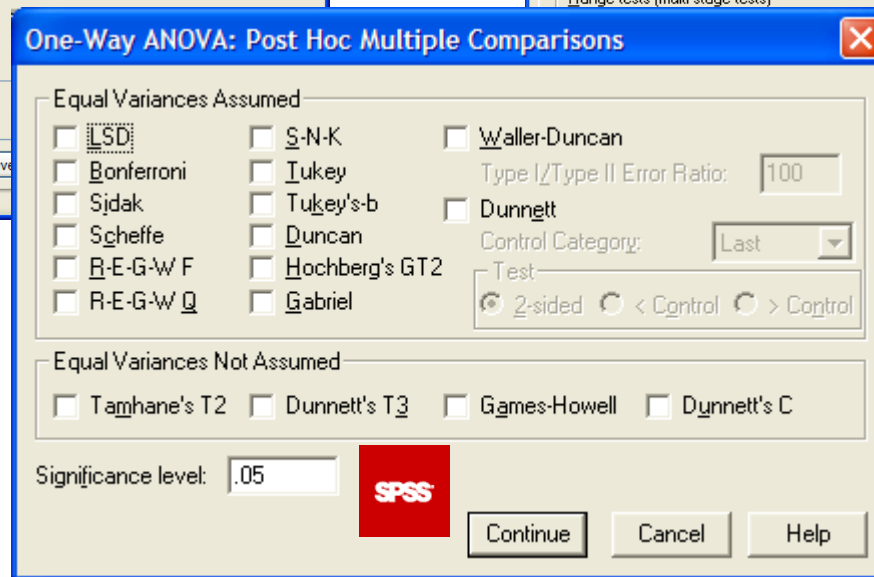
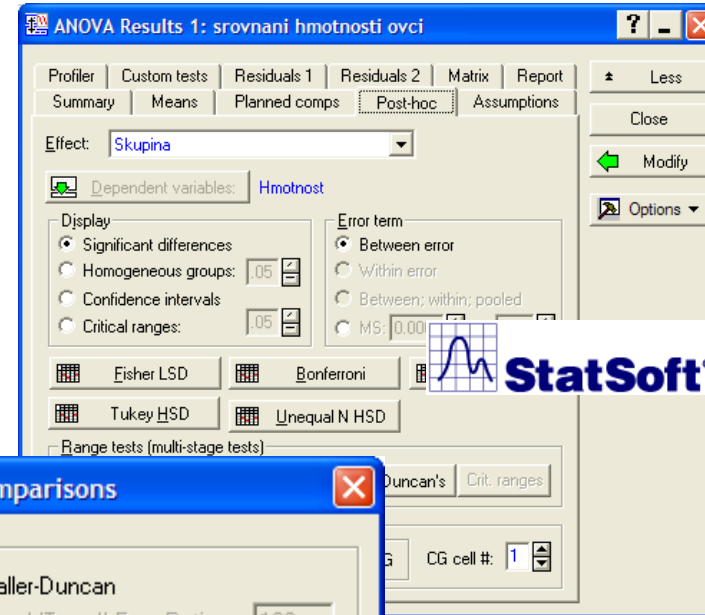
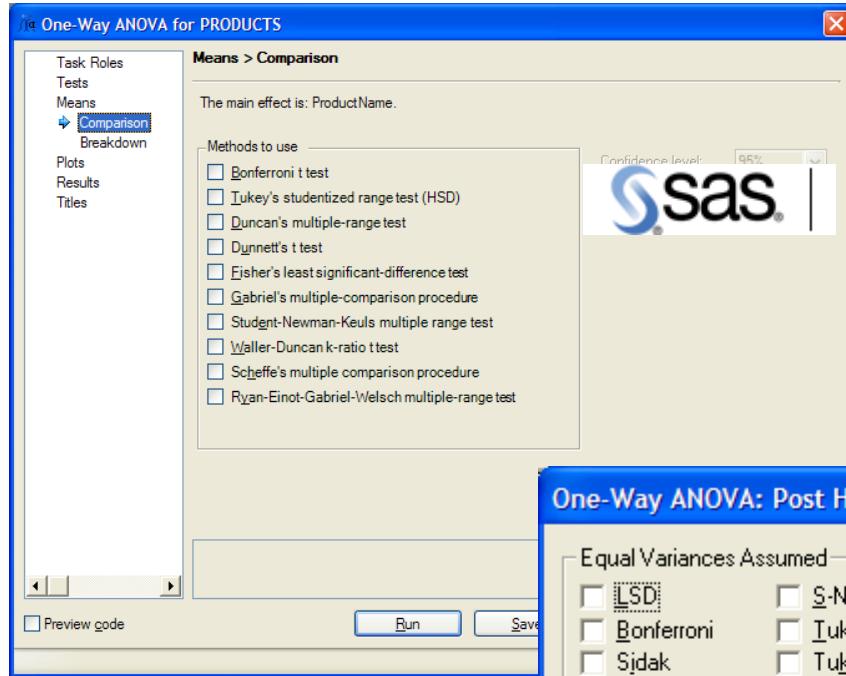


Řešením jsou různé procedury korigující hodnotu p (např. Bonferroniho korekce, FWR, FDR procedury apod.)

Analýza rozptylu - testy kontrastů



Řada různých post-hoc testů



Příklad: Anova - One way

Dávka rostlinného stimulatoru (0, 4, 8, 12 mg/l)

A = 4 ; n = 8

I. ANOVA

Bartlett's test: P = 0,9847

K-S test: P = 0,482 - 0,6525 pro jednotlivé kategorie

Source	D.f.	SS	MS	F	p
Between	3	305.8	101.9	8.56	<0.001
Within	28	322.2	11.9		
Total	31	638			

II. Multiple Range Test (NKS –test)

Level	Average	Homogeneous groups		
0	34.8	x		
4	41.4		x	
12	41.8		x	
8	52.6			x