



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

**Modifikace profilu absolventa biologických studijních oborů na PŘF UP:  
rozšíření praktické výuky a molekulárních, evolučních a cytogenetických oborů**

CZ.1.07/2.2.00/28.0158

# Biostatistika III.

## **Úvod do statistického usuzování**

Testování statistických hypotéz  
(klasická frekvenční statistika)

Statistické testy pro 1, 2 a více souborů, ANOVA

**Martin Duchoslav**

*Katedra botaniky PŘF UP*

*Olomouc 2012-2015*

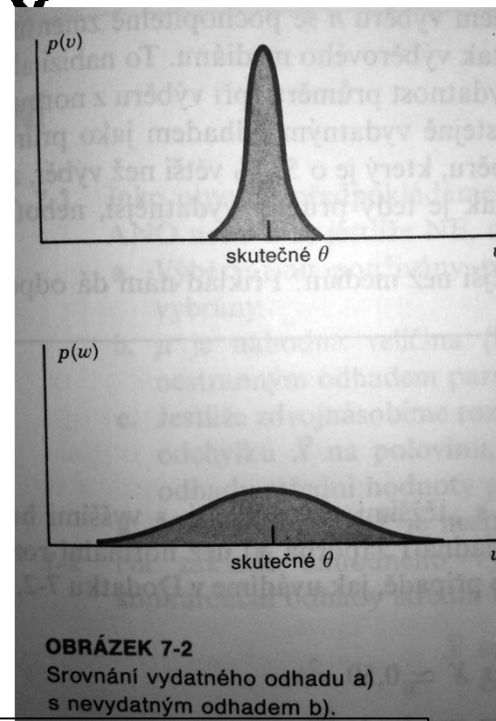
# Statistické usuzování

- provedení zobecnění z náhodného výběru na populaci
- existuje řada různých postupů, jak provést statistické usuzování, v rámci dvou typů:
  - **odhadování parametru** buď **bodově** nebo **číselným intervalem**
  - **statistické testy**
- techniky vycházejí z výběrových rozdělení, vyžadují určitý pravděpodobnostní model dat

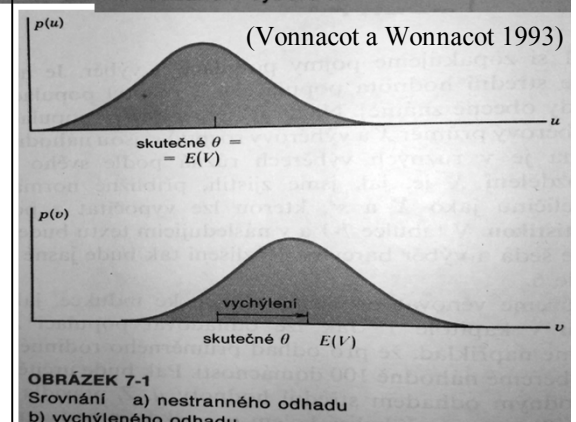
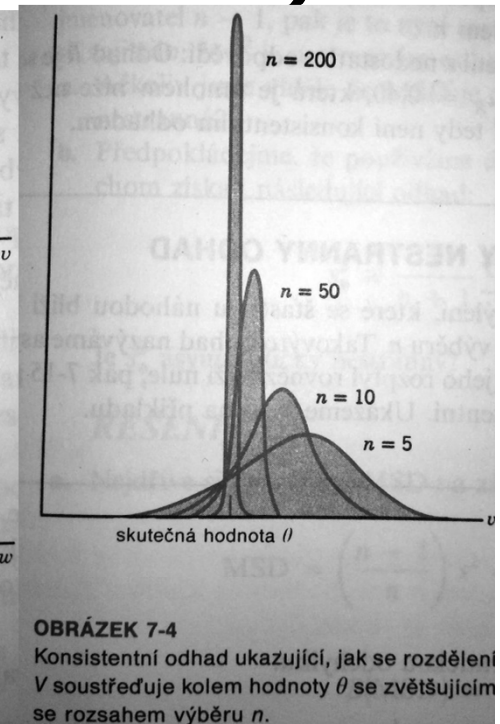
Popisná statistika	Statistické usuzování
Zkoumáme data z různých úhlů (těžíme data)	Odpovídáme na otázku definovanou ještě před sběrem dat
Závěry platí jen na výběr	Generalizujeme závěry na populaci
Závěry jsou neformální, založené jen na datech	Závěry jsou formální, uvádíme jejich spolehlivost

# Odhadování parametru: bodový odhad (*point estimate*)

- jednou hodnotou se snažíme co nejlépe odhadnout popul. parametr
- kvalitu odhadu posuzujeme dle hledisek:
  - **Konzistence** = konzistentní odhad je takový, který se soustřeďuje ve stále užších a užších hranicích kolem hledané hodnoty s rostoucím  $n$
  - **Nestrannost** =  $V$  je nestranným odhadem parametru  $\Theta$  pokud  $E(V) = \Theta$  (při opak. výběrech kolísá odhad symetricky kolem teoret. hodnoty)
  - **Vydatnost** = **eficience** = odhad je vysoce koncentrovaný, tj. s malým rozptylem
  - **Rezistence** = odlehlé hodnoty nemají vliv na hodnotu odhadu



Symetrické rozdělení, odhad  $\mu$ : aritmetický průměr nejvydatnějším odhadem, medián horším, ale ne vždy  $\rightarrow$  vliv odlehlých pozorování (medián je rezistentnější, viz Laplaceovo rozdělení)

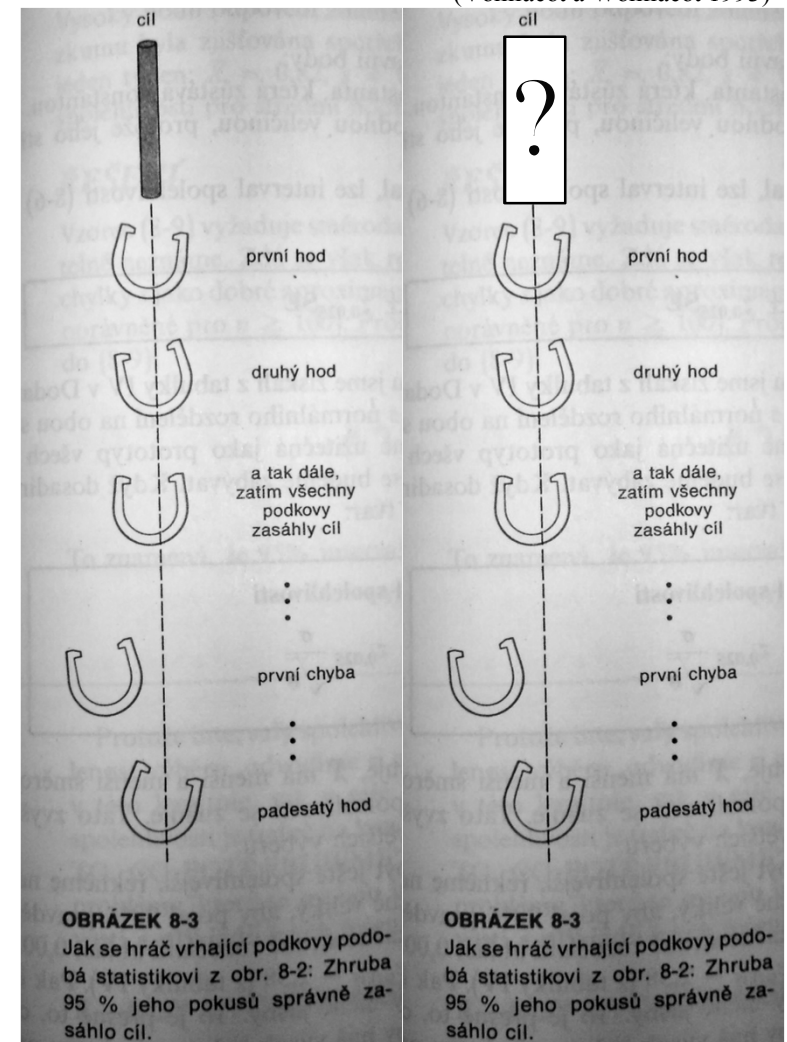


# Odhadování parametru: interval spolehlivosti (*confidence interval*)

(Vonnacot a Wonnacot 1993)

## Interval spolehlivosti

- reprezentuje rozsah hodnot (interval) kolem bodového odhadu, který s předem zvolenou pravděpodobností (např. 95%) zahrnuje parametr populace při opakovaném provádění výběru
- šířka charakterizuje přesnost odhadu parametru a poskytuje informaci o velikosti rozdílu či odchylky od normy + poskytuje informaci k provedení testu stat. hypotézy („množina přijatelných hypotéz“) → podrobný rozbor u kapitoly t-test





# Testování statistických hypotéz

Jak je možné a jakým způsobem odvodit z výsledků zjištěných na jednom výběrovém souboru informaci o celé populaci?






- Statistická hypotéza se týká základního souboru, který neznáme.
- Úkolem matematické statistiky je rozhodnout na základě  $1-n$  výběrových souborů o „platnosti“ určité hypotézy
- Toto rozhodování se provádí tzv. testem statistické hypotézy
  - jedná se rozhodovací pravidlo, které každé realizaci náhodného výběru přiřadí jedno ze dvou rozhodnutí:

**zamítnutí nebo nezamítnutí hypotézy**

*Zakladatelé: R. Fisher, J. Neyman, E. Pearson*

# Fisher *versus* Neyman & Pearson: dva rozdílné přístupy k testování hypotéz

Fisher (1954, 1956): testování signifikance (významnosti)

1. Formulujeme nulovou hypotézu ( $H_0$ )  Co to je  $H_0$ ?
2. Provedeme experiment (pozorování) a rozhodneme na základě hypotézy o
3. volbě testové charakteristiky (kriteria, *test statistics*).  Co to je kriteria?
4. Zjistíme tzv. P-value a rozhodneme o platnosti  $H_0$   Co to je P-value?

# Co to je statistická nulová hypotéza?

Obecně se preferuje nejjednodušší vysvětlení (*viz Princip parsimonie*) → nejčastěji se pak definuje nulová hypotéza jako:

→ „**Není rozdíl mezi...**“ nebo „**X a Y se rovnají...**“

(např. „*Není rozdíl mezi délkou života kuřáků a nekuřáků*“)

to ale ve skutečnosti znamená:

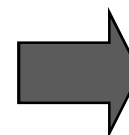


→ „**Rozdíly mezi skupinami nejsou větší než by bylo možné očekávat v důsledku náhodné variability.**“

Takové tvrzení nazýváme statistická nulová hypotéza ( $H_0$ ) (*null hypothesis*), protože hypotéza vylučuje jakoukoliv další sílu (= vliv) mimo náhodnou variabilitu.

# Kdy zamítám nulovou hypotézu?

**Nulovou hypotézu zamítám,  
pokud dostanu uspořádání dat,  
které je velmi nepravděpodobné  
za předpokladu platnosti  
nulové hypotézy!**



*pokračuj...*

# Co to je testová statistika (testové kritérium)?

- většinou se jedná o specifický vzorec, do kterého dosazují numerické hodnoty charakteristik zjištěných na výběru
- **vzorce jsou navrženy tak, že při platnosti  $H_0$  má testové kritérium rozdělení pravděpodobnosti shodné s některým z teoretických modelů** (např. normální rozdělení, Studentovo t- rozdělení, F rozdělení, aj.)
- pro teoretické modely jsou sestaveny kvantilové tabulky

# P-hodnota a „signifikantní výsledek“ (Fisherův přístup)

**Pravděpodobnost** výskytu uspořádání dat z provedeného experimentu (či pozorování) a extrémnějšího pak měří tzv. hodnota významnosti

$P = P\text{-value}$  → *pravděpodobnost, že pozorujeme naše data nebo [teoreticky] data stejně či více extrémní za předpokladu pravdivosti nulové hypotézy:  $P = P(\text{data} / H_0)$ .*

Na čem závisí  $P$  ?

(1) na velikosti výběru (s rostoucím  $n$  klesá), (2) na velikosti odchylky skutečné hodnoty parametru od hodnoty předpokládaného hypotézou (čím dále, tím je menší) a (3) na variabilitě uvnitř výběrů (menší variabilita → menší  $P$ )

Kdy je  $P$  dostatečně malé? Čím menší, tím silnější důkaz proti  $H_0$ .

**Konvencí** tehdy § →

když  $P \leq 0,05$  §§ – pak zamítám  $H_0$  a říkám, že výsledek je statisticky „signifikantní“ (= významný)

když  $P > 0,05$  – pak nezamítám  $H_0$  a říkám, že výsledek je statisticky „nesignifikantní“ (= nevýznamný)

§ Obecně neexistuje žádný důvod proč volit právě takovou hodnotu. Jde pouze o konvenci. Alternativně se volí hraniční hodnoty  $P = 0,01$  a  $P = 0,001$ . Dosažená hladina významnosti testu se často označuje hvězdičkami: \*  $0,01 < P \leq 0,05$ , \*\*  $0,001 < P \leq 0,01$ , \*\*\*  $P \leq 0,001$

§§ „Either an exceptionally rare chance has occurred or the theory is not true“ (Fisher 1959, p. 39)

# Neyman a Pearson (1928, 1933): testování hypotéz

1. Formulujeme nulovou hypotézu\* ( $H_0$ ) a k ní alternativu ( $H_A$ ). [ $H_A$  musí platit, když zamítneme  $H_0$ ]
2. Zvolíme riziko (hladinu významnosti = *significance level*,  $\alpha$ ) v soulase s důsledky, které by mohla mít chyba I. druhu.
3. Zvolíme rozsah výběru.
4. Provedeme experiment (pozorování) a rozhodneme na základě hypotézy o volbě testové charakteristiky (kriteria, *test statistics*,  $T$ ). V tabulkách nalezneme kritickou hodnotu (*critical value*,  $k_\alpha$ ) testové charakteristiky na zvolené hladině významnosti.

Jestliže  $|T| \leq k_\alpha$ , *nezamítneme* nulovou hypotézu (Pozor !  $H_0$  nemusí být pravdivá !!!\*\*). Jestliže  $|T| > k_\alpha$ , *zamítneme (reject)* nulovou hypotézu a říkáme, že platí  $H_A$ .      \*\* „Absence důkazu není důkazem absence“  
(C. Sagan)

5. Spočítáme silofunkci (sílu, *power*) testu (je-li to možné) a rozhodneme, zda pravděpodobnost s níž zamítáme nulovou hypotézu když neplatí je dostatečně velká, tedy pravděpodobnost chyby II. druhu je dostatečně malá.




# Co to je statistická alternativní hypotéza ?

Alternativní (statistická) hypotéza  $H_A$  (*alternative hypothesis*)

(ve smyslu Neymana a Pearsona):

→ buď explicitně definována jako „ne  $H_0$ “ (nejčastěji, pokryje tak nejvíce alternativ) nebo konkrétní hodnotou/intervalem

→ vědecká hypotéza se dostává do pozice alternativní hypotézy, předpokládající nenulový účinek (ale nemusí to být pravidlem!)

→ statistická hypotéza se zabývá strukturou dat, ne mechanismem, který ho způsobil (často je náš pokus černou skříňkou – *black box*) → **výzkumník musí v dalším kroku usoudit na mechanismus z pozorovaného výsledku** 



# Hladina významnosti (Neyman-Pearsonův přístup)

**Hladina významnosti ( $\alpha$ )**, např.  $\alpha=0,05$ :

proporce chybně zamítnutých  $H_0$  za užití pravidla

„když  $P \leq \alpha$ , pak zamítám  $H_0$ “, **pokud by se experiment opakoval mnohokrát a  $H_0$  byla pravdivá** (např. 5 ze 100 při  $\alpha=0,05$ )

*P*-value dle těchto autorů jen říká, zda-li se zamítá nebo nezamítá  $H_0$  na námi zvolené hladině významnosti ( $\alpha$ ) ... nic víc... (Oakes 1986)

# Možná rozhodnutí při testování statistických hypotéz

Skutečnost	Rozhodnutí statistického testu	
	Nezamítneme $H_0$	Zamítneme $H_0$
$H_0$ je platná	Správné rozhodnutí	<b>Chyba I. druhu</b>
$H_0$ neplatí	<b>Chyba II. druhu</b>	Správné rozhodnutí

# Chyba I. Druhu (*Type I error*)

- je spojena se zamítnutím nulové hypotézy, která ve skutečnosti platí
- její pravděpodobnost se nazývá a značí  $\alpha$  = RIZIKO =  
*je to pravděpodobnost, že se dopustíme chyby I. druhu falešným zamítnutím nulové hypotézy* [=  $P(\text{chyba I. druhu} | H_0 \text{ platí})$ ]

*= falešně pozitivní výsledek; = producer error*

- $1 - \alpha$  = spolehlivost testu
- velikost této chyby se volí malá (obvykle  $\alpha = 0,05; 0,01; 0,001$ ) a její hodnotu si stanovujeme před testováním

$\alpha = P ? \rightarrow$  pozor: ačkoliv se  $P$  a  $\alpha$  často ztotožňují, jejich význam je odlišný

více např. Schervish MJ. (1996), Hubbard R. & Bayarri MJ. (2003)

# Chyba II. Druhu (*Type II error*)

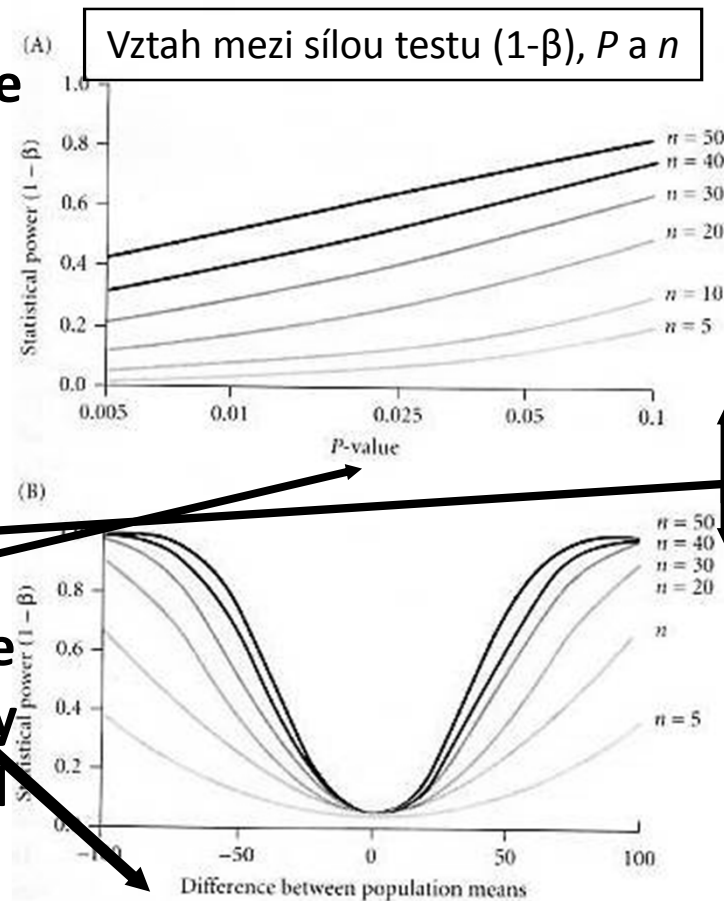
- je pravděpodobnost nesprávného přijetí nulové hypotézy

Značí se  $\beta$  = falešně negativní výsledek; = consumer error

[= P(chyba II. druhu |  $H_0$  neplatí)]

- $1-\beta$  = síla testu = jedná se o pravděpodobnost, že zamítneme nulovou hypotézu, která ve skutečnosti neplatí (měla by být alespoň 0,8, tj. pak máme dobrou šanci detekovat signifikantní rozdíl v datech, pokud je tento přítomný)

- $\beta$  závisí na velikosti výběru (klesá), velikosti  $\alpha$  (se snižující se  $\alpha$  roste  $\beta$ ) a na velikosti odchylky skutečné hodnoty parametru od hodnoty předpokládaného hypotézou (čím dále, tím je nižší)



**Figure 4.5** The relationship between statistical power,  $P$ -values, and observable effect sizes as a function of sample size. (A) The  $P$ -value is the probability of incorrectly rejecting a true null hypothesis, whereas statistical power is the probability of correctly rejecting a false null hypothesis. The general result is that the lower the  $P$ -value used for rejection of the null hypothesis, the lower the statistical power of correctly detecting a treatment effect. At a given  $P$ -value, statistical power is greater when the sample size is larger. (B) The smaller the observable effect of the treatment (i.e., the smaller the difference between the treatment group and the control group), the larger the sample size necessary for good statistical power to detect a treatment effect.<sup>21</sup>

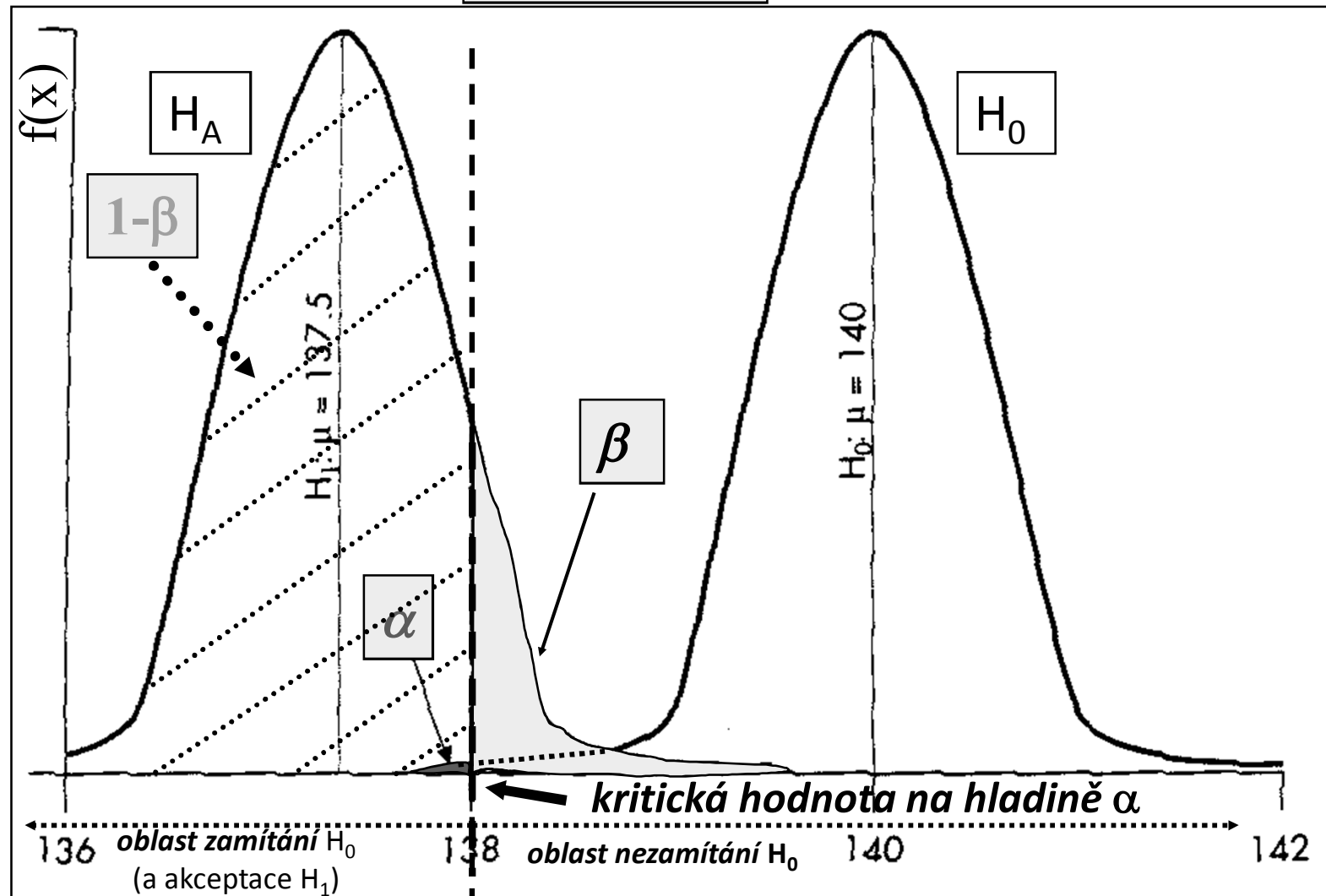
(Gotelli & Ellison 2004, str.103)

# Copak jsou ty chyby ? (II)

\*

$$H_0: \mu = 140,0$$

$$H_A: \mu = 137,5$$

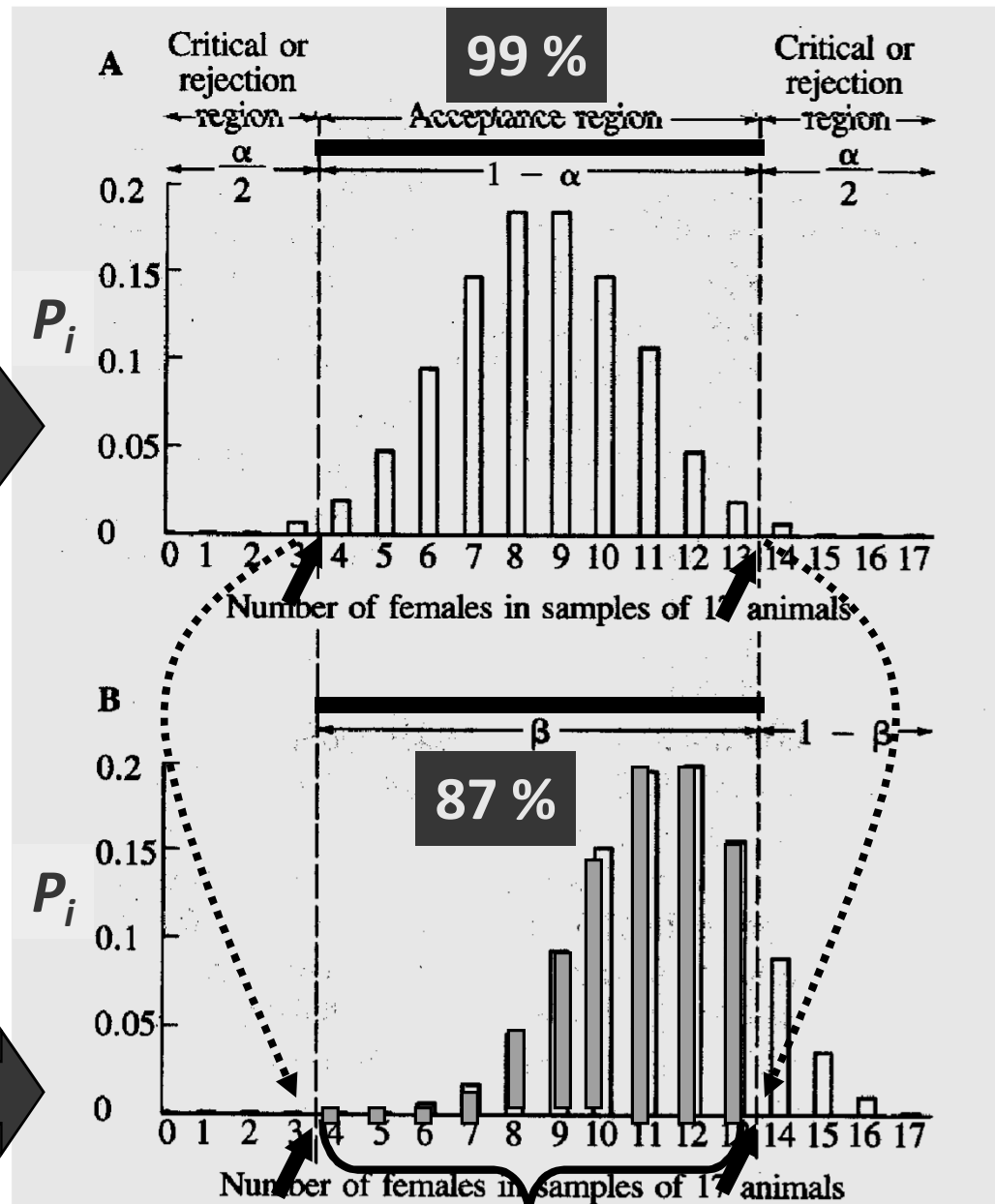


# Vysvětlení chyby II. druhu

$$H_0: p_{\text{♀}} = p_{\text{♂}} = 0,5$$


Očekávané  
pravděpodobnosti výskytů  
různého počtu samic druhu  
X, pokud jsme vybrali  
náhodně 17 zvířat z  
populace druhu X.

$$H_A: p_{\text{♂}} : p_{\text{♀}} = 1:2$$



Suma  $P_i$  hodnot ( )  $H_A$  ohraničených kritickými hodnotami ( / ) v případě platnosti  $H_0$  = pravděpodobnost chyby II. druhu

# Hybridní přístup užívaný v současnosti při testování hypotéz

1. Formulujeme nulovou hypotézu ( $H_0$ ) a k ní alternativu ( $H_A$ ).
2. Zvolíme riziko = hladinu významnosti ( $\alpha$ ) v souladu s důsledky, které by mohla mít chyba I. druhu. 
3. Zvolíme rozsah výběru.
4. Provedeme experiment (pozorování) a rozhodneme na základě hypotézy o volbě testové charakteristiky (kriteria, *test statistics*,  $T$ ). V tabulkách nalezneme kritickou hodnotu (*critical value*,  $k_\alpha$ ).  
Jestliže  $|T| < k_\alpha$ , nezamítneme (*not reject*) nulovou hypotézu.  
Jestliže  $|T| \geq k_\alpha$ , zamítneme (*reject*) hypotézu a říkáme, že odchylky od hypotézy jsou statisticky významné.

**V čem je ta „hybridnost“ ?** Prezентujeme často aktuálně zjištěné hodnoty  $P$  (např.  $P = 0,0123$ ,  $P = 0,00018$ ) jako „sílu důkazu“ proti  $H_0$  (= „*strength of evidence against the  $H_0$* “).

# Vědecká (biologická) versus statistická významnost I

- Velké výběry vykazují menší vliv náhody ➡ i velmi malé populační rozdíly budou signifikantní, pokud je výběr dostatečně velký.
- Malé výběry vykazují větší citlivost na náhodu ➡ i velmi velké populační rozdíly nebudou signifikantní, pokud je výběr malý.

Trial	No. preferring A	No. preferring B	% preferring A
1	15	5	75.0
2	114	86	57.0
3	1,046	954	52.3
4	1,001,455	998,555	50.07

(Hubbard & Lindsay 2008: 77)

P-value je u všech testů stejná = 0.041

- **P-hodnota (statistická významnost, signifikance) neříká, zda-li je efekt dostatečně velký, aby byl prakticky (biologicky) významný.**  
**Statistická signifikance není totéž jako praktická (biologická) signifikance.**

- **Je velmi užitečné uvádět konfidenční intervaly pro námi studovaný parametr.** Konfidenční interval udává **velikost účinku**, než jen odpověď, zda-li je rozdíl dostatečně velký, aby to nebylo náhodou.



# Biologická versus statistická významnost II

Hodnocení velikosti účinku (*effect size*; *ES*)

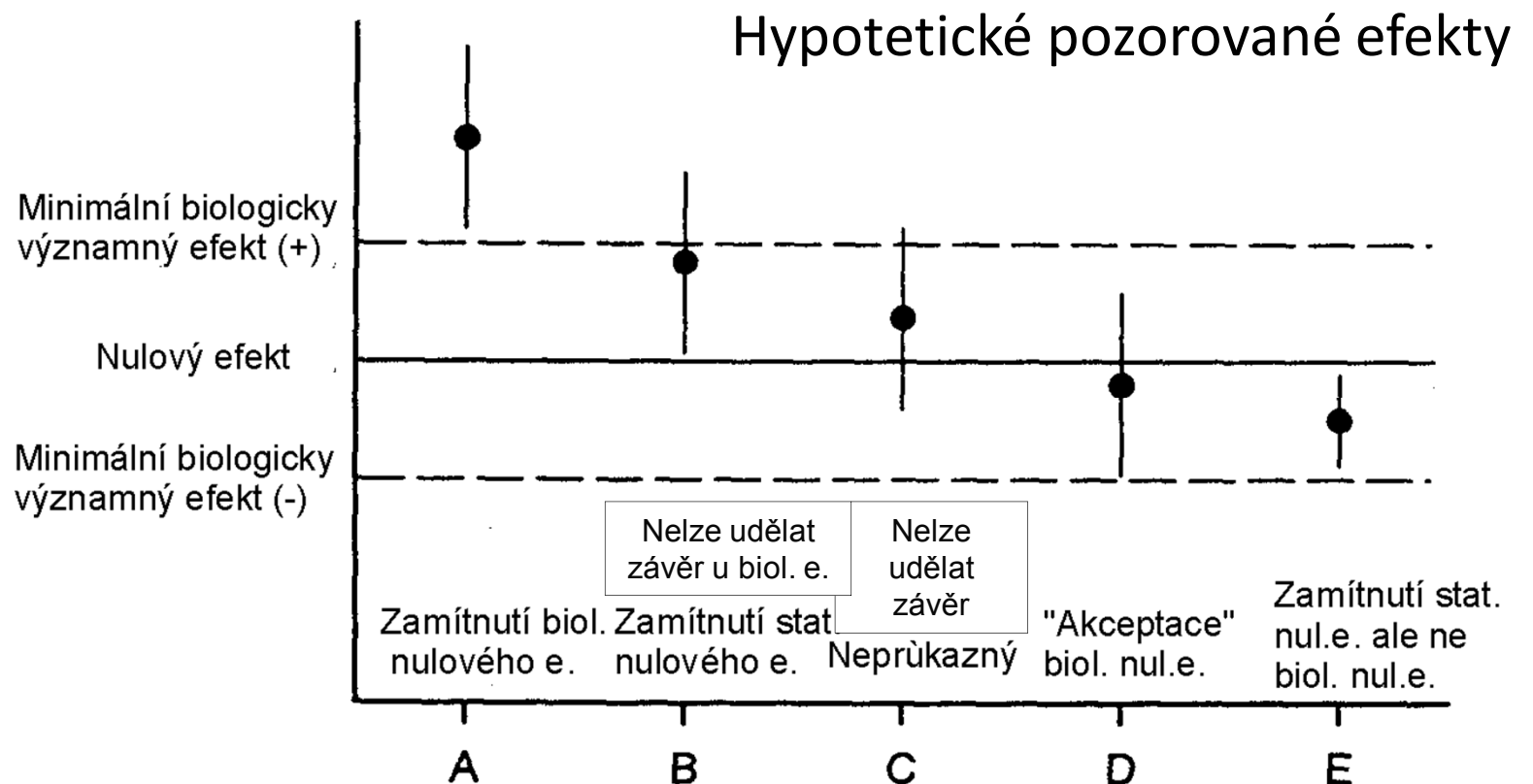
**Cohenův koeficient účinku  $d$**

- jde o standardizovaný rozdíl průměrů;  $|d| > 0.8$ ...velký efekt,  $|d| = (0.2, 0.8)$ ... střední e.,  $|d| < 0.2$  ... malý e.

$$*ES = d = \frac{\bar{x}_1 - \bar{x}_2}{s}$$

kde  $s$  = směr. odchyl. rozdílu měření

\* jeden z možných koeficientů, mj. pro dvě skupiny



(Schenier & Gurevitch 2001, upraveno)

(průměry = body, a jejich 95% konfidenční intervaly= úsečky):

*možné výsledky, které mohou v praxi nastat*

# Biologická versus statistická významnost III

Hodnoty  $P$  a konfidenční intervaly pro různě velké výběry ze dvou populací pacientů užívajících buď „placebo“ nebo „lék na snížení krevního tlaku“ při konstantním rozdílu průměrů (10 mm Hg),

$$\sigma=s_1^2=s_2^2= 400 \text{ mm Hg}^2 \text{ a testu } H_0: \mu_2-\mu_1=0$$

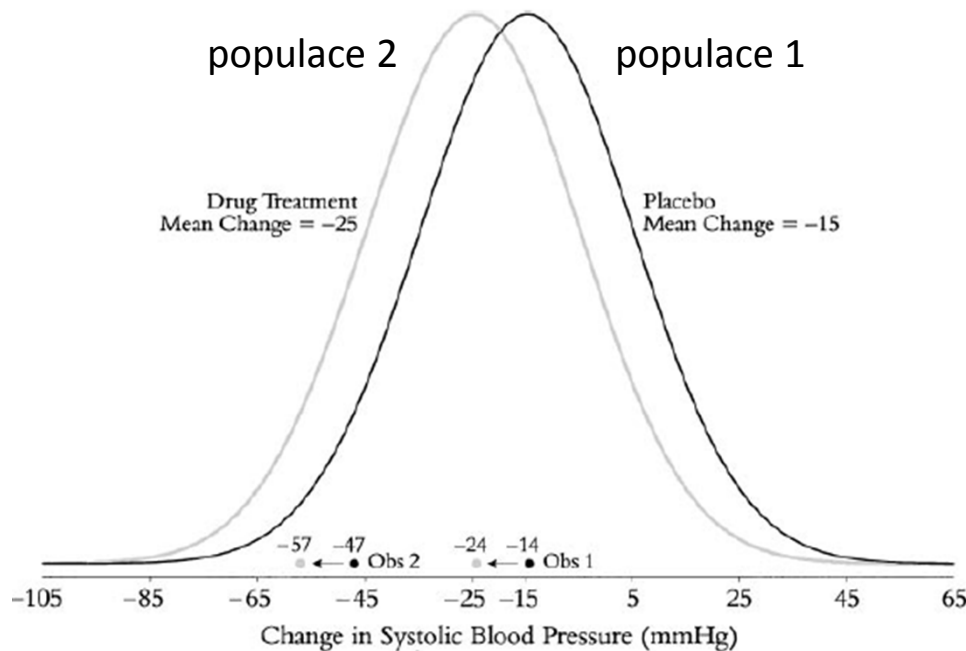


Fig. 5. Statistical and scientific significance differ: placebo (black) and drug-treatment (gray) populations. The populations represent theoretical distributions of changes in systolic blood pressure during year 5 of the Systolic Hypertension in the Elderly Program clinical trial (see Ref. 45). The distributions are described by the normal probability density function (Eq. 1) in which the sample means and variances,  $\bar{y}_i$  and  $s_i^2$ , are substituted for the population means and variances. To generate samples of size  $n$  from each population, observations (Obs) were drawn at random from the placebo population; corresponding observations from the drug-treatment population were obtained by subtracting 10 from each placebo observation. The sampling procedure is illustrated for  $n = 2$ .

**Velikost účinku  
v situaci, kdy  
všechny SD  
budou stejné a  
konstantní (20),  
bude pro všechny  
případy stejná !!!  
 $d=(y_2-y_1)/s=-10/20=(-)0,5$**

Table 2. *Statistical and scientific significance differ: statistical results*

$n$	$\bar{y}_2 - \bar{y}_1$	SE $[\bar{y}_2 - \bar{y}_1]$	95% Confidence Interval*	$t^\dagger$	Pr $[\mu_2 - \mu_1 = 0]^\ddagger$
2	-10	23.1	-110 to +90	-0.43	0.71
4	-10	22.6	-65 to +45	-0.44	0.67
8	-10	12.3	-36 to +16	-0.81	0.43
10	-10	10.1	-31 to +11	-0.99	0.34
15	-10	7.3	-25 to +5	-1.38	0.18
20	-10	5.8	-22 to +2	-1.74	0.09
25	-10	5.3	-21 to +1	-1.88	0.07
32	-10	4.7	-19 to -1	-2.13	0.04
64	-10	3.7	-17 to -3	-2.70	<0.01
128	-10	2.4	-15 to -5	-4.25	<0.001

$n$ , Sample size drawn from placebo (population 1) and drug treatment (population 2) populations (see Fig. 5). \*Confidence interval for the difference between population means,  $\mu_2 - \mu_1$  (see Eq. A2).  $^\dagger$ Test statistic used to evaluate statistical significance of the difference  $\bar{y}_2 - \bar{y}_1$  (see Eq. A3).  $^\ddagger$ Probability (2-tailed) that  $\mu_2 - \mu_1 = 0$ ; this is the significance level  $P$  for the null hypothesis  $H_0: \mu_2 - \mu_1 = 0$ . The difference  $\bar{y}_2 - \bar{y}_1$  and the 95% confidence interval for the difference  $\mu_2 - \mu_1$  reflect the magnitude and uncertainty of the experimental results. The test statistic  $t$  and its associated  $P$  value reflect statistical significance. An increase in the no. of observations drawn from each population decreases SE  $[\bar{y}_2 - \bar{y}_1]$ ; as a consequence, the statistical significance increases (irregularly, because of random sampling), but the estimated difference between population means remains constant at  $\bar{y}_2 - \bar{y}_1 = -10$ . The APPENDIX details the statistical equations required to perform this sampling exercise.

Fundamental concepts in statistics:  
elucidation and illustration

DOUGLAS CURRAN-EVERETT, SUE TAYLOR, AND KAREN KAFADAR  
Departments of Pediatrics and of Preventive Medicine and Biometrics, School of Medicine,  
University of Colorado Health Sciences Center, Denver, 80262; and Department of Mathematics,  
University of Colorado at Denver, Denver, Colorado 80217-3364

Curran-Everett, Douglas, Sue Taylor, and Karen Kafadar. Fundamental concepts in statistics: elucidation and illustration. *J. Appl. Physiol.* 85(3): 775-786, 1998.—Fundamental concepts in statistics form

# Biologická versus statistická významnost IV

## Hodnocení velikosti účinku a meta-analýza- Hedges'g

Cooper, H. & Hedges, L.V. (1994) *The Handbook of Research Synthesis*. Russel Sage Foundation, New York, NY.

*Journal of Applied Ecology* 2005  
42, 261–269

**The effects of organic agriculture on biodiversity and abundance: a meta-analysis**

JANNE BENGTTSSON, JOHAN AHNSTRÖM and  
ANN-CHRISTIN WEIBULL\*

*Department of Ecology and Crop Production Science, Section for Landscape Ecology, SLU, Box 7043, S-750 07 Uppsala, Sweden*

$$g = \frac{\bar{X}_{org} - \bar{X}_{conv}}{S} \times \left(1 - \frac{3}{4m - 1}\right) \quad \text{where } m = n_{org} + n_{conv} - 2.$$

**Meta-analýza =**  
usiluje  
o zobecnění  
výsledků výzkumu  
ze souboru studií  
o konkrétní  
vědecké otázce  
**Příklad:**  
meta-analýza 66  
publikací, které  
porovnávaly vliv  
konvenčního  
a organického  
zemědělství  
na druhovou  
bohatost.  
**Publikační**  
zkreslení  
(*publication bias*)

**Table 1.** A meta-analysis of the effects of organic agricultural methods on species richness. Positive effect sizes indicate higher species richness in organic farming systems. See text for calculation of the Hedges'g statistic. The studies in the meta-analysis are listed in Appendices 1 and 3, and also available from the web site <http://www.cul.slu.se>. Only studies published before December 2002 are included. The number of positive studies column and the associated *n* column include all studies, and those in which no quantitative effect size, only increases or decreases in species richness, was given. When *Q* (heterogeneity of effect sizes among studies) is significant, the results are shown for a random effects model. \**P* < 0.05 for average effect size ≠ 0, for heterogeneity of effect sizes (*Q*) and for binomial test of the number of positive studies; CL, confidence limit

	Average effect size (Hedges'g) (± 95% CL)	<i>n</i>	<i>Q</i>	No. of positive studies	<i>n</i>
Total	1.152 (± 0.524)*	32	170*	53*	63
By organism group					
Birds	1.495 (± 1.236)*	2	0	3	3
Arthropods	0.929 (± 0.589)*	19	71.7*	21*	28
Predatory insects	0.843 (± 0.590)*	15	43.8*	15	21
Carabidae	0.941 (± 0.861)*	11	34.7*	10	13
Non-predatory arthropods	1.046 (± 1.982)	4	26.2*	6	7
Soil organisms	0.306 (± 0.559)	5	3.3	7	10
Plants	2.684 (± 1.976)*	6	81.6*	22*	22
By scale of study					
Plot or single field	2.917 (± 1.769)*	8	55.8*	15*	17
Field on farm	0.703 (± 0.550)*	11	19.3*	24*	27
Field/farm in matched landscape	0.818 (± 0.791)*	13	79.6*	14	19

# Jedna nebo dvě strany?

Obecný tvar hypotézy může mít dvě formy.

Bud' Vás zajímá

- obecně, je-li (specifický) rozdíl mezi populačním parametrem ( $x_1$ ) a hypotetizovanou hodnotou ( $x_0$ )  
➡ = oboustranná hypotéza (testuje se oboustranným testem, *two-tailing testing*)

$$H_0: x_1 = x_0$$

$$H_A: x_1 \neq x_0$$

(= *simple hypothesis, point hypothesis*)

nebo

- je-li specifický *směr* rozdílu mezi populačním parametrem a hypotetizovanou hodnotou  
➡ = jednostranná hypotéza (jednostranný test, *one-tailing testing*)

$$H_0: x_1 \leq x_0$$

$$H_A: x_1 > x_0$$

(či obráceně)

(= *directional hypothesis, composite hypothesis*)

# Výběr směru hypotézy + příklad

Pravidlo:

Je legitimní použít jednostrannou alternativu  $H_A$  pouze pokud je  $H_A$  formulována před analýzou dat.

Příklad:

Zkoumám vliv 2 druhů hnojiv na produkci pšenice.

Pak se mohu ptát:

(A)  $H_0$ : Oba druhy hnojiv mají stejný účinek na výnos.

$H_A$ : Oba druhy hnojiv mají různý účinek na výnos.

nebo

(B)  $H_0$ : První druh hnojiva má stejný nebo menší účinek na výnos.

$H_A$ : První druh hnojiva má větší účinek na výnos.

*(nebo nerovnosti obrátím)*

# Kritická hodnota testu

Kritická hodnota = hodnota kvantilu hraniční pro oblast zamítání  $H_0$  na zvolené  $P$   nebo  $\alpha$  

- V případě *oboustranného testu*:

musíme rozdělit danou hladinu významnosti na dvě části reprezentující dva možné konce distribuce.

Značíme  $k_{\alpha(2)}$ , např.  $t_{0,05(2)}$

- V případě *jednostranného testu*:

uvažujeme pouze jeden konec distribuce a danou hladinu významnosti proto nedělíme.

Značíme  $k_{\alpha(1)}$ , např.  $t_{0,05(1)}$

Platí, že kritická hodnota pro jednostranný test je vždy méně extrémnější než kritická hodnota pro oboustranný test, tj. jednostranný test má větší sílu než test oboustranný !

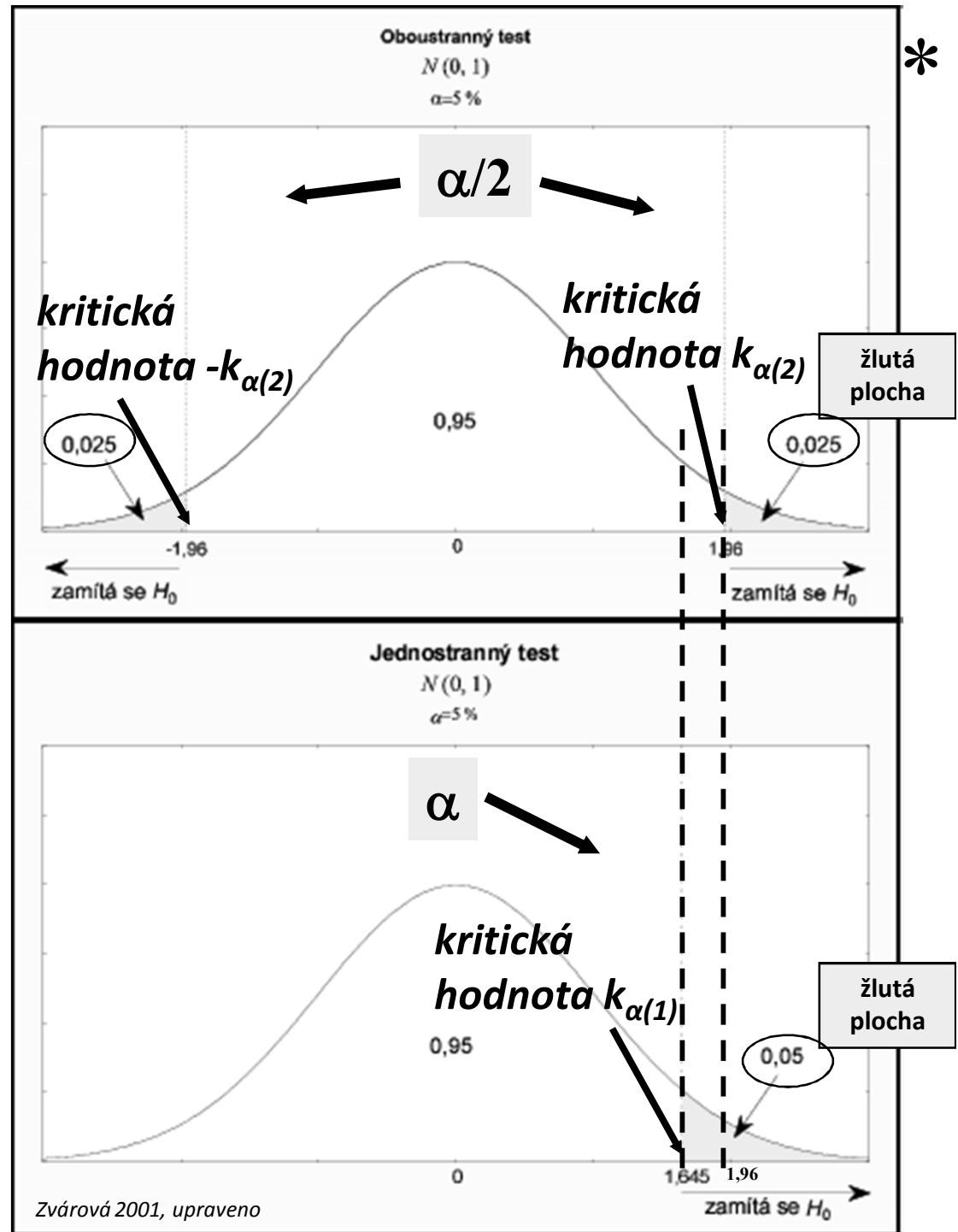
# Jedna nebo dvě strany II

## Oboustranný test

### Příklad:

testuji průměr kvantitativní spojité proměnné (např. výšku, hmotnost) z výběru proti očekané hodnotě při  $\alpha = 0,05$ .

## Jednostranný test



# Kritika testování nulových hypotéz

❑ **běžné používání „klamných, hloupých, banálních“  $H_0$**  (= *false  $H_0$ , trivial  $H_0$ , silly nulls*)

→ ideální  $H_0$  má být taková, že její zamítnutí má logické důsledky, které vedou k lepšímu pochopení studovaného problému

→ bohužel ve většině případů lze platnost  $H_0$  ( $H_0$ : není rozdíl) zamítnout apriorně bez sběru dat („Je těžko představitelné, aby nebyl rozdíl“)

→ důvod užívání „banálních“  $H_0$  („nulový rozdíl“): je mnohdy obtížné vytvořit  $H_0$  s nenulovým efektem: stávající teorie nemusí být dostatečně rozvinutá, aby se taková  $H_0$  mohla SMYSLUPLNĚ formulovat

❑ **hodnota  $P$  je založena i na pravděpodobnosti dat extrémnějších než je výsledek pozorování** (a které nebyly pozorovány!)

→ většina  $H_0$  predikuje unimodální rozdělení pravděpodobností (normální a odvozené od normálního rozdělení) → vliv „nepozorovaných extrémnějších výsledků“ je obvykle malý

❑ **v případě zamítnutí  $H_0$  dochází k „akceptování“  $H_A$  bez toho, jak dobře souhlasí s daty** (→ případ Sally Clark)

❑ **neschopnost zamítnout  $H_0$ , když je  $P$  velké** → příčinou může totiž být špatný design studie, např. malý počet opakování, je důležité provádět stanovení síly testu!



# Další čtení

- Fisher R.A. (1935): The design of experiments.- Oliver & Boyd, Edinburgh.
- Fisher R.A. (1954): Statistical methods for research workers.- Oliver & Boyd, Edinburgh.
- Fisher R.A. (1956): Statistical methods and scientific inference.- Oliver & Boyd, Edinburgh.
- Fox G., Negrete-Yankelevich S., Sosa V.J. (2015): Ecological statistics.- Oxford Univ. Press.
- Hubbard R. & Bayarri M.J. (2003): Confusion over measures of evidence ( $p$ 's) versus errors ( $\alpha$ 's) in classical statistical testing.- *Amer. Statistician* 57: 171-182
- McCarthy M.A. (2007): Bayesian methods for ecology.- Cambridge University Press.
- Neyman J., Pearson E. (1928): On the use and interpretation of certain test criteria for purposes of statistical inference, part I. – *Biometrika* 20A: 175-240.
- Neyman J., Pearson E. (1933): On the problem of the most efficient tests of statistical hypotheses.- *Phil. Trans. Royal. Soc. London, Ser. A* 231: 289-337.
- Oakes M. (1986): Statistical inference: a commentary for the social and behavioural sciences.- Wiley, Chichester.
- Schervish MJ. (1996): P values: what they are and what they are not.- *Amer. Statistician* 50: 203-206.
- Tkadlec E. (2011): Strategie a metody vědecké práce v přírodních vědách.- Olomouc.

Testování statistických  
(nulových) hypotéz

-

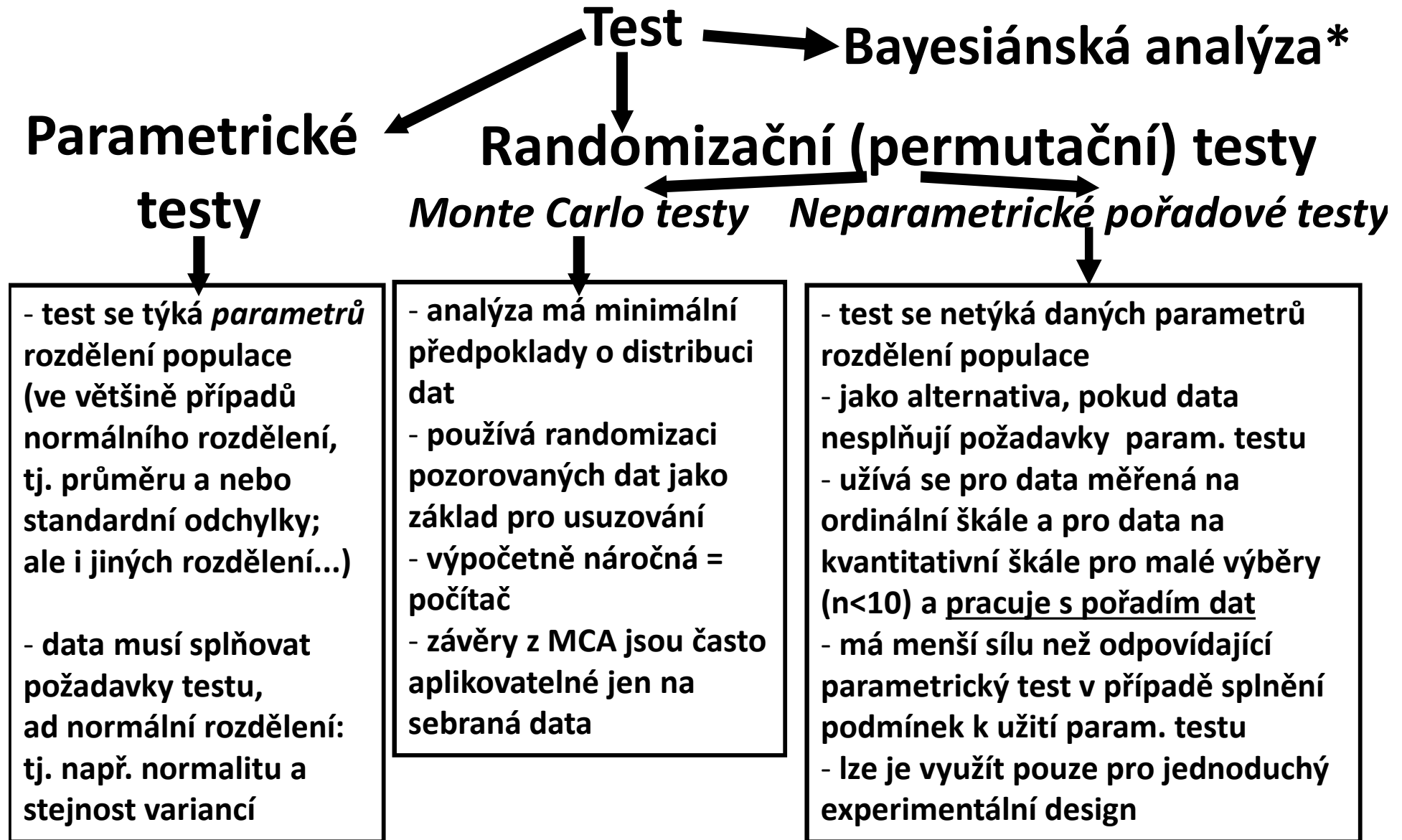
obecné poznámky

# Testování statistických hypotéz - dej si pozor na ...

**Na co si dát při testování pozor aneb jaký test zvolit?**

- je nutno volit testy podle toho, na jaké škále je studovaná proměnná měřena
  - statistické testy mají **specifické předpoklady a jistá omezení**
    - **při jejich nesplnění či nedbání omezení je nesprávné tyto testy používat - získané výsledky jsou (mohou být) nepřesné až chybné a tedy nepravdivé**
    - testovat tutéž hypotézu (parametr) lze často více testy
- ➡ možné problémy s interpretací při různých závěrech těchto testů (**liberární** vs. **kozervativní testy**)

# Základní dělení testů hypotéz



# Postupné kroky při testování statistické hypotézy

## Monte Carlo analýza

(permutační testy)

- 1) specifikuj testovou statistiku
- 2) vytvoř (simuluj) distribuci testové statistiky za předpokladu platnosti  $H_0$
- 3) zvol jedno či oboustranný test
- 4) porovnej pozorovanou statistiku s distribucí simulovaných hodnot a stanov přibližné  $P$

## Parametrická analýza

- 1) specifikuj testovou statistiku
- 2) specifikuj nulovou distribuci
- 3) zvol jedno či oboustranný test
- 4) vypočti  $P$

## Neparametrická analýza

- 1) specifikuj testovou statistiku
- 2) zvol jedno či oboustranný test
- 3) vypočti  $P$

# Permutační testy

Ludbrook J, Dudley H. Why permutation tests are superior to *t* and *F* tests in biomedical research. *Am Stat* 52: 127–132, 1998.

**Příklad:** Liší se koncentrace cholesterolu v krevní plazmě u skupiny lidí stravující se pouze masem z ryby (fish;  $n_1 = 7$ ) versus jiným druhem masa (meat;  $n_2 = 5$ ) po 1 roce experimentu?

*Adv Physiol Educ* 36: 181–187, 2012;  
doi:10.1152/advan.00072.2012.

Staying Current

## Explorations in statistics: permutation methods

Douglas Curran-Everett

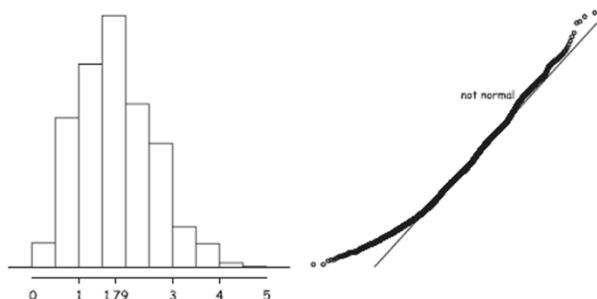
Division of Biostatistics and Bioinformatics, National Jewish Health, Denver, Colorado; and Department of Biostatistics and Informatics, Colorado School of Public Health, University of Colorado Denver, Denver, Colorado

1 Fish	2 Meat
5.42	6.51
5.86	7.56
6.16	7.61
6.55	7.84
6.80	11.50
7.00	
7.11	
6.41	8.20
$\bar{y}_1$	$\bar{y}_2$

Values represent plasma cholesterol (in mmol/l) after 1 year on a diet of fish or meat. Simulated data are from Ludbrook and Dudley (35).

$$t = \frac{\bar{y}_2 - \bar{y}_1}{SE\{\bar{y}_2 - \bar{y}_1\}} = \frac{8.20 - 6.41}{0.8875} = 2.017$$

dvouvýběrový t-test:  $P = 0,104 \rightarrow$  **nezamítám  $H_0$**



Data ale nemají normální rozdělení

**Randomizační (permutační) test:** 1) kolika způsoby můžeme přeuspořádat data (12 hodnot) dvou skupin (=prvků) po 7 a 5 členech (=opakování) (**permutace s opakováním**)?

$$\binom{12}{5} = \frac{12!}{5!7!} = \frac{12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7!}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 7!} = 792$$

2) Kolik případů z těchto 792 uspořádání má rozdíl mezi průměry souborů ( $y_1 - y_2$ ) stejný nebo vyšší (**absolutně**) než je ten pozorovaný (8,2-6,41=1,79)?  $7 \rightarrow P = 7/792 = 0,009 \rightarrow$  **zamítám  $H_0$**

	originalní nasbíraná data					
	Arrangement					
	1	2	3	7	22	57
1 Fish	5.42	5.42	5.42	5.42	5.42	5.42
	5.86	5.86	5.86	5.86	5.86	5.86
	6.16	6.16	6.16	6.16	6.16	6.16
	6.55	6.55	6.55	6.55	6.51	6.55
	6.80	6.80	6.80	6.51	6.80	6.80
	7.00	6.51	6.51	7.00	7.00	7.00
	6.51	7.11	7.56	7.11	7.11	7.11
2 Meat	7.11	7.00	7.00	6.80	6.55	6.51
	7.56	7.56	7.11	7.56	7.56	7.56
	7.61	7.61	7.61	7.61	7.61	7.61
	7.84	7.84	7.84	7.84	7.84	7.84
	11.50	11.50	11.50	11.50	11.50	11.50
$\Delta\bar{y}_j^*$	2.00	1.96	1.80	1.89	1.80	1.79

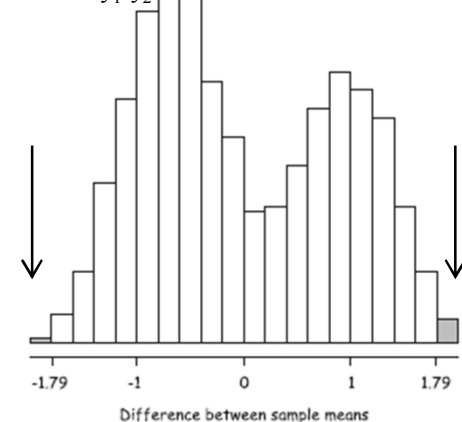
Arrangements of the data in Table 2 for which the magnitude of the difference  $\Delta\bar{y}_j^*$  between sample means is at least as extreme as the observed difference of 1.79. Arrangement 57 holds the observed sample data.

**Příklad** všech možných uspořádání 6 hodnot do dvou skupin po třech a příslušný rozdíl průměrů

původní nasbíraná data	Arrangement	Group 1	Group 2	$\Delta\bar{y}_j^*$
→	1	1 2 3	4 5 6	3
	2	1 2 4	3 5 6	2.3
	3	1 2 5	4 3 6	1.7
	4	1 2 6	4 5 3	1
	5	1 4 3	2 5 6	1.7
	6	1 5 3	4 2 6	1
	7	1 6 3	4 5 2	0.3
	8	4 2 3	1 5 6	1
	9	5 2 3	4 1 6	0.3
	10	6 2 3	4 5 1	-0.3
	11	1 4 5	2 3 6	0.3
	12	1 4 6	2 5 3	-0.3
	13	1 5 6	4 2 3	-1
	14	4 2 5	1 3 6	-0.3
	15	4 2 6	1 5 3	-1
	16	5 2 6	4 1 3	-1.7
	17	4 5 3	1 2 6	-1
	18	4 6 3	1 5 2	-1.7
	19	5 6 3	4 2 2	-2.3
	20	4 5 6	1 2 3	-3

For arrangement *j*, the difference  $\Delta\bar{y}_j^*$  between sample means is  $\Delta\bar{y}_j^* = \bar{y}_2 - \bar{y}_1$ . Arrangement 1 holds the observed sample data.

Rozdělení rozdílů průměrů mezi soubory  $y_1 - y_2$  v 792 randomizacích



# Základní typy uspořádání pokusu

**zcela znáhodněné**  
(*completely randomized*)

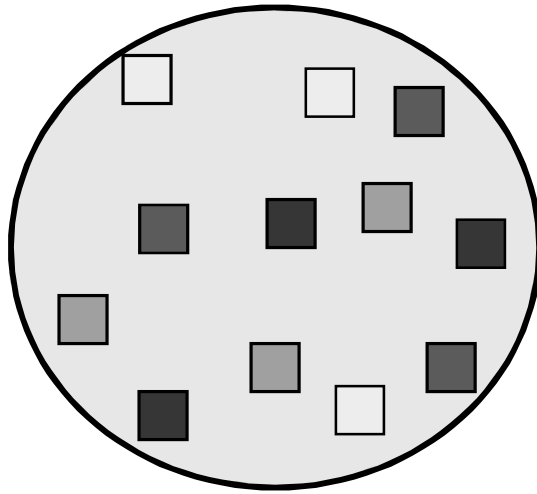
- z populace vybereme náhodně  $n$  jedinců (prvků)
- z takto vybraných  $n$  jedinců (prvků) náhodně vybereme  $n_1$  jednotek, na které aplikujeme první pokusný zásah, ze zbývajících náhodně vybereme  $n_2$  jedinců, na které aplikujeme druhý pokusný zásah atd.

**blokové**  
(*block*)

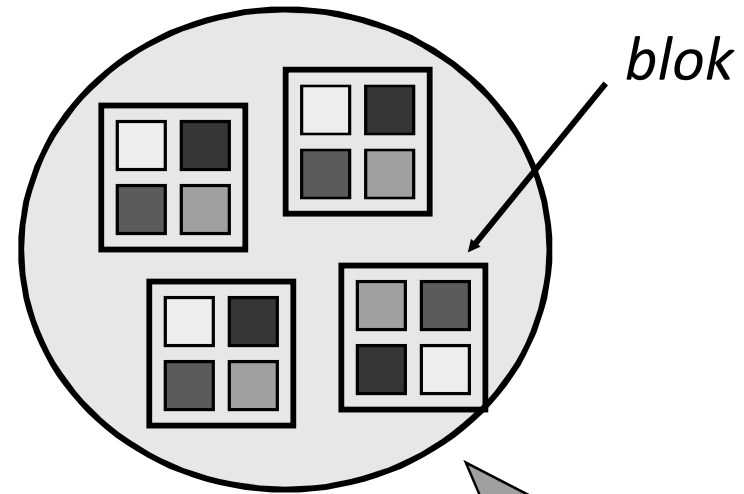
- blok je vytvářen jedinci (prvky), kteří jsou si v nějakém smyslu *blízcí* (prostorově, časově, geneticky atp.)
- blokovým uspořádáním můžeme vyloučit (odfiltrovat) nekontrolovatelné či náhodné vlivy při experimentu (rozdíl mezi bloky nás nezajímá)
- nejjednodušším typem je tzv. párové uspořádání (*paired*)

# Nejtypičtější uspořádání experimentů\*

Příklad: 3 různé zásahy (treatments: ■, ■, □) a kontrola (control: ■)



**zcela znáhodněné**  
**(completely randomized)**



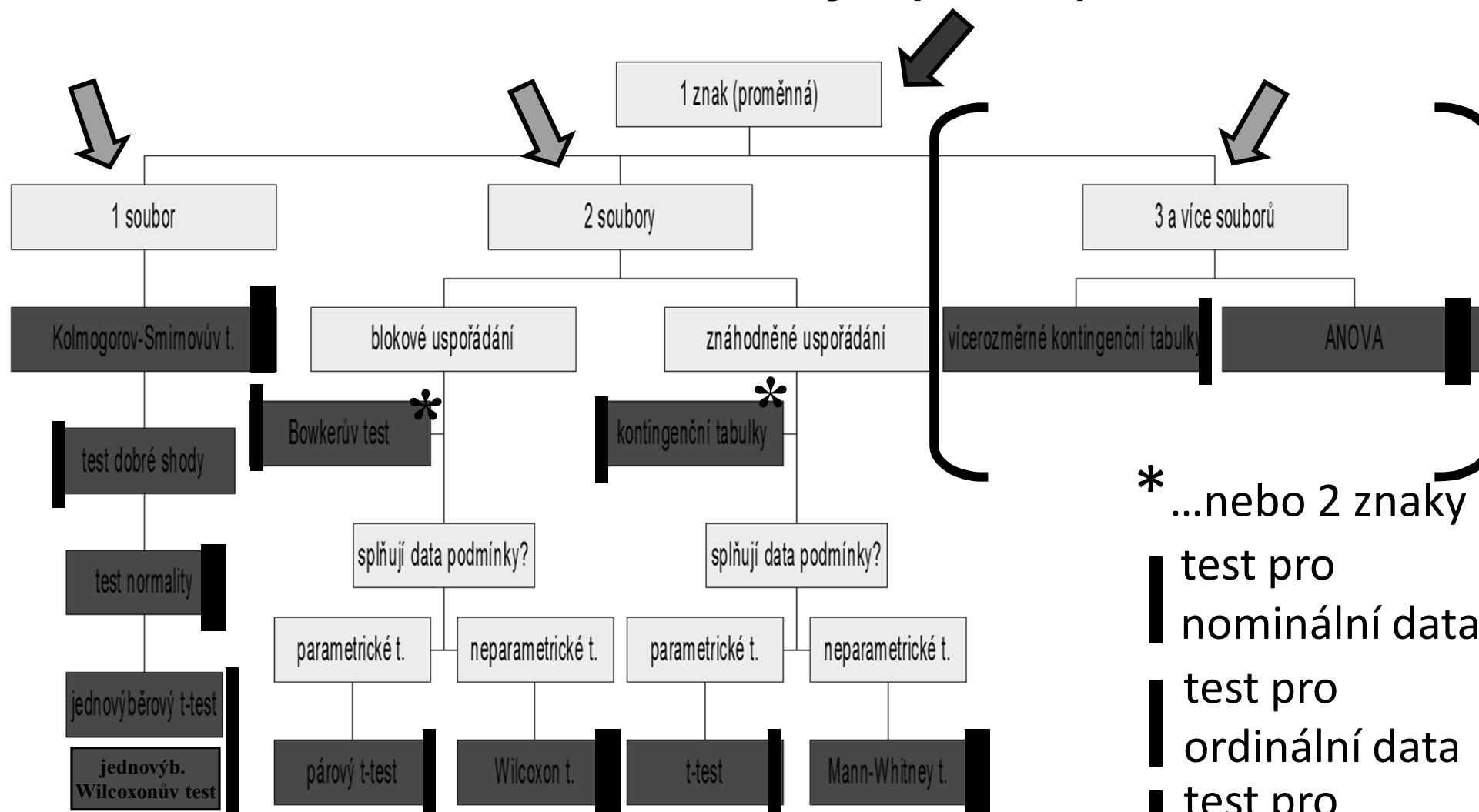
**blokové**  
**(block)**

umožňuje  
dělit  
variabilitu na  
komponenty

*\*existují složitější typy uspořádání - více později*



# Testování hypotéz – rozhodovací strom aneb jaký test použít ?



*Pozn.: schéma reprezentuje výběr nejužívanějších testů pro testování hypotéz o střední hodnotě, shodě dat s očekáváním či testů normality pro 1 znak a 1-2 výběry (výjimky viz výše) - není proto úplné a je možné k němu mít výhrady...:-)*

Testy hypotéz  
týkajících se  
kvalitativních (kategoriálních)  
proměnných

# Analýza frekvencí - test dobré shody (*Goodness of fit*)

Získáme soubor kategoriálních\* dat a přejeme si vědět, zda-li populace, ze které tento soubor pochází, se shoduje s předpokládanou teoretickou distribucí.

V nejjednodušším případě srovnáváme skutečně zjištěné a očekávané četnosti jevů

Sledujeme-li více jak jeden jev, zajímá nás,  
jak se liší pozorované četnosti od očekávaných

(tedy: zda-li je shoda mezi „dosaženým“ a očekávaným“ nebo ne!)

- jako výsledek experimentu (pozorování) sledujeme výskyt náhodného jevu A, o jehož pravděpodobnosti předpokládáme, že je rovna danému číslu  $P(A)$
- pozorovaná četnost jevu A v  $n$  nezávisle opakovaných pokusech je  $f_i$  a relativní četnost pak  $p=f_i/n$

*pak...*

# Test dobré shody

(pro 1 proměnnou s 2 a více kategoriemi) (*Goodness of fit*)

Za míru odchylek „dosažených“ od „očekávaných“ odchylek je považována míra:

$$\chi^2_{k-1} = \sum_{i=1}^k \frac{(f_i - \hat{f}_i)^2}{\hat{f}_i}$$

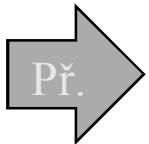
čti: „chí-kvadrát“

$\hat{f}_i$  ... očekávaná absolutní četnost jevu  $i$  v kategorii  $k$

$f_i$  ... skutečně zjištěná absolutní četnost jevu  $i$  v kategorii  $k$

Míra  $\chi^2_{k-1}$  má přibližně Pearsonovo rozdělení o  $n = k-1$  stupních volnosti

**Předpoklady:** 1) pozorování jsou klasifikována do kategorií vzájemně nezávisle.  
2) Ne více jak 20% kategorií má očekávané frekvence nižší než 5.



Je-li vypočítané  $\chi^2$  větší než kritická hodnota  $\chi^2_{(\alpha, n)}$ , zamítáme  $H_0$  o shodě zjištěných a očekávaných frekvencí na zvolené hladině významnosti ( $\alpha$ ).

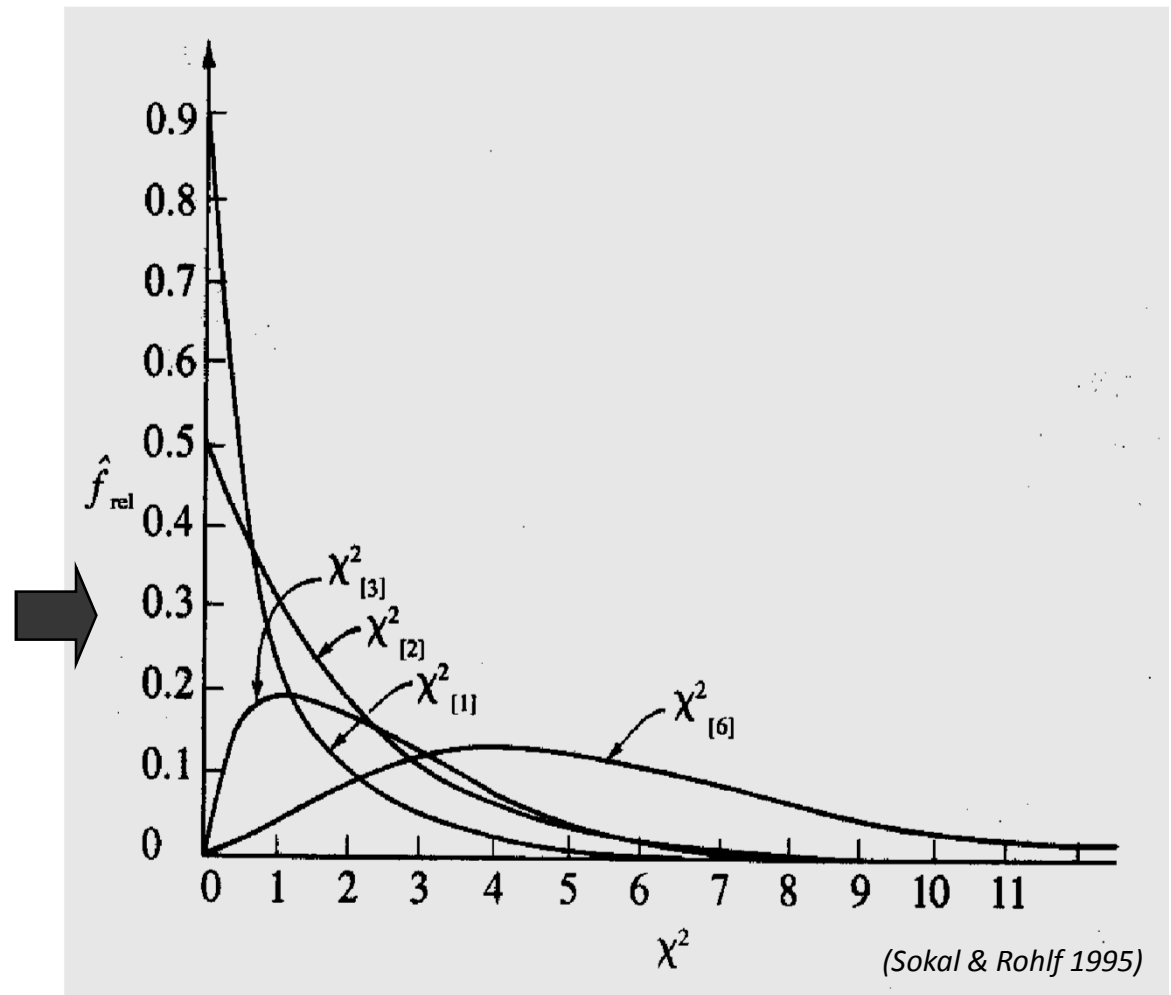
# Jak vypadá Pearsonovo ( $\chi^2$ ) rozdělení ?

-je to funkce, jejíž hodnoty  
kolísají v intervalu  $\langle 0, +\infty \rangle$

- má pouze 1 parametr:

$\nu$  [ný] ... stupně volnosti  
(= počet kategorií – 1)

Hustoty pravděpodobnosti  
pro Pearsonovo rozdělení s  
1, 2, 3 a 6 stupni volnosti.



# Tabulka kritických hodnot $\chi^2$ distribuce

df	$\alpha$		
	0.05	0.01	0.001
1	3.841	6.635	10.827
2	5.991	9.210	13.815
3	7.815	11.345	16.268
4	9.488	13.277	18.465
5	11.070	15.086	20.517
6	12.592	16.812	22.457
7	14.067	18.475	24.322
8	15.507	20.090	26.125
9	16.919	21.666	27.877
10	18.307	23.209	29.588
11	19.675	24.725	31.264
12	21.026	26.217	32.909
13	22.362	27.688	34.528
14	23.685	29.141	36.123
15	24.996	30.578	37.697
16	26.296	32.000	39.252
17	27.587	33.409	40.790
18	28.869	34.805	42.312
19	30.144	36.191	43.820
20	31.410	34.170	37.566

(Lepš 1996)

# Složená vs. jednoduchá $H_0$

Je-li kategorická proměnná



vícestavová



- $H_0$  je složená (*compound  $H_0$* )  
= obsahuje více jak jeden nezávislý výrok
- $H_A$  je nesměrovaná (*omni- nebo  $nondirectional$* )
- v případě zamítnutí  $H_0$  test neposkytuje směrované rozhodnutí

Př.:  $H_0$ : Poměr krevních skupin (A, B, AB, 0) v populaci je 45:20:7:28

$H_A$ : Alespoň jeden poměr se odlišuje od předpokladu

dichotomická (binární)



- $H_0$  je jednoduchá (*simple  $H_0$* )  
= obsahuje právě jeden nezávislý výrok
- $H_A$  může být i směrovaná (jednostranná) (*directional*)
- v případě zamítnutí  $H_0$  test poskytuje směrované rozhodnutí
- lze provést i test jednostranné  $H$  (jdou-li data ve směru  $H_A$ , dělíme  $P$  dvěma)

Př.: viz následující snímek

# Příklad



Genetik získal 100 potomků při křížení a ptá se, zda-li jeho výsledek odpovídá teoreticky očekávanému fenotypovému poměru žlutě a červeně kvetoucích rostlin 3:1. Získal 84:16, ale teoreticky měl získat 75:25. Je výsledek experimentu signifikantně odlišný od předpokladu?

$H_0$ : poměr žlutá:červená odpovídá 3:1

$H_A$ : poměr je jiný (*alternativa*: poměr je více jak 3:1 ve prospěch žluté barvy)

2 kategorie;  $k = 2$

$$\chi^2_{k-1} = \sum_{i=1}^k \frac{(f_i - \hat{f}_i)^2}{\hat{f}_i} = \left( \frac{(84 - 75)^2}{75} \right) + \left( \frac{(16 - 25)^2}{25} \right) = 1,08 + 3,24 = 4,32$$

kritická hodnota na  $\alpha = 0,05$ :

$$\chi_{(\alpha, k-1)}^2 = \chi_{(0,05, 1-1)}^2 = \chi_{(0,05; 1)}^2 = 3,841$$

protože

$$\chi_{(0,05; 1)}^2 = 3,841 < \chi_{test} = 4,32$$



zamítáme  $H_0$

(*pro jednostranný test*: skutečná proporce žlutých rostlin (84 ze 100=84%) je vychýlená ve směru  $H_A$  (oproti předpokladu  $H_0=75\%$ ), a tedy  $P$  dělíme dvěma)





# Test heterogeneity

(více souborů jedné kategoriální proměnné)

Naším cílem může být zjistit, zda-li  
můžeme zkombinovat výsledky jednotlivých (dílčích) experimentů  
(pozorování) do jednoho souboru?!

**$H_0$ : všechny vzorky (soubory) pocházejí z téže populace**



## *Postup:*

1. Spočítáme testy dobré shody pro všechny vzorky zvlášť.
2. Sečteme všechny dosažené a očekávané četnosti pro jednotlivé vzorky a spočítáme test dobré shody pro takto vzniklý jeden soubor.
3. Pokud jsou vzorky homogenní, měly by být obě hodnoty přibližně stejné – rozdíl mezi nimi je též proměnná s Pearsonovým rozdělením.
4. Hodnotu této statistiky porovnáme s kritickou hodnotou na příslušné  $\alpha$  a DF (rozdíl DF výše zmíněných prom.) a rozhodnu  $H_0$

### Příklad:

Mendel experimentoval s křížením hrachu a zabýval se barvou semen. Celkem provedl 10 experimentů s křížením homozygota dominantního (AA; žlutá semena) a recesivního (aa; zelená s.).

Předpokladem bylo, že poměr fenotypů semen získaných křížením bude 3:1 ve prospěch žlutých semen.

$H_0$ : Experimenty jsou homogenní (tj. pocházejí z téže populace).

$H_A$ : Experimenty jsou heterogenní (tj. pocházejí z různých populací)

(Zar 1999, str. 468)

**EXAMPLE 21.5** An example of heterogeneity chi-square analysis. The seed color frequencies resulting from pea breeding experiments of Gregor Mendel (1865).

The null hypothesis for each experiment is that the population sampled has a 3:1 ratio of yellow-to-green seeds.

The null hypothesis for heterogeneity chi-square testing is that all ten samples could have come from the same population (i.e., there is homogeneity).

For each experiment, the observed frequencies,  $f_i$ , are given, with the frequencies predicted by the null hypothesis,  $\hat{f}_i$ , in parentheses.

Experiment	Yellow Seeds	Green seeds	Total seeds (n)	Uncorrected chi-square*	DF
1	25 (27.0000)	11 (9.0000)	36	0.5926	1
2	32 (29.2500)	7 (9.7500)	39	1.0342	1
3	14 (14.2500)	5 (4.7500)	19	0.0175	1
4	70 (72.7500)	27 (24.2500)	97	0.4158	1
5	24 (27.7500)	13 (9.2500)	37	2.0270	1
6	20 (19.5000)	6 (6.5000)	26	0.0513	1
7	32 (33.7500)	13 (11.2500)	45	0.3630	1
8	44 (39.7500)	9 (13.2500)	53	1.8176	1
9	50 (48.0000)	14 (16.0000)	64	0.3333	1
10	44 (46.5000)	18 (15.5000)	62	0.5376	1
Total of chi-squares				7.1899	10
Chi-square of totals (i.e., pooled chi-square)				0.1367	1
Heterogeneity chi-square				7.0532	9

$\chi^2_{0.05,9} = 16.1919$ .

Do not reject the homogeneity null hypothesis.  $0.50 < P < 0.75$  [ $P = 0.63$ ]

\*In heterogeneity analysis, chi-square is computed *without* correction for continuity.

# Kontingenční tabulky (*Contingency tables*)

„Nástavba“ testu dobré shody.

## Užívají se pro zjišťování

a) vztahů mezi 2 a více znaky nominálními (*každý s 2 a více kategoriemi*)

b) vlivu více kategoriálních proměnných (tzv. *prediktorů*) na jednu „odpovědní“ kategoriální proměnnou

(*popř. diskrétních kvantitativních nebo spojitých kvantitativních proměnných s hodnotami sloučenými do skupin !!!*)


## Jaké hlavní typy hypotéz umožňují testovat kontingenční tabulky?

1. Hypotézu o shodnosti struktury (1 znaku ve 2 a více výběrech)
2. Hypotézu o nezávislosti (2 znaků v 1 výběru)
3. Hypotézu o symetrii (2 znaků či opakovaných měření v 1 výběru)

# Kontingenční tabulky - příklady

**Příklad č.1:** Byl studován výskyt mihulí v tocích České republiky. Předběžné výsledky ukázaly, že jejich přítomnost/nepřítomnost v toku není určena současným stupněm znečištění ani znečištěním v minulosti (nelze ale vyloučit jednorázovou intoxikaci). Byly tedy studovány další vlastnosti jednotlivých toků, zvl. mechanické zábrany, které mohou limitovat pohyb kruhoústých a ryb v toku. Toky byly klasifikovány do 2 typů: *a*) s přítomností jezů a splavů zabraňujících zpětnému návratu vodních obratlovců a *b*) bez přítomnosti jezů a splavů. Bylo celkem vyšetřeno 100 toků. Z nich bylo 50 s jezy a 50 bez jezů. Z toků typu *a*) byly mihule nalezeny v 10 případech, v tocích typů *b*) ve 40 případech. Je poměr toků s výskytem/absencí mihulí shodný v obou typech toků (tj. v tocích s bariérami/bez bariér)?

**Příklad č. 2:** Zkoumáme vzájemný výskyt dvou druhů na skalní stepi. Celkem jsme na plochu rozmístili náhodně 100 plošek o rozměru 1x1 m. Na každé ploše jsme zaznamenali přítomnost/nepřítomnost druhu *A* a druhu *B*. Oba druhy se vyskytovaly v 36 čtvercích, ani jeden ve 20 čtvercích, pouze druh *A* se vyskytoval ve 30 čtvercích. Vyskytují se druhy vzájemně nezávisle?

**Příklad č. 3:** Sledujeme skupinu 20 pacientů, kteří byli léčeni dvěma různými hypertenzivy *A* a *B*. Každý pacient dostával po dobu 1 měsíce lék *A* a po odeznění případných účinků po dobu 1 měsíce lék *B*. Výsledek byl klasifikován jako úspěch (tlak snížen o více než 15 mm Hg) či neúspěch. Liší se léky v účinku? 

# Kontingenční tabulky II.

- v mnoha situacích sbíráme data simultánně pro 2 (a více) proměnných a tak by bylo zajímavé zjistit, zda-li frekvence výskytu v různých kategoriích jedné proměnné jsou nezávislé na frekvencích další proměnné ( $= H_0$ )

***Příklad: dvě nominální proměnné (X, Y), každá se dvěma kategoriemi (+, -)***

A, B, C, D = **absolutní četnosti kombinací znaků** = **BUŇKA**,  $n$  = velikost výběru

	Y+	Y-	Celkem	
X+	A	B	A+B	} $R_i$
X-	C	D	C+D	
Celkem	A+C	B+D	$n$	

$$\sum C_j = \sum R_i = n$$

 $C_j$  $C_j$ 

**Marginální součty**  
(marginální distribuce)

# Kontingenční tabulky III.

Jak ale spočítat očekávané frekvence,  
pokud by byly jevy nezávislé?



**Opakování:** pravděpodobnost současného výskytu dvou nezávislých jevů A a B je rovna součinu jejich pravděpodobností, tj.  $P(A) \cdot P(B)$

Obecně pro  
kontingenční  
tabulku:

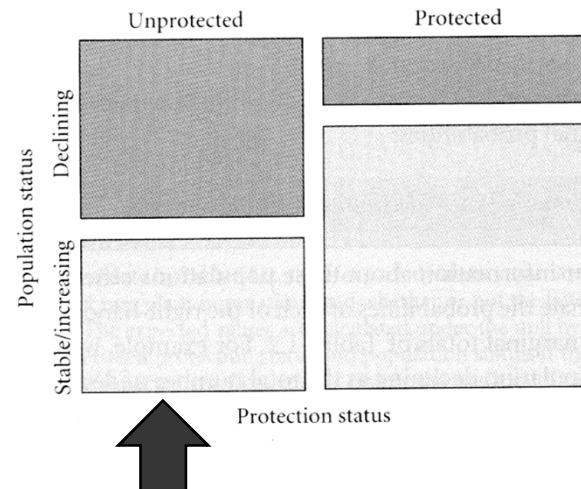
$$\hat{f}_{ij} = \left( \frac{R_i}{n} \right) \left( \frac{C_j}{n} \right) n = \frac{R_i C_j}{n}$$

Test dobré shody pro kontingenční tabulku se pak vypočítá:

$$\chi^2 = \sum \sum \frac{(f_{ij} - \hat{f}_{ij})^2}{\hat{f}_{ij}} = n \left( \sum \sum \frac{f_{ij}^2}{R_i C_j} - 1 \right)$$

se stupni volnosti  $DF = (r-1)(c-1)$

# Grafická vizualizace kontingenční tabulky



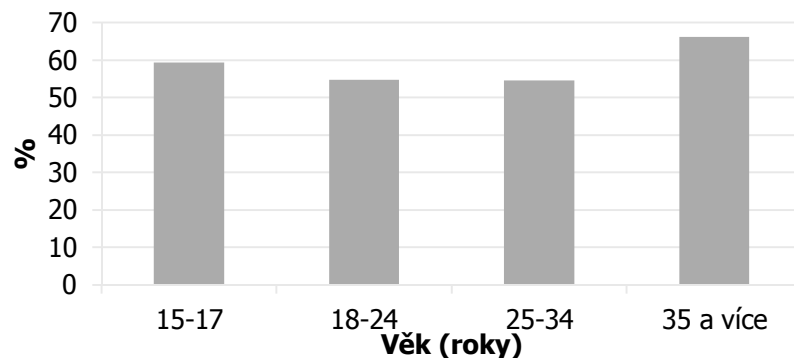
**Figure 11.1** Illustration of a mosaic plot. This plot portrays the relationship between the status of rare plant populations and whether or not the land on which they occur is protected (Farnsworth 2004). The cell frequencies (from Table 11.2) are represented by “tiles” whose size is proportional to their relative frequency in the dataset. The width of the columns is proportional to the column totals. For example, because 40 of the 73 populations are protected, the protected column occupies 55% (or 40/73) of the width of the plot. The height of each tile is proportional to the cell frequency. Thus, 32 of the 40 protected populations are stable or increasing, so this tile occupies 32/40, or 80%, of the right-hand column.

(Gotelli & Ellison 2004)

**Mozaikový diagram:** frekvence buněk jsou reprezentovány „dlaždicemi“, jejichž plocha je proporční jejich relativní frekvenci v souboru

- šířka sloupečku je proporční jeho podílu na celku (viz  $C_j$ )
- výška každé dlaždice je proporční frekvenci buňky [např.  $A/(A+C)$ ]

% žen studujících  
vysokou školu  
ve 4 věkových kategoriích



**Sloupcový diagram relativních frekvencí:**

srovnání **podmíněných** (%) četností jedné proměnné [např.  $A/(A+B)$ ] pro separátní kategorie druhé proměnné

(Ize-li určit směr vlivu, pak na ose x vynášíme kategorie vysvětlující proměnné a na ose y četnosti kategorie (i) vysvětlované proměnné)

--- tzv. *Row percents* nebo *Column percents*

# Kontingenční tabulky:

## 2x2 tabulky (= *čtyřpolní tabulky*)

- jde o nejjednodušší typ kontingenčních tabulek s  $r=2$  a  $c=2$  a s  $DF=1$

Lze rozlišit celkem 3 typy experimentálního designu, který vede k analýze prostřednictvím čtyřpolních tabulek:

### A. Kategorie 1

- je fixováno  $n$ , ale nejsou fixovány marginální součty
- pro analýzu užíváme Yatesovu korekci\* nebo výpočet:

$$\chi^2 = \frac{n(ad - bc)^2}{R_1 R_2 C_1 C_2}$$

### B. Kategorie 2 („*srovnávací pokus*“)

- před analýzou je určen jeden nebo druhý okraj (absolutní četnosti)
- buď užíj Yatesovu korekci, lépe Haberovu korekci\* nebo nejlépe Fisherův exaktní test\*

### C. Kategorie 3 („*independence trial*“)

- oba okraje jsou fixovány
- použij Yatesovu korekci testu dobré shody\* nebo Haberovu korekci\*





# Čtyřpolní tabulky – příklady 1

**Příklady (kategorie 1):** sloupce a řádky jsou zaměnitelné

- (1) Zkoumáme vzájemný výskyt dvou druhů na skalní stepi. Celkem jsme na plochu rozmístili náhodně 100 plošek o rozměru 1x1 m. Na každé ploše jsme zaznamenali přítomnost/nepřítomnost druhu A a druhu B. Oba druhy se vyskytovaly v 36 čtvercích, ani jeden ve 20 čtvercích, pouze druh A se vyskytoval ve 30 čtvercích. *Vyskytují se druhy vzájemně nezávisle?*
- (2) Při studiu vztahu mezi barvou vlasů a očí v populaci Němců antropolog pozoroval náhodný výběr 6800 lidí s těmito výsledky:

		Barva vlasů (V)		
		Tmavá (T)	Světlá (S)	Celkem
Barva očí (O)	Tmavá (T)	726	131	857
	Světlá (S)	3129	2814	5943
Celkem		3855	2945	6800

Jak chápat statistickou hypotézu? – jako tzv. **statistickou nezávislost**

$H_0$ : Barva očí je nezávislá na barvě vlasů

=  $H_0$ : Barva vlasů je nezávislá na barvě očí.

=  $H_0$ : Barva očí a barva vlasů jsou vzájemně nezávislé.

**obecně  $H_0: P(C|G1)=P(C|G2)$ ,  
kde G1 a G2 jsou dvě skupiny  
srovnávané ve vztahu ke znaku C**

(př.:  $H_0: TO|TV=TO|SV \rightarrow 726/3855=131/2945$ )  
nebo

$H_0: TV|TO=TV|SO \rightarrow 726/857=3129/5943$ )

# Čtyřpolní tabulky – příklady 2

**Příklad č. 3** (kategorie 2): Studuji promořenost populací 2 druhů myšic vnitřními parazity. Prohlédnu 100 zvířat 1. druhu na přítomnost / nepřítomnost parazitů (parazité přítomni u 50 jedinců) a totéž provedu pro 2. druh (prohlédnu 50 jedinců; paraziti u 20 jedinců).

Ptám se: *liší se promořenost parazity u těchto druhů? (tj. liší se proporce populace infikované parazity u druhů nebo ne?,  $H_0: p_1 = p_2$ )*

**Příklad č. 4** (kategorie 3): Testuji schopnost 2 druhů vodních plžů odolávat vodnímu proudu. Náhodně vyberu 20 jedinců prvního druhu a 10 jedinců druhého druhu, nechám je přitisknout k podložce a pustím vodu. Experiment ukončím ve chvíli, kdy polovina všech jedinců (bez ohledu na druh) je odnesena proudem.

*Jsou oba druhy schopné odolávat vodnímu proudu stejně za daných podmínek nebo ne?*

# Čtyřpolní tabulky II

## Yatesova korekce na kontinuitu

$$\chi_c^2 = \frac{n \left( |ad - bc| - \frac{n}{2} \right)^2}{R_1 R_2 C_1 C_2}$$

## Haberova korekce

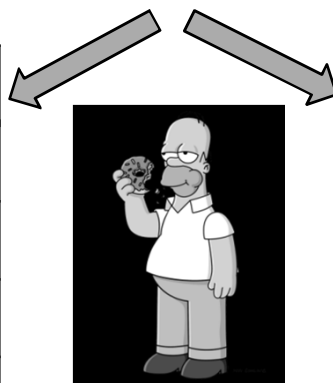
- složitější, počítá jinak číselník vzorce pro čtyřpolní tabulky, výpočet viz Zar (1999), str. 491-492.

# Vliv skryté proměnné na analýzu frekvencí – Simpsonův paradox

Jsou záchranné helikoptéry úspěšnější v záchraně životů než sanitky?

	helikoptéra	sanitka
Postižený zemře	64 (32% z Celkem)	260 (24% z Celkem)
Postižený přežije	136	840
Celkem	200	1100

Vážná nehoda		
	helikoptéra	sanitka
Postižený zemře	48 (48%)	60 (60%)
Postižený přežije	52	40
Celkem	100	100



Lehká nehoda		
	helikoptéra	sanitka
Postižený zemře	16 (16%)	200 (20%)
Postižený přežije	84	800
Celkem	100	1000

Jev, kdy vliv další = skrytá(é) (= rušivá(é)) proměnná(é) změni směr působení nebo obrátí závislost v kontingenční tabulce (při slučování dílčích skupin do jedné).

# Fisherův exaktní test

- vhodný pro analýzu čtyřpolních tabulek kategorie 2 - srovnávací pokus
- založen na hypergeometrické pravděpodobnosti
- počítá přímo pravděpodobnost výskytu dané čtyřpolní tabulky

$$P = \frac{R_1! R_2! C_1! C_2!}{f_{11}! f_{21}! f_{12}! f_{22}! n!}$$

- možnost jednostranného a oboustranného\* testu
- zvláště vhodný, když  $n < 30$

## *Jednostranný test:*

- pro testování  $H_0$  je třeba, aby získaná data směřovala k  $H_A$  a pak sečíst pravděpodobnosti výskytu všech čtyřpolních tabulek s extrémnějším počtem  $f_{11}$  než byl v získané tabulce (tj. ve směru alternativní hypotézy) při zachování konstantních marginálních součtů
- proto „exaktní“
- nejdou-li data ve směru  $H_A$ , netestujeme a říkáme, že nelze zamítnout  $H_0$

\*složitější výpočet

# Fisherův exaktní test - příklad

V populaci lidí zkoumám, zda-li je levo/pravorukost nezávislá/závislá na pohlaví. Vybral jsem náhodně 34 mužů a 36 žen a prozkoumal jejich „rukost“ - viz tabulka:

	Muži	Ženy	Celkem
Levoručí	6	12	18
Pravoručí	28	24	52
Celkem	34	36	70

$H_0$ : Levorukost je stejně či více rozšířená mezi chlapci než mezi dívkami.

$H_A$ : Levorukost je méně běžnější u chlapců než u dívek.

Řešení? FET

(a) For the data of Example 22.2, if

\*

$H_0$ : In the sampled population, left-handedness is not less common among boys than among girls.

$H_A$ : In the sampled population, left-handedness is less common among boys than among girls.

Expressing sample left-handedness as a proportion for each sex:

	Boys	Girls
Left-handed	0.18	0.33
Right-handed	0.82	0.67
Total	1.00	1.00

The sample data are in the direction of  $H_A$ , but are they significantly so?

Jak vypadají  
tyto tabulky?  
takto...

$f_{11}$	$P$
6	0.07257
5	0.02772
4	0.00759
3	0.00144
2	0.00018
1	0.00001
0	0.00000
one-tailed $P$	0.10951

For  $\alpha = 0.05$ , do not reject  $H_0$ . **Nezamítám  $H_0$**

If the hypotheses had been  $H_0$ : In the sampled population, left-handedness is not less common among girls than among boys and  $H_A$ : In the sampled population, left-handedness is less common among girls than among boys, we would have observed that the sample data are *not* in the direction of  $H_A$  and would not reject  $H_0$ , without even computing probabilities.

# Tabulky ve směru $H_a$ pro Fisherův exaktní test - příklad

$f_{11}=6$	Muži	Ženy	Celkem
Levoručí	6	12	18
Pravoručí	28	24	52
Celkem	34	36	70

$f_{11}=5$	Muži	Ženy	Celkem
Levoručí	5	13	18
Pravoručí	29	23	52
Celkem	34	36	70

$f_{11}=4$	Muži	Ženy	Celkem
Levoručí	4	14	18
Pravoručí	30	22	52
Celkem	34	36	70

Pozn.: Data jsou ve směru  
alternativní hypotézy až po  
vyčerpání možností buňky  $f_{11}$ .  
Marginální součty se nemění!

$f_{11}=3$	Muži	Ženy	Celkem
Levoručí	3	15	18
Pravoručí	31	21	52
Celkem	34	36	70

$f_{11}=2$	Muži	Ženy	Celkem
Levoručí	2	16	18
Pravoručí	32	20	52
Celkem	34	36	70

$f_{11}=1$	Muži	Ženy	Celkem
Levoručí	1	17	18
Pravoručí	33	19	52
Celkem	34	36	70

$f_{11}=0$	Muži	Ženy	Celkem
Levoručí	0	18	18
Pravoručí	34	18	52
Celkem	34	36	70

# Párové testování dat na nominální škále

## McNamarův test = test symetrie\*

- jedná o typ designu, kdy na stejném objektu provádíme (postupně) buď dva zásahy (*treatmenty*) či na něm zjišťujeme dvě nominální veličiny

$H_0: (b-c)=0 \Rightarrow b:c=1:1$  = „úspěšnost (či lépe neúspěšnost) je stejná“

		Pokus (Proměnná) 1	
		+	-
Pokus (Proměnná) 2	+	a	b
	-	c	d

**Konkordantní** ( $a+d$ ; souhlasná reakce) a **diskordantní** ( $b+c$ ; rozdílná reakce) **páry**.

$$\chi^2 = \frac{\left(b - \frac{(b+c)}{2}\right)^2}{\frac{(b+c)}{2}} + \frac{\left(c - \frac{(b+c)}{2}\right)^2}{\frac{(b+c)}{2}} = \frac{(b-c)^2}{b+c}$$

**DF=1**

Pokud je platná  $H_0$ , pak očekávaný počet  $b$  a  $c$  je  $(b+c)/2$

**POZOR: pozorování nejsou vzájemně nezávislá, nelze užít klasický test dobré shody !!!**

\* Alternativa: **binomický** test s  $H_0: p(b)=p(c)=0,5$ , kde  $n=b+c$

Př.

- obdobně lze provádět test pro více jak 2x2 kategorie = Bowkerův test



**Příklad:** Sledujeme skupinu 18 pacientů, kteří byli léčeni dvěma různými hypertenzivy *A* a *B*. Každý pacient dostával po dobu 1 měsíce lék *A* a po odeznění případných účinků po dobu 1 měsíce lék *B*. Výsledek byl klasifikován jako úspěch (tlak snížen o více než 15 mm Hg) či neúspěch. Liší se léky v účinku?

**Tabulka** - nesprávné uspořádání

	Lék A	Lék B	Celkem
Úspěch	10	4	14
Neúspěch	8	14	22
Celkem	18	18	36

**Tabulka** - správné uspořádání

	Lék B		
Lék A	Úspěch	Neúspěch	Celkem
Úspěch	1	3	4
Neúspěch	9	5	14
Celkem	10	8	18

# McNamarův test - příklad

\*

- $H_0$ : Úspěšnost obou léků je shodná  
 $H_A$ : Úspěšnost léků není shodná

Zaměříme se na pacienty vykazující u každého léku jiné výsledky (= diskordantní páry; *b* a *c*).

$$\chi^2 = (b-c)^2 / (b+c) = (3-9)^2 / (3+9) = 3,00$$

$$\chi^2_{KRIT} = 3,84 \rightarrow \chi^2 < \chi^2_{KRIT}$$

Závěr: nezamítáme  $H_0$

# Poměr proporcí (relativní risk) a poměr šancí

- alternativní míry závislosti v kontingenční tabulce

**Poměr pravděpodobností (populačních proporcí):**  $p_1/p_2$ , je-li výsledek škodlivý (diagnostika), pak se poměr nazývá **relativní risk** (*relative risk*)

**Poměr šancí** (*odds ratio*;  $\theta$ ): poměr šancí za dvou rozdílných podmínek

**šance** (*odds*):  $p_1/(1-p_1)$ , tj. poměr pravděpodobnosti, že jev nastane, k pravděpodobnosti, že nenastane

$$\theta = \frac{\frac{p_1}{1-p_1}}{\frac{p_2}{1-p_2}}$$

Příklad (2x2 tabulky, typ 1)		Status kouření matky	
		kuřačka	nekuřačka
Hmotnost novorozeněte	nízká	237 $(p_1)=237/3726=0,064$	197 $(p_2)=197/6067=0,032$
	normální	3489	5870
	celkem	3726	6067

- Relativní risk mít novorozeně s menší hmotností je u matek-kuřaček 2x větší než u nekuřaček:  $p_1/p_2 = 0,064/0,032 = 2$
- Poměr šancí mít novorozeně s menší hmotností u matek-kuřaček je cca 2x větší než u nekuřaček:  $\theta = [p_1/(1-p_1)]/[p_2/(1-p_2)]=0,0679/0,03356=2,023 \approx 2$

# Testování shody empir. a teoret. rozdělení ordinální a kvantit. prom.

Jak zjistit, zda-li existuje shoda mezi očekávanou a skutečně zjištěnou  
distribucí v pokusu, kde hodnotíme  
ordinální či kvantitativní proměnnou (tj. kategorie lze seřadit !!!)



## Kolmogorov-Smirnovův test (*Kolmogorov-Smirnov test; K-S test*)

- test testuje shodu pozorované a očekávané kumulativní frekvenční distribuce
- pro každou kategorii  $i$  určíme absolutní rozdíl mezi oběma distribucemi:

$$|d_i| = |F_i - \hat{F}_i|$$

- největší rozdíl  $d_{\text{MAX}}$  je považován za testovou statistiku, kterou porovnáme s tabulkovou hodnotou s parametry  $n$  (velikost souboru) a  $k$  (počet kategorií)

**Pokud  $|d_{\text{max}}| \geq d_{\text{KRIT}}$ , zamítáme  $H_0$ .**

Poznámka: postup pro kvantitativní proměnnou je složitější, než je uvedeno.

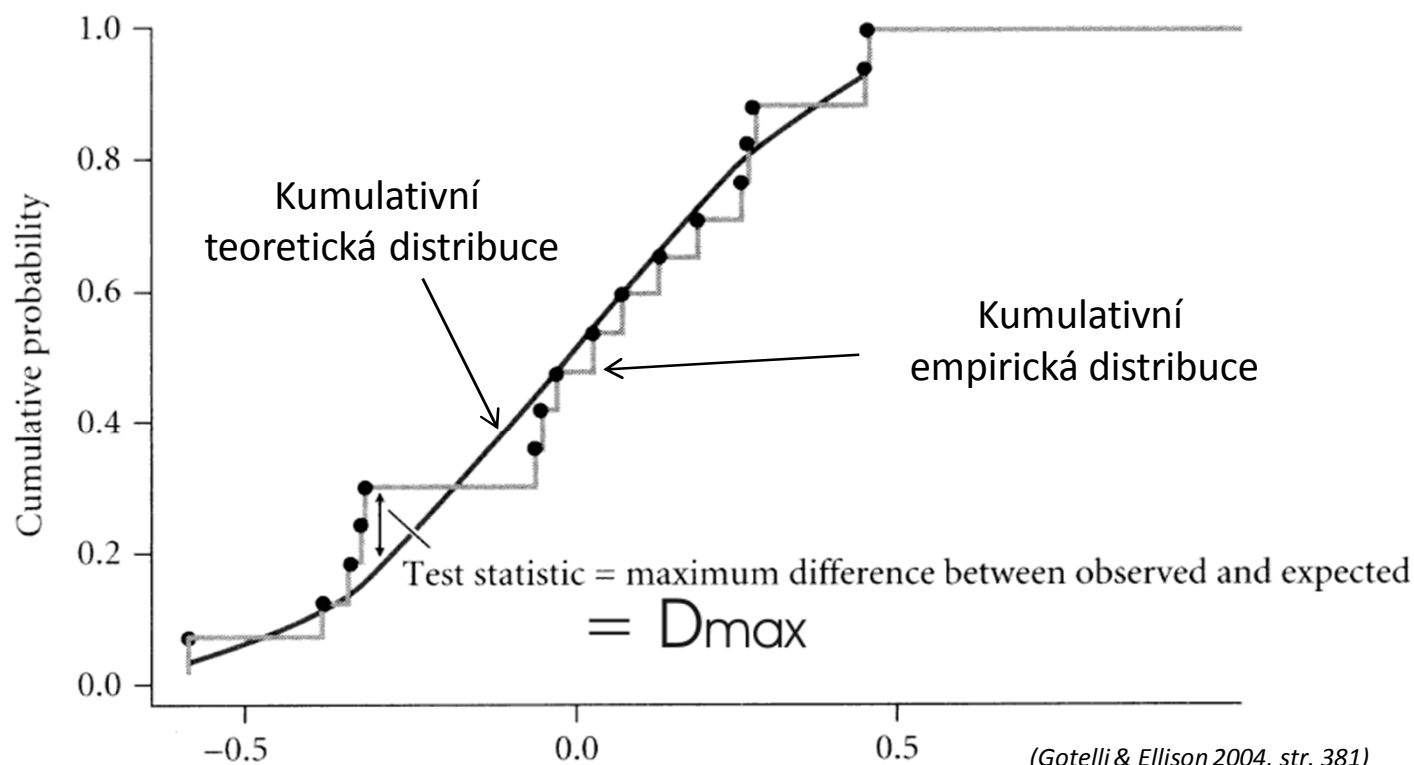
Lze testovat i  $H_0$ : dvě proměnné mají shodnou distribuci !!! A modifikací využít test pro zjištění normality rozdělení dat (tzv. Lillieforsův test).

# Kolmogorov-Smirnovův test

## - tabulky kritických hodnot (část)

**TABLE B.8 (cont.)** Critical Values of  $d_{max}$  for the Kolmogorov-Smirnov Goodness of Fit Test for Discrete or Grouped Data

$k$	$n$	$\alpha < 0.50$	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
4	72	5	7	9	10	11	12	13	14	15
4	76	5	8	9	10	11	12	13	14	15
4	80	5	8	9	10	11	12	13	15	15
4	84	5	7	9	10	12	13	14	15	16
4	88	5	7	9	10	12	13	14	15	16
4	92	5	7	9	10	12	13	14	16	16
4	96	5	7	9	10	12	13	14	16	17
4	100	5	8	9	11	12	13	14	16	17
5	5	2	3	3	3	4	4	4	4	4
5	10	3	3	4	4	5	5	5	6	6
5	15	3	4	5	5	6	6	7	7	7
5	20	4	5	5	6	7	7	7	8	8
5	25	4	5	6	6	7	8	8	9	9
5	30	4	5	6	7	8	8	9	10	10
5	35	4	6	7	7	8	9	10	10	11
5	40	5	6	7	8	9	10	10	11	12
5	45	5	6	7	8	9	10	11	12	12
5	50	5	7	8	9	10	11	11	12	13



12	12	13	14
12	12	13	14
12	13	14	15
13	14	15	15
13	14	15	16
13	14	15	16
13	14	15	16
13	14	15	16
13	14	15	16
13	14	16	17
14	15	16	17
4	4	5	5
6	6	6	7
7	7	8	8
8	8	9	9
9	9	10	10
9	10	11	11
10	11	11	12
11	11	12	13
11	12	13	14
12	13	13	14
12	13	14	15
13	13	14	15
13	14	15	16
13	14	15	16
14	15	16	16
14	15	16	17
5	5	5	5

# Kolmogorov-Smirnovův test - příklad \*

**Příklad:** Studuji chutnost 5 typů stravy pro kočky lišící se pouze obsahem vody. Obsah je kvantifikován ordinální stupnicí o 5 stupních: 1 (mokrý) až 5 (suchý). Celkem jsem testoval těchto 5 typů stravy na 35 kočkách tak, že jsem každé kočce dal na výběr z těchto 5 druhů stravy a pozoroval její reakci (výběr).

**Ptám se:** *preferují kočky nějaký typ nebo ne?*

$H_0$ : Cats show no preference among the five food moisture contents (i.e., cats prefer all five equally).

$H_A$ : Cats do show preference among the five food moisture contents (i.e., cats do not prefer all five equally).

	Moisture class ( $X_i$ )					$n$
	(Moist) 1	2	3	4	(Dry) 5	
$f_i$	8	13	6	6	2	35
$\hat{f}_i$	7	7	7	7	7	35
$F_i$	8	21	27	33	35	
$\hat{F}_i$	7	14	21	28	35	
$ d_i $	1	7	6	5	0	

$$d_{\max} = \text{maximum } |d_i| = 7$$

$$(d_{\max})_{0.05, 5, 35} = 7$$

Therefore, reject  $H_0$ .

$$0.02 < P < 0.05$$

(Zar 1999)

# Statistické testy hypotéz týkající se kvantitativní a ordinální proměnné

# Hypotézy o jednom výběru

*(One-sample hypotheses)*

# Neparametrický znaménkový test *(the Sign test)*

= Binomický test na  $p=0,5$  (*Binomial test*)

- uvažujeme náhodný výběr o rozsahu  $n$  ze spojitého rozdělení s mediánem  $M$ ; *test je vhodný pokud je proměnná silně asymetrická*
- nejobecněji testujeme, zda-li medián určený  $H_0$  (tedy nabývající hodnoty  $a$ ) leží ve středu výběru nebo ne (tj.  $H_0: M=a$ )
- test je slabý, zvl. pro malé  $n$
- *testové kritérium:  $S^+ =$  počet rozdílů  $x_i - a$  s kladným znaménkem*  
→ platí-li  $H_0$ , má binomické rozdělení  $Bi(n; 0,5)$  (tj.  $H_0$  je ekvivalentní:  $p=0,5$ ; = binomický test [Binomial test]).
- POZOR !!!:  $n$  = původní  $n$  minus počet vyřazených hodnot  
(= hodnoty rovnající se hypotetizovanému mediánu)
- *oboustranný test:  $H_0$  zamítáme pokud  $S^+ \leq S_{\alpha(2), n}$  nebo  $S^+ \geq n - S_{\alpha(2), n}$*   
(viz tabulky, nad  $n > 25$  lze aproximovat normálním r.)
- *jednostranný test:  $H_0$  zamítáme pro (a)  $H_A: M < a$ , když  $S^+ < S_{\alpha(1), n}$  nebo pro (b)  $H_A: M > a$ , když  $S^+ > n - S_{\alpha(1), n}$*



# Znaménkový test - příklad

***Byla měřena teplota těla (ve stupních Celsia) 25 náhodně vybraných krabů osídlujících příbojovou zónu vystavených prostředí se stálou experimentální teplotou vzduchu 24,3 °C:***

25,8; 24,6; 22,9; 25,1; 27,3; 24,0; 24,5; 23,9; 26,2; 24,3; 24,6; 23,3; 25,5; 28,1; 24,8; 23,5; 26,3; 25,4; 25,5; 23,9; 27,0; 24,8; 22,9; 25,4

***Je teplota krabů shodná s teplotou vnějšího prostředí?***

$H_0: M = 24,3$

$H_A: M \neq 24,3$

Počet kladných rozdílů  $x_i - 24,3$ : 17

Počet záporných rozdílů  $x_i - 24,3$ : 7

Počet vyřazených hodnot (nulových rozdílů): 1

pak  $n = 24$

kritická hodnota  $S$  pro  $n = 24$  a oboustranný test na  $\alpha = 0,05$  je: 6 a 18 (pro jednostranný test na  $\alpha = 0,05$  pak 7 a 17)

Závěr: 17 je méně extrémní než 18 - nezamítáme  $H_0$  (oboustranný test)

# Tabulky kritických hodnot pro znaménkový test - část

Critical Values of  $C$  for the Sign Test or the Binomial Test with  $p = 0.5$

$\alpha(2):$	.50	.20	.10	.05	.02	.01	.005	.002	.001
$\alpha(1):$	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
$n$									
2	---	---	---	---	---	---	---	---	---
3	0	---	---	---	---	---	---	---	---
4	0	0	---	---	---	---	---	---	---
5	1	0	0	---	---	---	---	---	---
6	1	0	0	0	---	---	---	---	---
7	2	1	0	0	0	---	---	---	---
8	2	1	1	0	0	0	---	---	---
9	2	2	1	1	0	0	0	---	---
10	3	2	1	1	0	0	0	0	---
11	3	2	2	1	1	0	0	0	0
12	4	3	2	2	1	1	0	0	0
13	4	3	3	2	1	1	1	0	0
14	5	4	3	2	2	1	1	1	0
15	5	4	3	3	2	2	1	1	1
16	6	4	4	3	2	2	2	1	1
17	6	5	4	4	3	2	2	1	1
18	7	5	5	4	3	3	2	2	1
19	7	6	5	4	4	3	3	2	2
20	7	6	5	5	4	3	3	2	2
21	8	7	6	5	4	4	3	3	2
22	8	7	6	5	5	4	4	3	3
23	9	7	7	6	5	4	4	3	3
24	9	8	7	6	5	5	4	4	3
25	10	8	7	7	6	5	5	4	4
26	10	9	8	7	6	6	5	4	4
27	11	9	8	7	7	6	5	5	4
28	11	10	9	8	7	6	6	5	5
29	12	10	9	8	7	7	6	5	5
30	12	10	10	9	8	7	6	6	5
31	13	11	10	9	8	7	7	6	6
32	13	11	10	9	8	8	7	6	6
33	14	12	11	10	9	8	8	7	6
34	14	12	11	10	9	9	8	7	7
35	15	13	12	11	10	9	8	8	7
36	15	13	12	11	10	9	9	8	7
37	15	14	13	12	10	10	9	8	8
38	16	14	13	12	11	10	9	9	8
39	16	15	13	12	11	11	10	9	8
40	17	15	14	13	12	11	10	9	9
41	17	15	14	13	12	11	11	10	9
42	18	16	15	14	13	12	11	10	10
43	18	16	15	14	13	12	11	11	10
44	19	17	16	15	13	13	12	11	10
45	19	17	16	15	14	13	12	11	11
46	20	18	16	15	14	13	13	12	11
47	20	18	17	16	15	14	13	12	11
48	21	19	17	16	15	14	13	12	12
49	21	19	18	17	15	15	14	13	12
50	22	19	18	17	16	15	14	13	13

$\alpha(2):$	.50	.20	.10	.05	.02	.01	.005	.002	.001
$\alpha(1):$	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
$n$									
51	22	20	19	18	16	15	15	14	13
52	23	20	19	18	17	16	15	14	13
53	23	21	20	18	17	16	15	14	14
54	24	21	20	19	18	17	16	15	14
55	24	22	20	19	18	17	16	15	14
56	24	22	21	20	18	17	17	16	15
57	25	23	21	20	19	18	17	16	15
58	25	23	22	21	19	18	17	16	16
59	26	24	22	21	20	19	18	17	16
60	26	24	23	21	20	19	18	17	16
61	27	24	23	22	20	20	19	18	17
62	27	25	24	22	21	20	19	18	17
63	28	25	24	23	21	20	19	18	18
64	28	26	24	23	22	21	20	19	18
65	29	26	25	24	22	21	20	19	18
66	29	27	25	24	23	22	21	20	19
67	30	27	26	25	23	22	21	20	19
68	30	28	26	25	23	22	22	20	20
69	31	28	27	25	24	23	22	21	20
70	31	29	27	26	24	23	22	21	20
71	32	29	28	26	25	24	23	22	21
72	32	30	28	27	25	24	23	22	21
73	33	30	28	27	26	25	24	22	22
74	33	30	29	28	26	25	24	23	22
75	34	31	29	28	26	25	24	23	22
76	34	31	30	28	27	26	25	24	23
77	35	32	30	29	27	26	25	24	23
78	35	32	31	29	28	27	26	24	24
79	36	33	31	30	28	27	26	25	24
80	36	33	32	30	28	27	26	25	24
81	36	34	32	31	29	28	27	26	25
82	37	34	33	31	30	28	27	26	25
83	37	35	33	32	30	29	28	27	26
84	38	35	33	32	30	29	28	27	26
85	38	36	34	32	31	30	29	27	26
86	39	36	34	33	31	30	29	28	27
87	39	37	35	33	32	31	29	28	27
88	40	37	35	34	32	31	30	29	28
89	40	37	36	34	33	31	30	29	28
90	41	38	36	35	33	32	31	29	29
91	41	38	37	35	33	32	31	30	29
92	42	39	37	36	34	33	32	30	29
93	42	39	38	36	34	33	32	31	30
94	43	40	38	37	35	34	32	31	30
95	43	40	38	37	35	34	33	32	31
96	44	41	39	37	36	34	33	32	31
97	44	41	39	38	36	35	34	32	31
98	45	42	40	38	37	35	34	33	32
99	45	42	40	39	37	36	35	33	32
100	46	43	41	39	37	36	35	34	33

# Parametrický Studentův $t$ -test (*One-sample $t$ -test*)

**Podmínky použití testu:** 1. rozdělení sledované náhodné veličiny ve výběru by nemělo být příliš odlišné od normálního (test je ale robustní na odchylky→); 2. prostý náhodný výběr

## Testování oboustranné hypotézy o průměru

$$\begin{aligned} H_0: \bar{x} &= \mu_0 \\ H_A: \bar{x} &\neq \mu_0 \end{aligned}$$

$$t = \frac{\bar{x} - \mu_0}{s_{\bar{x}}} \quad \text{s DF} = n-1$$

**Pokud  $|t| \geq t_{\alpha(2),v}$**   
→ zamítáme  $H_0$

## Testování jednostranné hypotézy o průměru

$$\begin{aligned} H_0: \bar{x} &\geq \mu_0 \\ H_A: \bar{x} &< \mu_0 \end{aligned}$$

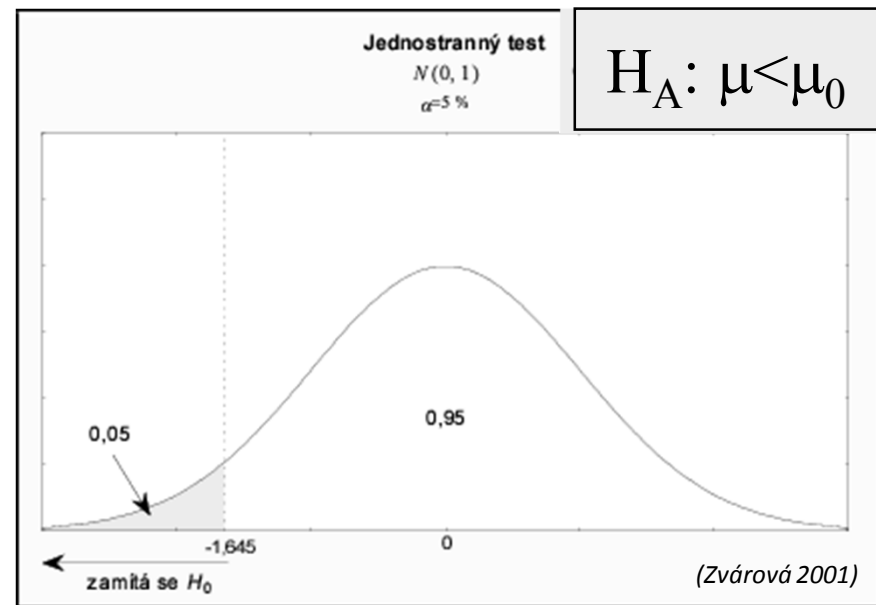
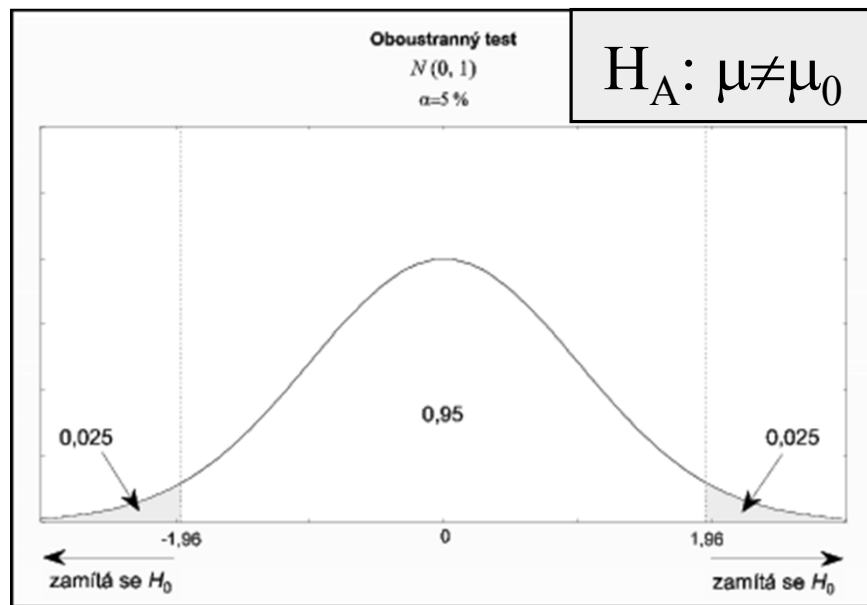
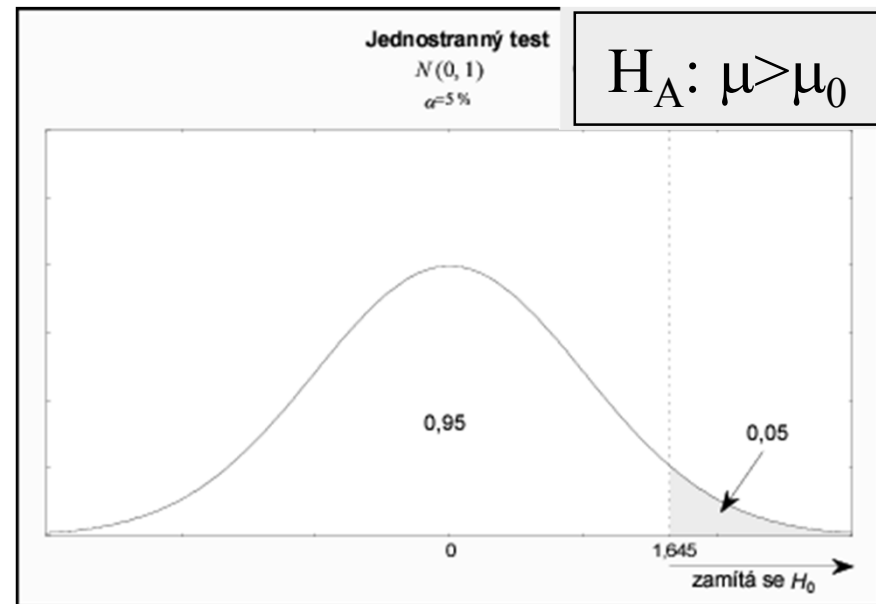
$$t = \frac{\bar{x} - \mu_0}{s_{\bar{x}}} \quad \text{s DF} = n-1$$

**Pokud  $t \leq -t_{\alpha(1),v}$**   
→ zamítáme  $H_0$

$t \geq t_{\alpha(1),v}$  v případě  $H_A: \bar{x} > \mu_0$

# Hypotézy o jednom výběru

Vztah tvaru hypotézy, hladiny významnosti a rozhodování



Příklad:  $\alpha=0,05$ , normální rozdělení  $N(0;1)$

# Doporučení pro užití t-testu

- Mimo případ malých rozsahů souboru je důležitější podmínka prostého náhodného výběru, než že populace má normální rozdělení.
- $n < 15$ : použít t-test pokud data mají  $\pm$  normální rozdělení (zhruba symetrická, jeden vrchol, žádné odlehlé hodnoty), jinak ne
- $n \geq 15$ : užít t-test mimo případy silně šikmého rozdělení či přítomnosti odlehlých hodnot
- velké výběry: t-test lze užít i pro zešikmená data pokud je  $n \geq 40$

# Jednovýběrový *t*-test - příklad

Body temperatures (measured in °C) of twenty-five intertidal crabs placed in air at 24.3°C: 25.8, 24.6, 26.1, 22.9, 25.1, 27.3, 24.0, 24.5, 23.9, 26.2, 24.3, 24.6, 23.3, 25.5, 28.1, 24.8, 23.5, 26.3, 25.4, 25.5, 23.9, 27.0, 24.8, 22.9, 25.4.

$$H_0: \mu = 24.3^\circ\text{C}$$

$$H_A: \mu \neq 24.3^\circ\text{C}$$

$$\alpha = 0.05$$

$$n = 25$$

$$\bar{X} = 25.03^\circ\text{C}$$

$$s^2 = 1.80(^\circ\text{C})^2$$

$$s_{\bar{X}} = \sqrt{\frac{1.80(^\circ\text{C})^2}{25}} = 0.27^\circ\text{C}$$

$$t = \frac{\bar{X} - \mu}{s_{\bar{X}}} = \frac{25.03^\circ\text{C} - 24.3^\circ\text{C}}{0.27^\circ\text{C}} = 2.704$$

$$\nu = 24$$

$$t_{0.05(2), 24} = 2.064$$

As  $|t| > t_{0.05(2), 24}$ , reject  $H_0$  and conclude that the sample of twenty-five body temperatures came from a population whose mean is not 24.3°C.

$$0.01 < P < 0.02 \quad [P = 0.012]$$

# Interval spolehlivosti pro populační průměr I.

Víme, že 95% všech možných průměrů výběrů (o velikosti  $n$ ) z populace s  $\mu$  je přítomno v t-rozdělení v rozmezí  $-t_{0,05(2),\nu}$  a  $t_{0,05(2),\nu}$

$$P\left(-t_{0,05(2),\nu} \leq \frac{\bar{x} - \mu}{s_{\bar{x}}} \leq t_{0,05(2),\nu}\right) = 0,95$$



$$P(\bar{x} - t_{0,05(2),\nu} s_{\bar{x}} \leq \mu \leq \bar{x} + t_{0,05(2),\nu} s_{\bar{x}}) = 0,95$$

= (95%) konfidenční interval pro průměr (*Confidence limits*)

**Na čem závisí šíře intervalu?**

- $s \uparrow$   $n \downarrow$  šíře intervalu
- $s \uparrow$   $s_{\bar{x}} \uparrow$  šíře intervalu
- $s \downarrow$   $\alpha \uparrow$  šíře intervalu

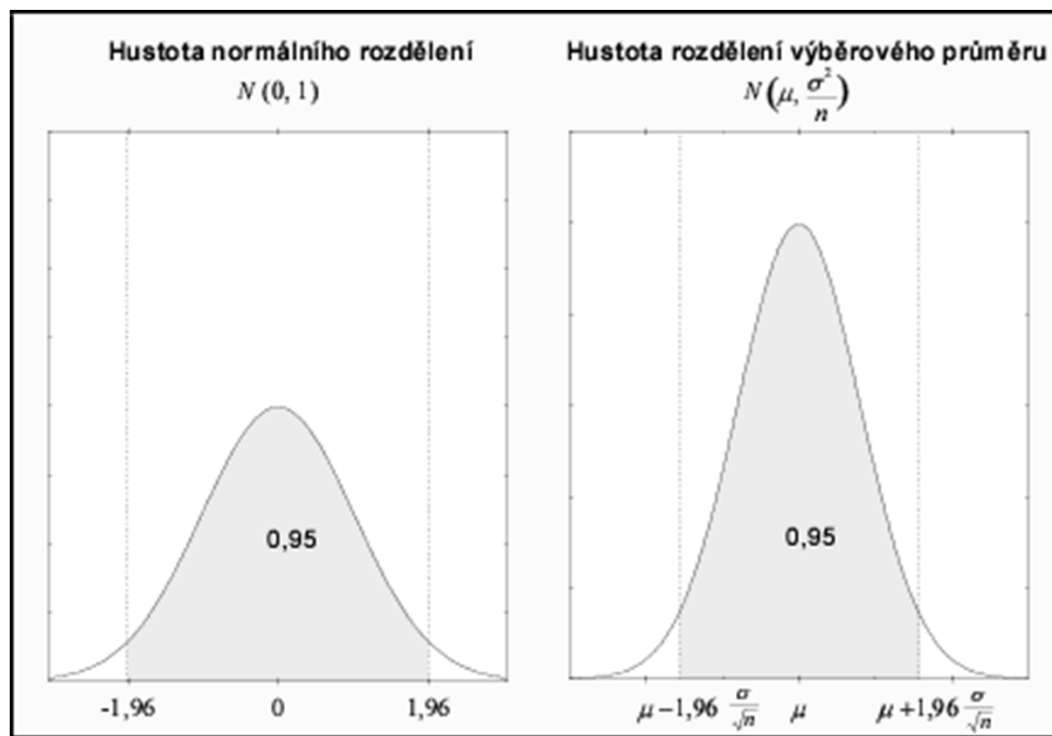
# Interval spolehlivosti II.

Kolik intervalů spolehlivosti nepokrývá populační průměr?

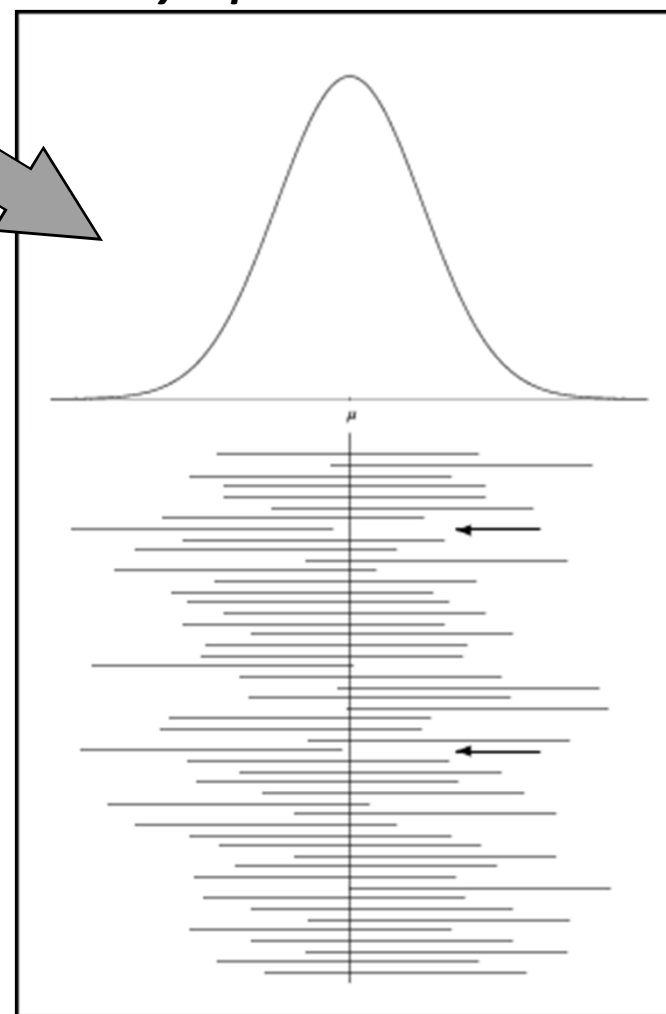
Odpověď: „ $\alpha$ -procent“

*Vygenerováno náhodně 50 výběrů o  $n=10$  z populace se známým průměrem*

Za předpokladu, že známe  $\sigma$ , pak



(Zvárová 2001)





# Interval spolehlivosti III.

## Ecology and Evolution

Open Access

### Vegetation in clear-cuts depends on previous land use: a century-old grassland legacy

Dennis Jonason, Mathias Ibbe, Per Milberg, Albert Tunér, Lars Westerberg & Karl-Olof Bergman

IFM Biology, Conservation Ecology Group, Linköping University, Linköping SE-581 83, Sweden

#### Keywords

Extinction debt, habitat fragmentation, historical maps, plant traits, production forestry, remnant populations, seminatural grasslands.

#### Correspondence

Dennis Jonason, IFM Biology, Conservation Ecology Group, Linköping University, SE-581 83 Linköping, Sweden.  
Tel: 0046 13 282 583;  
Fax: 0046 13 137 568;  
E-mail: denjo@ifm.liu.se

#### Funding Information

The project was funded by a grant from Skogsällskapet.

Received: 24 June 2014; Revised: 16 September 2014; Accepted: 23 September 2014

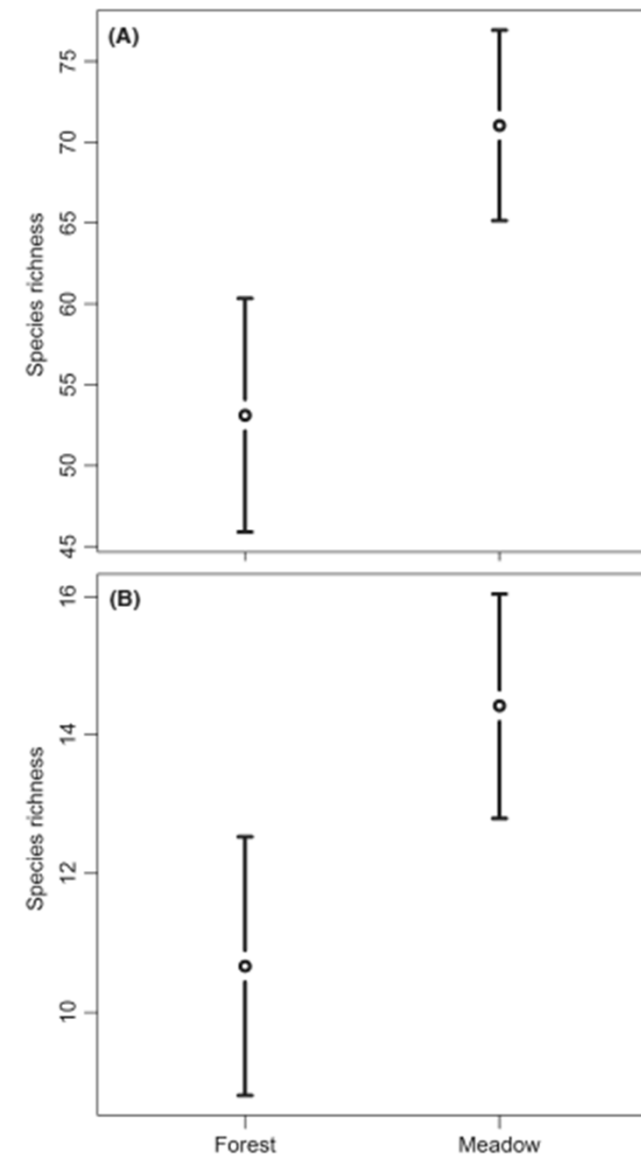
*Ecology and Evolution* 2014; 4(22): 4287–4295

doi: 10.1002/ece3.1288

#### Abstract

Plant species richness in central and northern European seminatural grasslands is often more closely linked to past than present habitat configuration, which is indicative of an extinction debt. In this study, we investigate whether signs of historical grassland management can be found in clear-cuts after at least 80 years as coniferous production forest by comparing floras between clear-cuts with a history as meadow and as forest in the 1870s in Sweden. Study sites were selected using old land-use maps and data on present-day clear-cuts. Species traits reflecting high capacities for dispersal and persistence were used to explain any possible links between the plants and the historical land use. Clear-cuts that were formerly meadow had, on average, 36% higher species richness and 35% higher richness of grassland indicator species, as well as a larger overall seed mass and lower anemochory, compared to clear-cuts with history as forest. We suggest that the plants in former meadows never disappeared after afforestation but survived as remnant populations. Many contemporary forests in Sweden were managed as grasslands in the 1800s. As conservation of remaining grassland fragments will not be enough to reduce the existing extinction debts of the flora, these young forests offer opportunities for grassland restoration at large scales. Our study supports the concept of remnant populations and highlights the importance of considering historical land use for understanding the distribution of grassland plant species in fragmented landscapes, as well as for policy-making and conservation.

Clear-cuts with a history as meadow had, on average, 36% higher species richness compared to clearcuts with a long history as forest (Fig. 1A) and a 35% higher richness of grassland indicator species (Fig. 1B).



**Figure 1.** Average species richness with CI 95% of (A) all plants and of (B) plants classified as grassland indicator species in clear-cuts historically managed as forest and as meadow.

# Tabulky Studentova rozdělení

## Tabulka Studentova t-rozdělení

Pravděpodobnosti hodnot  $\geq$   
než hodnota  $t$  (vypočtená z  $t$ -transformace)  
v tabulce

v	Critical Values of the $t$ Distribution									
	$\alpha(2)$ : 0.50	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
1	1.000	3.078	6.314	12.706	31.821	63.657	127.321	318.309	636.619	
2	0.816	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.599	
3	0.765	1.638	2.353	3.182	4.541	5.841	7.453	10.215	12.924	
4	0.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610	
5	0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869	
6	0.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959	
7	0.711	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408	
8	0.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041	
9	0.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781	
10	0.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587	
11	0.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437	
12	0.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318	
13	0.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221	
14	0.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140	
15	0.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073	
16	0.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015	
17	0.689	1.333	1.740	2.110	2.567	2.898	3.222	3.648	3.965	
18	0.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922	
19	0.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883	
20	0.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850	
21	0.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819	
22	0.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792	
23	0.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.768	
24	0.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745	
25	0.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725	
26	0.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707	
27	0.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690	
28	0.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674	
29	0.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659	
30	0.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646	
31	0.682	1.309	1.696	2.040	2.453	2.744	3.022	3.375	3.633	
32	0.682	1.309	1.694	2.037	2.449	2.738	3.015	3.365	3.622	
33	0.682	1.308	1.692	2.035	2.445	2.733	3.008	3.356	3.611	
34	0.682	1.307	1.691	2.032	2.441	2.728	3.002	3.348	3.601	
35	0.682	1.306	1.690	2.030	2.438	2.724	2.996	3.340	3.591	
36	0.681	1.306	1.688	2.028	2.434	2.719	2.990	3.333	3.582	
37	0.681	1.305	1.687	2.026	2.431	2.715	2.985	3.326	3.574	
38	0.681	1.304	1.686	2.024	2.429	2.712	2.980	3.319	3.566	
39	0.681	1.304	1.685	2.023	2.426	2.708	2.976	3.313	3.558	
40	0.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551	
41	0.681	1.303	1.683	2.020	2.421	2.701	2.967	3.301	3.544	
42	0.680	1.302	1.682	2.018	2.418	2.698	2.963	3.296	3.538	
43	0.680	1.302	1.681	2.017	2.416	2.695	2.959	3.291	3.532	
44	0.680	1.301	1.680	2.015	2.414	2.692	2.956	3.286	3.526	
45	0.680	1.301	1.679	2.014	2.412	2.690	2.952	3.281	3.520	
46	0.680	1.300	1.679	2.013	2.410	2.687	2.949	3.277	3.515	
47	0.680	1.300	1.678	2.012	2.408	2.685	2.946	3.273	3.510	
48	0.680	1.299	1.677	2.011	2.407	2.682	2.943	3.269	3.505	
49	0.680	1.299	1.677	2.010	2.405	2.680	2.940	3.265	3.500	
50	0.679	1.299	1.676	2.009	2.403	2.678	2.937	3.261	3.496	

1. ř. - pro  
oboustranný test, 2.  
pro jednostranný test

Hodnoty  $t$

Stupně volnosti (n-1)

This table was prepared using Equations 26.7.3 and 26.7.4 of Zelen and Severo (1964), except for the values at infinity degrees of freedom, which are adapted from White (1970). Except for the values at infinity degrees of freedom,  $t$  was calculated to eight decimal places and then rounded to three decimal places.

Examples:

(Zar 1996, upraveno)

$t_{0.05(2), 13} = 2.160$  and  $t_{0.01(1), 19} = 2.539$

# Testování proporce výsledku („úspěchů“) v populaci

- zabýváme se proporcí  $p$  nějakého výsledku v populaci („úspěch“)
- protože pracujeme s výběrem,  
pak odhadem  $p$  je výběrová proporce

$$\hat{p} = \frac{n_{\text{úspěch}}}{n}$$

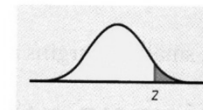
- lze ukázat, že s rostoucím  $n$  se rozdělení výběrové proporce blíží rozdělení normálnímu s parametry  $\mu = p$  a  $s = \sqrt{p(1-p)/n}$
- podmínky užití: prostý náhodný výběr, dostatečně velký výběr

**Statistický test** pro proporcí  $H_0: p = p_0$  využívá z-statistiku mající standardizované normální rozdělení:

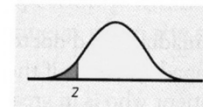
$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

- použít jedině když  $np_0$  a  $n(1-p_0) \geq 10$
- P-hodnoty pro test  $H_0$  jsou:

$$H_a: p > p_0 \quad \text{is} \quad P(Z \geq z)$$



$$H_a: p < p_0 \quad \text{is} \quad P(Z \leq z)$$



$$H_a: p \neq p_0 \quad \text{is} \quad 2P(Z \geq |z|)$$



# Proporce populace - příklad

Náhodný výběr mezi novorozenci zjistil mezi 25468 prověřenými 13173 chlapců. Publikované poznatky ukazují, že zastoupení chlapců a dívek není 1:1, ale že chlapci jsou v populaci četnější než dívky? Jak se shoduje pozorování s předpokladem?

$$H_0: p = 0.5$$

$$H_A: p > 0.5$$

$$\hat{p} = \frac{13172}{25468} = 0.5172$$

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.5172 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{25468}}} = 5.49$$

Z tabulek z-rozdělení vyčteme, že  $P(Z \geq 5.49)$  je hodnota menší než 0.002.

Zamítáme  $H_0$  a říkáme, že chlapců je v populaci novorozenců více než 50%.

# Statistická tabulka normovaného normálního rozdělení (z-rozdělení)

*(tabulka uvádí  
proporci normály,  
která leží za  
hodnotou z;  
(tj. je více  
extrémní)*

**TABLE B.2** Proportions of the Normal Curve (One-Tailed)

This table gives the proportion of the normal curve that lies beyond (i.e., is more extreme than) a given normal deviate; e.g.,  $Z = (X_i - \mu)/\sigma$  or  $Z = (X - \mu)/\sigma_{\bar{X}}$ . For example, the proportion of a normal distribution for which  $Z \geq 1.51$  is 0.0655.

z	0	1	2	3	4	5	6	7	8	9	z
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641	0.0
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247	0.1
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859	0.2
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483	0.3
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121	0.4
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776	0.5
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451	0.6
0.7	0.2420	0.2389	0.2358	0.2327	0.2297	0.2266	0.2236	0.2207	0.2177	0.2148	0.7
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867	0.8
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611	0.9
1.0	0.1587	0.1562	0.1538	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379	1.0
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170	1.1
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985	1.2
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823	1.3
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681	1.4
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559	1.5
1.6	0.0548	0.0537	0.0526	0.0515	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455	1.6
1.7	0.0446	0.0436	0.0427	0.0419	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367	1.7
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294	1.8
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0238	0.0233	1.9
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183	2.0
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143	2.1
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110	2.2
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084	2.3
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064	2.4
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048	2.5
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036	2.6
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026	2.7
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019	2.8
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014	2.9
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010	3.0
3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007	3.1
3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	3.2
3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003	3.3
3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002	3.4
3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	3.5
3.6	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	3.6
3.7	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	3.7
3.8	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	3.8

# Kvantitativní a ordinální proměnné

Hypotézy o dvou výběrech –  
znáhodněné uspořádání  
*(Two sample hypotheses)*

# Parametrický dvouvýběrový t-test (*two-sample t-test*)

Testování rozdílů mezi dvěma průměry: *oboustranná hypotéza*

$$\begin{array}{l} H_0: \mu_1 - \mu_2 = 0 \\ H_A: \mu_1 - \mu_2 \neq 0 \end{array}$$



$$\begin{array}{l} H_0: \mu_1 = \mu_2 \\ H_A: \mu_1 \neq \mu_2 \end{array}$$



$$t = \frac{\bar{x}_1 - \bar{x}_2}{S_{\bar{x}_1 - \bar{x}_2}}$$

variance rozdílu mezi 2 nezávislými proměnnými je rovna součtu variancí těchto proměnných: t-test vyžaduje, aby  $\sigma_1^2 = \sigma_2^{2*}$ , pak střední chyba rozdílu průměrů je:

$$S_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$$

Jak spočítat společnou varianci?  
(=sdružený odhad rozptylu):

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

pokud  $|t| \geq t_{\alpha(2),v}$  .... zamítáme  $H_0$

\* otestovat F-testem

(pro jednostranný test: je důležité sledovat orientaci  $H_0$  a tedy i znaménko  $t$ -hodnoty, která není v absolutní hodnotě, viz jednostranný t-test)



# Dvouvýběrový t-test

## Podmínky užití t-testu:

### Nezávislost výběrů

- nejsou-li výběry na sobě nezávislé, tj. oba např. obsahují měření stejného jedince, pak uvedený postup nemůžeme použít!

### Normální rozdělení

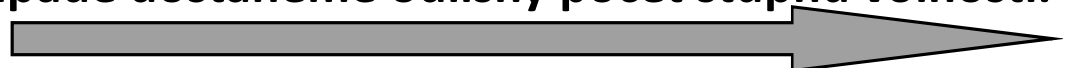
- rozdělení sledované náhodné veličiny v obou skupinách by nemělo být příliš odlišné od normálního

### Prosté náhodné výběry

- pokud by pozorování v některém výběru nebyla nezávislá, pak by mohlo dojít k neoprávněnému zmenšení odhadu rozptylu  $s^2$ . Mohl by tedy vyjít významný rozdíl, přestože by ve skutečnosti rozdíl nebyl.

### Shodné rozptyly ve skupinách

- pokud se rozptyly v obou skupinách výrazně liší, můžeme použít modifikaci dvouvýběrového t testu. V tomto případě dostaneme odlišný počet stupňů volnosti.

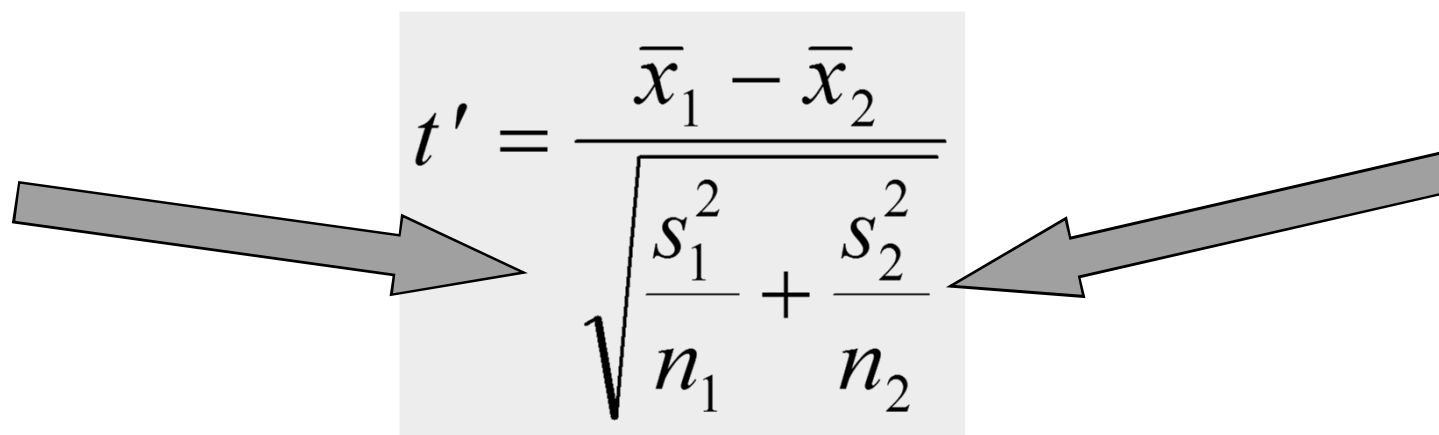




# A co když není splněna podmínka rovnosti variancí ?

užíváme

**Welchovu\* aproximaci t-testu**


$$t' = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

**se speciálně počítanými stupni volnosti**

$$df = \frac{(SE_1^2 + SE_2^2)^2}{SE_1^4 / (n_1 - 1) + SE_2^4 / (n_2 - 1)}$$

*SE = střední chyba průměru*

*(\*též známa pod názvem „Satterthwaite's methods“)*

# Dvouvýběrový t-test: jednostranná hypotéza; příklad 1

**EXAMPLE 8.2** A two-sample  $t$  test for the one-tailed hypotheses,  $H_0: \mu_1 \geq \mu_2$  and  $H_A: \mu_1 < \mu_2$  (which could also be stated as  $H_0: \mu_1 - \mu_2 \geq 0$  and  $H_A: \mu_1 - \mu_2 < 0$ ). The data are heights of plants, each grown with one of two different fertilizers.

$$H_0: \mu \geq \mu_2$$

$$H_A: \mu_1 < \mu_2$$

Present fertilizer	Newer fertilizer
48.2 cm	52.3 cm
54.6	57.4
58.3	55.6
47.8	53.2
51.4	61.3
52.0	58.0
55.2	59.8
49.1	54.8
49.9	
52.6	
$n_1 = 10$	$n_2 = 8$
$\nu_1 = 9$	$\nu_2 = 7$
$\bar{X}_1 = 51.91$ cm	$\bar{X}_2 = 56.55$ cm
$SS_1 = 102.23$ cm <sup>2</sup>	$SS_2 = 69.20$ cm <sup>2</sup>

$$s_p^2 = \frac{102.23 + 69.20}{9 + 7} = \frac{171.43}{16} = 10.71 \text{ cm}^2$$

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{10.71}{10} + \frac{10.71}{8}} = \sqrt{2.41} = 1.55 \text{ cm}$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}} = \frac{51.91 - 56.55}{1.55} = \frac{-4.64}{1.55} = -2.99$$

$$t_{0.05(1), 16} = 1.746$$

As  $t$  of  $-2.99$  is less than  $-1.746$ ,  $H_0$  is rejected.

$$0.0025 < P < 0.005 \quad [P = 0.0043]$$

# Dvouvýběrový t-test; příklad 2

Existují rozdíly v hmotnosti žen a mužů studujících 2. ročník biologie na UP?

## Descriptive Statistics Section

Variable	Count	Mean	Standard Deviation	Standard Error	95.0% LCL of Mean	95.0% UCL of Mean
pohlavix=m	16	70.3125	8.364359	2.09109	65.85545	74.76955
pohlavix=ž	12	59.41667	8.061788	2.327238	54.29445	64.53888

## Tests of Assumptions Section

Assumption	Value	Probability	Decision(.050)
Skewness Normality (pohlavix=m)	-1.4629	0.143505	Cannot reject normality
Kurtosis Normality (pohlavix=m)	0.9918	0.321314	Cannot reject normality
Omnibus Normality (pohlavix=m)	3.1236	0.209763	Cannot reject normality
Skewness Normality (pohlavix=ž)	0.0858	0.931622	Cannot reject normality
Kurtosis Normality (pohlavix=ž)	0.4145	0.678507	Cannot reject normality
Omnibus Normality (pohlavix=ž)	0.1792	0.914309	Cannot reject normality
Variance-Ratio Equal-Variance Test	1.0765	0.920826	Cannot reject equal variances
Modified-Levene Equal-Variance Test	0.003	0.956722	Cannot reject equal variances

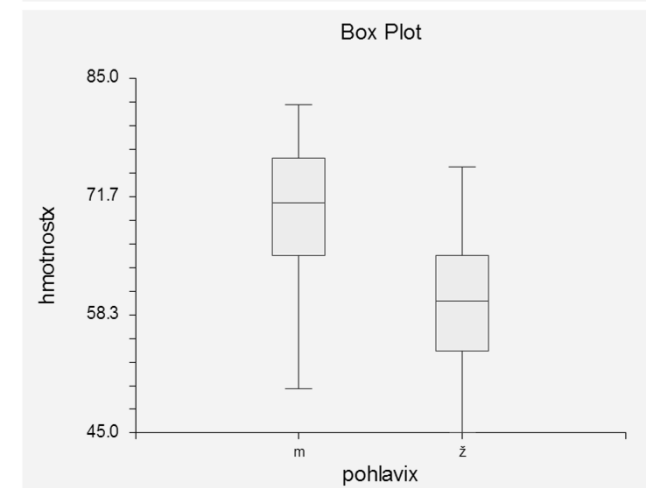
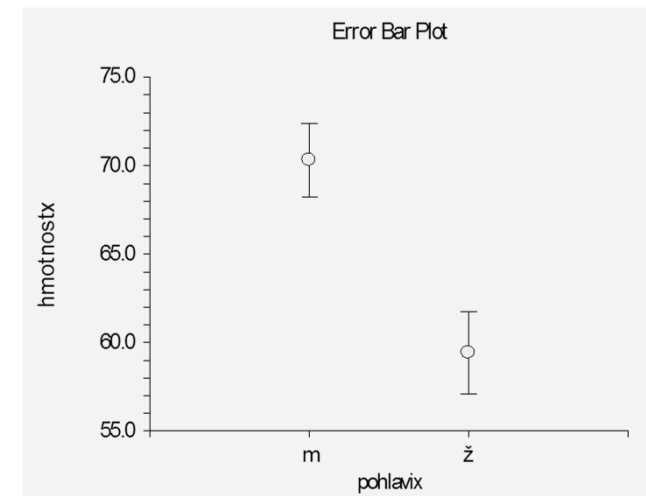
## Equal-Variance T-Test Section

Alternative Hypothesis	T-Value	Prob Level	Reject H0 at .050
Difference <> 0	3.4636	0.00186	Yes
Difference < 0	3.4636	0.99907	No
Difference > 0	3.4636	0.00093	Yes
Difference: (pohlavix=m)-(pohlavix=ž)			

## Aspin-Welch Unequal-Variance Test Section

Alternative Hypothesis	T-Value	Prob Level	Reject H0 at .050
Difference <> 0	3.4826	0.001898	Yes
Difference < 0	3.4826	0.999051	No
Difference > 0	3.4826	0.000949	Yes

Proveden náhodný výběr 28 osob (16 mužů a 12 žen)



# Test shody rozptylů (variancí): F-test

$$H_0: \sigma_1 = \sigma_2$$
$$H_A: \sigma_1 \neq \sigma_2$$

- Použijeme pouze na data kvantitativní.
- Lze použít i jednostranný test

**F-test\***

$$F = \frac{s_1^2}{s_2^2}$$

kde v čitateli je větší z obou  $s^2$  !!!!!

Studovaná veličina (F) má tzv. Fisher-Snedecorovo (F) rozdělení se dvěma parametry: stupni volnosti čitatele a jmenovatele

**je-li  $F_{\text{test}} \geq F_{\alpha(2), v1, v2}$  ..... zamítáme  $H_0$**

- test je **slabý**, velmi náchylný na nedodržení normálního rozdělení srovnávaných souborů, zvl. pro malé soubory

\*Alternativou jak porovnat rozptýlenost ve skupinách měření je užít **Levenův test** (*Levene test*): ten lze užít najednou i pro více jak 2 soubory, je robustní na odchylky od normálního rozdělení, více zde.

# F-rozdělení- tabulky (část)

$F_{0,95}(\nu_1, \nu_2)$	$\nu_1$							
	1	2	3	4	5	6	8	10
$\nu_2$	1	2	3	4	5	6	8	10
	1	2	3	4	5	6	8	10
1	161,45	199,50	215,71	224,58	230,16	233,99	238,88	241,88
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,40
3	10,13	9,55	9,28	9,12	9,01	8,94	8,85	8,79
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,96
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,74
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,06
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,64
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,35
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,14
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,98
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,85
12	4,75	3,89	3,49	3,26	3,11	3,00	2,85	2,75
13	4,67	3,81	3,41	3,18	3,03	2,92	2,77	2,67
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,60
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,54
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,49
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,45
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,41
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,38
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,35
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,32
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,30
23	4,28	3,42	3,03	2,80	2,64	2,53	2,37	2,27
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,25
25	4,24	3,39	2,99	2,76	2,60	2,49	2,34	2,24
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,16
35	4,12	3,27	2,87	2,64	2,49	2,37	2,22	2,11
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,08
45	4,06	3,20	2,81	2,58	2,42	2,31	2,15	2,05
50	4,03	3,18	2,79	2,56	2,40	2,29	2,13	2,03
60	4,00	3,15	2,76	2,53	2,37	2,25	2,10	1,99
70	3,98	3,13	2,74	2,50	2,35	2,23	2,07	1,97
80	3,96	3,11	2,72	2,49	2,33	2,21	2,06	1,95
90	3,95	3,10	2,71	2,47	2,32	2,20	2,04	1,94
100	3,94	3,09	2,70	2,46	2,31	2,19	2,03	1,93

# Interval spolehlivosti pro rozdíl mezi dvěma průměry

$$\bar{d} \pm t_{\frac{\alpha}{2}, n_1 + n_2 - 2} s_{\bar{x}_1 - \bar{x}_2} \quad \bar{d} = \mu_1 - \mu_2$$

**Pokud neexistuje významný rozdíl mezi průměry výběrů,  
lze očekávat,  
že interval spolehlivosti pro rozdíl středních hodnot  
bude zahrnovat nulu.**

# Neparametrický Mann-Whitney U-test

- místo změřených hodnot pracujeme s jejich pořadím
- data seřadíme sestupně či vzestupně (zde sestupně) bez ohledu na různé soubory
- Co testujeme:

$H_0$ : Rozdělení obou skupin je shodné (mediány se rovnají\*).

$H_A$ : Rozdělení obou skupin se liší (mediány se liší\*).

$$U = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

$$U' = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2$$

(stejným hodnotám dáváme průměrné pořadí)

\* tato hypotéza má jeden specifický předpoklad: rozdělení hodnot (distribuce) je shodné u obou skupin, a liší se jen pozicí „mediánu“, což většinou nebývá splněno, proto se používá obecnější hypotéza

$R_i$  = součet pořadí v souboru  $i$

$n_i$  = počet prvků v souboru  $i$

$U + U' = n_1 n_2$ ; hodnoty se pohybují v intervalu  $\langle 0; n_1 n_2 \rangle$

větší z obou  $U$  porovnáme s kritickou hodnotou  $U_{\alpha(2), n_1, n_2}$

je-li  $U$  či  $U' > U_{\text{krit}}$ , zamítáme  $H_0$

(v případě řazení vzestupného hledáme menší z obou  $U$ )

# U-test: příklad

Příklad. Výkon 18 gymnastek byl ohodnocen stanovením jejich pořadí od nejlepší (pořadí 1) po nejslabší (pořadí 18). V této skupině bylo  $n_1 = 11$  žákyň trenérky A a  $n_2 = 7$  žákyň trenérky B. Na základě výsledků (pořadí) shrnutých v tabulce se má posoudit nulová hypotéza  $H_0$ : "účinnost výukových metod obou trenérek se neliší".

Trenérka:

A	B
1	2
4	3
5	6
7	9
8	12
10	15
11	18
13	
14	
16	
17	
$n_1 = 11$	$n_2 = 7$

Kontrola:

$$R_1 = 106 \quad R_2 = 65$$

$$R_1 + R_2 = \frac{19 \cdot 18}{2} = 171$$

$$U_1 = 7+5+5+4+4+3+3+2+2+1+1=37, \text{ anebo}$$

$$U_1 = 11 \cdot 7 + 11 \cdot 12 / 2 - 106 = 37 \text{ (kontrola)}$$

$$U_2 = 10+10+8+6+4+2+0=40, \text{ anebo}$$

$$U_2 = 11 \cdot 7 + 7 \cdot 8 / 2 - 65 = 40 \text{ (kontrola)}$$

V tabulce U-rozdělení najdeme - pro výběry rozsahu 7 a 11,  $\alpha=0,05$  a oboustranný test - kritickou hodnotu  $U_{0,05}(7,11)=61$ .



# U-test: tabulky - část

**TABLE B.11 (cont.)** Critical Values of the Mann-Whitney  $U$  Distribution

		$\alpha(2):$ $\alpha(1):$	0.20 0.10	0.10 0.05	0.05 0.025	0.02 0.01	0.01 0.005	0.005 0.0025	0.002 0.001	0.001 0.0005
$n_1$	$n_2$									
6	9	39	42	44	47	49	50	52	53	
	10	43	46	49	52	54	55	57	58	
	11	47	50	53	57	59	60	62	64	
	12	51	55	58	61	63	65	68	69	
	13	55	59	62	66	68	70	73	74	
	14	59	63	67	71	73	75	78	79	
	15	63	67	71	75	78	80	83	85	
	16	67	71	75	80	83	85	88	90	
	17	71	76	80	84	87	90	93	95	
	18	74	80	84	89	92	95	98	100	
	19	78	84	89	94	97	100	103	106	
	20	82	88	93	98	102	105	108	111	
	21	86	92	97	103	107	110	114	116	
	22	90	96	102	108	111	115	119	121	
	23	94	101	106	112	116	120	124	126	
	24	98	105	111	117	121	125	129	132	
	25	102	109	115	121	126	130	134	137	
	26	106	113	119	126	131	134	139	142	
	27	110	117	124	131	135	139	144	147	
	28	114	122	128	135	140	144	149	152	
	29	118	126	132	140	145	149	154	157	
	30	122	130	137	145	150	154	159	163	
6	31	125	134	141	149	154	159	164	168	
	32	129	138	146	154	159	164	169	173	
	33	133	142	150	158	164	169	174	178	
	34	137	147	154	163	169	174	179	183	
	35	141	151	159	168	173	179	185	188	
	36	145	155	163	172	178	184	190	194	
	37	149	159	167	177	183	188	195	199	
	38	153	163	172	182	188	193	200	204	
	39	157	167	176	186	193	198	205	209	
	40	161	172	181	191	197	203	210	214	
7	7	36	38	41	43	45	46	48	49	
	8	40	43	46	49	50	52	54	55	
	9	45	48	51	54	56	58	60	61	
	10	49	53	56	59	61	63	65	67	
	11	54	58	61	65	67	69	71	73	
	12	58	63	66	70	72	75	77	79	
	13	63	67	71	75	78	80	83	85	
	14	67	72	76	81	83	86	89	91	
	15	72	77	81	86	89	92	95	97	
	16	76	82	86	91	94	97	101	103	
	17	81	86	91	96	100	103	106	109	

# Kvantitativní a ordinální proměnné

Hypotézy o dvou výběrech

—

blokové (= párové) uspořádání

# Testování průměrného rozdílu: parametrický párový t-test (*paired t-test*)

- v testu neužíváme původní změřené hodnoty,  
ale rozdíly příslušných párů pozorování (měření;  $d_i$ )

Oboustranný test:

$$\begin{aligned} H_0: \mu_d &= 0 \\ H_A: \mu_d &\neq 0 \end{aligned}$$

Jednostranný test:

$$\begin{aligned} H_0: \mu_d &\geq 0 \\ H_A: \mu_d &< 0 \end{aligned}$$

(nebo obráceně)

$$t = \frac{\bar{d}}{s_{\bar{d}}}$$

$$v = n-1 \text{ párů}$$

je-li  $|t_{\text{test}}| > t_{\alpha(2),v}$   
→ zamítáme  $H_0$

Pokud  $t \leq -t_{\alpha(1),v}$   
→ zamítáme  $H_0$   
 $t \geq t_{\alpha(1),v}$  v případě  $H_A: \mu_d > 0$

# Párový t-test: příklad

□ **Příklad:** Osmi rostlinám tabáku byl odebrán druhý list. Jedna náhodně vybraná polovina listu byla ošetřena přípravkem A, druhá přípravkem B. Potom byly listy potřeny suspenzí agresora a byl sledován počet skvrn na každé polovině.

Výsledky pokusu jsou

Rostlina	Počet skvrn		Rozdíl
	Přípravek A	Přípravek B	
1	9	10	-1
2	17	11	6
3	31	18	13
4	18	14	4
5	7	6	1
6	8	7	1
7	20	17	3
8	10	5	5

$H_0$ : Není rozdíl v účinku přípravků A a B.

$H_A$ : Je rozdíl v účinku přípravků A a B.

Postup:

1. spočítáme průměr rozdílů

2. spočítáme standardní odchylku a standardní chybu:

Průměr	s	$s_x$
4	4.309458	1.52362

3. dosadíme do vzorce t-testu:

$$t = 4/1,52 = 2,625$$

4. porovnáme s kritickou hodnotou  $t_{0,05(2), 7} = 2,365.....$   $t > t_{krit}$  - zamítáme  $H_0$

# Neparametrický Wilcoxonův pořadový test

## Co se testuje:

$H_0$ : **Není systematická difference uvnitř párů**  
(medián rozdílů  $M$  je nulový).

$H_1$ : **Je systematická difference uvnitř párů**  
(medián rozdílů  $M$  je různý od nuly).

- testovací procedura zahrnuje počítání rozdílů, jimž se přiřadí pořadí bez ohledu na znaménko od nejmenšího po největší
  - sečtou se pořadí se znaménky + a - zvlášť

➡ **výsledek jsou testové statistiky  $T_+$  a  $T_-$**

**Oboustranný test:** je-li menší z obou  $T < T_{\alpha(2),n}$  ....zamítáme  $H_0$

**Jednostranný test:**  $H_0$  zamítáme pro

(a)  $H_A: M < a$ , když  $T^+ < S_{\alpha(1),n}$  nebo pro

(b)  $H_A: M > a$ , když  $T^+ > n - S_{\alpha(1),n}$

# Wilcoxonův pořadový test - příklad

□ **Příklad:** Osmi rostlinám tabáku byl odebrán druhý list. Jedna náhodně vybraná polovina listu byla ošetřena přípravkem A, druhá přípravkem B. Potom byly listy potřeny suspenzí agresora a byl sledován počet skvrn na každé polovině.

Výsledky pokusu jsou

Rostlina	Počet skvrn		Rozdíl
	Přípravek A	Přípravek B	
1	9	10	-1
2	17	11	6
3	31	18	13
4	18	14	4
5	7	6	1
6	8	7	1
7	20	17	3
8	10	5	5

Uspořádáme rozdíly:

Uspořádané rozdíly	1	-1	1	3	4	5	6	13
Pořadí rozdílů	2	2	2	4	5	6	7	8

Součet pořadí záporných rozdílů  $W_- = 2$  a protože platí  $W_+ + W_- = n(n+1)/2 \Rightarrow W_+ = 8 * 9/2 - 2 = 34$ .  $W = \min(2, 34) = 2$ .

Počet nenulových rozdílů je  $n = 8$  a pro  $\alpha = 0,05$  je kritická hodnota 3,7. Protože  $W = 2 < 3,7$  zamítáme hypotézu stejné účinnosti přípravků A a B.

# Wilcoxon test: tabulky - část

**TABLE B.12** Critical Values of the Wilcoxon  $T$  Distribution

	$\alpha(2) = 0.50$ $\alpha(1) = 0.25$	0.20 0.10	0.10 0.05	0.05 0.025	0.02 0.01	0.01 0.005	0.005 0.0025	0.001 0.0005
$n$								
4	2	0						
5	4	2	0					
6	6	3	2	0				
7	9	5	3	2	0			
8	12	8	5	3	1	0		
9	16	10	8	5	3	1	0	
10	20	14	10	8	5	3	1	
11	24	17	13	10	7	5	3	0
12	29	21	17	13	9	7	5	1
13	35	26	21	17	12	9	7	2
14	40	31	25	21	15	12	9	4
15	47	36	30	25	19	15	12	6
16	54	42	35	29	23	19	15	8
17	61	48	41	34	27	23	19	11
18	69	55	47	40	32	27	23	14
19	77	62	53	46	37	32	27	18
20	86	69	60	52	43	37	32	21
21	95	77	67	58	49	42	37	25
22	104	86	75	65	55	48	42	30
23	114	94	83	73	62	54	48	35
24	125	104	91	81	69	61	54	40
25	136	113	100	89	76	68	60	45
26	148	124	110	98	84	75	67	51
27	160	134	119	107	92	83	74	57
28	172	145	130	116	101	91	82	64
29	185	157	140	126	110	100	90	71
30	198	169	151	137	120	109	98	78
31	212	181	163	147	130	118	107	86
32	226	194	175	159	140	128	116	94
33	241	207	187	170	151	138	126	102
34	257	221	200	182	162	148	136	111
35	272	235	213	195	173	159	146	120
36	289	250	227	208	185	171	157	130
37	305	265	241	221	198	182	168	140
38	323	281	256	235	211	194	180	150
39	340	297	271	249	224	207	192	161
40	358	313	286	264	238	220	204	172

Tří a více souborů ordinální  
nebo kvantitativní proměnné

-


statistické testy



# Analýza variance (*Analysis of variance*; ANOVA)

-máme-li více než dva výběry, testování rozdílů mezi průměry provádíme vždy analýzou variance,  
ne testováním dílčích hypotéz vždy pro dva výběry → PROČ?

Nejsme schopni udržet chybu I. druhu na požadované hladině pro všechny prováděné testy !!!



Počet testů	Pravděpodobnost výskytu alespoň 1 chyby I. druhu ( $\alpha=0,05$ )
3	0,14
10	0,40
45	0,90
C	$1-(1-\alpha)^C$

- rozlišujeme jedno (*one-way*) a více-cestné (*multiple, multifactor*) ANOVy → rozumíme tím (simultánní) analýzu vlivu jedné či více kategoriálních proměnných (faktorů, prediktorů) na „závislou“ kvantitativní nebo ordinální proměnnou

→ různé úrovně daného faktoru se nazývají hladiny (*treatment level*)  
(zde na okraj: existují i jiná - složitější - uspořádání ANOVy!!!)

# ANOVA - příklady

## Příklad č. 1:

liší se průměrný počet květů jedinců vstavače pletového mezi 5 populacemi?

## *Řešení:*

jednocestná ANOVA, faktor: populace, počet hladin: 5 (populací)

## Příklad č. 2:

jaký je vliv 3 druhů hnojiv a 3 intenzit zálivky na růst vybraného druhu?

## *Možné (vhodnější) řešení:*

jeden pokus (a ne dva nezávislé pokusy!!!), dvoucestná ANOVA, faktor č. 1: hnojení, počet hladin: 3 (např. hnojivo 1, hnojivo 2, hnojivo 3), faktor č. 2: zálivka, počet hladin: 3 (např. kontrola, zálivka malým a větším množstvím vody) – jejich kombinace 3x3 hladiny = 9

# Jednocestná ANOVA (*One-way ANOVA*)

- testujeme vliv alespoň 2 (*nejčastěji 3 a více*) hladin jednoho faktoru na kvantitativní proměnnou

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_k$$

$k$  = počet hladin faktoru (tj. počet srovnávaných souborů)

## PŘEDPOKLADY ANOVY:

- výběry pocházejí z téhož normálního rozdělení (a mají tedy stejný parametr = střední hodnotu), vyplývá z toho též rovnost (jejich) variancí  $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$ \*
- znáhodněné uspořádání pokusu (hodnoty jsou vzájemně nezávislé) !!!
- reziduály jsou normálně distribuovány
- vzorky jsou korektně klasifikovány
- hlavní efekty jsou aditivní
- (+ !!! předpoklad pro náš výklad:  $n_1 = n_2 = \dots = n_k$ )

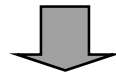
\*Doporučuje se otestovat tuto hypotézu pomocí např. **Bartletova testu**

# Jednocestná ANOVA

Pokud platí  $H_0$ , měly by být výběry identické. Pak lze na všechny výběry pohlížet jako na výběry vybrané z jedné a téže populace.

Pak máme dva alternativní způsoby, jak odhadnout varianci  $\sigma^2$  této populace:

1. průměrný rozptyl všech výběrů  $s_p^2$ ,  
nebo
2. usuzovat ze zjištěného  $s_x^2$  – zjištěného rozptylu mezi  
průměry výběrů

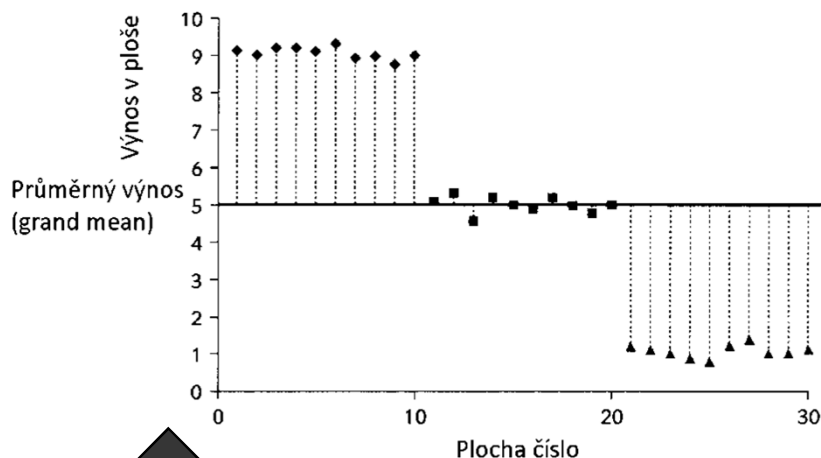


Je-li  $H_0$  pravdivá,  
tyto dva odhady by měly být stejné !!!

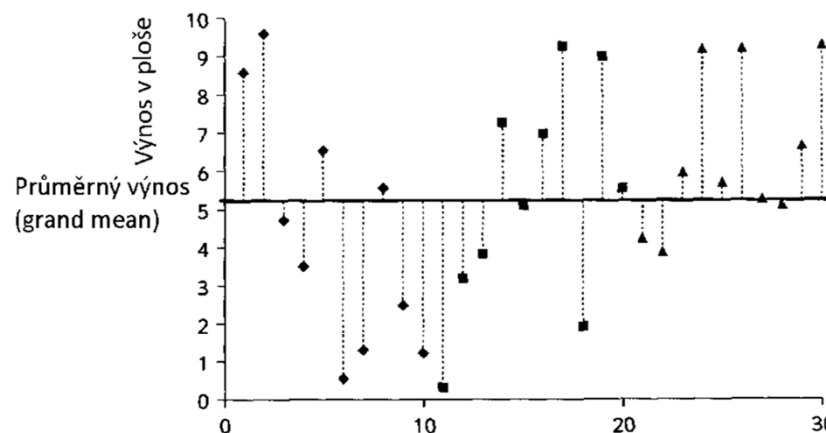
# Dělení variability v Anově

Vždy 3 soubory ploch pod vlivem 3 hladin hnojiva: A (plocha 1-10), B (11-20) a C (21-30).

Datový příklad 1

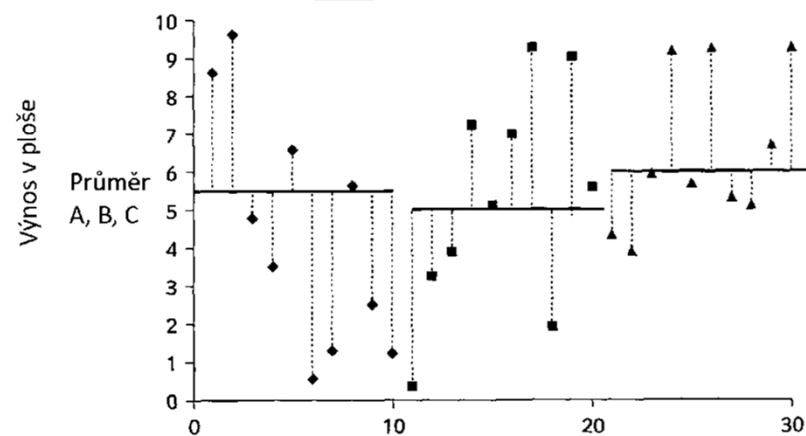
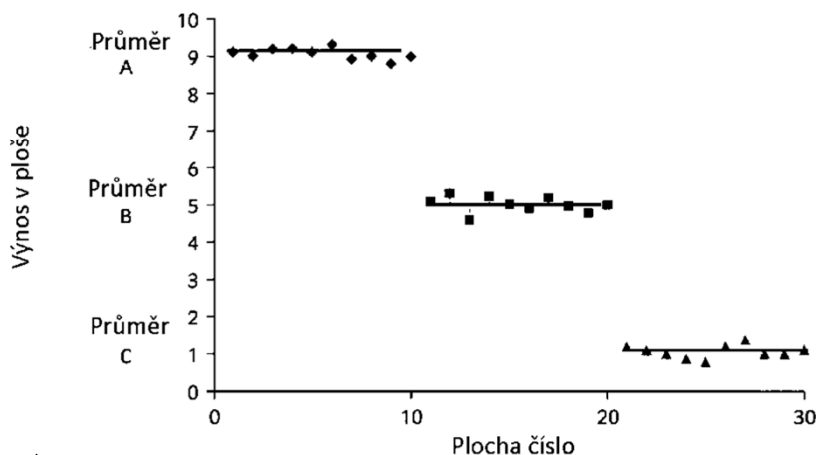


Datový příklad 2



$$\bar{x} \approx \mu$$

↑ Variabilita kolem celkového průměru  $\bar{x}$  (grand mean) ↑



$$\bar{x}_i$$

↑ Variabilita kolem průměrů skupin  $\bar{x}_i$  (souborů, treatmentů) ↑

# ANOVA – jak je to s těmi variancemi

Každá odchylka konkrétní hodnoty od celkového průměru

lze rozložit na:  $(x_{ij} - \bar{x}) = (x_{ij} - \bar{x}_i) + (\bar{x}_i - \bar{x})$

Pak celkový součet čtverců  $SS_{TOT}$  reprezentuje

celkovou variabilitu souboru dat =  $SS_{TOT} = \sum_{i=1}^n (x_{ij} - \bar{x})^2$

= součet čtverců uvnitř skupin  $SS_E$  + součet čtverců mezi skupinami  $SS_G$

$$SS_E = \sum_{i=1}^k \left[ \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 \right]$$

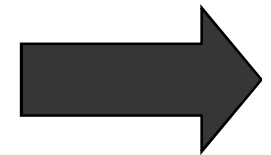
$$SS_G = \sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2$$

Stupně volnosti:

$$DF_{TOTAL} = DF_E + DF_G$$

$$DF_E = \sum_{i=1}^k (n_i - 1) = N - k$$

$$DF_G = k - 1$$



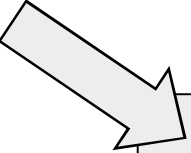
kde  $k$ =počet skupin,  $n_i$ =velikost skupiny  $i$ ,  $N$ =počet všech hodnot (tj. suma  $n_i$ )

# ANOVA – jak je to konečně s těmi variancemi

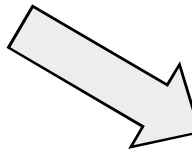
$$\text{Model: } X_{ij} = \mu + A_i + \varepsilon_{ij}$$

$$\text{Model v případě platnosti } H_0: X_{ij} = \mu + \varepsilon_{ij}$$

Průměrný čtverec uvnitř skupin  
(=  $s_p^2$ )


$$MS_E = \frac{SS_E}{DF_E} = \frac{SS_E}{N - k}$$

Průměrný čtverec mezi  
skupinami (=  $s_x^2$ )


$$MS_G = \frac{SS_G}{DF_G} = \frac{SS_G}{k - 1}$$

Připomeneme: Je-li  $H_0$  pravdivá,  
tyto dva odhady by měly být stejné !!!

Jejich poměr by tak měl být roven zhruba jedné.

Toto porovnání provedeme pomocí F-testu:

$$F = \frac{MS_G}{MS_E}$$

je-li  $F \geq F_{\alpha(1), (k-1), (N-k)}$  zamítáme  $H_0$

# Jak chápat poměr variancí?

Celkovou variabilitu kolem střední hodnoty (celkového průměru) můžeme rozložit na variabilitu:

variabilitu nevysvětlenou  
(= „náhodnou“)



vysvětlenou  
(= „různými“ skupinami)



*Jiný pohled  
na ANOVu*

$$F = \frac{MS_G}{MS_E} = \frac{\text{nevysvětlený} + \text{vysvětlený rozptyl}}{\text{nevysvětlený rozptyl}}$$

Jestliže  $H_0$  není pravdivá, čítec vzroste.

Jmenovatel však bude stále odhadem  $\sigma^2$ .

Co to tedy znamená, když zamítneme nulovou hypotézu:  
že **ALESPON** jedna skupina (průměr) se odlišuje od ostatních.  
(Ale která?)





# Jednoduchá ANOVA - příklad

Byly zjišťovány váhy u tří plemen králíků chovaných za standardních podmínek. Od každého plemene bylo chováno 5 kusů (tj.  $n_i=5$  pro všechna  $i$ ). Cílem bylo zjistit, zda se za daných podmínek budou váhy tří porovnávaných plemen lišit. Na konci pokusu byly zjištěny následující váhy (v kg):

Plemeno 1: 3,3,4,5,5

Plemeno 2: 4,4,5,6,6

Plemeno 3: 5,5,6,7,7

Z toho spočteme průměry:  $\bar{X}_1 = 4, \bar{X}_2 = 5, \bar{X}_3 = 6$ , celkový průměr  $\bar{X} = 5$ . Počet stupňů volnosti v každé skupině je čtyři, tzn.  $DF_e = 4 + 4 + 4 = 12$ ,  $DF_G = 3 - 1 = 2$ . Součet čtverců odchylek uvnitř skupin je podle 8.1  $SS_e = (3-4)^2 + (3-4)^2 + (4-4)^2 + (5-4)^2 + (5-4)^2 + (4-5)^2 + (4-5)^2 + (5-5)^2 + (6-5)^2 + (6-5)^2 + (5-6)^2 + (5-6)^2 + (6-6)^2 + (7-6)^2 + (7-6)^2 = 12$ , součet čtverců odchylek mezi skupinami je podle 8.4  $SS_G = 5 \times (4-5)^2 + 5 \times (5-5)^2 + 5 \times (6-5)^2 = 10$ . Po vydělení součtu čtverců počtem stupňů volnosti dostáváme  $MS_G = 10/2 = 5$ ,  $MS_e = 12/12 = 1$ .  $F = MS_G / MS_e = 5/1 = 5$ . Protože kritická hodnota  $F$  pro jednostranný test na pětiprocentní hladině významnosti při 2 a 12 stupních volnosti je 3.89, zamítáme nulovou hypotézu o rovnosti středních hodnot na 5%-ní hladině významnosti.

# Jednocestná ANOVA: počítačový výstup

**Příklad:** Existuje rozdíl ve výšce rostlin mezi 3 populacemi?  
Musí následovat otázka: jak jsou vybrány příslušné populace? To odpoví na otázku, zda-li po případném zamítnutí  $H_0$  provádět mnohonásobná porovnávání.

*Faktor (nezávislá proměnná)*

*Zdroj variability*

**Analysis of Variance Table**

Source	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level
A:Habitat	2	108.45	54.22	0.70	0.499
S(A)	73	5646.62	77.35		
Total	75	5755.07			

*reziduální*

$DF_G$   $DF_E$   $SS_G$   $SS_{TOT}$   $SS_E$   $MS_E$   $MS_G$   $F = MS_G / MS_E$

$DF_{TOT}$

*P-hodnota*

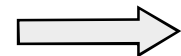
**Závěr:** Nelze zamítnout  $H_0$  o stejné průměrné výšce rostlin ve 3 populacích.

# Modely ANOVy

Chceme-li vědět více o tom, které soubory se od sebe liší (zamítli jsme  $H_0$ ), musíme nejdříve upozornit na to, že v ANOVě rozlišujeme 3 modely a podle toho se liší i hypotézy, které testujeme:

1. Model I. - model s pevnými efekty (*fixed effects*) – lze jít dále  
 $H_0$ : není rozdíl mezi průměry srovnávaných hladin
2. Model II. - model s náhodnými efekty (*random effects*) – analýzu ukončíme po F-testu  
 $H_0$ : variabilita uvnitř souborů je stejná či vyšší než mezi soubory
3. Model III. - smíšený model (*mixed model*) – kombinuje 1. a 2.  
(tento model je možné uvažovat až u dvou- a více-cestné ANOVy)

*A jak že se to zjišťuje?*



# Mnohonásobná porovnání

(*Multiple comparison tests*)

- testy nám umožní říci, které soubory (průměry) se od sebe liší
- problémy s udržení chyby I. druhu na zvolené hladině  
(řešení → **family(=experiment)-wise**  $\alpha$  [ $=\alpha(f)$ ] versus **comparison-wise**  $\alpha$  [ $\alpha(c)$ ]);
- chyba II. druhu se objevuje častěji než v F-testu ANOVy
- **a priori** (plánovaná porovnávání) a
- **a posteriori** (porovnávání po ANOVě) testy

**Máme opět několik možností jak srovnávat:**

A) každého s každým:

např.  $H_0: \mu_A = \mu_B$

**Tukey (HSD) test** (konzervativní)

**Student-Newman-Keuls test**

**Bonferroni test, Duncan test, LSD-test** (nejliberálnější)

B) každého s kontrolou

**Dunnet test**

C) mnohonásobné kontrasty (*multiple contrast*)

porovnáváme vybraný zásah oproti „průměrnému“ zásahu ostatních skupin: **Scheffé test**

**Bonferroni procedura**

$\alpha/c$

**Dunn-Šidák procedura**

$1-(1-\alpha)^{1/c}$

**Sekvenční Bonferroni**

Nejmenší P testuji  $\alpha/c$ ,  
druhé nejmenší P pak  
 $\alpha/(c-1)$ , další P... $\alpha/(c-2)$ ,.

$c$  = počet testů

$$q = \frac{\bar{x}_A - \bar{x}_B}{SE}, SE = \sqrt{\frac{s^2}{2} \left( \frac{1}{n_A} + \frac{1}{n_B} \right)}$$

$SE$  = střední chyba odhadu rozdílu  
průměrů A a B

$s^2 = MS_E$

$n$  = počet hodnot ve skupině A+B

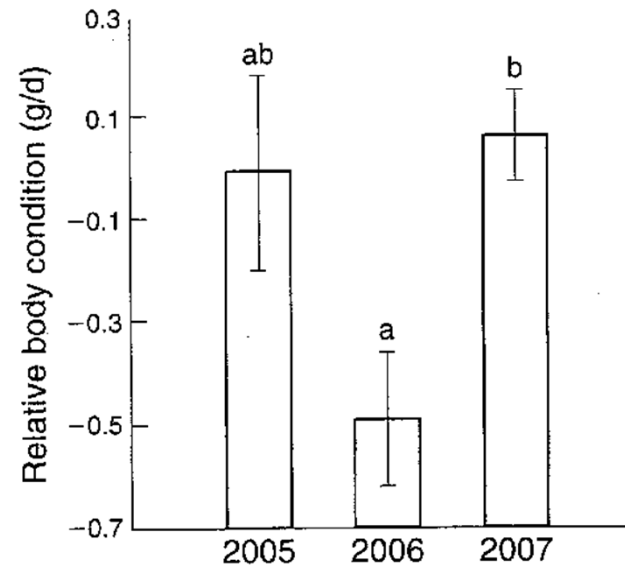
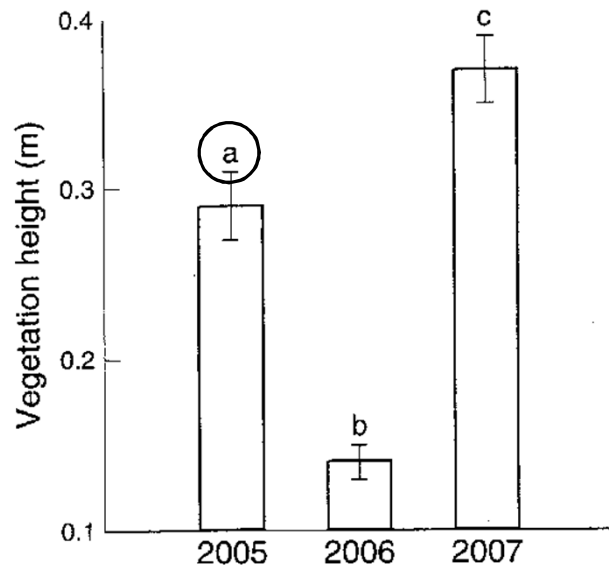
$q_{a,v,k}$  = „Studentized range“,

$v = DF_E$

$k$  = počet skupin

$\alpha$  = hladina významnosti

# Mnohonásobná porovnání: možnosti zobrazení výsledků



(Sperry & Weatherhead 2008, p. 2773)

Různá písmena (indexy) nad sloupečky (nebo u hodnot v tabulce) označují signifikantní rozdíly mezi průměry skupin ( $\pm$  SE)

**Table 4.** Mean above-ground biomass ( $\pm$  SE) of *Geum reptans* individuals with only stolons, only flowers, both reproductive modes, or no reproduction at all in a competition and a temperature experiment

Experiment	Mean above-ground biomass (g) of individuals with:			
	Stolons	Flowers	Stolons and flowers	No reproduction
Competition	1.13 $\pm$ 0.07 <sup>a</sup>	1.53 $\pm$ 0.13 <sup>a</sup>	1.57 $\pm$ 0.07 <sup>a</sup>	0.63 $\pm$ 0.14 <sup>c</sup>
Temperature	1.47 $\pm$ 0.05 <sup>b</sup>	1.63 $\pm$ 0.14 <sup>a</sup>	1.72 $\pm$ 0.04 <sup>a</sup>	1.34 $\pm$ 0.13 <sup>b</sup>

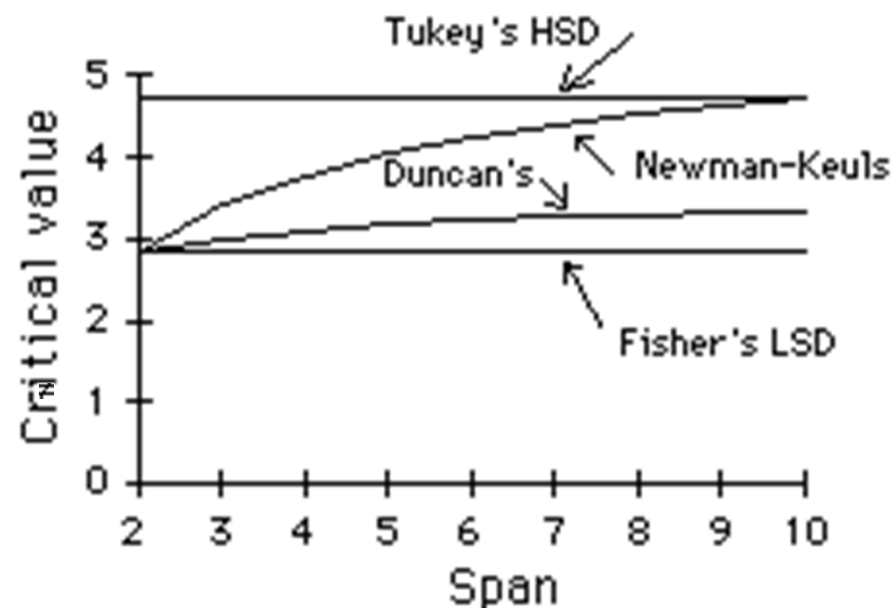
Different superscripts indicate significant differences among means,  $P < 0.05$ .

(Pluess & Stöcklin 2005, p. 234)

# „Každý s každým“ testy v ANOVĚ: doporučení

Př.: Kritické hodnoty pro experiment s 10 hladinami (4 opakování na zásah) jako funkce „vzdálenosti“ pořadí srovnávaných průměrů.

- **rozpětí průměrů neovlivňuje kritickou hodnotu ( $k$ ) pro Tukey's HSD**  
(nejkonzervativnější = nejvyšší  $k$  pro všechny testy; tj. zamítne nejméně  $H_0$ ) **a Fisher's LSD** (nejliberálnější = nejnižší  $k$  pro všechna srovnání; tj. zamítne nejvíce  $H_0$ )
- **Newman-Keuls a Duncan procedury zvyšují  $k$  jako funkci „vzdálenosti“ pořadí mezi průměry**



## Možné důsledky

Př.: Pokud je  $H_0$  akceptována pro 2 porovnání ze 3, např. takto: 1.  $\mu_1 \neq \mu_3$ , 2.  $\mu_1 = \mu_2$ , 3.  $\mu_2 = \mu_3 \rightarrow$  to ale není možné! Správný závěr: buď (a)  $\mu_1 \neq \mu_2$  nebo  $\mu_2 \neq \mu_3$  nebo (b) ani  $\mu_1$  ani  $\mu_3$  se rovná  $\mu_2$ .  
Potřebujete více dat pro rozhodnutí mezi těmito alternativami...

# Neparametrická jednocestná ANOVA

## *Kruskal-Wallisův test*

- lze ji aplikovat na data ze znáhodněného experimentálního uspořádání v případě nesplnění podmínek pro užití „parametrické“ jednocestné ANOVy
- je založena na pořadí dat, ne na jejich hodnotách

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1)$$

$n_i$  ... počet pozorování ve skupině  $i$ ,  $N$  ... je počet všech pozorování

$R_i$  ... součet pořadí pozorování skupiny  $i$ ,  $k$  ... počet skupin

- pro  $k < 6$  existují speciální tabulky, pro  $k > 5$  lze distribuci  $H$  aproximovat Pearsonovým ( $\chi^2$ ) rozdělením s  $k-1$  stupni volnosti
- lze provádět mnohonásobná porovnávání (např. K-W Z-test)

# Kruskal-Wallisův test-příklad

Entomolog studuje vertikální distribuci určitého druhu mouchy v listnatém lese a získal pět souborů jedinců z každé ze tří vegetačních vrstev (bylinné patro, keřové patro, stromové patro).

$H_0$ : Abundance druhu je stejná ve všech třech vegetačních patrech

$H_A$ : Abundance druhu není stejná ve všech třech vegetačních patrech

$\alpha = 0.05$

Data jsou následující (s pořadím - ranky - hodnot v závorkách):

Počet jedinců / m <sup>3</sup> patra		
Bylinné patro	Keřové patro	Stromové patro
14.0 (15)	8.4 (11)	6.9 (8)
12.1 (14)	5.1 (2)	7.3 (9)
9.6 (12)	5.5 (4)	5.8 (5)
8.2 (10)	6.6 (7)	4.1 (1)
10.2 (13)	6.3 (6)	5.4 (3)

$$n_1 = 5$$

$$R_1 = 64$$

$$n_2 = 5$$

$$R_2 = 30$$

$$n_3 = 5$$

$$R_3 = 26$$

$$N = 5 + 5 + 5 = 15$$

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1) = \frac{12}{15(16)} \left[ \frac{64^2}{5} + \frac{30^2}{5} + \frac{26^2}{5} \right] - 3(16) = \frac{12}{240} [1134.400] - 48 = 8.720$$

$$H_{0.05,5,5,5} = 5.780 \text{ Zamítáme tedy } H_0. 0.005 < P < 0.01$$



# Dvoucestná ANOVA

## (*Two-way ANOVA*)

Nejobecněji

- simultánní analýza vlivu dvou faktorů na studovanou proměnnou
- výhodná: stačí 1 experiment na analýzu vlivu obou faktorů + navíc umožňuje studium interakcí (*interactions*) mezi faktory

Typy dvoucestné ANOVy:

1. „vyvážená“ dvoufakt. ANOVA (*balanced, orthogonal two-way ANOVA*)
2. „nevyvážená“ dvoufakt. ANOVA (*unbalanced two-way ANOVA*)
3. dvoufakt. ANOVA bez replikací (*ANOVA without replications*)
4. znáhodněné bloky (*two-way ANOVA with randomised blocks*)
5. opakované měření (*two-way ANOVA with repeated-measures*)

# Dvoufaktorová ANOVA z pohledu počtu opakování

**vyvážená ANOVA**  
(balanced design)

Factor A

	Level 1	Level 2	Level 3	Level 4	No. of data per row
Level 1	XXX	XXX	XXX	XXX	12
Level 2	XXX	XXX	XXX	XXX	12
Level 3	XXX	XXX	XXX	XXX	12
No. of data per column	9	9	9	9	N = 36

(a) Equal replication per cell.

XXX	XXX	XXX	XXX	12
XXXX	XXXX	XXXX	XXXX	16
XX	XX	XX	XX	8
9	9	9	9	N = 36

(b) Equal replication within rows: proportional replication within columns.

**nevyvážená  
proporční  
ANOVA**  
(unbalanced  
proportional  
ANOVA)

**nevyvážená  
proporční  
ANOVA**  
(unbalanced  
proportional  
ANOVA)

XXX	XXX XXX	XXX XXX XXX	XXX XXX	24
XXXX	XXXX XXXX	XXXX XXXX XXXX	XXXX XXXX	32
XX	XX XX	XX XX XX	XX XX	16
9	18	27	18	N = 72

(c) Proportional replication within rows and within columns

XXX	XX	XX	XXXX	11
XXXX	XX	XX	XXX	11
XXXX	XXX	XXXX	XXX	14
11	7	8	10	N = 36

(d) Disproportional replication

**nevyvážená  
ANOVA**  
(unbalanced  
ANOVA)

**Dvoucestná  
ANOVA,**  
**1.(A) i 2.(B)**  
**faktor fixní,**  
**křížky=počet**  
**opakování**

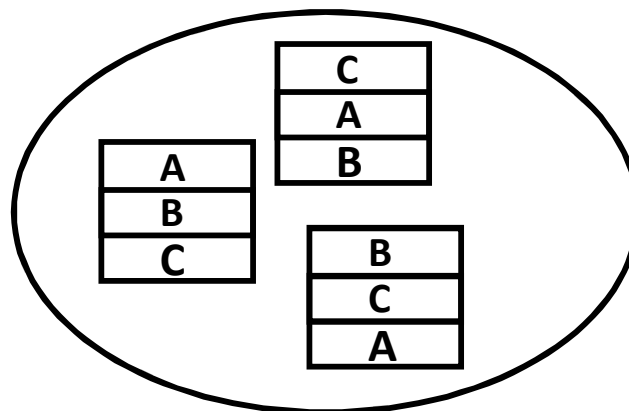
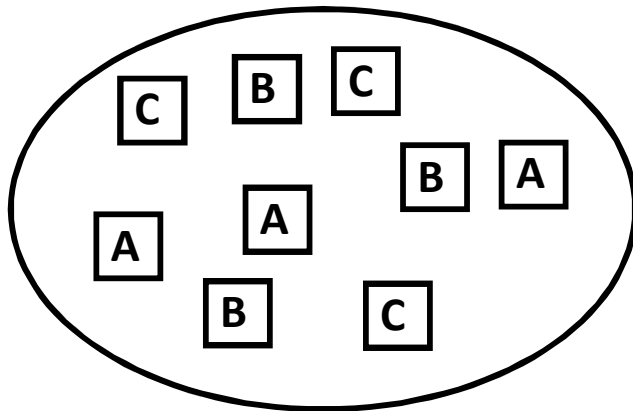
X	X	X	X	4
X	X	X	X	4
X	X	X	X	4
3	3	3	3	N = 12

(e) No replication

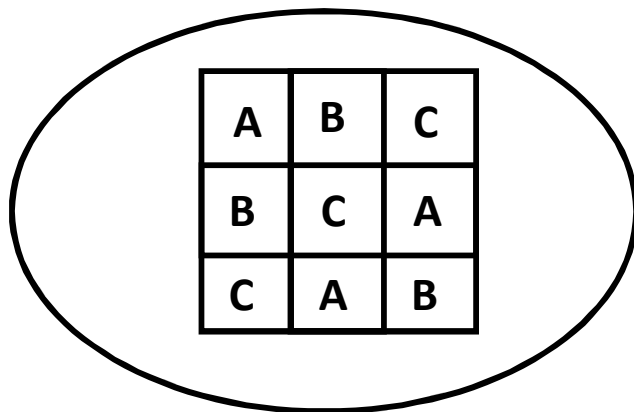
**ANOVA**  
**bez replikací**

# Možná uspořádání experimentů - ANOVA\*

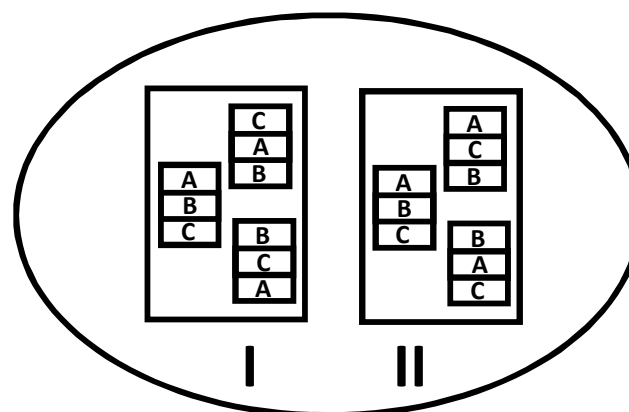
*Completely randomized design    Randomized block design*



*Latin square design*



*Split plot design*



A, B, C = zásahy = treatments

# Vyvážená dvoufaktorová ANOVA

- faktory jsou překřížené (*crossed*) - všechny hladiny faktoru 1 existují v kombinaci se všemi hladinami faktoru 2
- faktory jsou s pevnými efekty
- každá kombinace hladin obou faktorů se nazývá buňka (*cell*), počet měření ve všech buňkách je shodný (stejný počet opakování = *replications*)
- pokus je uspořádán jako „*completely randomized design*“
- analýza testuje celkem 3 hypotézy: dvě podobné jako v případě jednocestné ANOVy (*ALE: vliv faktoru nezávisle na faktoru druhém*), třetí reprezentuje testování interakce (*interaction*) mezi těmito dvěma faktory

## Model ANOVy:

pozorování = celkový průměr + vliv faktoru *A* + vliv faktoru *B* + vliv interakce *A* a *B* + náhodná variabilita

[výpočet (a interpretace) se liší s ohledem na model I., II. či III. ANOVy a v závislosti na vyváženosti modelu]

# Co je to interakce?

**Interakce (*interaction*):** rozumíme tím, že společný vliv faktorů je více než aditivní. Znamená to, že vliv faktoru A je závislý na presenci dílčí hladiny faktoru B (a opačně).

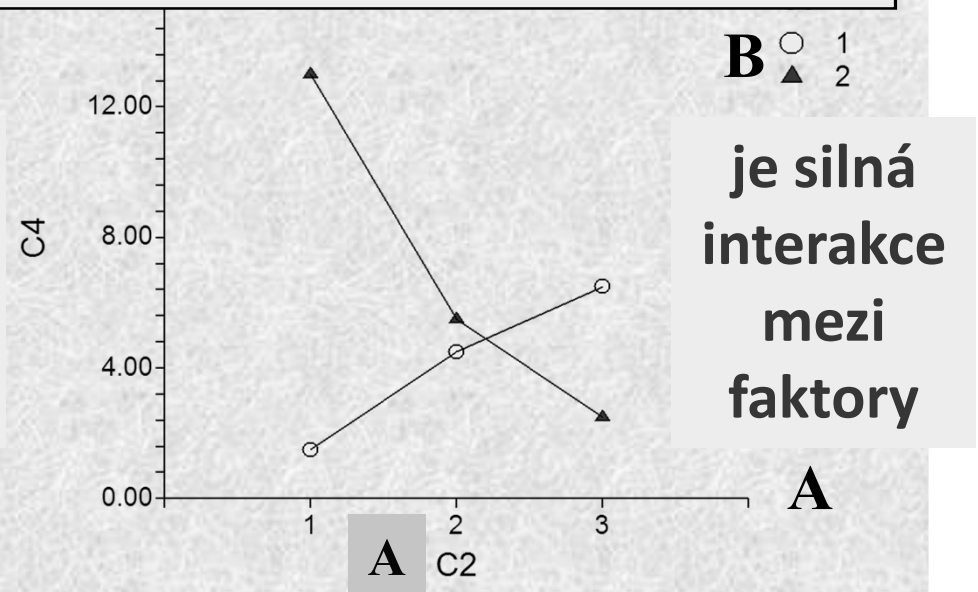
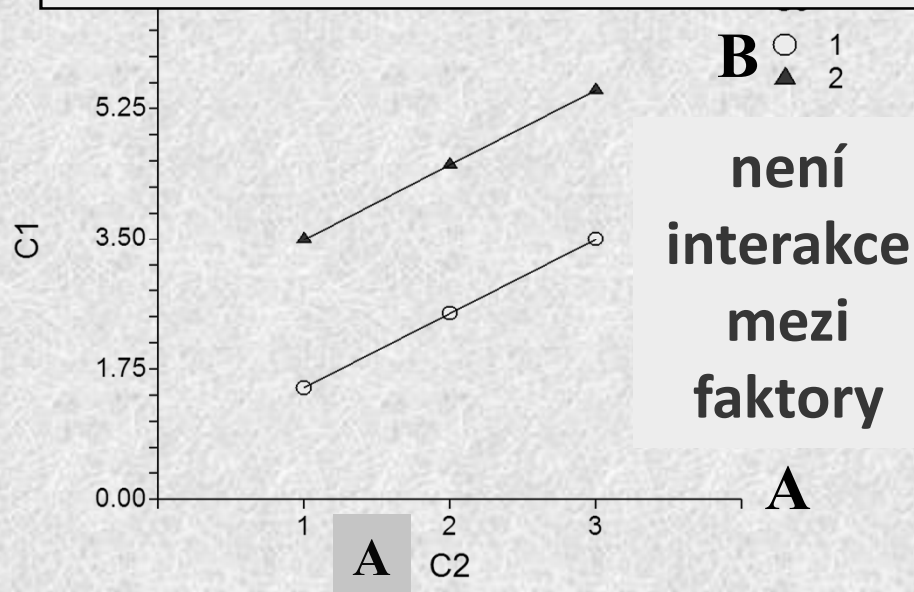
	Faktor A		
Faktor B	Hladina 1	Hladina 2	Hladina 3
Hladina 1	10	20	25
Hladina 2	20	30	35

	Faktor A		
Faktor B	Hladina 1	Hladina 2	Hladina 3
Hladina 1	10	20	25
Hladina 2	50	20	10

Je-li interakce, pak variabilita mezi buňkami není rovna součtu variability faktorů A a B

$$SS_{INT} = SS_{CELL} - SS_A - SS_B$$

$$DF_{INT} = (a-1)(b-1)$$



# Možné efekty dvou faktorů a jejich interakce

*Symboly značí průměry  
hladin faktorů*

$A_{1,2}$  ... faktor A 

$B_{1,2}$  ... faktor B 

I ... interakce

*Vliv faktoru:*

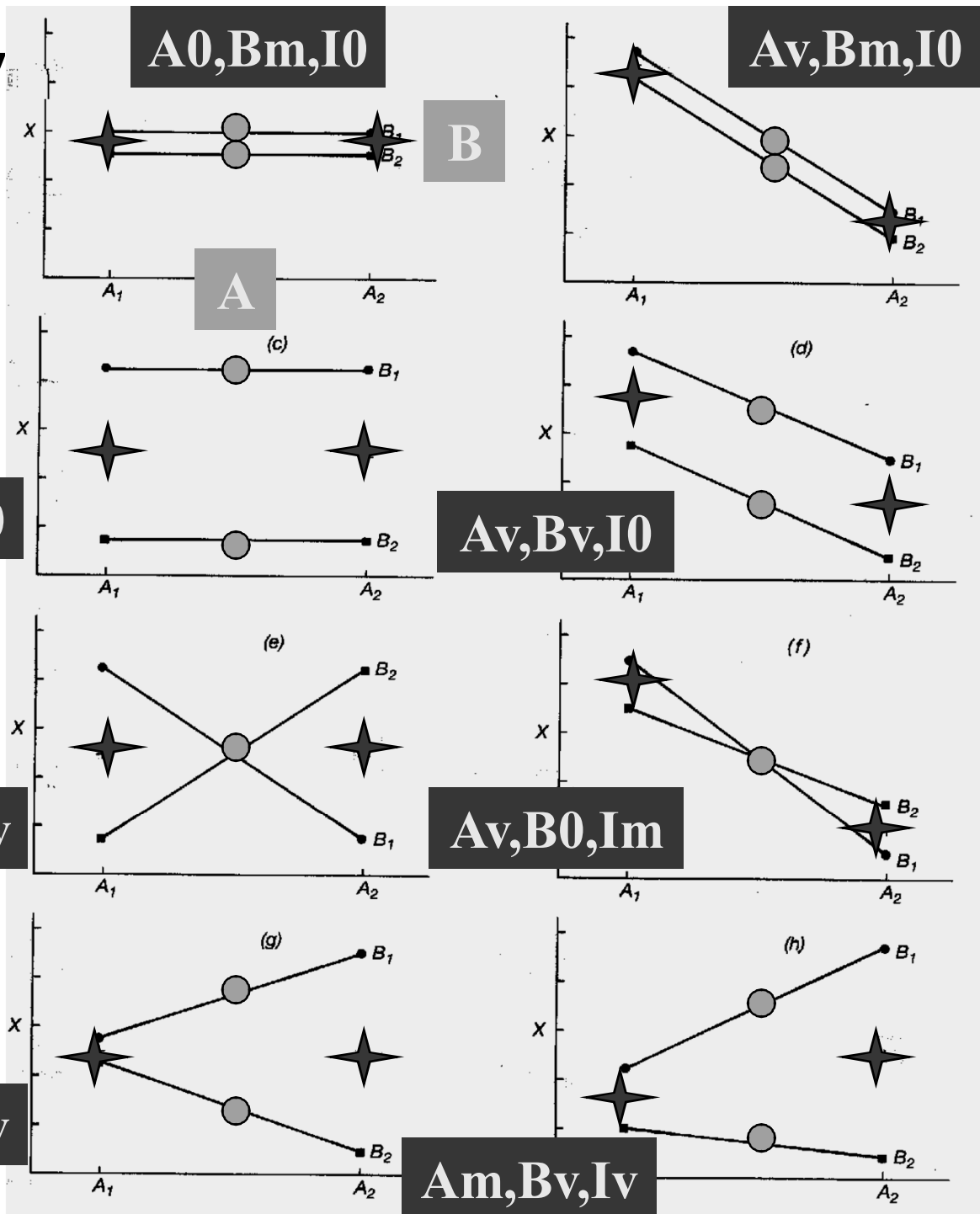
m ... malý efekt

v ... velký efekt

0 ... žádný efekt

*2 faktory, každý o dvou  
hladinách, faktoriální  
uspořádání*

*Zar  
(1996: 245),  
upraveno*



# Jak se počítá F-test pro různé modely dvoucestné ANOVy

**TABLE 12.3** Computation of the  $F$  Statistic for Tests of Significance in a Two-Factor ANOVA with Replication

Hypothesized effect	Model I (factors $A$ and $B$ both fixed)	Model II (factors $A$ and $B$ both random)	Model III (factor $A$ fixed: factor $B$ random)
Factor $A$	$\frac{\text{factor } A \text{ MS}}{\text{error MS}}$	$\frac{\text{factor } A \text{ MS}}{A \times B \text{ MS}}$	$\frac{\text{factor } A \text{ MS}}{A \times B \text{ MS}}$
Factor $B$	$\frac{\text{factor } B \text{ MS}}{\text{error MS}}$	$\frac{\text{factor } B \text{ MS}}{A \times B \text{ MS}}$	$\frac{\text{factor } B \text{ MS}}{\text{error MS}}$
$A \times B$ interaction	$\frac{A \times B \text{ MS}}{\text{error MS}}$	$\frac{A \times B \text{ MS}}{\text{error MS}}$	$\frac{A \times B \text{ MS}}{\text{error MS}}$

Testy modelu I. ANOVy jsou silnější než modelu II. a III., mám-li model III., síla testu pro fixní faktor  $A$  roste s počtem hladin náhodného (random) faktoru  $B$

# Dvoucestná ANOVA - příklad



Analysis of Variance Summary Table			
Source of variation	SS	DF	MS
Total	1827.6975	19	
Cells	1461.3255	3	
Factor A (hormone)	1386.1125	1	1386.1125
Factor B (sex)	70.3125	1	70.3125
A × B	4.9005	1	4.9005
Within cells (error)	366.3720	16	22.8982

For  $H_0$ : There is no effect of hormone treatment on the mean plasma calcium concentration of birds in the population sampled.

$$F = \frac{\text{hormone MS}}{\text{within-cells MS}} = \frac{1386.1125}{22.8982} = 60.5$$

$$F_{0.05(1), 1, 16} = 4.49$$

Therefore, reject  $H_0$ .

$$P < 0.0005 \quad [P = 0.00000080]$$

For  $H_0$ : There is no difference in mean plasma calcium concentration between male and female birds in the population sampled.

$$F = \frac{\text{sex MS}}{\text{within-cells MS}} = \frac{70.3125}{22.8982} = 3.07$$

$$F_{0.05(1), 1, 16} = 4.49$$

Therefore, do not reject  $H_0$ .

$$0.05 < P < 0.10 \quad [P = 0.099]$$

For  $H_0$ : There is no interaction of sex and hormone treatment affecting the mean plasma calcium concentration of birds in the population sampled.

$$F = \frac{\text{hormone} \times \text{sex interaction MS}}{\text{within-cells MS}} = \frac{4.9005}{22.8982} = 0.214$$

$$F_{0.05(1), 1, 16} = 4.49$$

Therefore, do not reject  $H_0$ .

$$P > 0.25 \quad [P = 0.65]$$

and data for a Model I, two-factor analysis of variance with equal

ations (in mg/100 ml) of birds of both sexes, half of the birds of each one and half not treated with the hormone.

ct of hormone treatment on the mean plasma calcium concentration of  $\mu_{\text{hormone}} = \mu_{\text{hormone}} \text{ or } \mu_{1.} = \mu_{2.}$ .

ct of hormone treatment on the mean plasma calcium concentration of  $\mu_{\text{hormone}} \neq \mu_{\text{hormone}} \text{ or } \mu_{1.} \neq \mu_{2.}$ .

rence in mean plasma calcium concentration between female and male  $\mu_{\text{female}} = \mu_{\text{male}} \text{ or } \mu_{.1} = \mu_{.2}$ .

ice in mean plasma calcium concentration between female and male  $\mu_{\text{female}} \neq \mu_{\text{male}} \text{ or } \mu_{.1} \neq \mu_{.2}$ .

raction of sex and hormone treatment on the mean plasma calcium birds.

ion of sex and hormone treatment on the mean plasma calcium con- ds.

Hormone Treatment		Hormone Treatment	
	Male	Female	Male
	14.5	39.1	32.0
	11.0	26.2	23.8
	10.8	21.3	28.8
	14.3	35.8	25.0
	10.0	40.2	29.3
74.4	$\sum_{l=1}^5 X_{12l} = 60.6$	$\sum_{l=1}^5 X_{21l} = 162.6$	$\sum_{l=1}^5 X_{22l} = 138.9$
.88	$\bar{X}_{12} = 12.12$	$\bar{X}_{21} = 32.52$	$\bar{X}_{22} = 27.78$



# Jak vypadá počítačový výstup?

Analysis of Variance Table – dvoufaktorová ANOVA, oba faktory s pevnými efekty

**Faktor A**  
**Faktor B**  
**Interakce**

Source			Sum of	Mean		Prob	Power
Term	DF		Squares	Square	F-Ratio	Level	(Alpha=0.05)
A	2		16	8	24.00	0.000008*	0.999968
B	1		24	24	72.00	0.000000*	1.000000
AB	2		0	0	0.00	1.000000	0.050000
S	18		6	0.3333333			
Total (Adjust)	23		46				
Total	24						

\* Term significant at alpha = 0.05

Analysis of Variance Table

**Faktor A**  
**Faktor B**  
**Interakce**

Source			Sum of	Mean		Prob	Power
Term	DF		Squares	Square	F-Ratio	Level	(Alpha=0.05)
A	2		34.33333	17.16667	34.33	0.000001*	1.000000
B	1		48.16667	48.16667	96.33	0.000000*	1.000000
AB	2		250.3333	125.1667	250.33	0.000000*	1.000000
S	18		9	0.5			
Total (Adjust)	23		341.8333				
Total	24						

\* Term significant at alpha = 0.05

# Další typy dvoucestné ANOVy (1)

## 2. Nevyvážená ANOVA

- může existovat ve dvou typech: proporční replikace vs. disproporční replikace
- výpočet mnohem složitější než u vyrovnané ANOVy

## 3. Dvoufaktorová ANOVA bez replikací

- zvláštní typ ANOVy, kde v každé buňce je jen po jedné hodnotě
- analýza podobná vyrovnané dvoucestné ANOVě, ale řada omezení (zvláště: nelze testovat interakci apod.)

# Další typy dvoucestné ANOVy (2)

## 4. Znáhodněné bloky (*two-way ANOVA with randomised blocks*)

- blok (*block*) = vymezená plocha (či časový úsek) s relativně homogenními podmínkami prostředí
- Model III. ANOVy, jeden faktor je fixní (takový, jehož vliv chceme studovat) a druhý je náhodný (blok)
- v bloku mám vždy po jednom\* opakování každé hladiny faktoru
- hlavní test (faktoru):  
$$F = MS_{\text{faktor}} / MS_{\text{reminder}}, \text{ přičemž } SS_{\text{reminder}} = SS_{\text{tot}} - SS_{\text{faktor}} - SS_{\text{blok}}$$
- rozdíly mezi bloky jsou využity pro odfiltrování části variability, která nás nezajímá (→ předpoklad: není interakce mezi blokem a faktorem – možný problém\* !!!)
- síla testu faktoru bude vyšší s více bloky a s výraznějším vlivem bloku (tedy s většími rozdíly mezi bloky, mj. v heterogenním prostředí !)

\* Pokud je více opakování hladin na blok, lze studovat i interakci blok x faktor

# Znáhodněné bloky

## *(randomised complete block design)*

**Příklad:** Byly testovány tři způsoby ošetření listu cykasu (voda, spóry houby *Beauveria bassania* ve vodní suspenzi a olej) proti rostlinnému patogenu vlnatce, sající na listech. Byly náhodně vybráno 5 jedinců cykasů, na každém náhodně tři větve, spočítali se počty jedinců vlnatek na 2 ploškách o jednotkovém rozměru. Pak došlo k ošetření větví jedním z testovaných ošetření (vždy náhodně jedna z větví obdržela ošetření vodou atp.). Po jisté době se spočítaly počty jedinců vlnatky a odečetly se od původního počtu před zásahem (viz tabulka):

	Plant				
	1	2	3	4	5
Water	-9	18	10	9	-6
	-6	5	9	0	13
Spores	-4	29	4	-2	11
	7	10	-1	6	-1
Oil	4	29	14	14	7
	11	36	16	18	15

	DF	SS	MS	F-stat	p-value
Blocks	4	686.4	171.60		
Treatments	2	432.03	216.02	12.2	.0037
Error	8	141.8	17.725		

*Oehlert (2010), p. 317*

Analysis of Variance Table (výstup z GLM ANOVA z NCSS 9, na zprůměrovaných datech)

Source	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level
A: blok	4	686.4	171.6		
B: zasah	2	432.0333	216.0167	12.19	0.003729
S	8	141.8	17.725		
Total (Adjusted)	14	1260.233			
Total	15				

Analysis of Variance Table (výstup jednocestné ANOVA z GLM ANOVA z NCSS 9, na zprůměrovaných datech)

Source	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level
A: zasah	2	432.0333	216.0167	3.13	0.080557
S	12	828.2	69.01667		
Total (Adjusted)	14	1260.233			
Total	15				

Analysis of Variance Table (výstup z GLM ANOVA, když se počítá s jednotlivými ploškami)

Source	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level
A: blok	4	1372.8	343.2		
B: zasah	2	864.0667	432.0333	9.53	0.000966
S	23	1042.6	45.33043		
Total (Adjusted)	29	3279.467			
Total	30				

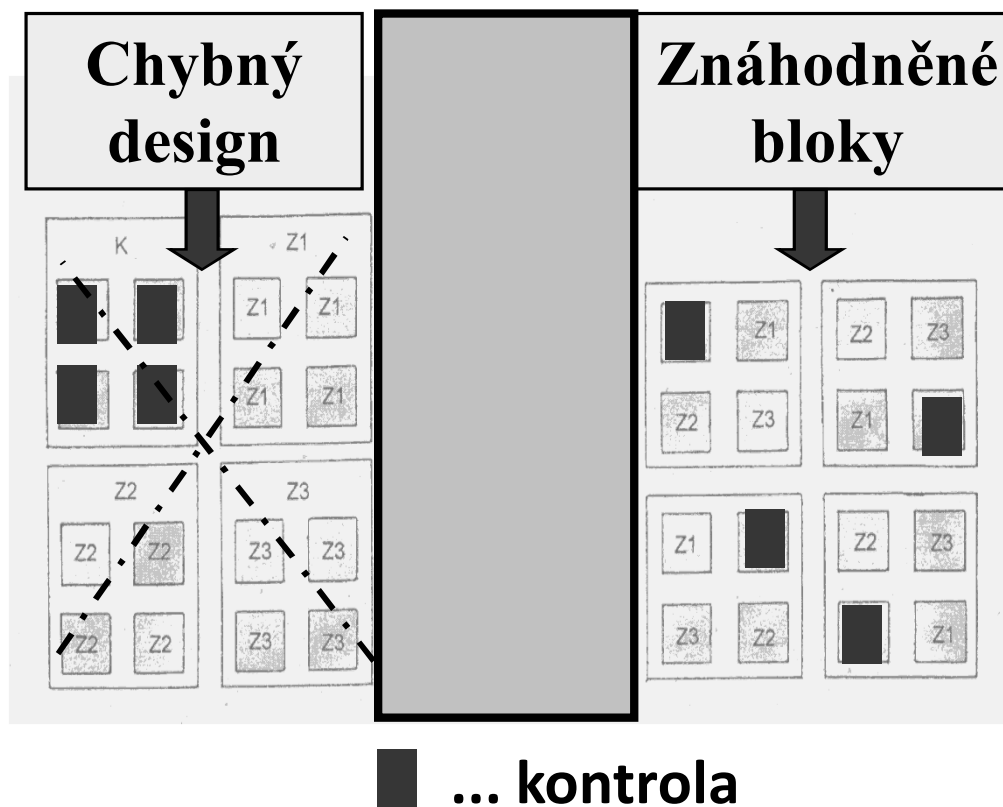
Analysis of Variance Table (výstup jednocestné ANOVA z GLM ANOVA z NCSS 9, když se počítá s jednotlivými ploškami)

Source	DF	Sum of Squares	Mean Square	F-Ratio	Prob Level
A: zasah	2	864.0667	432.0333	4.83	0.016107
S	27	2415.4	89.45926		
Total (Adjusted)	29	3279.467			
Total	30				

# Znáhodněné bloky - špatný a dobrý design

Příklad:

*1. faktor fixní (4 hladiny), 2. náhodný (blok), 4 opakování*



# Další typy dvoucestné ANOVy (3)

## 5. Opakovaná měření (*two-way ANOVA with repeated measures*)

- replikování vzniká tak, že na *témže subjektu provádím postupně zásahy odpovídající hladinám testovaného faktoru (a tedy měříme opakovaně danou proměnnou), standardně v náhodném pořadí\**
- výpočet analogický jako v případě dvoucestné ANOVy bez replikací
- není náš zájem studovat rozdíly mezi subjekty (*náhodný faktor*), rozdíly mezi subjekty jsou využity pro odfiltrování části variability, která nás nezajímá
- řada výhod (bývá ekonomická) a nevýhod (zvl. *carryover effect* - řešení -> *counterbalancing*; problém statistické cirkularity [*circularity, sphericity*])

*\* nebo pouze zaznamenávám proměnnou opakovaně v čase (zde jde pak o analogii split-plot designu)*

# Neparametrická dvoucestná ANOVA

*(Friedmanův test; Friedman test)*

- vhodná pouze pro analýzu designu znáhodněných bloků či opakovaných měření (jeden faktor fixní, druhý náhodný)
- pracuje s pořadím, vhodný na data, kde nelze dodržet podmínky normality a homoskedality

# Třícestná ANOVA

## *(Three-way ANOVA)*

-zkoumáme vliv 3 faktorů:

### -1. třífaktoriální ANOVA

-můžeme testovat (v případě Modelu I.) celkem 3 základní  $H_0$  o faktorech (jako v jednocestné ANOVě), 3  $H_0$  o interakcích mezi 2-ma faktory a 1  $H_0$  o interakci všech 3 faktorů

Možná další uspořádání pokusu:

### 2. Latinský čtverec:

2 zdroje variability pro blokování jednoho faktoru  
(= 1 fixní a 2 náhodné faktory (řádky a sloupce))

### 3. split-plot design:

stejný blok je aplikován na pouze některé (tj. ne všechny) kombinace faktorů (tj. uvnitř každé z  $a$  úrovní faktoru  $A$  je  $n$  různých bloků se všemi  $b$  úrovněmi faktoru  $B$ )



# Hierarchická ANOVA (*Nested ANOVA*)

- uspořádání faktorů není faktoriální, ale hierarchické  
(neexistují zde všechny možné kombinace hladin faktorů)

## Příklad:

Studuji velikost genomu rostliny ze dvou kontrastních stanovišť (faktor A), v každém odeberu náhodně 3 rostliny (faktor B) a z každé rostliny 2 listy (subvzorky).

**Faktor A má 2 hladiny**



**Faktor B má 6 hladin**

(tyto hladiny jsou všechny různé, tj. neexistují tu 2 stejné hladiny faktoru B pro 2 různé hladiny faktoru A; faktor B je NESTED

Počet opakování v nejnižší hierarchické jednotce (zde po 2)

v A)  
- zvyšuje přesnost odhadu replikací díky většímu počtu podvzorků (subsamples)

- lze testovat jak vliv jednotlivých faktorů (nejčastěji nejvyšší faktor je fixní, ostatní hierarchické=nested)

- lze poměřovat podíl variability vysvětlené jednotlivými úrovněmi

# Hierarchická ANOVA *II.*

**Příklad:** Čtyři faktory (A, B, C, D), plně nested = hierarchicky podřízené – podřazené. Např. faktor B je podřazený faktoru A, tedy pro každou z hladin faktoru A existují jiné hladiny faktoru B, atd.

Source	DF	EMS
A	$a - 1$	$\sigma^2 + n\sigma_\delta^2 + nd\sigma_\gamma^2 + ncd\sigma_\beta^2 + nbcd\sigma_\alpha^2$
B(A)	$a(b - 1)$	$\sigma^2 + n\sigma_\delta^2 + nd\sigma_\gamma^2 + ncd\sigma_\beta^2$
C(AB)	$ab(c - 1)$	$\sigma^2 + n\sigma_\delta^2 + nd\sigma_\gamma^2$
D(ABC)	$abc(d - 1)$	$\sigma^2 + n\sigma_\delta^2$
Error	$abcd(n - 1)$	$\sigma^2$

Component	Estimate
$\sigma_\alpha^2$	$(MS_A - MS_B)/(nbcd)$
$\sigma_\beta^2$	$(MS_B - MS_C)/(ncd)$
$\sigma_\gamma^2$	$(MS_C - MS_D)/(nd)$
$\sigma_\delta^2$	$(MS_D - MS_E)/n$
$\sigma^2$	$MS_E$

Oehlert (2010), p. 281-2

Testování vlivu daného faktoru:

$$F\text{-testem} = \frac{MS_{\text{faktor}}}{MS_{\text{nejbližší podřízený faktor}}}$$

# Hierarchická ANOVA *III.*

**Jak zapsat data na analýzu a provést analýzu? Příklad:** Dva druhy – z každého dvě populace – z každé populace 2-3 kolonie a z každé kolonie 2 jedinci, každý změřený 2x

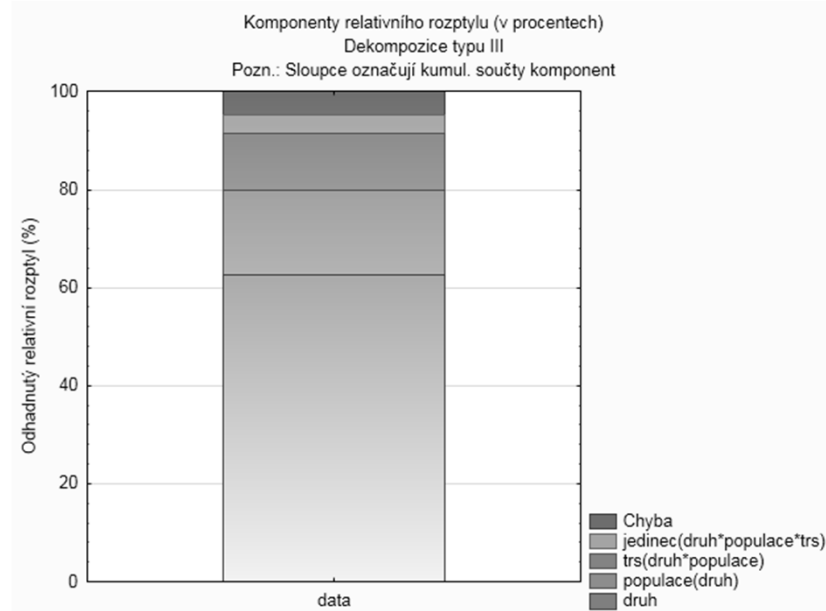
Faktor A	Faktor B	Faktor C	Faktor D	Data
<i>Druh</i>	<i>Populace</i>	<i>Kolonie</i>	<i>Jedinec</i>	
1	1	1	1	1
1	1	1	1	2
1	1	1	2	2
1	1	1	2	3
1	1	2	3	1
1	1	2	3	3
1	1	2	4	2
1	1	2	4	4
1	1	3	5	3
1	1	3	5	5
1	1	3	6	2
1	2	3	6	4
1	2	4	7	6
1	2	4	7	8
1	2	4	8	9
1	2	4	8	8
1	2	5	9	10
1	2	5	9	9
1	2	5	10	10
1	2	5	10	9
2	3	6	11	5
2	3	6	11	6
2	3	6	12	8
2	3	6	12	9
2	3	7	13	12
2	3	7	13	13
2	3	7	14	14
2	3	7	14	15
2	3	8	15	10
2	3	8	15	12
2	4	8	16	13
2	4	8	16	12
2	4	9	17	14
2	4	9	17	15
2	4	9	18	10
2	4	9	18	12
2	4	10	19	12
2	4	10	19	13
2	4	10	20	14
2	4	10	20	10

Hierarchická ANOVA

Faktor	SS	DF	MS	F	P
A	391.066	1	391.066	6.545355	0.124861
B(A)	119.758	2	59.879	4.358402	0.052124
C(AB)	112.664	8	14.083	4.039381	0.024276
D(ABC)	32.417	9	3.602	2.632123	0.036255
Chyba	26.000	19	1.368		

Jednocestná ANOVA s jediným faktorem: A (druh) na stejných datech

A	409.600	1	409.600	44.7393	0.000000
Chyba	347.900	38	9.155		



Transformace dat

# Transformace dat

- v případě, že vliv hladin faktorů není aditivní, popř. data nemají normální rozdělení, je třeba provést matematickou operaci s daty před vlastní analýzou = transformaci, tj. změnu původních dat na jiné pomocí funkce.

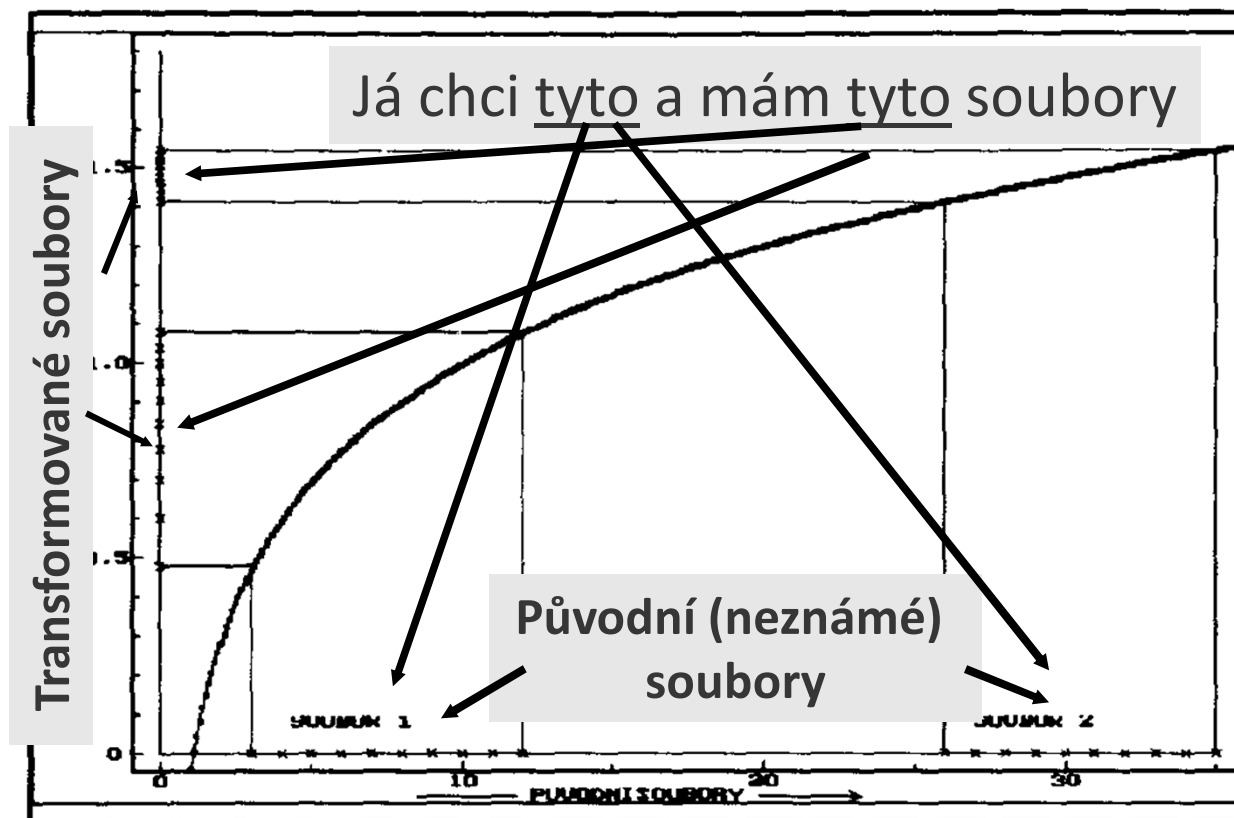
## Obecně:

- výběry vznikly nám neznámou transformací  $f$  z neznámých zákl. symetrických souborů se stejnou mírou variability

- hledáme takovou inverzní funkci  $f^{-1}$ , která provede zpětnou transformaci.

**Cíl:** více symetrické soubory s +- stejnou variabilitou

**Problém:** zvolení vhodné transformace



# Transformace dat (2)

Základní typy transformací:

1. logaritmická transformace (*logarithmic tr.*)

- původní multiplikativní vliv faktoru se převede na aditivní

$$x' = \log(x+1)$$

2. odmocninová transformace (*the square-root tr.*)

- vhodná pro data s Poissonovým rozdělením

$$x' = \sqrt{x}$$
$$x' = \sqrt{x + 0,5}$$
$$x' = \sqrt{x + \frac{3}{8}}$$

3. arcsinová transformace (angulární, *angular tr.*)

- vhodné pro proporce a procenta (před výpočtem nejdříve převést % na rozsah 0-1)

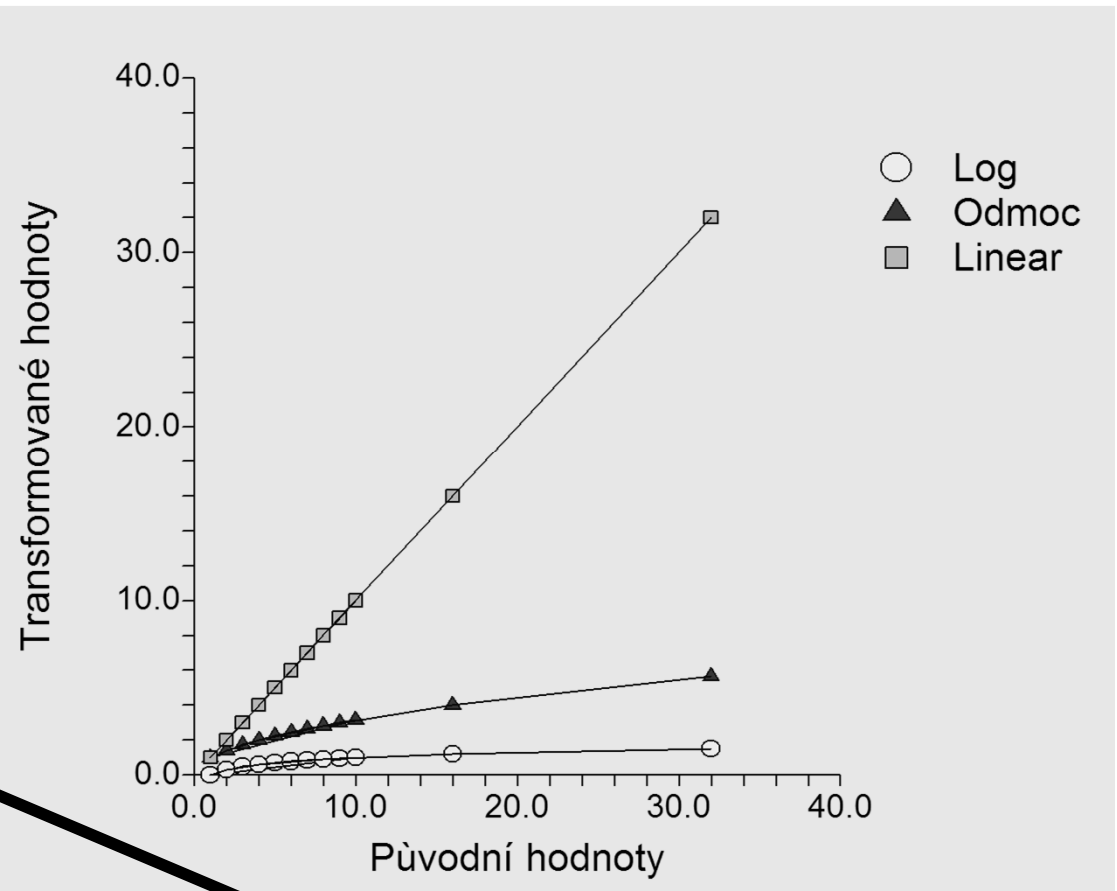
$$p' = \arcsin \sqrt{p}$$

4. Box-Cox transformace (*Box-Cox tr.*)

- rodina transformací zahrnující jako speciální případy výše uvedené transformace, výpočet iterativní metodou

# Transformace dat - ukázka

Originální data	Log tr.	Odmoc. tr.	Lineární tr.
1	0.000	1.000	1.000
2	0.301	1.414	2.000
3	0.477	1.732	3.000
4	0.602	2.000	4.000
5	0.699	2.236	5.000
6	0.778	2.449	6.000
7	0.845	2.646	7.000
8	0.903	2.828	8.000
9	0.954	3.000	9.000
10	1.000	3.162	10.000
16	1.204	4.000	16.000
32	1.505	5.657	32.000



$$x' = \log(x)$$

$$x' = \sqrt{x}$$

$$x' = x$$

