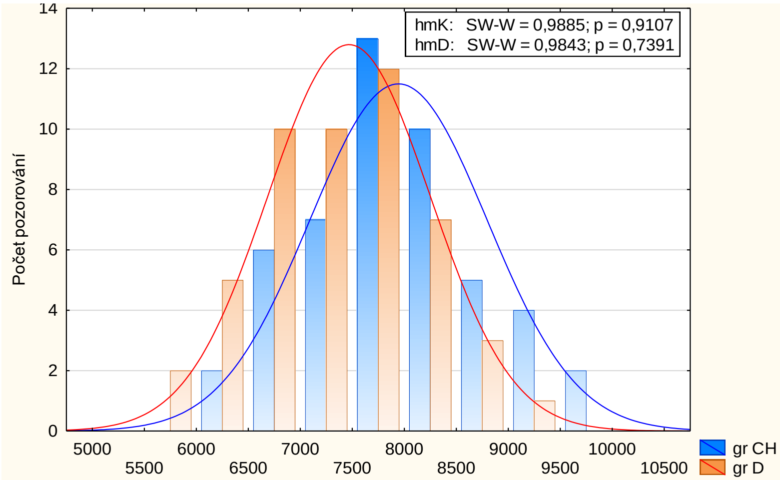
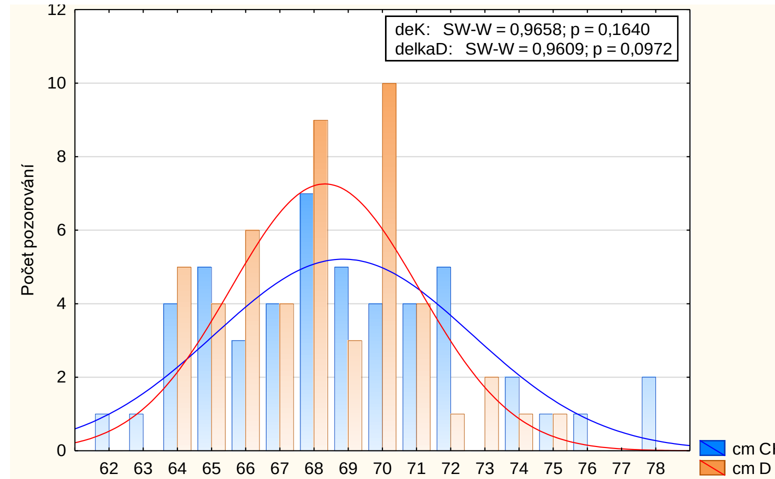
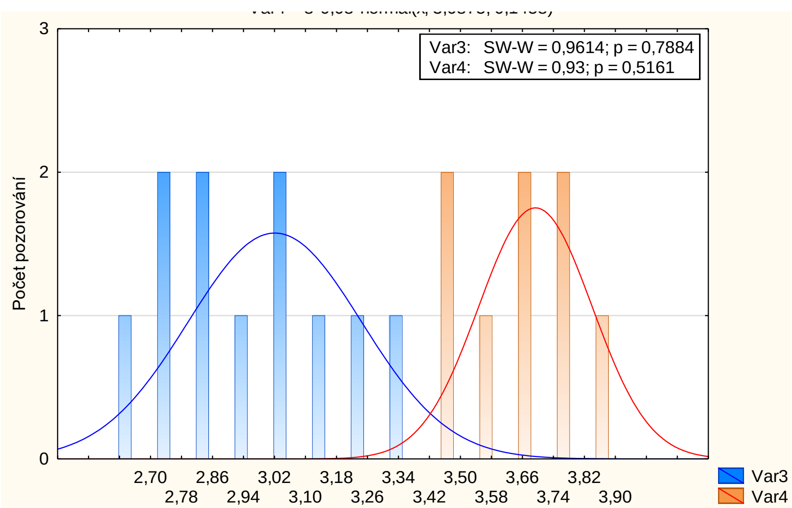
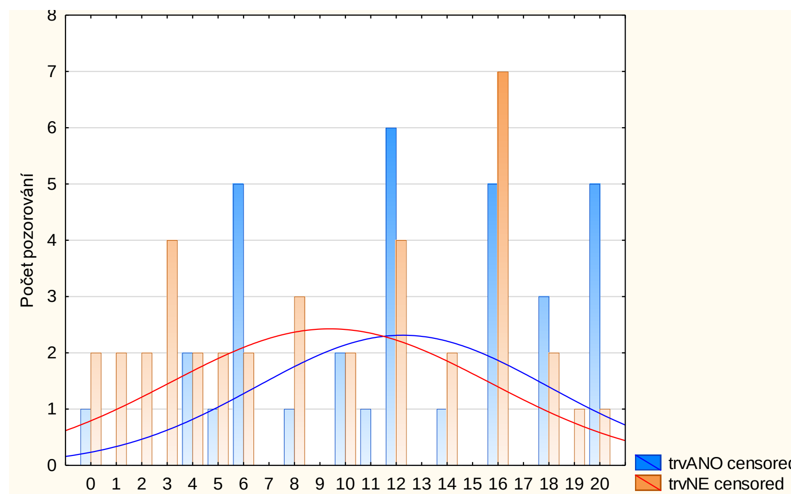
# Dva výběry


hmK: SW-W = 0,9885; p = 0,9107
hmD: SW-W = 0,9843; p = 0,7391
Počet pozorování
gr CH
gr D


deK: SW-W = 0,9658; p = 0,1640
delkaD: SW-W = 0,9609; p = 0,0972
Počet pozorování
cm CH
cm D


Var3: SW-W = 0,9614; p = 0,7884
Var4: SW-W = 0,93; p = 0,5161
Počet pozorování
Var3
Var4


Počet pozorování
trvANO censored
trvNE censored

# Dva výběry: F-test shody variancí [homogeneity of variances]

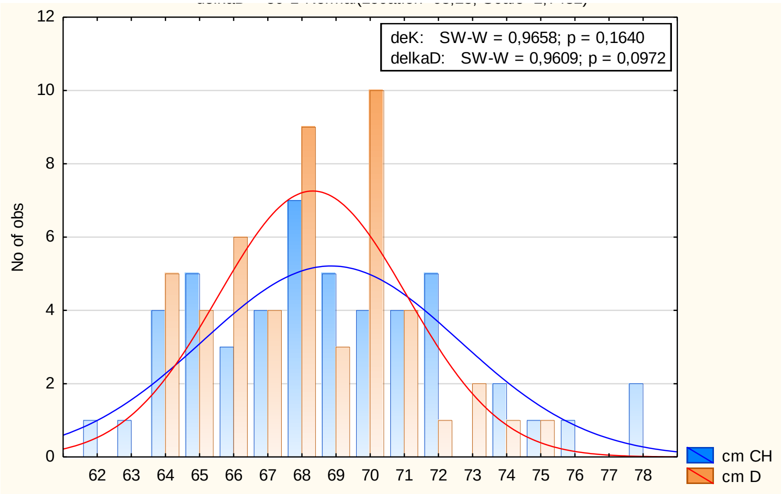
<u>Předpoklady testu</u>:

- $(X_1, X_2, \ldots, X_k)$ a $(Y_1, Y_2, \ldots, Y_m)$ všechno nezávislé; $n_X = k, n_Y = m$
- $\boldsymbol{X_i \sim N(\mu_X, \sigma^2{}_X), Y_i \sim N(\mu_Y, \sigma^2{}_Y)}$, parametry neznáme.

<u>Hypotéza</u>: $\boldsymbol{H_0\colon \sigma^2{}_X = \sigma^2{}_Y}$

ale testujeme $\boldsymbol{\dfrac{\sigma^2{}_X}{\sigma^2{}_Y}} = 1$

alternativa $\boldsymbol{H_1\colon \sigma^2{}_X \neq \sigma^2{}_Y}$




<u>Testová statistika</u>:

$$\boldsymbol{F} = \frac{\dfrac{S_X{}^2}{\sigma_X{}^2}}{\dfrac{S_Y{}^2}{\sigma_Y{}^2}} = \frac{S_X{}^2}{\sigma_X{}^2}\cdot\frac{\sigma_Y{}^2}{S_Y{}^2} = \frac{S_X{}^2}{S_Y{}^2}\cdot\frac{\sigma_Y{}^2}{\sigma_X{}^2} =_{H_0} \boldsymbol{\frac{S_X{}^2}{S_Y{}^2}} \sim_{\boldsymbol{H_0}} \boldsymbol{F_{(n_X-1, n_Y-1)}}$$

Za platnosti hypotézy je zlomek roven 1

# Podrobnější odvození testové statistiky F:

Teoreticky má být $\boldsymbol{F} = \dfrac{\dfrac{\sum_{i=1}^{k} V_i^{\,2}}{k}}{\dfrac{\sum_{j=1}^{m} W_j^{\,2}}{m}}$

$V_i \sim N(0,1)$ a tedy $\sum_{i=1}^{k} V_i^{\,2} \sim \chi^2{}_k$

$W_j \sim N(0,1)$ a tedy $\sum_{j=1}^{m} W_j^{\,2} \sim \chi^2{}_m$

Máme: $\dfrac{S_X{}^2}{\sigma_X{}^2} = \dfrac{1}{\sigma_X{}^2} \cdot \dfrac{\sum_{i=1}^{k}(X_i-\bar{X})^2}{k-1} = \dfrac{1}{k-1} \cdot \dfrac{\sum_{i=1}^{k}(X_i-\bar{X})^2}{\sigma_X{}^2} = \dfrac{1}{k-1} \cdot \sum_{i=1}^{k}\left(\dfrac{X_i-\bar{X}}{\sigma_X}\right)^2$

Jeden stupeň volnosti ztrácím odhadem $\mu_X = \bar{X}$

$\sim N(0,1)$

Podobně pro $S^2{}_Y$ .

Proto má test. statistika F-rozdělení
s (k-1) a (m-1) stupni volnosti:
s ($n_X$-1) a ($n_Y$-1)

$$\boldsymbol{F} = \frac{\dfrac{S_X{}^2}{\sigma_X{}^2}}{\dfrac{S_Y{}^2}{\sigma_Y{}^2}} \sim F_{(k-1,\,m-1)}$$

# Dva výběry: F-test shody variancí

<u>Testová statistika</u>: $\boldsymbol{F = \frac{S_X{}^2}{S_Y{}^2} \sim_{H_0} F_{(n_X-1, n_Y-1)}}$

<u>Kritérium</u>: typicky uvažujeme oboustrannou alternativu $\boldsymbol{H_1: \sigma^2{}_X \neq \sigma^2{}_Y}$, zajímá nás jen otázka shody či neshody variancí.

Proti nulové hypotéze svědčí dvě situace:

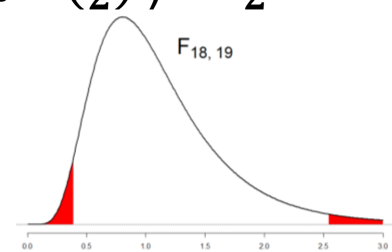
buď $\boldsymbol{S^2{}_X \gg S^2{}_Y}$ a $\boldsymbol{F = \frac{S_X{}^2}{S_Y{}^2}}$ leží na pravém chvostu, srovnám s $\boldsymbol{F_{f1,f2}\left(1-\frac{\alpha}{2}\right)}$

nebo $\boldsymbol{S^2{}_X \ll S^2{}_Y}$ a $\boldsymbol{F = \frac{S_X{}^2}{S_Y{}^2}}$ leží na levém chvostu, srovnám s $\boldsymbol{F_{f1,f2}\left(\frac{\alpha}{2}\right)}$

tedy buď $\boldsymbol{P\left(F \geq F_{f1,f2}\left(1-\frac{\alpha}{2}\right)\right) = \frac{\alpha}{2}}$ nebo $\boldsymbol{P\left(F \leq F_{f1,f2}\left(\frac{\alpha}{2}\right)\right) = \frac{\alpha}{2}}$

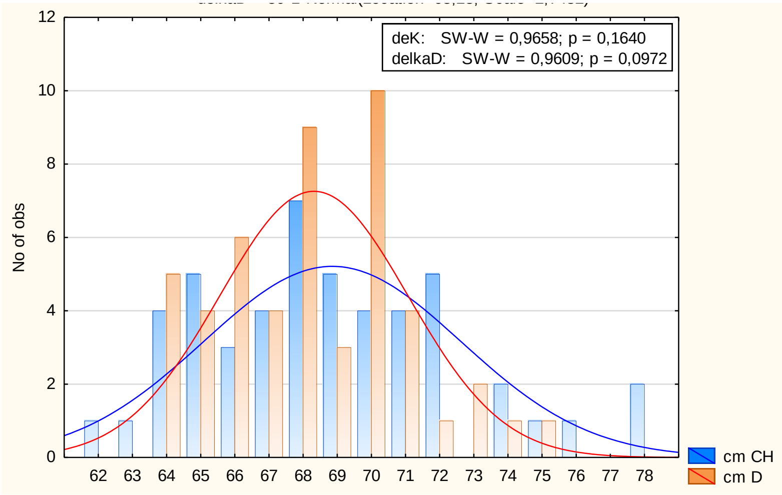
F-rozdělení <u>není souměrné</u>, ale platí, že

$$\boldsymbol{P\left(F \leq F_{f1,f2}\left(\frac{\alpha}{2}\right)\right) = P\left(\frac{1}{F} \geq F_{f2,f1}\left(1-\frac{\alpha}{2}\right)\right) = \frac{\alpha}{2}}$$




<u>Proto kritérium zní</u>: $F = \frac{\textit{větší z odhadů rozptylu}}{\textit{menší z odhadů rozptylu}} \geq \boldsymbol{F}_{\textbf{čitat,jmenov}}\left(1-\frac{\alpha}{2}\right)$

# Dva výběry: F-test shody variancí




Příklad: délky miminek ve 24. týdnu

Rozlišujeme Chlapce a Dívky

$H_0$: $\sigma^2_X = \sigma^2_Y$    $H_1$: $\sigma^2_X \neq \sigma^2_Y$

R:
```
var.test(x, y, ratio=1,
  alternative=c("two.sided","less",
  "greater"), conf.level=0.95, ...)
```

**Numerator degrees of freedom**
Stupně volnosti v čitateli (nahoře)
**Denominator degr. of freedom**
St. volnosti ve jmenovateli (dole).

- F test to compare two variances

data:  delka[Hoch == "ano"] and delka[Hoch == "ne"]

F = 1.8629, **num df** = 48, **denom df** = 49, p-value = 0.03225

alternative hypothesis: true ratio of variances is not equal to 1

95 percent confidence interval:  1.054853     3.295444

sample estimates:  ratio of variances       1.862886

Odhady rozptylů: ${S^2}_{CH} = 14.07,\ {S^2}_{D} = 7.55$

# Příklad: F-test shody variancí

Jaký výsledek dostaneme, když zadáme proměnné v opačném pořadí?

R: **var.test**(delka[Hoch=="ne"], delka[Hoch=="ano"])

- F test to compare two variances

data: delka[Hoch == "**ne**"] and delka[Hoch == "**ano**"]

**F = 0.5368**, num df = **49**, denom df = **48**, p-value = **0.03225**

alternative hyp.: true ratio of variances is not equal to 1

95 percent confidence interval: **0.3034493** **0.9479990**

sample estimates: ratio of variances 0.5368016

$$\frac{14.07}{7.55} = 1.86$$

$$\frac{7.55}{14.07} = \frac{1}{1.86} = 0.538$$

$$F_{49,\,48}(0.025) = 0.566$$
$$F_{48,\,49}(0.975) = 1.766$$

První zadání:

- F test to compare two variances

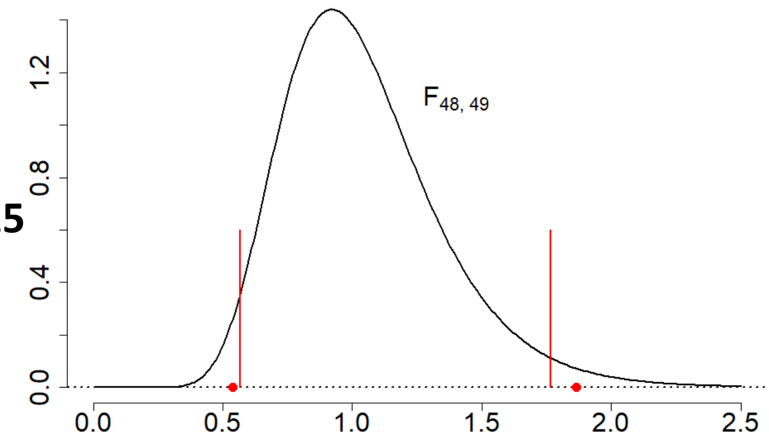
data: delka[Hoch == "**ano**"] and delka[Hoch == "**ne**"]

**F = 1.8629**, num df = **48**, denom df = **49**, p-value = **0.03225**

alternative hyp.: true ratio of variances is not equal to 1

95 percent confidence interval: **1.054853** **3.295444**

sample estimates: ratio of variances 1.862886



# Příklad: F-test shody variancí

STAT: test shody variancí provádí v rámci t-testu rovnosti středních hodnot

*Statistiky → Základní statistiky → t-test, nezávislé, dle skupin* (*dle proměn.*)

t-testy; grupováno: kluk (kojeni podle KLUKa)
Skup. 1: ano
Skup. 2: ne

| Proměnná | Průměr ano | Průměr ne | t | sv | p | Poč.plat ano | Poč.plat. ne | Sm.odch. ano | Sm.odch. ne | F-poměr Rozptyly | p Rozptyly |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **delka** | 68,81633 | 68,28000 | 0,812707 | 97 | 0,418377 | 49 | 50 | 3,750964 | 2,748209 | 1,862886 | 0,032245 |

T-test pro nezávislé vzorky (kojeni podle KLUKa)
Pozn.: Proměnné byly brány jako nezávislé vzorky

| Skup. 1 vs. skup. 2 | Průměr skup. 1 | Průměr skup. 2 | Hodnota t | sv | p | Poč.plat. skup. 1 | Poč.plat. skup. 2 | Sm.odch. skup. 1 | Sm.odch. skup. 2 | F-poměr Rozptyly | p Rozptyly |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **delkaD vs. delkaCH** | 68,28000 | 68,81633 | -0,812707 | 97 | 0,418377 | 50 | 49 | 2,748209 | 3,750964 | 1,862886 | 0,032245 |

# Konfidenční interval pro poměr variancí $\frac{\sigma^2{}_X}{\sigma^2{}_Y}$

Je nesouměrný, protože F-rozdělení je nesouměrné:

$$F_{n_X-1,n_Y-1}\left(\frac{\alpha}{2}\right) \le \frac{S^2{}_X}{S^2{}_Y}\cdot\frac{\sigma^2{}_Y}{\sigma^2{}_X} \le F_{n_X-1,n_Y-1}\left(1-\frac{\alpha}{2}\right)$$

zapracujeme znalost $P\left(F \le F_{f1,f2}\left(\frac{\alpha}{2}\right)\right) = P\left(\frac{1}{F} \ge F_{f2,f1}\left(1-\frac{\alpha}{2}\right)\right)$

$$F_{n_Y-1,n_X-1}\left(1-\frac{\alpha}{2}\right) \ge \frac{S^2{}_Y}{S^2{}_X}\cdot\frac{\sigma^2{}_X}{\sigma^2{}_Y} \ge F_{n_Y-1,n_X-1}\left(\frac{\alpha}{2}\right)$$

Uspořádáme logicky menší < větší a převedeme poměr odhadů rozptylů:

$$F_{n_Y-1,n_X-1}\left(\frac{\alpha}{2}\right)\cdot\frac{S^2{}_X}{S^2{}_Y} \le \frac{\sigma^2{}_X}{\sigma^2{}_Y} \le F_{n_Y-1,n_X-1}\left(1-\frac{\alpha}{2}\right)\cdot\frac{S^2{}_X}{S^2{}_Y}$$

<u>Ad příklad</u>: 0.566 * 1.863 ≤ var(X)/ var(Y) ≤ 1.766 * 1.863

1.055 ≤ var(X)/ var(Y) ≤ 3.295 … nepokrývá 1, $H_0$ zamítáme

# F-test shody variancí – další poznámky

- Pokud nezamítám $H_0$ (rozptyly jsou shodné), počítám odhad společného rozptylu [pooled variance] takto:

$$S^2 = \frac{\sum_{i=1}^{n_X}(X_i - \bar{X})^2 + \sum_{j=1}^{n_Y}\left(Y_j - \bar{Y}\right)^2}{n_X + n_Y - 2} = \frac{(n_X - 1){S_X}^2 + (n_Y - 1){S_Y}^2}{n_X + n_Y - 2}$$

- F-test je dost slabý test, zvláště při malých četnostech výběrů. Uvažte, že pro výběry velikosti 10 (běžné počty v biologii) srovnáváme statistiku F s kvantilem $F_{9,\,9}$ (0.975) = 4.026, tzn. že musí být $S_X^2 > 4 {*} S_Y^2$, čtyřikrát větší, aby test zamítnul $H_0$. Proto je pravděpod. β chyby 2. druhu velká.

- Excel (verze 2010, modul Analýza dat) udává p-hodnotu jednostranného testu. Skutečná p-hodnota testové statistiky je dvojnásobná! Excel počítá $P\left(náh.\, vel. \geq \frac{S^2{}_X}{S^2{}_Y}\right) = p$, ale musím přidat také „druhý chvost“.

# Další testy shody variancí

**Leveneův test** – bude u analýzy rozptylu; pracuje s odchylkami od průměrů (nebo lépe mediánů). Předpokládá normální rozdělení dat.

STAT: volitelný test v záložce *Možnosti* při definování t-testu.

R: balík **`car`** (**C**ompanion to **A**pplied **R**egression), funkce **`leveneTest`**`(x, y, …)`

**Brown & Forsythe test** – modifikace Leveneova testu na mediány, takto je test robustnější vůči odchylce od normálního rozdělení dat.
STAT: tamtéž.

**Ansari-Bradleyův dvouvýběrový test** - pořadový test, neparametrický. Testuje hodnotu ***s***, což je zde poměr směrodatných odchylek (scales).
R: **`ansari.test`**`(x, y, …)`

R obsahuje také **Moodyho test**, který je ovšem komentován takto:
„existují užitečnější testy pro tento problém“ ☺

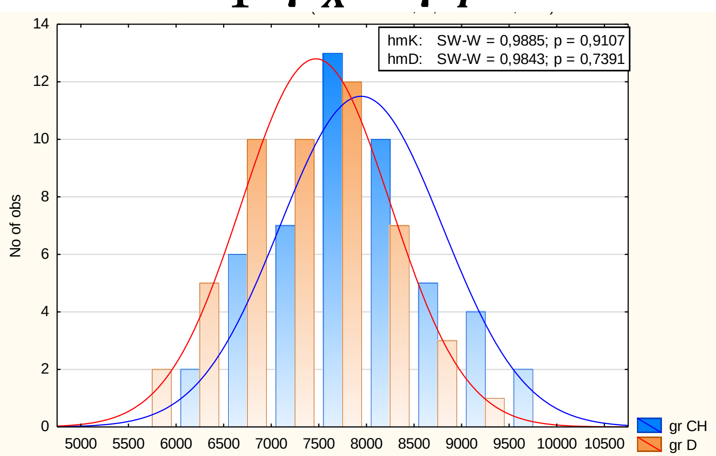
# Dva výběry: t-test shody průměrů

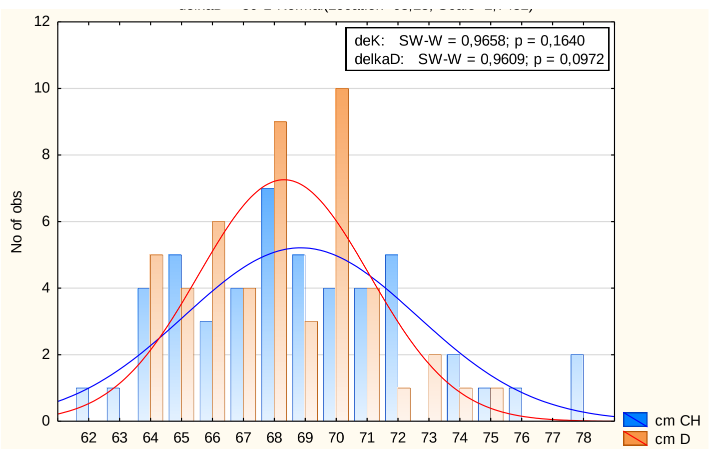
Předpoklady testu:

- $(X_1, X_2, \dots, X_k)$ a $(Y_1, Y_2, \dots, Y_m)$ všechno nezávislé
- $\boldsymbol{X_i \sim N(\mu_X, \sigma_X{}^2), Y_i \sim N(\mu_Y, \sigma_Y{}^2)}$, parametry neznáme.
- výběry mají stejnou varianci, liší se tedy jen posunutím střední hodnoty; $\boldsymbol{\sigma_X{}^2 = \sigma_Y{}^2}$ ozn. $\boldsymbol{\sigma^2}$
- pokud předpoklad stejných rozptylů není splněn, máme Welchův test (dále)

Hypotéza: $\boldsymbol{H_0}: \boldsymbol{\mu_X = \mu_Y}$ testujeme $\mu_X - \mu_Y = 0$

alternativa $\boldsymbol{H_1}: \boldsymbol{\mu_X \neq \mu_Y}$



# Dva výběry: t-test shody průměrů

Hypotéza: $\boldsymbol{H_0}: \boldsymbol{\mu_X = \mu_Y}$ testujeme $\mu_X - \mu_Y = 0$

alternativa $\boldsymbol{H_1}: \boldsymbol{\mu_X \neq \mu_Y}$

Použijeme odhad: $$S^2 = \frac{\sum_{i=1}^{n_X}(X_i-\bar{X})^2 + \sum_{j=1}^{n_Y}\left(Y_j-\bar{Y}\right)^2}{n_X+n_Y-2} = \frac{(n_X-1){S_X}^2+(n_Y-1){S_Y}^2}{n_X+n_Y-2}$$

Odhad pro společný rozptyl [pooled variance]

Testová statistika:

$$\boldsymbol{T = \frac{\bar{X}-\bar{Y}-0}{S.E.(\bar{X}-\bar{Y})} = \frac{\bar{X}-\bar{Y}}{S}\sqrt{\frac{n_X n_Y}{n_X+n_Y}}} \sim_{H_0} \boldsymbol{t_{n_X+n_Y-2}}$$

čti: *má za platnosti hypotézy $H_0$ rozdělení …*

Kritéria podle $H_1$:

$$\mu_X \neq \mu_Y \;\rightarrow\; \boldsymbol{|T| \geq t_{n_X+n_Y-2}(1-{}^{\alpha}\!/_{2})}$$

$$\mu_X > \mu_Y \;\rightarrow\; \boldsymbol{T \geq t_{n_X+n_Y-2}(1-\alpha)}$$

$$\mu_X < \mu_Y \;\rightarrow\; \boldsymbol{T \leq t_{n_X+n_Y-2}(\alpha)} = -t_{n_X+n_Y-2}(1-\alpha)$$

# Pro zájemce odvození rovnosti

$S.E.(\bar{X}-\bar{Y}) = S \cdot \sqrt{\frac{n_X+n_Y}{n_X n_Y}}$, tedy $\boldsymbol{var}\,(\bar{\boldsymbol{X}}-\bar{\boldsymbol{Y}}) = \left(\frac{\boldsymbol{n_X+n_Y}}{\boldsymbol{n_X n_Y}}\right)\boldsymbol{S^2}$

- $\boldsymbol{var}(\bar{\boldsymbol{X}}-\bar{\boldsymbol{Y}}) = E\{(\bar{X}-\bar{Y})-E(\bar{X}-\bar{Y})\}^2 = E\{\bar{X}-E\bar{X}-\bar{Y}+E\bar{Y}\}^2 =$ $E\{(\bar{X}-E\bar{X})-(\bar{Y}-E\bar{Y})\}^2 = E\{(\bar{X}-E\bar{X})^2 - 2\cdot(\bar{X}-E\bar{X})(\bar{Y}-E\bar{Y}) +$

- Všimněte si rozdílu proti párovému t-testu:

Párový t-test **$H_0$: $\mu_X = \mu_Y$** a počítám $X_i = U_i - V_i$ … průměr rozdílů, tj. má smysl počítat rozdíl v páru pozorování

Dvouvýběrový **$H_0$: $\mu_X = \mu_Y$** ($H_0$ stejná) $\bar{X} - \bar{Y}$ … rozdíl průměrů
Zde není žádný vztah mezi $X_i$ a $Y_i$, navíc počet $n_X$ a $n_Y$ se může lišit.

- **t-test je celkem robustní vůči narušení předpokladů**, zvlášť pokud máme dostatek pozorování **a** výběry jsou zhruba stejně početné (uplatní se CLV).

**PŘESTO** je-li podezření na nestejnost variancí (normálního rozdělení) **nebo** se $n_X$ a $n_Y$ značně liší, použijeme lépe

**Welchův přibližný t-test:**

$$T = \frac{\bar{X} - \bar{Y} - 0}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}} \sim_{H_0} t_f \quad , kde \;\; f = \frac{\left(\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}\right)^2}{\frac{\left(\frac{S_X^2}{n_X}\right)^2}{n_X - 1} + \frac{\left(\frac{S_Y^2}{n_Y}\right)^2}{n_Y - 1}}$$

Počet stupňů volnosti $f$ může být i desetinné číslo! Rozdělení testové statistiky známe jen přibližně, proto také p-hodnota je jen přibližná.

# T-test - zadání v softwaru:

STAT: *Statistiky → Základní statistiky → t-test, nezávislé, dle skupin*

nebo → *t-test, nezávislé, dle proměnných*

Welchův test: v záložce *Možnosti* zvolit *test se samostatnými odhady rozptylu* (*Test with separate variance estimates*).

R:
```
t.test(x, y, alternative=c("two.sided","less",
    "greater"), mu=0, paired=FALSE, var.equal=FALSE,
    conf.level=0.95, ...)

t.test(měření ~ skupiny, data, subset, …)
```

Přednastavené nestejné rozptyly, počítá rovnou Welchův přibližný test.

# t-test – příklad první:

Délka miminek ve 24. týdnu

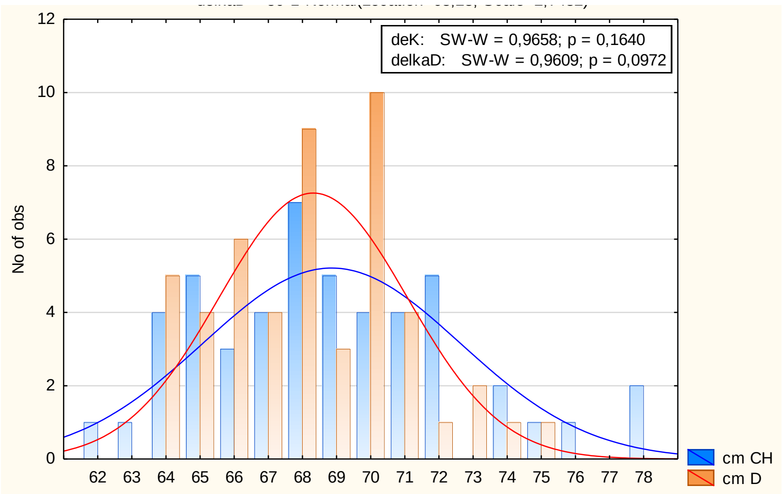
Rozlišujeme Chlapce a Dívky

$H_0$: $\mu_X = \mu_Y$ $H_1$: $\mu_X \neq \mu_Y$

Předpoklady:

Normalita slušná, navíc máme dost pozorování v obou skupinách (49 a 50).

Variance se různí → Welchův přibližný t-test.




STAT:

T-test pro nezávislé vzorky (kojeni podle KLUKa)
Pozn.: Proměnné byly brány jako nezávislé vzorky

Welchovo ***t*** zde

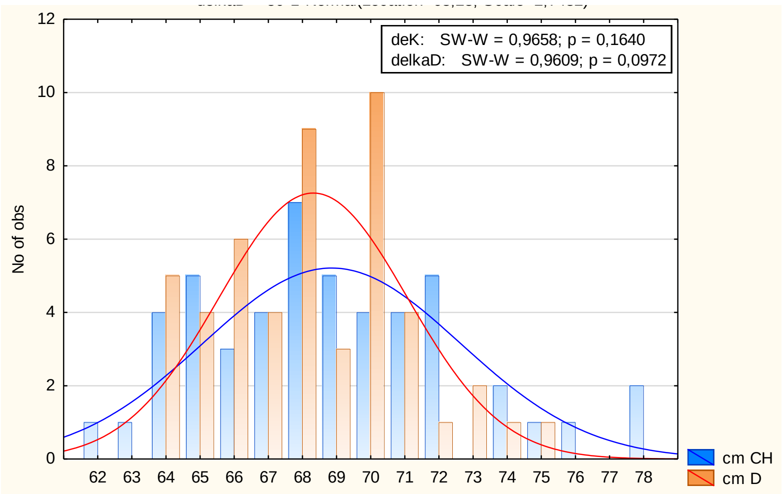
| Skup. 1 vs. skup. 2 | Průměr skup. 1 | Průměr skup. 2 | Hodnota t | sv | p | t samost. odh.rozp. | sv | p oboustr. |
|---|---|---|---|---|---|---|---|---|
| delkaD vs. delkaCH | 68,28000 | 68,81633 | -0,812707 | 97 | 0,418377 | -0,810211 | 87,94454 | 0,420006 |

Zde zadání „dle proměnných“.

pokračování horní tabulky

| Poč.plat. skup. 1 | Poč.plat. skup. 2 | Sm.odch. skup. 1 | Sm.odch. skup. 2 | F-poměr Rozptyly | p Rozptyly |
|---|---|---|---|---|---|
| 50 | 49 | 2,748209 | 3,750964 | 1,862886 | 0,032245 |

# t-test – příklad první:




Délka miminek ve 24. týdnu

Rozlišujeme Chlapce a Dívky

$H_0$: $\mu_X = \mu_Y$ $H_1$: $\mu_X \neq \mu_Y$

Předpoklady:

Normalita slušná, navíc máme dost pozorování v obou skupinách (49 a 50).

Variance se různí → Welchův přibližný t-test.

R: `t.test(delka ~ Hoch, data=kojeni, var.equal=FALSE)`

Welch Two Sample t-test

data: delka by Hoch

t = 0.81021, df = 87.945, p-value = 0.42

Nezamítám hypotézu o shodných průměrech

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval: -0.7791904 1.8518435

sample estimates: mean in group ano mean in group ne

68.81633 68.28000

# t-test – příklad druhý:

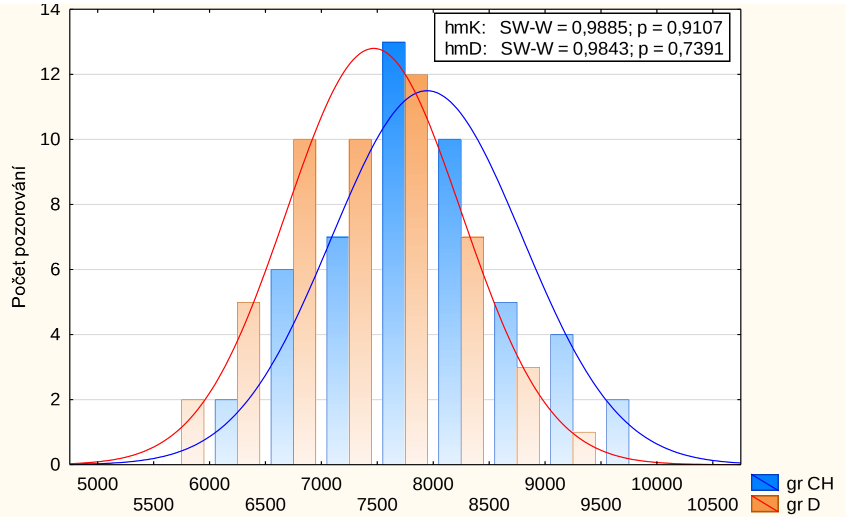
Hmotnost miminek ve 24. týdnu

Rozlišujeme Chlapce a Dívky

$H_0$: $\mu_X = \mu_Y$    $H_1$: $\mu_X \neq \mu_Y$

Předpoklady:

Normalita výborná, navíc máme dost pozorování v obou skupinách (49 a 50).




STAT:

t-testy; grupováno: kluk (kojeni podle KLUKa)
Skup. 1: ano
Skup. 2: ne

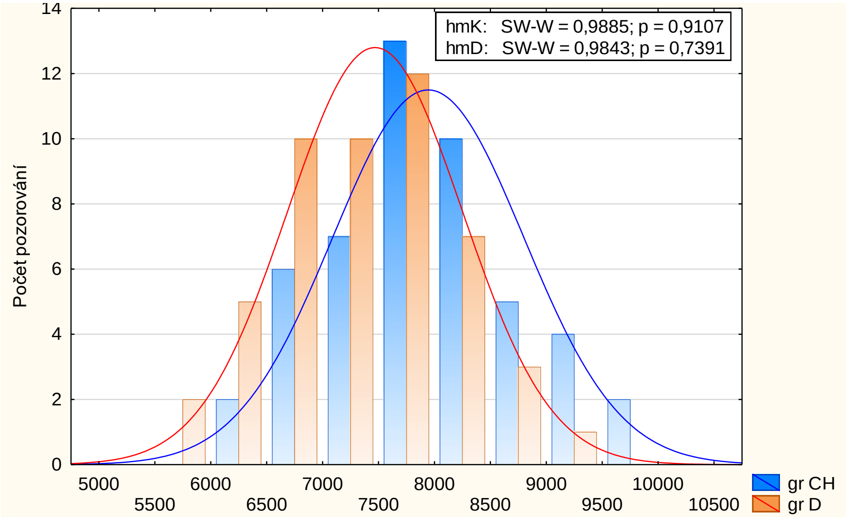
| Proměnná | Průměr ano | Průměr ne | t | sv | p | Poč.plat ano | Poč.plat. ne | Sm.odch. ano | Sm.odch. ne | F-poměr Rozptyly | p Rozptyly |
|---|---|---|---|---|---|---|---|---|---|---|---|
| hmotnost | 7928,571 | 7455,260 | 2,888674 | 97 | 0,004772 | 49 | 50 | 850,1127 | 779,2897 | 1,190023 | 0,546220 |

Zadání „dle skupin“.

Zamítám hypotézu o shodných průměrech

# t-test – příklad druhý:

Hmotnost miminek ve 24. týdnu
Rozlišujeme Chlapce a Dívky
$H_0$: $\mu_X = \mu_Y$ $H_1$: $\mu_X \neq \mu_Y$
Předpoklady:
Normalita výborná, navíc máme dost pozorování v obou skupinách (49 a 50).




R: `var.test(hmotnost~Hoch,data=kojeni)`

F test to compare two variances
data: hmotnost by Hoch
F = 1.19, num df = 48, denom df = 49, p-value = 0.5462 … variance OK

`t.test(hmotnost ~ Hoch, data=kojeni, var.equal=TRUE)`

Two Sample t-test
data: hmotnost by Hoch
t = 2.8887, df = 97, p-value = 0.004772
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval: 148.1130 798.5098
sample estimates:
mean in group ano mean in group ne
7928.571 7455.260

Zamítám hypotézu o shodných průměrech

# Neparametrický test rovnosti středních hodnot

## Mannův-Whitneyův test alias dvouvýběrový Wilcoxonův test

Předpoklady:

$(X_1, X_2, \ldots, X_k)$ a $(Y_1, Y_2, \ldots, Y_m)$ všechno nezávislé ze spojitého rozdělení.

Nulová hypotéza: rozdělení pravděpodobností náh. veličin ***X*** a ***Y*** je stejné (tedy ani posunutí, ani velký rozdíl ve variabilitě)

→ Za platnosti nulové hypotézy jsou stejné i populační mediány

Princip Wilcoxonova pořadového testu:

Sesypeme oba výběry dohromady a hodnoty seřadíme (neřešíme +/- jako u párového testu). Když platí $H_0$, měly by se zhruba pravidelně střídat hodnoty z ***X*** a z ***Y***. Součet pořadí by tedy měl být srovnatelný, zhruba polovina z $\frac{(n_X+n_Y)\cdot(n_X+n_Y+1)}{2}$, tedy $\frac{(n_X+n_Y)\cdot(n_X+n_Y+1)}{4}$.

Protože ale velikosti výběrů $\boldsymbol{n_X}$ a $\boldsymbol{n_Y}$ nemusí být stejné, musím spočtený součet porovnávat s poměrnou částí celého součtu, viz dále:

# Dvouvýběrový Wilcoxonův test

Princip Wilcoxonova pořadového testu:

- $W_X = \sum_{i=1}^{n_X} R(X_i)$ … součet pořadí hodnot z výběru $X$
- $W_Y = \sum_{j=1}^{n_Y} R(Y_j)$ … součet pořadí hodnot z výběru $Y$

~ multinomické rozdělení

- Součet všech pořadí: $W_X + W_Y = \frac{(n_x+n_Y)\cdot(n_x+n_Y+1)}{2}$
- Za platnosti $H_0$ je očekávaný součet

*Poměrná část vůči celkovému počtu hodnot*

$$EW_X = \frac{n_x}{n_x+n_Y}\frac{(n_x+n_Y)\cdot(n_x+n_Y+1)}{2} = \frac{n_x\cdot(n_x+n_Y+1)}{2}$$

$$var\, W_X = \frac{n_x\cdot n_Y\cdot(n_x+n_Y+1)}{12}$$

- Při větším množství shod v pořadí ještě úprava výběrového rozptylu…
- Pro menší rozsahy výběrů $n_X$ a $n_Y$ lze počítat přesné pravděpodobnosti, pro větší $n$ se používá aproximace normálním rozdělením.

# Dvouvýběrový Mannův-Whitneyův test

Test je založen na zdánlivě jiné myšlence: porovnává všechny možné dvojice $(X_i, Y_j)$ a počítá, v kolika případech je hodnota *X* menší než *Y*.

Jsou-li distribuce veličin srovnatelné ($H_0$), bude to zhruba polovina dvojic.

- Označme $\boldsymbol{U_X}$ počet dvojic, kde $\boldsymbol{(X_i < Y_j), U_X \sim_{H_0} Bi(n_X \cdot n_Y, 0.5)}$
- Označme $\boldsymbol{U_Y}$ počet dvojic, kde $\boldsymbol{(X_i > Y_j)}$
- Případy $\boldsymbol{(X_i = Y_j)}$ započítáme polovinou k $\boldsymbol{U_X}$ a polovinou k $\boldsymbol{U_Y}$
- Kritický obor je zpravidla popisován pomocí $U = \min(U_X, U_Y)$.

- Tyto rovnice ukazují souvislost mezi testovými statistikami:

$$\boldsymbol{W_X = n_x n_Y + \frac{n_x \cdot (n_x + 1)}{2} - U_X}$$

$$\boldsymbol{W_Y = n_x n_Y + \frac{n_Y \cdot (n_Y + 1)}{2} - U_Y}$$

## Poznámky k Mann-Whitneyovu *U* testu a Wilcoxonovu testu

- Oba testy prověřují nulovou hypotézu o shodě rozdělení, ze kterého pocházejí porovnávané výběry. Pokud testujeme nulovou hypotézu o shodě polohy (mediánu), musíme předpokládat, že se distribuce příliš neliší tvarem.
- Statistika ***U*** zahrnuje úpravu pro shody v pořadí (tied values).
- Uvádí se, že pro výběry početnější než 20 pozorování se rozdělení ***U*** statistiky rychle blíží normálnímu rozdělení, proto je ve výsledcích testů také ***Z*** statistika a ***Z*** statistika s opravou na spojitost (Yatesova korekce).
- Zároveň je pro menší ***n*** spočtena přesná (exact) pravděpodobnost, že náhodná veličina bude větší než spočtená statistika ***U***, tedy prst. na pravém chvostu. Tato p-hodnota je vynásobena *2, což odpovídá pravděpodobnosti oboustranné alternativní hypotézy. Nicméně ve výpočtu není zahrnuta oprava na shody v pořadí, proto R takovou zkreslenou „exact“ pravděpodobnost vůbec nenabízí.

# Neparametrické dvouvýběrové testy v softwaru:

<u>STAT</u>: *Statistiky → Neparametrické statistiky → Porovnání dvou nezávislých vzorků (skupiny)*

tzn. data musí být uspořádány jako hodnoty (Var1) a kódování (Var2)

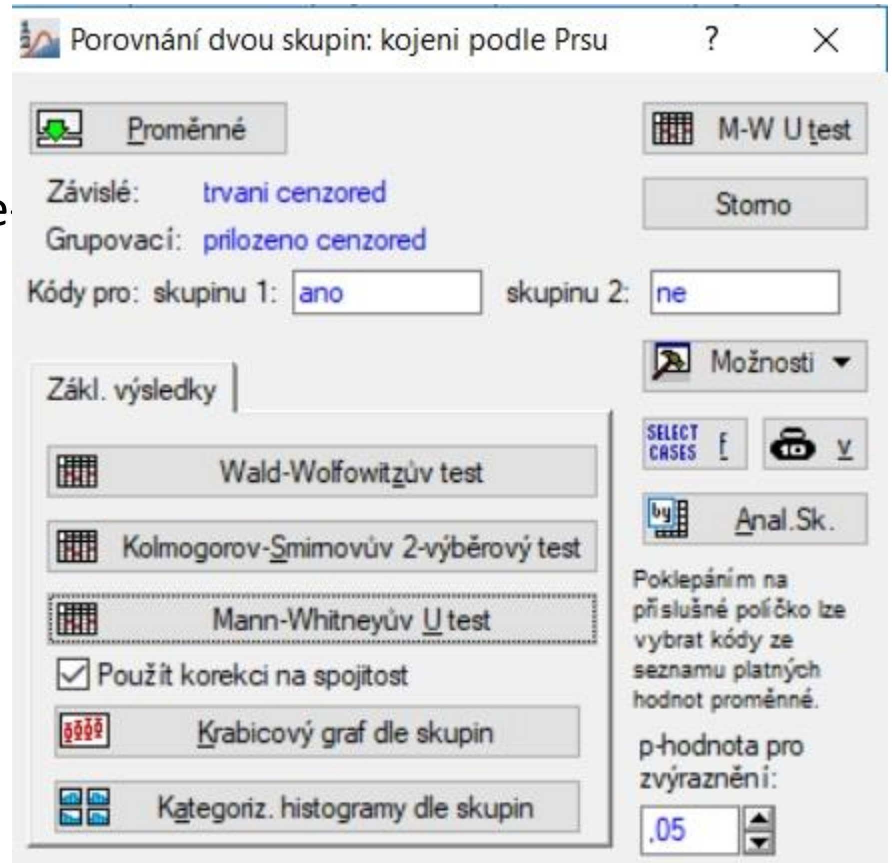
<u>Nabídka</u>:

**Wald-Wolfowitz runs test** – testuje rozdíly ve tvaru distribuční funkce, nedá se interpretovat přímo na posunutí středních hodnot.

**Kolmogorov-Smirnov dvouvýběrový test** – totéž

**Mann-Whitney U test** – může být interpretován stejně jako t-test, tedy jako test shody (nebo posunutí) středních hodnot.

**Yatesova korekce na spojitost** – pro velké rozsahy výběrů vypnout.



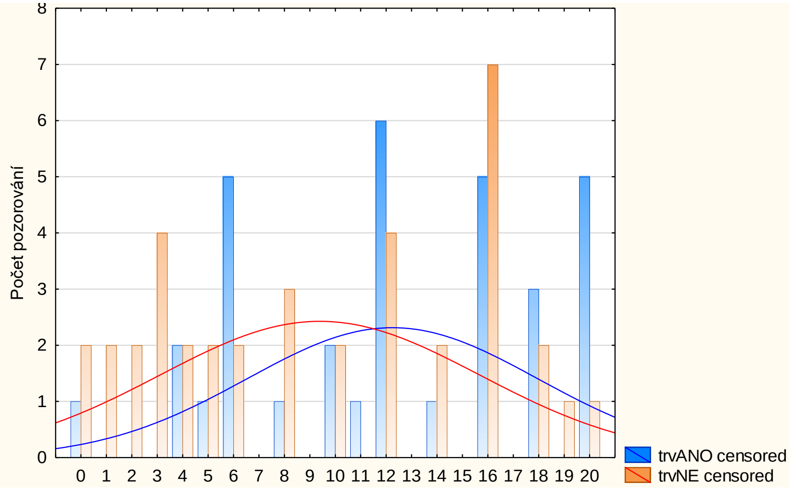
# Mann-Whitney alias Wilcoxon - příklad:

Příklad: data o délce kojení (v týdnech) a záznam, zda bylo dítě do 30 minut po porodu přiloženo k prsu. Z výběru byly „uříznuty“ údaje „24 týdnů“, protože toto číslo zahrnuje všechny matky, které kojily ve 24. týdnu, ale neříkají nic o tom, jak dlouho ještě kojily po tomto datu.

Mann-Whitneyův U Test (w/ oprava na spojitost) (kojeni podle Prsu)
Dle proměn. prilozeno cenzored
Označené testy jsou významné na hladině p <,05000

| Proměnná | Sčt poř. ano | Sčt poř. ne | U | Z | p-hodn. | Z upravené | p-hodn. | N platn. ano | N platn. ne | 2*1str. přesné p |
|---|---|---|---|---|---|---|---|---|---|---|
| trvani cenzored | 1359,500 | 1196,500 | 455,5000 | 1,971386 | 0,048681 | 1,981647 | 0,047520 | 33 | 38 | 0,047578 |

R: `wilcox.test(x, y, alternative=c("two.sided","less","greater"), mu=0, paired=FALSE, exact=NULL, correct=TRUE, conf.int=FALSE, conf.level=0.95, ...)`




Pozn. STAT a R se v číslech zcela neshodnou…

# Mediánový test

Existuje, je však velmi slabý.

# Porovnání dvou pravděpodobností

tj. mám data o dvou znacích nominální proměnné a ptám se,
zda se dva výběry shodnou v pravděpodobnosti, že nastane znak A.
Toto odpovídá čtyřpolní kontingenční tabulce.
(je-li sledovaných znaků více, testuji obecnou kontingenční tabulku)

Předpoklad: mám dvě série vzájemně nezávislých pokusů, ve kterých zjišťuji, zda nastal znak (jev) A. Prst. znaku A v jedné sérii je stejný.

- $Y_1$ = počet pokusů, kdy nastal znak A v první sérii (celkem $n_1$)
- $Y_2$ = počet pokusů, kdy nastal znak A ve druhé sérii (celkem $n_2$)
- $Y_1 \sim Bi(n_1, p_1) \qquad Y_2 \sim Bi(n_2, p_2)$

Nulová hypotéza: obě pravděpodobnosti jsou shodné, $p_1 = p_2 = p$

Odhady: $\hat{p}_1 = \frac{Y_1}{n_1} \qquad \hat{p}_2 = \frac{Y_2}{n_2}$

$$var(\hat{p}_1 - \hat{p}_2) = \frac{p_1 \cdot (1 - p_1)}{n_1} + \frac{p_2 \cdot (1 - p_2)}{n_2} =_{H_0} p \cdot (1 - p) \cdot \left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

# Porovnání dvou pravděpodobností

$\boldsymbol{Y_1 \sim Bi(n_1, p_1)}$ $\boldsymbol{Y_2 \sim Bi(n_2, p_2)}$

Nulová hypotéza: obě pravděpodobnosti jsou shodné, $\boldsymbol{p_1 = p_2 = p}$

Testovat můžeme trojím způsobem:

1) Přesný binomický test → R: binom.test
2) Přibližný test přes aproximaci normálním rozdělením

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{S.E.(\hat{p}_1 - \hat{p}_2)} = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0,1)$$

STAT: *Statistiky →Základní statistiky → Testy rozdílů*

3) Chí-kvadrát test

R: prop.test

chisq.test

STAT: *Statistiky →Neparametrické statistiky → 2x2 tabulky*