

# Osnova přednášky Přehled t-testů

1. Jednovýběrový t-test

2. Párový t-test

3. Dvouvýběrový t-test

- ověření hypotézy o shodě rozptylů (F-test)

- Cohenův koeficient věcného účinku

4. Provedení t-testů v systému STATISTICA pomocí aplikace Testy rozdílů

## Přehled t-testů

### 1. Jednovýběrový t-test

**Označení:** Necht'  $X_1, \dots, X_n$  je náhodný výběr,  $n \geq 2$ .

$$M = \frac{1}{n} \sum_{i=1}^n X_i \quad \dots \text{výběrový průměr,}$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - M)^2 \quad \dots \text{výběrový rozptyl,}$$

$$S = \sqrt{S^2} \quad \dots \text{výběrová směrodatná odchylka}$$

**Definice jednovýběrového t-testu:** Necht'  $X_1, \dots, X_n$  je náhodný výběr z  $N(\mu, \sigma^2)$ , kde rozptyl  $\sigma^2$  neznáme. Necht'  $n \geq 2$  a  $c$  je reálná konstanta. Test  $H_0: \mu = c$  proti  $H_1: \mu \neq c$  (resp.  $H_0: \mu = c$  proti  $H_1: \mu < c$  resp.  $H_0: \mu = c$  proti  $H_1: \mu > c$ ) se nazývá **jednovýběrový t-test**.

**Poznámka:** Jednovýběrový t-test je založen na pivotové statistice  $T = \frac{M - \mu}{\frac{S}{\sqrt{n}}} \sim t(n-1)$ .

Dosadíme-li za  $\mu$  konstantu  $c$ , dostaneme testovou statistiku:  $T_0 = \frac{M - c}{\frac{S}{\sqrt{n}}}$ , která se v případě platnosti  $H_0$  řídí Studentovým rozložením  $t(n-1)$ .

### Provedení jednovýběrového t-testu pomocí kritického oboru

Vypočteme realizaci testové statistiky  $t_0 = \frac{m - c}{\frac{s}{\sqrt{n}}}$ . Stanovíme kritický obor  $W$ . Pokud  $t_0 \in W$ ,

$H_0$  zamítáme na hladině významnosti  $\alpha$  a přijímáme  $H_1$ .

**Oboustranný test:** Testujeme  $H_0: \mu = c$  proti  $H_1: \mu \neq c$ . Kritický obor má tvar:

$$W = (-\infty, -t_{1-\alpha/2}(n-1)) \cup (t_{1-\alpha/2}(n-1), \infty).$$

**Levostranný test:** Testujeme  $H_0: \mu = c$  proti  $H_1: \mu < c$ . Kritický obor má tvar:  $W = (-\infty, -t_{1-\alpha}(n-1))$ .

**Pravostranný test:** Testujeme  $H_0: \mu = c$  proti  $H_1: \mu > c$ . Kritický obor má tvar:  $W = (t_{1-\alpha}(n-1), \infty)$ .

## Provedení jednovýběrového t-testu pomocí intervalu spolehlivosti

Pro **oboustranný test** sestrojíme oboustranný  $100(1-\alpha)\%$  interval spolehlivosti pro střední hodnotu  $\mu$  při neznámém rozptylu  $\sigma^2$ :

$$(d, h) = \left( m - \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1), m + \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1) \right)$$

Pro **levostranný test** sestrojíme pravostranný  $100(1-\alpha)\%$  interval spolehlivosti pro střední hodnotu  $\mu$  při neznámém rozptylu  $\sigma^2$ :

$$(-\infty, h) = \left( -\infty, m + \frac{s}{\sqrt{n}} t_{1-\alpha}(n-1) \right)$$

Pro **pravostranný test** sestrojíme levostranný  $100(1-\alpha)\%$  interval spolehlivosti pro střední hodnotu  $\mu$  při neznámém rozptylu  $\sigma^2$ :

$$(d, \infty) = \left( m - \frac{s}{\sqrt{n}} t_{1-\alpha}(n-1), \infty \right)$$

Pokud číslo  $c$  padne do intervalu spolehlivosti, nulovou hypotézu nezamítáme na hladině významnosti  $\alpha$ . V opačném případě nulovou hypotézu zamítáme na hladině významnosti  $\alpha$  a přijímáme alternativní hypotézu.

## Provedení jednovýběrového t-testu pomocí p-hodnoty

Vypočteme p-hodnotu (nejlépe pomocí statistického software) a porovnáme ji s hladinou významnosti  $\alpha$ . Jestliže  $p \leq \alpha$ , pak nulovou hypotézu zamítáme na hladině významnosti  $\alpha$  a přijímáme alternativní hypotézu. Je-li  $p > \alpha$ , pak nulovou hypotézu nezamítáme na hladině významnosti  $\alpha$ .

Způsob výpočtu p-hodnoty:

Označme  $\Phi$  distribuční funkci Studentova rozložení  $t(n-1)$ .

Pro oboustrannou alternativu  $p = 2 \min\{\Phi(t_0), 1 - \Phi(t_0)\}$ .

Pro levostrannou alternativu  $p = \Phi(t_0)$ .

Pro pravostrannou alternativu  $p = 1 - \Phi(t_0)$ .

### **Příklad na jednovýběrový t-test:**

Bylo prováděno sledování obsahu vitamínu C ve vzorcích mrkve, která byla zakoupena na biofarmě. Celkem bylo provedeno analytické stanovení obsahu vitamínu C ve 20 vzorcích mrkve a byly zjištěny následující koncentrace (v mg/kg):

41,1; 32,6; 28,9; 19,6; 23,6; 35,0; 36,7; 45,9; 49,6; 33,6; 17,8; 24,6; 29,6; 47,7; 41,6; 39,8; 15,6; 34,1; 44,0 a 55,8

Průměrný obsah vitamínu C v mrkvi, který je uváděn v literatuře, je 35 mg/kg.

Liší se obsah vitamínu C stanoveného ve vzorcích mrkve z biofarmy od průměrné hodnoty uváděné v literatuře?

### **Řešení pomocí systému STATISTICA:**

Otevřeme datový soubor mrkev.sta s jednou proměnnou X a 20 případy. V proměnné X jsou zapsány zjištěné hodnoty obsahu vitamínu C.

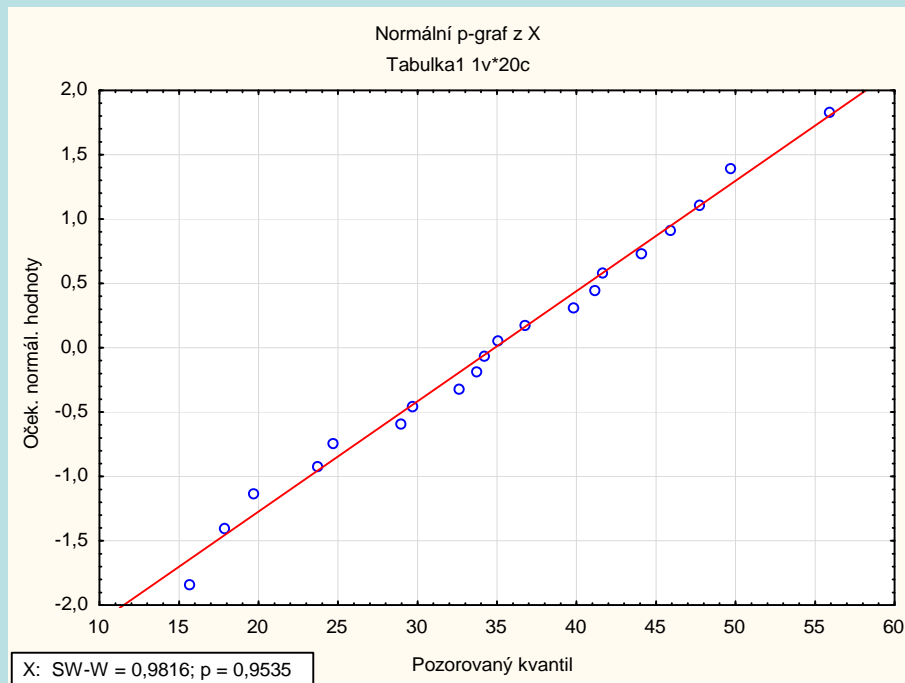
Na hladině významnosti 0,05 testujeme hypotézu  $H_0: \mu = 35$  proti alternativě  $H_1: \mu \neq 35$ .

Jde o úlohu na jednovýběrový t-test.

Nejprve pomocí N-P grafu a Shapirova – Wilkova testu ověříme, zda data pocházejí z normálního rozložení.

**Vytvoření N-P plotu:** Grafy – 2D Grafy – Normální pravděpodobnostní grafy – Proměnná X – OK - odškrtneme Neurčovat průměrnou pozici svázaných pozorování – zaškrtneme Shapiro – Wilkův test - OK.

## Ověření normality pomocí N-P grafu a S-W testu:



Body v N-P grafu jsou blízko ideální přímky.

S-W test poskytl p-hodnotu 0,9535, tedy na hladině významnosti 0,05 nezamítáme hypotézu o normalitě.

## Provedení jednovýběrového t-testu:

Statistiky – Základní statistiky/tabulky – t-test, samost. vzorek – OK – Proměnné X – OK. Do Referenční hodnoty napíšeme 35, na záložce Možnosti zaškrtneme Výpočet mezí spolehl. – Výpočet. Dostaneme tabulku:

Proměnná	Test průměrů vůči referenční konstantě (hodnotě) (mrkev.sta)									
	Průměr	Sm.odch.	N	Sm.chyba	Int. spolehl. -95,000%	Int. spolehl. +95,000%	Referenční konstanta	t	SV	p
X: obsah vitamínu C v mrkvi	34,86000	11,08719	20	2,479170	29,67104	40,04896	35,00000	-0,056471	19	0,955557

Vidíme, že průměrný obsah vitamínu C ve sledovaných 20 vzorcích je 34,86 mg/kg se směrodatnou odchylkou 11,09 mg/kg.

**Test pomocí intervalu spolehlivosti:** S pravděpodobností 95 % se neznámá střední hodnota obsahu vitamínu C nachází v intervalu 29,67 mg/kg až 40,05 mg/kg. Protože referenční konstanta 35 mg/kg se nachází v tomto 95% intervalu spolehlivosti, hypotézu  $H_0: \mu = 35$  nezamítáme na hladině významnosti 0,05.

**Test pomocí kritického oboru:** Testová statistika nabývá hodnoty -0,0565. Kritický obor je

$$W = (-\infty, -t_{1-\alpha/2}(n-1)) \cup (t_{1-\alpha/2}(n-1), \infty) = (-\infty, -t_{0,975}(19)) \cup (t_{0,975}(19), \infty) = (-\infty, -2,093) \cup (2,093, \infty)$$

Testová statistika se nerealizuje v kritickém oboru, tedy hypotézu  $H_0: \mu = 35$  nezamítáme na hladině významnosti 0,05.

**Upozornění:** Kvantil  $t_{0,975}(19)$  najdeme v systému STATISTICA pomocí Pravděpodobnostního kalkulátoru: Statistiky – Pravděpodobnostní kalkulátor - vybereme Rozdělení t (Studentovo). Do okénka sv. napíšeme 19 a do okénka p napíšeme 0,975. V okénku t se objeví 2,093.

**Test pomocí p-hodnoty:** Protože p-hodnota je 0,9556, což je větší než hladina významnosti 0,05, hypotézu  $H_0: \mu = 35$  nezamítáme na hladině významnosti 0,05.



## 2. Párový t-test

**Označení:** Necht'  $\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \dots, \begin{pmatrix} X_n \\ Y_n \end{pmatrix}$  je náhodný výběr z dvourozměrného rozložení, přičemž  $n \geq 2$ .

Označíme  $\mu = \mu_1 - \mu_2$  a zavedeme **rozdílový náhodný výběr**  $Z_1 = X_1 - Y_1, \dots, Z_n = X_n - Y_n$ , o němž předpokládáme, že se řídí normálním rozložením. Označíme

$$M = \frac{1}{n} \sum_{i=1}^n Z_i \quad \dots \text{výběrový průměr rozdílového náhodného výběru,}$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (Z_i - M)^2 \quad \dots \text{výběrový rozptyl rozdílového náhodného výběru.}$$

### Definice párového t-testu:

Necht'  $\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \dots, \begin{pmatrix} X_n \\ Y_n \end{pmatrix}$  je náhodný výběr z dvourozměrného rozložení, přičemž  $n \geq 2$ . Necht'  $c$  je reálná konstanta. Testujeme  $H_0: \mu_1 - \mu_2 = c$  (tj.  $\mu = c$ ) proti  $H_1: \mu_1 - \mu_2 \neq c$  (tj.  $\mu \neq c$ ) nebo testujeme nulovou hypotézu proti jedné z jednostranných alternativ. Tento test se nazývá **párový t-test**.

**Provedení párového t-testu:** Párový t-test se provádí stejně jako jednovýběrový t-test aplikovaný na rozdílový náhodný výběr  $Z_1, \dots, Z_n$ .

### **Příklad na párový t-test:**

U 22 vzorků dřeva byla změřena jejich vlhkost v procentech dvěma metodami. Výsledky máme v tabulce:

6	5,3	4,5	6,6	7,5	7,4	7,4	5	7,5	7,2	7,7	7,8	7,5	7,6	7,7	6,7	7,9	7,9	6,9	6,7	7	7
8,8	8,3	6,7	9,1	8,7	8	7,9	7,6	6,2	5,9	8,5	9	8,1	7,1	8	7,5	7,4	7,3	7	6,8	6,4	6,4

Na hladině významnosti 0,05 testujte hypotézu, že střední hodnoty vlhkosti dřeva získané oběma metodami se neliší.

### **Řešení pomocí systému STATISTICA:**

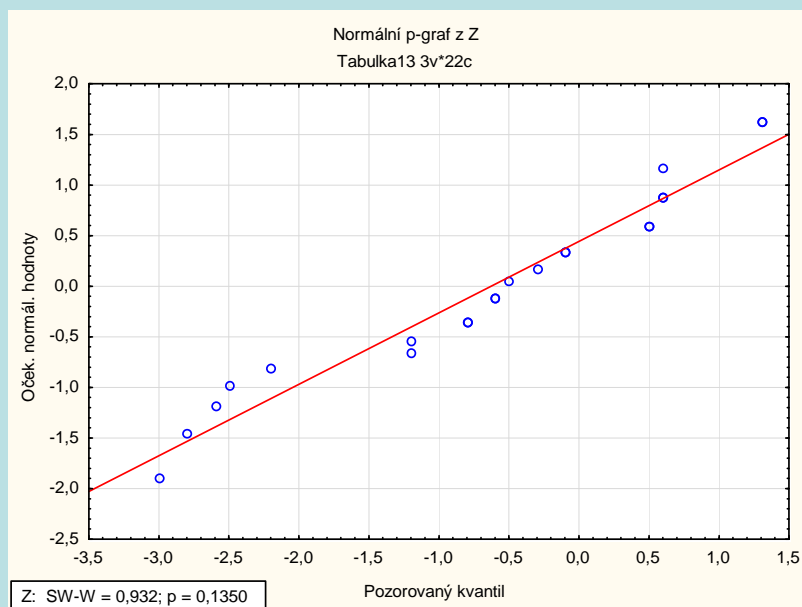
Otevřeme datový soubor vlhkost\_dreva.sta s dvěma proměnnými X, Y a 22 případy. Datový soubor doplníme o proměnnou Z, do níž uložíme rozdíl  $X - Y$ .

Na hladině významnosti 0,05 testujeme hypotézu  $H_0: \mu_1 - \mu_2 = 0$  proti alternativě  $H_1: \mu_1 - \mu_2 \neq 0$ .

Jde o úlohu na párový t-test.

Nejprve pomocí N-P grafu a Shapirova – Wilkova testu ověříme, zda data v proměnné Z pocházejí z normálního rozložení.

**Vytvoření N-P plotu:** Grafy – 2D Grafy – Normální pravděpodobnostní grafy – Proměnná Z – OK - odškrtneme Neurčovat průměrnou pozici svázaných pozorování – zaškrtneme Shapiro – Wilkův test - OK.



Vzhled N-P grafu svědčí o mírném porušení normality, které však není významné na hladině významnosti 0,05, protože p-hodnota S – W testu je 0,135, což je větší než 0,05.

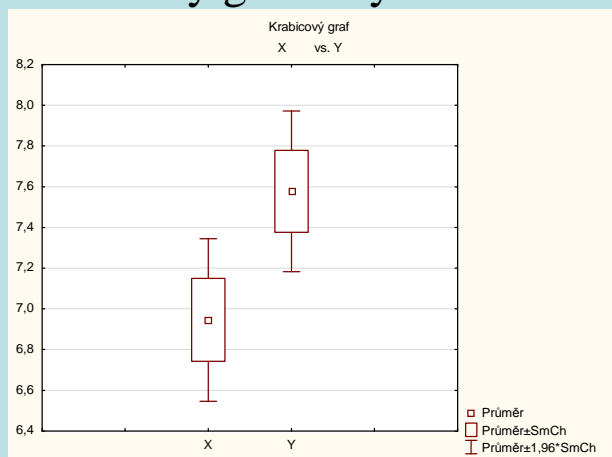
## Provedení párového t-testu:

Statistiky – Základní statistiky/tabulky – t-test, závislé vzorky – OK – Proměnné – 1. seznam proměnných X, 2. seznam proměnných Y – OK - Výpočet. Dostaneme tabulku:

t-test pro závislé vzorky (vlhkost_dreva.sta)										
Označ. rozdíly jsou významné na hlad. $p < ,05000$										
Proměnná	Průměr	Sm.odch.	N	Rozdíl	Sm.odch. rozdílu	t	sv	p	Int. spolehl. -95,000%	Int. spolehl. +95,000%
X	6,945455	0,954552								
Y	7,577273	0,944625	22	-0,631818	1,311265	-2,26002	21	0,034569	-1,21320	-0,050436

Testová statistika párového t-testu se realizuje číslem -2,26, odpovídající p-hodnota je 0,0346, tedy na hladině významnosti 0,05 zamítáme nulovou hypotézu. S rizikem omylu nejvýše 5 % jsme prokázali, že střední hodnoty vlhkosti dřeva získané oběma metodami se liší.

Tabulku ještě doplníme krabicovými diagramy: Na záložce Detailní výsledky zvolíme Krabicový graf a vybereme možnost Prům./SmCh/1,96\*SmCh.



### 3. Dvouvýběrový t-test

**Označení:** Máme dva nezávislé náhodné výběry, první je  $X_{11}, \dots, X_{1n_1}$ , pochází z rozložení  $N(\mu_1, \sigma^2)$ , jeho rozsah  $n_1 \geq 2$ , druhý je  $X_{21}, \dots, X_{2n_2}$ , pochází z rozložení  $N(\mu_2, \sigma^2)$  a jeho rozsah  $n_2 \geq 2$ .

$$M_1 = \frac{1}{n} \sum_{i=1}^{n_1} X_{1i}, \quad M_2 = \frac{1}{n} \sum_{i=1}^{n_2} X_{2i} \quad \dots \text{výběrové průměry 1. a 2. výběru,}$$

$$S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_{1i} - M_1)^2, \quad S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (X_{2i} - M_2)^2 \quad \dots \text{výběrové rozptyly 1. a 2. výběru,}$$

$$S_*^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \quad \dots \text{vážený průměr výběrových rozptylů.}$$

**Definice dvouvýběrového t-testu:** Necht'  $X_{11}, \dots, X_{1n_1}$  je náhodný výběr z rozložení  $N(\mu_1, \sigma^2)$  a  $X_{21}, \dots, X_{2n_2}$  je na něm nezávislý náhodný výběr rozložení  $N(\mu_2, \sigma^2)$ , přičemž  $n_1 \geq 2$  a  $n_2 \geq 2$  a  $\sigma^2$  neznáme. Necht'  $c$  je konstanta.

Test  $H_0: \mu_1 - \mu_2 = c$  proti  $H_1: \mu_1 - \mu_2 \neq c$  (resp.  $H_0: \mu_1 - \mu_2 = c$  proti  $H_1: \mu_1 - \mu_2 < c$  resp.

$H_0: \mu_1 - \mu_2 = c$  proti  $H_1: \mu_1 - \mu_2 > c$ ) se nazývá **dvouvýběrový t-test**.

**Poznámka:** Dvouvýběrový t-test je založen na pivotové statistice  $T = \frac{(M_1 - M_2) - (\mu_1 - \mu_2)}{S_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t(n_1 + n_2 - 2)$ .

Dosadíme-li za  $\mu_1 - \mu_2$  konstantu  $c$ , dostaneme testovou statistku:  $T_0 = \frac{(M_1 - M_2) - c}{S_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ , která se

v případě platnosti  $H_0$  řídí Studentovým rozložením  $t(n_1 + n_2 - 2)$ .

### Provedení dvouvýběrového t-testu pomocí kritického oboru

Vypočteme realizaci testové statistiky  $t_0 = \frac{(m_1 - m_2) - c}{s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ . Stanovíme kritický obor  $W$ .

Pokud  $t_0 \in W$ ,  $H_0$  zamítáme na hladině významnosti  $\alpha$  a přijímáme  $H_1$ .

**Oboustranný test:** Testujeme  $H_0: \mu_1 - \mu_2 = c$  proti  $H_1: \mu_1 - \mu_2 \neq c$ . Kritický obor má tvar:

$$W = (-\infty, -t_{1-\alpha/2}(n_1 + n_2 - 1)) \cup (t_{1-\alpha/2}(n_1 + n_2 - 1), \infty).$$

**Levostranný test:** Testujeme  $H_0: \mu_1 - \mu_2 = c$  proti  $H_1: \mu_1 - \mu_2 < c$ . Kritický obor má tvar:

$$W = (-\infty, -t_{1-\alpha}(n_1 + n_2 - 1)).$$

**Pravostranný test:** Testujeme  $H_0: \mu_1 - \mu_2 = c$  proti  $H_1: \mu_1 - \mu_2 > c$ . Kritický obor má tvar:

$$W = (t_{1-\alpha}(n_1 + n_2 - 1), \infty).$$

## Provedení dvouvýběrového t-testu pomocí intervalu spolehlivosti

Pro **oboustranný test** sestrojíme oboustranný  $100(1-\alpha)\%$  interval spolehlivosti pro rozdíl středních hodnot  $\mu_1 - \mu_2$  při společném neznámém rozptylu  $\sigma^2$ :

$$(d, h) = (m_1 - m_2 - s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{1-\alpha/2}(n_1+n_2-2), m_1 - m_2 + s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{1-\alpha/2}(n_1+n_2-2))$$

Pro **levostranný test** sestrojíme pravostranný  $100(1-\alpha)\%$  interval spolehlivosti pro střední hodnotu  $\mu$  při společném neznámém rozptylu  $\sigma^2$ :

$$(-\infty, h) = (-\infty, m_1 - m_2 + s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{1-\alpha}(n_1+n_2-2))$$

Pro **pravostranný test** sestrojíme levostranný  $100(1-\alpha)\%$  interval spolehlivosti pro střední hodnotu  $\mu$  při společném neznámém rozptylu  $\sigma^2$ :

$$(d, \infty) = (m_1 - m_2 - s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} t_{1-\alpha}(n_1+n_2-2), \infty)$$

Pokud číslo  $c$  padne do tohoto intervalu, nulovou hypotézu nezamítáme na hladině významnosti  $\alpha$ . V opačném případě nulovou hypotézu zamítáme na hladině významnosti  $\alpha$  a přijímáme alternativní hypotézu.

## Provedení dvouvýběrového t-testu pomocí p-hodnoty

Vypočteme p-hodnotu (nejlépe pomocí statistického software) a porovnáme ji s hladinou významnosti  $\alpha$ . Jestliže  $p \leq \alpha$ , pak nulovou hypotézu zamítáme na hladině významnosti  $\alpha$  a přijímáme alternativní hypotézu. Je-li  $p > \alpha$ , pak nulovou hypotézu nezamítáme na hladině významnosti  $\alpha$ .

Způsob výpočtu p-hodnoty:

Označme  $\Phi$  distribuční funkci Studentova rozložení  $t(n_1+n_2-2)$ .

Pro oboustrannou alternativu  $p = 2 \min\{\Phi(t_0), 1 - \Phi(t_0)\}$ .

Pro levostrannou alternativu  $p = \Phi(t_0)$ .

Pro pravostrannou alternativu  $p = 1 - \Phi(t_0)$ .

**Upozornění:** Před provedením dvouvýběrového t-testu je zapotřebí testovat hypotézu o shodě rozptylů daných dvou rozložení (tzv. hypotézu homoskedasticity), tj. hypotézu  $H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$  proti

alternativě  $H_1: \frac{\sigma_1^2}{\sigma_2^2} \neq 1$ . K tomu slouží **F-test**. Je založen na pivotové statistice

$F = \frac{S_1^2/S_2^2}{\sigma_1^2/\sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$ . Nahradíme-li podíl  $\frac{\sigma_1^2}{\sigma_2^2}$  předpokládanou hodnotou 1, dostaneme

testovou statistiku  $\tau_0 = \frac{S_1^2}{S_2^2}$ , která se v případě platnosti  $H_0$  řídí rozložením  $F(n_1-1, n_2-1)$ .



## Provedení F-testu pomocí kritického oboru

Vypočteme realizaci testové statistiky  $t_0 = \frac{s_1^2}{s_2^2}$ . Stanovíme kritický obor  $W$ .

Pokud  $t_0 \in W$ ,  $H_0$  zamítáme na hladině významnosti  $\alpha$  a přijímáme  $H_1$ .

**Oboustranný test:** Testujeme  $H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$  proti  $H_1: \frac{\sigma_1^2}{\sigma_2^2} \neq 1$ . Kritický obor má tvar:

$$W = (0, F_{\alpha/2}(n_1 - 1, n_2 - 1)) \cup (F_{1-\alpha/2}(n_1 - 1, n_2 - 1), \infty).$$

**Levostranný test:** Testujeme  $H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$  proti  $H_1: \frac{\sigma_1^2}{\sigma_2^2} < 1$ . Kritický obor má tvar:

$$W = (0, F_{\alpha}(n_1 - 1, n_2 - 1)).$$

**Pravostranný test:** Testujeme  $H_0: \frac{\sigma_1^2}{\sigma_2^2} = 1$  proti  $H_1: \frac{\sigma_1^2}{\sigma_2^2} > 1$ . Kritický obor má tvar:

$$W = (F_{1-\alpha}(n_1 - 1, n_2 - 1), \infty).$$

## Provedení F-testu pomocí intervalu spolehlivosti

Pro **oboustranný test** sestrojíme oboustranný  $100(1-\alpha)\%$  interval spolehlivosti pro podíl rozptylů  $\frac{\sigma_1^2}{\sigma_2^2}$ :

$$(d, h) = \left( \frac{s_1^2 / s_2^2}{F_{1-\alpha/2}(n_1 - 1, n_2 - 1)}, \frac{s_1^2 / s_2^2}{F_{\alpha/2}(n_1 - 1, n_2 - 1)} \right)$$

Pro **levostranný test** sestrojíme pravostranný  $100(1-\alpha)\%$  interval spolehlivosti pro podíl rozptylů  $\frac{\sigma_1^2}{\sigma_2^2}$ :

$$(0, h) = \left( 0, \frac{s_1^2 / s_2^2}{F_{\alpha}(n_1 - 1, n_2 - 1)} \right)$$

Pro **pravostranný test** sestrojíme levostranný  $100(1-\alpha)\%$  interval spolehlivosti pro podíl rozptylů  $\frac{\sigma_1^2}{\sigma_2^2}$ :

$$(d, \infty) = \left( \frac{s_1^2 / s_2^2}{F_{1-\alpha}(n_1 - 1, n_2 - 1)}, \infty \right).$$

Pokud číslo  $c$  padne do intervalu spolehlivosti, nulovou hypotézu nezamítáme na hladině významnosti  $\alpha$ . V opačném případě nulovou hypotézu zamítáme na hladině významnosti  $\alpha$  a přijímáme alternativní hypotézu.

## Provedení F-testu pomocí p-hodnoty

Vypočteme p-hodnotu (nejlépe pomocí statistického software) a porovnáme ji s hladinou významnosti  $\alpha$ . Jestliže  $p \leq \alpha$ , pak nulovou hypotézu zamítáme na hladině významnosti  $\alpha$  a přijímáme alternativní hypotézu. Je-li  $p > \alpha$ , pak nulovou hypotézu nezamítáme na hladině významnosti  $\alpha$ .

Způsob výpočtu p-hodnoty:

Označme  $\Phi$  distribuční funkci Fisherova - Snedecorova rozložení  $F(n_1-1, n_2-1)$ .

Pro oboustrannou alternativu  $p = 2 \min\{\Phi(t_0), 1 - \Phi(t_0)\}$ .

Pro levostrannou alternativu  $p = \Phi(t_0)$ .

Pro pravostrannou alternativu  $p = 1 - \Phi(t_0)$ .

### **Příklad na dvouvýběrový t-test:**

V restauraci "U bílého koníčka" měřili ve 20 případech čas obsluhy zákazníka. Výsledky v minutách: 6, 8, 11, 4, 7, 6, 10, 6, 9, 8, 5, 12, 13, 10, 9, 8, 7, 11, 10, 5. V restauraci "Zlatý lev" bylo dané pozorování uskutečněno v 15 případech s těmito výsledky: 9, 11, 10, 7, 6, 4, 8, 13, 5, 15, 8, 5, 6, 8, 7. Na hladině významnosti 0,05 testujte hypotézu, že střední hodnoty doby obsluhy jsou v obou restauracích stejné.

### **Řešení pomocí systému STATISTICA:**

Otevřeme datový soubor restaurace.sta s dvěma proměnnými X a ID a 35 případy. V proměnné X jsou zapsány zjištěné doby obsluhy. Proměnná ID slouží k rozlišení restaurací – nabývá hodnoty 1 pro restauraci "U bílého koníčka" a hodnoty 2 pro restauraci „Zlatý lev“.

Na hladině významnosti 0,05 testujeme nulovou hypotézu  $H_0: \mu_1 - \mu_2 = 0$  proti oboustranné alternativě  $H_1: \mu_1 - \mu_2 \neq 0$ . Je to úloha na dvouvýběrový t-test.

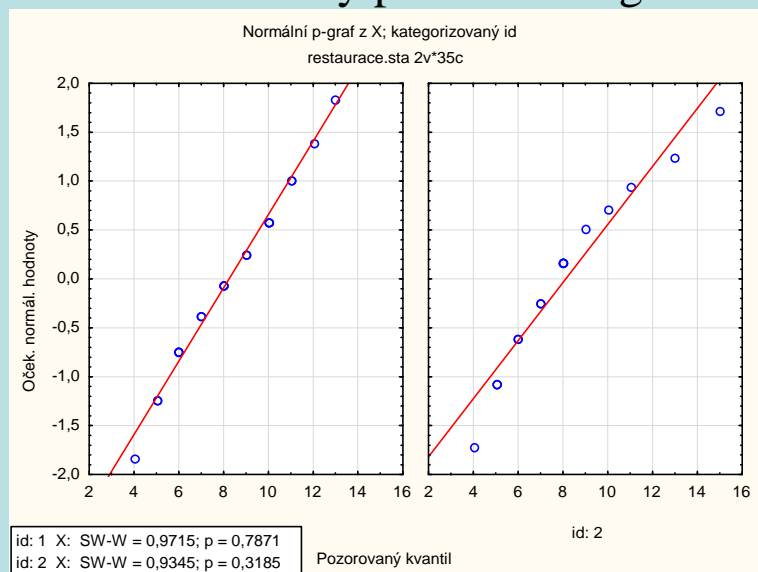
Nejprve ověříme předpoklady dvouvýběrového t-testu.

Nezávislost obou náhodných výběrů: splněno, plyne přímo ze způsobu pořízení dat.

Normalita obou nezávislých náhodných výběrů: ověříme pomocí N-P grafů a Shapirova – Wilkova testu.

**Vytvoření N-P plotu pro oba výběry:** Grafy – 2D Grafy – Normální pravděpodobnostní grafy – Proměnná X – OK - odškrtneme Neurčit průměrnou pozici svázaných pozorování – zaškrtneme Shapiro – Wilkův test – na záložce Kategorizovaný zapneme Kategorie X – Změnit proměnnou - ID – OK – OK.

Ověření normality pomocí N-P grafu a S-W testu:



V obou případech jsou body v N-P grafu blízko ideální přímky.

Pro první výběr S-W test poskytl p-hodnotu 0,7871, pro druhý výběr 0,3185, tedy na hladině významnosti 0,05 nezamítáme ani v jednom případě hypotézu o normalitě.

## Testování hypotézy o shodě rozptylů:

Statistiky – Základní statistiky /tabulky – t-test, nezávislé, dle skupin – OK, Proměnné –Závislé proměnné X, Grupovací proměnná ID – OK.

Po kliknutí na tlačítko Výpočet dostaneme tabulku:

Proměnná	t-testy; grupováno: ID (restaurace.sta)										
	Průměr 1	Průměr 2	t	sv	p	Poč.plat 1	Poč.plat. 2	Sm.odch. 1	Sm.odch. 2	F-poměr Rozptyly	p Rozptyly
X	8,250000	8,133333	0,123730	33	0,902279	20	15	2,510504	3,067495	1,492952	0,410440

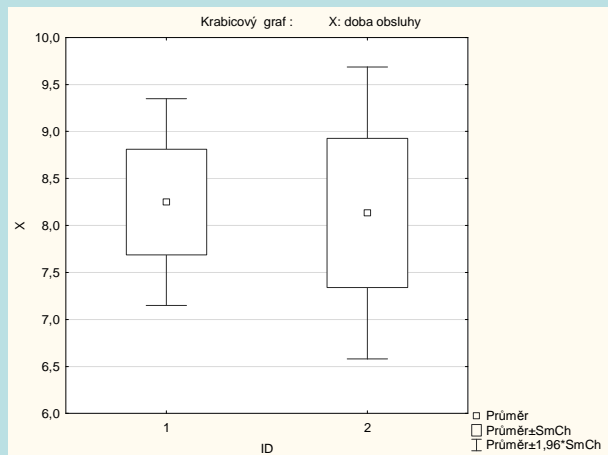
Vidíme, že testová statistika pro test shody rozptylů se realizuje hodnotou 1,492952, odpovídající p-hodnota je 0,41044, tedy na hladině významnosti 0,05 nezamítáme hypotézu o shodě rozptylů.

(Upozornění: v případě zamítnutí hypotézy o shodě rozptylů je zapotřebí v tabulce t-testu pro nezávislé vzorky dle skupin zaškrtnout volbu Test se samostatnými odhady rozptylu.)

## Testování hypotézy o shodě středních hodnot (provedení dvouvýběrového t-testu):

Z výše uvedené výstupní tabulky plyne, že testová statistika pro test shody středních hodnot se realizuje hodnotou 0,12373, počet stupňů volnosti je 33, odpovídající p-hodnota 0,902279, tedy hypotézu o shodě středních hodnot nezamítáme na hladině významnosti 0,05. Neprokázal se rozdíl ve středních hodnotách dob obsluhy v restauracích "U bílého koníčka" a „Zlatý lev“.

Tabulku ještě doplníme krabicovými diagramy. Na záložce Detaily zaškrtneme krabicový graf a vybereme volbu Průměr/SmCh/1,96\*SmCh.



Z grafu je vidět, že průměrná doba obsluhy v první restauraci je nepatrně delší a má menší variabilitu než ve druhé restauraci.

## Cohenův koeficient věcného účinku – doplnění významu dvouvýběrového t-testu:

Cohenův koeficient  $d$  vypočteme podle vzorce:  $d = \frac{|m_1 - m_2|}{s_*}$ .

Tento koeficient slouží k posouzení velikosti rozdílu průměrů, který je standardizován pomocí odmocniny z váženého průměru výběrových rozptylů. Jedná se o tzv. **věcnou významnost** neboli **velikost účinku** skupiny na variabilitu hodnot sledované náhodné veličiny. Velikost účinku hodnotíme podle následující tabulky:

Hodnota $d$	účinek
aspoň 0,8	velký
mezi 0,5 až 0,8	střední
mezi 0,2 až 0,5	malý
pod 0,2	zanedbatelný

(Uvedené hodnoty nemají samozřejmě absolutní platnost, posouzení, jaký účinek považujeme za velký či malý, závisí na kontextu.)

Je zapotřebí si uvědomit, že při dostatečně velkých rozsazích náhodných výběrů i malý rozdíl ve výběrových průměrech způsobí zamítnutí nulové hypotézy na hladině významnosti  $\alpha$ , i když z věcného hlediska tak malý rozdíl nemá význam. Naopak, máme-li výběry malých rozsahů, pak i značně velký rozdíl ve výběrových průměrech nemusí vést k zamítnutí nulové hypotézy na hladině významnosti  $\alpha$ .

V předešlém příkladě  $d = 0,0423$ , tedy účinek restaurace na průměrnou dobu obsluhy je zcela zanedbatelný.



## 4. Provedení t-testů v systému STATISTICA pomocí aplikace Testy rozdílů

Předpokládejme, že jsou splněny podmínky použití t-testů (normalita, v případě dvou výběrů nezávislost a homoskedasticita). Ze zadaných dat zjistíme rozsahy, průměry a směrodatné odchylky.

### Jednovýběrový t-test:

Použijeme příklad s obsahem vitamínu C ve 20 vzorcích mrkve.

V Základních statistikách a tabulkách vypočteme průměr ( $m = 34,86$ ) a směrodatnou odchylku ( $s = 11,087$ ) obsahu vitamínu C. Pak použijeme Testy rozdílů: r, %, průměry – OK – vybereme Rozdíl mezi dvěma průměry (normální rozdělení) – zaškrtneme Výběrový průměr vs. Střední hodnota – do políčka Pr1 napíšeme 34,86, do políčka SmOd1 napíšeme 11,087, do políčka N1 napíšeme 20, do políčka Pr2 napíšeme 35 - Výpočet. Dostaneme p-hodnotu 0,9556, tedy nezamítáme nulovou hypotézu na hladině významnosti 0,05.

Testy rozdílů: r, %, průměry: mrkev.sta

Poslat/tisknout výsledky každ. výpočtu do okna protokolu Storno

Rozdíl mezi dvěma korelačními koeficienty

r1: 0,00 N1: 10 r2: 0,00 N2: 10 p: 1,0000  Jednostr.  Oboustr. Výpočet

Rozdíl mezi dvěma průměry (normální rozdělení)

Pr1: 34,86 SmOd1: 11,087 N1: 20 p: ,9556 Výpočet

Pr2: 35 SmOd2: 1 N2: 10  Jednostr.  Oboustr.

Výběrový průměr vs. střední hodnota

Rozdíl mezi dvěma poměry

P 1: ,500000 N1: 10 P 2: ,500000 N2: 10 p: 1,0000  Jednostr.  Oboustr. Výpočet

## Párový t-test:

Použijeme příklad s měřením vlhkosti 22 vzorků dřeva dvěma metodami.

V Základních statistikách a tabulkách vypočteme průměr ( $m = -0,6318$ ) a směrodatnou odchylku ( $s = 1,3113$ ) rozdílů výsledků obou metod. Pak použijeme Testy rozdílů: r, %, průměry – OK – vybereme Rozdíl mezi dvěma průměry (normální rozdělení) – zaškrtneme Výběrový průměr vs. Střední hodnota – do políčka Pr1 napíšeme  $-0,6318$ , do políčka SmOd1 napíšeme  $1,3113$ , do políčka N1 napíšeme 22, do políčka Pr2 napíšeme 0 - Výpočet. Dostaneme p-hodnotu 0,0346, tedy zamítáme nulovou hypotézu na hladině významnosti 0,05.

The screenshot shows a software window titled "Testy rozdílů: r, %, průměry: vlhkost\_dreva.sta". It contains three sections for different types of difference tests, each with input fields and a "Výpočet" button.

- Top section:** "Rozdíl mezi dvěma korelačními koeficienty". Fields: r1: 0,00, N1: 10, r2: 0,00, N2: 10, p: 1,0000. Radio buttons for "Jednostr." and "Oboustr." (selected).
- Middle section:** "Rozdíl mezi dvěma průměry (normální rozdělení)". Fields: Pr1: -,6318, SmOd1: 1,3113, N1: 22, p: ,0346, Pr2: 0, SmOd2: 1, N2: 10. Radio buttons for "Jednostr." and "Oboustr." (selected). A checked checkbox "Výběrový průměr vs. střední hodnota".
- Bottom section:** "Rozdíl mezi dvěma poměry". Fields: P 1: ,500000, N1: 10, P 2: ,500000, N2: 10, p: 1,0000. Radio buttons for "Jednostr." and "Oboustr." (selected).

## Dvouvýběrový t-test:

Použijeme příklad s dobami obsluhy zákazníků ve dvou restauracích.

V Základních statistikách a tabulkách vypočteme průměry ( $m_1 = 8,25$ ,  $m_2 = 8,1313$ ) a směrodatné odchylky dob obsluhy v obou restauracích ( $s_1 = 2,5105$ ,  $s_2 = 3,0675$ ).

Pak použijeme Testy rozdílů: r, %, průměry – OK – vybereme Rozdíl mezi dvěma průměry (normální rozdělení) – do políčka Pr1 napíšeme 8,25, do políčka SmOd1 napíšeme 2,5105, do políčka N1 napíšeme 20, do políčka Pr2 napíšeme 8,1313, do políčka SmOd2 napíšeme 3,0675, do políčka N2 napíšeme 15 – Výpočet. Dostaneme p-hodnotu 0,9006, tedy nezamítáme nulovou hypotézu na hladině významnosti 0,05.

Testy rozdílů: r, %, průměry: restaurace.sta

Poslat/tisknout výsledky každ. výpočtu do okna protokolu Storno

Rozdíl mezi dvěma korelačními koeficienty

r1: 0,00 N1: 10 p: 1,0000  Jednostr. Výpočet  
r2: 0,00 N2: 10  Oboustr.

Rozdíl mezi dvěma průměry (normální rozdělení)

Pr1: 8,25 SmOd1: 2,5105 N1: 20 p: ,9006 Výpočet  
Pr2: 8,1313 SmOd2: 3,0675 N2: 15  Jednostr.  
 Oboustr.

Výběrový průměr vs. střední hodnota

Rozdíl mezi dvěma poměry

P 1: ,500000 N1: 10 p: 1,0000  Jednostr. Výpočet  
P 2: ,500000 N2: 10  Oboustr.