

# **ZÁKLADY STATISTIKY**

## **3. ANALYTICKÁ STATISTIKA**

**3.1 Opakování: výzkumné soubory,**

**3.2 Hypotéz nulová/alternativní**

**3.3 Věcná a statistická významnost**

**3.4 Testování statistických hypotéz**

# **(MATEMATICKÁ) STATISTIKA**

```
graph TD; A["(MATEMATICKÁ) STATISTIKA"] --> B["DESKRIPTIVNÍ  
(popisná)  
zpracování a popis  
dat"]; A --> C["ANALYTICKÁ  
(inferentní, induktivní)  
analýza a vyhodnocení  
dat"];
```

## **DESKRIPTIVNÍ**

**(popisná)**

**zpracování a popis  
dat**

## **ANALYTICKÁ**

**(inferentní, induktivní)  
analýza a vyhodnocení  
dat**

**Využití analytické statistiky:**

**(1) prokázat významnost či nevýznamnost vlivu**

**intervence mezi výsledky testu vytrvalosti dvou  
tréninkových skupin (tréninková metoda),**

**(2) Prokázat významnost intersexuálních diferencí síly  
mezi soubory tenistů a tenistek 11-12 let (gender).**

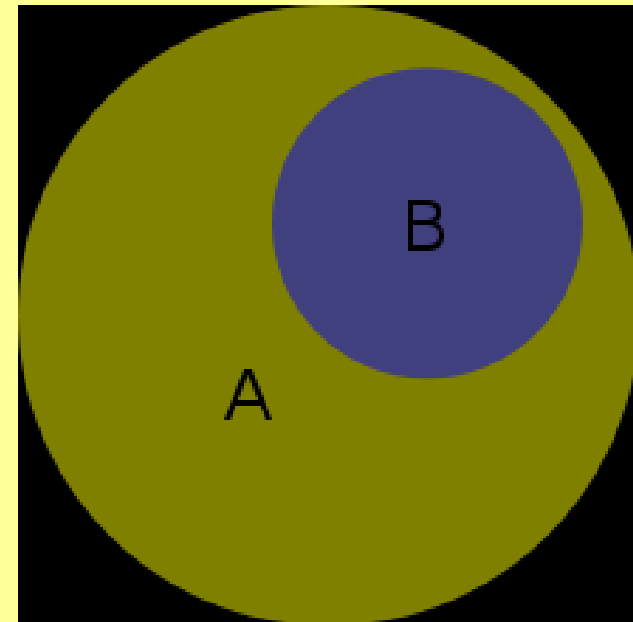
## 3.1 Stručné opakování

### TYPY VÝZKUMNÝCH SOUBORŮ

#### ZÁKLADNÍ SOUBOR (ZS)

(generální soubor, population, Grundgesamtheit) **je soubor všech jedinců, u kterých bychom teoreticky měli šetření provádět.**

**ZS je obvykle není dostupný, výzkum je možný pouze s omezeným počtem jedinců (objektů), soubor nazýváme **výběrový soubor** (sample, Stichprobe).**



**VÝBĚROVÝ SOUBOR** získaný náhodným, resp. záměrným výběrem je podmnožinou prvků základního souboru.

Z poznatků zjištěných u **náhodně vybraného** výběrového souboru, můžeme (při splnění určitých statistických požadavků) činit **závěry platné pro základní soubor.**

## **ZÁVISLÉ SOUBORY**

(test hod na koš, družstvo A 1., 2. pokusy)

## **NEZÁVISLÉ SOUBORY**

(test hod na koš, družstvo A, družstvo B)

## 3.2 HYPOTÉZA

je podmíněný výrok o vztahu mezi dvěma nebo více proměnnými (Kerlinger, 1972).

Hypotézy jsou důležité, nepostradatelné prostředky vědeckého výzkumu; jsou to pracovní nástroje teorie.

### Kritéria dobrých hypotéz

1. hypotézy jsou *výroky o vztazích* mezi proměnnými
2. hypotézy obsahují *jasné implikace* pro ověřování předpokládaných vztahů (např. jestliže ..., pak ...).

**Hypotéza formuluje předpokládaný vztah mezi proměnnými, který se pomocí testování hypotéz zamítá nebo nelze zamítnout.**

**Druhy hypotéz (Röthig, 1992)**

**1. Pracovní hypotéza - subjektivní domněnky o předmětu výzkumného problému.**

Pracovní hypotéza je formulována všeobecně, je základem pro realizaci předvýzkumu.

**2. Výzkumná (věcná) hypotéza – zdůvodněný předpoklad o existenci vztahu mezi dvěma či více proměnnými.**

Zpřesněná formulace, ověřujeme testováním statistických hypotéz.

**3. Statistická hypotéza** - hypotetické tvrzení vyjádřené ve **statistických termínech** o relacích, vyvozených z předpokládaných vztahů ve věcné H.

$$H_0: \mu = \mu_0 \quad H_A: \mu \neq \mu_0 \quad ; \quad H_A: \mu > \mu_0 \quad ; \quad H_A: \mu < \mu_0$$

Stupeň obecnosti ověřovaného tvrzení (hypotézy) klesá (od pracovní H → ke statistické H).

Stupeň přesnosti ověřovaného tvrzení (hypotézy) vzrůstá (od pracovní H → ke statistické H).

Hypotéza je testována pomocí tzv. **testovacích metod** (testů), hypotézu **zamítáme**, resp. **nezamítáme**.

# HYPOTÉZA NULOVÁ

Základním typem úvahy při statistickém testování tzv. *nulová hypotéza* ( $H_0$ ).

Podstatou *nulové hypotézy* je odůvodněný předpoklad, že mezi dvěma jevy **není statisticky významný rozdíl** (rozdíl je nulový, resp. velmi malý).

Jako *nulová hypotéza* se označuje domněnka, že dva statistické soubory **se shodují** v určitých statistických parametrech (např.  $M$ ,  $r$ ).

$$H_0: \mu = \mu_0 \quad (M_1 = M_2; r_1 = r_2)$$



**Nepravděpodobný výsledek ( $H_0$ )** má být stanoven **předem** (*tělesná výška tenistů/-tek U14 je stejná*).

Výsledky testování hypotéz jsou posuzovány na tzv. **hladině významnosti ( $p$ ,  $\alpha$ )**, která vyjadřuje **pravděpodobnost chyby I. druhu** (tedy chybné zamítnutí testované hypotézy).

Úroveň hladiny významnosti  **$p = 0,05$  ( $0,01$ )** znamená, že nulová hypotéza se **zamítá**, když pravděpodobnost platnosti nulové hypotézy **je menší než 5% (1%)**.

# HYPOTÉZA ALTERNATIVNÍ

Předpokládáme-li, že mezi dvěma jevy **existuje významný rozdíl**, formulujeme tzv. **alternativní hypotézu  $H_A$** .

Hypotéza  **$H_A$**  popírá platnost nulové hypotézy ( $H_0$ ), vymezuje situaci, když se  $H_0$  zamítá.

**Výsledek pravděpodobný** (TV  $M \neq \check{Z}$ ; U14,  $H_A$ ), resp. **nepravděpodobný** (TV  $M = \check{Z}$ ; U14,  $H_0$ ) **musí být stanoveno předem**.

**H: oboustranná      resp.      H: jednostranná**

**$H_A: \mu \neq \mu_0$    ;    $H_A: \mu > \mu_0$    ;    $H_A: \mu < \mu_0$**

## 3.2 STATISTICKÁ A VĚCNÁ VÝZNAMNOST

**Brownlee**, J. (2020). *A Gentle Introduction to Statistical Power and Power Analysis in Python*. Retrieved from <https://machinelearningmastery.com/statistical-power-and-power-analysis-in-python/>

**Cohen**, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

**Hopkins**, W. (2016). *A New View of Statistics*. <http://www.sportsci.org/resource/stats/index.html>

**Cumming**, et al. (2012). The statistical recommendations of the American Psychological Association Publication Manual: Effect sizes, confidence intervals, and meta-analysis. *Australian Journal of Psychology*, 64, 138–146. doi:10.1111/j.1742-9536.2011.00037.x

**Soukup**, P. (2013). Věcná významnost výsledků a její možnosti měření. *Data a výzkum*, 7(2), 125–148. <http://dx.doi.org/10.13060/23362391.2013.127.2.41>.

## **3.2 STATISTICKÁ A VĚCNÁ VÝZNAMNOST**

### **(STATISTIC CALCULATORS)**

<https://www.socscistatistics.com/>

<https://www.statskingdom.com/index.html>

[https://www.psychometrica.de/effect\\_size.html](https://www.psychometrica.de/effect_size.html)

<https://effect-size-calculator.herokuapp.com/>

# TESTOVÁNÍ STATISTICKÝCH HYPOTÉZ

*V souladu s názory řady autorů (Brownlee, 2020; Cohen, 1988; Cumming, 2013; Ellis, 2010; Hoppkins, 2016):*

- 1. nejprve posuzujeme **statistickou významnost**, tedy (jde-li o náhodný výběr, resp. randomizovaný výzkum) testujeme **nulovou hypotézu**, jakožto kritérium pro posouzení rizika zobecnění (např. pomocí t-testu),*
- 2. V případě, že nulová hypotéza se zamítá, zhodnotíme **věcnou významnost** (např. pomocí ES koeficient  $d$ ).*

Ukázka: <http://www.socscistatistics.com/effectsize/Default3.aspx>

## A) STATISTICKÁ VÝZNAMNOST

**Pouze statistická významnost výsledků** = dlouhodobá kritika zneužívání tohoto postupu.

Smysluplné použití možné jen pro **reprezentativní soubory** získané metodami **náhodného výběru** a pro **randomizované řízené experimenty**.

**Hlavní nevýhoda** testování  $H_0$  pomocí **statistické významnosti** je její **vazba na rozsah souboru** ( $n$ ):

1) u **velkých výběrů** jsou i nepatrné rozdíly, resp. asociace (korelace) statisticky významné,

2) u **malých výběrů** jsou i velké rozdíly, resp. velká asociace (korelace) statisticky nevýznamné (*tabulka*).



## Statistické tabulky

15

**Tabulka VIII** – Kritické hodnoty pro Pearsonův korelační koeficient (oboustranný test)

	$\alpha$			$\alpha$			$\alpha$	
$n$	0,05	0,01	$n$	0,05	0,01	$n$	0,05	0,01
3	0,9969	0,9999	14	0,5324	0,6614	25	0,3961	0,5052
4	0,9500	0,9900	15	0,5140	0,6411	30	0,3610	0,4629
5	0,8783	0,9587	16	0,4973	0,6226	35	0,3338	0,4296
6	0,8114	0,9172	17	0,4822	0,6055	40	0,3120	0,4026
7	0,7545	0,8745	18	0,4683	0,5897	45	0,2940	0,3801
8	0,7067	0,8343	19	0,4555	0,5751	50	0,2787	0,3610
9	0,6664	0,7977	20	0,4438	0,5614	60	0,2542	0,3301
10	0,6319	0,7646	21	0,4329	0,5487	70	0,2352	0,3060
11	0,6021	0,7348	22	0,4227	0,5368	80	0,2199	0,2864
12	0,5760	0,7079	23	0,4132	0,5256	90	0,2072	0,2702
13	0,5529	0,6835	24	0,4044	0,5151	100	0,1966	0,2565

Zdroj: Anděl, Jiří. *Statistické metody*. 2. vyd. Praha: MATFYZPRESS, 2003

## B) VĚCNÁ VÝZNAMNOST

Posuzovat významnost **rozdílů** či **vztahů** pomocí tzv.

**věcné významnosti** („size of effect“, „effect size“, „velikost efektu“ pomocí ES indexů; Cohen, 1988) se proto doporučuje u **nenáhodných výběrů**, resp. při **zamítnutí nulové hypotézy** (statistická významnost).

*Výhodou použití věcné významnosti je malá závislost na rozsahu souboru ( $n$ ).*

<http://www.socscistatistics.com/effectsize/Default3.aspx>

<https://www.statskingdom.com/index.html>

[https://stats.libretexts.org/Learning\\_Objects/02%3A\\_Interactive\\_Statistics](https://stats.libretexts.org/Learning_Objects/02%3A_Interactive_Statistics)



**Použití věcné významnosti je požadováno jak**

metodology, tak i vědeckými časopisy.

Značný počet výzkumů obsahuje **nesprávnou interpretací výsledků**, z důvodu **používání pouze statistické významnosti**, neboť ji nabízí statistické software.

**Řada autorů** (Brownlee, 2020; Cohen, 1988; Cumming, 2013; Ellis, 2010; Hoppkins, 2016) proto doporučuje zjišťování **velikosti efektu** (effect size, ES), což má význam zejména v případě, že **nulová hypotéza se zamítá**.

# POSUZOVÁNÍ VĚCNÉ VÝZNAMNOSTI

**(1) Cohen** (1988, 1992). **Indexy velikosti efektu**  
(hodnoty pro malé, střední a velké efekty).

Test	Effect size		
	small	medium	large
<i>d</i>	.20	.50	.80
<i>r</i>	.10	.30	.50
$\chi^2$	.10	.30	.50

Vysvětlivky:

d = pro difference středních hodnot

R = pro korelace

$\chi^2$  = pro chí kvadrát

# POSUZOVÁNÍ VĚCNÉ VÝZNAMNOSTI

**(2) Soukup (2013). Effect size po úpravě do intervalů**

Test	small	medium	large
<b>d</b>	0,2-0, 49	0,5-0,79	větší než 0,8
<b>r</b>	0,1-0,29	0,3-0,49	větší než 0,5
<b>Chi2</b>	0,1-0,29	0,3-0,49	větší než 0,5

**Př. 1: Formulace: nulová hypotéza ( $H_0$ )**  
 **$H_{01}$ : intersexuální rozdíly** somatických a motorických předpokladů mezi tenisty ( $n=221$ ) a tenistkami ( $n=193$ ) ve věkové kategorii **11 -12 let jsou nevýznamné.**

Soubor/SC H	Tenisté		Tenistky		Cohen's d, hodnocení efektu
	M	SD	M	SD	
Výška (cm)	155,10	7,62	154,60	6,94	0,07 (žádný)
Hmotnost (kg)	43,50	6,68	43,49	7,17	0,00 (žádný)
MS (kp)	25,14	4,60	23,08	4,61	0,45 (malý)
RS	0,58	0,09	0,53	0,09	0,56 (střední)

**Př. 2: Formulace: alternativní hypotéza ( $H_A$ ,  $H_1$ )**  
 **$H_{A1}$ : intersexuální rozdíly** somatických a motorických předpokladů mezi tenisty ( $n=157$ ) a tenistkami ( $n=163$ ) ve věkové kategorii **13 -14 let** jsou významné.

Category	M (male)	SD	M (female)	SD	Cohen's d
Height (cm)	169.79	9.27	164.93	5.80	0.63 (med)
Weight (kg)	57.05	9.26	53.57	6.31	0.44 (small)
MHSL (kp)	34.64	7.53	29.09	3.84	0.94 (large)
RHSL	0.61	0.10	0.55	0.06	0.73 (med)

# TESTOVÁNÍ STATISTICKÝCH HYPOTÉZ

## POMOCÍ STATISTICKÉ VÝZNAMNOSTI

Nejčastěji posuzujeme

1. Významnost **diferencí středních hodnot** ( $M_1, M_2$ )  
dvou výběrových souborů ( $n_1, n_2$ ),
2. **Míru závislosti** dvou či více jevů (proměnných).

# 1. NOMINÁLNÍ DATA - STATISTICKÉ TESTOVACÍ METODY

## 1. Lyžaři



## 2. Lyžaři



## Znak - kouření

PŘEDPOKLAD	PROBLÉM	TESTOVACÍ METODA
Dva <b>nezávislé soubory</b> (znaky nabývají právě dvou hodnot)	Zkouška významnosti rozdílů souborů	$X^2$ -čtyřpolní test (Fischerův test, čtyřpolní tabulka)
Dva <b>nezávislé soubory</b> (znaky nabývají více hodnot)	Zkouška významnosti rozdílů souborů	$X^2$ -vícepolní test (kontingenční tabulka)
Dva <b>závislé soubory</b> (znaky nabývají právě dvou hodnot)	Zkouška významnosti změn	$X^2$ -Mc Nemarův test
Dva <b>závislé soubory</b>	Hodnocení závislosti	Koef. kontingence C

## 2. ORDINÁLNÍ DATA - STATISTICKÉ TESTOVACÍ METODY

### 1. Gymnasté A



### 2. Gymnasté B



**Znak - body**

PŘEDPOKLAD	PROBLÉM	TESTOVACÍ METODA
Dva <b>nezávislé soubory</b>	Test rovnosti centrálních tendencí	Medianový test (jednoduchý), U-test Mann-Whitneyho, Kolmogorov-Smirnovův test, Marshallův test
Dva <b>závislé soubory</b>	Test rovnosti centrálních tendencí	Znaménkový test, Wilcoxonův test
Více <b>nezávislých souborů</b>	Test rovnosti centrálních tendencí	Medianový test (rozšířený), H-test Kruskal-Wallisův (analýza rozptylu)
Dva <b>závislé soubory</b>	Hodnocení míry závislosti	Spearmanův resp. Kendallův koeficient korelace
Více <b>závislých souborů</b>	Hodnocení míry závislosti	Friedmanova analýza rozptylu



### 3. METRICKÁ DATA - STATISTICKÉ TESTOVACÍ METODY I

#### Tenisté



#### Tenistky



**Znak:  
TV**

PŘEDPOKLAD	PROBLÉM	TESTOVACÍ METODA
Dva <b>nezávislé</b> <b>soubory</b>	Zkouška rovnosti rozptylů (homogenita)	F-test
Dva <b>nezávislé</b> <b>soubory</b>	Zkouška rovnosti středních hodnot	t-test
Dva <b>nezávislé</b> <b>soubory</b>	Zkouška nezávislosti korelací	Korelační test
Dva <b>závislé</b> <b>soubory</b>	Zkouška rovnosti rozptylů (homogenita)	F-test

### 3. METRICKÁ DATA - STATISTICKÉ TESTOVACÍ METODY II

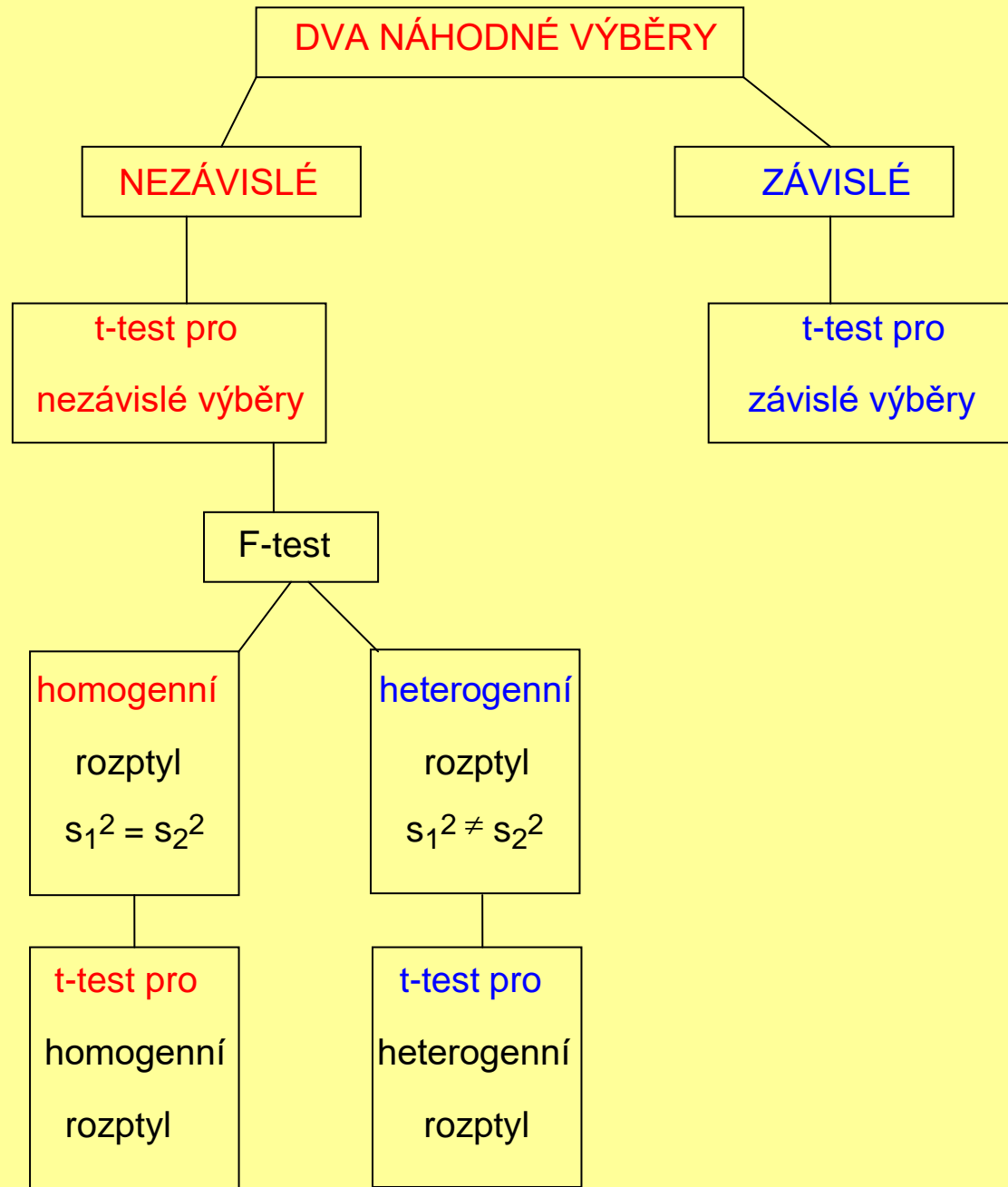
**Tenisté**

**Tenistky**

PŘEDPOKLAD	PROBLÉM	TESTOVACÍ METODA
Dva závislé soubory	Zkouška rovnosti středních hodnot	Diferenční t-test (párový)
Dva závislé soubory	Hodnocení závislosti	Koef. součinné korelace a regrese
Více nezávislých souborů	Zkouška rovnosti průměrů	Analýza rozptylu, Duncanův test pořadí, Bartlettův test
Více nezávislých souborů	Zkouška rovnosti korelačních koeficientů	Test homogeneity

**Znak:  
TV**

# ROZHODOVACÍ DIAGRAM PRO UŽITÍ t-TESTU



# STATISTICKÉ TESTOVACÍ METODY

## Párový t - test

- dva závislé soubory
- zkouška rovnosti středních hodnot

**PŘÍKLAD** – Zjistěte, zda se na automobilu určité značky sjíždějí obě přední pneumatiky stejně rychle

číslo automobilu	1	2	3	4	5	6
pravá pneumatika	1,8	1	2,2	0,9	1,5	1,6
leva pneumatika	1,5	1,1	2	1,1	1,4	1,4
rozdíl	0,3	-0,1	0,2	-0,2	0,1	0,2

$$H_0 : \mu = \mu_1 - \mu_2 = 0$$

$$H_A : \mu = \mu_1 - \mu_2 \neq 0$$

$$T = \frac{|\bar{X} - \mu|}{s} \sqrt{n}$$

$$T < t_{n-1; 1-\frac{\alpha}{2}} \Rightarrow$$

hypotézu nelze  
zamítnou

# STATISTICKÉ TESTOVACÍ METODY

## Párový t - test

číslo automobilu	1	2	3	4	5	6
pravá pneumatika	1,8	1	2,2	0,9	1,5	1,6
leva pneumatika	1,5	1,1	2	1,1	1,4	1,4
rozdíl	0,3	-0,1	0,2	-0,2	0,1	0,2

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{6} (0,3 - 0,1 + 0,2 - 0,2 + 0,1 + 0,2) = \frac{0,5}{6} = \underline{\underline{0,0833}}$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{0,2167^2 + (-0,1833)^2 + 0,1167^2 + (-0,2833)^2 + 0,0167^2 + 0,1167^2}{5} =$$
$$= \frac{0,18833}{5} = 0,0377$$

$$s = \sqrt{s^2} = \sqrt{0,0377} = \underline{\underline{0,1941}}$$

# STATISTICKÉ TESTOVACÍ METODY

## Párový t - test

$$t_{n-1; 1-\frac{\alpha}{2}} = t_{6-1; 1-\frac{0,05}{2}} = t_{5; 0,975} = 2,571$$

= > **z tabulek**

$$T = \frac{|X - \mu|}{s} \sqrt{n} = \frac{0,0833 - 0}{0,1941} \sqrt{6} = 1,0518 < 2,571$$

**Protože  $1,0518 < 2,571$ , nelze na základě získaných dat zamítnout hypotézu, že se obě přední pneumatiky sjíždějí stejně rychle.**

# STATISTICKÉ TESTOVACÍ METODY


## Párový t - test


### Pomocí Excelu – Analýza dat – Dvouvýběrový párový t-test na střední hodnotu

	A	B	C	D	E	F	G	H
1	číslo automobilu	1	2	3	4	5	6	
2	pravá pneumatika	1,8	1	2,2	0,9	1,5	1,6	
3	leva pneumatika	1,5	1,1	2	1,1	1,4	1,4	
4	rozdíl	0,3	-0,1	0,2	-0,2	0,1	0,2	
5								

**Dvouvýběrový párový t-test na střední hodnotu**

**Vstup**

1. soubor:  


2. soubor:  

Hypotetický rozdíl středních hodnot:

☒ Popisky

Alfa:

**Možnosti výstupu**

☐ Výstupní oblast:  

☒ Nový list:

☐ Nový sešit

OK Storno Nápořád

Dvouvýběrový párový t-test na střední hodnotu

	<i>pravá pneumatika</i>	<i>leva pneumatika</i>
Stř. hodnota	1,5	1,41666667
Rozptyl	0,24	0,10966667
Pozorování	6	6
Pears. korelace	0,961571662	
Hyp. rozdíl stř. hodnot	0	
Rozdíl	5	
t Stat	1,051757905	
P(T<=t) (1)	0,17053101	
t krit (1)	2,015048372	
P(T<=t) (2)	0,34106202	
t krit (2)	2,570581835	

# STATISTICKÉ TESTOVACÍ METODY

## Dvouvýběrový t - test

- dva nezávislé soubory
- test rovnosti středních hodnot

**PŘÍKLAD** – U studentů rozdělených do dvou skupin byl zaznamenán počet leh-sedů za 1 minutu. Jsou obě skupiny stejně výkonné?

1. skupina	62	54	55	60	53	58
2. skupina	52	56	49	50	51	

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{(n-1)s_X^2 + (m-1)s_Y^2}} \sqrt{\frac{nm(n+m-2)}{n+m}}$$

$$H_0: \mu_1 = \mu_2$$

$$H_A: \mu_1 \neq \mu_2$$

$$|T| < t_{n+m-2; 1-\frac{\alpha}{2}} \Rightarrow \text{hypotézu nelze zamítnout}$$



# STATISTICKÉ TESTOVACÍ METODY

## Dvouvýběrový t - test

1. skupina	62	54	55	60	53	58
2. skupina	52	56	49	50	51	

$$n_1=6 \quad n_2=5 \quad AP_X=57 \quad AP_Y=51,6 \quad s_X^2=12,8 \quad s_Y^2=7,3$$

$$\begin{aligned} T &= \frac{\bar{X} - \bar{Y}}{\sqrt{(n-1)s_X^2 + (m-1)s_Y^2}} \sqrt{\frac{nm(n+m-2)}{n+m}} = \\ &= \frac{57 - 51,6}{\sqrt{(6-1)12,8 + (5-1)7,3}} \sqrt{\frac{6 \cdot 5 \cdot (6+5-2)}{6+5}} = \\ &= \frac{5,4}{\sqrt{62,5 + 29,2}} \sqrt{24,55} = 2,79 \end{aligned}$$

# STATISTICKÉ TESTOVACÍ METODY

## Dvouvýběrový t -test

$$t_{m+m-2;1-\frac{\alpha}{2}} = t_{6+5-2;1-\frac{0,05}{2}} = t_{9;0,975} = 2,262 = > \text{ z tabulek}$$

$$|T| = 2,79 \geq 2,262$$

**Protože  $2,79 \geq 2,262$  zamítáme hypotézu, že se obě skupiny studentů jsou stejně výkonné.**

# STATISTICKÉ TESTOVACÍ METODY

## Dvouvýběrový t - test

### Pomocí Excelu – Analýza dat – Dvouvýběrový t-test s rovností rozptylů

	A	B	C	D	E	F	G	H
1	1. skupina	62	54	55	60	53	58	
2	2. skupina	52	56	49	50	51		

**Dvouvýběrový t-test s rovností rozptylů**

Vstup

1. soubor:

2. soubor:

Hypotetický rozdíl středních hodnot:

☒ Popisky

Alfa:

Možnosti výstupu

☐ Výstupní oblast:

☒ Nový list:

☐ Nový sešit

OK Storno Nápořěda

#### Dvouvýběrový t-test s rovností rozptylů

	1. skupina	2. skupina
Stř. hodnota	57	51,6
Rozptyl	12,8	7,3
Pozorování	6	5
Společný rozptyl	10,35555556	
Hyp. rozdíl stř. hodnot	0	
Rozdíl	9	
t Stat	2,77122216	
P(T<=t) (1)	0,010855041	
t krit (1)	1,833112923	
P(T<=t) (2)	0,021710083	
t krit (2)	2,262157158	

# STATISTICKÉ TESTOVACÍ METODY

## F - test

- dva nezávislé soubory
- zkouška rovnosti rozptylů

**PŘÍKLAD** – Na základě dat uvedených v předchozím příkladě rozhodněte, zda oba základní soubory mají stejné rozptyly.

1. skupina	62	54	55	60	53	58
2. skupina	52	56	49	50	51	

$$Z = \frac{s_X^2}{s_Y^2} \quad \text{volím tak, aby } Z > 1$$

$$H_0: \sigma_X^2 = \sigma_Y^2$$

$$H_A: \sigma_X^2 \neq \sigma_Y^2$$

$$Z < F_{n-1, m-1; 1-\frac{\alpha}{2}} \Rightarrow \text{hypotézu nelze zamítnou}$$

# STATISTICKÉ TESTOVACÍ METODY

## F - test

1. skupina	62	54	55	60	53	58
2. skupina	52	56	49	50	51	

$$n=6 \quad m=5 \quad s_x^2=12,8 \quad s_y^2=7,3$$

$$Z = \frac{s_X^2}{s_Y^2} = \frac{12,8}{7,3} = 1,753$$

$$F_{n-1, m-1; 1-\frac{\alpha}{2}} = F_{6-1, 5-1; 1-\frac{0,05}{2}} = F_{5, 4; 0,975} = 9,36 = > \text{ z tabulek}$$

$$Z = 1,753 < 9,36$$

**Protože  $1,753 < 9,36$  nelze zamítnout hypotézu o shodnosti rozptylů.**

# STATISTICKÉ TESTOVACÍ METODY

## F - test

### Pomocí Excelu – Analýza dat – Dvouvýběrový F-test pro rozptyl

	A	B	C	D	E	F	G	H
1	1. skupina	62	54	55	60	53	58	
2	2. skupina	52	56	49	50	51		

**Dvouvýběrový F-test pro rozptyl**

**Vstup**

1. soubor: \$A\$1:\$G\$1

2. soubor: \$A\$2:\$F\$2

☒ Popisky

Alfa: 0.05

**Možnosti výstupu**

☐ Výstupní oblast:

☒ Nový list:

☐ Nový sešit

OK

Storno

Nápověda

Dvouvýběrový F-test pro rozptyl

	1. skupina	2. skupina
Stř. hodnota	57	51.6
Rozptyl	12.8	7.3
Pozorování	6	5
Rozdíl	5	4
F	1.753424658	
P(F<=f) (1)	0.303172533	
F krit (1)	6.256056502	

**Příklady výpočtů v souboru:**

**A\_statisticka analyza dat\_priklady.docx**



***Děkuji  
za pozornost***