Analýza dat pro Neurovědy

RNDr. Eva Koriťáková, Ph.D.
doc. RNDr. Ladislav Dušek, Dr.

Blok 3

Jak a kdy použít parametrické a neparametrické testy I.

Osnova

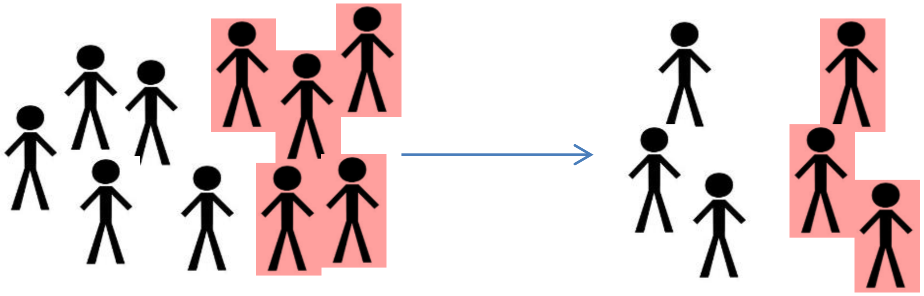
1. Vhodná volba typu testu v různých situacích
2. Jednovýběrové testy
3. Párové testy
4. Dvouvýběrové testy
5. Neparametrické testy

Statistické testování - postup

1. **Sestavíme hypotézu** k ověření (např. chceme ověřit, jestli se pacienti a zdravá populace liší v nějakém parametru)

2. **Vybereme** vhodný statistický test

3. **Stanovíme velikost** vzorku a **provedeme výběr z populace** (např. vybereme pacienty a zdravé lidi a naměříme zkoumaný parametr)

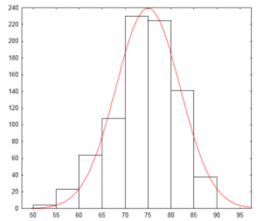


4. **Aplikujeme** vhodný **statistický test** a **rozhodneme**, jestli hypotézu zamítáme, nebo ne

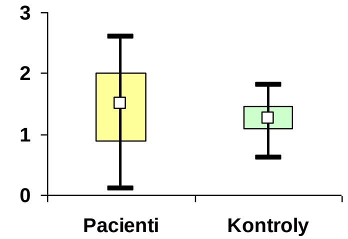
1. Vhodná volba testu v různých situacích

Výběr statistického testu se provádí na základě

- **Typu dat** – ordinální, nominální data, nebo kvantitativní data?
- **Rozdělení** dat – u **parametrických** testů.
  - **Normalita** předpokladem mnoha testů



- **Homogenity rozptylu** srovnávaných skupin – tzn. předpokladu, že rozptyl ve skupinách je přibližně stejný.
  - mnoho testů vyžaduje homogenitu rozptylu



- **Typu hypotézy (srovnání):**
  - 1 skupina vs. referenční hodnota (jednovýběrový test)
  - 1 skupina před a po (párový test)
  - 2 skupiny mezi sebou (dvouvýběrový test)
  - Více skupin mezi sebou
- **Typu alternativní hypotézy:** oboustranná vs. jednostranná

Předpoklady statistického testu

- Výše uvedené podmínky pro výběr statistického testu jsou zároveň **předpoklady použití statistického testu**

- Další předpoklad: **vyrovnané počty subjektů** ve srovnávaných skupinách – aby byly odhady ve srovnávaných skupinách podobně přesné a spolehlivé

- Splnění **všech předpokladů** je důležité pro **použití statistického testu**

V případě, že tyto předpoklady nejsou splněny, nemůžeme důvěřovat výsledkům testu !!!

Parametrické a neparametrické testy

- **Parametrické testy:**
  - Mají předpoklady o rozdělení vstupních dat (např. předpoklad normálního rozdělení), protože se zabývají testováním tvrzení o neznámých parametrech rozdělení (např. střední hodnoty)
  - Mají větší sílu než neparametrické testy

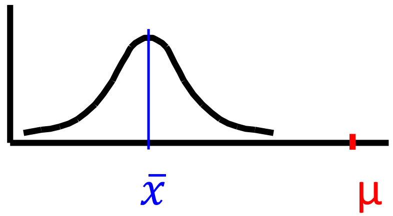
- **Neparametrické testy:**
  - Nemají předpoklady o rozdělení vstupních dat
  - Možné je použít při **asymetrickém** rozdělení nebo **odlehlých** hodnotách
  - Nevyužívají původní hodnoty, ale nejčastěji pouze jejich pořadí – tím dochází **k redukci informační hodnoty** původních dat, a proto mají **menší sílu**
  - Menší sílu testu je možné vykompenzovat větší velikostí vzorku
  - Používání neparametrických testů je „bezpečnější“

Parametrické a neparametrické testy pro kvantitativní data – přehled

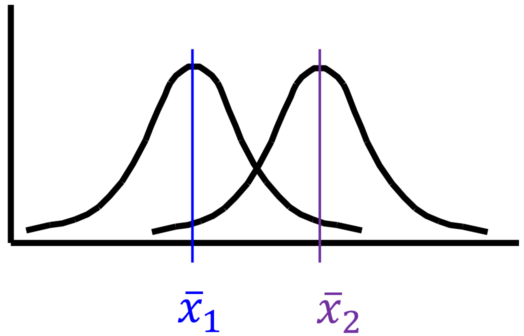
Typ srovnání	Parametrický test	Neparametrický test
1 skupina dat s referenční hodnotou – jednovýběrové testy:	Jednovýběrový t-test, jednovýběrový z-test	Wilcoxonův test
2 skupiny dat párově – párové testy:	Párový t-test	Wilcoxonův test, znaménkový test
2 skupiny dat nepárově – dvouvýběrové testy:	Dvouvýběrový t-test	Mannův-Whitneyův test, mediánový test
Více skupin nepárově:	ANOVA	Kruskalův- Wallisův test

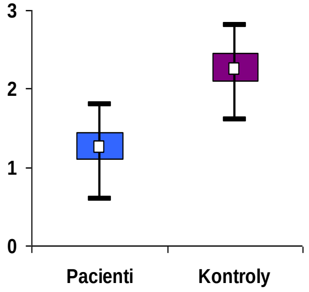
Jednovýběrové a dvouvýběrové testy

- **Jednovýběrové testy:**
  - **Srovnávají jeden vzorek s referenční hodnotou** (popřípadě se statistickým parametrem cílové populace)
  - Průměrný objem hipokampu u 406 pacientů s MCI v našem souboru vs. 6575 mm^3 zjištěným při populačním epidemiologickém průzkumu.



- **Dvouvýběrové testy:**
  - **Srovnáváme dvě skupiny dat**
  - Příklady: srovnání objem hipokampu u mužů a u žen, srovnání kognitivního výkonu podle dvou kategorií věku.

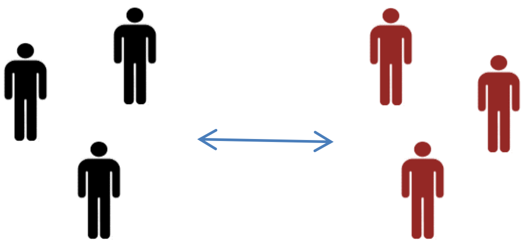




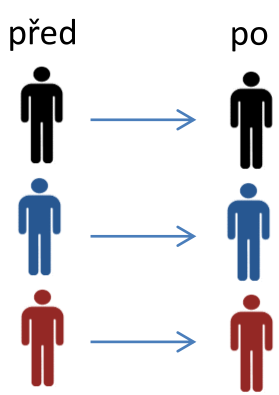
Rozdělení na základě typu srovnání II:

Párové a nepárové testy

- **Nepárové testy:**
  - **Srovnáváme dvě skupiny dat**, které jsou na sobě **nezávislé** – mezi objekty neexistuje vazba.
  - Příklady: srovnání objem hipokampu u mužů a u žen, srovnání kognitivního výkonu podle dvou kategorií věku.



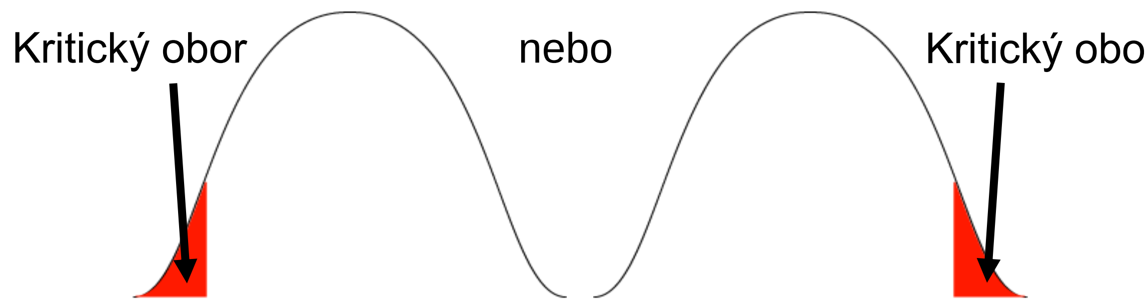
- **Párové testy:**
  - **Srovnáváme dvě skupiny dat**, které jsou na sobě **závislé** – mezi objekty existuje vazba.
  - Příklady: hodnota krevního tlaku **před** začátkem léčby a **po** ukončení léčby



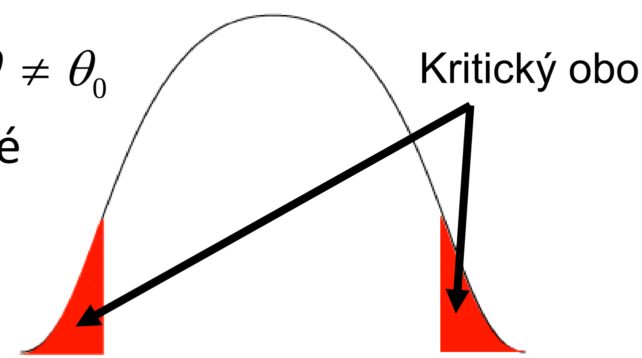
Rozdělení na základě typu alternativní hypotézy:

Jednostranné a oboustranné testy

- **Jednostranné („One-Tailed“) testy:**
  - Jednostranná alternativní hyp.: $H_1 : \theta < \theta_0$ $H_1 : \theta > \theta_0$
  - Např. testujeme, zda je objem mozkové struktury menší u žen než u mužů či zda je průměrná spotřeba tišících léků větší u pacientů než populační průměr apod.



- **Oboustranné („Two-Tailed“) testy:**
  - Oboustranná alternativní hyp.: $H_1 : \theta \neq \theta_0$
  - Např. testujeme, zda se objem mozkové struktury liší u žen a mužů apod.



Shrnutí zásad při testování

1. **Znát základní typy testů** a vědět, pro jaká data se používají.
2. **Ověřit předpoklady testu** – smysl má pouze aplikace „správného“ testu na „správná“ data.
3. Posoudit, zda je výsledek **významný i z klinického hlediska**.
4. Být si vědom toho, že **statistický test není nic víc než matematický vzorec** aplikovaný na data, tedy existuje nenulová pravděpodobnost, že výsledek bude chybný (viz chyba I. a II. druhu). Ovlivnit výsledky testu můžeme například změnou velikosti vzorku.

Parametrické a neparametrické testy pro kvantitativní data – přehled

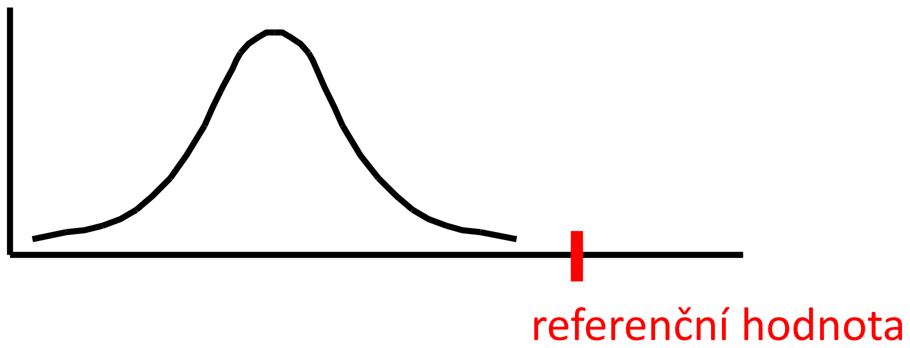
Typ srovnání	Parametrický test	Neparametrický test
1 skupina dat s referenční hodnotou – jednovýběrové testy:	**Jednovýběrový t-test, jednovýběrový z-test**	Wilcoxonův test
2 skupiny dat párově – párové testy:	Párový t-test	Wilcoxonův test, znaménkový test
2 skupiny dat nepárově – dvouvýběrové testy:	Dvouvýběrový t-test	Mannův-Whitneyův test, mediánový test
Více skupin nepárově:	ANOVA	Kruskalův- Wallisův test

2. Jednovýběrové testy

Jednovýběrové („One-Sample“) testy

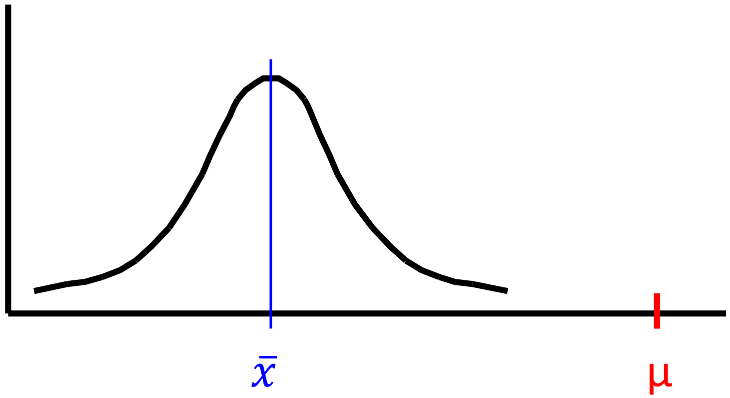
- **Srovnávají jeden vzorek („one sample“) s referenční hodnotou** (popřípadě se statistickým parametrem cílové populace).
- V testu je tedy srovnáváno rozložení hodnot (vzorek) s jediným číslem (referenční hodnota, hodnota cílové populace).
- Otázka položená v testu může být vztažena k průměru, rozptylu,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

- Parametrické jednovýběrové testy, kterým se budeme věnovat:
  - **jednovýběrový t-test** (test o střední hodnotě při neznámém rozptylu)
  - **jednovýběrový z-test** (test o střední hodnotě při známém rozptylu)



Jednovýběrový t-test

- **Srovnáváme střední hodnotu jednoho výběru s referenční hodnotou.**
- Jde o test o střední hodnotě při **neznámém** rozptylu – tzn. testujeme, zda se průměr dané proměnné v našem výběru liší od referenční hodnoty (často populačního průměru), přičemž rozptyl dané proměnné počítáme z našeho výběru.



- Předpoklad: **normalita dat**
- Testová statistika: $T = \dfrac{\bar{x} - \mu}{s / \sqrt{n}}$

Jednovýběrový t-test

- **Příklad:** Chceme srovnat průměrný objem hipokampu u 406 pacientů s MCI v našem souboru s průměrným objemem hipokampu 6575 mm^3 zjištěným při populačním epidemiologickém průzkumu.

- Tzn. hypotézy budou mít tvar: $H_0: \bar{x} = 6575$ a $H_1: \bar{x} \neq 6575$

- **Postup:**
  1. Ověření normality – vykreslíme histogram objemu hipokampu pacientů s MCI.
  2. Aplikujeme statistický test – 3 možnosti:
     I. Testování pomocí p-hodnoty
     II. Testování pomocí kritického oboru
     III. Testování pomocí intervalu spolehlivosti
  3. Nulovou hypotézu zamítneme nebo nezamítneme.

Testování pomocí p-hodnoty (v SPSS)

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Hippocampus_volume (mm3)	406	6552.614	176.1691	8.7431

One-Sample Test

	Test Value = 6575					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Hippocampus_volume (mm3)	-2.560	405	.011	-22.3861	-39.574	-5.199

Protože p-hodnota 0,011 < 0,05 → zamítáme nulovou hypotézu → **Průměrný objem hipokampu u pacientů s MCI v našem souboru se statisticky významně liší od populačního průměru.**

Z popisné sumarizace vidíme, že je menší než populační průměr (konkrétně o 22,4 mm^3).

Testování pomocí p-hodnoty (ručně)

Příklad: Chceme srovnat průměrný objem hipokampu u 406 pacientů s MCI v našem souboru s průměrným objemem hipokampu 6575 mm^3 zjištěným při populačním epidemiologickém průzkumu.

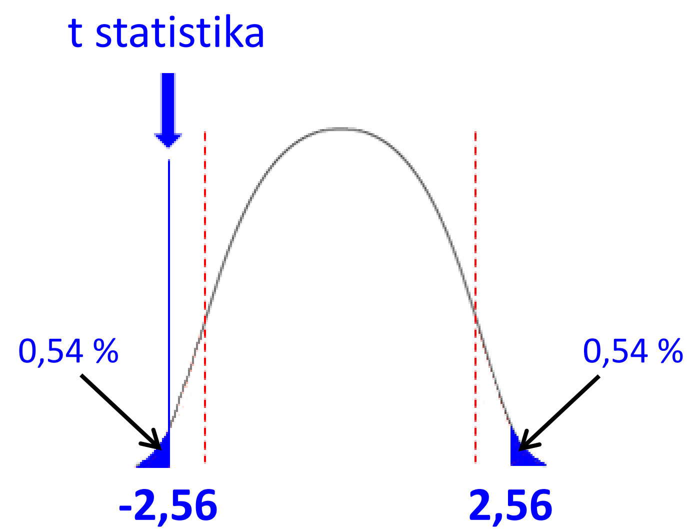
Výpočet testové statistiky:

$n = 406$

$\bar{x} = 6552{,}6\ \text{mm}^3$

$s = 176{,}2\ \text{mm}^3$

$$t = \frac{\bar{x}-\mu}{s/\sqrt{n}} = \frac{6552{,}6-6575}{176{,}2/\sqrt{406}} = -2{,}56$$



Výpočet p-hodnoty:

$$p = 2\cdot\left(P\left(T \le -2{,}56\right)\right) = 2\cdot 0{,}0054 = 0{,}0108$$

Protože p-hodnota 0,0108 < 0,05 → zamítáme nulovou hypotézu → **Průměrný objem hipokampu u pacientů s MCI v našem souboru se statisticky významně liší od populačního průměru.**

Testování pomocí kritického oboru

Příklad: Chceme srovnat průměrný objem hipokampu u 406 pacientů s MCI v našem souboru s průměrným objemem hipokampu 6575 mm^3 zjištěným při populačním epidemiologickém průzkumu.

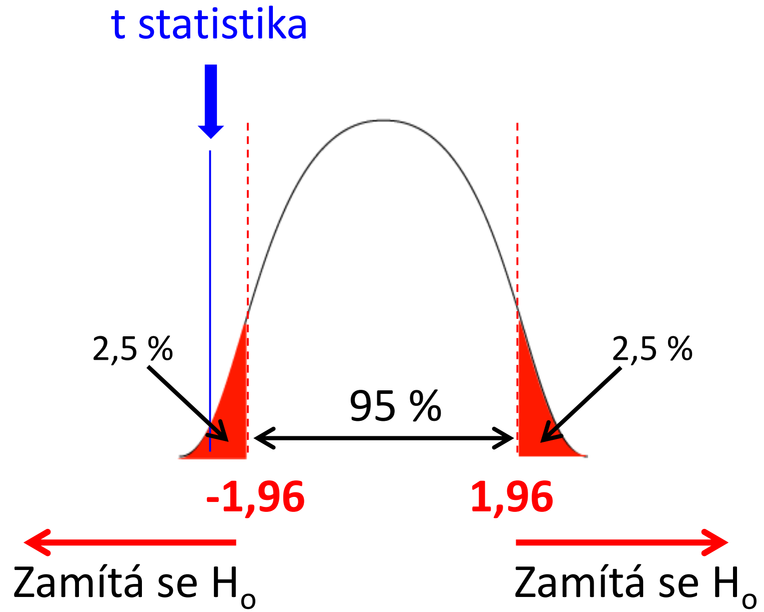
Výpočet testové statistiky:

$n = 406$

$\bar{x} = 6552{,}6 \text{ mm}^3$

$s = 176{,}2 \text{ mm}^3$

$$t = \frac{\bar{x}-\mu}{s/\sqrt{n}} = \frac{6552{,}6-6575}{176{,}2/\sqrt{406}} = -2{,}56$$



Stanovení kritického oboru:

kritické hodnoty: $t_{\alpha/2}(405) \cong -1{,}96$

$t_{1-\alpha/2}(405) \cong 1{,}96$

Protože testová statistika t=-2,56 leží v kritickém oboru → zamítáme nulovou hypotézu → **Průměrný objem hipokampu u pacientů s MCI v našem souboru se statisticky významně liší od populačního průměru.**

Testování pomocí intervalu spolehlivosti

Příklad: Chceme srovnat průměrný objem hipokampu u 406 pacientů s MCI v našem souboru s průměrným objemem hipokampu 6575 mm^3 zjištěným při populačním epidemiologickém průzkumu.

Výpočet intervalu spolehlivosti:

$n = 406$

$\bar{x} = 6552{,}6$ mm^3

$s = 176{,}2$ mm^3

$$\bar{x} - \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1) \leq \mu \leq \bar{x} + \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1)$$

$$6552{,}6 - \frac{176{,}2}{\sqrt{406}} t_{1-0{,}05/2}(406-1) \leq \mu \leq 6552{,}6 + \frac{176{,}2}{\sqrt{406}} t_{1-0{,}05/2}(406-1)$$

$$6535{,}4 \leq \mu \leq 6569{,}8$$

Protože 95% interval spolehlivosti (6535,4; 6569,8) neobsahuje populační průměr 6575 → zamítáme nulovou hypotézu → **Průměrný objem hipokampu u pacientů s MCI v našem souboru se statisticky významně liší od populačního průměru.**

Zmenšení N

N = 406

Mean	Std.Dv.	**N**	Std.Err.	Lower CI	Upper CI	Reference	t-value	df	**p**
6552,6	176,2	**406**	8,7	6535,4	6569,8	6575	-2,56	405	**0,0108**

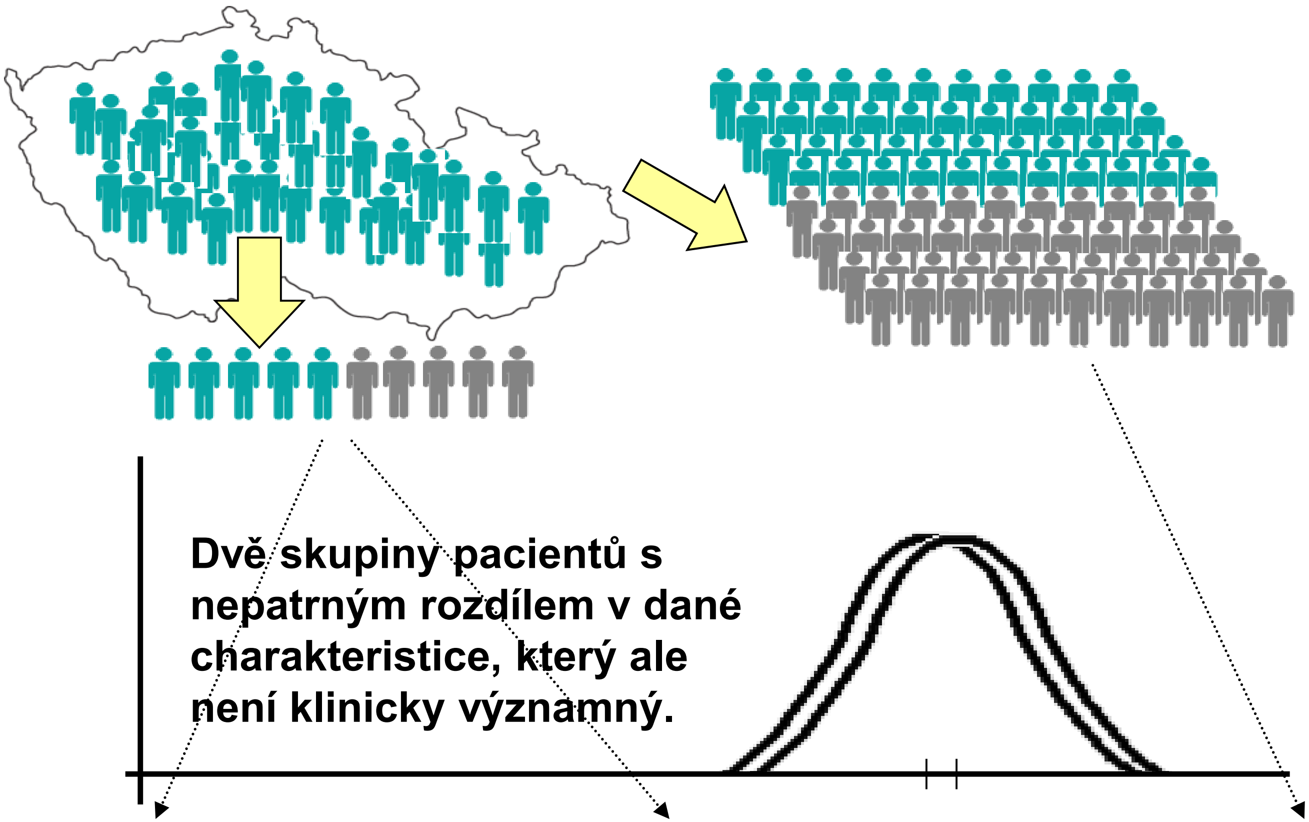
p=0,0108 < 0,05 → zamítáme nulovou hypotézu

N = 100

Mean	Std.Dv.	**N**	Std.Err.	Lower CI	Upper CI	Reference	t-value	df	**p**
6552,2	171,4	**100**	17,1	6518,2	6586,2	6575	-1,33	99	**0,1865**

p=0,1865 > 0,05 → nezamítáme nulovou hypotézu

Vliv velikosti vzorku na výsledky testování - opakování



n_1 = 10, n_2 = 10	**n_1 = 100, n_2 = 100**	**n_1 = 1000, n_2 = 1000**
p = 0,797	p = 0,140	p < 0,001

Statistická významnost způsobená velkým N

Oboustranný vs. jednostranný jednovýběrový t-test

Oboustranný jednovýběrový t-test:

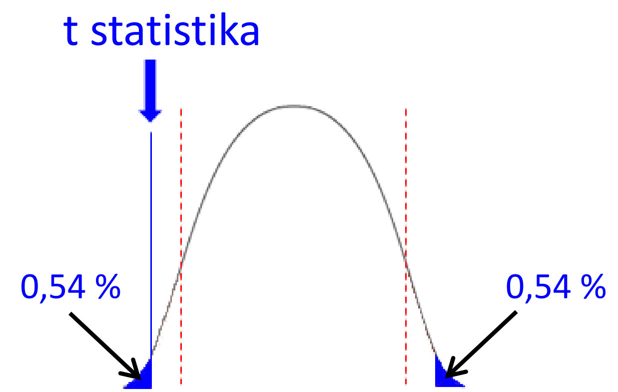
Příklad: Chceme srovnat objem hipokampu u pac. s MCI s populačním průměrem. Tzn. chceme ověřit, zda se objem hipokampu u pac. s MCI v našem souboru **liší** od populačního průměru.

Alternativní hypotéza: $H_1 : \bar{x} \neq \mu$

p = 0,0108

$\bar{x} = 6552{,}6$ mm^3

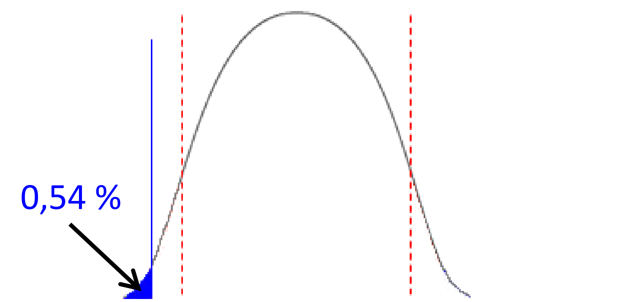
$\mu = 6575$ mm^3



Jednostranný jednovýběrový t-test:

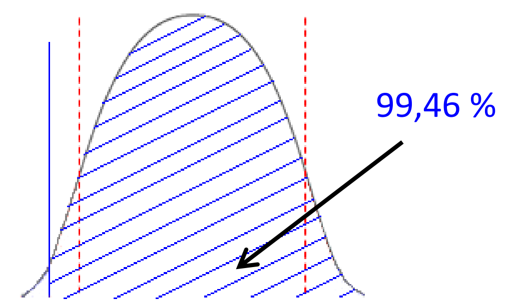
1. **Levostranný** – příklad: Chceme ověřit, zda je objem hipokampu u pac. s MCI v našem souboru **menší** než populační průměr: $H_1 : \bar{x} < \mu$

   p = 0,0108/2 = 0,0054



2. **Pravostranný** – příklad: Chceme ověřit, zda je objem hipokampu u pac. s MCI v našem souboru **větší** než populační průměr: $H_1 : \bar{x} > \mu$

   p = 1 - 0,0108/2 = 0,9946

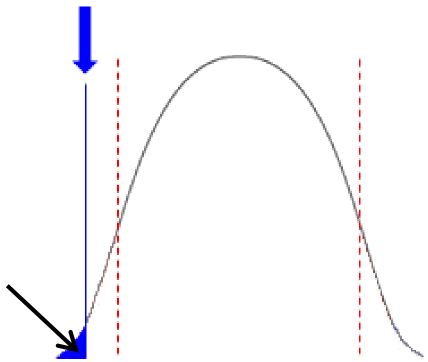


Jednostranný jednovýběrový t-test

Skutečnost: $\overline{x} < \mu$

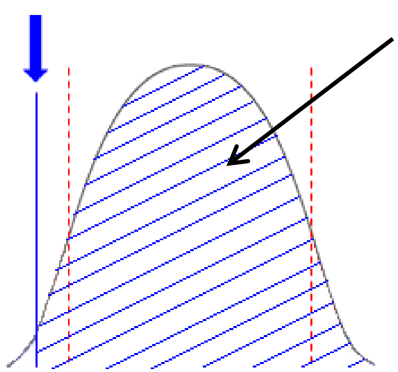
Levostranný jednovýběrový t-test:

$H_1\colon\ \bar{x} < \mu$



Pravostranný jednovýběrový t-test:

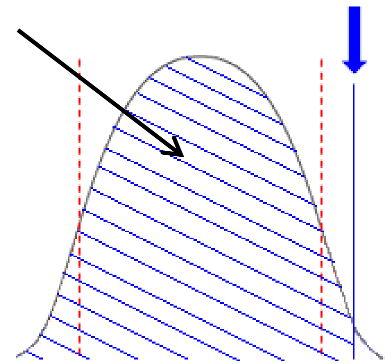
$H_1\colon\ \bar{x} > \mu$



Skutečnost: $\overline{x} > \mu$

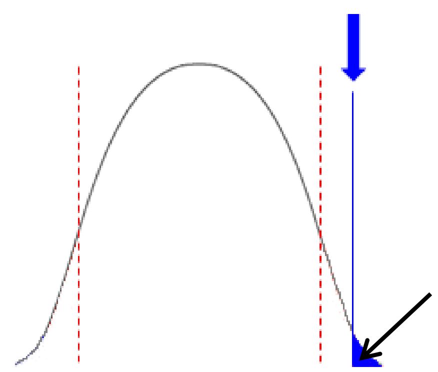
Levostranný jednovýběrový t-test:

$H_1\colon\ \bar{x} < \mu$



Pravostranný jednovýběrový t-test:

$H_1\colon\ \bar{x} > \mu$



Úkol 1

Zadání: Zjistěte, zda se liší průměrný objem amygdaly u mužů v našem souboru od populačního průměrného objemu 2800 mm^3 (nezapomeňte ověřit předpoklady).

Řešení:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Amygdala_volume (mm3)	482	2786,29	239,360	10,903

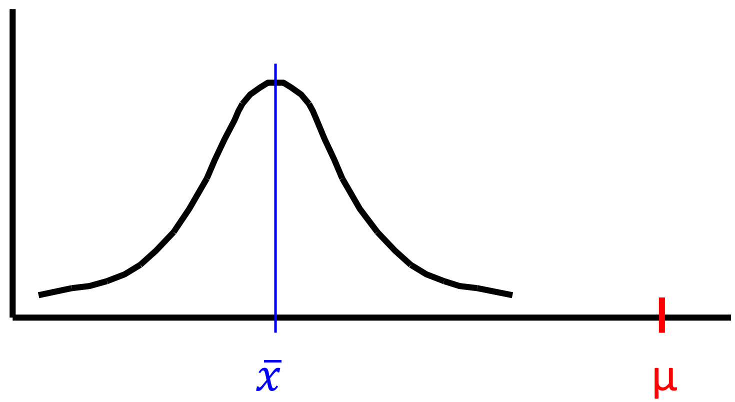
One-Sample Test

	Test Value = 2800					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Amygdala_volume (mm3)	-1,257	481	,209	-13,708	-35,13	7,71

Závěr: Neprokázali jsme na základě našich dat, že by se průměrný objem amygdaly u mužů statisticky významně lišil od hodnoty 2800 mm^3.

Z-test

- **Srovnáváme střední hodnotu jednoho výběru s referenční hodnotou.**
- Jde o test o střední hodnotě při **známém** rozptylu – tzn. testujeme, zda se průměr dané proměnné v našem výběru liší od referenční hodnoty (často populačního průměru), přičemž známe rozptyl dané proměnné pro celou populaci.



- Předpoklad: **normalita dat**
- Testová statistika: $Z = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$

Z-test

- **Příklad:** Při populačním průzkumu bylo zjištěno, že průměrná hodnota MMSE skóre je 27,5 (SD = 4). Chceme zjistit, zda se průměrná hodnota MMSE skóre u 406 pacientů s MCI v našem souboru liší od populační průměrné hodnoty.
- Tzn. hypotézy budou mít tvar: $H_0 : \bar{x} = 27{,}5$ a $H_1 : \bar{x} \neq 27{,}5$
- **Postup:**

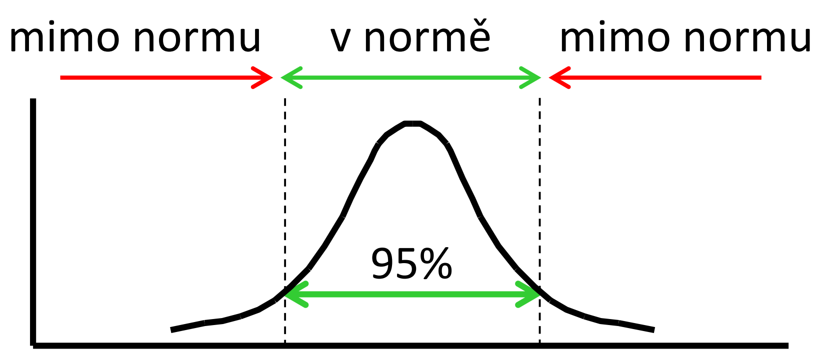
1. Ověření normality – vykreslíme histogram MMSE skóre u pacientů s MCI, abychom ověřili, že průměr je dobrý ukazatel středu hodnot.
2. Aplikujeme statistický test – vypočítáme p-hodnotu:
    - v Excelu:
      =2*MIN(Z.TEST(A1:A406;27,5;4);1-Z.TEST(A1:A406;27,5;4))
    - v Matlabu: [H,P] = ztest(X,27.5,4)
3. Nulovou hypotézu zamítneme nebo nezamítneme:
   p=0,013 < 0,05 → zamítáme nulovou hypotézu → **Průměrná hodnota MMSE skóre u pacientů s MCI v našem souboru se statisticky významně liší od populačního průměru.**

Z-skóre

- Odečtení populačního průměru (μ) a vydělení populační směrodatnou odchylkou (σ): $u_i = \frac{x_i - \mu}{\sigma}$

- Souvislost se standardizací: $u_i = \frac{x_i - \bar{x}}{s}$

- Často při hodnocení různých skóre – určuje se, kteří lidé jsou mimo normu.



Parametrické a neparametrické testy pro kvantitativní data – přehled

Typ srovnání	Parametrický test	Neparametrický test
1 skupina dat s referenční hodnotou – jednovýběrové testy:	Jednovýběrový t-test, jednovýběrový z-test	Wilcoxonův test
2 skupiny dat párově – párové testy:	**Párový t-test**	Wilcoxonův test, znaménkový test
2 skupiny dat nepárově – dvouvýběrové testy:	Dvouvýběrový t-test	Mannův-Whitneyův test, mediánový test
Více skupin nepárově:	ANOVA	Kruskalův- Wallisův test

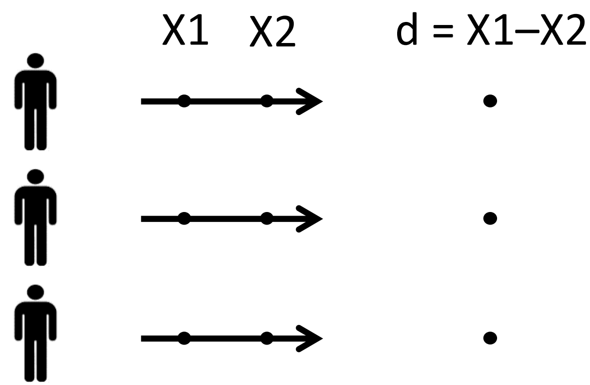
3. Párové testy

Párový t-test

- **Srovnáváme dvě skupiny dat, které ale na sobě nejsou nezávislé – mezi objekty existuje vazba** (např. člověk před a po operaci, stejný kmen krys)

- Příklady: srovnání objemu hipokampu na začátku léčby a 1 rok po zahájení léčby, srovnání kognitivního výkonu pacientů před a po léčbě

- Test je v podstatě prováděn na **diferencích skupin** (rozdílech původních hodnot), nikoliv na původních datech → **obě skupiny tedy musí mít shodný počet hodnot** (všechna měření v jedné skupině musí být spárována s měřením v druhé skupině!)



- Předpoklad: **normalita diferencí** (rozdílů původních hodnot)

- Testová statistika: $T = \frac{\bar{d} - d_0}{s_d / \sqrt{n}}$, kde $\bar{d}$ je průměrný rozdíl, d_0 je referenční hodnota (většinou 0), s_d je směrodatná odchylka rozdílů

Párový t-test

- **Příklad:** Chceme srovnat, zda se liší objem hipokampu u pacientů s Alzheimerovou chorobou při vstupu do studie a 2 roky po zahájení studie (tzn. chceme zjistit, zda došlo ke změně objemu hipokampu).
- Tzn. hypotézy budou mít tvar: $H_0 : \bar{d} = 0$ a $H_1 : \bar{d} \neq 0$
- **Postup:**
  1. Ověření normality rozdílů – vytvoříme novou proměnnou, která bude obsahovat rozdíly objemů hipokampu, a vykreslíme histogram.
  2. Aplikujeme statistický test.
  3. Nulovou hypotézu zamítneme nebo nezamítneme:
     p<0,001 < 0,05 → zamítáme nulovou hypotézu → **Rozdíl v objemu hipokampu u pacientů s AD při vstupu do studie a 2 roky po zahájení studie je statisticky významný.**
- **Poznámka:** Stejné výsledky dostaneme, pokud použijeme jednovýběrový t-test a jako vstupní proměnnou vezmeme proměnnou s rozdílem objemů.

Úkol 2

Zadání: Zjistěte, zda se liší MMSE skóre u kontrolních subjektů (CN) při vstupu do studie a dva roky po zahájení studie (nezapomeňte ověřit předpoklady).

Řešení:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	MMSE	29,19	150	,910	,074
	MMSE_24	28,98	150	1,207	,099

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	MMSE - MMSE_24	,213	1,256	,103	,011	,416	2,080	149	,039

Závěr: U kontrolních subjektů došlo po dvou letech ke statisticky významnému snížení hodnot MMSE skóre.

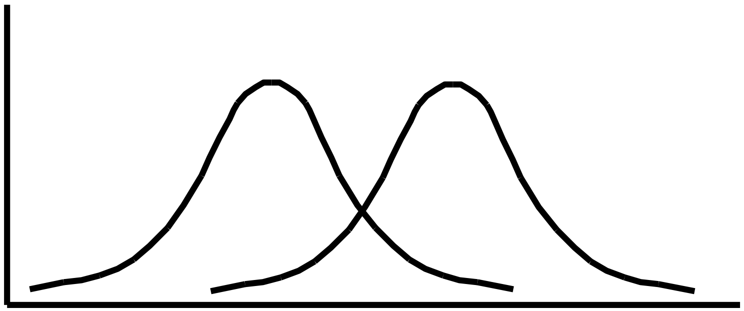
Parametrické a neparametrické testy pro kvantitativní data – přehled

Typ srovnání	Parametrický test	Neparametrický test
1 skupina dat s referenční hodnotou – jednovýběrové testy:	Jednovýběrový t-test, jednovýběrový z-test	Wilcoxonův test
2 skupiny dat párově – párové testy:	Párový t-test	Wilcoxonův test, znaménkový test
2 skupiny dat nepárově – dvouvýběrové testy:	**Dvouvýběrový t-test**	Mannův-Whitneyův test, mediánový test
Více skupin nepárově:	ANOVA	Kruskalův- Wallisův test

4. Dvouvýběrové testy

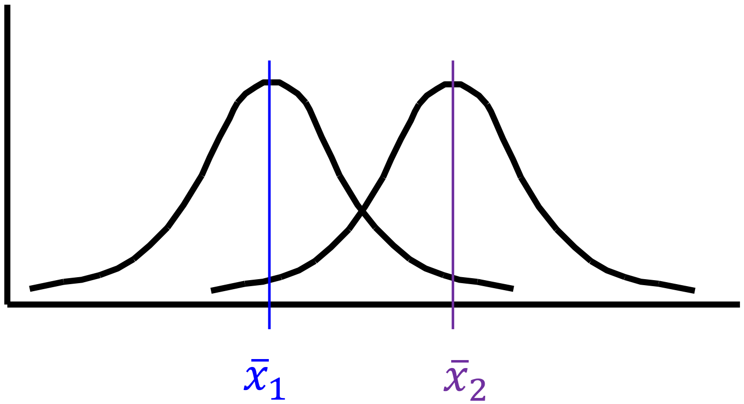
Dvouvýběrové („Two-Sample“) testy

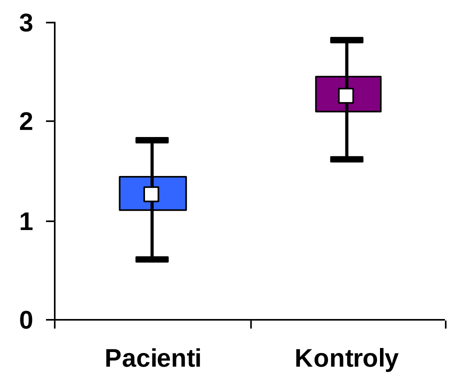
- Srovnávají navzájem dva nezávislé vzorky („two samples“).
- V testu jsou srovnávána dvě rozložení hodnot.
- Otázka položená v testu může být opět vztažena k průměru, rozptylu, podílu hodnot i dalším statistickým parametrům popisujícím vzorek.
- Parametrické dvouvýběrové testy, kterým se budeme věnovat:
  - **dvouvýběrový t-test** (test o rozdílu průměrů dvou nezávislých vzorků)
  - **F-test** (test o shodnosti rozptylů dvou nezávislých vzorků)



Dvouvýběrový t-test

- **Srovnáváme dvě skupiny dat, které jsou na sobě nezávislé – mezi objekty neexistuje vazba.**
- Příklady: srovnání objemu hipokampu u mužů a u žen, srovnání kognitivního výkonu podle dvou kategorií věku.





- Předpoklad: **normalita dat v OBOU skupinách, shodnost (homogenita) rozptylů** v obou skupinách
- Testová statistika: $T = \dfrac{\overline{X}_1 - \overline{X}_2 - c}{s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$, kde s_* je vážená směrodatná odchylka,

  c je konstanta, o kterou se rozdíl průměrů má lišit (většinou rovna 0)

Dvouvýběrový t-test

- **Příklad:** Chceme srovnat, zda se liší objem putamenu podle pohlaví.
- Tzn. hypotézy budou mít tvar: $H_0 : \bar{x}_1 - \bar{x}_2 = 0$ a $H_1 : \bar{x}_1 - \bar{x}_2 \neq 0$
- **Postup:**
  1. Popisná sumarizace objemu putamenu podle pohlaví.
  2. Ověření normality hodnot v OBOU skupinách pomocí histogramu (tzn. vykreslíme histogram zvlášť pro muže a zvlášť pro ženy).
  3. Ověření shodnosti rozptylů – vizuálně pomocí krabicových grafů a pomocí F-testu či Leveneova testu.
  4. Aplikujeme statistický test.
  5. Nulovou hypotézu zamítneme nebo nezamítneme:
     p<0,097 > 0,05 → nezamítáme nulovou hypotézu → Neprokázali jsme rozdíl objemu putamenu podle pohlaví v našem souboru (na hladině významnosti α=0,05.)

Dvouvýběrový t-test

Group Statistics

	Gender_rek	N	Mean	Std. Deviation	Std. Error Mean
Putamen_volume (mm3)	F	351	11175.292	198.4214	10.5910
	M	482	11152.113	199.2849	9.0772

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Putamen_volume (mm3)	Equal variances assumed	.088	$\boxed{.767}$	1.661	831	.097	23.1794	13.9581	-4.2180	50.5767
	Equal variances not assumed			1.662	756.160	.097	23.1794	13.9486	-4.2032	50.5619

Pokud p-hodnota u Leveneova testu ≥ 0,05

Pokud p-hodnota u Leveneova testu < 0,05

V tomto případě p=0,767 > 0,05 → čteme p-hodnotu z 1. řádku.

Úkol 3

Zadání: Zjistěte, zda se liší objem thalamu podle pohlaví (nezapomeňte ověřit předpoklady).

Řešení:

Group Statistics

	Gender_rek	N	Mean	Std. Deviation	Std. Error Mean
Thalamus_volume (mm3)	M	482	12828,02	194,93	8,88
	F	351	12469,89	201,72	10,77

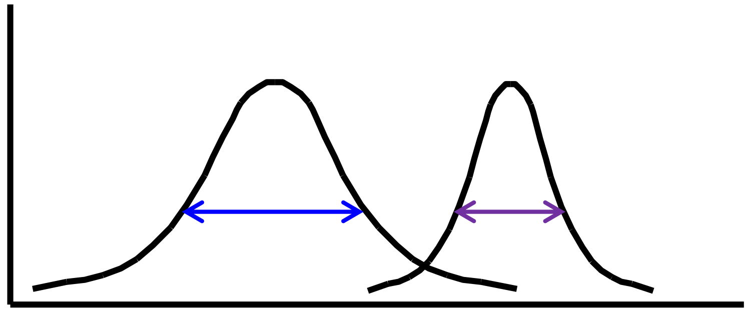
Independent Samples Test

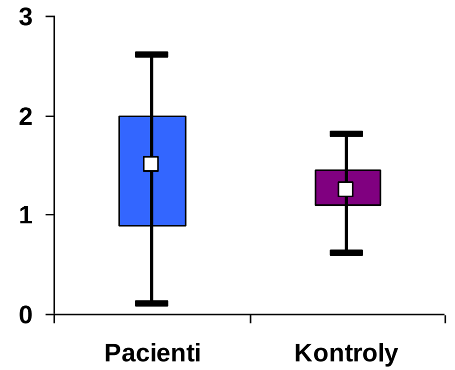
		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Thalamus_volume (mm3)	Equal variances assumed	,053	,819	25,801	831	$\boxed{,000}$	358,135	13,881	330,890	385,380
	Equal variances not assumed			25,662	739,150	,000	358,135	13,956	330,737	385,532

Závěr: Objem thalamu se statisticky významně liší podle pohlaví, přičemž ženy mají statisticky významně menší objem thalamu než muži.

F-test

- **Srovnáváme rozptyly (variabilitu) dvou skupin dat, které jsou na sobě nezávislé** (mezi objekty neexistuje vazba).
- F-test patří mezi dvouvýběrové parametrické testy.
- Příklady: srovnání variability objemu hipokampu u pacientů s AD a kontrol.
- Použití: ověření předpokladu shodnosti (homogenity) rozptylů u dvouvýběrového t-testu.





- Předpoklad: normalita dat v OBOU skupinách.
- Testová statistika: $F = \frac{s_1^2}{s_2^2}$, kde s_1^2 je rozptyl prvního výběru a s_2^2 je rozptyl druhého výběru

F-test

- **Příklad:** Chceme srovnat, zda se liší variabilita objemu putamenu podle pohlaví.
- Tzn. hypotézy budou mít tvar: $H_0 : \sigma_M^2 = \sigma_Z^2$ a $H_1 : \sigma_M^2 \neq \sigma_Z^2$
- **Postup:**
  1. Ověření normality hodnot v OBOU skupinách pomocí histogramu (tzn. vykreslíme histogram zvlášť pro muže a zvlášť pro ženy).
  2. Vykreslení krabicových grafů, které nám napoví, zda máme očekávat shodu nebo neshodu rozptylů.
  3. Aplikujeme statistický test (F-test je součástí dvouvýběrového t-testu v softwaru STATISTICA; v software SPSS můžeme použít Leveneův test).
  4. Nulovou hypotézu zamítneme nebo nezamítneme:
     p=0,934 > 0,05 → nezamítáme nulovou hypotézu → Neprokázali jsme rozdíl ve variabilitě objemu putamenu podle pohlaví v našem souboru (na hladině významnosti α=0,05.)

5. Neparametrické testy

Parametrické a neparametrické testy pro kvantitativní data – přehled

Typ srovnání	Parametrický test	Neparametrický test
1 skupina dat s referenční hodnotou – jednovýběrové testy:	Jednovýběrový t-test, jednovýběrový z-test	**Wilcoxonův test**
2 skupiny dat párově – párové testy:	Párový t-test	**Wilcoxonův test, znaménkový test**
2 skupiny dat nepárově – dvouvýběrové testy:	Dvouvýběrový t-test	**Mannův-Whitneyův test, mediánový test**
Více skupin nepárově:	ANOVA	Kruskalův- Wallisův test

Neparametrické testy

- **Nemají předpoklady** o rozdělení vstupních dat, je tedy možné je použít při asymetrickém rozdělení nebo odlehlých hodnotách.
- Používání neparametrických testů je „bezpečnější“.
- Mají však **menší sílu**, protože dochází k redukci informační hodnoty původních dat z důvodu, že neparametrické testy nevyužívají původní hodnoty, ale nejčastěji pouze jejich pořadí („rank“).
- Menší sílu testu je možné vykompenzovat větší velikostí vzorku.

- Neparametrické testy:
  - **Wilcoxonův test –** jednovýběrový i párový test
  - Znaménkový test – párový test
  - **Mannův-Whitneyův test –** dvouvýběrový test
  - Mediánový test – dvouvýběrový test

Wilcoxonův test

- Neparametrická alternativa jednovýběrového i párového t-testu a z-testu.
- Je testem o mediánu – hypotézy mají tvar: $H_0 : \tilde{x} = c$ a $H_1 : \tilde{x} \neq c$

- Princip Wilcoxonova testu:
  1. Spočítáme diference všech hodnot $x_1, x_2, \dots, x_n$ od *c.*
  2. Podíváme se, jestli je zhruba ½ diferencí kladných a ½ záporných. (To je ekvivalentní s tím, že zhruba polovina hodnot $x_1, x_2, \dots, x_n$ je menších než c a polovina hodnot $x_1, x_2, \dots, x_n$ je větších než c).

- Je zřejmé, že odlehlé hodnoty nebudou v tomto testu problém, protože nehodnotíme velikost diferencí, ale pouze, zda je zhruba ½ z nich kladných a ½ záporných.

Wilcoxonův test jako jednovýběrový test

- **Příklad:** Chceme zjistit, zda se hodnoty MMSE skóre u 197 pacientů s Alzheimerovou chorobou v našem souboru liší od populačního mediánu 27,5.
- Tzn. hypotézy budou mít tvar: $H_0 : \tilde{x} = 27{,}5$ a $H_1 : \tilde{x} \neq 27{,}5$
- **Postup:**
  1. Vykreslíme histogram a spočítáme popisnou statistiku, abychom viděli, že u MMSE skóre u pacientů s AD není splněn předpoklad normálního rozdělení → proto použijeme neparametrický test.
  2. Aplikujeme statistický test.*
  3. Nulovou hypotézu zamítneme nebo nezamítneme:
     p<0,001 < 0,05 → zamítáme nulovou hypotézu → **Medián MMSE skóre u pacientů s AD v našem souboru se statisticky významně liší od populačního mediánu (konkrétně je statisticky významně nižší než populační medián).**

*Software STATISTICA neumožňuje počítat jednovýběrový Wilcoxonův test přímo. Lze to však obejít vytvořením nové proměnné, která ve všech řádcích bude mít hodnotu 27,5, a použitím párového Wilcoxonova testu

Wilcoxonův test jako párový test

- **Příklad:** Chceme srovnat, zda se liší MMSE skóre u pacientů s MCI při vstupu do studie a 2 roky po zahájení studie.
- Tzn. hypotézy budou mít tvar: $H_0 : \tilde{d} = 0$ a $H_1 : \tilde{d} \neq 0$
- **Postup:**
  1. Vykreslení histogramu nové proměnné s rozdíly MMSE skóre, abychom viděli, že u rozdílů není splněn předpoklad normálního rozdělení → proto použijeme neparametrický test.
  2. Aplikujeme statistický test.
  3. Nulovou hypotézu zamítneme nebo nezamítneme:
     p<0,001 < 0,05 → zamítáme nulovou hypotézu → **Rozdíl MMSE skóre u pacientů s MCI při vstupu do studie a 2 roky po zahájení studie je statisticky významný, přičemž došlo ke statisticky významnému poklesu v hodnotách MMSE skóre po 2 letech.**

Úkol 4

Zadání: Zjistěte, zda se liší váha u mužů v našem souboru od populačního mediánu 75 kg.

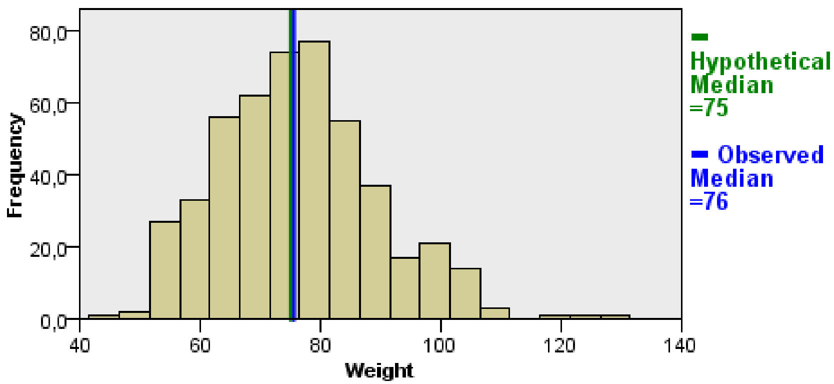
Řešení:

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of Weight equals 75.	One-Sample Wilcoxon Signed Rank Test	295,000	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

One-Sample Wilcoxon Signed Rank Test



Total N	482
Test Statistic	57 939,000
Standard Error	2 926,481
Standardized Test Statistic	1,048
Asymptotic Sig. (2-sided test)	,295

Závěr: Neprokázali jsme na základě našich dat, že by se váha u mužů statisticky významně lišila od populačního mediánu 75 kg.

Mannův-Whitneyův (U) test

- Někdy nazýván jako dvouvýběrový Wilcoxonův test.
- Neparametrická alternativa dvouvýběrového t-testu.
- Testuje se, zda jsou srovnatelné distribuční funkce.
- Hypotézy mají tvar: $H_0 : F(x) = F(y)$ a $H_1 : F(x) \neq F(y)$

- Princip Mannova-Whitneyova testu:
  1. Všechny hodnoty z obou výběrů dohromady (tedy n_1+n_2 hodnot) uspořádáme vzestupně podle velikosti → každé hodnotě přiřadíme pořadí.
  2. Spočítáme součet pořadí hodnot prvního výběru a součet pořadí hodnot druhého výběru.
  3. Na základě těchto dvou součtů vypočteme testové statistiky.

- Je zřejmé, že odlehlé hodnoty nebudou v tomto testu problém, protože pracujeme s pořadími namísto původních hodnot.

Mannův-Whitneyův (U) test

- **Příklad:** Chceme srovnat, zda se liší objem hipokampu podle pohlaví.
- Tzn. hypotézy budou mít tvar: $H_0 : F(x) = F(y)$ a $H_1 : F(x) \neq F(y)$
- **Postup:**
  1. Popisná sumarizace objemu hipokampu podle pohlaví.
  2. Vykreslení histogramů objemu hipokampu u mužů a u žen, abychom viděli, že není splněn předpoklad normálního rozdělení → proto použijeme neparametrický test.
  3. Aplikujeme statistický test.
  4. Nulovou hypotézu zamítneme nebo nezamítneme:
     p<0,001 < 0,05 → zamítáme nulovou hypotézu → **Objem hipokampu je u mužů a u žen statisticky významně odlišný, přičemž u žen je statisticky významně nižší než u mužů.**

Úkol 5

Zadání: Zjistěte, zda se liší MMSE skóre u kontrolních subjektů a pacientů s AD.

Řešení:

Ranks

	Group_3kat	N	Mean Rank	Sum of Ranks
MMSE	CN	230	311,92	71742,50
	AD	197	99,67	19635,50
	Total	427		

Test Statistics[a]

	MMSE
Mann-Whitney U	132,500
Wilcoxon W	19635,500
Z	-17,916
Asymp. Sig. (2-tailed)	,000

a. Grouping Variable: Group_3kat

Závěr: MMSE skóre se statisticky významně liší u skupin subjektů, přičemž pacienti s AD mají statisticky významně nižší hodnoty MMSE skóre než kontrolní subjekty.

Poznámka 1

- Všechny dosud uvedené testy se zabývají hodnocením **spojitých náhodných veličin** (mohou nabývat jakýchkoliv hodnot v určitém rozmezí).
- Příklady: výška, váha, vzdálenost, čas, teplota.

- Uvedené testy lze ale použít i pro hodnocení diskrétních náhodných veličin – ale **musí to být odůvodnitelné** (např. velký počet možných hodnot).
- Příklady: počet krevních buněk, počet hospitalizací, počet krvácivých epizod za rok.

Poznámka 2

- **Parametrické a neparametrické testy nemusí vycházet stejně**. Důvody:
  1. Nesplněné předpoklady parametrického testu.
  2. Malá síla neparametrického testu.

- Jsou-li však splněny předpoklady parametrického testu a je-li dostatek dat, bude to vycházet stejně.

- Měli bychom preferovat parametrické testy, ALE pouze po důkladném ověření jejich předpokladů!

Úkol 6

- **Zadání:** Chceme ověřit, zda se liší objem jednotlivých mozkových struktur podle pohlaví. Vykreslete histogramy a rozmyslete si, jaký test (jaké testy) byste použili.

- **Řešení:** Protože u některých proměnných (např. objem hipokampu) není splněn předpoklad normálního rozdělení, musíme všechny proměnné testovat neparametricky (tzn. Mannovým-Whitneyovým testem), aby p-hodnoty byly srovnatelné. Tzn. nebylo by vhodné mít část testů provedených parametrickým testem (dvouvýběrový t-test) a část testů neparametrickým testem, když se jedná o proměnné se stejným významem (jsou to všechno objemy mozkových struktur).

Poděkování…

Příprava výukových materiálů předmětu „DSAN01 Analýza dat pro Neurovědy “ byla finančně podporována prostředky projektu FRVŠ č. 942/2013 „Inovace materiálů pro interaktivní výuku a samostudium předmětu Analýza dat pro Neurovědy“

FOND ROZVOJE VYSOKÝCH ŠKOL