

Analýza dat pro Neurovědy



RNDr. Eva Janoušová
doc. RNDr. Ladislav Dušek, Dr.

Jaro 2014

Blok 3

Jak a kdy použít parametrické a
neparametrické testy I.

Osnova

1. Dvouvýběrové testy
2. F-test
3. Neparametrické testy

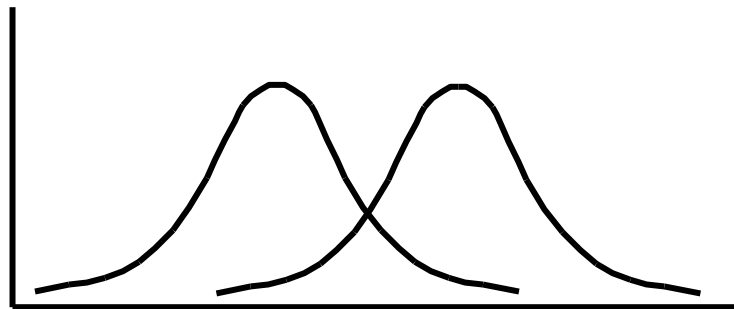
Parametrické a neparametrické testy pro kvantitativní data – přehled

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

1. Dvouvýběrové testy

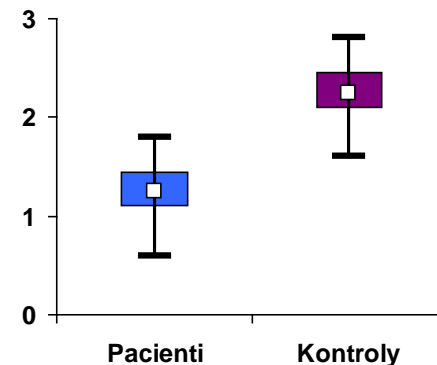
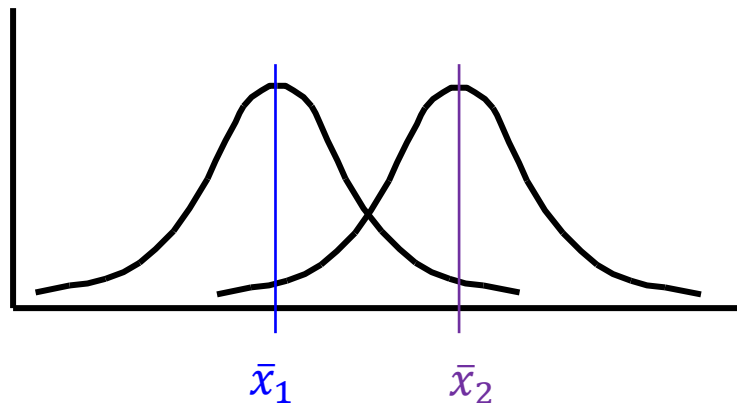
Dvouvýběrové („Two-Sample“) testy

- Srovnávají navzájem dva nezávislé vzorky („two samples“).
- V testu jsou srovnávány 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.



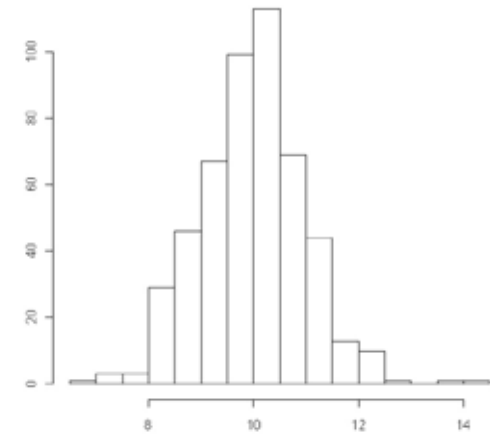
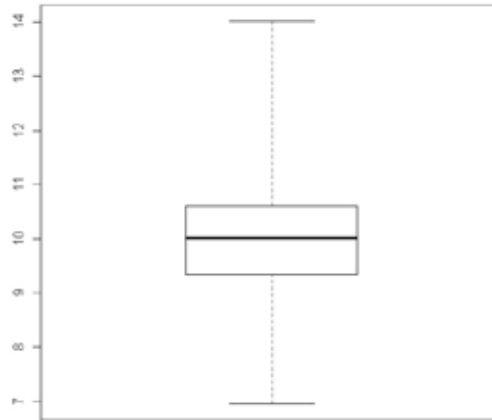
- Předpoklad: **normalita dat** v **OBOU skupinách**, **shodnost (homogenita) rozptylů** v obou skupinách
- Testová statistika: $T = \frac{\bar{x}_1 - \bar{x}_2 - c}{s_* \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$, kde s_* je vážená směrodatná odchylka, c je konstanta, o kterou se rozdíl průměrů má lišit (většinou rovna 0)

Ověření normality dat

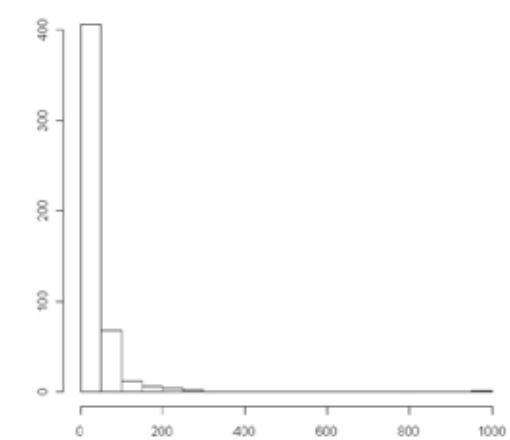
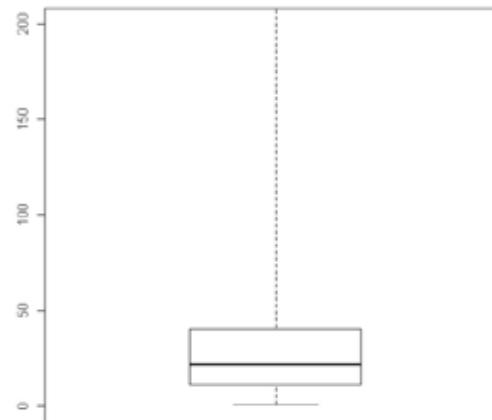
- **Graficky:**
 - histogram
 - krabicový graf (box-plot)
 - Q-Q graf
- **Testy normality:**
 - Shapirův-Wilkův test
 - Kolmogorovův-Smirnovův test
- **Testy nejsou vždy nejlepším nástrojem! Vždy je důležité se podívat i očima!**
- Pokud o sledované veličině prokazatelně víme, že v cílové populaci nabývá normální rozdělení (např. výška lidské postavy), ale v daném souboru normální rozdělení nepotvrdíme, **pak s naším náhodným výběrem není něco v pořádku** – např. není reprezentativní.

Ověření normality graficky – krabicový graf a histogram

- Normální rozdělení

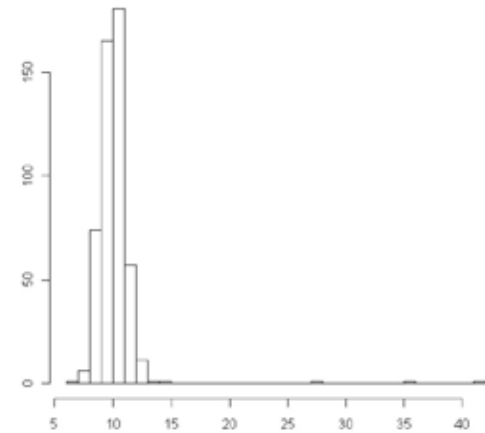
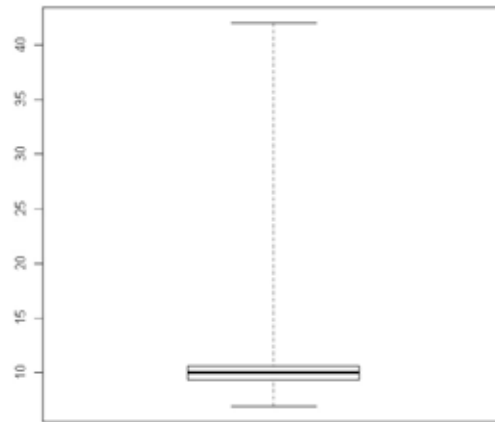


- Log-normální rozdělení

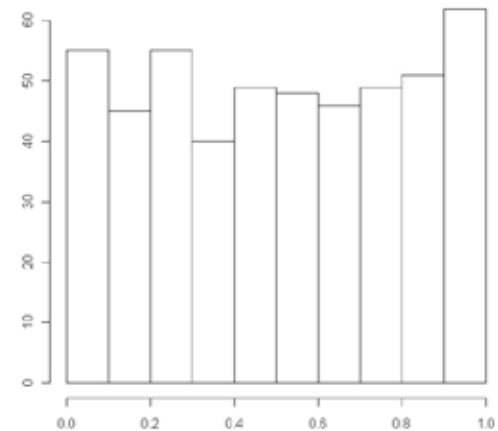
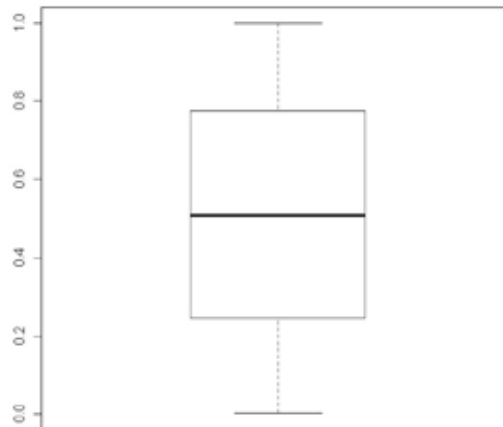


Ověření normality graficky – krabicový graf a histogram

- Normální rozdělení s odlehlými hodnotami

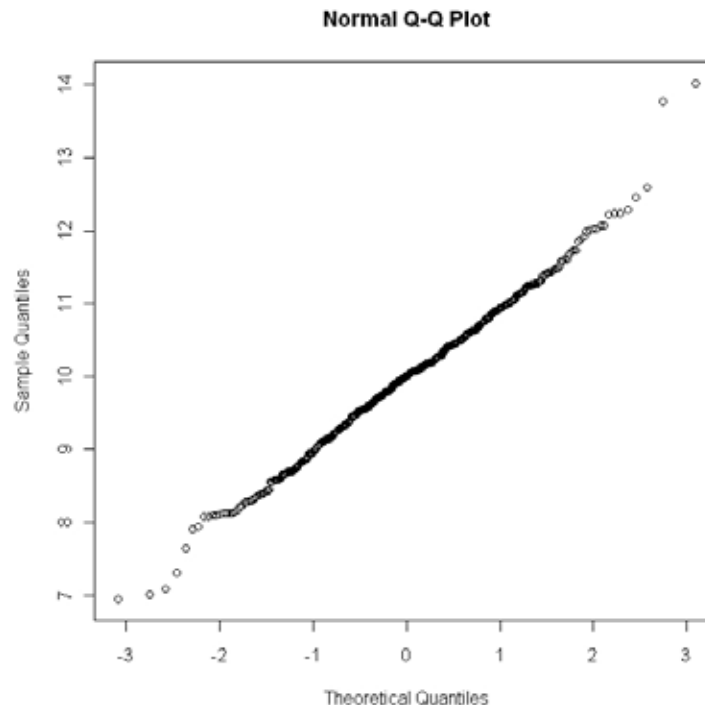


- Rovnoměrně spojité rozdělení



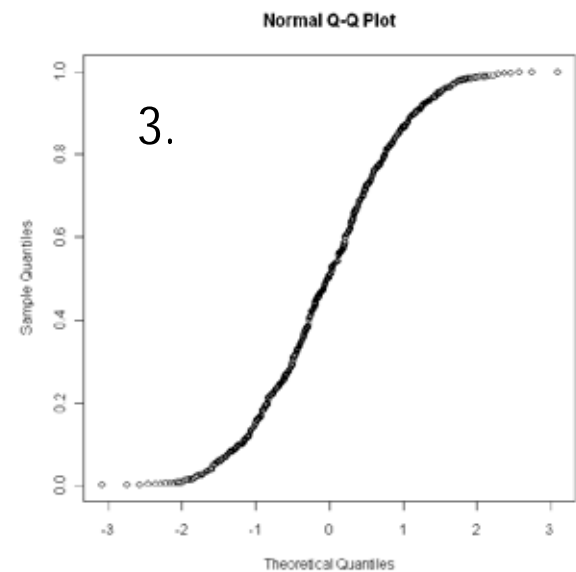
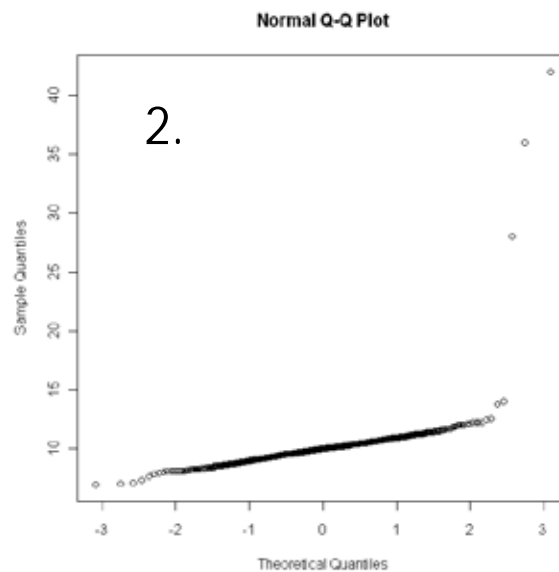
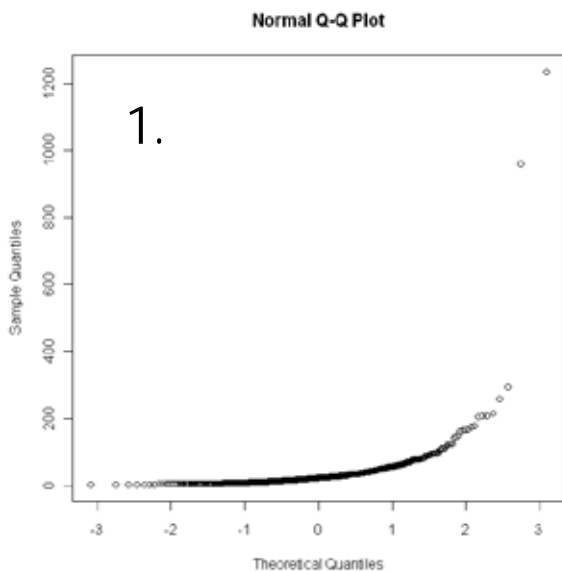
Ověření normality graficky – Q-Q graf

- Q-Q graf proti sobě zobrazuje kvantily pozorovaných hodnot a kvantily teoretického rozdělení pravděpodobnosti (zde normálního rozdělení).
- V případě shody leží všechny body na přímce.
- Normální rozdělení:



Ověření normality graficky – Q-Q graf

1. Log-normální rozdělení
2. Normální rozdělení s odlehlými hodnotami
3. Rovnoměrně spojité rozdělení

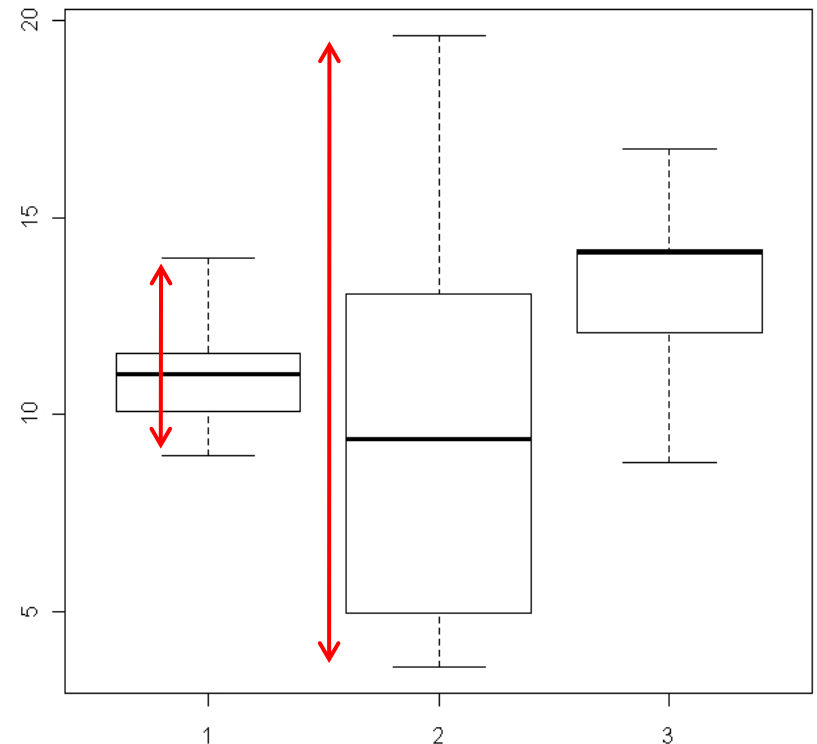


Ověření normality pomocí testů

- **Shapirův-Wilkův test** – v podstatě se jedná o proložení seřazených hodnot regresní přímkou vzhledem k očekávaným hodnotám normálního rozdělení. Má tedy přímý vztah k Q-Q plotu – vyhodnocuje, jak moc se Q-Q plot liší od ideální přímky. **Doporučován pro menší vzorky, může být „moc“ přísný pro velké vzorky.**
- **Kolmogorovův-Smirnovův test** – založen na srovnání výběrové distribuční funkce s teoretickou distribuční funkcí odpovídající normálnímu rozdělení. K-S test hodnotí maximální vzdálenost mezi těmito dvěma distribučními funkcemi. V praxi se používá korekce dle Lillieforse.

Ověření shody (homogeneity) rozptylů

- **Grafické ověření** – krabicový graf, histogram.
- **F-test** (testování shody rozptylů dvou vzorků)
- **Leveneův test** – často používaný (testování shody rozptylů dvou a více vzorků)
- **Bartlettův test**



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.
 4. Aplikujeme statistický test (v softwaru STATISTICA: t-test, independent, by groups).
 5. Nulovou hypotézu zamítneme nebo nezamítneme:
 $p=0,097 > 0,05 \rightarrow$ nezamítáme nulovou hypotézu \rightarrow Neprokázali jsme rozdíl objemu putamenu podle pohlaví (na hladině významnosti $\alpha=0,05$.)

Úkol 3.

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

- **Řešení:**

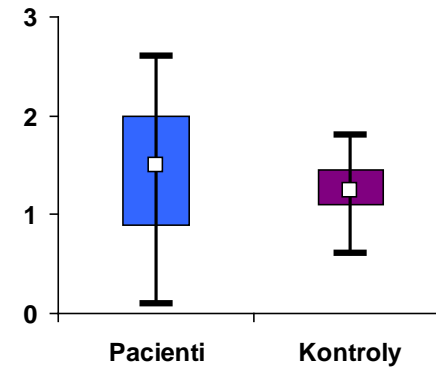
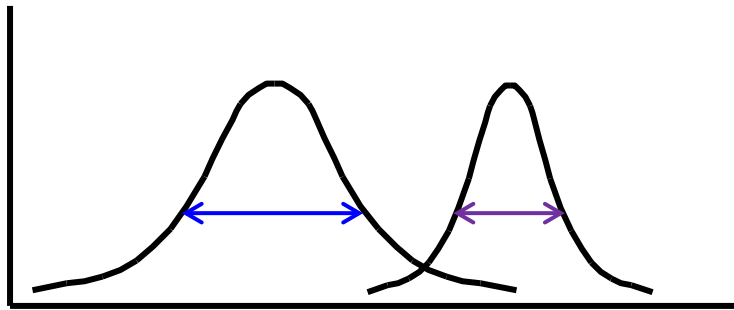
T-tests; Grouping: Gender_rek (Data_neuro_vycistena2)									
Group 1: M									
Group 2: F									
Variable	Mean M	Mean F	t-value	df	p	Valid N M	Valid N F	Std.Dev. M	Std.Dev. F
Thalamus_volume (mm3)	12828,02	12469,89	25,80103	831	0,00	482	351	194,9292	201,7188

F-ratio Variances	p Variances	Levene F(1,df)	df Levene	p Levene
1,070876	0,486954	0,052658	831	0,818557

2. F-test

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 thalamu podle pohlaví.
- Tzn. hypotézy budou mít tvar: $H_0 : s_M^2 = s_Z^2$ a $H_1 : s_M^2 \neq s_Z^2$
- **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 (tedy zvolíme t-test, independent, by groups)).
 4. Nulovou hypotézu zamítneme nebo nezamítneme:
 $p=0,487 > 0,05 \rightarrow$ nezamítáme nulovou hypotézu \rightarrow Neprokázáli jsme rozdíl ve variabilitě objemu thalamu podle pohlaví (na hladině významnosti $\alpha=0,05$.)

3. 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 difference všech hodnot x_1, x_2, \dots, x_n od c .
 2. Podíváme se, jestli je zhruba $\frac{1}{2}$ diferencí kladných a $\frac{1}{2}$ 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 problémem, protože nehodnotíme velikost diferencí, ale pouze, zda je zhruba $\frac{1}{2}$ z nich kladných a $\frac{1}{2}$ záporných.

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. Ověření existence vazby mezi oběma skupinami dat pomocí tečkového grafu.
 2. Vykreslení histogramu nové proměnné s rozdíly MMSE skóre, abychom viděli, že u rozdílů není splněn předpoklad normálního rozdělení → proto použijeme neparametrický test.
 3. Aplikujeme statistický test.
 4. Nulovou hypotézu zamítneme nebo nezamítneme:
 $p < 0,001 < 0,05 \rightarrow$ **zamítáme** nulovou hypotézu → **Rozdíl MMSE skóre u pacientů s MCI při vstupu do studie a 2 roky po zahájení studie je statisticky významný.**

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 (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).
 3. Nulovou hypotézu zamítneme nebo nezamítneme:
 $p < 0,001 < 0,05 \rightarrow$ **zamítáme** nulovou hypotézu → **Medián MMSE skóre u pacientů s AD v našem souboru se statisticky významně liší od populačního mediánu.**

Úkol 4.

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

		Wilcoxon Matched Pairs Test (Data_neuro)			
		Marked tests are significant at p < .05000			
		Include condition: v8="M"			
Pair of Variables	Valid N	T	Z	p-value	
Weight & weight_konst	468	51807.00	1.047366	0.294931	

Mannův-Whitneyův (U) test

- Někdy nazýván jako dvouvýběrový Wilcoxonův test.
- Neparametrická alternativa dvouvýběrového t-testu.
- Testuje se, zda jsou srovnatelné distribuční funkce (tzn. zda mediány obou výběrů jsou srovnatelné).
- 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 problémem, 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:
 $p < 0,001 < 0,05 \rightarrow$ **zamítáme** nulovou hypotézu → **Objem hipokampu je u mužů a u žen statisticky významně odlišný.**

Úkol 5.

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

Mann-Whitney U Test (Data_neuro_vycistena2)									
By variable Group_3kat									
Marked tests are significant at p <,05000									
variable	Rank Sum CN	Rank Sum AD	U	Z	p-value	Z adjusted	p-value	Valid N CN	Valid N AD
MMSE	71742,50	19635,50	132,5000	17,71653	0,00	17,91596	0,00	230	197

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.

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“

