

7. Párové testy



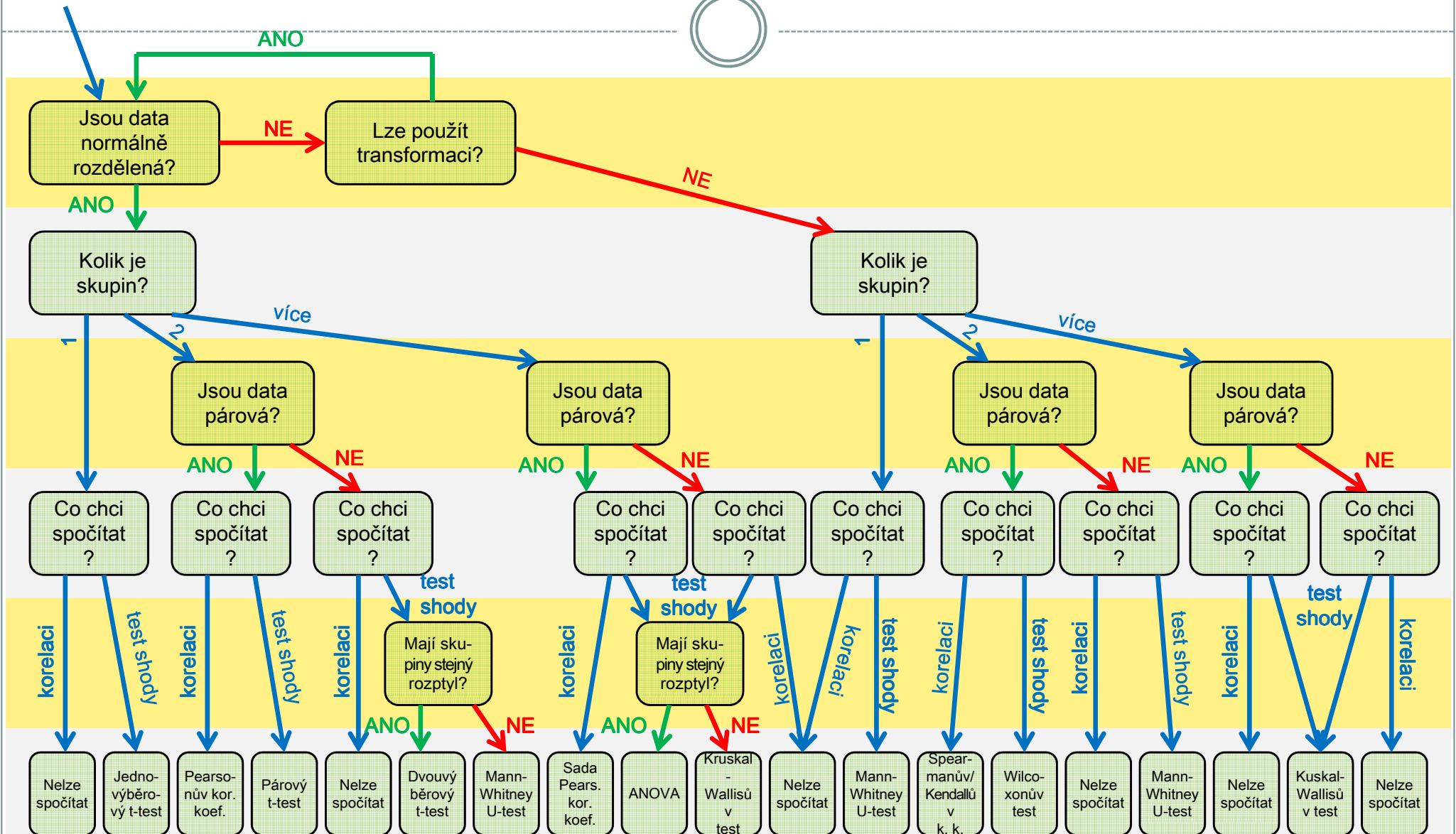
Párový t-test
Wilcoxonův test

Shrnutí statistických testů



Typ srovnání	Nulová hypotéza	Parametrický test	Neparametrický test
1 skupina dat vs. etalon	Střední hodnota je rovna hodnotě etalonu.	jednovýběrový t-test	Mann-Whitneyův test
2 skupiny dat nepárově	Obě skupiny hodnot pochází ze stejného rozdělení.	nepárový t-test	Mann-Whitneyův test
2 skupiny dat párově	Zkoumaný efekt mezi páry hodnot je nulový.	Párový t-test	Wilcoxonův test; znaménkový test
shoda rozdělení	rozdělení dat ve skupině odpovídá teoretickému (vybranému) rozdělení.	Shapiro-Wilkův test; Kolmogorovův-Smirnovův test; Lilieforsův test	χ^2 test, test dobré shody
homoskedasticita (shoda rozptylů)	rozptyl obou (všech) skupin je shodný.	F-test, Levenův test	Levenův test, Brown-Forsythův test
více skupin nepárově	Zkoumaný efekt mezi skupinami hodnot je nulový.	ANOVA	Kruskal-Wallisův test
korelace	Neexistuje (příčinná, důsledková) vazba mezi skupinami hodnot.	Pearsonův koeficient	Spearmanův koeficient; Kendallův koeficient

Shrnutí statistických testů



F test



- Parametrický test sloužící k rozhodnutí, zda mají dva nebo více vzorků stejný rozptyl, někdy nazýván Fisherův test.
- H_0 : rozptyl je stejný.
 H_A : rozptyl se liší.
- Testová statistika:

$$\frac{\frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (X_{2i} - \bar{X}_2)^2}{\frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_{1i} - \bar{X}_1)^2}$$

n_1 je počet hodnot v 1. skupině

n_2 je počet hodnot ve 2. skupině

Levenův test



- Neparametrický test sloužící k rozhodnutí, zda mají dva nebo více vzorků stejný rozptyl.
- H_0 : rozptyl je stejný.
 H_A : rozptyl se liší.
- Testová statistika:

$$W = \frac{(N - k) \sum_{i=1}^k N_i (Z_i - Z)^2}{(k - 1) \sum_{i=1}^k \sum_{j=1}^{N_i} N_i (Z_{i,j} - Z_i)^2}$$

N je celkový počet hodnot

N_i je počet hodnot v i -té skupině

k je počet skupin

\bar{x}_i je průměr hodnot i -té skupiny (resp. medián)

$Z_{i,j} = |x_{i,j} - \bar{x}_i|$

Z_i je průměr $Z_{i,j}$

Z je průměr všech $Z_{i,j}$

Párové dvouvýběrové testy – předpoklady



- Skupiny dat jsou spojeny přes objekt měření, příkladem může být měření parametrů pacienta před léčbou a po léčbě (nemusí jít přímo o stejný objekt, dalším příkladem mohou být např. krysy ze stejné linie).
- Oba soubory musí mít shodný počet hodnot, protože všechna měření v jednom souboru musí být spárována s měřením v druhém souboru. Při vlastním výpočtu se potom počítá se změnou hodnot (diferencí) subjektů v obou souborech.
- Před párovým testem je vhodné ověřit si zda existuje vazba mezi oběma skupinami – vynesení do grafu, korelace.

Existuje několik možných designů experimentu, stručně lze sumarizovat:

1. pokus je párový a jako párový se projeví
2. párové provedení pokusu – párově se neprojeví
 - možná párovost není
 - špatně provedený pokus – malé n , velká variabilita, špatný výběr jedinců
3. čekali jsme nezávislé a jsou
4. čekali jsem nezávislé a nejsou
 - vazba
 - náhoda

Párový dvouvýběrový t-test



- Tento test nemá žádné předpoklady o rozložení vstupních dat, protože je počítán až na základě jejich diferencí.
- Tyto difference by měly být normálně rozloženy a otázkou v párovém t-testu je, zda se průměrná hodnota diferencí rovná nějakému číslu, typicky jde o srovnání s nulou jako důkaz neexistence změny mezi oběma spárovanými skupinami.
- V podstatě jde o one sample t-test, kde místo rozdílu průměru vzorku a cílové populace je uveden průměr diferencí a srovnávané číslo (0 v případě otázky, zda není rozdíl mezi vzorky).

- Pro srovnání s 0 (testovou statistikou je t rozložení):
$$t = \frac{\bar{D}}{s} \sqrt{n} \quad v = n - 1$$

- Někdy je obtížné rozhodnout, zda jde nebo nejde o párové uspořádání, párový test by měl být použit pouze v případě, že můžeme potvrdit vazbu (korelace, vynesení do grafu), jedním z důvodů proč toto ověřovat je fakt, že v případě párového t-testu není nutné brát ohled na variabilitu původních dvou souborů, tento předpoklad však platí pouze v případě vazby mezi proměnnými. Výpočet obou typů testů se vlastně liší v použité s, jednou jde o s diferencí, v druhém případě o složený odhad rozptylu obou souborů.

- Zda je párové uspořádání efektivnější lze určit na základě:

- Síly vazby
- Je-li s_D výrazně menší než $s_{x_1-x_2}$

- Závislost je možné rozepsat pomocí vzorce:
$$s_D^2 \cong \sigma_{x_1}^2 + \sigma_{x_2}^2 - 2Cov(x_1; x_2)$$

- v případě $Cov=0$, tedy v případě neexistence vazby pak s_D^2 odpovídá součtu původních rozptylů, tedy přibližně $S_{x_1-x_2}$.

Párový dvouvýběrový t-test – příklad

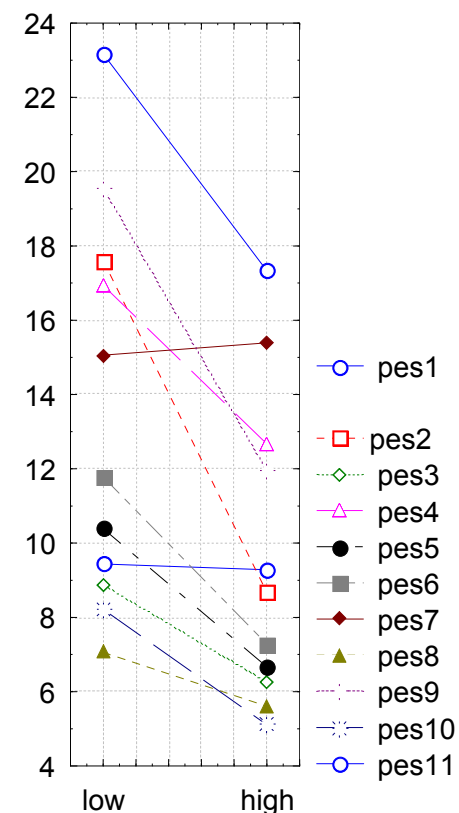


Byl prováděn pokus s dietou 11 diabetických psů, každý pes byl vystaven dvěma dietám s odlišným typem sacharidů (snadno vstřebatelné X pozvolna se rozkládající na glukózu), hodnoty krevní glukózy v průběhu jednotlivých diet mají být srovnány pro zjištění vlivu diety na hladinu krevní glukózy. Protože každý pes absolvoval obě diety, jde o párové uspořádání, kdy výsledky hodnoty v obou pokusech jsou spojeny přes pokusné zvíře.

- 1. Nulová hypotéza zní, že skutečný průměrný rozdíl mezi oběma dietami je 0, alternativní hypotéza zní, že to není 0.**
- 2. Pro každého psa je spočítán rozdíl mezi jeho hladinou glukózy při obou dietách a měly by být ověřeny předpoklady pro one sample t-test – tedy alespoň přibližně normální rozložení.**
- 3. Je spočítána testová charakteristika, výpočet vlastně probíhá jako one-sample t-test, kde je zjišťována významnost průměru diferencí obou souborů jako rozdíl mezi touto hodnotou a nulou (nula je hodnota, kterou by průměrná diference měla nabývat, pokud platí nulová hypotéza). $T=4.37$ s 10 stupni volnosti, skutečná hodnota $p=0,0014$ a tedy na hladině $p=0,05$ můžeme nulovou hypotézu zamítnout**

$$t = \frac{\text{rozdíl}_\text{průměru}_\text{vzorku}_\text{a}_\text{populace}}{SE(\text{průměru})} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{\bar{x} - \mu}{s} \sqrt{n}$$

- 4. Závěrem můžeme říci, že nulová hypotéza neexistence rozdílu mezi oběma dietami byla zamítnuta, což znamená, že high-fibre dieta má významný vliv na snížení hladiny krevní glukózy.**



Neparametrická obdoba párového t-testu



Wilcoxon test

- Jsou vytvořeny difference mezi soubory, je vytvořeno jejich pořadí bez ohledu na znaménko a poté je sečteno pořadí kladných a pořadí záporných rozdílů. Menší z těchto dvou hodnot je srovnána s kritickou hodnotou testu a pokud je menší než kritická hodnota testu, pak zamítáme hypotézu shody obou souborů hodnot. Pro test existuje aproximace na normální rozložení, ale pouze pro velká $n > 25$.

$$t = \frac{\text{Menší_suma_diferencí} - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

Před zásahem	Po zásahu	Změna	Absolutní pořadí
6	2	4	10
2,5	3	-0,5	1,5
6,3	5	1,3	6
8,1	9	-0,9	5
1,5	2	-0,5	1,5
3,4	4	-0,6	3
2,5	1	1,5	8
1,11	2	0,89	4
2,6	4	-1,4	7
1	3	-2	9

Wilcoxonův test – příklad I

člověk	A	B	diference	pořadí
1	142	138	4	4,5
2	140	136	4	4,5
3	144	147	-3	3
4	144	139	5	7
5	142	143	-1	1
6	146	141	5	7
7	149	143	6	9,5
8	150	145	5	7
9	142	136	6	9,5
10	148	146	2	2

A.....parametr krve před podáním léku

B.....parametr krve po podání léku

W_+ Σ pořadí kladných rozdílů = 51

W_- = 4

**$W = \min(W_+; W_-) = 4$
počet párů = $n = 10$**

Pokud je W menší než kritická hodnota testu, pak zamítáme hypotézu shody distribučních funkcí obou skupin.

Wilcoxonův test – příklad II



Byla testována nová dieta pro laboratorní krysy, při pokusu byl zjišťován její vliv na různých liniích krys, bylo proto zvoleno párové uspořádání kdy krysy v obou dietách jsou spojeny přes svoji linii, tj. na začátku byly dvojice krys stejné linie, jedna z nich byla náhodně přiřazena k dietě, druhá z dvojice pak do druhé diety.

1. nulová hypotéza je, že váha krys není ovlivněna použitou dietou, alternativní, že ovlivnění dietou existuje
2. spočítáme difference – tyto difference jsou nenormální a proto je vhodné využít neparametrický test
3. Spočítáme sumu pořadí kladných a záporných diferencí, zde je menší suma záporných diferencí – 31
4. výsledkem výpočtu je $p > 0,05$ a tedy nemáme dostatečné důkazy pro zamítnutí nulové hypotézy, nelze říci, že by nová dieta byla efektivnější než stará
5. pro doplnění výsledků je vhodné zjistit také skutečnou velikost rozdílu hmotností ve skupinách, např. ve formě mediánu