9. Parametrické testy

Kolomogorovův-Smirnovův (Lilieforsův) test

Shapiro-Wilkův test

Jednovýběrový a dvouvýběrový t-test

Párový t-test

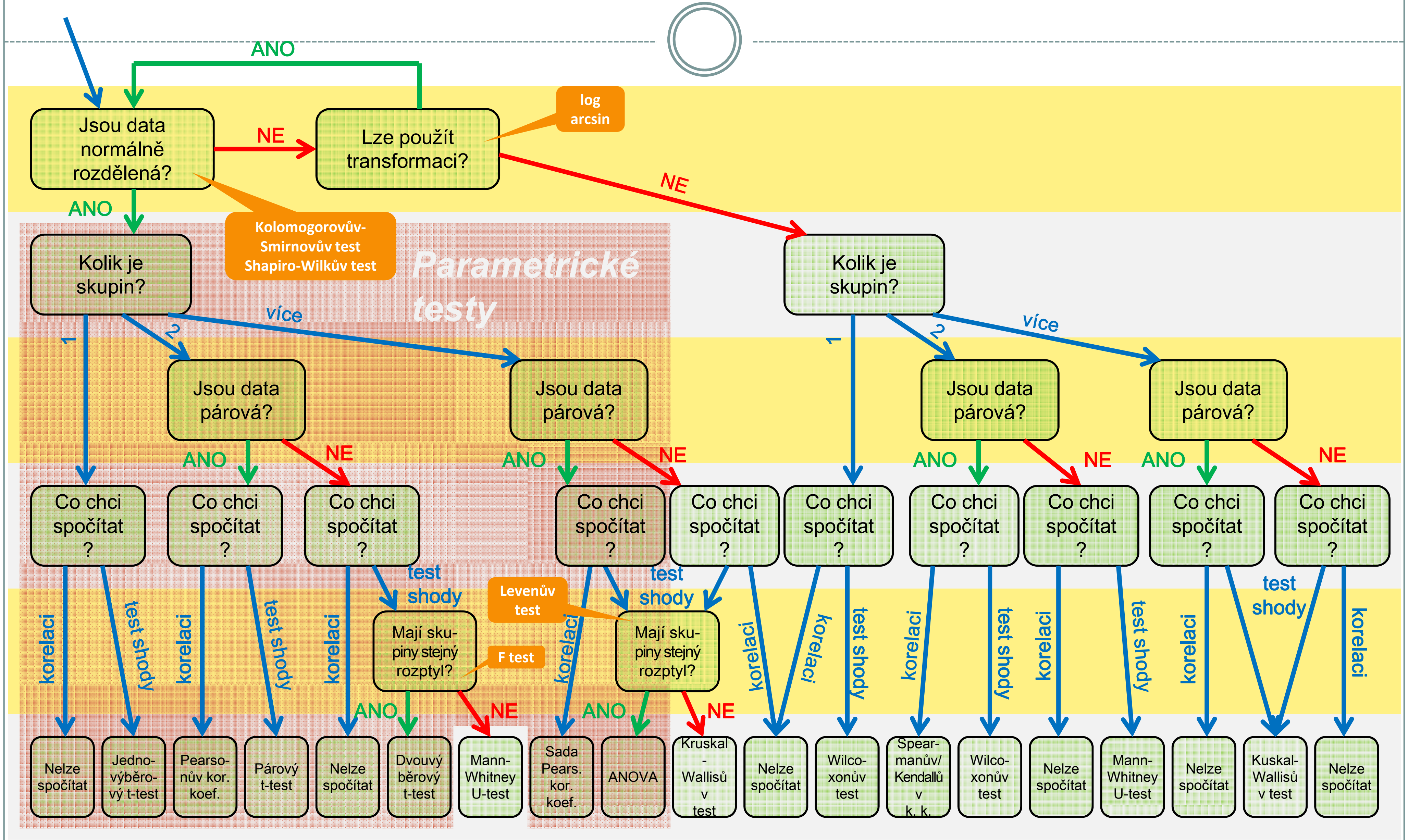
F-test a Levenův test

Vytvořil Institut biostatistiky a analýz, Masarykova univerzita
J. Jarkovský, L. Dušek, J. Kalina

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	Wilcoxonův test; znaménkový 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-Wilksův test; Kolmogorovův-Smirnovův test	$\chi 2$ test, test dobré shody
homoskedasticita (shoda rozptylů)	**rozptyl obou (všech) skupin je shodný.**	F test; Levenův test	Bartlettův test
více skupin nepárově	**Zkoumaný efekt mezi skupinami hodnot je nulový.**	ANOVA (analýza rozptylu)	Kruskal- Wallisův test
více skupin závisle	**Zkoumaný efekt mezi závislými hodnotami je nulový.**	ANOVA opakovaných měření	Friedmannův test
korelace	**Neexistuje (příčinná, důsledková) vazba mezi skupinami hodnot.**	Pearsonův koeficient	Spearmanův koeficient; Kendallův koeficient

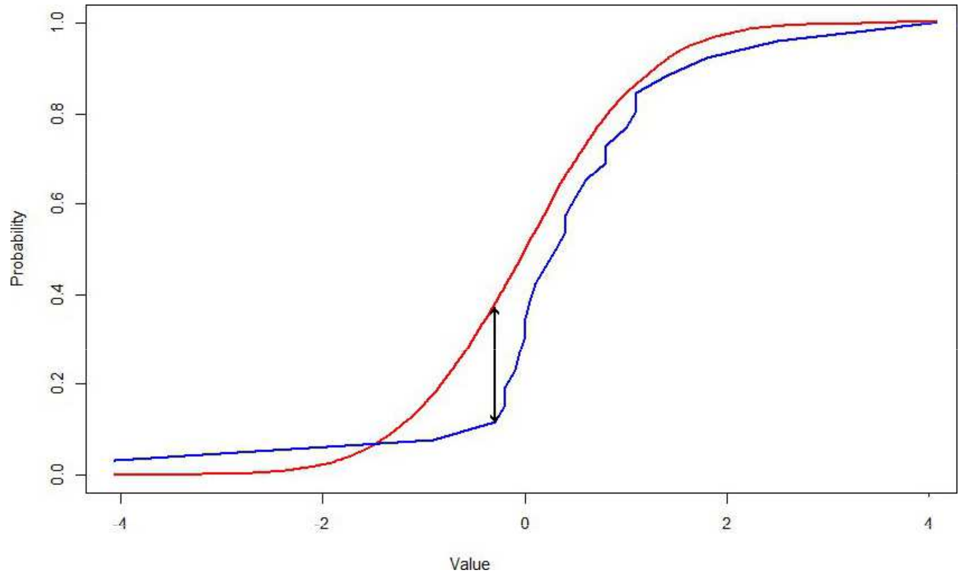
Shrnutí statistických testů



Testy normality

Kolgomorovův-Smirnovův test

Tento test je často používán, dokáže dobře najít odlehlé hodnoty, ale počítá spíše se symetrií hodnot než přímo s normalitou. Jde o neparametrický test pro srovnání rozdílu dvou rozložení. Je založen na zjištění rozdílu mezi reálným kumulativním rozložením (vzorek) a teoretickým kumulativním rozložením. Měl by být počítán pouze v případě, že známe průměr a směrodatnou odchylku hypotetického rozložení, pokud tyto hodnoty neznáme, měla by být použita jeho modifikace – Lilieforsův test.



Shapiro-Wilkův test

Jde o neparametrický test použitelný i při velmi malých n (10) s dobrou sílou testu, zvláště ve srovnání s alternativními typy testů, je zaměřen na testování symetrie.

t-Test

Tři varianty **parametrického** t-testu: jednovýběrový
dvouvýběrový
párový

Předpoklad: **Měřená náhodná veličina má normální rozdělení.**

↓

Výběrový průměr má normální rozdělení se stejnou střední hodnotou, skutečný rozptyl ovšem neznáme.

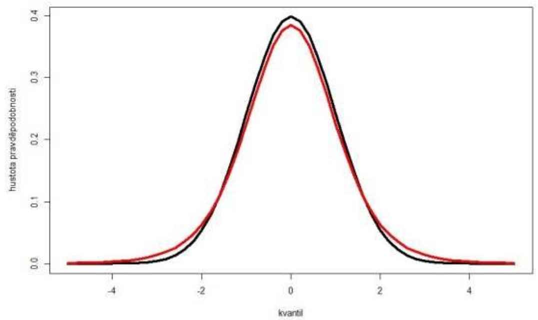
↓

Rozdíl výběrového průměru od skutečné střední hodnoty má také normální rozdělení.

↓

Při využití výběrového rozptylu má rozdíl t-rozdělení.

Kvantil	-3	-2	-1	0	1	2	3
Norm(0,1)	0,00	0,05	0,24	0,40	0,24	0,05	0,00
$t_7(0,1)$	0,01	0,06	0,23	0,38	0,23	0,06	0,01



t-Test

Princip: podle určené hladiny pravděpodobnosti se stanoví maximální přípustná velikost rozdílu výběrového průměru a skutečné střední hodnoty. Testuje se velikost rozdílu.

Postup: Výpočet normalizovaného rozdílu a jeho porovnání s tabelovanou hodnotou (jednostranná a dvoustranná varianta):

$$t = \frac{\bar{x} - \mu}{s}\sqrt{n}$$

H_0	H_A	Testová statistika	Interval spolehlivosti
$\bar{x} \leq \mu$	$\bar{x} > \mu$	t	t > t
$\bar{x} \geq \mu$	$\bar{x} < \mu$	t	t < t
$\bar{x} = \mu$	$\bar{x} \neq \mu$	t	|t| > t

t-Test

Koncentrace antibiotika v cílovém orgánu

Při 1000 měřeních antibiotika byla zjištěna v cílovém orgánu průměrná koncentrace 202,5 jednotek a směrodatná odchylka 44 jednotek.

Požadovaná koncentrace antibiotika je 200 jednotek.

1) ***Je daný rozdíl 2,5 významný vzhledem k variabilitě znaku na hladině významnosti 5 %?***
2) ***Jaká je skutečná hladina významnosti?***

$$t = \frac{\bar{x} - \mu}{s}\sqrt{n} = \frac{2,5}{44}\sqrt{1000} = 1,797$$

F test

- Parametrický test sloužící k rozhodnutí, zda mají dva nebo více vzorků stejný rozptyl, někdy nazýván Fisherův test.
- H0: rozptyl je stejný.
  HA: rozptyl se liší.
- Testová statistika:
- F=

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

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.
- H0: rozptyl je stejný.
  HA: 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 $Z_{i,j}$

Párový t-test – 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 t-test

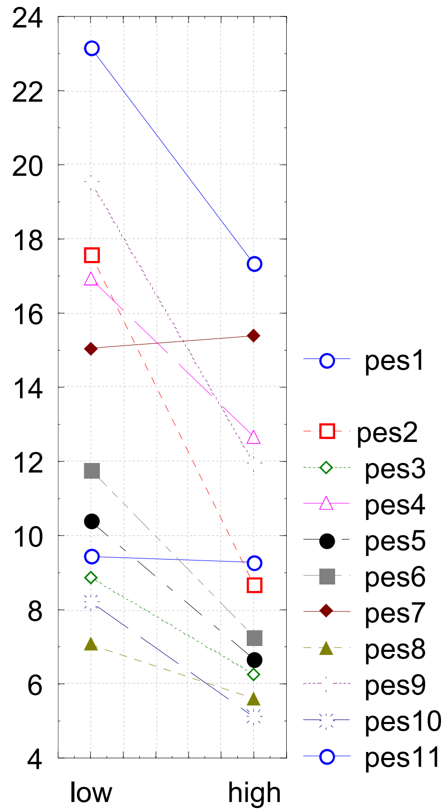
- Tento test nemá žádné předpoklady o rozložení vstupních dat, protože je počítán až na základě jejich diferencí.
- Tyto diference by měly být normálně rozloženy a otázkou v párovém t-testu je, zda se průměrná hodnota diferencí rovná nějakému číslu, typicky jde o srovnání s nulou jako důkaz neexistence 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): $t = \frac{\bar{D}}{s}\sqrt{n} \qquad \upsilon = n-1$
- 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_{x1-x2}
- 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_{x1-x2}.

Párový 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 zamítnou***

$$t = \frac{rozdíl\_průměru\_vzorku\_a\_populace}{SE(průměru)} = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} = \frac{\bar{x}-\mu}{s}\sqrt{n}$$



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.***