Testy statistických hypotéz

- **Statistická hypotéza** je jakýkoliv předpoklad o rozdělení pravděpodobnosti jedné nebo několika náhodných veličin.

- Na základě **náhodného výběru**, který je **reprezentativním vzorkem základního souboru** (který neznáme, k němuž se ale daná statistická hypotéza váže), potřebujeme ověřit, zda základní soubor je či není ve shodě s testovanou hypotézou. (Meloun & Militký, 2013)

- Statistický test každému náhodnému výběru přiřadí jedno ze dvou rozhodnutí
  - **zamítnutí hypotézy**

    nebo
  - **nezamítnutí hypotézy**

Meloun, M., Militký, J. (2013). *Kompendium statistického zpracování dat*. Praha: Karolinum.

Chyba prvního a druhého druhu

Vždy existuje **riziko**, že naše tvrzení nebude v souladu se skutečností, tedy **že** buď

- zamítneme hypotézu, která ve skutečnosti platí
  takovou chybu označme α (tzv. **chyba 1. druhu**

  nebo

- nezamítneme hypotézu, která ve skutečnosti neplatí
  takovou chybu označme β (tzv. **chyba 2. druhu**).

Zmenšení α vede za jinak nezměněných podmínek ke zvětšení β a naopak.

(Meloun & Militký, 2013)

- Hodnotu α volíme nejčastěji **0,05**; **0,01**; **0,005**; **0,001**.
- Když hypotézu zamítneme, znamená to, že téměř jistě (s pravděpodobností $1 - \alpha$) neplatí.

Meloun, M., Militký, J. (2013). *Kompendium statistického zpracování dat*. Praha: Karolinum.

Testy statistických hypotéz – postup

Rozlišujeme **nulovou hypotézu H_0** a **alternativní hypotézu H_1**.

- o nulové hypotéze máme rozhodnout, zda ji zamítneme nebo nezamítneme (alternativní hypotézu přijmeme v případě, když zamítneme nulovou hypotézu)
- nulová hypotéza vždy předpokládá, že pozorovaný jev je pouze dílem náhody (tzv. **testování "na nulu"**)

K ověření hypotézy slouží **výběrová charakteristika** (též **statistika**), která má při platnosti **H_0** známé rozdělení pravděpodobnosti →

najdeme oblast hodnot, které se za předpokladu testované hypotézy vyskytnou jen s malou pravděpodobností – tzv. **kritický obor**.

Jestliže hodnota testovaného parametru padne do kritického oboru, nulovou hypotézu zamítneme a předpokládáme, že platí alternativní hypotéza.

(Meloun & Militký, 2013)

Meloun, M., Militký, J. (2013). *Kompendium statistického zpracování dat*. Praha: Karolinum.

Testy jednostranné a oboustranné

Mohou nastat tři případy formulace nulové a alternativní hypotézy:

1. H_0: $t \leq t_0$, H_1: $t > t_0$,
2. H_0: $t \geq t_0$, H_1: $t < t_0$,
3. H_0: $t = t_0$, H_1: $t \neq t_0$.

Pro jednoduchost budeme používat jen **oboustranné testy** (3.).

Shrnutí postupu při testování:

1. Formulace nulové, resp. alternativní hypotézy (H_0, resp. H_1).
2. Volba hladiny významnosti α.
3. Nalezení vhodné výběrové charakteristiky, určení kritického oboru.
4. Výpočet testové charakteristikyna základě náhodného výběru.
5. Rozhodnutí.

(Meloun & Militký, 2013)

Meloun, M., Militký, J. (2013). *Kompendium statistického zpracování dat*. Praha: Karolinum.

Test střední hodnoty normálního rozdělení

H_0: **Střední hodnota** souboru s normálním rozdělením, ze kterého byl proveden výběr, **je μ_0**.

H_1: Střední hodnota souboru s normálním rozdělením, ze kterého byl proveden výběr, není μ_0.

Testovací kritérium:

$$\boxed{t = \frac{\left|\hat{\mu} - \mu_0\right|}{\widehat{\sigma_{\hat{\mu}}}}}$$

$\hat{\mu}$ a $\widehat{\sigma_{\hat{\mu}}}$ ...výběrová střední hodnota a její výběr. směrodatná odchylka

Kritický obor: $|t| > t_{1-\alpha}(N\text{-}1)$

$t_{1-\alpha}(N\text{-}1)$ ... kvantily Studentova rozdělení s N-1 stupni volnosti pro zvolenou hladinu významnosti α, které najdeme ve statistických tabulkách nebo vypočítáme pomocí funkce

(Lepš & Šmilauer, 2016)

=T.INV.2T(α,N-1)

Lepš, J., Šmilauer, P. (2016). *Biostatistika*. České Budějovice: EPISTEME.

Test rovnosti dvou středních hodnot normálního rozdělení

H_0: **Střední hodnoty dvou souborů** s normálním rozdělením, ze kterých byl proveden výběr, **se rovnají** $(\widehat{\mu_1} = \widehat{\mu_2})$.

H_1: Střední hodnoty dvou souborů s normálním rozdělením, ze kterých byl proveden výběr, se nerovnají $(\widehat{\mu_1} \neq \widehat{\mu_2})$.

Testovací kritérium:

$$t = \frac{|\widehat{\mu_1} - \widehat{\mu_2}|}{\sqrt{\frac{\widehat{\sigma_1}^2}{N_1} + \frac{\widehat{\sigma_2}^2}{N_2}}}$$

$\widehat{\mu_1}$ a $\widehat{\sigma_1}$ resp. $\widehat{\mu_2}$ a $\widehat{\sigma_2}$ ... výběr. střední hodnota a výběr. směrodatná odchylka jednoho měření 1. resp. 2. souboru

Kritický obor: $|t| > t_{1-\alpha}(N_1+N_2-2)$

$t_{1-\alpha}(N_1+N_2-2)$ … kvantily Studentova rozdělení s N_1+N_2-2 stupni volnosti pro zvolenou hladinu významnosti α,

=T.INV.2T(α,N_1+N_2-2)

(Lepš & Šmilauer, 2016)

Lepš, J., Šmilauer, P. (2016). *Biostatistika*. České Budějovice: EPISTEME.

Bonferonniho korekce

Test rovnosti středních hodnot se používá pro srovnání **dvou** středních hodnot.

Ale co v případě, že je hodnot třeba 5?

- Srovnáme-li každé dvě, tj. 10 testů rovnosti dvou středních hodnot.
- Je-li hladina významnosti p = 0,05, máme výraznou šanci, že některé rozdíly vyjdou falešně významné.

Problém řeší **Bonferonniho korekce** - za významné nepovažujeme rozdíly, pro které $t > t_{1-\alpha}(N_1+N_2-2)$, ale rozdíly, pro které $t > t_{1-\alpha/n}(N_1+N_2-2)$, kde ***n*** je počet srovnávaných středních hodnot.

(Lepš & Šmilauer, 2016)

Jedná se spíš o nouzové řešení, lepší je použít ANOVA.

Lepš, J., Šmilauer, P. (2016). *Biostatistika*. České Budějovice: EPISTEME.

Grubbsův test odlehlých hodnot pro normální rozdělení

Jako míra odlehlosti hodnoty slouží její vzdálenost testované hodnoty od aritmetického průměru výběru dat, vztažená ke směrodatné odchylce (NE výběrové, ale směrodatné odchylce celého uvažovaného souboru).

Testovací kritérium:

$$\boxed{T = \frac{|x_i - \hat{\mu}|}{S}}, \text{ kde } S = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{\mu})^2}$$

Je-li testovací kritérium T větší než kritická hodnota $T_{N,\alpha}$, vyloučíme testovanou hodnotu ze souboru.

(Budíková et al.,2010)

Kritické hodnoty Grubbsova T-rozdělení (pro α = 0,05 a 0,01)

N	3	4	5	7	10	15	20	30	50	70	100	200
$T_{N,0,05}$	1,15	1,48	1,72	2,02	2,29	2,55	2,71	2,91	3,13	3,26	3,38	3,61
$T_{N,0,01}$	1,15	1,49	1,76	2,14	2,48	2,81	3,00	3,24	3,48	3,62	3,75	3,98

Budíková, M., Králová, M., Maroš, B. (2010). *Průvodce základními statistickými metodami.* Praha: GRADA

Test korelačního koeficientu

H_0: korelační koeficient je nulový

(lineární závislost mezi x a y neexistuje)

H_1: korelační koeficient je nenulový

(závislost mezi x a y existuje).

Testovací kritérium:

$$t = \frac{|r|}{\sqrt{1-r^2}}\sqrt{N-2}$$

Kritický obor: $|t| > t_{1-\alpha}(N-2)$

$t_{1-\alpha}(N-2)$ jsou kvantily Studentova rozdělení s N-2 stupni volnosti pro zvolenou hladinu významnosti α, které najdeme ve statistických tabulkách nebo vypočítáme pomocí funkce

(Meloun & Militký, 2013)

=T.INV.2T(α,N-2)

Meloun, M., Militký, J. (2013). *Kompendium statistického zpracování dat*. Praha: Karolinum.

Test rovnosti dvou středních hodnot pro párové hodnoty

Předpokládejme, že testujeme účinky preparátu na zlepšení paměti. Pokusné osoby nejdříve absolvovaly test paměti, pak dostaly preparát a absolvovaly test paměti ještě jednou.

Pro testování bychom mohli použít **test rovnosti dvou středních hodnot** z minulé kapitoly. Lze však očekávat, že výsledky testu budou mít velkou variabilitu, která může překrýt případné malé zlepšení.

Nabízí se proto možnost

1. spočítat pro každou osobu **rozdíl** obou testů paměti,
2. **testovat, zda je střední hodnota** rozdílů mezi testy **nulová** nebo různá od nuly (**test střední hodnoty** pro $\mu_0 = 0$).

(Lepš & Šmilauer, 2016)

Lepš, J., Šmilauer, P. (2016). *Biostatistika*. České Budějovice: EPISTEME.

Testy normality
- orientační

Pro rychlou orientaci, jestli má výběrový soubor normální rozdělení, lze porovnat **průměr μ** a **medián $x_{0,50}$**. U souboru hodnot s normálním rozdělením by se obě veličiny neměly lišit o víc než desetinu, tj.

$$0,9 < \frac{\mu}{x_{0,50}} < 1,1$$

Tímto testem vlastně ověřujeme, jestli rozdělení není příliš šikmé.
(Lepš & Šmilauer, 2016)

Diskuse o normalitě výběrového souboru má smysl pouze pokud je **soubor dostatečně velký** - máme-li méně než 10 hodnot, nelze z nich o rozdělení říct téměř nic.
Rozumný počet hodnot je **větší než 100**, lépe větší než 200.

Lepš, J., Šmilauer, P. (2016). *Biostatistika*. České Budějovice: EPISTEME.

Test normality
- test šikmosti a špičatosti

Normální rozdělení má **nulovou** **šikmost** i **špičatost**. →

H$_0$: šikmost a špičatost jsou nulové.

Označme: a_3 šikmost, a_3^* šikmost podle Excelu =SKEW

a_4 špičatost, a_4^* špičatost podle Excelu =KURT

Testovací kritérium pro šikmost (Meloun & Militký, 2013)

$$u_3 = \frac{a_3}{\sqrt{\frac{6(N-2)}{(N+1)(N+3)}}} = \frac{\sqrt{(N+1)(N-2)(N+3)}}{\sqrt{N(N-1)}} a_3^*$$

Testovací kritérium pro špičatost (Meloun & Militký, 2013)

$$u_4 = \frac{a_4 + \frac{6}{N+1}}{\sqrt{\frac{24N(N-2)(N-3)}{(N+1)^2(N+3)(N+5)}}} = \sqrt{\frac{(N-2)(N-3)(N+3)(N+5)}{24N(N^2+1)}} a_4^*$$

Meloun, M., Militký, J. (2013). *Kompendium statistického zpracování dat*. Praha: Karolinum.

Test normality
- test šikmosti a špičatosti

Nulovou hypotézu, že šikmost je nulová ($a_3 = 0$) resp. špičatost je nulová ($a_4 = 0$), zamítáme v případě, že

$\boxed{u_3 > u_{1-\alpha,0.05}}$ resp. $\boxed{u_4 > u_{1-\alpha,0.05}}$

$u_{1-\alpha,0.05}$ ...kvantily normálního rozdělení N(0, 1) pro zvolenou hladinu významnosti α, které najdeme ve statistických tabulkách nebo vypočítáme pomocí funkce

=NORMINV(1-*α*/2; 0; 1)

Šikmost a špičatost – kritické hodnoty

N	α = 0,05				α = 0,01			
	a_3	a_3^*	a_4	a_4^*	a_3	a_3^*	a_4	a_4^*
20	0.927	1.004	1.206	1.984	1.218	1.319	1.675	2.595
30	0.794	0.837	1.179	1.648	1.044	1.100	1.610	2.162
40	0.705	0.733	1.116	1.444	0.926	0.963	1.513	1.896
50	0.640	0.660	1.054	1.303	0.841	0.867	1.422	1.710
70	0.550	0.562	0.948	1.113	0.723	0.739	1.272	1.461
100	0.466	0.473	0.832	0.939	0.612	0.622	1.122	1.233
150	0.384	0.388	0.706	0.772	0.505	0.510	0.941	1.014
200	0.334	0.337	0.624	0.671	0.440	0.443	0.830	0.882
250	0.300	0.302	0.565	0.602	0.394	0.397	0.751	0.791
300	0.274	0.276	0.521	0.550	0.361	0.362	0.691	0.723
400	0.238	0.239	0.456	0.477	0.313	0.314	0.604	0.627
500	0.213	0.214	0.411	0.427	0.280	0.281	0.544	0.562
700	0.181	0.181	0.350	0.362	0.237	0.238	0.463	0.475
1000	0.151	0.152	0.295	0.303	0.199	0.199	0.390	0.398
2000	0.107	0.107	0.211	0.214	0.141	0.141	0.278	0.282
3000	0.088	0.088	0.173	0.175	0.115	0.115	0.230	0.230
4000	0.076	0.076	0.150	0.152	0.100	0.100	0.198	0.199