Stručný úvod do testování statistických hypotéz

1. Formulujeme hypotézu
   (předpokládáme, že pozorovaný jev je pouze náhodný).
2. Zvolíme hladinu významnosti testu α, tj. riziko, s nímž jsme ochotni se smířit. Jedná se o riziko, že zamítneme hypotézu, která je ve skutečnosti správná.
3. Spočítáme příslušné testovací kritérium a porovnáme ho s příslušnou kritickou hodnotou.
4. Nulovou hypotézu buď nezamítneme (testovací kritérium je menší než kritická hodnota) nebo zamítneme
   (testovací kritérium je větší než kritická hodnota).

Chyba 1. a 2. 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 že 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.

Hodnotu α volíme nejčastěji 0,05; 0,01; 0,005; 0,001 (pro jiné riziko nemáme většinou k dispozici kritické hodnoty).
Když hypotézu zamítneme, znamená to, že téměř jistě (s pravděpodobností $1 - \alpha$) neplatí.

Stručný úvod do testování statistických hypotéz

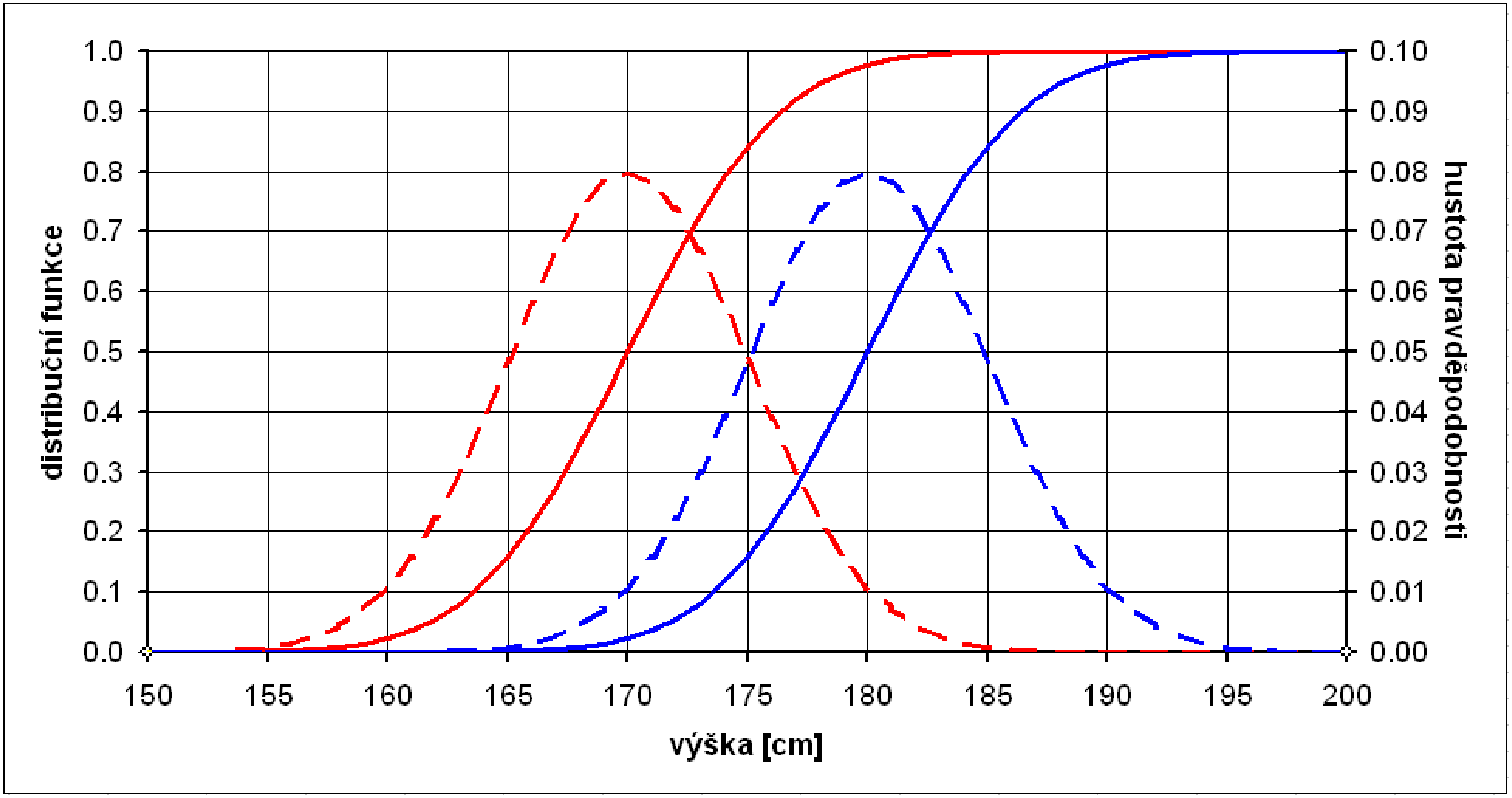
		Výsledek testu	
		platí H_0	neplatí H_0
Skutečnost	platí H_0	správné rozhodnutí pravděpodobnost **1 -** α **spolehlivost testu**	chyba I. druhu pravděpodobnost α **hladina významnosti**
	neplatí H_0	chyba II. druhu pravděpodobnost β **hladina významnosti**	správné rozhodnutí pravděpodobnost **1 -** β **síla testu**

Stručný úvod do testování statistických hypotéz

Chceme testovat podle výšky, zda neznámá osoba je muž.
Muži mají μ=180 cm, σ=5 cm; ženy μ=170 cm, σ=5 cm.
Nulová hypotéza: Osoba je žena.
Jakou výšku musíme požadovat na hladině významnosti 0.05 (5 % riziko, že řekneme, že je to muž a bude to žena)?
Jakou výšku musíme požadovat na hladině významnosti 0.01?
Jaká je pravděpodobnost chyby 2. druhu?
Jaká je síla testu?

Stručný úvod do testování statistických hypotéz

Muži jsou modře, ženy červeně; distribuční funkce je plnou čarou, hustota pravděpodobnosti čárkovaně.



Grubbsův test odlehlých hodnot

Jako míra odlehlosti hodnoty slouží její vzdálenost od aritmetického průměru výběru dat s **normálním rozdělením**, vztažená k výběrové směrodatné odchylce.

Testovací statistika má tvar
(Meloun & Militký, 2013)

$$T = \frac{|x_{ext} - \hat{\mu}|}{\sigma}$$

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

Poznámky:

1. Testujeme jednu hodnotu, která je nejvzdálenější od průměru X_{ext}.
2. V literatuře existují i varianty testující současně minimální i maximální hodnotu, případně používající nevýběrovou směrodatnou odchylku. Tyto varianty používají jiné kritické hodnoty.

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

N	3	4	5	7	10	15	20	30	50	70	100	200
$T_{N,0,05}$	0,94	1,28	1,53	1,87	2,17	2,46	2,64	2,86	3,10	3,23	3,37	3,60
$T_{N,0,01}$	0,94	1,30	1,58	1,98	2,35	2,71	2,92	3,18	3,45	3,60	3,74	3,97

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

Test střední hodnoty normálního rozdělení

Na začátku předpokládáme, že střední hodnota souboru s normálním rozdělením, ze kterého byl proveden výběr, je μ_0.

Testovací kritérium:

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

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

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

Test střední hodnoty normálního rozdělení

Hodnoty:
5.34; 4.03; 5.94; 3.24; 5.21;
5.07; 4.31; 7.47; 6.81; 4.59.

$\mu = 5{,}20$

$\sigma = 1{,}28$

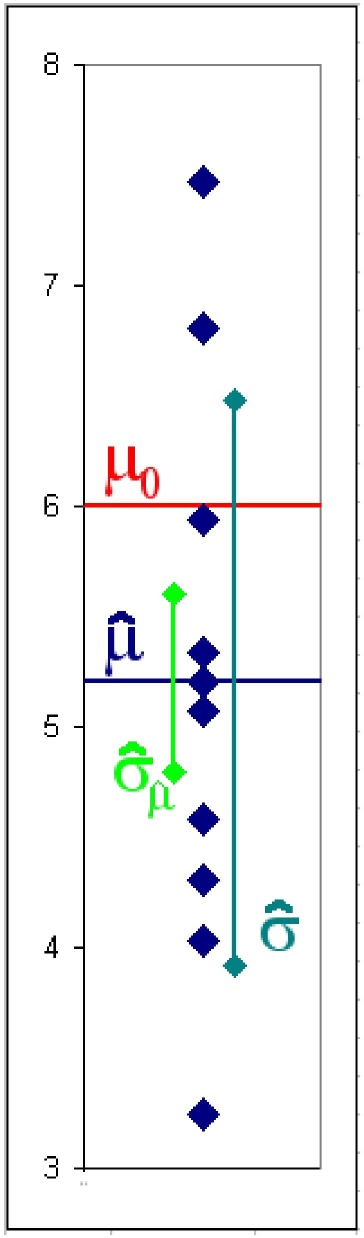
$\sigma_{\mu} = 0{,}40$

$\mu_0 = 6{,}00$

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

$t = 2{,}00$

$t_{krit}(\alpha=0{,}05) = 2{,}26.$



Test rovnosti dvou středních hodnot normálního rozdělení

Na začátku předpokládáme, že střední hodnoty dvou souborů s normálním rozdělením, ze kterých byl proveden výběr, se rovnají.

Testovací kritérium:

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

kde $\widehat{\mu_1}$ a $\widehat{\sigma_1}$ resp. $\widehat{\mu_2}$ a $\widehat{\sigma_2}$ jsou výběrová střední hodnota a výběrová směrodatná odchylka 1. souboru resp. 2. souboru.

Kritickou hodnotou $t_{1-\alpha}(N_1+N_2-2)$ jsou kvantily Studentova rozdělení s N_1+N_2-2 stupni volnosti pro zvolenou hladinu významnosti α, které najdeme ve statistických tabulkách nebo vypočítáme pomocí funkce =TINV(α, N_1+N_2-2).

Test rozdílu dvou středních hodnot normálního rozdělení

Hodnoty:

1. sada: 7.42; 5.43; 5.88; 6.53; 7.71; 5.72; 5.77; 4.89; 5.65; 5.92.

2. sada: 4.79; 4.87; 5.60; 4.58; 6.59; 4.55; 3.71; 3.85; 3.46; 5.36.

$\mu_1 = 6{,}09$ $\mu_2 = 4{,}74$

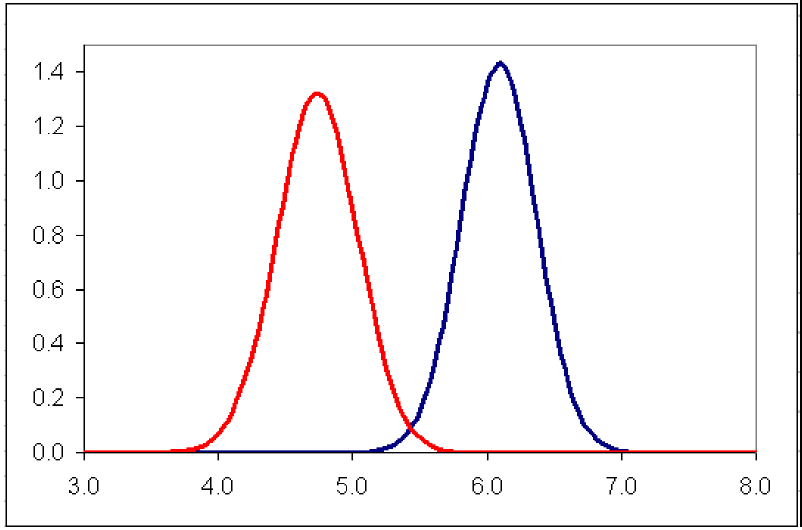
$\sigma_1 = 0{,}88$ $\sigma_2 = 0{,}95$

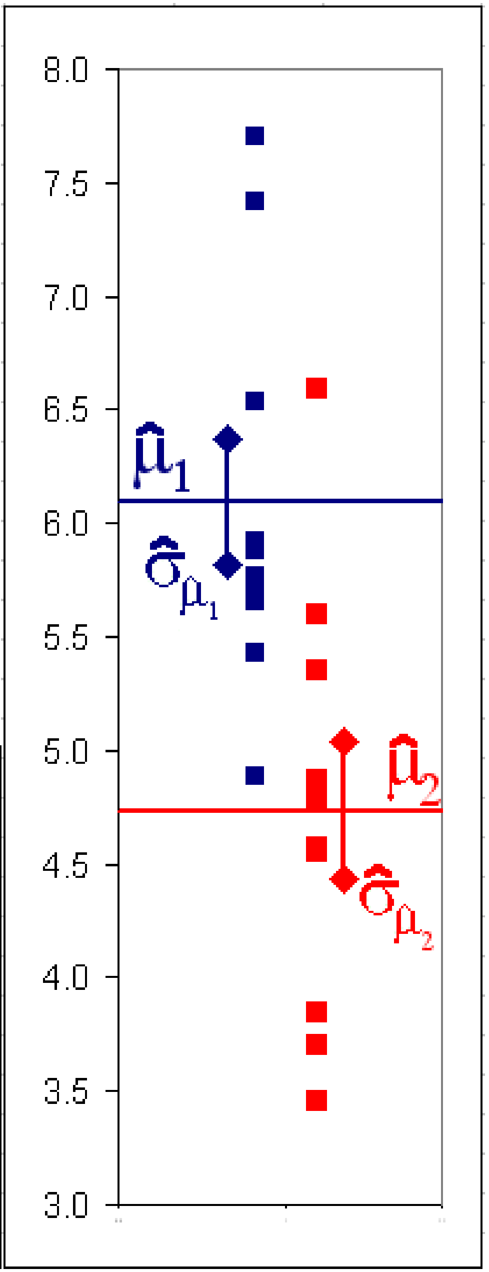
$\sigma_{\mu 1} = 0{,}28$ $\sigma_{\mu 2} = 0{,}30$

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

$t = 3{,}31$

$t_{krit}(\alpha=0{,}01) = 2{,}88.$





Test rozdílu 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 rozdílu 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 spočítat pro každou osobu rozdíl obou testů paměti a testovat, zda je střední rozdíl mezi testy nulový nebo různý od nuly.

Orientační test normality

Pro rychlou orientaci, jestli má výběrový soubor normální rozdělení, lze porovnat průměr a medián. U souboru hodnot s normálním rozdělením by se obě veličiny neměly lišit o víc než desetinu:

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

kde μ je průměr a $x_{0,50}$ medián výběrového souboru.
Tímto testem vlastně ověřujeme, jestli rozdělení není příliš šikmé.

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.

Test normality

Normální rozdělení má nulovou šikmost i špičatost.
Při testování normality tedy budeme testovat nulové hypotézy že šikmost a špičatost jsou nulové.

a_3 je šikmost, a_3^* je šikmost podle Excelu

a_4 je špičatost, a_4^* je špičatost podle Excelu

Testovací kritérium pro šikmost

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

Testovací kritérium pro špičatost

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

Test normality

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

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

kde $u_{1-\alpha,0.05}$ jsou 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).

Kritické hodnoty šikmosti a špičatosti

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