

4. cvičení: Úvod do testování hypotéz. Testy normality.

Příklad 1.: Systematická chyba měřicího přístroje se eliminuje nastavením přístroje a měřením etalonu, jehož správná hodnota je $\mu = 10,00$. Nezávislými měřeními za stejných podmínek byly získány hodnoty: 10,24 10,12 9,91 10,19 9,78 10,14 9,86 10,17 10,05, které považujeme za realizace náhodného výběru rozsahu 9 z rozložení $N(\mu, \sigma^2)$, kde střední hodnotu μ neznáme a výrobce přístroje garantuje směrodatnou odchylku $\sigma = 0,15$. Budeme testovat hypotézu, že $\mu = 10$.

Proti nulové hypotéze $H_0: \mu = 10$ postavíme oboustrannou alternativu $H_1: \mu \neq 10$. Na hladině významnosti 0,05 testujte H_0 proti H_1

- a) pomocí kritického oboru
- b) pomocí intervalu spolehlivosti
- c) pomocí p-hodnoty

Výsledky:

$$t_0 = 1,022$$

Ad a) $W = (-\infty, -1,96) \cup (1,96, \infty)$. Protože $1,022 \notin W$, H_0 nezamítáme na hladině významnosti 0,05.

Ad b) $(d, h) = (9,953; 10,149)$. Protože $10 \in (-\infty; 10,133)$, H_0 nezamítáme na hladině významnosti 0,05.

Ad c) $p = 0,3077$. Protože $0,3077 > 0,05$, nulovou hypotézu nezamítáme na hladině významnosti 0,05.

Příklad 2.: Uvažme data z 1. příkladu. Proti nulové hypotéze $H_0: \mu = 10$ postavíme levostrannou alternativu $H_1: \mu < 10$. Na hladině významnosti 0,05 testujte H_0 proti H_1

- a) pomocí kritického oboru
- b) pomocí intervalu spolehlivosti
- c) pomocí p-hodnoty

Výsledky:

Ad a) $W = (-\infty, -1,64)$. Protože $1,022 \notin W$, H_0 nezamítáme na hladině významnosti 0,05.

Ad b) $(-\infty, h) = (-\infty; 10,133)$. Protože $10 \in (9,953; 10,149)$, H_0 nezamítáme na hladině významnosti 0,05.

Ad c) $p = 0,3077$. Protože $0,3077 > 0,05$, nulovou hypotézu nezamítáme na hladině významnosti 0,05.

Příklad 3.: Uvažme data z 1. příkladu. Proti nulové hypotéze $H_0: \mu = 10$ postavíme pravostrannou alternativu $H_1: \mu > 10$. Na hladině významnosti 0,05 testujte H_0 proti H_1

- a) pomocí kritického oboru
- b) pomocí intervalu spolehlivosti
- c) pomocí p-hodnoty

Výsledky:

Ad a) $W = (1,64, \infty)$. Protože $1,022 \notin W$, H_0 nezamítáme na hladině významnosti 0,05.

Ad b) $(d, \infty) = (9,969; \infty)$. Protože $10 \in (9,969; \infty)$, H_0 nezamítáme na hladině významnosti 0,05.

Ad c) $p = 0,15386$. Protože $0,15386 > 0,05$, nulovou hypotézu nezamítáme na hladině významnosti 0,05.

Příklad 4.: V sedmi náhodně vybraných prodejnách byly zjišťovány ceny určitého druhu zboží (v Kč): 35 29 30 33 45 33 36. Pomocí Lillieforsovy varianty K-S testu zjistěte, zda na hladině významnosti 0,05 lze tyto hodnoty považovat za realizace výběru z normálního rozložení.

Výsledky:

Testová statistika: $D_7 = 0,23923$, modifikovaná kritická hodnota pro $n = 7$, $\alpha = 0,05$ je 0,3. Protože $0,23923 < 0,3$, hypotézu o normalitě nezamítáme na hladině významnosti 0,05.

Výpočet pomocí systému STATISTICA:

Otevřeme nový datový soubor o jedné proměnné (nazveme ji X) a 7 případech. Do proměnné X napíšeme zjištěné ceny.

Statistika – Základní statistiky/tabulky – Tabulky četností – OK – Proměnné X – OK – zvolíme záložku Normalita – odškrtneme K – S test – Testy normality. Dostaneme tabulku:

Proměnná	Testy normality (ceny_zbozi.sta)		
	N	max D	Lilliefors p
X	7	0,240290	p > .20

Testová statistika nabývá hodnoty 0,24 (rozdíl oproti ručnímu výpočtu je způsoben zaokrouhlováním), odpovídající p-hodnota je větší než 0,05, tedy na hladině významnosti 0,05 nezamítáme hypotézu o normalitě cen daného druhu zboží.

Doplnění: V systému STATISTICA lze snadno provést i Shapiro -Wilkův test a Andersonův – Darlingův test.

Provedení Shapiro – Wilkova testu:

V menu vybereme Statistiku – Základní statistiky/tabulky – Tabulky četností – OK, Proměnné X – OK. Na záložce zvolíme Normalita a zaškrtneme Shapiro – Wilkův W test – Testy normality.

Proměnná	Testy normality (ceny_zbozi.sta)		
	N	W	p
X	7	0,868661	0,180679

Vidíme, že testová statistika S-W testu je $W = 0,8687$, odpovídající p-hodnota je 0,1807, tedy hypotézu o normalitě nezamítáme na hladině významnosti 0,05.

Provedení A - D testu:

Statistiky – Rozdělení & simulace – proložení dat rozděleními – OK – Proměnné Spojité: X – na záložce Spojité proměnné ponecháme zaškrtnuté pouze Normální, na záložce Možnosti vybereme Anderson – Darling – OK – Souhrnné statistiky rozdělení.

	Souhrn rozdělení for Proměnná: X (ceny_zbozi.sta)	
	AD stat.	AD p-hodn.
Normální (poloha,měřítko)	0,455557	0,786985

Testová statistika A – D testu je 0,4556, odpovídající p-hodnota je 0,787, tedy hypotézu o normalitě nezamítáme na hladině významnosti 0,05.