Úvod do testování hypotéz
17.5. a 21.5. 2018

Provádíme průzkum, zda v nejmenované hospodě okrádají své hosty. Zakoupíme proto 10 piv a změříme jejich objem. Obdrželi jsme následující hodnoty (v litrech):

$$0.510,\ 0.462,\ 0.491,\ 0.466,\ 0.461,\ 0.503,\ 0.495,\ 0.488,\ 0.512,\ 0.505.$$

(Vychází $\overline{X}_{10} = 0.4893$, $S_{10} = \sqrt{S_{10}^2} = 0.01973$.) Předpokládejme, že data jsou realizace nezávislých náhodných veličin s normálním rozdělením.

(a) Odhadněte bodově střední hodnotu objemu jednoho piva. Lze z tohoto odhadu něco usuzovat o tom, zda v hospodě točí pivo „správně“?

(b) Odhadněte střední hodnotu objemu jednoho piva intervalově.
- Uveďte 95%-ní oboustranný odhad. Zamyslete se, kdy pro obecná data $X_1, \dots, X_n$ překrývá tento interval hodnotu 0.5.
- Uveďte 95%-ní horní odhad. Opět se zamyslete, pro která $X_1, \dots, X_n$ překrývá tento interval hodnotu 0.5.

(c) Máme podezření, že střední hodnota natočeného piva je menší než 0.5 l. Ověřte na základě naměřených hodnot pravdivost jeho výroku.
- Uveďte uvažovaný model a zformulujte nulovou a alternativní hypotézu.
- Uvědomte si, jaké naměřené hodnoty budou svědčit proti nulové hypotéze (ve prospěch alternativní). Na základě této úvahy nalezněte vhodný kritický obor.
- Proveďte test na hladině $\alpha = 0.05$ pro naše data. Jaký je závěr? Jak jej budeme interpretovat?
- Jaká je souvislost mezi výsledkem tohoto testu a horním intervalovým odhadem z (b)?

(d) Chamtivého majitele dané hospody ale spíše zajímá, zda hostinský netočí více piva než by měl. Dal nám tedy za úkol otestovat na základě našich dat , zda tomu tak není.
- Formulujte nulovou a alternativní hypotézu pro tuto situaci.
- Jaká naměřená data by nyní svědčila proti nulové hypotéze ve prospěch alternativy? Na základě této úvahy nalezněte vhodný kritický obor.
- Proveďte test pro naše data. Jaký je závěr?

(e) Hostinský tvrdí, že střední hodnota natočeného piva je přesně 0.5 l. Ověřte tedy ještě i toto jeho tvrzení.
- Opět napište nulovou a alternativní hypotézu.
- Které hodnoty nyní svědčí proti H_0? Jak bude vypadat kritický obor?
- Jaká je souvislost s oboustranným intervalovým odhadem z (b)?

Testování hypotéz = ověřování platnosti nějakého výroku pomocí naměřených dat.

$\hookrightarrow$ **Předpokládaný model**: $X_1, \dots, X_n$ je náhodný výběr z určitého rozdělení F_θ, kde $\theta \in \Theta$ neznáme. Naše data jsou konkrétní realizací tohoto výběru.

$\hookrightarrow$ **Hypotéza** je výrok týkající se neznámé hodnoty θ, o jehož platnosti chceme rozhodnout na základě nasbíraných dat.
Testuje se vždy tzv. **nulová hypotéza** $H_0 : \theta \in \Theta_0$ proti **alternativní hypotéze** $H_1 : \theta \in \Theta_1$, kde Θ_0 a Θ_1 jsou disjunktní podmnožiny Θ.

$\hookrightarrow$ **Test** je rozhodovací pravidlo, na jehož základě zamítáme nebo nezamítáme H_0.
Možná rozhodnutí:
- zamítáme H_0 ve prospěch H_1 (*naše data svědčí proti H_0, prokazujeme platnost H_1*)
- nezamítáme H_0 (*na základě našich dat nelze H_0 zamítnout, naše data nejsou v rozporu s H_0*)

$\leftrightsquigarrow$ **nesymetrie** mezi H_0 a H_1!

$\hookrightarrow$ Většinou nemůžeme rozhodnout s absolutní jistotou, která z hypotéz je platná $\rightsquigarrow$ můžeme se dopustit **chyby**. Mohou nastat tyto možnosti:

	Skutečnost	
Rozhodnutí	H_0 platí	H_1 platí
zamítáme H_0	chyba 1.druhu	OK
nezamítáme H_0	OK	chyba 2.druhu

Chyba 1.druhu je závažnější (falešně něco prokazujeme) $\rightsquigarrow$ její pravděpodobnost chceme kontrolovat.
Zvolíme α = maximální přípustná pravděpodobnost chyby 1.druhu (většinou $\alpha = 0.05$ nebo 0.01) a chceme

$$\mathsf{P}(\text{chyba 1. druhu}) = \mathsf{P}(\text{zamítáme } H_0 |\ H_0 \text{ platí}) \leq \alpha.$$

$\hookrightarrow$ Test je popsán **kritickým oborem** W = množina výsledků pokusů, pro které H_0 zamítáme

- Je-li $(X_1, \dots, X_n) \in W$, pak H_0 zamítáme.
- Je-li $(X_1, \dots, X_n) \notin W$, pak H_0 nezamítáme.

Musí platit:

$$\mathsf{P}_\theta\big((X_1, \dots, X_n) \in W\big) \leq \alpha \quad \text{pro všechna } \theta \in \Theta_0$$

a mluvíme pak o testu na **hladině** α.