12. cvičení z STP

6. května 2021

Příklad 12.1 (asymptotický test střední hodnoty)

U 64 praktických lékařů byl naměřen výběrový průměr počtu zakazniku za den 23, výběrový rozptyl pak byl roven 36. Rozdělení počtu zakazniku není známé.

(a) Sestrojte (asymptotický) oboustranný interval pro střední hodnotu počtu zakazniku o spolehlivosti 95%.

(b) Otestujte na hladině 5%, zda skutečná střední hodnota počtu zakazniku za den může být považována za rovnu 25.

Řešení:

Máme veličiny

$$X_i = \text{“\textit{počet pacientů u i-tého lékaře za den}”}$$

pro $i = 1, \ldots, n$, kde $n = 64$, které budeme pokládat za nezávislé. Jejich rozdělení není známé.

(a) Asymptotický oboustranný interval pro střední hodnotu o spolehlivosti $1 - \alpha$ se bude podobat tomu, který se odvozuje, pokud X_i mají normální rozdělení. Rozdíl bude jen v tom, že kvantil $t_{1-\frac{\alpha}{2};\, n-1}$ pro Studentovo rozdělení s $n-1$ stupni volnosti (který používáme, když X_i mají normální rozdělení) se nahradí asymptotickou hodnotou když $n \to \infty$, která odpovídá kvantilu $u_{1-\frac{\alpha}{2}} = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$ pro rozdělení $N(0,1)$.

Příslušný asymptotický interval o spolehlivosti $1 - \alpha = 95\%$ tedy je:

$$\langle \mu_L,\ \mu_U \rangle := \left\langle \ \overline{\boldsymbol{x}} - \frac{s_{\boldsymbol{x}}}{\sqrt{n}} \cdot u_{1-\frac{\alpha}{2}} \quad , \quad \overline{\boldsymbol{x}} + \frac{s_{\boldsymbol{x}}}{\sqrt{n}} \cdot u_{1-\frac{\alpha}{2}} \ \right\rangle$$

kde $\overline{\boldsymbol{x}} = 23$ je výběrový průměr a $s^2_{\boldsymbol{x}} = 36$ je výběrový rozptyl. Pro $\alpha = 5\%$ máme hodnotu kvantilu $u_{1-\frac{\alpha}{2}} = u_{0.975} = \Phi^{-1}(0.975) \doteq 1.96$. Po dosazení máme tedy

$$\langle \mu_L,\ \mu_U \rangle := \left\langle \ 23 - \frac{\sqrt{36}}{\sqrt{64}} \cdot 1.96 \quad , \quad 23 + \frac{\sqrt{36}}{\sqrt{64}} \cdot 1.96 \ \right\rangle \doteq \langle\ 22.28,\ 23.72\ \rangle$$

Tento interval se pochopitelně MĚNÍ s každým měřením (protože je závislý na naměřených vstupech), a jeho smysl je ten, že skutečná hodnota $\mu = E(X)$ (která se NEMĚNÍ!) bude obsažena v tomto (obecně proměnném intervalu) s pravděpodobností 95%.

(b) Podle zadání máme na hladině $\alpha = 5\%$ otestovat hypotézu o střední hodnotě $\mu = E(X)$ tvaru

$\mathbf{H}_0 : \mu = \mu_0$

proti alternativní hypotéze:

$\mathbf{H}_A : \mu \neq \mu_0$.

kde $\mu_0 = 25$.

Pomocí intervalového odhadu:

Využijeme už spočítaného asymptotického oboustranného intervalu $\langle \mu_L,\ \mu_U \rangle$ pro střední hodnotu o spolehlivosti 95%. Podle toho, co jsme uvedli výše v části (a), je pravděpodobnost, že střední hodnota

$\mu = E(X)$ bude obsažena v (proměnném) intervalu $\langle \mu_L,\ \mu_U \rangle$, rovna 95%. Tedy mimo tento interval se ocitne jen v 5% případů.

Jestliže předpokládáme, že $\mu = \mu_0$ (tj. hypotézu $\mathbf{H}_0$), bude kritérium pro její zamítnutí na hladině α přirozeně tvaru:

$$\text{zamítáme } \mathbf{H}_0 \text{ (na hladině } \alpha) \quad \Leftrightarrow \quad \mu_0 \notin \langle \mu_L,\ \mu_U \rangle \ .$$

A protože skutečně nakonec máme, že $\mu_0 = 25 \notin \langle\ 22.28,\ 23.72\ \rangle = \langle \mu_L,\ \mu_U \rangle$, tak hypotézu $\mathbf{H}_0$ **ZAMÍTÁME** na hladině 5%.

Pomocí testovací statistiky:

Podmínka pro zamítnutí $\mathbf{H}_0$ na hladině α se dá z formy pro interval spolehlivosti

$$\mu_0 \notin \left\langle \ \overline{\boldsymbol{x}} - \frac{s_{\boldsymbol{x}}}{\sqrt{n}} \cdot u_{1-\frac{\alpha}{2}} \quad , \quad \overline{\boldsymbol{x}} + \frac{s_{\boldsymbol{x}}}{\sqrt{n}} \cdot u_{1-\frac{\alpha}{2}} \ \right\rangle$$

ekvivalentně přepsat jako

$$|t| > u_{1-\frac{\alpha}{2}}$$

kde $t = \frac{\overline{\boldsymbol{x}}-\mu_0}{s_{\boldsymbol{x}}}\sqrt{n}$, což je (podobně jako v **Příkladu 11.4**) hodnota testovací veličiny (tzv. *statistiky*):

$$T = \frac{\overline{\boldsymbol{X}} - \mu_0}{S_{\boldsymbol{X}}}\sqrt{n}$$

Protože však neznáme rozdělení veličin X_i, neznáme ani *přesné* rozdělení této statistiky T. To ale na druhou stranu nevadí, protože pro dost velká n nakonec bude mít veličina T přibližné rozdělení $N(0, 1)$ (bez ohledu na počáteční rozdělení rozdělení veličin X_i). To je tedy důvod, proč se pak v zamítacím kritériu objevují kvantily pro norm. rozdělení. Co přesně znamená "dost velká n", závisí pochopitelně na tom, jak "divoké" je rozdělení veličin X_i. Když si teď připustíme, že X_i by mohly mít Poissonovo rozdělení, tak i relativně malé hodnoty n (my máme $n = 64$) mohou být dostatečné pro použití asymptotiky.

Shrňme si to tedy tak, že kritérium pro zamítnutí $\mathbf{H}_0$ (na hladině α) je tvaru

$$\text{zamítáme } \mathbf{H}_0 \text{ (na hladině } \alpha) \quad \Leftrightarrow \quad |t| > u_{1-\frac{\alpha}{2}} \ .$$

Při konkrétním dosazení máme

$$t \ = \ \frac{\overline{\boldsymbol{x}} - \mu_0}{s_{\boldsymbol{x}}}\sqrt{n} \ = \ \frac{23-25}{\sqrt{36}}\sqrt{64} \ = \ -\frac{16}{3} \ \doteq \ -5.33$$

a tudíž

$$|t| \doteq |5.33| > 1.96 \doteq u_{0.975}$$

což znamená, že hypotézu $\mathbf{H}_0$ (opět) **ZAMÍTÁME** na hladině 5%.

(Výsledek musel samozřejmě dopadnout stejně jako pomoci intervalu spolehlivosti, protože je to ekvivalentní princip.)

Poznámky k testu dobré shody: Chceme otestovat (na hladině α), jestli daná veličina X s konečně mnoha (navzájem různými!) hodnotami $a_1, \ldots, a_k$ (ne nutně číselnými) má předepsané pravděpodobnosti $(p_1, \ldots, p_k)$, tedy nulovou hypotézu

$\mathbf{H}_0 : P(X = a_i) = p_i$ pro všechna $i \in \{1, \ldots, k\}$

proti alternativní hypotéze:

$\mathbf{H}_A : P(X = a_{i_0}) \neq p_{i_0}$, pro alespoň jedno $i_0 \in \{1, \ldots, k\}$

Při n pokusech s veličinou X si pro $i = 1, \ldots, k$ označme veličiny

$$N_i = \text{“}\textit{počet výskytů případu } X = a_i \textit{ při } n \textit{ pokusech}\text{”} .$$

Máme tedy náhodný vektor

$$\boldsymbol{N} = (N_1, \ldots, N_k)$$

a vztah $N_1 + \cdots + N_k = n$. Už z toho vidíme, že veličiny N_i nejsou nezávislé, ale zase k té nezávislosti tak daleko nemají. Náhodný vektor $\boldsymbol{N}$ má tzv. multinomické rozdělení a jednotlivá marginální rozdělení veličin jsou binomická, konkrétně $N_i \sim \mathrm{Bi}(n, p_i)$. Speciálně tedy $E(N_i) = n \cdot p_i$.

Jako testovací veličinu zde používáme:

$$T = \sum_{i=1}^{k} \frac{(N_i - n \cdot p_i)^2}{n \cdot p_i}$$

která má asymptoticky (tj. pro $n \to \infty$) tzv. χ^2-rozdělení s $k-1$ stupni volnosti. Pro praktické použití této asymptotiky se obvykle požaduje, aby platilo, že

$$n \cdot p_i \geq 5 \quad \text{pro všechna} \quad i \in \{1, \ldots, k\} \ .$$

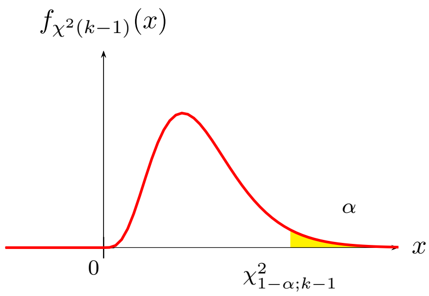
Hodnoty $n \cdot p_i$ se označují jako tzv. *teoretické četnosti*.

Pokud tedy platí nulová hypotéza $\mathbf{H}_0$, měly by být hodnoty veličiny T malé. Jestliže hodnoty T budou příliš velké, bude to důvod k zamítnutí nulové hypotézy.

Jak určit hranici, kde už nastane zamítnutí: veličina T má (přibližně) $\chi^2_{(k-1)}$ rozdělení, tedy platí

$$P_{(H_0 \text{ platí})}\left(T > \chi^2_{1-\alpha;\, k-1}\right) \doteq \alpha$$

kde $\chi^2_{1-\alpha;\, k-1}$ je hodnota kvantilu pro $\chi^2_{(k-1)}$ rozdělení (viz obrázek, kde α je velikost žluté plochy pod hustotou $f_{\chi^2(k-1)}(x)$ pro $\chi^2_{(k-1)}$ rozdělení).



Kritérium pro **ZAMÍTNUTÍ** $\mathbf{H}_0$ (na hladině α) proto volíme jako

$$t > \chi^2_{1-\alpha;\, k-1} \quad \Leftrightarrow \quad \text{zamítáme } \mathbf{H}_0 \text{ (na hladině } \alpha) \ .$$

Z definice chyby 1. druhu, tj.

$$\text{nastává chyba 1. druhu} \quad \Leftrightarrow \quad (\text{hypotéza } \mathbf{H}_0 \text{ platí} \quad \& \quad \text{my ji zamítáme})$$

pak totiž máme, že

$$P_{(H_0 \text{ platí})}\Big(\text{nastává chyba 1. druhu}\Big) = P_{(H_0 \text{ platí})}\Big(\text{zamítáme } \mathbf{H}_0 \text{ (na hladině } \alpha)\Big) =$$

$$= P_{(H_0 \text{ platí})}\left(T > \chi^2_{1-\alpha;\, k-1}\right) \doteq \alpha$$

neboli pravděpodobnost chyby 1. druhu (ovšem za předpokladu platnosti $\mathbf{H}_0$!) je pak omezena hodnotou α.

Příklad 12.2 *Firma má 3 pobočky. Dva roky bylo sledováno, která z nich zaznamenala nejvyšší měsíční výnos. Bylo zjištěno, že nejvýnosnější byla první pobočka $10\times$, druhá $6\times$ a třetí $8\times$.*
Je možné říct, že první pobočka je nejvýnosnější $2\times$ častěji než každá ze zbylých dvou? Testujte na hladině 5%.

Řešení:
Máme tedy veličinu

$$X = \text{“\textit{číslo pobočky, která je zrovna (tj. v daném měsíci) nejvýnosnější}”}$$

s $k = 3$ hodnotami $\{\textit{první, druhá, třetí}\}$.

Nejdříve si potřebujeme zjistit, jaké rozdělení

$$P(X = \textit{první}) = p_1, \quad P(X = \textit{druhá}) = p_2, \quad P(X = \textit{třetí}) = p_3$$

vlastně předpokládáme. Z požadavku máme

$$p_1 = 2 \cdot p_2, \quad p_1 = 2 \cdot p_3, \quad p_1 + p_2 + p_3 = 1$$

z čehož dostáváme

$$(p_1, p_2, p_3) = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$$

Naše hypotéza tedy je

$\mathbf{H}_0$: veličina X má rozdělení s pravděpodobnostmi $(p_1, p_2, p_3) = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$,

a alternativní hypotéza bude:

$\mathbf{H}_A$: veličina X má rozdělení *jiné* než $\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$.

Využijeme test dobré shody. Celkový počet měření (tj. počet měsíců) je $n = 10 + 6 + 8 = 24$. Pro přehlednost si vypíšeme tabulku s jednotlivými četnostmi (pozorovanými i teoretickými):

i (pobočky)	první	druhá	třetí
n_i (pozorované četnosti)	10	8	6
p_i (teoretické pravděpodobnosti)	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$
$n \cdot p_i$ (teoretické četnosti)	$24 \cdot \frac{1}{2} = 12$	$24 \cdot \frac{1}{4} = 6$	$24 \cdot \frac{1}{4} = 6$

Vidíme, že všechny teoretické četnosti jsou ≥ 5, takže skutečně můžeme použít asymptotické přiblížení pro testovací statistiku T (ta tedy bude mít χ^2-rozdělení). Teď už si jen spočítáme hodnotu této statistiky

$$t = \sum_{i=1}^{3} \frac{(n_i - np_i)^2}{np_i} = \frac{(10-12)^2}{12} + \frac{(8-6)^2}{6} + \frac{(6-6)^2}{6} = \frac{1}{3} + \frac{2}{3} + 0 = 1$$

a porovnáme s kvantilem χ^2-rozdělení s $k - 1 = 3 - 1 = 2$ stupni volnosti:

$$t = 1 \not> 5.99 \doteq \chi^2_{0.95;\,2} = \chi^2_{1-\alpha;\,k-1}$$

Protože zamítací kritérium NENÍ splněno, tak $\mathbf{H}_0$ **NEZAMÍTÁME** (na hladině α).

Příklad 12.3 (test dobré shody - geometrické rozdělení)

Při házení mincí jsme zaznamenávali, kdy nám poprvé padne líc. Veličina X měřící počet neúspěchů před prvním úspěchem měla následující četnosti výsledků:

hodnota	0	1	2	3	4	5	6
pozorovaná četnost	29	15	10	5	3	0	2

Posuďte na hladině významnosti $\alpha = 0.05$ hypotézu, že náhodná veličina X má geometrické rozdělení $\mathrm{Geom}(\frac{1}{2})$*, tj.*

$$P(X = i) = \left(\tfrac{1}{2}\right)^{i+1}, \qquad i \in \mathbb{N}_0$$

.

Řešení:

Veličina s geometrickým rozdělením nabývá nekonečně mnoha hodnot. Test dobré shody je ale možné dělat jen s veličinou s *konečně* mnoha hodnotami. Proto musíme některé hodnoty sloučit do jediné skupiny. Zde se přirozeně nabízí udělat to pro hodnoty 6 a výše. Pravděpodobnost pro tuto skupinu je pak součet pravděpodobností jednotlivých hodnot v této skupině. V našem případě je

$$P(X \geq 6) = 1 - P(X < 6) = 1 - \sum_{i=0}^{5} P(X = i) = 1 - \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \frac{1}{64}\right) = \frac{1}{64}\ .$$

Při testu dobré shody porovnáváme naměřené četnosti s očekávanými četnostmi. Rozsah souboru (tj. počet měření) je $n = 29 + 15 + 10 + 5 + 3 + 0 + 2 = 64$. Naší tabulku tedy zpřesníme a doplníme o teoretické pravděpodobnosti p_i a teoretické (tj. očekávané) četnosti $n \cdot p_i$:

položka i	0	1	2	3	4	5	≥ 6
pozorovaná četnost n_i	29	15	10	5	3	0	2
teoretická pravděpodobnost p_i	1/2	1/4	1/8	1/16	1/32	1/64	1/64
teoretická četnost $n \cdot p_i$	32	16	8	4	2	1	1

Další podmínkou pro test dobré shody je to, aby jednotlivé položky měly TEORETICKÉ četnosti $n \cdot p_i \geq 5$. Pokud tomu tak není, je potřeba položky vhodně sloučit tak, abychom této hranice dosáhli. Zde se opět nabízí udělat to pro hodnoty $i \geq 3$.

Původní veličinu X tedy nakonec nahradíme veličinou X' popsanou následující tabulkou:

položka i	0	1	2	≥ 3
pozorovaná četnost n_i	29	15	10	10
teoretická pravděpodobnost p_i	1/2	1/4	1/8	1/8
teoretická četnost $n \cdot p_i$	32	16	8	8

Nyní už můžeme zformulovat naši nulovou hypotézu

$\mathbf{H}_0$: pro pravděpodobnosti hodnot veličiny X' platí $\left(p_0,\ p_1,\ p_2,\ p_{\geq 3}\right) = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\right)$,

kterou budeme testovat proti alternativní hypotéze:

$\mathbf{H}_1$: pro pravděpodobnosti hodnot veličiny X' platí $\left(p_0,\ p_1,\ p_2,\ p_{\geq 3}\right) \neq \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\right)$.

Množina K všech položek (tj. hodnot) veličiny X' má tedy velikost $k = 4$. Pro ni si (pomocí *poslední* tabulky) spočítáme realizaci testovací statistiky

$$t = \sum_{i \in K} \frac{(n_i - n \cdot p_i)^2}{n \cdot p_i} = \frac{(29-32)^2}{32} + \frac{(15-16)^2}{16} + \frac{(10-8)^2}{8} + \frac{(10-8)^2}{8} \doteq 1.34$$

Kritérium pro **ZAMÍTNUTÍ** (viz opět poznámka výše) je tvaru

$$t > \chi^2_{1-\alpha;\, k-1} \quad \Leftrightarrow \quad \text{zamítáme } \mathbf{H}_0 \text{ (na dané hladině } \alpha) \ .$$

V našem případě je

$$\chi^2_{1-\alpha;\, k-1} = \chi^2_{0.95;\, 3} \doteq 7.815 \ .$$

Protože

$$t \doteq 1.34 \not> 7.815 \doteq \chi^2_{0.95;\, 3} \ ,$$

nulovou hypotézu $\mathbf{H}_0$ pro veličinu X' **NEZAMÍTÁME**. Tento výsledek interpretujeme tak, že hypotézu

X má geometrické rozdělení s parametrem $q = 1/2$,

rovněž **NEZAMÍTÁME**.