

ANALÝZA DAT V R

Intervaly spolehlivosti a testování hypotéz

– intervaly spolehlivosti a testy hypotéz spolu úzce souvisejí

1. Interval spolehlivosti a test pro střední hodnotu normálního rozdělení při známém rozptylu

Nechť $X_1, X_2, \dots, X_n$ je náhodný výběr z normálního rozdělení $\mathcal{N}(\mu, \sigma^2)$, kde σ^2 je známý.

Oboustranný interval o spolehlivosti $(1 - \alpha) \cdot 100\%$ pro střední hodnotu je

$$\left(\bar{x} - \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \cdot \frac{\sigma}{\sqrt{n}}; \bar{x} + \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \cdot \frac{\sigma}{\sqrt{n}} \right).$$

Testová statistika testu střední hodnoty je

$$T = \frac{\bar{x} - \mu_0}{\sigma} \sqrt{n} \stackrel{H_0}{\approx} \mathcal{N}(0, 1).$$

Funkce v R je ve tvaru `z.test(x, alternative=, mu=, sigma.x=, conf.level=)`, kde argumenty jsou následující:

- `x` ... data (vektor, nebo matice o jednom sloupci)
- `alternative` ... zvolená alternativa:
 - oboustranná $H_1 : \mu \neq \mu_0$... `alternative="two.sided"`
 - pravostranná $H_1 : \mu \geq \mu_0$... `alternative="greater"`
 - levostranná $H_1 : \mu \leq \mu_0$... `alternative="less"`
- `mu` ... testovaná hodnota, tj. μ_0
- `sigma.x` ... známá směrodatná odchylka, tj. σ
- `conf.level` ... spolehlivost, tj. $1 - \alpha$

Pozor! Tato funkce je dostupná pouze po nainstalování balíku **BSDA**.

Výstupem testování hypotéz je mimo jiné i tzv. p-hodnota, což je pravděpodobnost za platnosti nulové hypotézy, že dostaneme stejnou nebo ještě extrémnější hodnotu testové statistiky ve srovnání s právě vypočítanou hodnotou. Stačí tedy porovnat zvolenou hladinu významnosti testu α s p-hodnotou: je-li p-hodnota menší nebo rovna α , pak nulovou hypotézu zamítáme, v opačném případě ji nezamítáme.

111 116 105 111 110 114 108 106 112 108 112 111 105 111 108 110

Příklad 1

U náhodně vybraných žáků 8. třídy bylo změřeno IQ:

Předpokládejme, že data pocházejí z normálního rozdělení s neznámou střední hodnotou μ a rozptylem $\sigma^2 = 9$.

- (a) Určete bodový i 95% intervalový odhad střední hodnoty IQ žáků 8. třídy.
- (b) Určete 95% levostranný intervalový odhad střední hodnoty IQ žáků 8. třídy.
- (c) Odborník tvrdí, že je střední hodnota IQ žáků 8. třídy rovna 110. Proveďte test na hladině $\alpha = 0,05$.

2. Interval spolehlivosti a test pro střední hodnotu normálního rozdělení při neznámém rozptylu

Nechť $X_1, X_2, \dots, X_n$ je náhodný výběr z normálního rozdělení $\mathcal{N}(\mu, \sigma^2)$, kde σ^2 je neznámý.

Oboustranný interval o spolehlivosti $(1 - \alpha) \cdot 100\%$ pro střední hodnotu je

$$\left(\bar{x} - t_{n-1} \left(1 - \frac{\alpha}{2} \right) \cdot \frac{s}{\sqrt{n}}; \bar{x} + t_{n-1} \left(1 - \frac{\alpha}{2} \right) \cdot \frac{s}{\sqrt{n}} \right).$$

Testová statistika testu střední hodnoty je

$$T = \frac{\bar{x} - \mu_0}{s} \sqrt{n} \stackrel{H_0}{\sim} t_{n-1}.$$

Funkce v R je ve tvaru `t.test(x, alternative=, mu=, paired=FALSE, conf.level=)`, kde argumenty jsou následující:

- `x` ... data (vektor, nebo matice o jednom sloupci)
- `alternative` ... zvolená alternativa viz bod 1
- `mu` ... testovaná hodnota, tj. μ_0
- `paired` ... typ t-testu (jednovýběrový, párový, dvouvýběrový) - v tomto případě jednovýběrový
- `conf.level` ... spolehlivost, tj. $1 - \alpha$

Příklad 2

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

- (a) Odhadněte bodově střední hodnotu objemu jednoho piva.

0,510 0,462 0,491 0,466 0,461 0,503 0,495 0,488 0,512 0,505

(b) Zkonstruuje intervalový odhad pro střední hodnotu objemu jednoho piva o spolehlivosti 95%.

(c) Otestujte, zda je hostinský skutečně nepoctivý. Volte $\alpha = 0,05$.

3. Interval spolehlivosti a test pro rozptyl normálního rozdělení

Nechť $X_1, X_2, \dots, X_n$ je náhodný výběr z normálního rozdělení $\mathcal{N}(\mu, \sigma^2)$.

Oboustranný interval o spolehlivosti $(1 - \alpha) \cdot 100\%$ pro rozptyl je

$$\left(\frac{(n-1)s^2}{\chi_{n-1}^2(1-\frac{\alpha}{2})}; \frac{(n-1)s^2}{\chi_{n-1}^2(\frac{\alpha}{2})} \right).$$

Testová statistika testu rozptylu je

$$T = \frac{(n-1)s^2}{\sigma_0^2} \stackrel{H_0}{\sim} \chi_{n-1}^2.$$

Funkce v R je ve tvaru `varTest(x, alternative=, sigma.squared=, conf.level=)`, kde argumenty jsou následující:

- `x` ... data (vektor, nebo matice o jednom sloupci)
- `alternative` ... zvolená alternativa viz bod 1
- `sigma.squared` ... testovaná hodnota, tj. σ_0^2
- `conf.level` ... spolehlivost, tj. $1 - \alpha$

Pozor! Tato funkce je dostupná pouze po nainstalování balíku `EnvStats`.

Příklad 3

Vraťme se k příkladu 2.

(a) Odhadněte bodově rozptyl objemu jednoho piva.

(b) Zkonstruuje intervalový odhad pro rozptyl objemu jednoho piva o spolehlivosti 90%.

4. Interval spolehlivosti a test pro relativní četnost

Nechť $X_1, X_2, \dots, X_n$ je náhodný výběr.

Oboustranný interval o spolehlivosti $(1 - \alpha) \cdot 100\%$ pro relativní četnost je

$$\left(\hat{p} - \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right); \hat{p} + \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \cdot \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \right).$$

Testová statistika testu relativní četnosti je

$$T = \frac{\hat{p} - p_0}{\sqrt{\hat{p}(1 - \hat{p})}} \sqrt{n} \stackrel{H_0}{\sim} N(0, 1).$$

Funkce v R je ve tvaru `binom.test(x, n, alternative=, p=, conf.level=)`, kde argumenty jsou následující:

- `x` ... počet úspěchů
- `n` ... celkový počet pokusů
- `alternative` ... zvolená alternativa viz bod 1
- `p` ... testovaná hodnota, tj. p_0
- `conf.level` ... spolehlivost, tj. $1 - \alpha$

Příklad 4

Z náhodného vzorku 300 lidí mělo 60 modré oči.

- (a) V jakém intervalu můžeme s pravděpodobností 95% očekávat skutečný podíl modrookých v populaci?
- (b) Na hladině 1 % testujte, zda je podíl modrookých v populaci menší než 25 %.

5. Interval spolehlivosti a test pro rozdíl středních hodnot ze dvou závislých výběrů

Nechť $(X_1, Y_1)^T, (X_2, Y_2)^T, \dots, (X_n, Y_n)^T$ je náhodný výběr z dvourozměrného rozdělení se střední hodnotou $(\mu_X, \mu_Y)^T$. Položme $Z_i = X_i - Y_i, i = 1, 2, \dots, n$, a nechť $Z_1, Z_2, \dots, Z_n$ je náhodný výběr z normálního rozdělení $\mathcal{N}(\mu_Z = \mu_X - \mu_Y, \sigma_Z^2)$.

Oboustranný interval o spolehlivosti $(1 - \alpha) \cdot 100\%$ pro rozdíl středních hodnot je

$$\left(\bar{z} - t_{n-1} \left(1 - \frac{\alpha}{2} \right) \cdot \frac{s_Z}{\sqrt{n}}; \bar{z} + t_{n-1} \left(1 - \frac{\alpha}{2} \right) \cdot \frac{s_Z}{\sqrt{n}} \right).$$

Testová statistika testu rozdílu středních hodnot je

$$T = \frac{\bar{z} - \mu_0}{s_Z} \sqrt{n} \stackrel{H_0}{\sim} t_{n-1}.$$

Funkce v R je ve tvaru `t.test(x, y, alternative=, mu=, paired=TRUE, conf.level=)`, kde argumenty jsou následující:

- `x, y` ... data (dva vektory, nebo matice o dvou sloupcích)
- `alternative` ... zvolená alternativa viz bod 1

- `mu` ... testovaná hodnota, tj. μ_0
- `paired` ... typ t-testu (jednovýběrový, párový, dvouvýběrový) - v tomto případě párový
- `conf.level` ... spolehlivost, tj. $1 - \alpha$

Příklad 5

U několika osob byla měřena síla stisku pravé a levé ruky:

pravá	140	90	125	130	95	121	85	97	131	110
levá	138	87	110	132	96	120	86	90	129	100

Potvrzují data domněnku, že pravá ruka je silnější? Volte $\alpha = 0,05$.

Řešení

Pro každou osobu máme dvojici hodnot (X = síla stisku pravé ruky, Y = síla stisku levé ruky), tj. párová měření. Jde tedy o náhodný výběr z dvourozměrného rozdělení se střední hodnotou $(\mu_X, \mu_Y)^T$, kde μ_X je střední hodnota síly stisku pravé ruky a μ_Y je střední hodnota síly stisku levé ruky. Testujeme $H_0 : \mu_X = \mu_Y$ (neboli $H_0 : \mu_X - \mu_Y = 0$) proti $H_1 : \mu_X > \mu_Y$ (neboli $H_0 : \mu_X - \mu_Y > 0$) na hladině $\alpha = 0,05$. Položíme-li $Z_i = X_i - Y_i, i = 1, 2, \dots, 10$, pak $H_0 : \mu_Z = 0$ a $H_1 : \mu_Z > 0$. Předpokladem pro použití párového t-testu je, že $Z_1, Z_2, \dots, Z_{10}$ je náhodný výběr z normálního rozdělení – to můžeme ověřit např. použitím Shapirova-Wilkova testu (p-hodnota je 0,1169). Jelikož jsou splněny předpoklady párového t-testu, můžeme jej použít. V R použijeme funkci `t.test`.

```
x <- c(140, 90, 125, 130, 95, 121, 85, 97, 131, 110)
y <- c(138, 87, 110, 132, 96, 120, 86, 90, 129, 100)
z <- x-y
shapiro.test(z)
t.test(x, y, alternative="greater", mu=0, paired=TRUE, conf.level=0.95)
```

Jelikož p-hodnota párového t-testu je 0,0334 (je menší než stanovená hladina významnosti α), zamítáme nulovou hypotézu a přijímáme alternativní hypotézu. Pravá ruka je tedy silnější než levá.

6. Interval spolehlivosti a test pro rozdíl středních hodnot ze dvou nezávislých výběrů

Nechť $X_1, X_2, \dots, X_n$ je náhodný výběr z normálního rozdělení $\mathcal{N}(\mu_X, \sigma^2)$ a $Y_1, Y_2, \dots, Y_m$ je náhodný výběr z normálního rozdělení $\mathcal{N}(\mu_Y, \sigma^2)$.

Oboustranný interval o spolehlivosti $(1 - \alpha) \cdot 100\%$ pro rozdíl středních hodnot je

$$\left(\bar{x} - \bar{y} - t_{n+m-2} \left(1 - \frac{\alpha}{2} \right) \cdot s_* \sqrt{\frac{n+m}{nm}}; \bar{x} - \bar{y} + t_{n+m-2} \left(1 - \frac{\alpha}{2} \right) \cdot s_* \sqrt{\frac{n+m}{nm}} \right),$$

kde $s_*^2 = \frac{1}{n+m-2}[(n-1)s_X^2 + (m-1)s_Y^2]$.

Testová statistika testu rozdílu středních hodnot je

$$T = \frac{\bar{x} - \bar{y} - \mu_0}{s_*} \sqrt{\frac{nm}{n+m}} \stackrel{H_0}{\sim} t_{n+m-2}.$$

Funkce v R je ve tvaru `t.test(x, y, alternative=, mu=, paired=FALSE, var.equal=TRUE, conf.level=)`, kde argumenty jsou následující:

- `x, y` ... data (dva vektory, nebo matice o dvou sloupcích)
- `alternative` ... zvolená alternativa viz bod 1
- `mu` ... testovaná hodnota, tj. μ_0
- `paired` ... typ t-testu (jednovýběrový, párový, dvouvýběrový) - v tomto případě dvouvýběrový
- `var.equal` ... shoda rozptylů - v tomto případě ano (existuje i test pro rozdíl středních hodnot, kde rozptyly shodné nejsou)
- `conf.level` ... spolehlivost, tj. $1 - \alpha$

Příklad 6

Byly zjištěny následující výšky žáků 5. třídy:

chlapci	130	140	136	141	139	133	149	151	139	136	138	142	127	139	147
dívky	135	141	143	132	146	146	151	141	141	131	142	141			

Testujte, zda chlapci a dívky jsou v průměru stejně vysocí. Volte $\alpha = 0,05$.

Řešení

Jedná se o dvě nezávislé skupiny měření – skupina chlapců a skupina dívek. Pro každou osobu máme jednu hodnotu, označme X výšku chlapce a Y výšku dívky. Jde tedy o dva nezávislé náhodné výběry, přičemž první výběr má rozsah 15 a pochází z populace se střední hodnotou μ_X a rozptylem σ_X^2 , druhý výběr má rozsah 12 a pochází z populace se střední hodnotou μ_Y a

rozptylem σ_Y^2 . Testujeme $H_0 : \mu_X = \mu_Y$ (neboli $H_0 : \mu_X - \mu_Y = 0$) proti $H_1 : \mu_X \neq \mu_Y$ (neboli $H_0 : \mu_X - \mu_Y \neq 0$) na hladině $\alpha = 0,05$. Předpokladem pro použití dvouvýběrového t-testu je, že máme dva nezávislé náhodné výběry z normálního rozdělení se stejným rozptylem. Normalitu můžeme ověřit např. použitím Shapirova-Wilkova testu (p-hodnoty jsou 0,7998 a 0,3835). Porovnat rozptyly můžeme např. použitím F-testu (p-hodnota je 0,6977). Jelikož jsou splněny předpoklady dvouvýběrového t-testu, můžeme jej použít. V R použijeme funkci `t.test`.

```
x <- c(130, 140, 136, 141, 139, 133, 149, 151, 139, 136, 138, 142, 127, 139, 147)
y <- c(135, 141, 143, 132, 146, 146, 151, 141, 141, 131, 142, 141)

shapiro.test(x)
shapiro.test(y)

var.test(x, y, conf.level=0.95) t.test(x, y, alternative="two.sided", mu=0, paired=FALSE,
var.equal=TRUE, conf.level=0.95)
```

Jelikož p-hodnota dvouvýběrového t-testu je 0,4883 (je větší než stanovená hladina významnosti α), nezamítáme nulovou hypotézu. Je možné, že chlapci a dívky mají v průměru stejnou výšku.

INTERVALY SPOLEHLIVOSTI A TESTY HYPOTÉZ	
1.	<code>z.test(x, alternative=, mu=, sigma.x=, conf.level=)</code>
2.	<code>t.test(x, alternative=, mu=, paired=FALSE, conf.level=)</code>
3.	<code>varTest(x, alternative=, sigma.squared=, conf.level=)</code>
4.	<code>binom.test(x, n, alternative=, p=, conf.level=)</code>
5.	<code>t.test(x, alternative=, mu=, paired=TRUE, conf.level=)</code>
6.	<code>t.test(x, alternative=, mu=, paired=FALSE, var.equal=TRUE, conf.level=)</code>

Ověření normality

1. Grafické posouzení

- např. histogram, q-q graf: `qqnorm(x)`, `qqline(x)`

2. Statistické testy

- např. Shapirův-Wilkův test: `shapiro.test(x)`

Test shody rozptylů

- např. F-test shody rozptylů: `var.test(x, y, conf.level=)`