

6. Parametrické úlohy o jednom náhodném výběru z normálního rozložení

6.1. Motivace

6.2. Věta: Rozložení statistik odvozených z výběrového průměru a výběrového rozptylu

6.3. Příklad: Hmotnost balíčku krystalového cukru baleného na automatické lince se řídí normálním rozložením se střední hodnotou 1002 g a směrodatnou odchylkou 8 g. Kontrolor náhodně vybírá 9 balíčků z jedné série a zjišťuje, zda jejich průměrná hmotnost je alespoň 999 g. Pokud ne, podnik musí zaplatit pokutu 20 000 Kč. Jaká je pravděpodobnost, že podnik bude muset zaplatit pokutu?

Řešení:

$$X \sim N(1002, 64), M \sim N\left(1002, \frac{64}{9}\right)$$

$$\begin{aligned} P(M \leq 999) &= P\left(\frac{M-1002}{\sqrt{\frac{64}{9}}} \leq \frac{999-1002}{\sqrt{\frac{64}{9}}}\right) = P\left(U \leq -\frac{9}{8}\right) = \Phi\left(-\frac{9}{8}\right) = 1 - \Phi\left(\frac{9}{8}\right) = 1 - \Phi(1,125) = \\ &= 1 - \Phi(1,13) = 1 - 0,87076 = 0,12924 \end{aligned}$$

Pravděpodobnost, že podnik bude platit pokutu, je asi 12,9%.

6.4. Věta: Vzorce pro meze 100(1- α)% empirických intervalů spolehlivosti pro μ a σ^2

- Interval spolehlivosti pro μ , když σ^2 známe
- Interval spolehlivosti pro μ , když σ^2 neznáme
- Interval spolehlivosti pro σ^2 , když μ neznáme
- Interval spolehlivosti pro σ^2 , když μ známe

6.5. Příklad: 10 krát nezávisle na sobě byla změřena jistá konstanta μ . Výsledky měření byly: 2 1,8 2,1 2,4 1,9 2,1 2 1,8 2,3 2,2. Tyto výsledky považujeme za číselné realizace náhodného výběru X_1, \dots, X_{10} z rozložení $N(\mu, \sigma^2)$, kde parametry μ, σ^2 neznáme. Najděte 95% empirický interval spolehlivosti pro μ , a to

- oboustranný,
- levostranný,
- pravostranný.

Řešení:

Vypočteme realizaci výběrového průměru: $m = 2,06$, výběrového rozptylu: $s^2 = 0,0404$ a výběrové směrodatné odchylky: $s = 0,2011$. Riziko α je 0,05. Jde o situaci popsanou v bodě (b), kde využíváme pivotovou statistiku T , která se řídí Studentovým rozložením $t(9)$. V tabulkách najdeme kvantil $t_{0,975}(9) = 2,2622$ pro oboustranný interval spolehlivosti a kvantil $t_{0,95}(9) = 1,8331$ pro jednostranné intervaly spolehlivosti.

$$\text{ad a) } d = m - \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1) = 2,06 - \frac{0,2011}{\sqrt{10}} 2,2622 = 1,92$$

$$h = m + \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1) = 2,06 + \frac{0,2011}{\sqrt{10}} 2,2622 = 2,20$$

$1,92 < \mu < 2,20$ s pravděpodobností aspoň 0,95.

$$\text{ad b) } d = m - \frac{s}{\sqrt{n}} t_{1-\alpha}(n-1) = 2,06 - \frac{0,2011}{\sqrt{10}} 1,8331 = 1,94$$

$1,94 < \mu$ s pravděpodobností aspoň 0,95.

$$\text{ad c) } h = m + \frac{s}{\sqrt{n}} t_{1-\alpha}(n-1) = 2,06 + \frac{0,2011}{\sqrt{10}} 1,8331 = 2,18$$

$\mu < 2,18$ s pravděpodobností aspoň 0,95.

6.6. Definice: Jednotlivé typy testů pro parametry normálního rozložení

a) Jednovýběrový z-test

b) Jednovýběrový t-test

c) Test o rozptylu

6.7. Příklad: Podle údajů na obalu čokolády by její čistá hmotnost měla být 125 g. Výrobce dostal několik stížností od kupujících, ve kterých tvrdili, že hmotnost čokolád je nižší než deklarovaných 125 g. Z tohoto důvodu oddělení kontroly náhodně vybralo 50 čokolád a zjistilo, že jejich průměrná hmotnost je 122 g a směrodatná odchylka 8,6 g. Za předpokladu, že hmotnost čokolád se řídí normálním rozložením, můžeme na hladině významnosti 0,01 považovat stížnosti kupujících za oprávněné?

Řešení: X_1, \dots, X_{50} je náhodný výběr z $N(\mu, \sigma^2)$. Testujeme hypotézu $H_0: \mu = 125$ proti levostranné alternativě $H_1: \mu < 125$. Protože neznáme rozptyl σ^2 , použijeme jednovýběrový t-

test. Realizace testového kritéria $t_0 = \frac{m - c}{\frac{s}{\sqrt{n}}} = \frac{122 - 125}{\frac{8,6}{\sqrt{50}}} = -2,4667$. Kritický obor: $W =$

$(-\infty, -t_{1-\alpha}(n-1)) = (-\infty, -t_{0,99}(49)) = (-\infty, -2,4044)$. Protože se testová statistika realizuje v kritickém oboru, zamítáme nulovou hypotézu na hladině významnosti 0,01. Stížnosti kupujících tedy lze považovat za oprávněné.

6.8. Definice: Definice rozdílového náhodného výběru na základě náhodného výběru z dvourozměrného normálního rozložení.

6.9. Věta: Vzorec pro meze $100(1-\alpha)\%$ empirického intervalu spolehlivosti pro střední hodnotu rozdílového náhodného výběru

6.10. Příklad: Dvěma rozdílnými laboratorními metodami se zjišťoval obsah chemické látky v roztoku (v procentech). Bylo vybráno 5 vzorků a proměřeno oběma metodami. Výsledky měření jsou obsaženy v tabulce:

číslo vzorku	1	2	3	4	5
1. metoda	2,3	1,9	2,1	2,4	2,6
2. metoda	2,4	2,0	2,0	2,3	2,5

Za předpokladu, že data mají normální rozložení, sestrojte 90% empirický interval spolehlivosti pro rozdíl středních hodnot výsledků obou metod.

Řešení:

Přejdeme k rozdílovému náhodnému výběru, jehož realizace jsou: -0,1 -0,1 0,1 0,1 0,1.

Vypočteme $m = 0,02$, $s^2 = 0,012$, $s = 0,109545$. Předpokládáme, že tato data pocházejí z normálního rozložení $N(\mu, \sigma^2)$. Vypočteme meze 90% oboustranného intervalu spolehlivosti pro μ při neznámém σ :

$$d = m - \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1) = 0,02 - \frac{0,109545}{\sqrt{5}} t_{0,95}(4) = 0,02 - \frac{0,109545}{\sqrt{5}} 2,1318 = -0,0844$$

$$h = m + \frac{s}{\sqrt{n}} t_{1-\alpha/2}(n-1) = 0,02 + \frac{0,109545}{\sqrt{5}} t_{0,95}(4) = 0,02 + \frac{0,109545}{\sqrt{5}} 2,1318 = 0,1244$$

$-0,0844 < \mu < 0,1244$ s pravděpodobností aspoň 0,9.

6.11. Definice: Definice párového t-testu

6.12. Příklad: V následující tabulce jsou údaje o výnosnosti dosažené 12 náhodně vybranými firmami při investování do mezinárodního podnikání (veličina X) a do domácího podnikání (veličina Y):

č.firmy	1	2	3	4	5	6	7	8	9	10	11	12
X	10	12	14	12	12	17	9	15	9	11	7	15
Y	11	14	15	11	13	16	10	13	11	17	9	19

(Výnosnost je vyjádřena v procentech a představuje podíl na zisku vložených investic za rok.) Za předpokladu, že data pocházejí z dvourozměrného normálního rozložení, na hladině významnosti 0,1 testujte hypotézu, že neexistuje rozdíl mezi střední hodnotou výnosnosti investic do mezinárodního a zahraničního podnikání proti oboustranné alternativě.

Testování proved'te

a) pomocí intervalu spolehlivosti

b) pomocí kritického oboru.

(Pro úsporu času máte uvedeny realizace výběrového průměru $m = -1,3$ a výběrového rozptylu $s^2 = 4,78$ rozdílového náhodného výběru $Z_i = X_i - Y_i$, $i = 1, \dots, 12$.)

Řešení:

Testujeme $H_0: \mu = 0$ proti $H_1: \mu \neq 0$

ad a)

90% interval spolehlivosti pro střední hodnotu μ při neznámém rozptylu σ^2 má meze:

$$d = m - \frac{s}{\sqrt{n}} t_{0,95}(n-1) = -1,3 - \frac{\sqrt{4,78}}{\sqrt{12}} 1,7959 = -2,4677$$

$$h = m + \frac{s}{\sqrt{n}} t_{0,95}(n-1) = -1,3 + \frac{\sqrt{4,78}}{\sqrt{12}} 1,7959 = -0,1989$$

Protože číslo $c = 0$ neleží v intervalu $(-2,4677; -0,1989)$, H_0 zamítáme na hladině významnosti 0,1.

ad b)

$$\text{Vypočítáme realizaci testové statistiky } t_0 = \frac{m - c}{\frac{s}{\sqrt{n}}} = \frac{-1,3}{\frac{\sqrt{4,78}}{\sqrt{12}}} = -2,11085$$

Stanovíme kritický obor $W = (-\infty, -t_{0,95}(11)) \cup (t_{0,95}(11), \infty) = (-\infty, -1,7959) \cup (1,7959, \infty)$

Protože testová statistika se realizuje v kritickém oboru, H_0 zamítáme na hladině významnosti 0,1.