

## Testovanie hypotéz

**Uvedieme zoznam niektorých testov. Podrobnosti sú uvedené v časti \Program Files\GNU Octave 2.1.73\usr\share\octave\2.1.73\m\statistics\tests\**

**[pval, z] = z\_test (x, m, v, alt)** % test zhody strednej hodnoty vzorky so zadanou hodnotou pri známom rozptyle

**[pval, t, df] = t\_test (x, m, alt)** % t test zhody strednej hodnoty vzorky so zadanou hodnotou pri neznámom rozptyle

**[pval, z] = z\_test\_2 (x, y, v\_x, v\_y, alt)** % t test zhody strednej hodnoty dvoch vzoriek pri známych rozptyloch

**[pval, t, df] = t\_test\_2 (x, y, alt)** % t test zhody strednej hodnoty dvoch vzoriek pri neznámych rozptyloch

**[pval, f, df\_num, df\_den] = var\_test (x, y, alt)** % test zhody disperzií dvoch vzoriek

**[pval, Tsq] = hotelling\_test (x, m)** % test zhody strednej hodnoty vzorky so zadanou hodnotou

**[pval, Tsq] = hotelling\_test\_2 (x, y)** % test zhody stredných hodnôt dvoch vzoriek

**[pval, f, df\_b, df\_w] = anova (y, g)** % test ANOVA rozdielu populácií

**manova (Y, g)** % test MANOVA (viacrozmerný test ANOVA)

**[pval, t, df] = t\_test\_regression (y, X, R, r, alt)** % t test pre klasický regresný model

**t = cor\_test (X, Y, ALTERNATIVE, METHOD)** % testuje, či vzorky patria nekorelovaným populáciám

**[pval, ks] = kolmogorov\_smirnov\_test (x, dist, varargin)** % Kolmogorov-Smirnovov test rozdelenia

**[pval, ks, d] = kolmogorov\_smirnov\_test\_2 (x, y, alt)** % Kolmogorov-Smirnovov test zhody rozdelenia dvoch vzoriek

**[pval, chisq] = run\_test (x)** % test nezávislosti údajov

**[pval, z] = prop\_test\_2 (x1, n1, x2, n2, alt)** % test rovnosti pravdepodobnosti dvoch vzoriek

**[pval, z] = u\_test (x, y, alt)** % Mannov-Whitneyov znamienkový test

**[pval, chisq, df] = chisquare\_test\_independence (X)** %  $\chi^2$  test nezávislosti

Označenie:

"!=" alebo "<>" použijeme pre obojstranný test

">" použijeme pre pravostranný test

"<" použijeme pre ľavostranný test

Možnosti použitia niektorých testov uvedieme v nasledujúcich príkladoch.

**Príklad** Zo základného súboru s normálnym rozdelením sme urobili náhodný výber s realizáciami: 22,4; 28,0; 20,1; 27,4; 23,9; 24,8; 26,4; 27,0; 25,4; 25,6 . Určme obojstranný (ľavostranný, pravostranný) test pre strednú hodnotu  $\mu = 25,1$ , resp.  $\mu = 30$  pri neznámom rozptyle ( $H_0 : \mu = \mu_0$ ).

**Riešenie:** Testujeme základnú hypotézu  $H_0 : \mu = \mu_0$  proti alternatívnej hypotéze  $H_1 : \mu \neq \mu_0$

```
octave:25> x=[22.4,28.0,20.1,27.4,23.9,24.8,26.4,27.0,25.4,25.6]
```

```
x =
```

```
22.400 28.000 20.100 27.400 23.900 24.800 26.400 27.000 25.400 25.600
```

```
octave:26> n=length(x)
```

```
n = 10
```

```
octave:27> std(x)
```

% modifikovaná výberová smerodajná odchýlka

```
ans = 2.4304
```

```
octave:28> mean(x)
```

% výberový priemer

```
ans = 25.100
```

```
octave:32> t_test(x,25.1,"<>")
```

% "<>" – obojstranná alternatívna hypotéza  $H_1 : \mu \neq \mu_0$

```
pval: 1
```

```
ans = 1
```

% pravdepodobnosť, že  $\mu = 25,1$

```
octave:33> t_test(x,30,"<>")
```

```
pval: 0.000128914
```

```
ans = 0.00012891
```

% pravdepodobnosť, že  $\mu = 30$

```
octave:34> t_test(x,30,"<")
```

% "<" – ľavostranná alternatívna hypotéza  $H_1 : \mu < \mu_0$

```
pval: 6.44568e-05
```

```
ans = 6.4457e-05
```

% pravdepodobnosť, že  $\mu > 30$

```
octave:35> t_test(x,30,">")
```

% ">" – pravostranná alternatívna hypotéza  $H_1 : \mu > \mu_0$

```
pval: 0.999936
```

```
ans = 0.99994
```

% pravdepodobnosť, že  $\mu < 30$

**Príklad** Vzorky chemickej látky sme analyzovali dvoma metódami: a) polarografickou metódou, b) titračnou metódou. Dostali sme tieto výsledky:

a)	38,2	36,4	37,7	36,1	37,9	37,8	-	-
b)	39,5	38,7	37,8	38,6	39,2	39,1	38,9	39,2

Na hladine významnosti 5% testujte hypotézu o rovnocennosti oboch metód, t.j.  $H_0 : \mu_1 = \mu_2$  proti  $H_1 : \mu_1 \neq \mu_2$ .

*Riešenie:*

```
octave:9> [pval,t,df]=t_test_2(x,y,"<>")
pval = 0.0015087          % Záver: Hypotézu  $H_0$  zamietame.
```

```
t = -4.0864              % testovacia charakteristika
df = 12
```

**Príklad** Určitý výrobok sa vyrába dvoma technologickými postupmi. Kontrolným meraním sme zistili pri náhodne vybraných výrobkoch tieto údaje o určitej kvalitatívnej vlastnosti.

Postup A: 13,15,15,14,13.

Postup B: 13,12,14,13,13,15,16.

Rozhodnime, či sa disperzia kvality pri oboch technologických procesoch významne líši.

Predpokladáme normálne rozdelenie oboch výberových súborov.

*Riešenie:* Testujeme  $H_0 : \sigma_1^2 = \sigma_2^2$  proti  $H_0 : \sigma_1^2 \neq \sigma_2^2$ .

```
octave:1> A=[13,15,15,14,13];B=[13,12,14,13,13,15,16];
octave:2> var_test(A,B,"<>")
pval: 0.554873
ans = 0.55487 Teda pravdepodobnosť pval je taká, že hypotézu  $H_0$  nezamietame.
```

```
octave:3> sA=std(A)          % modifikovaná výberová smerodajná odchýlka
```

```
sA = 1
```

```
octave:4> sB=std(B)
```

```
sB = 1.3801
```

```
octave:5> F=sA^2/sB^2          % testovacia charakteristika  $F = (sA)^2 / (sB)^2$ 
```

```
F = 0.52500
```

```
octave:6> b=f_inv(0.025,4,6)
```

```
b = 0.10873
```

```
octave:7> c=f_inv(0.975,4,6)
```

```
c = 6.2272
```

```
octave:8> bsigma=sqrt(b)
```

```
bsigma = 0.32974
```

```
octave:9> csigma=sqrt(c)
```

```
csigma = 2.4954
```

Teda kritická oblasť pre  $\sigma^2$  je

$$K_\alpha = (0; f_{\alpha/2, n_1-1, n_2-1}) \cup (f_{1-\alpha/2, n_1-1, n_2-1}; \infty) = (0; b) \cup (c; \infty) = (0; 0.10873) \cup (6.2272; \infty).$$

Pretože  $F \notin K_\alpha$ , hypotézu  $H_0$  nezamietame, t.j. rozdiel vo variabilite hrúbky dodávanej suroviny od oboch dodávateľov považujeme za štatisticky nevýznamný.