

Test dobré shody

Test dobré shody umožňuje ověřit, zda má náhodná veličina určité předem dané rozdělení pravděpodobnosti.

Často se používá **Pearsonův chí-kvadrát test**.

Použití: Ověření, že naměřené četnosti odpovídají očekávaným četnostem.

Omezení: Test lze použít za předpokladu, že všechny očekávané četnosti jsou aspoň 5. V opačném případě je třeba skupiny s malými očekávanými četnostmi sloučit.

Nulová hypotéza: Naměřené četnosti odpovídají očekávaným četnostem.

Postup

- 1) Obor všech hodnot se rozdělí na k nepřekrývajících se částí.
- 2) Pro každou část se stanoví pravděpodobnost p_i , že náhodná veličina nabyde hodnoty z i -té části.
- 3) Proveďte se N pokusů a zjistí se, kolikrát z těchto pokusů nabyla náhodná veličina hodnoty z 1., 2., ... k -té části. Tyto četnosti se označí X_1, X_2, \dots, X_k .
- 4) Porovnají se očekávané četnosti v jednotlivých částech (Np_i) se skutečnými četnostmi (X_i) pomocí vzorce (testovací kritérium):

$$\chi^2 = \frac{\sum_{i=1}^k (X_i - Np_i)^2}{Np_i}$$

- 5) Jestliže bylo rozdělení dáno včetně všech parametrů, je počet stupňů volnosti $k-1$; jestliže byl některý parametr neznámý, snižuje se počet stupňů volnosti za každý neznámý parametr.
- 6) Kritická hodnota: =CHISQ.INV.RT(α ; $k-1$)
- 7) Je-li testovací kritérium větší než kritické hodnota, nulovou hypotézu zamítneme na hladině významnosti α .

Příklad 1:

Při 48 hodech kostkou padly následující četnosti hodnot:

hodnota	1	2	3	4	5	6
četnost	10	8	8	4	6	12

Ověřte, zda je kostka pravidelná.

$$p_i = 1/6; Np_i = 48/6 = 8$$

očekávaná čet.	8	8	8	8	8	8
$\frac{(X_i - Np_i)^2}{Np_i}$	0,5	0	0	2	0,5	2

$$\chi^2 = 5$$

$$\chi_{krit}^2(0,05; 5) = 11,07$$

Příklad 2

U studentů byla zjištěna dominantní ruka. Zjistěte, zda se zastoupení leváků liší podle pohlaví.

	A	B	C
1		muž	žena
2	levák	11	6
3	pravák	73	49
4	suma	84	55

	A	B	C	D
1	realita	muž	žena	suma
2	levák	11	6	17
3	pravák	73	49	122
4	suma	84	55	139
5				
6	očekávání	muž	žena	suma
7	levák	10.3	6.7	17
8	pravák	73.7	48.3	122
9	suma	84	55	139
10				
11	výpočet	muž	žena	suma
12	levák	0.051	0.078	0.130
13	pravák	0.007	0.011	0.018
14	suma	0.059	0.089	0.148

	A	B	C	D
1	realita	muž	žena	suma
2	levák	11	6	=SUMA(B2:C2)
3	pravák	73	49	=SUMA(B3:C3)
4	suma	=SUMA(B2:B3)	=SUMA(C2:C3)	=SUMA(B4:C4)
5				
6	očekávání	muž	žena	suma
7	levák	=D2*B\$4/\$D\$4	=D2*C\$4/\$D\$4	=SUMA(B7:C7)
8	pravák	=D3*B\$4/\$D\$4	=D3*C\$4/\$D\$4	=SUMA(B8:C8)
9	suma	=SUMA(B7:B8)	=SUMA(C7:C8)	=SUMA(B9:C9)
10				
11	výpočet	muž	žena	suma
12	levák	=(B2-B7)^2/B7	=(C2-C7)^2/C7	=SUMA(B12:C12)
13	pravák	=(B3-B8)^2/B8	=(C3-C8)^2/C8	=SUMA(B13:C13)
14	suma	=SUMA(B12:B13)	=SUMA(C12:C13)	=SUMA(B14:C14)

$$\chi^2 = 0,148$$

$$\chi_{krit}^2(0,05; 1) = 3,84$$

Příklad 3

Byly změřeny výšky studentů ZE I v letech 2022 a 2023.
Otestujte, zda mají normální rozdělení.

164,0;164,0;168,0;168,5;168,5;168,5;169,0;169,4;170,0;170,0;170,0;170,0;
171,0;171,1;171,5;172,0;172,0;173,5;174,0;174,0;174,0;174,5;174,5;175,0;
175,0;175,5;175,5;175,5;175,5;176,0;176,0;176,5;176,5;177,0;177,0;177,0;
177,5;177,5;177,5;177,5;177,5;178,0;178,0;178,0;178,5;178,5;178,5;178,5;
178,5;179,0;179,5;179,5;180,0;180,0;180,0;180,0;180,0;180,0;180,0;180,0;
180,5;180,5;180,5;180,5;180,5;181,5;181,5;181,5;181,5;182,0;182,0;182,0;
182,0;182,0;182,5;182,5;182,5;182,5;183,0;183,0;183,5;183,5;184,0;184,0;
184,5;184,5;184,5;184,5;185,0;185,0;185,0;185,0;185,5;185,5;186,5;186,5;
186,5;186,5;187,5;187,5;188,0;188,0;188,5;188,5;189,0;189,0;190,0;190,0;
190,0;190,0;190,0;190,5;190,5;191,0;192,0;192,0;192,5;192,5;193,5;195,0;
195,0;196,5;198,5;199,0;201,5

$$n = 125; \mu = 181.14 \text{ cm}; \sigma = 7.49 \text{ cm}$$

výška [cm]	četnost
162	2
166	6
170	10
174	23
178	28
182	25
186	12
190	13
194	3
198	3
202	

výška [cm]	četnost
162	8
170	10
174	23
178	28
182	25
186	12
190	13
194	6
202	

Příklad 3

$n=125$; $\mu=181.14$ cm; $\sigma=7.49$ cm

výška [cm]	exper. četnost	oček. četnost
162	8	7.9
170	10	12.7
174	23	20.9
178	28	26.0
182	25	24.5
186	12	17.5
190	13	9.4
194	6	5.0
202		

$$\chi^2 = 4,20 \quad \chi_{krit}^2(0,05; 5) = 11,07$$

Excel - CHISQ.TEST()

Excel obsahuje i funkci =CHISQ.TEST(naměřené; očekávané), která po zadání matice naměřených a matice očekávaných hodnot vypočítá p -hodnotu. Zde není třeba přemýšlet o stupních volnosti.

Použijme data z příkladu 1:

	A	B	C	D	E	F	G
1	hodnota	1	2	3	4	5	6
2	naměřené	10	8	8	4	6	12
3	očekávané	8	8	8	8	8	8

=CHISQ.TEST(B2:G2;B3:G3)
vrátí $p = 0,42$

Použijme data z příkladu 2:

	A	B	C	D
1	realita	muž	žena	suma
2	levák	11	6	17
3	pravák	73	49	122
4	suma	84	55	139
5				
6	očekávání	muž	žena	suma
7	levák	10.3	6.7	17
8	pravák	73.7	48.3	122
9	suma	84	55	139
10				
11	výpočet	muž	žena	suma
12	levák	0.051	0.078	0.130
13	pravák	0.007	0.011	0.018
14	suma	0.059	0.089	0.148

=CHISQ.TEST(B2:C3;B7:C8)
vrátí $p = 0,700$