

Robert Hončariv a Kveta Hončarivová
Expertný systém biometrickej genetiky.
UPJŠ Košice 2000

Anotácia.

Encyklopedické spracovanie problematiky problémov a metodík

plánovania, zakladania a hodnotenie biologických pokusov, kvantitatívnej a populačnej genetiky. V súčasnosti obsahuje 2000 hesiel v päť jazyčnej verzii s príkladmi. Na CD Rome . dotazy, pripomienky a námety na e-mailovej adrese honcariv@kosice.upjs.sk.

DEMO 2

<i>vážený klzavý priemer</i>	vážený klouzavý průměr	Gewogenes Gleitendes Mittel	weighted moving average	взвешаное скользящее средне
<i>aritmetický priemer sumárnej skupiny</i>	aritmetický průměr sumární skupiny			
<i>extrémna hodnota</i>	odlehlá hodnota, extrémní hodnota	Ausreißer	outlier	выброс, аномальное значение
<i>arcus sinusová transformácia</i>	arcsinová transformace	Arc-sinus Transformation	arc sine transformation	аарсинус преобразование
<i>variáčné rozpätie</i>				
<i>Dixonov test</i>	Dixonův test	Dixon-Test	Dixon test	критерий Диксона
<i>klzavý priemer</i>	klouzavý průměr	Gleitendes Mittel	moving average	скользящее средне
<i>relatívna početnosť</i>	relativní četnost, poměrná četnost	Relative Häufigkeit	relative frequency, proportional frequency	относительная частота
<i>percentá</i>	relativní četnost, poměrná četnost	Relative Häufigkeit	relative frequency, proportional frequency	относительная частота
<i>t-test (relatívne frekvencie)</i>	t - test - relativní početnosť	t- test - Relative Häufigkeit	t- test - relative frequency, proportional frequency	t-критерий- относительная частота
<i>F test</i>	F test	F Test	F test	F-критерий
<i>relatívna chyba</i>	relativní chyba	Relativer Fehler	relative error	относительная ошибка
<i>kvadratická odchýlka</i>	kvadratická odchylka	Quadratische Abweichung	square deviation	квадратическое отклонение
<i>variáčny koeficint</i>	variáčny koeficient	Variationkoeffizient	coefficient of variation	коэффициент вариации
<i>Cochranov test</i>	Cochranov test	Cochran-test	Cochran's test	критерий Кохрана
<i>kvadratická odchýlka</i>	kvadratická odchylka	Quadratische Abweichung		

square deviation
 квадратическое отклонение
 synonymum pre rozptyl

$$s^2 = \frac{\sum (x - \bar{x})^2}{n - 1}$$

variačný koeficient
 variačný koeficient
 Variationkoeffizient
 coefficient of variation
 коэффициент вариации

predstavuje percentuálne vyjadrenie smerodajnej odchýlky

$$C.V = \frac{s \cdot 100}{\bar{x}}$$

často sa vyjadruje v percentách.. V biometrike sa používa relatívna chyba pre vyjadrenie variačného koeficienta (*variačný koeficient*).

relatívna chyba
 relativní chyba
 Relativer Fehler
 relative error
 относительная ошибка

predstavuje absolútnu hodnotu rozdielu medzi pozorovanou hodnotou a hodnotou a

$$\frac{|x - a|}{a}$$

F test

F test
 F Test
 F test
 F-критерий

pomocou tohto testu sa hodnotia významnosť rozdielov aritmetických priemerov podľa vzťahu

$$F = \frac{(\bar{x}_1 - \bar{x}_2)^2}{s_z^2} \cdot \frac{n_1 n_2}{n_1 + n_2} \geq F_{st} \begin{cases} d_{f_1} = 1 \\ d_{f_2} = n_1 + n_2 - 2 \end{cases}$$

počítačové procedúra
 počítačová procedúra
 Časť softveru, ktorá realizuje určitý výpočet..

Cochranov test

Cochranov test
 Cochran-test
 Cochran's test
 критерий Кохрана

Cochranov test sa používa pre porovnaní heterogenity alebo homogenity viac ako dvoch variancií. Používa sa pri extrémne veľkých varianciách podľa vzťahu

$$G_{\max} = \frac{s_{\max}^2}{s_1^2 + s_2^2 + \dots + s_k^2}$$

hodnoty sú tabelované.

percentá

vid' relatívna početnosť

relatívna početnosť

relativní četnost, poměrná četnost

Relative Häufigkeit

relative frequency, proportional frequency

относительная частота

Vyjadrenie kvalitatívneho znaku formou relatívni početnosti percentami (*percentá*)

Je vyjadrená hodnotami v intervale 0 - 1. Stredná chyba relatívnej početnosti sa vypočíta zo vzťahu

$$s_p = \frac{\bar{\sigma}}{\sqrt{n}} = \sqrt{\frac{pq}{n}}$$

p - pomer prípadov, s kladný znakom; q - pomer prípadov so záporným znakom.

vážený klzavý priemer

vážený klouzavý průměr

Gewogenes Gleitendes Mittel

weighted moving average

взвешанное скользящее среднее

sa používa pri presnejšom *vyrovňovaní krivky* pre regresnej závislosti (regresná závislosť)

Prvú a poslednú hodnotu grafu je potrebné vypočítať na základe vzťahu

$$x_{+1} = \frac{2x_1 + x_2 - x_4}{2}$$

extrémna hodnota

odlehlá hodnota, extrémní hodnota

Ausreißer

outlier

выброс, аномальное значение

Za extrémnu hodnotu sa považuje hodnota, ktorá sa vymyká variačnému rozpätiu (*variačné rozpätie*). Môže vzniknúť v dôsledku technickej chyby, ale reálne existuje vo výberovom súbore (*výberový súbor*) napríklad v dôsledku mutácie (*mutácia*). Jej výskyt vo výberovom súbore môže značným spôsobom ovplyvniť hodnotu aritmetického priemeru (*aritmetický priemer*). O jej existencii je potrebné sa objektívne presvedčiť. Možno použiť Cochranov Test (*Cochranov Test*) alebo Dixonov test (*Dixonov test*). Počítačové procedúry (*počítačová procedúra*) riešia túto úlohu vypočítaním normovanej odchýlky (*normovaná odchýlka*). O extrémnej hodnote sa možno presvedčiť aj jednoduchým ručným výpočtom podľa algoritmu. Existuje *relatívna extrémna hodnota* a *absolútna extrémna hodnota*

$$T = \frac{\hat{x} - \bar{x}}{s} \geq T_{st}$$

T kritérium artefaktu

\hat{X} hodnota predpokladaného artefaktu

T_{st} tabulková hodnota štandardu
 s smeradajná odchýlka

n	T_{st}	n	T_{st}	n	T_{st}	n	T_{st}
2	2.0	16 - 20	2.4	47 - 66	2.8	125 - 174	3.2
3 - 4	2.1	21 - 28	2.5	67 - 84	2.9	175 - 349	3.3
5 - 9	2.2	29 - 24	2.6	85 - 104	3.0	350 - 599	3.4
10 - 15	2.3	35 - 46	2.7	105 - 124	3.1	600 - 1500	3.5

Dixonov test

Dixonův test

Dixon-Test

Dixon test

критерий Диксона

Testové kritérium na stanovenie extrémnej hodnoty. Používa sa pri výberových súboroch s na väčším ako 7

arcus sinusová transformácia

arcsinová transformace

Arc-sinus Transformation

arc sine transformation

арксинус преобразование

V biologickom experimente často stojíme pred problémom vyjadrovať výsledky v relatívnych hodnotách percent. Pri porovnávaní a testovaní výsledku nemôžeme použiť *parametrické testy* *t*-test, analýzu variancie (*analýza variancie*), pretože percenta nepodliehajú normálnemu rozdeleniu. Preto sa hodnoty musia transformovať na *uhlové* alebo *radianové stupne*.

Pri hodnotení rozdielov percent pomocou *t*-testu, môže dôjsť nesprávnemu hodnoteniu

Príklad 1: Pri sledovaní frekvencie mutácií sa zistilo, že v jednej populácii sa vyskytuje z 5000 jedincov v 3 prípadoch, v druhej populácii z 500 prípadov tiež 3 prípady. Je treba zistiť, či existuje štatisticky významný rozdiel medzi sledovanými populáciami. Podľa *t*-testu rozdielu percent (*t*-test rozdielu percent) bude výpočet nasledovný

$$t_d = \frac{0.006 - 0.0006}{\sqrt{\frac{0.006 \cdot 0.994}{499} + \frac{0.0006 \cdot 0.9994}{4999}}} = 1.5 < 1.96$$

To znamená, že rozdiel nie je štatisticky významný. Iný výsledok dostaneme pri hodnotení po uhlovej transformácii (vid' príklad 2)

Uhlové stupne sa vyjadrujú arcus sinusovou transformáciou.

$$\varphi^0 = \arcsin \sqrt{p}$$

príčom nezávisí od hodnoty uhla, ale od veľkosti výberového súboru

$$s_\varphi = \frac{1}{\sqrt{n}}$$

hodnotu

$$s_\varphi^2 = \frac{1}{n}$$

preto ju možno ľahko vyhľadať v tabuľke inverzných hodnôt (*tabuľka inverzných hodnôt*)

Vhodnejšie je použiť transformáciu na radiánové stupne

$$x = 2 \frac{\pi}{180} \arcsin \sqrt{p}$$

$$x = 0.0349066 \arcsin \sqrt{p}$$

hodnoty radiánových stupňov možno nájsť v štatistických tabuľkách, ktoré značne uľahčia prácu. Po realizovaní testu, ak je to nevyhnutné, je potrebné previesť radiánové stupne späť na percentá.

Tabuľka radiánových stupňov

%	φ	%	φ	%	φ	%	φ	%	φ
0.01	0.2003	0.11	0.6761	0.21	0.9521	0.31	1.1810	0.41	1.3898
0.02	0.2838	0.12	0.7075	0.22	0.9764	0.32	1.2025	0.42	1.4101
0.03	0.3482	0.13	0.7377	0.23	1.0004	0.33	1.2239	0.43	1.4303
0.04	0.4027	0.14	0.7670	0.24	1.0240	0.34	1.2451	0.44	1.4505
0.05	0.4510	0.15	0.7954	0.25	1.0472	0.35	1.2661	0.45	1.4706
0.06	0.4949	0.16	0.8230	0.26	1.0701	0.36	1.2870	0.46	1.4907
0.07	0.5355	0.17	0.8500	0.27	1.0928	0.37	1.3078	0.47	1.5108
0.08	0.5735	0.18	0.8764	0.28	1.1152	0.38	1.3284	0.48	1.5308
0.09	0.6094	0.19	0.9021	0.29	1.1374	0.39	1.3490	0.49	1.5508
0.10	0.6435	0.20	0.9273	0.30	1.1593	0.40	1.3694	0.50	1.5708

%	φ	%	φ	%	φ	%	φ	%	φ
0.51	1.5908	0.61	1.7926	0.71	2.0042	0.81	2.2395	0.91	2.6322
0.52	1.6108	0.62	1.8132	0.72	2.0264	0.82	2.2653	0.92	2.5681
0.53	1.6308	0.63	1.8338	0.73	2.0488	0.83	2.2916	0.93	2.6061
0.54	1.6509	0.64	1.8546	0.74	2.0715	0.84	2.3186	0.94	2.6467
0.55	1.6710	0.65	1.8755	0.75	2.0944	0.85	2.3462	0.95	2.6906
0.56	1.6911	0.66	1.8965	0.76	2.1177	0.86	2.3746	0.96	2.7339
0.57	1.7113	0.67	1.9177	0.77	2.1412	0.87	2.4032	0.97	2.7934
0.58	1.7315	0.68	1.9391	0.78	2.1652	0.88	2.4341	0.98	2.8578
0.59	1.7618	0.69	1.9606	0.79	2.1895	0.89	2.4655	0.99	2.9413
0.60	1.7722	0.70	1.9623	0.80	2.2143	0.90	2.4981	1.00	3.1416

Príklad 1:

Pri pôsobení herbicídneho prípravku z 1000 rastlín uhynulo 900.

Aké je možné predpokladať hranice pôsobenia tohto preparátu s pravdepodobnosťou P = 0.99?

n = 1000; t = 2.96 ;

koeficient nepresnosti sa rovná

$$\Delta = \frac{2.6}{\sqrt{1000}} = 0.082$$

$$p = 0.90$$

$$\varphi_{ab} = 2.4981$$

$$\bar{\varphi} = 2.4981 \pm 0.082 = \underline{2.5801} \Leftrightarrow 2.4981 \text{C}$$

späťne podľa tabuliek

$$\phi_{\max} = 2.5801 \approx 0.93$$

$$\phi_{\min} = 2.4161 \approx 0.88$$

To znamená, že možno očakávať účinnosť v hraniciach 88 - 93 percent.

Príklad 2.

Zadanie príkladu 1

F-kriterium vypočítame za vzťahu

$$F_d = \frac{p_1 - p_2}{\phi_1 - \phi_2} \frac{n_1 n_2}{n_1 + n_2} \geq F_{tab}$$

pri hodnotách stupňov voľnosti (*stupeň voľnosti*)

$$df_1 = 1; df_2 = n_1 + n_2 - 2$$

riešenie:

$$n_1 = 5000 \quad p_1 = \frac{3}{5000} = 0.0006 \quad \phi_1 = 0.049$$

$$n_2 = 500 \quad p_2 = \frac{3}{500} = 0.006 \quad \phi_2 = 0.155$$

$$d = 0.106 \quad d^2 = 0.0112$$

$$F_d = 0.0112 \frac{5000 \cdot 500}{5500} = 5.1$$

$$F_{tab} = 3.8$$

to znamená, že rozdiel je štatisticky významný na hladine významnosti $P = 0.05$

aritmetický priemer súmarnej skupiny

aritmetický priemer sumárnej skupiny

$$\bar{x}_{\Sigma} = \frac{\sum n_i \bar{x}_i}{\sum n_i}$$

smerodatná odchýlka priemeru sumárnej skupiny

$$s_{\Sigma} = \sqrt{\frac{\sum (n_i - 1) s_i^2 + \sum n_i (\bar{x}_i - \bar{x}_{\Sigma})^2}{\sum n_i - 1}}$$

