Correlation الارتباط

Simple Correlation :الارتباط البسيط

وهو مقياس لتحديد طبيعة العلاقة وقوتها بين متغيرين مستقلين مثل x و y، وتسمى العلاقة طردية او موجبة اذا كانت زيادة قيم أحد المتغيرين يصحبها زيادة قيم المتغير الاخر، واذا اخذت القيم اتجاهين متعاكسين تسمى العلاقة عكسية او سالبة.

Estimation of Correlation Coefficient :تقدير معامل الارتباط

ان معامل الارتباط والذي يرمز له بالرمز r لعينة عشوائية حجمها n لمتغيرين (x و y)

$$(x_1, y_1), (x_2, y_2), \ldots (x_n, y_n)$$

يحسب من المعادلة التالية

$$r = \frac{\sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}}{\sqrt{(\sum x_i^2 - \frac{(\sum x_i)^2}{n})(\sum y_i^2 - \frac{(\sum y_i)^2}{n})}}$$

تتراوح قيمة معامل الارتباط (r) بين +1 و -1 أو أي قيمة بينهما ($1 \geq r \geq -1$)، وهاتين القيمتين هما درجتا الارتباط التام، الطردي التام (+1) والعكسي التام (-1)، وتجدر الإشارة الى ان معامل الارتباط بين متغيرين هو مقياس وصفي للترابط الخطي بينهما، وعندما تكون r= 0 فان ذلك يعني عدم وجود ترابط خطي بينهما وليس بالضرورة عدم وجود علاقة بينهما.

Dr. Ahmed Shehab

مثال: احسب معامل الارتباط للبيانات التالية والتي تمثل طول وعرض الورقة لنبات ما.

عرض الورقة (x)	13	19	13	18	14	17	14	17	15	16
طول الورقة (y)	15	22	13	20	13	20	15	19	15	18

الحل: لأيجاد معامل الارتباط (r) حسب المعادلة السابقة نرتب البيانات في جدول كما يلي:

x_i	y_i	x_iy_i	x_i^2	y_i^2
13	15	195	169	225
19	22	418	361	484
13	13	169	169	169
18	20	360	324	400
14	13	182	196	169
17	20	340	289	400
14	15	210	196	225
17	19	323	289	361
15	15	225	225	225
16	18	288	256	324
$\sum x_i = 156$	$\sum y_i = 170$	$\sum x_iy_i = 2710$	$\sum x_i^2 = 2474$	$\sum y_i^2 = 2982$

$$r = \frac{\sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}}{\sqrt{(\sum x_i^2 - \frac{(\sum x_i)^2}{n})(\sum y_i^2 - \frac{(\sum y_i)^2}{n})}}$$

$$r = \frac{2710 - \frac{(156)(170)}{10}}{\sqrt{(2474 - \frac{(156)^2}{10})(2982 - \frac{(170)^2}{10})}}$$

$$= 0.95$$

نوع العلاقة: طردية (موجبة)

Dr. Ahmed Shehab

اختبار معنوية الارتباط:

ان الحكم على المعنوية الإحصائية لمعامل الارتباط يعتمد على اختبارات إحصائية محددة، وذلك بمقارنة القيمة المطلقة لـ r المحسوبة مع قيمة r الجدولية عند درجة حرية n-2، حيث n = عدد ازواج المشاهدات، وعند مستوى الاحتمالية المطلوب، فاذا كانت قيمة r المحسوبة اكبر من او تساوي r الجدولية فان هذا يعني ان الارتباط معنوي اما اذا كانت قيمة r المحسوبة اقل من r الجدولية فان هذا يعني ان الارتباط غير معنوي ففي المثال السابق نستخرج قيمة r الجدولية من جدول r عند درجة حرية 8 ومستوى احتمالية 0.05 فنجدها تساوي 0.63

الاستنتاج: بما أن قيمة r المحسوبة (0.95) هي اكبر من قيمة r الجدولية (0.63) نستنتج ان الارتباط معنوي.

df\α	0.2	0.1	0.05	0.02	0.01	0.001
1	0.951057	0.987688	0.996917	0.999507	0.999877	0.999999
2	0.800000	0.900000	0.950000	0.980000	0.990000	0.999000
3	0.687049	0.805384	0.878339	0.934333	0.958735	0.991139
4	0.608400	0.729299	0.811401	0.882194	0.917200	0.974068
5	0.550863	0.669439	0.754492	0.832874	0.874526	0.950883
6	0.506727	0.621489	0.706734	0.788720	0.834342	0.924904
7	0.471589	0.582206	0.666384	0.749776	0.797681	0.898260
8	0.442796	0.549357	0.631897	0.715459	0.764592	0.872115
9	0.418662	0.521404	0.602069	0.685095	0.734786	0.847047
10	0.398062	0.497265	0.575983	0.658070	0.707888	0.823305
11	0.380216	0.476156	0.552943	0.633863	0.683528	0.800962
12	0.364562	0.457500	0.532413	0.612047	0.661376	0.779998
13	0.350688	0.440861	0.513977	0.592270	0.641145	0.760351
14	0.338282	0.425902	0.497309	0.574245	0.622591	0.741934
15	0.327101	0.412360	0.482146	0.557737	0.605506	0.724657
16	0.316958	0.400027	0.468277	0.542548	0.589714	0.708429
17	0.307702	0.388733	0.455531	0.528517	0.575067	0.693163
18	0.299210	0.378341	0.443763	0.515505	0.561435	0.678781
19	0.291384	0.368737	0.432858	0.503397	0.548711	0.665208
20	0.284140	0.359827	0.422714	0.492094	0.536800	0.652378
21	0.277411	0.351531	0.413247	0.481512	0.525620	0.640230
22	0.271137	0.343783	0.404386	0.471579	0.515101	0.628710
23	0.265270	0.336524	0.396070	0.462231	0.505182	0.617768
24	0.259768	0.329705	0.388244	0.453413	0.495808	0.607360
25	0.254594	0.323283	0.380863	0.445078	0.486932	0.597446
26	0.249717	0.317223	0.373886	0.437184	0.478511	0.587988
27	0.245110	0.311490	0.367278	0.429693	0.470509	0.578956
28	0.240749	0.306057	0.361007	0.422572	0.462892	0.570317
29	0.236612	0.300898	0.355046	0.415792	0.455631	0.562047
30	0.232681	0.295991	0.349370	0.409327	0.448699	0.554119

df\α	0.2	0.1	0.05	0.02	0.01	0.00
35	0.215598	0.274611	0.324573	0.380976	0.418211	0.51889
40	0.201796	0.257278	0.304396	0.357787	0.393174	0.48957
45	0.190345	0.242859	0.287563	0.338367	0.372142	0.46467
50	0.180644	0.230620	0.273243	0.321796	0.354153	0.44320
60	0.164997	0.210832	0.250035	0.294846	0.324818	0.40786
70	0.152818	0.195394	0.231883	0.273695	0.301734	0.37979
80	0.142990	0.182916	0.217185	0.256525	0.282958	0.35681
90	0.134844	0.172558	0.204968	0.242227	0.267298	0.33754
100	0.127947	0.163782	0.194604	0.230079	0.253979	0.32109
125	0.114477	0.146617	0.174308	0.206245	0.227807	0.28860
150	0.104525	0.133919	0.159273	0.188552	0.208349	0.26431
175	0.096787	0.124036	0.147558	0.174749	0.193153	0.24528
200	0.090546	0.116060	0.138098	0.163592	0.180860	0.22984
250	0.081000	0.103852	0.123607	0.146483	0.161994	0.20607
300	0.073951	0.094831	0.112891	0.133819	0.148019	0.18843
350	0.068470	0.087814	0.104552	0.123957	0.137131	0.17465
400	0.064052	0.082155	0.097824	0.115997	0.128339	0.16352
450	0.060391	0.077466	0.092248	0.109397	0.121046	0.15427
500	0.057294	0.073497	0.087528	0.103808	0.114870	0.14643
600	0.052305	0.067103	0.079920	0.094798	0.104911	0.13378
700	0.048427	0.062132	0.074004	0.087789	0.097161	0.12393
800	0.045301	0.058123	0.069234	0.082135	0.090909	0.11598
900	0.042711	0.054802	0.065281	0.077450	0.085727	0.10938
1000	0.040520	0.051993	0.061935	0.073484	0.081340	0.10380
1500	0.033086	0.042458	0.050582	0.060022	0.066445	0.08482
2000	0.028654	0.036772	0.043811	0.051990	0.057557	0.07348
3000	0.023397	0.030027	0.035775	0.042457	0.047006	0.06002
4000	0.020262	0.026005	0.030984	0.036773	0.040713	0.05199
5000	0.018123	0.023260	0.027714	0.032892	0.036417	0.04651

جدول (r)

α: مستوى الاحتمالية

df: درجات الحرية

مثال: جد قيمة معامل الارتباط وبين نوع الارتباط مع اختبار معنوية الارتباط للمتغيرين x و y

X	6	1	8	10	2
y	5	7	4	3	9

الحل: لإيجاد معامل الارتباط (r) نرتب البيانات في جدول كما يلي:

x_i	y_i	x_iy_i	x_i^2	y_i^2
6	5	30	36	25
1	7	7	1	49
8	4	32	64	16
10	3	30	100	9
2	9	18	4	81
$\sum x_i=27$	$\sum y_i=28$	$\sum x_iy_i=117$	$\sum x_i^2=205$	$\sum y_i^2=180$

$$r = \frac{\sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}}{\sqrt{(\sum x_i^2 - \frac{(\sum x_i)^2}{n})(\sum y_i^2 - \frac{(\sum y_i)^2}{n})}}$$

$$r = \frac{117 - \frac{(27)(28)}{5}}{\sqrt{(205 - \frac{(27)^2}{5})(180 - \frac{(28)^2}{5})}}$$

$$r = \frac{-34.2}{\sqrt{(59.2)(23.2)}}$$

$$r = \frac{-34.2}{37.05}$$

$$r = -0.92$$

العلاقة عكسية (سالبة)

لاختبار معنوية الارتباط نستخرج قيمة r الجدولية من جدول r عند درجة حرية 3 (n-2) ومستوى احتمال 0.05 فنجدها تساوي 0.87

Dr. Ahmed Shehat

بما أن قيمة r المحسوبة (0.92-) قيمتها المطلقة (0.92) أكبر من قيمة r الجدولية (0.87) نستنتج ان الارتباط معنوي على مستوى احتمال 0.05. اما على مستوى احتمال 0.01 نجد ان قيمة r الجدولية تساوي 0.95، بما أن قيمة r المحسوبة (0.92-) قيمتها المطلقة (0.92) اقل من قيمة r الجدولية (0.95) نستنتج ان الارتباط غير معنوي على مستوى احتمال 0.01.

Dr. Ahmed Shehab

المصادر:

1. الراوي، خاشع محمود. 1989. المدخل الى الإحصاء. وزارة التعليم العالي والبحث العلمي. جامعة الموصل.
2. طبيه، احمد عبدالسميع. 2007. مبادئ الإحصاء، عمان. دار البداية. ر.أ: www.daralbedayah.com (2007/6/1709)
3. David, M. Lane. Introduction to Statistics. Online Edition.