

Rozdział 7

Miary korelacji/asocjacji

Współczynnik korelacji liniowej Pearsona (r_{xy})

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

Zadanie 1.

Na podstawie poniższych danych oblicz współczynnik korelacji a następnie wykreśl wykres rozrzutu (XY) i określ czy istnieje związek i jaki jest jego charakter (+ istotność statystyczna).

Tablica 7.1: Korelacja liniowa Pearson'a - dane

zestaw 1		zestaw 2						
x	y	x	y	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$
0	0	2	8					
1	3	3	4					
2	6	4	9					
3	8	5	2					
5	11	6	5					
7	13	7	6					
9	14	8	3					
12	15	9	1					
16	16	10	7					
20	16	20	17					

Istotność statystyczna współczynnika korelacji Pearson'a

Statystyka testowa (o rozkładzie t – Studenta o $n - 2$ stopniach swobody) stosowana jest dla niewielkich prób (<100 elementów)

$$t = \frac{r}{\sqrt{1 - r^2}} \sqrt{n - 2}$$

Dla dużych prób (>100 elementów) stosujemy statystykę Z o rozkładzie normalnym $N(\mu = 0, \sigma = 1)$

$$Z = \frac{r}{\sqrt{1 - r^2}} \sqrt{n}$$

Współczynnik korelacji rang Spearman'a (r_s)

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

Istotność statystyczna współczynnika korelacji jest testowana analogicznie jak w przypadku r_{xy}

$$t = \frac{r_s}{\sqrt{1 - r_s^2}} \sqrt{n - 2}$$

Zadanie 2.

Oblicz korelację rang Spearman'a i zweryfikuj hipotezę o jej istotności statystycznej

Tablica 7.2: Korelacja Spearman'a - dane

IQ, x_i	TVh, y_i	ranga x_i	ranga y_i	d_i	d_i^2
106	7				
86	0				
100	27				
101	50				
99	28				
103	29				
97	20				
113	12				
112	6				
110	17				
				Σ	

Tabele kontyngencji

Liczebności oczekiwane

$$\hat{n}_{ij} = n \left(\frac{n_{i.}}{n} \cdot \frac{n_{.j}}{n} \right) = \frac{n_{i.} \cdot n_{.j}}{n}$$

Statystyka testowa χ^2

$$\chi^2 = \sum_i \sum_j \frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$$

Tablica 7.3: Występowanie typów malarii w trzech regionach tropikalnych (n_{ij})

	Azja	Afryka	Ameryka Pd.	Σ
Malaria A	31	14	45	
Malaria B	2	5	53	
Malaria C	53	45	2	
Σ				

Tablica 7.4: Liczebności oczekiwane

	Azja	Afryka	Ameryka Pd.	Σ
Malaria A				
Malaria B				
Malaria C				
Σ				

Statystyka χ^2 przyjmuje wartości z przedziału $[0; n\sqrt{(w-1)(k-1)}]$ Wynik porównujemy z rozkładem χ^2 przy $(w-1)(k-1)$ stopniach swobody i założonym poziomie istotności.

Jeżeli nasza wartość jest większa od wartości krytycznej odrzucamy hipotezę zerową o braku asocjacji

Współczynniki kontyngencji

Współczynnik zbieżności T-Czupurowa

$$T = \sqrt{\frac{\chi^2}{n\sqrt{(w-1)(k-1)}}$$

Współczynnik V-Cramera

$$V = \sqrt{\frac{\chi^2}{n \cdot \min(w-1; k-1)}}$$

Współczynnik kontyngencji C-Pearsona

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}}$$

$$C_{max}(jeżeli w=k) = \sqrt{\frac{k-1}{k}}; C_{max}(jeżeli w \neq k) = \frac{1}{2} \left[\sqrt{\frac{k-1}{k}} + \sqrt{\frac{w-1}{w}} \right]$$

C skorygowane

$$C_{kor} = \frac{c}{c_{max}}$$

Tablica 7.5: Tabela pomocnicza

Liczebności obserwowane n_{ij}	Liczebności oczekiwane \hat{n}_{ij}	$n_{ij} - \hat{n}_{ij}$	$(n_{ij} - \hat{n}_{ij})^2$	$\frac{(n_{ij} - \hat{n}_{ij})^2}{\hat{n}_{ij}}$
31				
14				
45				
2				
5				
53				
53				
45				
2				
Σ				

Tablica 7.6: Tablica wartości krytycznych rozkładu χ^2 dla liczby stopni swobody n

n	0,10	0,05	0,01	0,01	0,001	n	n	0,10	0,05	0,01	0,01	0,001	n
1	2,71	3,84	6,63	7,88	10,83	1	51	64,30	68,67	77,39	80,75	87,97	51
2	4,61	5,99	9,21	10,60	13,82	2	52	65,42	69,83	78,62	82,00	89,27	52
3	6,25	7,81	11,34	12,84	16,27	3	53	66,55	70,99	79,84	83,25	90,57	53
4	7,78	9,49	13,28	14,86	18,47	4	54	67,67	72,15	81,07	84,50	91,87	54
5	9,24	11,07	15,09	16,75	20,52	5	55	68,80	73,31	82,29	85,75	93,17	55
6	10,64	12,59	16,81	18,55	22,46	6	56	69,92	74,47	83,51	86,99	94,46	56
7	12,02	14,07	18,48	20,28	24,32	7	57	71,04	75,62	84,73	88,24	95,75	57
8	13,36	15,51	20,09	21,96	26,12	8	58	72,16	76,78	85,95	89,48	97,04	58
9	14,68	16,92	21,67	23,59	27,88	9	59	73,28	77,93	87,17	90,72	98,32	59
10	15,99	18,31	23,21	25,19	29,59	10	60	74,40	79,08	88,38	91,95	99,61	60
11	17,28	19,68	24,73	26,76	31,26	11	61	75,51	80,23	89,59	93,19	100,89	61
12	18,55	21,03	26,22	28,30	32,91	12	62	76,63	81,38	90,80	94,42	102,17	62
13	19,81	22,36	27,69	29,82	34,53	13	63	77,75	82,53	92,01	95,65	103,44	63
14	21,06	23,68	29,14	31,32	36,12	14	64	78,86	83,68	93,22	96,88	104,72	64
15	22,31	25,00	30,58	32,80	37,70	15	65	79,97	84,82	94,42	98,11	105,99	65
16	23,54	26,30	32,00	34,27	39,25	16	66	81,09	85,96	95,63	99,33	107,26	66
17	24,77	27,59	33,41	35,72	40,79	17	67	82,20	87,11	96,83	100,55	108,53	67
18	25,99	28,87	34,81	37,16	42,31	18	68	83,31	88,25	98,03	101,78	109,79	68
19	27,20	30,14	36,19	38,58	43,82	19	69	84,42	89,39	99,23	103,00	111,06	69
20	28,41	31,41	37,57	40,00	45,31	20	70	85,53	90,53	100,43	104,22	112,32	70
21	29,62	32,67	38,93	41,40	46,80	21	71	86,64	91,67	101,62	105,43	113,58	71
22	30,81	33,92	40,29	42,80	48,27	22	72	87,74	92,81	102,82	106,65	114,84	72
23	32,01	35,17	41,64	44,18	49,73	23	73	88,85	93,95	104,01	107,86	116,09	73
24	33,20	36,42	42,98	45,56	51,18	24	74	89,96	95,08	105,20	109,07	117,35	74
25	34,38	37,65	44,31	46,93	52,62	25	75	91,06	96,22	106,39	110,29	118,60	75
26	35,56	38,89	45,64	48,29	54,05	26	76	92,17	97,35	107,58	111,50	119,85	76
27	36,74	40,11	46,96	49,64	55,48	27	77	93,27	98,48	108,77	112,70	121,10	77
28	37,92	41,34	48,28	50,99	56,89	28	78	94,37	99,62	109,96	113,91	122,35	78
29	39,09	42,56	49,59	52,34	58,30	29	79	95,48	100,75	111,14	115,12	123,59	79
30	40,26	43,77	50,89	53,67	59,70	30	80	96,58	101,88	112,33	116,32	124,84	80
31	41,42	44,99	52,19	55,00	61,10	31	81	97,68	103,01	113,51	117,52	126,08	81
32	42,58	46,19	53,49	56,33	62,49	32	82	98,78	104,14	114,70	118,73	127,32	82
33	43,75	47,40	54,78	57,65	63,87	33	83	99,88	105,27	115,88	119,93	128,57	83
34	44,90	48,60	56,06	58,96	65,25	34	84	100,98	106,40	117,06	121,13	129,80	84
35	46,06	49,80	57,34	60,27	66,62	35	85	102,08	107,52	118,24	122,33	131,04	85
36	47,21	51,00	58,62	61,58	67,99	36	86	103,18	108,65	119,41	123,52	132,28	86
37	48,36	52,19	59,89	62,88	69,35	37	87	104,28	109,77	120,59	124,72	133,51	87
38	49,51	53,38	61,16	64,18	70,70	38	88	105,37	110,90	121,77	125,91	134,75	88
39	50,66	54,57	62,43	65,48	72,05	39	89	106,47	112,02	122,94	127,11	135,98	89
40	51,81	55,76	63,69	66,77	73,40	40	90	107,57	113,15	124,12	128,30	137,21	90
41	52,95	56,94	64,95	68,05	74,75	41	91	108,66	114,27	125,29	129,49	138,44	91
42	54,09	58,12	66,21	69,34	76,08	42	92	109,76	115,39	126,46	130,68	139,67	92
43	55,23	59,30	67,46	70,62	77,42	43	93	110,85	116,51	127,63	131,87	140,89	93
44	56,37	60,48	68,71	71,89	78,75	44	94	111,94	117,63	128,80	133,06	142,12	94
45	57,51	61,66	69,96	73,17	80,08	45	95	113,04	118,75	129,97	134,25	143,34	95
46	58,64	62,83	71,20	74,44	81,40	46	96	114,13	119,87	131,14	135,43	144,57	96
47	59,77	64,00	72,44	75,70	82,72	47	97	115,22	120,99	132,31	136,62	145,79	97
48	60,91	65,17	73,68	76,97	84,04	48	98	116,32	122,11	133,48	137,80	147,01	98
49	62,04	66,34	74,92	78,23	85,35	49	99	117,41	123,23	134,64	138,99	148,23	99
50	63,17	67,50	76,15	79,49	86,66	50	100	118,50	124,34	135,81	140,17	149,45	100

Tablica 7.7: Wartości krytyczne rozkładu t – Studenta

kwantyl roz- kładu	0,9	0,95	0,975	0,98	0,99	0,995	0,999	0,9995
obszar kry- tyczny jed- nostronny, $\alpha =$	0,1	0,05	0,025	0,02	0,01	0,005	0,001	0,0005
obszar kry- tyczny dwu- stronny, $\alpha =$	0,2	0,1	0,05	0,04	0,02	0,01	0,002	0,001
n=1 (rozkład Cauchy'ego)	3,08	6,31	12,71	15,89	31,82	63,66	318,31	636,63
2	1,89	2,92	4,30	4,85	6,96	9,92	22,33	31,60
3	1,64	2,35	3,18	3,48	4,54	5,84	10,21	12,92
4	1,53	2,13	2,78	3,00	3,75	4,60	7,17	8,61
5	1,48	2,02	2,57	2,76	3,36	4,03	5,89	6,87
6	1,44	1,94	2,45	2,61	3,14	3,71	5,21	5,96
7	1,41	1,89	2,36	2,52	3,00	3,50	4,79	5,41
8	1,40	1,86	2,31	2,45	2,90	3,36	4,50	5,04
9	1,38	1,83	2,26	2,40	2,82	3,25	4,30	4,78
10	1,37	1,81	2,23	2,36	2,76	3,17	4,14	4,59
11	1,36	1,80	2,20	2,33	2,72	3,11	4,02	4,44
12	1,36	1,78	2,18	2,30	2,68	3,05	3,93	4,32
13	1,35	1,77	2,16	2,28	2,65	3,01	3,85	4,22
14	1,35	1,76	2,14	2,26	2,62	2,98	3,79	4,14
15	1,34	1,75	2,13	2,25	2,60	2,95	3,73	4,07
16	1,34	1,75	2,12	2,24	2,58	2,92	3,69	4,02
17	1,33	1,74	2,11	2,22	2,57	2,90	3,65	3,97
18	1,33	1,73	2,10	2,21	2,55	2,88	3,61	3,92
19	1,33	1,73	2,09	2,20	2,54	2,86	3,58	3,88
20	1,33	1,72	2,09	2,20	2,53	2,85	3,55	3,85
21	1,32	1,72	2,08	2,19	2,52	2,83	3,53	3,82
22	1,32	1,72	2,07	2,18	2,51	2,82	3,50	3,79
23	1,32	1,71	2,07	2,18	2,50	2,81	3,48	3,77
24	1,32	1,71	2,06	2,17	2,49	2,80	3,47	3,75
25	1,32	1,71	2,06	2,17	2,49	2,79	3,45	3,73
26	1,31	1,71	2,06	2,16	2,48	2,78	3,44	3,71
27	1,31	1,70	2,05	2,16	2,47	2,77	3,42	3,69
28	1,31	1,70	2,05	2,15	2,47	2,76	3,41	3,67
29	1,31	1,70	2,05	2,15	2,46	2,76	3,40	3,66
30	1,31	1,70	2,04	2,15	2,46	2,75	3,39	3,65
31	1,31	1,70	2,04	2,14	2,45	2,74	3,37	3,63
32	1,31	1,69	2,04	2,14	2,45	2,74	3,37	3,62
33	1,31	1,69	2,03	2,14	2,44	2,73	3,36	3,61
34	1,31	1,69	2,03	2,14	2,44	2,73	3,35	3,60
35	1,31	1,69	2,03	2,13	2,44	2,72	3,34	3,59
36	1,31	1,69	2,03	2,13	2,43	2,72	3,33	3,58
37	1,30	1,69	2,03	2,13	2,43	2,72	3,33	3,57
38	1,30	1,69	2,02	2,13	2,43	2,71	3,32	3,57
39	1,30	1,68	2,02	2,12	2,43	2,71	3,31	3,56
40	1,30	1,68	2,02	2,12	2,42	2,70	3,31	3,55
41	1,30	1,68	2,02	2,12	2,42	2,70	3,30	3,54
42	1,30	1,68	2,02	2,12	2,42	2,70	3,30	3,54
43	1,30	1,68	2,02	2,12	2,42	2,70	3,29	3,53
44	1,30	1,68	2,02	2,12	2,41	2,69	3,29	3,53
45	1,30	1,68	2,01	2,12	2,41	2,69	3,28	3,52
46	1,30	1,68	2,01	2,11	2,41	2,69	3,28	3,51
47	1,30	1,68	2,01	2,11	2,41	2,68	3,27	3,51
48	1,30	1,68	2,01	2,11	2,41	2,68	3,27	3,51
49	1,30	1,68	2,01	2,11	2,40	2,68	3,27	3,50
50	1,30	1,68	2,01	2,11	2,40	2,68	3,26	3,50
55	1,30	1,67	2,00	2,10	2,40	2,67	3,25	3,48
60	1,30	1,67	2,00	2,10	2,39	2,66	3,23	3,46
65	1,29	1,67	2,00	2,10	2,39	2,65	3,22	3,45
70	1,29	1,67	1,99	2,09	2,38	2,65	3,21	3,44
75	1,29	1,67	1,99	2,09	2,38	2,64	3,20	3,43
80	1,29	1,66	1,99	2,09	2,37	2,64	3,20	3,42
85	1,29	1,66	1,99	2,09	2,37	2,63	3,19	3,41
90	1,29	1,66	1,99	2,08	2,37	2,63	3,18	3,40
95	1,29	1,66	1,99	2,08	2,37	2,63	3,18	3,40
100	1,29	1,66	1,98	2,08	2,36	2,63	3,17	3,39
110	1,29	1,66	1,98	2,08	2,36	2,62	3,17	3,38
120	1,29	1,66	1,98	2,08	2,36	2,62	3,16	3,37
130	1,29	1,66	1,98	2,07	2,36	2,61	3,15	3,37
140	1,29	1,66	1,98	2,07	2,35	2,61	3,15	3,36
150	1,29	1,66	1,98	2,07	2,35	2,61	3,15	3,36
160	1,29	1,65	1,97	2,07	2,35	2,61	3,14	3,35
170	1,29	1,65	1,97	2,07	2,35	2,61	3,14	3,35
180	1,29	1,65	1,97	2,07	2,35	2,60	3,14	3,35
190	1,29	1,65	1,97	2,07	2,35	2,60	3,13	3,34
200	1,29	1,65	1,97	2,07	2,35	2,60	3,13	3,34
210	1,29	1,65	1,97	2,07	2,34	2,60	3,13	3,34
220	1,29	1,65	1,97	2,07	2,34	2,60	3,13	3,34
230	1,29	1,65	1,97	2,07	2,34	2,60	3,13	3,33
240	1,29	1,65	1,97	2,06	2,34	2,60	3,12	3,33
250	1,28	1,65	1,97	2,06	2,34	2,60	3,12	3,33
260	1,28	1,65	1,97	2,06	2,34	2,59	3,12	3,33
270	1,28	1,65	1,97	2,06	2,34	2,59	3,12	3,33
280	1,28	1,65	1,97	2,06	2,34	2,59	3,12	3,33
290	1,28	1,65	1,97	2,06	2,34	2,59	3,12	3,32
300	1,28	1,65	1,97	2,06	2,34	2,59	3,12	3,32
350	1,28	1,65	1,97	2,06	2,34	2,59	3,11	3,32
400	1,28	1,65	1,97	2,06	2,34	2,59	3,11	3,32
450	1,28	1,65	1,97	2,06	2,33	2,59	3,11	3,31
500	1,28	1,65	1,96	2,06	2,33	2,59	3,11	3,31
600	1,28	1,65	1,96	2,06	2,33	2,58	3,10	3,31
700	1,28	1,65	1,96	2,06	2,33	2,58	3,10	3,30
800	1,28	1,65	1,96	2,06	2,33	2,58	3,10	3,30
900	1,28	1,65	1,96	2,06	2,33	2,58	3,10	3,30
1000	1,28	1,65	1,96	2,06	2,33	2,58	3,10	3,30