

МИНИСТЕРСТВО ПРОСВЕЩЕНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКИЙ ГОСУДАРСТВЕННЫЙ ПЕДАГОГИЧЕСКИЙ УНИВЕРСИТЕТ им. А.И.
ГЕРЦЕНА»



Направление подготовки

09.03.01 – Информатика и вычислительная техника

Профиль «Технологии разработки программного обеспечения»

Лабораторная работа №7 часть 3

«Корреляция. Корреляционный анализ»

Работу выполнили студенты 2 курса 2-1 группы:

Зухир Амира

Крючкова Анастасия

Стецук Максим

Каргаполов Денис

САНКТ-ПЕТЕРБУРГ

2022

СОДЕРЖАНИЕ

Отчет Зухир Амиры	3
Отчет Крючковой Анастасии	7
Отчет Стецук Максима	11
Отчет Каргаполова Дениса	15

Лабораторная работа № 7 часть 2

Корреляция. Корреляционный анализ.

Вычисление множественного коэффициента корреляции.

Задание 1:

Изучили теория о вычислении множественного коэффициента корреляции.

Будут использованы данные формулы:

Выборочный множественный коэффициент корреляции:

$$R_i = \sqrt{1 - \frac{|A|}{A_{ii}}},$$

Из данной формулы для $z = 3$ получим:

$$R_i = \sqrt{\frac{r_{ij}^2 + r_{ik}^2 - 2r_{ij} r_{ik} r_{jk}}{1 - r_{jk}^2}}$$

Множественный коэффициент корреляции значимо отличается от нуля, если статистика

$$t = \frac{R^2(n-p)}{(1-R^2)(p-1)} > F_{кр}, \quad (4)$$

Коэффициенты r :

$$r = \frac{n \sum_{l=1}^n x_l y_l - (\sum_{l=1}^n x_l) (\sum_{j=1}^n y_j)}{\sqrt{n \sum_{l=1}^n x_l^2 - (\sum_{l=1}^n x_l)^2} \sqrt{n \sum_{j=1}^n y_j^2 - (\sum_{j=1}^n y_j)^2}}$$

Коэффициент корреляции:

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

Задание 2:

Задача 2.1. Проведено исследование шести фермерских хозяйств для изучения зависимости урожайности зерновых культур (Z , ц/га) от качества пашни (X , в баллах) и количества внесенных минеральных удобрений на 1 га пашни (Y , кг/га).

Результаты приведены в таблице 1.

Определить тесноту связи между величиной Z и величинами X и Y , используя множественный коэффициент корреляции. Проверить на уровне $\alpha = 0.05$ значимость R_z .

X	26	35	36	40	41	45
Y	2,1	2,3	2,4	2,6	2,9	3,0
Z	18,00	21,0	22,1	25,3	28,0	28,5

i	x_i	y_i	z_i	x_i^2	y_i^2	z_i^2	$x_i y_i$	$x_i z_i$	$y_i z_i$
1	26	2,1	18	676	4,41	324	54,6	468	37,8
2	35	2,3	21	1225	5,29	441	80,5	735	48,3
3	36	2,4	22,1	1296	5,76	488,41	86,4	795,6	53,04
4	40	2,6	25,3	1600	6,76	640,09	104	1012	65,78
5	41	2,9	28	1681	8,41	784	118,9	1148	81,2
6	45	3	28,5	2025	9	812,25	135	1282,5	85,5
Σ	223	15,3	142,9	8503	39,63	3489,75	579,4	5441,1	371,62
Средняя	37,16667	2,55	23,81667	1417,167	6,605	581,625	96,56667	906,85	61,93667

$x_i - x_{cp}$	$y_i - y_{cp}$	$z_i - z_{cp}$	$(x_i - x_{cp})^2$	$(y_i - y_{cp})^2$	$(z_i - z_{cp})^2$
-11,1667	-0,45	-5,81667	124,6944	0,2025	33,83361
-2,16667	-0,25	-2,81667	4,694444	0,0625	7,933611
-1,16667	-0,15	-1,71667	1,361111	0,0225	2,946944
2,833333	0,05	1,483333	8,027778	0,0025	2,200278
3,833333	0,35	4,183333	14,69444	0,1225	17,50028
7,833333	0,45	4,683333	61,36111	0,2025	21,93361
0	0	0	214,8333	0,615	86,34833

i	$(x_i - x_{cp}) * (y_i - y_{cp})$	$(x_i - x_{cp}) * (z_i - z_{cp})$	$(z_i - z_{cp}) * (y_i - y_{cp})$
1	5,025	64,953	2,618
2	0,542	6,103	0,704
3	0,175	2,003	0,258
4	0,142	4,203	0,074
5	1,342	16,036	1,464
6	3,525	36,686	2,108
Σ	10,75	129,9833333	7,225

r	64,5	35,90265	15,3	0,11742
---	------	----------	------	---------

r_{xy}	0,935234
r_{xz}	0,954355
r_{zy}	0,991456

матрица А		
1	0,935234	0,954355
0,935234	1	0,991456
0,954355	0,991456	1

R3	0,999301
Rz	0,993714
p	3
t	118,1844
k1	2
k2	3
α	0.05
$F_{кр}$	9,55

Рассчитали всё с помощью формул из задания 1 и получили: $t = 118,1844 > F_{кр} = 9,55$, а это значит, что коэффициент R является значимым.

Задание 3:

Задача 3.1 При изучении финансовой деятельности компании в течение некоторого времени был собран статистический материал. В таблице содержатся данные о ежемесячной прибыли Z (тыс. усл. ед.) , расходах на рекламу X (тыс. усл. ед.) и вложении капитала в ценные бумаги Y (тыс. усл. ед.).

Определить тесноту связи между переменной Z и переменными X и Y с помощью выборочного множественного коэффициента корреляции R и определить его значимость на уровне $\alpha = 0.05$.

Z	10	12	12	14	16	17	18
X	0,2	0,5	0,3	0,5	0,5	0,6	0,8
Y	0,80	0,2	1	1,2	0,9	1	1,1

i	x_i	y_i	z_i	x_i^2	y_i^2	z_i^2	$x_i y_i$	$x_i z_i$	$y_i z_i$
1	0,2	0,80	10	0,04	0,64	100	0,16	2	8
2	0,5	0,2	12	0,25	0,04	144	0,1	6	2,4
3	0,3	1	12	0,09	1	144	0,3	3,6	12
4	0,5	1,2	14	0,25	1,44	196	0,6	7	16,8
5	0,5	0,9	16	0,25	0,81	256	0,45	8	14,4
6	0,6	1	17	0,36	1	289	0,6	10,2	17
7	0,8	1,1	18	0,64	1,21	324	0,88	14,4	19,8
Σ	3,4	6,20	99	1,88	6,14	1453	3,09	51,2	90,4
Средняя	0,485714	0,885714	14,14286	0,268571	0,877143	207,5714	0,441429	7,314286	12,91429

$x_i - x_{cp}$	$y_i - y_{cp}$	$z_i - z_{cp}$	$(x_i - x_{cp})^2$	$(y_i - y_{cp})^2$	$(z_i - z_{cp})^2$
-0,28571	-0,08571	-4,14286	0,081633	0,007347	17,16327
0,014286	-0,68571	-2,14286	0,000204	0,470204	4,591837
-0,18571	0,114286	-2,14286	0,03449	0,013061	4,591837
0,014286	0,314286	-0,14286	0,000204	0,098776	0,020408
0,014286	0,014286	1,857143	0,000204	0,000204	3,44898
0,114286	0,114286	2,857143	0,013061	0,013061	8,163265
0,314286	0,214286	3,857143	0,098776	0,045918	14,87755
0	-1,2E-15	3,55E-15	0,228571	0,648571	52,85714

i	$(x_i - x_{cp}) * (y_i - y_{cp})$	$(x_i - x_{cp}) * (z_i - z_{cp})$	$(z_i - z_{cp}) * (y_i - y_{cp})$
1	0,024489796	1,184	0,355
2	-0,010	-0,031	1,469
3	-0,02122449	0,398	-0,245
4	0,004	-0,002	-0,045
5	0,000	0,027	0,027
6	0,013061224	0,327	0,327
7	0,067346939	1,212	0,827
Σ	0,078571429	3,114	2,714

r	-2,54	1,264911	6,2	-0,32388
---	-------	----------	-----	----------

r_{xy}	0,204068
r_{xz}	0,895974
r_{zy}	0,46358

матрица A		
1	0,204068	0,895974
0,204068	1	0,46358
0,895974	0,46358	1

R3	0,943292
Rz	0,887401
p	3
t	7,410929
k1	2
k2	4
α	0.05
$F_{кр}$	6,94

Рассчитали всё с помощью формул из задания 1 и получили: $t \approx 7,41 > F_{кр} = 6,94$, а это значит, что коэффициент R является значимым.

Лабораторная работа № 7 часть 2

Корреляция. Корреляционный анализ.

Вычисление множественного коэффициента корреляции.

Задание 1:

Изучили теория о вычислении множественного коэффициента корреляции.

Будут использованы данные формулы:

Выборочный множественный коэффициент корреляции:

$$R_i = \sqrt{1 - \frac{|A|}{A_{ii}}},$$

Из данной формулы для $z = 3$ получим:

$$R_i = \sqrt{\frac{r_{ij}^2 + r_{ik}^2 - 2r_{ij} r_{ik} r_{jk}}{1 - r_{jk}^2}}$$

Множественный коэффициент корреляции значимо отличается от нуля, если статистика

$$t = \frac{R^2(n-p)}{(1-R^2)(p-1)} > F_{кр}, \quad (4)$$

Коэффициенты r :

$$r = \frac{n \sum_{l=1}^n x_l y_l - (\sum_{l=1}^n x_l) (\sum_{j=1}^n y_j)}{\sqrt{n \sum_{l=1}^n x_l^2 - (\sum_{l=1}^n x_l)^2} \sqrt{n \sum_{j=1}^n y_j^2 - (\sum_{j=1}^n y_j)^2}}$$

Коэффициент корреляции:

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

Задание 2:

Задача 2.1. Проведено исследование шести фермерских хозяйств для изучения зависимости урожайности зерновых культур (Z , ц/га) от качества пашни (X , в баллах) и количества внесенных минеральных удобрений на 1 га пашни (Y , кг/га).

Результаты приведены в таблице 1.

Определить тесноту связи между величиной Z и величинами X и Y , используя множественный коэффициент корреляции. Проверить на уровне $\alpha = 0.05$ значимость R_z .

X	26	35	36	40	41	45
Y	2,1	2,3	2,4	2,6	2,9	3,0
Z	18,00	21,0	22,1	25,3	28,0	28,5

i	x_i	y_i	z_i	x_i^2	y_i^2	z_i^2	$x_i y_i$	$x_i z_i$	$y_i z_i$
1	26	2,1	18	676	4,41	324	54,6	468	37,8
2	35	2,3	21	1225	5,29	441	80,5	735	48,3
3	36	2,4	22,1	1296	5,76	488,41	86,4	795,6	53,04
4	40	2,6	25,3	1600	6,76	640,09	104	1012	65,78
5	41	2,9	28	1681	8,41	784	118,9	1148	81,2
6	45	3	28,5	2025	9	812,25	135	1282,5	85,5
Σ	223	15,3	142,9	8503	39,63	3489,75	579,4	5441,1	371,62
Средняя	37,16667	2,55	23,81667	1417,167	6,605	581,625	96,56667	906,85	61,93667

$x_i - x_{cp}$	$y_i - y_{cp}$	$z_i - z_{cp}$	$(x_i - x_{cp})^2$	$(y_i - y_{cp})^2$	$(z_i - z_{cp})^2$
-11,1667	-0,45	-5,81667	124,6944	0,2025	33,83361
-2,16667	-0,25	-2,81667	4,694444	0,0625	7,933611
-1,16667	-0,15	-1,71667	1,361111	0,0225	2,946944
2,833333	0,05	1,483333	8,027778	0,0025	2,200278
3,833333	0,35	4,183333	14,69444	0,1225	17,50028
7,833333	0,45	4,683333	61,36111	0,2025	21,93361
0	0	0	214,8333	0,615	86,34833

i	$(x_i - x_{cp}) * (y_i - y_{cp})$	$(x_i - x_{cp}) * (z_i - z_{cp})$	$(z_i - z_{cp}) * (y_i - y_{cp})$
1	5,025	64,953	2,618
2	0,542	6,103	0,704
3	0,175	2,003	0,258
4	0,142	4,203	0,074
5	1,342	16,036	1,464
6	3,525	36,686	2,108
Σ	10,75	129,9833333	7,225

r	64,5	35,90265	15,3	0,11742
---	------	----------	------	---------

r_{xy}	0,935234
r_{xz}	0,954355
r_{zy}	0,991456

матрица А		
1	0,935234	0,954355
0,935234	1	0,991456
0,954355	0,991456	1

R3	0,999301
Rz	0,993714
p	3
t	118,1844
k1	2
k2	3
α	0.05
$F_{кр}$	9,55

Рассчитали всё с помощью формул из задания 1 и получили: $t = 118,1844 > F_{кр} = 9,55$, а это значит, что коэффициент R является значимым.

Задание 3:

Задача 3.1 При изучении финансовой деятельности компании в течение некоторого времени был собран статистический материал. В таблице содержатся данные о ежемесячной прибыли Z (тыс. усл. ед.) , расходах на рекламу X (тыс. усл. ед.) и вложении капитала в ценные бумаги Y (тыс. усл. ед.).

Определить тесноту связи между переменной Z и переменными X и Y с помощью выборочного множественного коэффициента корреляции R и определить его значимость на уровне $\alpha = 0.05$.

Z	10	12	12	14	16	17	18
X	0,2	0,5	0,3	0,5	0,5	0,6	0,8
Y	0,80	0,2	1	1,2	0,9	1	1,1

i	x_i	y_i	z_i	x_i^2	y_i^2	z_i^2	$x_i y_i$	$x_i z_i$	$y_i z_i$
1	0,2	0,80	10	0,04	0,64	100	0,16	2	8
2	0,5	0,2	12	0,25	0,04	144	0,1	6	2,4
3	0,3	1	12	0,09	1	144	0,3	3,6	12
4	0,5	1,2	14	0,25	1,44	196	0,6	7	16,8
5	0,5	0,9	16	0,25	0,81	256	0,45	8	14,4
6	0,6	1	17	0,36	1	289	0,6	10,2	17
7	0,8	1,1	18	0,64	1,21	324	0,88	14,4	19,8
Σ	3,4	6,20	99	1,88	6,14	1453	3,09	51,2	90,4
Средняя	0,485714	0,885714	14,14286	0,268571	0,877143	207,5714	0,441429	7,314286	12,91429

$x_i - x_{cp}$	$y_i - y_{cp}$	$z_i - z_{cp}$	$(x_i - x_{cp})^2$	$(y_i - y_{cp})^2$	$(z_i - z_{cp})^2$
-0,28571	-0,08571	-4,14286	0,081633	0,007347	17,16327
0,014286	-0,68571	-2,14286	0,000204	0,470204	4,591837
-0,18571	0,114286	-2,14286	0,03449	0,013061	4,591837
0,014286	0,314286	-0,14286	0,000204	0,098776	0,020408
0,014286	0,014286	1,857143	0,000204	0,000204	3,44898
0,114286	0,114286	2,857143	0,013061	0,013061	8,163265
0,314286	0,214286	3,857143	0,098776	0,045918	14,87755
0	-1,2E-15	3,55E-15	0,228571	0,648571	52,85714

i	$(x_i - x_{cp}) * (y_i - y_{cp})$	$(x_i - x_{cp}) * (z_i - z_{cp})$	$(z_i - z_{cp}) * (y_i - y_{cp})$
1	0,024489796	1,184	0,355
2	-0,010	-0,031	1,469
3	-0,02122449	0,398	-0,245
4	0,004	-0,002	-0,045
5	0,000	0,027	0,027
6	0,013061224	0,327	0,327
7	0,067346939	1,212	0,827
Σ	0,078571429	3,114	2,714

r	-2,54	1,264911	6,2	-0,32388
---	-------	----------	-----	----------

r_{xy}	0,204068
r_{xz}	0,895974
r_{zy}	0,46358

матрица A		
1	0,204068	0,895974
0,204068	1	0,46358
0,895974	0,46358	1

R3	0,943292
Rz	0,887401
p	3
t	7,410929
k1	2
k2	4
α	0.05
$F_{кр}$	6,94

Рассчитали всё с помощью формул из задания 1 и получили: $t \approx 7,41 > F_{кр} = 6,94$, а это значит, что коэффициент R является значимым.

Лабораторная работа № 7 часть 2

Корреляция. Корреляционный анализ.

Вычисление множественного коэффициента корреляции.

Задание 1:

Изучили теория о вычислении множественного коэффициента корреляции.

Будут использованы данные формулы:

Выборочный множественный коэффициент корреляции:

$$R_i = \sqrt{1 - \frac{|A|}{A_{ii}}},$$

Из данной формулы для $z = 3$ получим:

$$R_i = \sqrt{\frac{r_{ij}^2 + r_{ik}^2 - 2r_{ij} r_{ik} r_{jk}}{1 - r_{jk}^2}}$$

Множественный коэффициент корреляции значимо отличается от нуля, если статистика

$$t = \frac{R^2(n-p)}{(1-R^2)(p-1)} > F_{кр}, \quad (4)$$

Коэффициенты r :

$$r = \frac{n \sum_{l=1}^n x_l y_l - (\sum_{l=1}^n x_l) (\sum_{j=1}^n y_j)}{\sqrt{n \sum_{l=1}^n x_l^2 - (\sum_{l=1}^n x_l)^2} \sqrt{n \sum_{j=1}^n y_j^2 - (\sum_{j=1}^n y_j)^2}}$$

Коэффициент корреляции:

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

Задание 2:

Задача 2.1. Проведено исследование шести фермерских хозяйств для изучения зависимости урожайности зерновых культур (Z , ц/га) от качества пашни (X , в баллах) и количества внесенных минеральных удобрений на 1 га пашни (Y , кг/га).

Результаты приведены в таблице 1.

Определить тесноту связи между величиной Z и величинами X и Y , используя множественный коэффициент корреляции. Проверить на уровне $\alpha = 0.05$ значимость R_z .

X	26	35	36	40	41	45
Y	2,1	2,3	2,4	2,6	2,9	3,0
Z	18,00	21,0	22,1	25,3	28,0	28,5

i	x_i	y_i	z_i	x_i^2	y_i^2	z_i^2	$x_i y_i$	$x_i z_i$	$y_i z_i$
1	26	2,1	18	676	4,41	324	54,6	468	37,8
2	35	2,3	21	1225	5,29	441	80,5	735	48,3
3	36	2,4	22,1	1296	5,76	488,41	86,4	795,6	53,04
4	40	2,6	25,3	1600	6,76	640,09	104	1012	65,78
5	41	2,9	28	1681	8,41	784	118,9	1148	81,2
6	45	3	28,5	2025	9	812,25	135	1282,5	85,5
Σ	223	15,3	142,9	8503	39,63	3489,75	579,4	5441,1	371,62
Средняя	37,16667	2,55	23,81667	1417,167	6,605	581,625	96,56667	906,85	61,93667

$x_i - x_{cp}$	$y_i - y_{cp}$	$z_i - z_{cp}$	$(x_i - x_{cp})^2$	$(y_i - y_{cp})^2$	$(z_i - z_{cp})^2$
-11,1667	-0,45	-5,81667	124,6944	0,2025	33,83361
-2,16667	-0,25	-2,81667	4,694444	0,0625	7,933611
-1,16667	-0,15	-1,71667	1,361111	0,0225	2,946944
2,833333	0,05	1,483333	8,027778	0,0025	2,200278
3,833333	0,35	4,183333	14,69444	0,1225	17,50028
7,833333	0,45	4,683333	61,36111	0,2025	21,93361
0	0	0	214,8333	0,615	86,34833

i	$(x_i - x_{cp}) * (y_i - y_{cp})$	$(x_i - x_{cp}) * (z_i - z_{cp})$	$(z_i - z_{cp}) * (y_i - y_{cp})$
1	5,025	64,953	2,618
2	0,542	6,103	0,704
3	0,175	2,003	0,258
4	0,142	4,203	0,074
5	1,342	16,036	1,464
6	3,525	36,686	2,108
Σ	10,75	129,9833333	7,225

r	64,5	35,90265	15,3	0,11742
---	------	----------	------	---------

r_{xy}	0,935234
r_{xz}	0,954355
r_{zy}	0,991456

матрица А		
1	0,935234	0,954355
0,935234	1	0,991456
0,954355	0,991456	1

R3	0,999301
Rz	0,993714
p	3
t	118,1844
k1	2
k2	3
α	0.05
$F_{кр}$	9,55

Рассчитали всё с помощью формул из задания 1 и получили: $t = 118,1844 > F_{кр} = 9,55$, а это значит, что коэффициент R является значимым.

Задание 3:

Задача 3.1 При изучении финансовой деятельности компании в течение некоторого времени был собран статистический материал. В таблице содержатся данные о ежемесячной прибыли Z (тыс. усл. ед.) , расходах на рекламу X (тыс. усл. ед.) и вложении капитала в ценные бумаги Y (тыс. усл. ед.).

Определить тесноту связи между переменной Z и переменными X и Y с помощью выборочного множественного коэффициента корреляции R и определить его значимость на уровне $\alpha = 0.05$.

Z	10	12	12	14	16	17	18
X	0,2	0,5	0,3	0,5	0,5	0,6	0,8
Y	0,80	0,2	1	1,2	0,9	1	1,1

i	x_i	y_i	z_i	x_i^2	y_i^2	z_i^2	$x_i y_i$	$x_i z_i$	$y_i z_i$
1	0,2	0,80	10	0,04	0,64	100	0,16	2	8
2	0,5	0,2	12	0,25	0,04	144	0,1	6	2,4
3	0,3	1	12	0,09	1	144	0,3	3,6	12
4	0,5	1,2	14	0,25	1,44	196	0,6	7	16,8
5	0,5	0,9	16	0,25	0,81	256	0,45	8	14,4
6	0,6	1	17	0,36	1	289	0,6	10,2	17
7	0,8	1,1	18	0,64	1,21	324	0,88	14,4	19,8
Σ	3,4	6,20	99	1,88	6,14	1453	3,09	51,2	90,4
Средняя	0,485714	0,885714	14,14286	0,268571	0,877143	207,5714	0,441429	7,314286	12,91429

$x_i - x_{cp}$	$y_i - y_{cp}$	$z_i - z_{cp}$	$(x_i - x_{cp})^2$	$(y_i - y_{cp})^2$	$(z_i - z_{cp})^2$
-0,28571	-0,08571	-4,14286	0,081633	0,007347	17,16327
0,014286	-0,68571	-2,14286	0,000204	0,470204	4,591837
-0,18571	0,114286	-2,14286	0,03449	0,013061	4,591837
0,014286	0,314286	-0,14286	0,000204	0,098776	0,020408
0,014286	0,014286	1,857143	0,000204	0,000204	3,44898
0,114286	0,114286	2,857143	0,013061	0,013061	8,163265
0,314286	0,214286	3,857143	0,098776	0,045918	14,87755
0	-1,2E-15	3,55E-15	0,228571	0,648571	52,85714

i	$(x_i - x_{cp}) * (y_i - y_{cp})$	$(x_i - x_{cp}) * (z_i - z_{cp})$	$(z_i - z_{cp}) * (y_i - y_{cp})$
1	0,024489796	1,184	0,355
2	-0,010	-0,031	1,469
3	-0,02122449	0,398	-0,245
4	0,004	-0,002	-0,045
5	0,000	0,027	0,027
6	0,013061224	0,327	0,327
7	0,067346939	1,212	0,827
Σ	0,078571429	3,114	2,714

r	-2,54	1,264911	6,2	-0,32388
---	-------	----------	-----	----------

r_{xy}	0,204068
r_{xz}	0,895974
r_{zy}	0,46358

матрица A		
1	0,204068	0,895974
0,204068	1	0,46358
0,895974	0,46358	1

R3	0,943292
Rz	0,887401
p	3
t	7,410929
k1	2
k2	4
α	0.05
$F_{кр}$	6,94

Рассчитали всё с помощью формул из задания 1 и получили: $t \approx 7,41 > F_{кр} = 6,94$, а это значит, что коэффициент R является значимым.

Лабораторная работа № 7 часть 2

Корреляция. Корреляционный анализ.

Вычисление множественного коэффициента корреляции.

Задание 1:

Изучили теория о вычислении множественного коэффициента корреляции.

Будут использованы данные формулы:

Выборочный множественный коэффициент корреляции:

$$R_i = \sqrt{1 - \frac{|A|}{A_{ii}}},$$

Из данной формулы для $z = 3$ получим:

$$R_i = \sqrt{\frac{r_{ij}^2 + r_{ik}^2 - 2r_{ij} r_{ik} r_{jk}}{1 - r_{jk}^2}}$$

Множественный коэффициент корреляции значимо отличается от нуля, если статистика

$$t = \frac{R^2(n-p)}{(1-R^2)(p-1)} > F_{кр}, \quad (4)$$

Коэффициенты r :

$$r = \frac{n \sum_{l=1}^n x_l y_l - (\sum_{l=1}^n x_l) (\sum_{j=1}^n y_j)}{\sqrt{n \sum_{l=1}^n x_l^2 - (\sum_{l=1}^n x_l)^2} \sqrt{n \sum_{j=1}^n y_j^2 - (\sum_{j=1}^n y_j)^2}}$$

Коэффициент корреляции:

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

Задание 2:

Задача 2.1. Проведено исследование шести фермерских хозяйств для изучения зависимости урожайности зерновых культур (Z , ц/га) от качества пашни (X , в баллах) и количества внесенных минеральных удобрений на 1 га пашни (Y , кг/га).

Результаты приведены в таблице 1.

Определить тесноту связи между величиной Z и величинами X и Y , используя множественный коэффициент корреляции. Проверить на уровне $\alpha = 0.05$ значимость R_z .

X	26	35	36	40	41	45
Y	2,1	2,3	2,4	2,6	2,9	3,0
Z	18,00	21,0	22,1	25,3	28,0	28,5

i	x_i	y_i	z_i	x_i^2	y_i^2	z_i^2	$x_i y_i$	$x_i z_i$	$y_i z_i$
1	26	2,1	18	676	4,41	324	54,6	468	37,8
2	35	2,3	21	1225	5,29	441	80,5	735	48,3
3	36	2,4	22,1	1296	5,76	488,41	86,4	795,6	53,04
4	40	2,6	25,3	1600	6,76	640,09	104	1012	65,78
5	41	2,9	28	1681	8,41	784	118,9	1148	81,2
6	45	3	28,5	2025	9	812,25	135	1282,5	85,5
Σ	223	15,3	142,9	8503	39,63	3489,75	579,4	5441,1	371,62
Средняя	37,16667	2,55	23,81667	1417,167	6,605	581,625	96,56667	906,85	61,93667

$x_i - x_{cp}$	$y_i - y_{cp}$	$z_i - z_{cp}$	$(x_i - x_{cp})^2$	$(y_i - y_{cp})^2$	$(z_i - z_{cp})^2$
-11,1667	-0,45	-5,81667	124,6944	0,2025	33,83361
-2,16667	-0,25	-2,81667	4,694444	0,0625	7,933611
-1,16667	-0,15	-1,71667	1,361111	0,0225	2,946944
2,833333	0,05	1,483333	8,027778	0,0025	2,200278
3,833333	0,35	4,183333	14,69444	0,1225	17,50028
7,833333	0,45	4,683333	61,36111	0,2025	21,93361
0	0	0	214,8333	0,615	86,34833

i	$(x_i - x_{cp}) * (y_i - y_{cp})$	$(x_i - x_{cp}) * (z_i - z_{cp})$	$(z_i - z_{cp}) * (y_i - y_{cp})$
1	5,025	64,953	2,618
2	0,542	6,103	0,704
3	0,175	2,003	0,258
4	0,142	4,203	0,074
5	1,342	16,036	1,464
6	3,525	36,686	2,108
Σ	10,75	129,9833333	7,225

r	64,5	35,90265	15,3	0,11742
---	------	----------	------	---------

r_{xy}	0,935234
r_{xz}	0,954355
r_{zy}	0,991456

матрица A		
1	0,935234	0,954355
0,935234	1	0,991456
0,954355	0,991456	1

R3	0,999301
Rz	0,993714
p	3
t	118,1844
k1	2
k2	3
α	0.05
$F_{кр}$	9,55

Рассчитали всё с помощью формул из задания 1 и получили: $t = 118,1844 > F_{кр} = 9,55$, а это значит, что коэффициент R является значимым.

Задание 3:

Задача 3.1 При изучении финансовой деятельности компании в течение некоторого времени был собран статистический материал. В таблице содержатся данные о ежемесячной прибыли Z (тыс. усл. ед.), расходах на рекламу X (тыс. усл. ед.) и вложении капитала в ценные бумаги Y (тыс. усл. ед.).

Определить тесноту связи между переменной Z и переменными X и Y с помощью выборочного множественного коэффициента корреляции R и определить его значимость на уровне $\alpha = 0.05$.

Z	10	12	12	14	16	17	18
X	0,2	0,5	0,3	0,5	0,5	0,6	0,8
Y	0,80	0,2	1	1,2	0,9	1	1,1

i	x_i	y_i	z_i	x_i^2	y_i^2	z_i^2	$x_i y_i$	$x_i z_i$	$y_i z_i$
1	0,2	0,80	10	0,04	0,64	100	0,16	2	8
2	0,5	0,2	12	0,25	0,04	144	0,1	6	2,4
3	0,3	1	12	0,09	1	144	0,3	3,6	12
4	0,5	1,2	14	0,25	1,44	196	0,6	7	16,8
5	0,5	0,9	16	0,25	0,81	256	0,45	8	14,4
6	0,6	1	17	0,36	1	289	0,6	10,2	17
7	0,8	1,1	18	0,64	1,21	324	0,88	14,4	19,8
Σ	3,4	6,20	99	1,88	6,14	1453	3,09	51,2	90,4
Средняя	0,485714	0,885714	14,14286	0,268571	0,877143	207,5714	0,441429	7,314286	12,91429

$x_i - x_{cp}$	$y_i - y_{cp}$	$z_i - z_{cp}$	$(x_i - x_{cp})^2$	$(y_i - y_{cp})^2$	$(z_i - z_{cp})^2$
-0,28571	-0,08571	-4,14286	0,081633	0,007347	17,16327
0,014286	-0,68571	-2,14286	0,000204	0,470204	4,591837
-0,18571	0,114286	-2,14286	0,03449	0,013061	4,591837
0,014286	0,314286	-0,14286	0,000204	0,098776	0,020408
0,014286	0,014286	1,857143	0,000204	0,000204	3,44898
0,114286	0,114286	2,857143	0,013061	0,013061	8,163265
0,314286	0,214286	3,857143	0,098776	0,045918	14,87755
0	-1,2E-15	3,55E-15	0,228571	0,648571	52,85714

i	$(x_i - x_{cp}) * (y_i - y_{cp})$	$(x_i - x_{cp}) * (z_i - z_{cp})$	$(z_i - z_{cp}) * (y_i - y_{cp})$
1	0,024489796	1,184	0,355
2	-0,010	-0,031	1,469
3	-0,02122449	0,398	-0,245
4	0,004	-0,002	-0,045
5	0,000	0,027	0,027
6	0,013061224	0,327	0,327
7	0,067346939	1,212	0,827
Σ	0,078571429	3,114	2,714

r	-2,54	1,264911	6,2	-0,32388
---	-------	----------	-----	----------

r_{xy}	0,204068
r_{xz}	0,895974
r_{zy}	0,46358

матрица A		
1	0,204068	0,895974
0,204068	1	0,46358
0,895974	0,46358	1

R3	0,943292
Rz	0,887401
p	3
t	7,410929
k1	2
k2	4
α	0.05
$F_{кр}$	6,94

Рассчитали всё с помощью формул из задания 1 и получили: $t \approx 7,41 > F_{кр} = 6,94$, а это значит, что коэффициент R является значимым.