

Практичне заняття № 4-5

Регресійний та кореляційний аналіз експериментальних даних

Мета роботи: виконати аналізу статистичних даних, отриманих при дослідження процесу різання чистовими торцевими фрезами, оснащеними надтвердими матеріалами.

Регресійний і кореляційний аналіз – статистичні методи дослідження. Це найпоширеніші способи показати залежність якого-небудь параметра від однієї або декількох незалежних змінних.

Теоретичні відомості

Регресійний і кореляційний аналіз – статистичні методи дослідження. Це найпоширеніші способи показати залежність якого-небудь параметра від однієї або декількох незалежних змінних.

Кореляційний аналіз допомагає встановити, є між показниками в одній або двох вибірках зв'язок чи ні. Наприклад, між часом роботи верстата та вартістю ремонту, ціною техніки та тривалістю експлуатації тощо.

Якщо зв'язок є, то чи тягне збільшення одного параметра підвищення (позитивна кореляція) або зменшення (негативна) іншого. Кореляційний аналіз допомагає аналітикові визначитися, чи можна за величиною одного показника передбачити можливе значення іншого.

Для розрахунку регресії та поля кореляції використовуємо програму Excel.

Коефіцієнт кореляції позначається r . Варіюється в межах від +1 до -1. Класифікація кореляційних зв'язків для різних сфер буде відрізнятися. При значенні коефіцієнта 0 лінійної залежності між вибірками не існує.

Розглянемо, як за допомогою засобів Excel знайти коефіцієнт кореляції.

Для знаходження парних коефіцієнтів застосовується функція КОРРЕЛ.

Приклад 1: Визначити, чи існує взаємозв'язок між часом роботи токарного верстата та вартістю його обслуговування.

	А	В
	Время работы станка, тыс.ч.	Ст-сть обслужив., тыс.руб.
1		
2	0,5	7,5
3	0,6	7,75
4	0,7	7,25
5	0,8	7,4
6	0,9	7,9
7	1	8
8	1,1	8,5
9	1,2	8,4
10	1,3	8,35
11	1,4	8,55
12	1,5	8,7
13	1,6	9,1
14	1,7	8,8

Ставимо курсор у будь-яке гніздо й натискаємо кнопку *fx*.

1. У категорії «Статистичні» вибираємо функцію КОРРЕЛ.
2. Аргумент «Масив 1» - перший діапазон значень – час роботи верстата: A2:A14.
3. Аргумент «Масив 2» - другий діапазон значень – вартість ремонту: B2:B14. Тиснемо ОК.

<i>fx</i> =КОРРЕЛ(A2:A14;B2:B14)		
В	С	Д
Ст-сть обслужив., тыс.руб.		
7,5	0,924771	

Щоб визначити тип зв'язку, потрібно подивитися абсолютне число коефіцієнта (для кожної сфери діяльності є своя шкала).

Для кореляційного аналізу декількох параметрів (більше 2) зручніше застосовувати «Аналіз даних» (надбудова «Пакет аналізу»). У списку потрібно вибрати кореляцію й позначити масив. Усе.

Отримані коефіцієнти відобразяться в кореляційній матриці. На зразок такої:

	А	В	С
1		Столбец 1	Столбец 2
2	Столбец 1	1	
3	Столбец 2	0,92477117	1

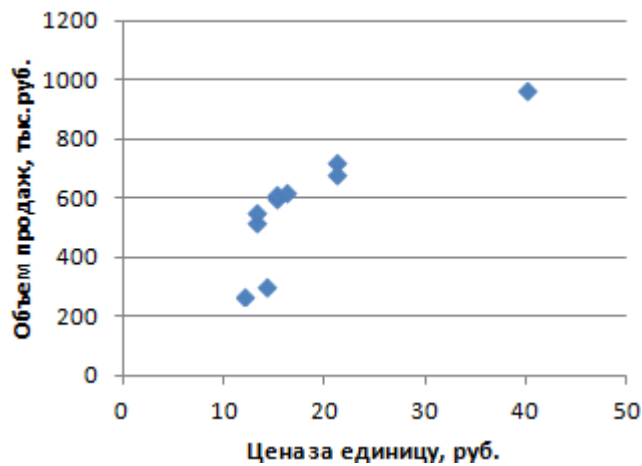
КОРЕЛЯЦІЙНО-РЕГРЕСІЙНИЙ АНАЛІ

За практикою ці дві методики часто застосовуються разом. Приклад.

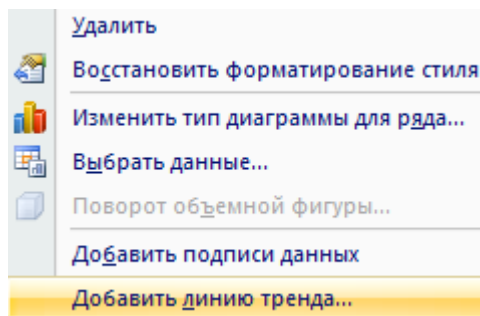
	А	В	С
1		Цена за ед., руб.	Объем продаж, тыс.руб.
2	n	X	Y
3	1	13	520
4	2	13	550
5	3	15	600
6	4	15	610
7	5	16	620
8	6	21	724
9	7	21	680
10	8	14	300
11	9	40	962
12	10	12	270

1. Будуємо кореляційне поле: «Вставка» - «Діаграма» - «Крапкова діаграма» (дає порівнювати пари). Діапазон значень – усі числові дані таблиці.

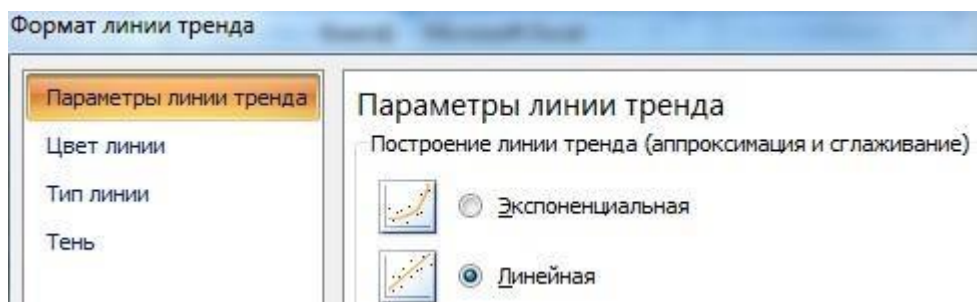
Поле кореляції



2. Клацаємо лівою кнопкою миші по будь-якій крапці на діаграмі. Потім правою. У меню, що відкрився, вибираємо «Додати лінію тренда».

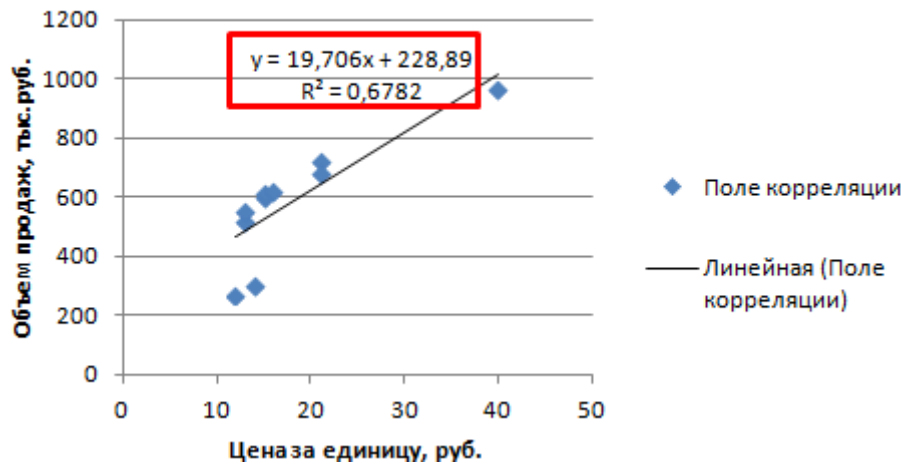


3. Призначаємо параметри для лінії. Тип – «Лінійна». Унизу – «Показати рівняння на діаграмі».



4. Тиснемо «Закрити».

Поле кореляции



Тепер стали видні й дані регресійного аналізу та її рівняння.

2. Регресійний аналіз в Excel показує вплив одних значень (самостійних, незалежних) на залежну змінну. Приміром, як залежить кількість активного населення від числа підприємств, величини заробітної плати й ін. параметрів. Або: як впливають параметри процесу різання на вихідні показники процесу обробки (сили, знос, шорсткість обробки, тощо).

Результат аналізу дозволяє виділяти пріоритети, і ґрунтуючись на головних факторах, прогнозувати, планувати розвиток процесів обробки.

Регресія буває:

- лінійною ($v = a + bx$);
- параболічною ($y = a + bx + cx^2$);
- експонентною ($y = a * \exp(bx)$);
- статичною ($y = a * x^b$);
- гіперболічною ($y = b/x + a$);
- логарифмічною ($y = b * \ln(x) + a$);
- показовою ($v = a * bx$)

Розглянемо на прикладі побудова регресійної моделі в Excel і інтерпретацію результатів. Виберемо лінійний тип регресії.

Завдання. На 6 підприємствах була проаналізована середньомісячна заробітна плата й кількість співробітників, що звільнилися. Необхідно визначити залежність числа співробітників, що звільнилися, від середньої зарплати.

	A	B	C
1		К-во ув.	З/п
2		y	x
3	1	60	100
4	2	35	150
5	3	20	200
6	4	20	250
7	5	15	300
8	6	15	350

Модель лінійної регресії має такий вигляд:

$$B = a_0 + a_1x_1 + \dots + a_kx_k.$$

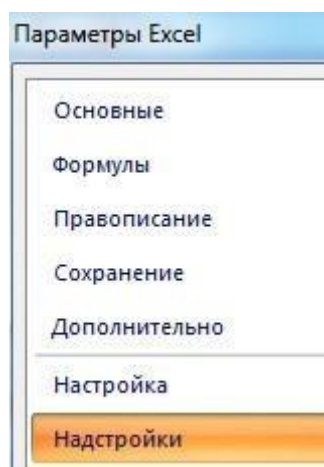
де a – коефіцієнти регресії, x_i – що впливають змінні, i – число факторів.

У нашому прикладі в якості Y виступає показник працівників, що звільнилися. фактор, що впливає, x – заробітна плата.

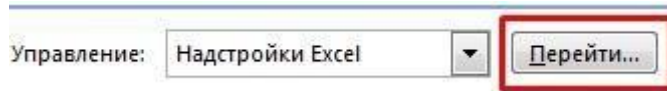
В Excel існують вбудовані функції, за допомогою яких можна розрахувати параметри моделі лінійної регресії. Але швидше це зробить надбудова «Пакет аналізу».

Активуємо потужний аналітичний інструмент:

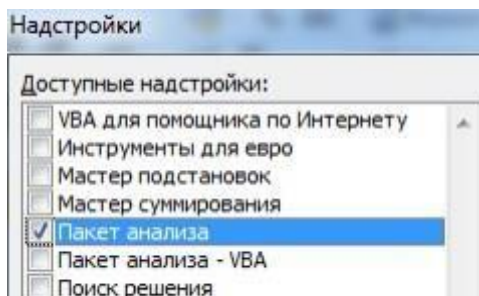
1. Натискаємо кнопку «Офіс» і переходимо на вкладку «Параметри Excel». «Надбудови».



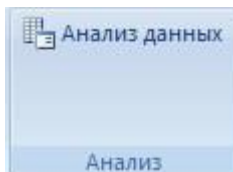
2. Унизу, під списком, що випадає, у поле «Керування» буде напис «Надбудови Excel» (якщо його немає, натисніть на прапорець праворуч і виберіть кнопку «Перейти»). Тиснемо.



3. Відкривається список доступних надбудов. Вибираємо «Пакет аналізу» і натискаємо ОК.

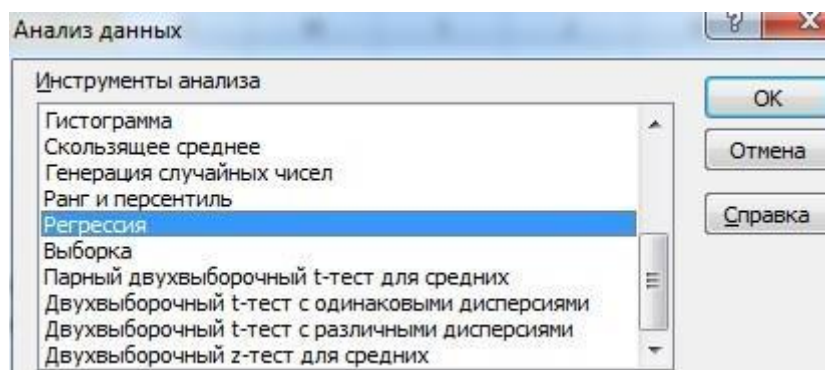


Після активації надбудова буде доступна на вкладці «Лані»



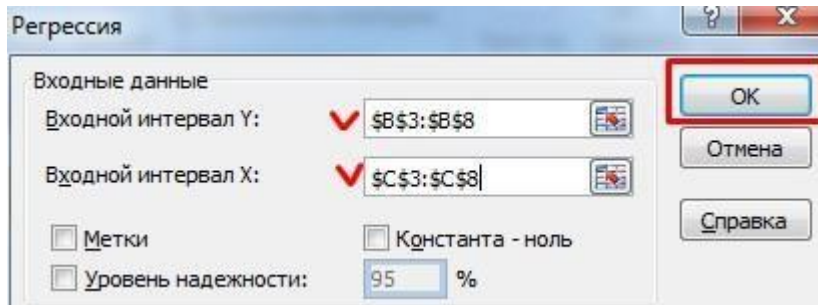
Тепер займемося безпосередньо регресійним аналізом.

1. Відкриваємо меню інструмента «Аналіз даних». Вибираємо «Регресія».



2. Відкриється меню для вибору вхідних значень і параметрів виводу (де відобразити результат). У полях для вихідних даних вказуємо діапазон

описуваного параметра (Y) фактора, що й впливає на нього (X). Решта можна й не заповнювати.



3. Після натискання ОК, програма відобразить розрахунки на новому аркуші (можна вибрати інтервал для відображення на поточному аркуші або призначити вивід у нову книгу).

	A	B	C	D	E	F	G
1	ВЫВОД ИТОГОВ						
3	<i>Регрессионная статистика</i>						
4	Множественный R	0,868736918					
5	R-квадрат	0,754703833					
6	Нормированный R-квадрат	0,693379791					
7	Стандартная ошибка	9,710083125					
8	Наблюдения	6					
10	<i>Дисперсионный анализ</i>						
11		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>ачимость F</i>	
12	Регрессия	1	1160,357143	1160,357143	12,30681818	0,024714	
13	Остаток	4	377,1428571	94,28571429			
14	Итого	5	1537,5				
16			<i>Стандарт</i>			<i>Нижние</i>	<i>Верх</i>
		<i>Коэффициент</i>	<i>ная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>95%</i>	<i>95%</i>
17	Y-пересечение	64,14285714	11,17212274	5,741331225	0,004560379	33,12407	95,16
18	Переменная X 1	-0,162857143	0,046423077	-3,508107493	0,024714164	-0,29175	-0,03

У першу чергу звертаємо увагу на R-Квадрат і коефіцієнти.

R-Квадрат – коефіцієнт детермінації. У нашому прикладі – 0,755, або 75,5%. Це означає, що розрахункові параметри моделі на 75,5% пояснюють залежність між досліджуваними параметрами. Чим вище коефіцієнт детермінації, тем якісніше модель. Добре – вище 0,8. Погано – менше 0,5 (такий аналіз чи навряд можна вважати резонним). У нашому прикладі – «непогано».

Коефіцієнт 64,1428 показує, яким буде Y, якщо всі змінні в розглянутій моделі будуть рівні 0. Тобто на значення аналізованого параметра впливають і інші фактори, не описані в моделі.

Коефіцієнт $-0,16285$ показує вагомість змінної X на Y . Тобто середньомісячна заробітна плата в межах даної моделі впливає на кількість, що звільнилися з вагою $-0,16285$ (це невеликий ступінь впливу). Знак «-» указує на негативний вплив: чим більше зарплата, тим менше працівників звільнилися.

2. Практична частина

Розрахунок регресії та коефіцієнта кореляції

1. Обрати варіант виконання завдання за даними експериментальних досліджень процесу торцевого фрезерування плоских поверхонь чистовими торцевими фрезами, оснащеними НТМ
2. Виконати розрахунки регресії та коефіцієнта кореляції для залежностей сил різання P_X, P_Y, P_Z від глибини різання t .
3. Навести графіки, рівняння, зробити висновки.

№ з/п	Подача на зуб S_z , мм/зуб	Глибина різання t , мм	Значення складових P_X, P_Y, P_Z сили різання та відношень P_X/P_Z і P_Y/P_Z					
			$P_{X, H}$	$P_{Y, H}$	$P_{Z, H}$	R, H	P_X/P_Z	P_Y/P_Z
1	2	3	4	5	6	7	8	9
1	Варіант 1 0,025	0,02	18	40	21	49	0,86	1,90
2		0,04	15	112	22	115	0,68	5,09
3		0,06	21	127	37	134	0,66	3,43
4		0,08	27	125	48	137	0,56	2,60
5		0,10	32	132	48	144	0,67	2,75
6		0,12	35	133	55	148	0,64	2,42
7	Варіант 2 0,032	0,02	16	44	24	53	0,67	1,83
8		0,04	17	120	26	124	0,65	4,61
9		0,06	22	131	37	138	0,59	3,54
10		0,08	29	155	49	165	0,59	3,16
11		0,10	34	188	49	197	0,69	3,84
12		0,12	40	186	62	200	0,65	3,00
13	Варіант 3 0,04	0,02	19	36	26	48	0,73	1,38
14		0,04	19	120	23	124	0,83	5,22
15		0,06	25	130	41	139	0,61	3,17

16		0,08	32	157	52	168	0,62	3,02
17		0,10	36	190	47	199	0,77	4,04
18		0,12	42	198	59	211	0,71	3,36

19	Варіант 4 0,05	0,02	17	41	29	53	0,59	1,41
20		0,04	19	120	27	123	0,7	4,44
21		0,06	29	152	44	161	0,66	3,45
22		0,08	32	174	54	185	0,59	3,22
23		0,10	40	189	56	201	0,71	3,38
24		0,12	41	200	62	213	0,66	3,22
25		Варіант 5 0,063	0,02	19	44	33	58	0,58
26	0,04		15	124	33	129	0,45	3,76
27	0,06		29	141	50	148	0,58	2,82
28	0,08		26	168	56	179	0,46	3
29	0,10		42	189	55	201	0,76	3,44
30	0,12		44	184	68	201	0,65	2,71
31	Варіант 6 0,08	0,02	18	49	34	62	0,53	1,44
32		0,04	21	125	33	131	0,64	3,79
33		0,06	37	152	50	164	0,74	3,04
34		0,08	35	163	62	178	0,57	2,63
35		0,10	40	185	61	199	0,66	3,03
36		0,12	49	190	68	208	0,72	2,79
37	Варіант 7 0,1	0,02	19	52	38	65	0,5	1,36
38		0,04	24	138	35	144	0,69	3,94
39		0,06	37	160	55	173	0,67	2,91
40		0,08	39	176	70	193	0,56	2,51
41		0,10	42	179	72	213	0,58	2,49
42		0,12	52	192	81	215	0,64	2,37

43	Варіант 8 0,125	0,02	24	52	33	66	0,73	1,58
44		0,04	20	149	36	155	0,56	4,14
45		0,06	38	157	59	172	0,64	2,66
46		0,08	39	171	77	192	0,51	2,22
47		0,10	51	184	79	207	0,65	2,33
48		0,12	55	196	82	219	0,67	2,39
49		Варіант 9 0,16	0,02	27	60	43	79	0,63
50	0,04		21	150	38	156	0,55	3,95
51	0,06		39	158	63	175	0,62	2,5
52	0,08		41	184	78	204	0,53	2,36
53	0,10		55	185	79	209	0,7	2,34
54	0,12		67	212	91	240	0,74	2,32
55	Варіант 10 0,2		0,02	24	60	43	78	0,56
56		0,04	27	150	58	163	0,47	2,59
57		0,06	44	155	68	175	0,65	2,28
58		0,08	42	182	87	206	0,48	2,09
59		0,10	61	193	88	221	0,69	2,19
60		0,12	67	204	98	236	0,68	2,08
61		Варіант 11 0,25	0,02	32	63	48	85	0,67
62	0,04		22	131	48	141	0,46	2,73
63	0,06		44	154	73	176	0,6	2,11
64	0,08		42	174	93	202	0,45	1,87
65	0,10		65	187	95	220	0,68	1,97
66	0,12		72	196	110	236	0,65	1,78
67	Варіант 12 0,315		0,02	33	70	53	94	0,62
68		0,04	26	135	55	148	0,47	2,45
69		0,06	53	133	75	162	0,71	1,77
70		0,08	58	178	98	211	0,59	1,82
71		0,10	73	181	104	221	0,7	1,74
72		0,12	87	195	116	243	0,75	1,68
73		Варіант 13	0,02	31	68	57	94	0,54

74	0,4	0,04	27	127	50	139	0,54	2,54
75		0,06	59	139	85	173	0,69	1,64
76		0,08	60	165	114	209	0,53	1,45
77		0,10	72	176	121	225	0,6	1,45
78		0,12	95	184	138	249	0,69	1,33
79	Варіант 14 0,5	0,02	32	71	63	100	0,51	1,12
80		0,04	26	124	55	138	0,47	2,25
81		0,06	60	133	104	179	0,58	1,28
82		0,08	66	157	138	219	0,48	1,14
83		0,10	90	176	142	243	0,63	1,24
84		0,12	112	196	170	283	0,66	1,15