

Корреляционные зависимости

Определение

- Зависимости между величинами, каждая из которых подвергается не контролируемому полностью разбросу, называются **корреляционными** зависимостями.
- Раздел математической статистики, который исследует такие зависимости, называется **корреляционным анализом**.
- Корреляционный анализ изучает усредненный закон поведения каждой из величин в зависимости от значений другой величины, а также **меру** такой зависимости.

Пусть важной характеристикой некоторой сложной системы является фактор **A**. На него могут оказывать влияние одновременно многие другие факторы: **B**, **C**, **D** и так далее.

Обычно требуется требуется определить:

- оказывает ли фактор **B** какое-либо заметное регулярное влияние на фактор **A**;
 - какие из факторов **B**, **C**, **D** и так далее оказывают наибольшее влияние на фактор **A**.
-

Пример для первого типа задач

Фактором **A** является средняя успеваемость учащихся школы, фактором **B** — финансовые расходы школы на хозяйственные нужды: ремонт здания, обновление мебели, эстетическое оформление помещения и т. п.

Здесь влияние фактора **B** на фактор **A** не очевидно.

Сбор статистики

1	А № п/п	В Затраты (руб/чел.)	С Успеваемость (ср.балл)
2	1	50	3,81
3	2	345	4,13
4	3	79	4,30
5	4	100	3,96
6	5	203	3,87
7	6	420	4,33
8	7	210	4
9	8	137	4,21
10	9	463	4,4
11	10	231	3,99
12	11	134	3,9
13	12	100	4,07
14	13	294	4,15
15	14	396	4,1
16	15	77	3,76
17	16	480	4,25
18	17	450	3,88
19	18	496	4,50
20	19	102	4,12
21	20	150	4,32



Коэффициент корреляции

- Мерой корреляционной зависимости является величина, которая называется *коэффициентом корреляции*.
 - Коэффициент корреляции (обычно обозначаемый греческой буквой **ρ**) есть число, заключенное в диапазоне от -1 до $+1$;
 - Если это число по модулю близко к $|1|$, то имеет место сильная корреляция, если к 0 , то слабая;
 - Близость ρ к $+1$ означает, что возрастанию одного набора значений соответствует возрастание другого набора, близость к -1 означает обратное;
-

-
- В Excel функция вычисления коэффициента корреляции называется **КОРРЕЛ** и входит в группу статистических функций.
 - Она запрашивает два диапазона значений и возвращает коэффициент корреляции.
 - Для нашего примера $r = 0,500273843$. Эта величина говорит о среднем уровне корреляции.
-

Пример для второго типа задач

Провести исследование по определению зависимости успеваемости учащихся старших классов от двух факторов: обеспеченности школьной библиотеки учебниками и обеспеченности школы компьютерами.

Сбор статистики

Обеспечение учебного процесса				
№	Обеспеченность учебниками (%)	Успеваемость (ср.балл)	Обеспеченность компьютерами (%)	Успеваемость (ср.балл)
1	50	3,81	10	3,98
2	78	4,15	25	4,01
3	94	4,69	19	4,34
4	65	4,37	78	4,41
5	99	4,53	45	3,94
6	87	4,23	32	3,62
7	100	4,73	90	4,6
8	63	3,69	21	4,24
9	79	4,08	34	4,36
10	94	4,2	45	3,99
11	93	4,32	67	4,5
	$\rho = 0,780931$		$\rho = 0,572465$	

-
- Корреляция между обеспеченностью учебниками и успеваемостью сильнее, чем корреляция между компьютерным обеспечением и успеваемостью.
 - Отсюда можно сделать вывод, что пока еще книга остается более значительным источником знаний, чем компьютер.
-

Вопросы

- Что такое корреляционная зависимость?
 - Что такое корреляционный анализ?
 - Какие типы задач можно решать с помощью корреляционного анализа?
 - Какая величина является количественной мерой корреляции? Какие значения она может принимать?
-