

Домашнее задание 12

Исследование зависимостей

Вы научитесь по наблюдениям выявлять статистическую зависимость между парами признаков и оценивать силу этой зависимости.

Задание

1. На любом языке программирования напишите программу, вычисляющую по заданной таблице сопряжённости признаков

- статистику хи-квадрат,
- Р-значение для критерия хи-квадрат Фишера-Пирсона,
- нормированные коэффициенты связи, основанные на статистике хи-квадрат,
- меры прогнозы Гутмана,
- коэффициенты контингенции и ассоциации (если таблица имеет размер 2×2).

С помощью написанной программы решите следующие задачи.

Задача 1.1

По данным исследования «Мониторинг социальных и экономических перемен в России» составлена следующая таблица сопряженности.

Вопрос А: «Как бы Вы оценили в настоящее время материальное положение вашей семьи?» (Хорошее, среднее, плохое).

Вопрос В: «Как бы Вы оценили в целом политическую обстановку в России?» (Благополучная, напряжённая, взрывоопасная критическая).

А \ В	Благоприятная	Напряженная	Взрывоопасная критическая
Хорошее	12	48	47
Среднее	20	478	666
Плохое	11	160	701

Имеется ли зависимость между признаками А и В? Прокомментируйте характер связи между А и В с помощью различных коэффициентов.

Задача 1.2

Изучается зависимость выбираемой абитуриентами специальности от пола. Было опрошено 480 человек и получены следующие результаты:

Специальность \ Пол	М	Ж
Естественные науки	168	92
Гуманитарные науки	85	135

Можно ли на уровне доверия 0,95 считать, что между полом и выбором специальности есть связь? Оцените силу связи с помощью различных коэффициентов.

2. На любом языке программирования напишите программу, которая по результатам наблюдений за двумя количественными признаками:

- вычисляет выборочный коэффициент корреляции и Р-значение для критерия проверки гипотезы о нулевом значении коэффициента корреляции;
- вычисляет коэффициент Спирмена и Р-значение для критерия Спирмена;
- вычисляет коэффициент Крамера и Р-значение для критерия Крамера;
- разбивая диапазон количественных признаков на несколько интервалов строит по наблюдениям таблицу сопряжённости, вычисляет нормированные коэффициенты связи (Пирсона, Крамера и др.), меры прогноза Гутмана и Р-значение для критерия хи-квадрат Фишера-Пирсона.

С помощью написанной программы решите следующие задачи.

Задача 2.1

Группа пловцов из 15 человек принимает участие в местной спортивной олимпиаде. В программе состязаний два заплыва: 50 м вольным стилем и 100 м баттерфляем. Результаты соревнований следующие:

№	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
50 м вольн.	22,49	22,56	23,45	22,58	24,3	24,2	23,47	23,5	24,48	25,02	23,04	23,24	25,2	24,61	26,02
100 м батт.	52,93	53,4	53,7	53,36	61,8	55,2	53,54	58,33	60,4	60,3	54,28	53,6	62,24	54,45	61,52

Определите степень связи между результатами в разных заплывах.

Задача 2.2

В таблице приведены данные об урожайности пшеницы и картофеля на соседних полях. Что можно сказать о зависимости показателей урожайности картофеля и пшеницы?

Годы	1926	1927	1928	1929	1930	1931	1932	1933	1934	1935	1936
Пшеница	20,1	23,6	26,3	19,9	16,7	23,2	31,4	33,5	28,2	35,3	29,3
Картофель	7,2	7,1	7,4	6,1	6,0	7,3	9,4	9,2	8,8	10,4	8,0