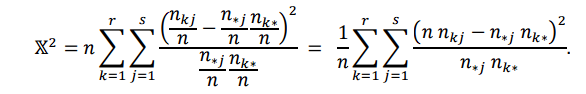
«Исследование статистических связей»

Шарифуллин Ринас Рамилевич гр. 09-131

Вариант – 13

**Задание 4\_1**

1. Говорят, что уровень успеваемости студентов по математической статистики зависит от их роста.
2. В выборку вошло n = 103 человек. В данных, представленных ниже, горизонтальный столбец таблицы показывает рост студента (в см.). Вертикальный – баллы учащегося за экзамен по математической статистике.
3. Рост студента есть случайная величина с функцией распределения 𝐹1, а 𝐹2 – функция распределения баллов.
4. Ожидается, что 𝐹1 и 𝐹2 зависимые случайные величины. Т.о., нулевая гипотеза 𝐇0: 𝐹1 и 𝐹2 – независимые сл. вел. при альтернативе 𝐇1: 𝐹1 и 𝐹2 – зависимые сл. вел.
5. Уровень значимости 𝛼 = 0.03.
6. Применим критерий сопряженности хи-квадрат. Области признаков разобьём на r = 3 и s = 5 интервалы. Ожидания будут подтверждены, если 𝑋2 примет маленькое значение, т.е. критическая область имеет вид {𝑋2 > 𝐶}.



1. При справедливости нулевой гипотезы функцию распределения статистики 𝑋2 можно приблизить функцией хи-квадрат распределения 𝕂𝕙𝕚(𝑥 | (𝑟 - 1) \* (s - 1)) = 𝐏H0{𝑋2 ≤ 𝑥) с (𝑟 - 1) \* (s - 1) = 8 степенями свободы.
2. Критическая константа 𝐶𝛼 находится как решение неравенства

𝐏H0 {𝑋2 ≥ 𝐶𝛼) = 1 − 𝕂𝕙𝕚(𝐶𝛼 | 8) = 0.03,

т.е. 𝐶𝛼 равна квантили порядка 0.97 хи-квадрат распределения с 8 ст. св. По таблице хи-квадрат распределения (процедурой Excel), находим 𝐶𝛼 = 17,01.

а. Вид критерия: гипотеза однородности отвергается, если {𝑋2 ≥ 17,01}.

9.

1. По представленным данным найдено

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 172,45 | 182,83 | >182,83 |  |
| >59 | 0 | 3 | 7 | 10 |
| 59 | 8 | 13 | 9 | 30 |
| 54,05 | 6 | 8 | 8 | 22 |
| 49,1 | 10 | 5 | 5 | 20 |
| 44,15 | 11 | 9 | 1 | 21 |
|  | 35 | 38 | 30 |  |
| Статистика | | | 19,74 | |
| степени свободы | | | 8 | |
| 3%-я критическая область | | | 17,01 | |
| Гипотеза независимости | | | отвергается | |
| с критическим уровнем значимости | | | *p-val* < 0,0041 | |

1. Критический уровень значимости p-value вычисляется по формуле *p-val* = 1 − 𝕂𝕙𝕚(19,74 | 8) = 0,0041. Есть все основания считать, что успеваемость студента по математической статистике зависит от его роста.

**Задание 4\_2**

1. Говорят, что уровень успеваемости студентов по математической статистики зависит от их роста.
2. В выборку вошло n = 103 человек. В данных, представленных ниже, горизонтальный столбец таблицы показывает рост студента (в см.). Вертикальный – баллы учащегося за экзамен по математической статистике.
3. Рост студента есть случайная величина с функцией распределения 𝐹1, а 𝐹2 – функция распределения баллов.
4. Ожидается, что 𝐹1 и 𝐹2 зависимые случайные величины. Т.о., нулевая гипотеза 𝐇0: 𝐹1 и 𝐹2 – независимые сл. вел. при альтернативе 𝐇1: 𝐹1 и 𝐹2 – зависимые сл. вел.
5. Уровень значимости 𝛼 = 0.01.
6. Применим критерий независимости компонент двумерного случайного вектора. Статистика Стьюдента вычисляется по следующей формуле:



Критическая область будет принимать вид:

1. Функция распределения тестовой статистики совпадает с функцией распределения Стьюдента 𝕊𝕥(𝑛−2) с 𝑛 − 2 = 101 степенями свободы.
2. Критическая константа 𝐶𝛼 находится из уравнения

2(1 – 𝕊𝕥101(𝐶)) = 0.01,

т.е. – равна верхней 0.005-квантили распределения Стьюдента с 101 степенями свободы.

1. Воспользовавшись таблицей (пакетом Excel), нашли, что 𝐶𝛼 = 2,87.
2. Окончательный вид критической области |t| ≥ 2,87

а. По представленным данным найдено

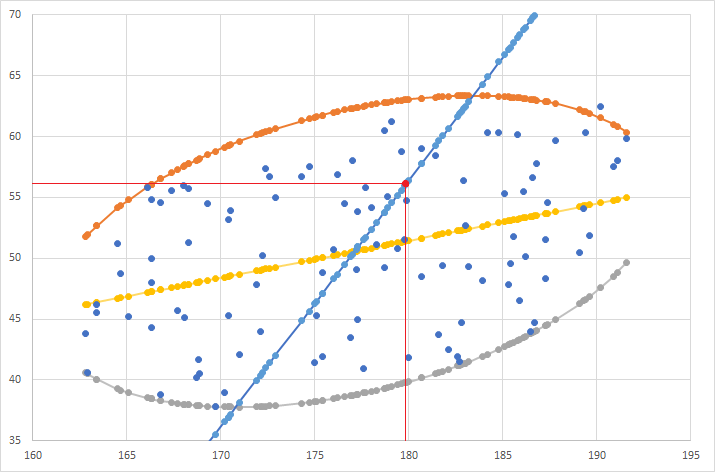
|  |  |  |
| --- | --- | --- |
|  | x | y |
| Среднее, | 177,12 | 50,56 |
| Дисперсия, s2 | 65,99 | 40,91 |
| Станд. отклонение, s | 8,12 | 6,39 |
| Объем выборки, n | 103 | 103 |
| Коэффициент корреляции, R | | 0,39 |
| Преобразование Стьюдента, t | | 4,28 |
| 1%-я критическая область | | 2,87 |
| Гипотеза независимости | | отвергается |
| с критическим уровнем значимости | | p-val < 0,000043 |

1. Критический уровень значимости p-value вычисляется по формуле *p-val* = 2(1 – 𝕊𝕥101(4,28)) = 0,000043. Есть все основания считать, что успеваемость студента по математической статистике зависит от его роста.

**Задание 4\_3**

1. По результатам независимых измерений значения баллов за экзамен по математической статистике и роста n = 103 студентов найти оценки коэффициентов линейной среднеквадратической регрессии роста учащегося () на полученную оценку (); представить график линии регрессии в поле всех данных; найти прогноз значения при значении = 56; дать оценку точности прогноза, изобразить эллипс рассеяния.
2. Для проверки предположения в эксперименте приняло участие n = 103 студента, у которых были измерены рост и уровень успеваемости по математической статистике.
3. По представленным данным найдено

|  |  |
| --- | --- |
| Коэффициент линейной регрессии | 0,49 |
| Уравнение регрессии на | 0,49 + 152,16 |
| Прогноз при = 56 | 179,8 |
| Коэфф.корреляции | 0,39 |
| Стандарт.отклонение наблюдений | 8,12 |
| Оценка стандарт. ошибки прогноза | 7,49 |



**Вывод.** При таком невысоком значении коэффициента корреляции (*R* = 0,39; стандартная ошибка прогноза равна 7,49) прогностические качества линии регрессии очень низкие.