

Это последняя домашка и больше не будет.

Никаких сегодня домашек — по домам, люди

Корреляционный анализ

1. Из 360 поступающих на DS-поток юношей было принято 97, а из 82 девушек — 40. Проверить гипотезу о независимости факта поступления от пола с $\alpha = 0.01$.
2. Группа Б05-024 перед написанием контрольной по статистике решила для храбрости постоять немного в планке. Ниже приведены результаты: в первой строке записано количество минут, которое человек простоял в планке, во второй — балл за контрольную.

0.5	2.5	8.5	2	6	10	0.5	1	0.5	3	6.5	0.5	2	3.5	6.0	1.5	1
3	0	2	2	2	3	0.5	2	1	2.5	3	3	2	2	3	0	3

Посчитайте коэффициент корреляции Спирмена вручную (можно использовать Python, но нельзя готовые реализации) и проверьте с помощью него гипотезу о том, что данные метрики независимы. Не забудьте использовать взвешенные ранги.

- 3*. Найдите дисперсию $D\tau$ коэффициента корреляции Кенделла при верности нулевой гипотезы о независимости выборок.
- 4*. Пусть $\hat{\rho}$ — коэффициент корреляции Пирсона, построенный по выборкам \mathbf{X} и \mathbf{Y} . Докажите, что есть эти выборки независимы и нормально распределены, то

$$\hat{\rho} \cdot \sqrt{\frac{n-2}{1-\hat{\rho}^2}} \sim T_{n-2}.$$

Указание. Найдите условное распределение $\hat{\rho}$ при фиксированной \mathbf{X} и покажите, что оно одинаково при любом условии.