

Практика проверки гипотез

5 вопросов

1
point

1.

По данным опроса, 75% работников ресторанов утверждают, что испытывают на работе существенный стресс, оказывающий негативное влияние на их личную жизнь. Крупная ресторанная сеть опрашивает 100 своих работников, чтобы выяснить, отличается ли уровень стресса работников в их ресторанах от среднего. 67 из 100 работников отметили высокий уровень стресса.

Посчитайте достигаемый уровень значимости, округлите ответ до четырёх знаков после десятичной точки.

0.0822

1
point

2.

Представим теперь, что в другой ресторанной сети только 22 из 50 работников испытывают существенный стресс. Гипотеза о том, что 22/50 соответствует 75% по всей популяции, методом, который вы использовали в предыдущей задаче, отвергается. Чем это может объясняться? Выберите все возможные варианты.



Сотрудники чрезмерно оптимистично отвечали на вопрос, потому что опрос не был анонимным



Эта сеть менее популярна, поэтому в её ресторанах меньше поток посетителей; отсюда меньший уровень стресса у работников



Сеть известна низкими зарплатами и большой текучкой кадров, поэтому туда идут работать люди, не склонные переживать о сохранении рабочего места



Условия работы в этой сети лучше, чем в среднем

1
point

3.

The Wage Tract — заповедник в округе Тома, Джорджия, США, деревья в котором не затронуты деятельностью человека со времён первых поселенцев. Для участка заповедника размером 200x200 м имеется информация о координатах сосен (sn — координата в направлении север-юг, we — в направлении запад-восток, обе от 0 до 200).

pin.es.txt

Проверим, можно ли пространственное распределение сосен считать равномерным, или они растут кластерами.

Загрузите данные, поделите участок на 5x5 одинаковых квадратов размера 40x40 м, посчитайте количество сосен в каждом квадрате (чтобы получить такой же результат, как у нас, используйте функцию `scipy.stats.binned_statistic_2d`).

Если сосны действительно растут равномерно, какое среднее ожидаемое количество сосен в каждом квадрате? В правильном ответе два знака после десятичной точки.

23.36

1
point

4.

Чтобы сравнить распределение сосен с равномерным, посчитайте значение статистики хи-квадрат для полученных 5x5 квадратов. Округлите ответ до двух знаков после десятичной точки.

150.59

1
point

5.

Насколько велико это значение? Если нулевая гипотеза справедлива, с какой вероятностью его можно было получить случайно?

Нулевое распределение статистики — хи-квадрат с $25 - 1 = 24$ степенями свободы (поскольку у равномерного распределения, с которым мы сравниваем данные, нет ни одного оцениваемого по выборке параметра, число степеней свободы $K - 1$, где K — количество интервалов).

Посчитайте достигаемый уровень значимости.

Если вы используете функцию `scipy.stats.chi2.cdf`, в качестве значения параметра `df` нужно взять 24 (это число степеней свободы); если функцию `scipy.stats.chisquare` — параметр `ddof` нужно брать равным 0 (это как раз количество параметров теоретического распределения, оцениваемых по выборке).

Отвергается ли гипотеза равномерности на уровне значимости 0.05?

- ☐ Достигаемый уровень значимости меньше 0.05, гипотеза равномерности отвергается.
- ☐ Достигаемый уровень значимости больше 0.05, гипотеза равномерности не отвергается.

☐

I, **לאניד גיאדי**, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account. Узнайте больше о Кодексе чести Coursera

Сдать тест

