

# Непараметрические критерии

7 вопросов

1  
point

1.

Какую из этих гипотез нельзя проверить с помощью критерия знаков?

- ☐ Гипотеза о равенстве средних в двух независимых выборках
- ☐ Гипотеза о равенстве среднего выборки некоторой константе
- ☐ Гипотеза о равенстве средних в двух связанных выборках

1  
point

2.

Какую из этих гипотез о средних двух независимых выборок проверяет критерий Манна-Уитни?

- ☐  $H_0: \mathbb{E}X_1 = \mathbb{E}X_2$  против  $H_1: \mathbb{E}X_1 < \neq > \mathbb{E}X_2$
- ☐  $H_0: P(X_1 > X_2) = \frac{1}{2}$  против  $H_1: P(X_1 > X_2) < \neq > \frac{1}{2}$
- ☐  $H_0: \text{med}X_1 = \text{med}X_2$  против  $H_1: \text{med}X_1 < \neq > \text{med}X_2$
- ☐  $H_0: F_{X_1}(x) = F_{X_2}(x)$  против  $H_1: F_{X_1}(x) = F_{X_2}(x + \Delta), \Delta < \neq > 0$

1  
point

3.

У вас есть две регрессионные модели и двадцать наборов данных. На  
обыкновенной выборке каждого набора вы построили обе модели и

обучающей подвыборке каждого набора вы настраиваете все модели и считаете среднеквадратичную ошибку на тестовой подвыборке набора. Какими из этих критериев **не стоит** проверять гипотезу о том, что среднее качество регрессионных моделей одинаково?

- ☐ Перестановочный критерий со статистикой  $T = \frac{1}{n} \sum_{i=1}^n X_{1i} - \frac{1}{n} \sum_{i=1}^n X_{2i}$  и  $C_{2n}^n$  вариантами расстановки объектов по выборкам
- ☐ Критерий знаковых рангов
- ☐ Критерий Манна-Уитни
- ☐ Критерий знаков
- ☐ Перестановочный критерий со статистикой  $T = \sum_{i=1}^n (X_{1i} - X_{2i})$  и  $2^n$  вариантами расстановки знаков у попарных разностей
- ☐ Критерий Стьюдента для двух независимых выборок

1  
point

4.

Давайте вернёмся к данным выживаемости пациентов с лейкоцитарной лимфомой из видео про критерий знаков:

49, 58, 75, 110, 112, 132, 151, 276, 281, 362\*

Измерено остаточное время жизни с момента начала наблюдения (в неделях); звёздочка обозначает цензурирование сверху — исследование длилось 7 лет, и остаточное время жизни одного пациента, который дожил до конца наблюдения, неизвестно.

Поскольку цензурировано только одно наблюдение, для проверки гипотезы  $H_0: medX = 200$  на этих данных можно использовать критерий знаковых рангов — можно считать, что время дожития последнего пациента в точности равно 362, на ранг этого наблюдения это никак не повлияет.

Критерием знаковых рангов проверьте эту гипотезу против двусторонней альтернативы, введите достигаемый уровень значимости, округлённый до четырёх знаков после десятичной точки.

Введите ответ здесь

1  
point

5.

В ходе исследования влияния лесозаготовки на биоразнообразие лесов острова Борнео собраны данные о количестве видов деревьев в 12 лесах, где вырубка не ведётся:

22, 22, 15, 13, 19, 19, 18, 20, 21, 13, 13, 15,

и в 9 лесах, где идёт вырубка:

17, 18, 18, 15, 12, 4, 14, 15, 10.

Проверьте гипотезу о равенстве среднего количества видов в двух типах лесов против односторонней альтернативы о снижении биоразнообразия в вырубаемых лесах. Используйте ранговый критерий. Чему равен достигаемый уровень значимости? Округлите до четырёх знаков после десятичной точки.

Введите ответ здесь

1  
point

6.

28 января 1986 года космический шаттл "Челленджер" взорвался при взлёте. Семь астронавтов, находившихся на борту, погибли. В ходе расследования причин катастрофы основной версией была неполадка с резиновыми уплотнительными кольцами в соединении с ракетными ускорителями. Для 23 предшествовавших катастрофе полётов "Челленджера" известны температура воздуха и появление повреждений хотя бы у одного из уплотнительных колец.

challenger.txt

С помощью бутстрепа постройте 95% доверительный интервал для разности средних температур воздуха при запусках, когда уплотнительные кольца повреждались, и запусках, когда повреждений не было. Чему равна его ближайшая к нулю граница? Округлите до четырёх знаков после запятой.

Чтобы получить в точности такой же доверительный интервал, как у нас:

- установите `random seed = 0` перед первым вызовом функции `get_bootstrap_samples`, один раз
- сделайте по 1000 псевдовыборок из каждой выборки.

Введите ответ здесь

1  
point

7.

На данных предыдущей задачи проверьте гипотезу об одинаковой средней температуре воздуха в дни, когда уплотнительные кольца повреждались, и дни, когда повреждений не было. Используйте перестановочный критерий и двустороннюю альтернативу. Чему равен достигаемый уровень значимости? Округлите до четырёх знаков после десятичной точки.

Чтобы получить такое же значение, как мы:

- установите `random seed = 0`;
- возьмите 10000 перестановок.

Введите ответ здесь



I, **Владислав Лисеев**, 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

4 вопросов без ответа

Сдать тест

