

ВОПРОСЫ НА ПОНИМАНИЕ

Упражнение 1 (10 баллов). Ответьте на следующие вопросы:

1. Какие гипотезы называются простыми, а какие сложными?
2. Чем отличаются параметрические критерии от непараметрических? В каких ситуациях можно использовать параметрические критерии, а в каких лучше использовать непараметрические?
3. Что такое уровень значимости? А фактический уровень значимости?
4. С помощью каких критериев можно проверить нормальность выборки?
5. Что такое ранг наблюдения? Когда предпочтительнее использовать ранговые критерии?

ЗАДАЧИ

КРИТЕРИИ СОГЛАСИЯ

Упражнение 2 (10 баллов). Для данных из файла `employees.xls` постройте гистограмму для переменной `SALARY`. Похоже ли распределение зарплат на нормальный закон? Проверьте нормальность с помощью изученных критериев.

Упражнение 3 (15 баллов). Методом перебора найдите наиболее подходящий тип распределения для выборки, содержащейся в файле `sample.csv`. Сначала постройте гистограмму, чтобы получить представление о форме плотности распределения. Затем изучите квантильные графики для разных распределений. Если будет такая возможность, попробуйте применить какие-то критерии согласия.

КРИТЕРИИ ОДНОРОДНОСТИ

В каждом следующем упражнении необходимо воспользоваться всеми изученными критериями, которые можно использовать в данных постановках. Кроме этого, решения необходимо сопроводить краткими комментариями к Вашим действиям и выводами, которые Вы сделали на основе анализа.

Упражнение 4 (15 баллов). В файле `anorexia.txt` записан вес пациентов до начала терапии анорексии и после ее окончания. Была ли терапия эффективна?

Упражнение 5 (15 баллов). В файле `seattle.txt` записаны цены на недвижимость в одном из районов Сиэтла в 2001 году и в 2002 году (объекты недвижимости выбирались случайно). Изменились ли в среднем цены в этом районе за год?

Упражнение 6 (15 баллов). В рамках исследования эффективности препарата метилфенидат пациенты с синдромом дефицита внимания и гиперактивности в течение недели принимали либо метилфенидат, либо плацебо. В конце недели каждый пациент проходили тест на способность к подавлению импульсивных поведенческих реакций. На втором этапе плацебо и препарат менялись, и после недельного курса каждый испытуемый проходили второй тест. Был ли эффект от применения препарата? Данные находятся в файле `methylphenidate.txt`.

Упражнение 7 (20 баллов). В файле `mtcars.csv` находятся данные из американского журнала Motor Trend 1974 года. Они описывают расход топлива в зависимости от 10 характеристик автомобиля (все автомобили 1973-1974 года). А именно, в данных вы найдете следующие столбцы:

- 1) `mpg` — расход топлива (миль/галлон);
- 2) `cyl` — количество цилиндров мотора;
- 3) `disp` — объем мотора (куб. дюйм);

- 4) **hp** — мощность мотора (в лошадиных силах);
- 5) **drat** — передаточное число редуктора заднего моста;
- 6) **wt** — вес (тысяч фунтов);
- 7) **qsec** — время, за которое автомобиль проезжает 1/4 мили;
- 8) **vs** — тип двигателя (0 = V-образный, 1 = рядный);
- 9) **am** — тип коробки передач (0 = автоматическая, 1 = ручная);
- 10) **gear** — количество передних передач;
- 11) **carb** — количество карбюраторов.

Влияет ли тип двигателя на расход топлива? А тип коробки передач?