

Библиотеки Python для Data Science: продолжение

Урок 2. Анализ данных и проверка статистических гипотез.

1. В чём различие между зависимыми и независимыми выборками?

Если можно установить гомоморфную пару, когда одному случаю из выборки X соответствует только один случай из выборки Y и наоборот, для каждого случая в двух выборках (основание взаимосвязи является важным для измеряемого на выборках признака), такие выборки называются: *Зависимыми* - выборки которые имеют временную связь через отношения, попарно связаны, в противоположность независимым. Для зависимых выборок принципиально сравнение начальной и конечной величин - их дельту, в то время как для независимых выборок принципиально сравнение их абсолютных величин.

2. Когда применяются параметрические статистические критерии, а когда — их непараметрические аналоги?

Поскольку параметрические методы оценки требуют знания закона распределения, а непараметрические такого знания не требуют, то и соответствующие им критерии применяются с учетом этих особенностей.

Например, для определения параметрического *F-критерия равенства дисперсий Фишера*, как следует из названия, требуется знание дисперсий, что отсылает нас к знанию закона распределения. А критерий *Хи-квадрат Пирсона* оперирует частотами наблюдаемых параметров, что не требует знания закона распределения.