**Статистическая мощность эксперимента в Excel**

Недавно прочитал книгу Майкла Херцога с соавторами [Статистика и планирование эксперимента для непосвященных](https://baguzin.ru/wp/?p=26128). В ней я в очередной раз встретился со *статистической мощностью*. До сих пор я относился к этой статистике несколько академически. Что называется, не чувствовал её на кончиках пальцев. Некоторые разделы книги Херцога меня заинтересовали, я построил пару моделей в Excel, и до меня дошла суть статистической мощности. Я понял, что статистическая мощность – вероятность получить в эксперименте статистически значимые результаты. Кроме того, я нашел формулу расчета статистической мощности в Excel.

Изображение выглядит как стол

Автоматически созданное описание

Рис. 1. Нормально распределенные случайные величины. Выборка 1 – N(0;1), Выборка 2 – N(1;1)

Если вы не знакомы со статистическим выводом, рекомендую начать с [t-статистики Стьюдента](http://baguzin.ru/wp/?p=19184).

#### Суть процедуры статистического вывода

Допустим, мы извлекаем две выборки по 20 элементов в каждой из двух генеральных совокупностей. Такая схема эксперимента называется двухвыборочной. Наша нулевая гипотеза Н0 состоит в том, что математические ожидания генеральных совокупностей равны μ1 = μ2. В качестве альтернативной выберем самую общую гипотезу μ1 ≠ μ2.

Но мы не знаем истинные значения μ1 и μ2. В нашем распоряжении только выборочные средние значения х̅1 и х̅2. Причем они не равны. Задача процедуры статистического вывода – ответить на вопрос: есть ли основания при заданном уровне значимости отвергнуть нулевую гипотезу? Т.е., можно ли объяснить различия между средними значениями х̅1 и х̅2 естественной вариабельностью при отборе значений из одной генеральной совокупности (μ1 = μ2). Или это настолько маловероятно, что различия между х̅1 и х̅2 скорее связаны с тем, что они извлечены из разных генеральных совокупностей (μ1 ≠ μ2).

#### Модель в Excel

Рассмотрим модель в Excel (см. рис. 1), в которой мы генерим нормально распределенные случайные величины из двух генеральных совокупностей: N(0;1) и N(1;1). Здесь N(0;1) – нормальное распределение с математическим ожиданием μ1 = 0 и σ1 = 1. N(1;1) – нормальное распределение с математическим ожиданием μ2 = 1 и σ2 = 1.

Для дальнейшей иллюстрации примера я сохранил сгенерированные формулами случайные числа, как значения. В приложенном Excel-файле на листе «Рис. 1» можно найти формулы, дающие случайные числа. Эти значения волатильны. Они изменяются после любого пересчета в Excel.

#### Двухвыборочный t-тест в Пакете анализа

В Excel есть средство, которое позволяет выполнить двухвыборочный t-тест. Пройдите *Данные* –> *Анализ данных* (кнопка расположена на правом краю ленты). Выберите *Двухвыборочный t-тест с одинаковыми дисперсиями* или *Двухвыборочный t-тест с различными дисперсиями*. Первый вид теста выполняется, если у нас есть основания предполагать, что неизвестные дисперсии двух выборок одинаковы.

Изображение выглядит как стол

Автоматически созданное описание

Рис. 2. Двухвыборочный t-тест с одинаковыми дисперсиями (надстройка *Пакет анализа*)

Если *Анализ данных* у вас на ленте не отображается, пройдите *Файл* –> *Параметры*. В окне *Параметры Excel* перейдите на вкладку *Надстройки*. В нижней части окна в поле *Управление* выберите *Надстройки* *Excel*. Кликните на кнопке *Перейти*. Поставьте галочку напротив *Пакет анализа*.

#### Двухвыборочный t-тест с помощью формул в Excel

В учебных целях такой подход я не могу рекомендовать, поскольку в нём все вычисления «находятся под капотом». Выполним шаги последовательно:

Изображение выглядит как стол

Автоматически созданное описание

Рис. 3. Двухвыборочный t-тест с одинаковыми дисперсиями (формулы)

#### Ошибка I-го рода

Выше мы сказали, что в качестве альтернативной рассматриваем самую общую гипотезу μ1 ≠ μ2. Такая гипотеза соответствует двустороннему тесту. Выполняя тест, мы определяем, куда попадет t-статистика, вычисленная на экспериментальных данных. В нашем примере t-статистика = 2,2341 больше критического tα равного 2,0244 для двустороннего теста с уровнем значимости α = 0,05.

Изображение выглядит как диаграмма

Автоматически созданное описание

Рис. 4. Графическое представление двухвыборочного t-теста

Двухвыборочный t-тест при заданном уровне значимости α = 0,05 позволяет отвергнуть нулевую гипотезу о равенстве математических ожиданий двух генеральных совокупностей. t-тест позволяет принять альтернативную гипотезу: Выборка 1 и Выборка 2 извлечены из разных генеральных совокупностей.

Заметим, что односторонний тест используется, если в качестве альтернативной принимается одна из направленных гипотез: μ1 > μ2 или μ1 < μ2.

#### Ошибка II-го рода

Выше мы задали уровень значимости α = 0,05. Уровень значимости также называют ошибкой I-го рода. Значение ошибки I-го рода, равное 0,05 означает, что с вероятностью 5% мы отвергнем нулевую гипотезу, хотя она будет верна. Ошибку I-го рода нельзя сделать очень маленькой, поскольку в этом случае возрастет частота ошибок II-го рода: мы будем придерживаться нулевой гипотезы, когда она перестала быть верной. Вероятность ошибка II-го рода соответствует площади под кривой распределения выборочных средних второй группы, отсекаемых t-значением.

Изображение выглядит как диаграмма

Автоматически созданное описание

Рис. 5. Ошибка II-рода. Расчеты см. в Excel-файле.

#### Статистическая мощность

Статистической мощностью называется величина *w = 1 – β*:

Изображение выглядит как диаграмма

Автоматически созданное описание

Рис. 6. Статистическая мощность

Статистическая мощность показывает, с какой вероятностью наш план эксперимента позволит отвергнуть нулевую гипотезу. Давайте проверим)) Проведем 1000 экспериментов с моделями, как на рис. 1, и подсчитаем t-значения. Можно ожидать, что около 870 t-значений будут более 2,0244. Распределение t-значений представлено на рис. 7. Этот анализ выполнен с помощью волатильных функций в Excel, так что вы можете перейти в Excel-файл на лист "Рис. 7", и понажимать клавишу F9, запуская перерасчет.

Изображение выглядит как диаграмма

Автоматически созданное описание

Рис. 7. Распределение t-значений в модельном эксперименте

Для представленных на рис. 7 данных 874 значения из 1000 лежат в области t > 2,0244.

#### Аналитическое решение для вычисления статистической мощности

В общем случае (для любой альтернативной гипотезы) вычислить ошибку II-го рода и статистическую мощность не представляется возможным. Это связано с тем, что для таких вычислений требуется предположение о размере эффекта d или значении μ2 (они взаимосвязаны). С другой стороны, оценку μ2 дает среднее по выборке х̅2. Основываясь на выборочной статистике можно оценить мощность предложенного плана эксперимента.

Что это нам дает? Оценив мощность эксперимента, мы получаем вероятность, с которой можем ожидать обнаружить эффект в последующих экспериментах.

Если в общем случае рассчитать статистическую мощность не получится, то для фиксированного размера эффекта в Excel есть аналитическое решение:

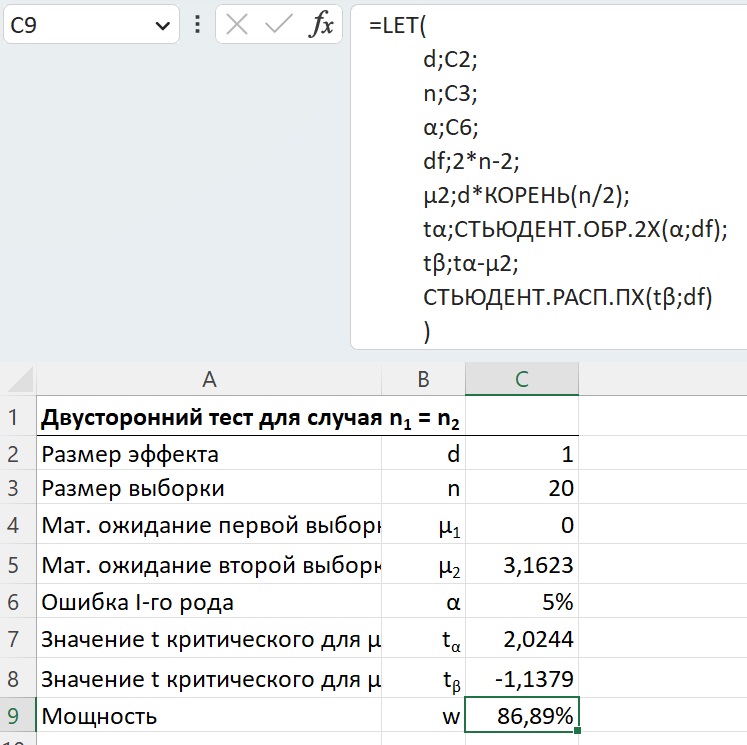


Рис. 8. Формула для расчета статистической мощности в Excel

Если вы не знакомы с функцией LET, рекомендую почитать [Упрощение формул Excel путем именования фрагментов с помощью функции LET](https://baguzin.ru/wp/?p=24081).

Здесь функция настроена на случай, когда размеры двух сравниваемых выборок равны: *n1 = n2 = n*. После объявления функции LET в последующих трех строках функция считывает с листа Excel три значения: размер эффекта d, размер выборки n, вероятность ошибки I-го рода α. А далее последовательно вычисляет число степеней свободы df, матожидание распределения, соответствующего альтернативной гипотезе μ2, критическое значение tα, значение tβ, соответствующее tα при переносе нуля в точку μ2, и наконец, площадь под кривой распределения μ2 справа от tβ. Вуаля! Это и есть статистическая мощность для заданных d, n, α.

Изображение выглядит как диаграмма

Автоматически созданное описание

Рис. 9. Обозначения на графике статистической мощности

#### Расчет размера выборки при заданной статистической мощности

Для экспериментатора может быть интересна и обратная задача: для заданных d, α, w рассчитать необходимый размер выборки n. Аналитическое решение в Excel я думаю не существует. Так как есть функции, возвращающие вероятность, например СТЬЮДЕНТ.РАСП()… есть функции, возвращающие t-значение, например, СТЬЮДЕНТ.ОБР(), но… нет функций, возвращающих число степеней свободы.

Это ограничение можно обойти, построив *Таблицу данных* (меню *Данные* –> *Прогноз* –> *Анализ "что если"* –> *Таблица данных*). Таблица и график позволяют определить *n*, соответствующее требуемой мощности:

Изображение выглядит как диаграмма

Автоматически созданное описание

Рис. 10. Графическое решение обратной задачи: для заданных d, α, w рассчитать необходимый размер выборки n

Понятно, что четыре переменные…

* размер эффекта d
* размер выборки n
* вероятность ошибки I-го рода α
* и статистическая мощность w

… однозначно связаны. Поэтому можно сформулировать и такой вопрос: при заданных n, α, w, какой размер эффекта можно обнаружить?

#### Калькулятор статистической мощности

И, наконец следует упомянуть, что в Инете есть бесплатные программы – калькуляторы, рассчитывающие статистическую мощность. Например, [G\*Power](https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower). На рис. 11 показан итог работы этой программы. На вход были поданы следующие условия:

* двусторонний t-критерий
* размер эффекта на генеральной совокупности d = 0,55
* частота ошибок I-го рода α = 0,05
* размеры выборок n1 = n2 = 40

На выходе получены значения

* параметр нецентральности δ; он же матожидание распределения альтернативной гипотезы μ2 = 2,4597
* критическое t-значение, tα = 1,9908
* число степеней свободы df = 78
* статистическая мощность w = 0,68

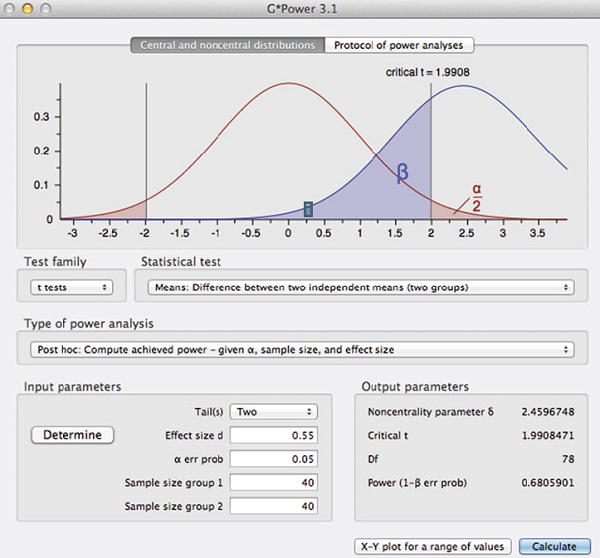


Рис. 11. Вывод программы G\*Power, вычисляющей мощность t-критерия

#### Использованные материалы

Майкл Херцог. [Статистика и планирование эксперимента для непосвященных](https://baguzin.ru/wp/?p=26128)

Statistical Power Analyses for Mac and Windows [G\*Power](https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower)

Конрад Карлберг. [Регрессионный анализ в Microsoft Excel](http://baguzin.ru/wp/?p=18972)

[t-статистика Стьюдента в Excel](http://baguzin.ru/wp/?p=19184)

Упрощение формул Excel путем именования фрагментов с помощью функции [LET](https://baguzin.ru/wp/?p=24081)