

Задание 1.

	Послали email	Открыли email
Группа А	10 000	2 200
Группа В	10 000	2 000

Маркетологи тестировали тему email-письма (разным группам послали email с разными темами). Целевой метрикой был процент открытий.

Какой теме стоит отдать предпочтение и почему? Объясните свое решение

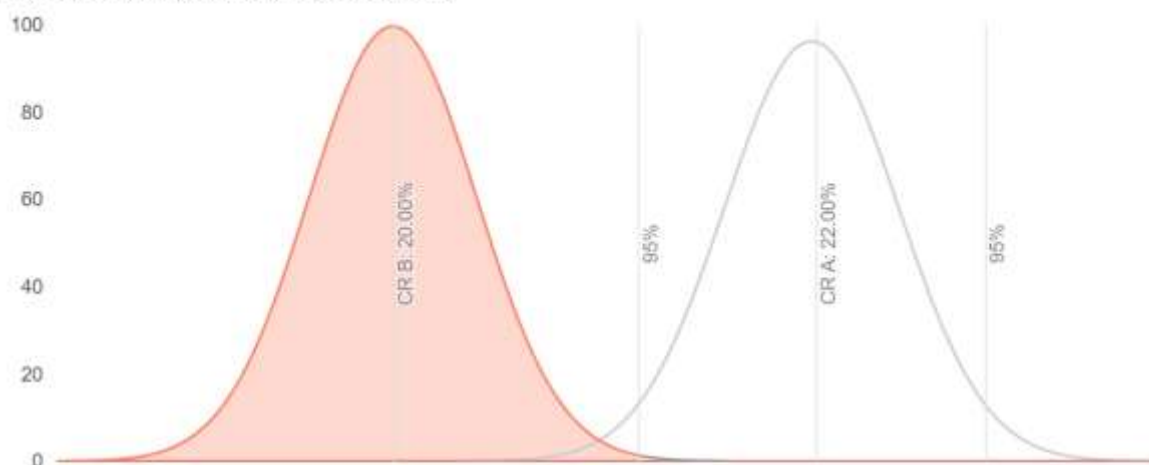
Решение:

Выдвинем H_0 – тема письма не оказывает никакого влияния (различие выборок неразличимо).

H_1 – значение различимо. В данном случае $p_value = 0.0003 < 0.05$, поэтому H_0 отвергаем.

Следовательно клиентам стоит направлять письма с темой, которые были направлены группе А.

Ожидаемые распределения вариации А и В.



Контроль скорости конверсии

Конверсии А / Посетители А

22,00%

Коэффициент конверсии В

Конверсии Б / Посетители Б

20,00%

Относительный рост коэффициента конверсии

$CR_B - CR_A / CR_A$

-9,09%

Наблюдаемая мощность

99,83%

значение р

0,0003

Z-оценка

$(CR_B - CR_A) / \sqrt{CR_A \cdot (1 - CR_A)}$

-3,4731

Чтобы активировать Windows, перейдите в раздел "Параметры".

Задание 2.

	Послали email	Открыли email
Группа А	10 000	2 070
Группа В	10 000	2 000

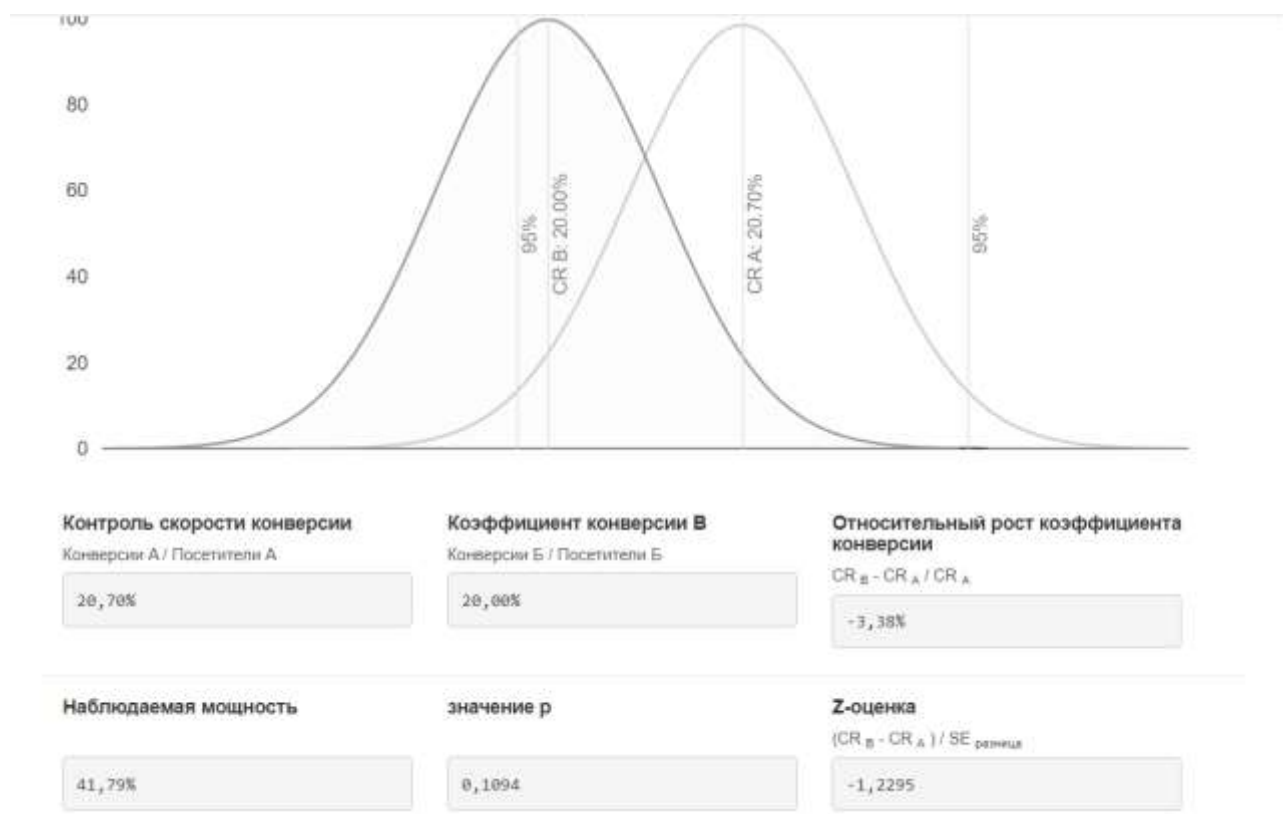
Кейс тот же, но итоги теста отличаются.

Решение:

Выдвинем H_0 – тема письма не оказывает никакого влияния (различие выборок неразличимо).

H_1 – значение различимо. В данном случае $p_value = 0.1094 > 0.05$, поэтому H_0 принимаем.

Следовательно стоит оставить все без изменений.



Задание 3.

	Размер группы	Совершили оплату
Тест	100	65
Контроль	100	50

Тестовой группе отправляется пуш с призывом совершить оплату, контрольной группе не отправляли. Является ли различие статистически значимым при уровне значимости 1%? А при 5%? H_0 наличие пуша не оказывает влияния

При уровне значимости в 1% имеем $p_value = 0.015 > 0.01 \Rightarrow H_0$ принимаем.

При уровне значимости в 5% имеем $p_value = 0.015 < 0.05 \Rightarrow H_0$ отвергаем.

И, как мы видим, результаты в 1 случае – не статзначимы, а в 2 – статзначимы.

Задание 4.

Хотим протестировать новую версию сайта. Ожидаем, что вырастет число регистраций.

Обычно у нас дневная конверсия из захода на сайт в регистрацию 50% и хотим поймать изменение на 5%-пунктов при уровне значимости 1% и мощности 80%.

Так же известно, что каждый день к нам на сайт приходит 100 уникальных пользователей.

Какой размер выборки нам нужен? Сколько дней требуется проводить АВ тест?

Решение:

Внесем наши данные в калькулятор.

Minimum sample size:

2001

unique visitors per test variation

Power

0.8

Confidence level

0.99

CR A

50

CR B

55

AB test duration

Minimum test duration 5.72 weeks *

Round up to a **AB-test period of 6 weeks** (discrete number of business cycles)

Minimum improvement needed

Minimum Conversion rate B of 55.97 % needed to run AB-test in max 4 weeks, with 700 unique visitors per week on your page *

This amounts to a **minimum relative improvement of 12% needed** in order to run the test at the required levels of power and confidence.

Результаты получил примерно такие же, как и по формулам.

Задание 5.

	Размер группы	Оставили заявку
Тест	1000	600
Контроль	1000	500

Проводим обзвон клиентов (тест – звоним, контроль – нет). Цель – чтоб клиент оставил заявку на продукт. Мы знаем, что средняя стоимость одного звонка 50 рублей (дозвонились или нет не важно), с одной заявки мы зарабатываем в среднем 180 рублей.

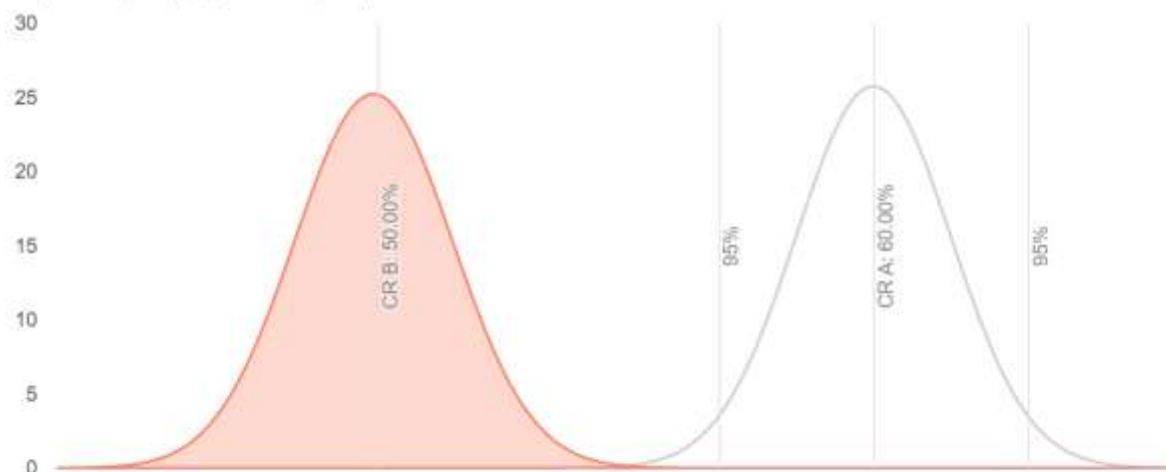
Проинтерпретируйте результаты AB теста

Решение:

На обзвон группы клиентов мы потратили 50 000. В тестовой группе удалось выручить 58 000, а в контрольной – 90 000. (с учетом издержек)

Несмотря на то, что обзвон в целом оказывает видимое влияние на число заявок:

Ожидаемые распределения вариации А и В.



Контроль скорости конверсии

Конверсии А / Посетители А

60,00%

Коэффициент конверсии В

Конверсии Б / Посетители Б

50,00%

Относительный рост коэффициента конверсии

$CR_B - CR_A / CR_A$

-16,67%

Наблюдаемая мощность

100,00%

значение р

0,0000

Z-оценка

$(CR_B - CR_A) / SE_{разница}$

Активация Windows

Чтобы активировать Windows, перейдите в раздел

Мы видим, что издержки достаточно большие. Поэтому логичнее отсавить все как есть (по крайней мере мы точно не потеряем доход).