

Один из студентов задал очень классный вопрос по поводу того, имеет ли смысл делать несбалансированные группы во время АВ-теста (например 10% контрол-группа и 90% тест-группа с новым функционалом). Ответ можно получить, если вспомнить как считается длительность АВ-теста:

Из объяснения расчетов длительности АВ-теста, которое тут было ранее:

$$d = (z_{\alpha/2} + z_{\beta})SE$$

Где:

$$SE = \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}} = \sqrt{\frac{\sigma^2}{n_{control}} + \frac{\sigma^2}{n}}$$

Фиксируем $n_{control}$ и решаем относительно n :

$$d = (z_{\alpha/2} + z_{\beta}) \sqrt{\frac{\sigma^2}{n_{control}} + \frac{\sigma^2}{n}}$$

$$n = \frac{(z_{\alpha/2} + z_{\beta})^2 n_{control} \sigma^2}{d^2 n_{control} - (z_{\alpha/2} + z_{\beta})^2 \sigma^2}$$

Теперь построим симуляцию – будем перебирать значение $n_{control}$ и посмотрим на размер тестовой группы и длительность теста в каждом конкретном случае (вдруг получится выигрыш по сравнению с 50/50).

Код симуляции:

```
control_group_sizes = range(100, 10000, 100)

p = 0.1
mean = p
sigma = p * (1 - p)
daily_users = 200

relative_significance = 0.5
d = mean * relative_significance

alpha = 0.01
power = 0.9
z = norm.ppf(power) + norm.ppf(1 - alpha / 2)

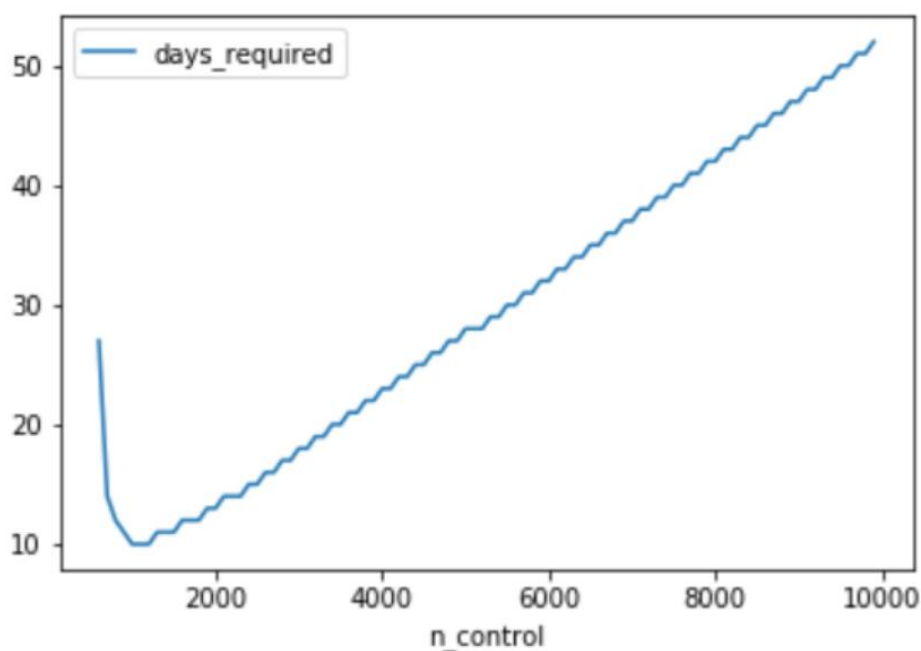
result = []
for n_control in control_group_sizes:
    n = (z ** 2 * n_control * sigma) / (d ** 2 * n_control - z ** 2 * sigma)
    result += [[n_control, n, n + n_control, (n + n_control)/daily_users]]

result = pd.DataFrame(result, columns = ['n_control',
                                         'n',
                                         'total_audience',
                                         'days_required']).round(2).astype(int)

result = result.query('total_audience > 0')
result
```

Результат:

	n_control	n	total_audience	days_required
5	600	4995	5595	27
6	700	2281	2981	14
7	800	1621	2421	12
8	900	1323	2223	11
9	1000	1153	2153	10
10	1100	1044	2144	10
11	1200	967	2167	10
12	1300	911	2211	11
13	1400	867	2267	11
14	1500	833	2333	11
15	1600	805	2405	12
16	1700	782	2482	12
17	1800	762	2562	12



Видно, что уменьшение контрол-группы приводит к росту размера тест-группы. Минимальная длительность теста достигается при соотношении 50/50. Вывод: уменьшение контрол-группы не дает никакого выигрыша.

Небольшое пояснение зачем это все нужно: обычно, когда перекашивают группы, руководствуются логикой "нам калькулятор сказал, что надо 10000 человек, давайте возьмем тестовую группу 90% от аудитории теста и наберем эти 10000 быстрее". А калькулятор, считал размеры групп, исходя из предположения, что группы будут одинаковые - по 50%.