A/B test.

- Can only test A or B, can't tell you what is missing.
- Need clear control, and clear metrics.

1. what A/B testing can do and can't do.

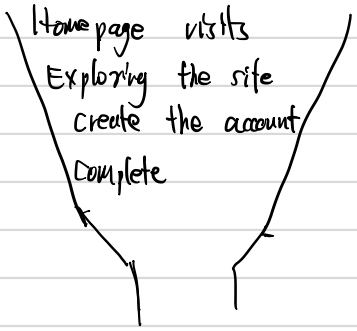
can't

- if too long or don't have data (buy a car, or refer)
- Suprisingly emotional. (New logo, needs long time to watch.)

can do

- test layout : clear control, clear metrics

- Customer funnel



- End to End.

1. Choose a metric

① Initial the hypothesis : 'changing the 'start now' button from orange to pink will increase the number of students exploring the site

② Refine the hypothesis :

- Which metric to use?
  - X. Total numbe of courses completed : will take too much time to get a result
  - X. Number of clicks : total number or fraction?
  - number of clicks / number of page views. : click-through rate / CTR.
  - ✓ unique visitors who click / unique visitors to page : click-through probability

③ updated hypothesis: Change the 'start now' button from orange to pink will increase the click-through-probability of the button

Click-through-rate vs click-through-probability
- if want to measure the usability of the button, use ··rate
  because users have a variety of different places on the page that they can choose to click on and rate will show how often they actually find the button
- If you want to know how they would get into the next level, use ·· probability.
  because you need to exclude the probability of double click, reload and so on

2. Review statistics

① which surprise you most? 150 900.? why?

a. Binomial distribution.:

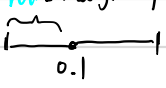
- with mean $= p$, standard deviation $= \sqrt{\frac{p(1-p)}{N}}$

When can use binomial
- 2 types of outcomes
- Independent.
- Identical distribution
  - p for all.

b. confidence interval

$\hat{p} = \frac{\#\text{ of users who clicked}}{\#\text{ of users.}} = \frac{100}{1000} = 0.1$ (N)

m = margin of error



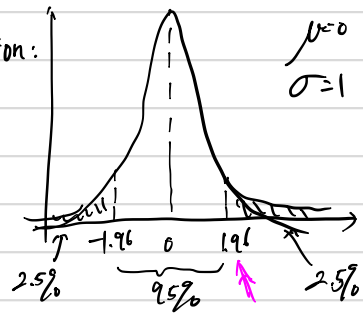
To use normal: check $N\cdot\hat{p} > 5$ and $N(1-\hat{p}) > 5$

$m = z^* SE$

$= z \times \sqrt{\frac{\hat{p}(1-\hat{p})}{N}}$

z: confidence level, SE: standard error

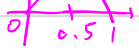
z distribution:



→ 1.96 for 97.5%

$m = 0.019$

$\hat{p}(1-\hat{p})$



- p away from 0.5 $p(1-p)$ ↓ m ↓
- N ↑ m ↓

two-tailed test ⇒ 99% → 1% → 0.5% ⇒ 99.5% ⇒ z = 2.58

exe: N=2000, X=300. confidence level = 99%. ⇒

$\hat{p} = \frac{300}{2000} = \frac{3}{20} = 0.15$

$m = 2.58 \times \sqrt{\frac{0.15 \times 0.85}{2000}} \approx 0.17$

② Establish statistical Significance

Hypothesis testing: How likely it is that your results occured by chance.

a. null hypothesis / baseline : no difference of probability between our experiment and control group

b. Alternative hypothesis :

⇓.

P_{con} P_{exp}.

a. Null hypothesis : H_0 $P_{con} = P_{exp} \Rightarrow P_{con} - P_{exp} = 0$

b. Alternative hypothesis H_A : $P_{con} \neq P_{exp} \Rightarrow P_{con} - P_{exp} \neq 0$

c. measure P_{con}, and P_{exp}
calculate $P(\hat{P}_{exp} - \hat{P}_{con} \mid H_0) = \alpha$

d. reject null if α is smeall enough
α is same type of significance threshold as a confidence interval
reject null if $\alpha < 0.05$