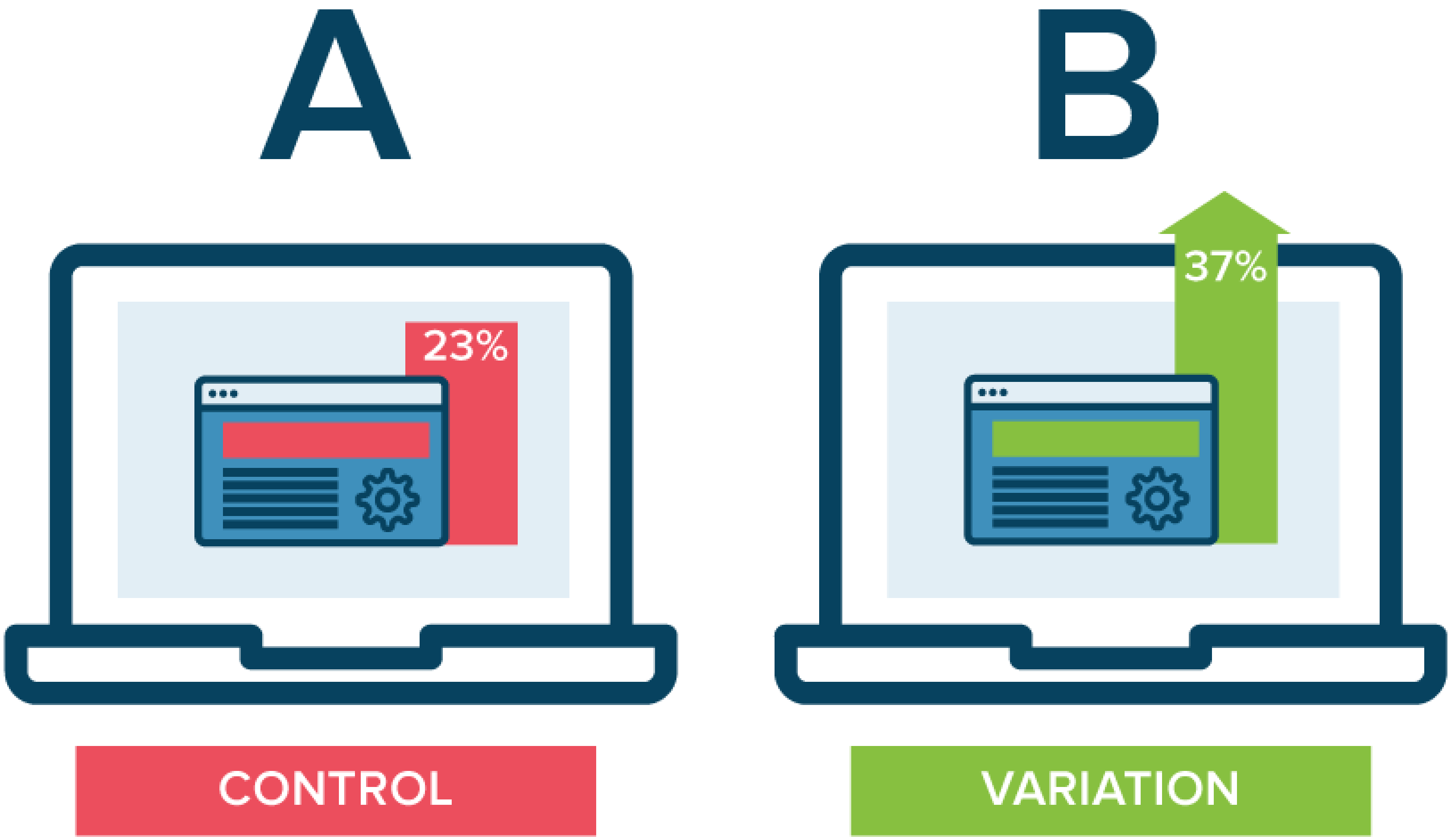
Bayesian A/B Testing



https://www.optimizely.com/optimization-glossary/ab-testing/

Bayes' Law for A/B testing

- We can do a t-test for A/B testing things like website design (e.g. red vs blue theme)
  - What are some problems with t-tests for large amounts of data like web traffic?

Bayes' Law for A/B testing

- We can do a t-test for A/B testing things like website design (e.g. red vs blue theme)
  - What are some problems with t-tests for large amounts of data like web traffic?
    - Everything looks significant because of the amount of data. t-tests were made for small data
    - Also assumes t-distributions (similar to normal distributions, but with slightly fatter tails)
    - t-stat inversely proportional to number of samples
  - Check out t.test.example.R in worldclass

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{{S_1}^2}{N_1} + \frac{{S_2}^2}{N_2}}}$$

http://rpsychologist.com/d3/NHST/

https://www.optimizely.com/optimization-glossary/ab-testing

https://ncalculators.com/math-worksheets/how-to-calculate-t-test.htm

R t-test and Bayesian A/B testing example

- gamma = (n – 1)!
- Mean = alpha / (alpha + beta)
- In R, dbeta(alpha, Beta)

$$f(x;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$$

$$= \frac{1}{\mathrm{B}(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}$$

R t-test and Bayesian A/B testing example

- Beta distribution good for A/B testing (binomial – 0 or 1 results)
- distribution of probability after S successes and F failures is given by Beta(S+1, F+1)
  - So assuming a 50/50 chance with 2 trials, we have Beta(2, 2). A completely blank prior would be Beta(1, 1), which would mean no trials

Beta Distribution

Parameters: $\alpha > 0$ and $\beta > 0$

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\alpha)} x^{\alpha-1}(1-x)^{\beta-1} \text{ for } 0 \le x \le 1$$

$$\mu = \frac{\alpha}{\alpha+\beta}$$

$$\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

$$x_{\text{mode}} = \frac{\alpha-1}{\alpha+\beta-2}$$

R t-test and Bayesian A/B testing example

- We can add the alpha/beta parameters from a prior and from data to get a posterior distribution
  - e.g. say we expect 30% conversion rate from a webpage for whatever reason, we could use beta(4, 8) as our prior distribution
  - Say we then find that 50 out of 450 people sign up, our evidence is showing a distribution of beta(51, 451)
  - We add the alpha/beta to get our posterior distribution, beta(55, 459). Then this becomes our prior the next time we re-evaluate the results

R t-test and Bayesian A/B testing example

- If we have one version of a webpage (A) that gets 50/450 to sign up, and B gets 37/425, which one is better and by how much?
- Add the data to our prior again, then randomly sample the two Beta distributions, divide the results of A by the results of B to get a distribution of how much better A is than B
- See bayes.ab.test.example.R file in worldclass