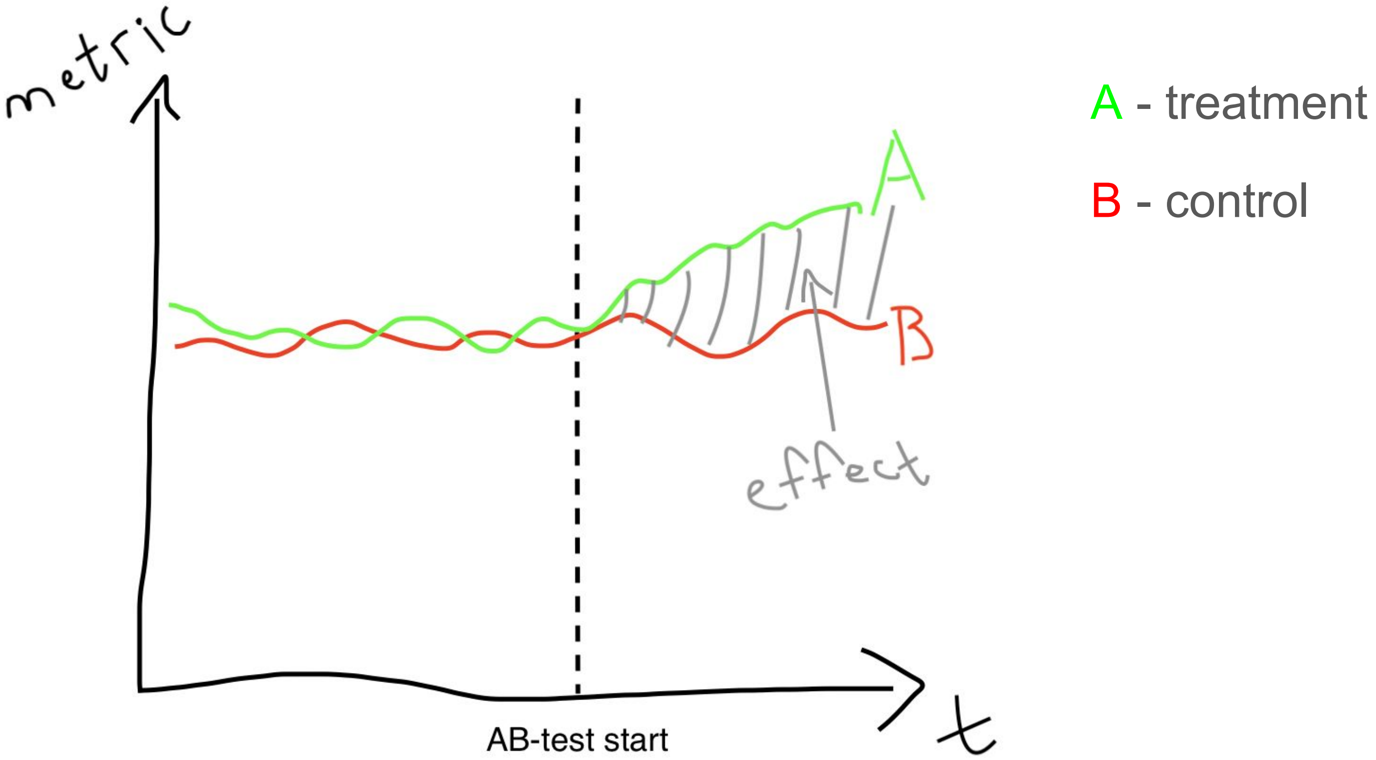
CUPED for faster AB-testing

Drambian David

What is an AB-test?

Control	Treatment
Register	Register (1 month free trial)

What happens in an ideal world?



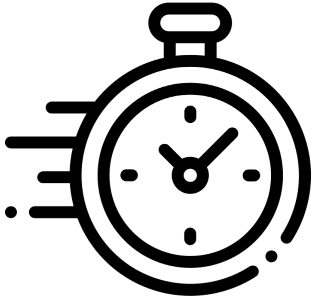
In other words, …

1. Run a test
2. Accumulate enough data
3. Calculate *the effect*
4. Evaluate if it’s significant

In other words, …

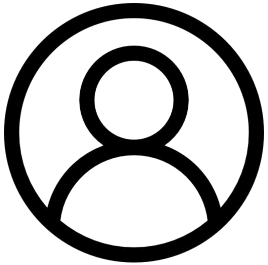
1. Run a test
2. Accumulate enough data ⚡ ⚡ ⚡
3. Calculate *the effect*
4. Evaluate if it’s significant

Why decrease sample size?



Short on time

Short on money



Not enough users

Minimum sample size

Confidence Interval for $(\mu_{treatment} - \mu_{control})$:

$$(\bar{x}_{treatment} - \bar{x}_{control}) \pm z_{\alpha/2} * \frac{\sigma}{\sqrt{n}} \tag{1}$$

$$we\ call \quad z_{\alpha/2} * \frac{\sigma}{\sqrt{n}} = \Delta = \text{Margin of Error} \tag{2}$$

$$n = \frac{z_{\alpha/2}^2 * \sigma^2}{\Delta^2} \tag{3}$$

where:

- $\boldsymbol{z_{\alpha/2}}$ is the score for a chosen significance level $\boldsymbol{\alpha}$
- $\boldsymbol{\Delta}$ (Margin of Error) is the minimum difference between means that we want to catch
- $\boldsymbol{\sigma}$ is the pooled estimate of standard deviation

Minimum sample size

Confidence Interval for $(\mu_{treatment} - \mu_{control})$:

$$(\bar{x}_{treatment} - \bar{x}_{control}) \pm z_{\alpha/2} * \frac{\sigma}{\sqrt{n}} \tag{1}$$

$$we\ call \quad z_{\alpha/2} * \frac{\sigma}{\sqrt{n}} = \Delta = \text{Margin of Error} \tag{2}$$

$$n = \frac{z_{\alpha/2}^2 * \sigma^2}{\Delta^2} \tag{3}$$

where:

- $\boldsymbol{z_{\alpha/2}}$ is the score for a chosen significance level $\boldsymbol{\alpha}$
- $\boldsymbol{\Delta}$ (Margin of Error) is the minimum difference between means that we want to catch
- $\boldsymbol{\sigma}$ is the pooled estimate of standard deviation

How to decrease Sample Size?

$$n = \frac{z_{\alpha/2}^2 * \sigma^2}{\Delta^2}$$

Make the test 'weaker' ⇒

⬆️ Significance Level (α) ⇒

⬇️ Score's Abs. Value ($|z_{\alpha/2}|$)

⬆️ Margin of Error (Δ)

⬇️ Variance (σ^2)

How to decrease Sample Size?

$$n = \frac{z^2_{\alpha/2} * \sigma^2}{\Delta^2}$$

Make … st 'weak…

⬆️ Signific… el (α) $\Rightarrow$

⬇️ Score'… ($|z_{\alpha/2}|$)

⬆️ Margin of Error (Δ)

⬇️ Variance (σ^2)

How to decrease Sample Size?

$$n = \frac{z^2_{\alpha/2} * \sigma^2}{\Delta^2}$$

Make [illegible]st ‘wea[illegible]

⬆️ Signific[illegible]el (α) ⇒

⬇️ Score’[illegible] ($|z_{\alpha/2}|$)

⬆️ Ma[illegible]or E[illegible](Δ)

⬇️ Variance (σ^2)

How to decrease Sample Size?

$$n = \frac{z^2_{\alpha/2} * \sigma^2}{\Delta^2}$$

Make … st 'weak …

⬆️ Signific … el (α) $\Rightarrow$

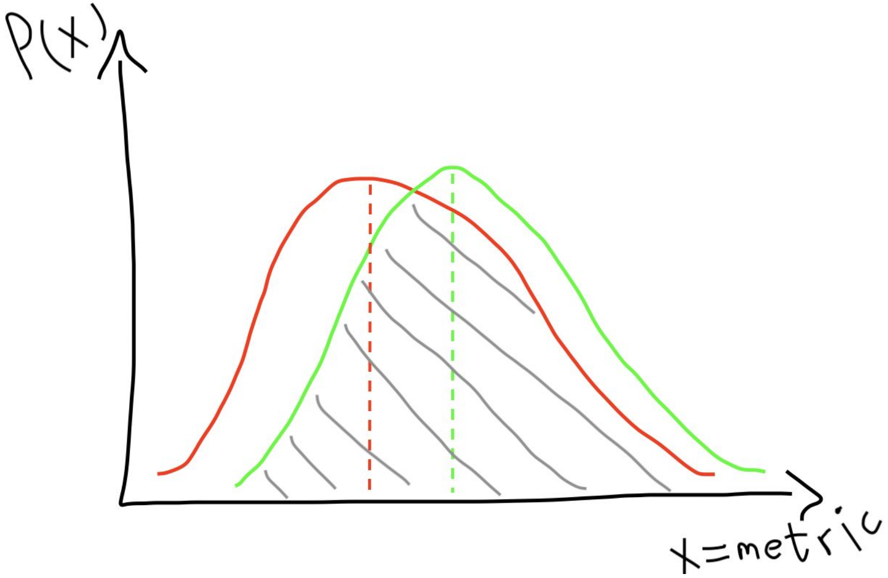
⬇️ Score' … ($|z_{\alpha/2}|$)

⬆️ Ma … of Er … (Δ)

⬇️ Variance (σ^2)

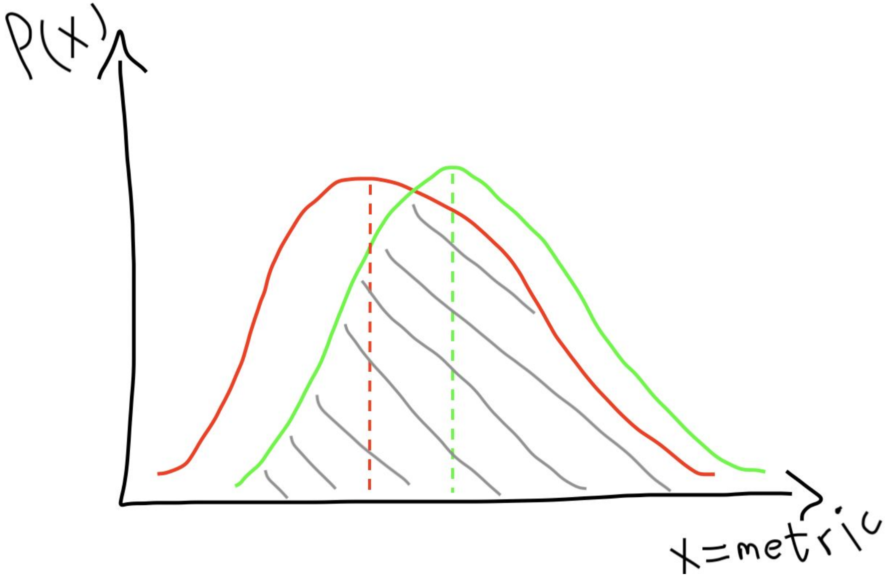
Intuition behind decreasing variance

High variance

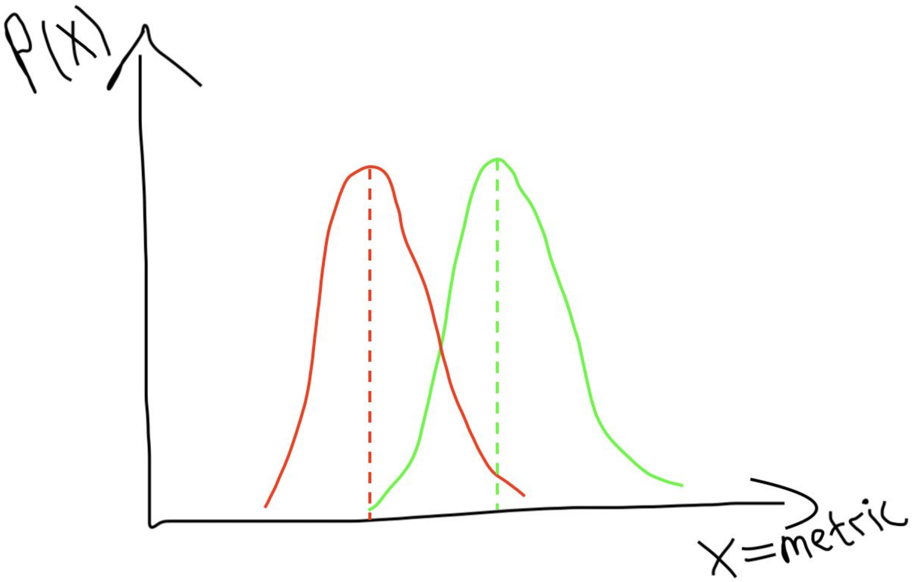


Intuition behind decreasing variance

High variance



Low Variance



Introducing ~~CUPID~~ **CUPED**

Controlled experiment **U**sing **P**re-**E**xisting **D**ata

1. Take a *covariate* *X* independent of the experiment
2. $Y_{CUPED} = Y - \theta X$

What happens to the *Variance*?

$$Y_{CUPED} = Y - \theta X \tag{4}$$

$$Var(Y_{CUPED}) = Var(Y - \theta X) = \underline{Var(Y) + \theta^2 Var(X) - 2\theta cov(Y, X)} \tag{5}$$

What happens to the *Variance*?

$$Y_{CUPED} = Y - \theta X \tag{4}$$

$$Var(Y_{CUPED}) = Var(Y - \theta X) = \underline{Var(Y) + \theta^2 Var(X) - 2\theta cov(Y, X)} \tag{5}$$

Trick: choose *X* such that it is dependant on *Y*, but independent on the experiment

What happens to the *Variance*?

$$Y_{CUPED} = Y - \theta X \tag{4}$$

$$Var(Y_{CUPED}) = Var(Y - \theta X) = Var(Y) + \theta^2 Var(X) - 2\theta cov(Y, X) \tag{5}$$

$$\theta_{min} = \frac{cov(Y, X)}{Var(X)} \tag{6}$$

What happens to the *Variance*?

$$Y_{CUPED} = Y - \theta X \tag{4}$$

$$Var(Y_{CUPED}) = Var(Y - \theta X) = Var(Y) + \theta^2 Var(X) - 2\theta cov(Y, X) \tag{5}$$

$$\theta_{min} = \frac{cov(Y, X)}{Var(X)} \tag{6}$$

$$Var(Y_{CUPED}) = Var(Y) + \frac{cov(Y, X)^2}{Var(X)^2} Var(X) - 2\frac{cov(Y, X)}{Var(X)} cov(Y, X) \tag{7}$$

$$\underline{Var(Y_{CUPED})} = Var(Y) + \frac{cov(Y, X)^2}{Var(X)} - 2\frac{cov(Y, X)^2}{Var(X)} = \underline{Var(Y)(1 - \rho(Y, X)^2)} \tag{8}$$

Intuition behind CUPED

The *covariate* **absorbs** part of the *Variance of Y*, without affecting the experiment

* Making Y_{CUPED} unbiased

$$\text{Unbiased, by definition: } \mathrm{E}(Y_{CUPED}) = \mu_Y \tag{9}$$

* Making Y_{CUPED} unbiased

$$\text{Unbiased, by definition: } \mathrm{E}(Y_{CUPED}) = \mu_Y \tag{9}$$

$$\text{But, } \mathrm{E}(Y_{CUPED}) = \mathrm{E}(Y - \theta X) = \mu_Y - \theta\,\mathrm{E}(X) \neq \mu_Y \tag{10}$$

* Making Y_{CUPED} unbiased

$$\text{Unbiased, by definition: } \mathrm{E}(Y_{CUPED}) = \mu_Y \tag{9}$$

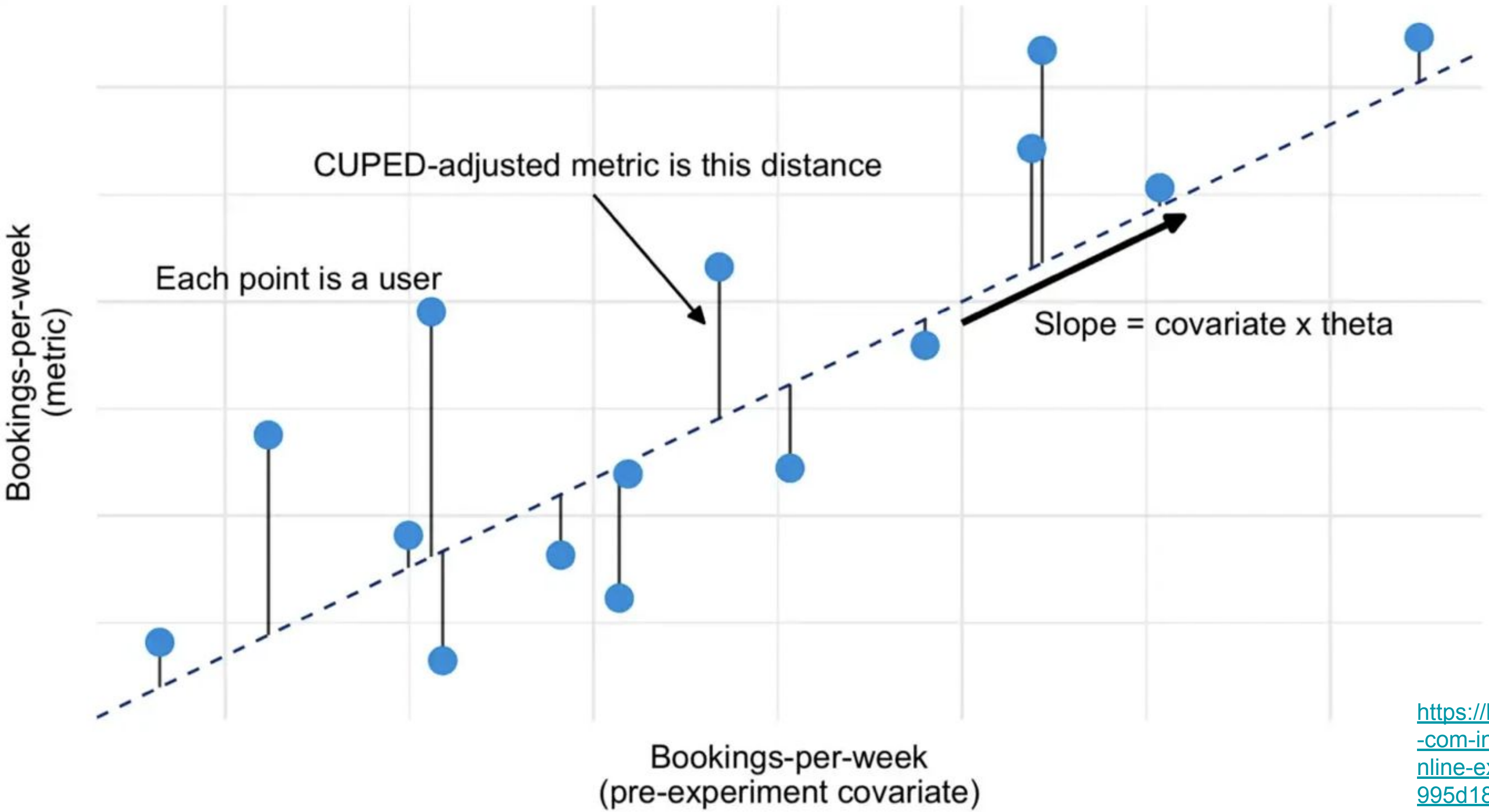
$$\text{But, } \mathrm{E}(Y_{CUPED}) = \mathrm{E}(Y - \theta X) = \mu_Y - \theta\,\mathrm{E}(X) \neq \mu_Y => \tag{10}$$

$$=> Y_{CUPED\ adjusted} = Y - \theta X + \theta\,\mathrm{E}(X) \tag{11}$$

How to choose a covariate?

- Absorbs the most variance of our metric (Y)
- Is independent from the experiment

How to choose a covariate?



https://booking.ai/how-booking-com-increases-the-power-of-online-experiments-with-cuped-995d186fff1d

Let’s practise

https://colab.research.google.com/drive/1M6B2jWZkRFe-g8_MYIAng3rKwq4sSwBa?usp=sharing

What we've learned?

- Decrease test sample size to save ⏱️, 💰 and 👤
- Use CUPED with covariate
  - Independent on experiment
  - Correlated with metric
- Absorb metric variance with covariate
- Have fun!

Be cool like him

dav1ddramb

<confidential-email>@gmail.com

@halftimestat

Useful links

https://www.researchgate.net/publication/237838291_Improving_the_Sensitivity_of_Online_Controlled_Experiments_by_Utilizing_Pre-Experiment_Data

https://www.kdd.org/kdd2016/papers/files/adp0945-xieA.pdf

https://towardsdatascience.com/how-to-double-a-b-testing-speed-with-cuped-f80460825a90

https://booking.ai/how-booking-com-increases-the-power-of-online-experiments-with-cuped-995d186fff1d

Other solutions

Stratification

LaTeX code

```
\documentclass{article}
\usepackage[utf8]{inputenc}
\usepackage{amsmath}
\usepackage[top=30pt,bottom=30pt,left=48pt,right=46pt]{geometry}
\usepackage{xcolor}
\DeclareMathOperator{\EX}{\mathbb{E}}
\title{CUPED for faster AB-testing}
\author{DD}
\date{December 2022}
\begin{document}
\maketitle
\section{Introduction}
\begin{Large}
\pmb{Confidence Interval} for $(\mu_{treatment} - \mu_{control})$ :
\begin{gather}
  (\Bar{x}_{treatment} - \Bar{x}_{control}) \pm z_{\alpha/2}*\frac{\sigma}{\sqrt{n}} \\
  we\;call\;\;\; z_{\alpha/2}*\frac{\sigma}{\sqrt{n}} = \Delta = \text{Margin of Error} \\
  n = \frac{z_{\alpha/2}^2 * \sigma^2}{\Delta^2}
\end{gather}
where:
\begin{itemize}
 \item \pmb{$z_{\alpha/2}$} is the score for a chosen significance level \pmb{$\alpha$}
 \item \pmb{$\Delta$} (Margin of Error) is the minimum difference between means that we want to catch
 \item \pmb{$\sigma$} is the pooled estimate of standard deviation
\end{itemize}
\end{Large}
\setlength{\jot}{15pt}
\section{CUPED}
\begin{Large}
\begin{gather}
  Y_{CUPED} = Y - \theta X \\
  Var(Y_{CUPED}) = Var(Y - \theta X) = Var(Y) + \theta^2 Var( X) - 2\theta cov(Y, X) \\
  \theta_{min} = \frac{cov(Y, X)}{Var(X)} \\
  Var(Y_{CUPED}) = Var(Y) + \frac{cov(Y, X)^2}{Var(X)^2} Var( X) - 2\frac{cov(Y, X)}{Var(X)} cov(Y, X) \\
  Var(Y_{CUPED}) = Var(Y) + \frac{cov(Y, X)^2}{Var(X)} - 2\frac{cov(Y, X)^2}{Var(X)} = Var(Y)(1-\rho(Y, X)^2)
\end{gather}
\end{Large}
\setlength{\jot}{20pt}
\section{Unbiased CUPED}
\begin{LARGE}
If we want $Y_{CUPED}$ to be unbiased:
\begin{gather}
  \text{Unbiased, by definition: } \EX(Y_{CUPED}) = \mu_{Y} \\
  \text{But, } \EX(Y_{CUPED}) = \EX(Y - \theta X) = \mu_Y - \theta \EX(X) \color{red} \neq \color{black}\mu_Y => \\
  => Y_{CUPED\;adjusted} = Y - \theta X + \theta \EX(X)
\end{gather}
\end{LARGE}

\end{document}
```