

# What's There to Know about A/B Testing?

Strata + Hadoop World, London  
6th May 2015

Noel Welsh, @noelwelsh



Everything is Change











# The Triangle of Trouble



People



Economics



Stats

People



People are *weird*

A small experiment

S \_ \_ P

S O A P



Decision making is  
malleable

Humans are risk  
adverse

# Economics

Business  $\neq$  Science



£££ ≠ knowledge

# Statistics

Null hypothesis  
significance test

p-value vs power  
Type I vs Type II error



# Bandit algorithms

Find the best choice  
(best arm selection)

Make the most money  
(regret minimisation)

How long have we got?  
(Time horizon)

Bayesian or  
Frequentist

# Issues

# Data Driven Procrastination

Test everything!



Tests are costly

Views

Development time

Cycle time

Statistical significance is  
not economic significance

What's the prior  
probability of improving  
KPIs?

How long will these  
changes live?

# The Uncanny Valley of False Precision



Know just enough  
theory to be dangerous

# Multiple comparisons problem

# Underpowered tests

Absence of evidence is  
not evidence of absence

Confirmation bias

Post-hoc reasoning

# Risk aversion

# A Useful Approach



Test everything ...

that is likely bring an  
economically meaningful  
effect ...

where the effect is small  
enough that it will be  
difficult to observe ...

and the test is cheap  
enough to run.

Use methods that allow  
for early stopping

or regret minimisation  
when horizon is small

I find Bayesian methods  
easier to analyse

Accept that a lot of  
your tests will fail



Thanks!  
Any questions?

# Photo Credits

All Photos CC licensed on Flickr

Foggy sunrise (Explore 2014-03-07)

Susanne Nilsson

Clear sky - no contrails

Björn Buhay

Grungy Dirty Dark Vintage Viewfinder Film Camera Lens

Glass Texture

D. Sharon Pruitt