

L5.1

2020年4月29日 星期三 下午4:52

2. Significance testing (显著性检验)

Checking significant materials - think about what metrics - invariant

	source count	# signif. increases	# certif. decreases	events	Time to complete
① Change order of course in course list	✓	✓	✓	hot topics	happened could be cause test affected e.g., easier classes before finish faster
course user (exp. count)	more	should be more early	more clear course flow	visit fewer pages	
eventually signif. user is more interested	more				
exist multiple courses exports	more				
others →	more				
② Change infrastructure to reduce load time while at diversion - current	✓	✓	✓	random	happened can't be cause test affected
multiple sources	more				
multiple metrics	more				
others →	more				
③ Change location of signature button to appear on course page	✓	✓	✓	your index taskbar	signature button to exp. random
→ open users to sign-in more	more				
use of different course	more				
consistent experience across sign-in sign-out boundary	more				
which metrics would make good improvements?	more				
✓ # events (sign-ups & logins all good)	more				
X OR user storage (only one metric)	less				
X add sign-in button to home page could affect this	less				
X probability of clicking this after course after sign-in	less				
X sign-in rate (this is what we're trying to change)	less				
X visit & sign-in complete sign-in process	less				
✓ video and other no backend changes	less				
user & control randomize - X after - X	less				

Checking materials

close (exp. count)

exp.

④ Run exp. for 2 weeks (iteration)

use of diversion - current

check: $\text{# of exp.} = \text{cont.}$

Weeks:

Day	# controls	# experiments	Weeks:		tot. no. of each group - each
			# exp.	exp.	
Mon	51	129	703	1307	
Tues	51	129	703	1307	1,047
Wed	51	129	703	1307	1,056
Thur	51	129	703	1307	1,046
Fri	51	129	703	1307	1,047
Sat	51	129	703	1307	1,047
Sun	51	129	703	1307	1,047
Total	51	129	703	1307	1,047

Total Cost = 16,434 → within expect? $\frac{1}{2}$ $\frac{1}{2}$

Total exp. = 1,818

How would you figure out whether this difference is within expectation?

Given: Each cookie randomly assigned to the control or experiment group with probability of 0.5

Like flipping a fair coin = heads = control

⑤ Compute standard deviation of binomial with probability of 0.5 of success

$P=0.5$ $N=61,434$ $b=61,434$ $\approx 12,287$ large Binomial \sim Normal

$SD = \sqrt{P(1-P)} = \sqrt{0.5(0.5)} = 0.25$

2. multiply by 2 - 3σ to get margin of error

$C = 3\sigma = 1.5$

$M = 2(C\sigma) = 0.5(1.5) = 0.75$

3. Compute CI around 0.5 (true STP)

$0.5 \pm M = 0.5 \pm 0.75 = (0.475, 0.525)$

4. Check whether observed fraction \gg margin interval

$0.51 \otimes 0.475 = 0.4975 < 0.525$ \Rightarrow set up incorrect

$P = \frac{61,434}{12,287} = 0.61$

Know what's going wrong - look day by day down a particular day - stand out \rightarrow cause problems

Overall pattern - what's the day?

1. Talk to the engineers - exp. set.

2. Try silencing to see if one particular slice is weird

3. Check age & cookies - does 1 group have more new cookies

Total events control = 15,348 Total events experiment = 15,312

Unit of diversion - 2 weeks

Unit of observation - 2 weeks

⑥ Run exp. for 7 days: Unit of diversion - 2 weeks

⑦ Run exp. for 7 days: Unit of diversion - 2 weeks

⑧ Run exp. for 7 days: Unit of diversion - 2 weeks

⑨ Run exp. for 7 days: Unit of diversion - 2 weeks

⑩ Run exp. for 7 days: Unit of diversion - 2 weeks

⑪ Run exp. for 7 days: Unit of diversion - 2 weeks

⑫ Run exp. for 7 days: Unit of diversion - 2 weeks

⑬ Run exp. for 7 days: Unit of diversion - 2 weeks

⑭ Run exp. for 7 days: Unit of diversion - 2 weeks

⑮ Run exp. for 7 days: Unit of diversion - 2 weeks

⑯ Run exp. for 7 days: Unit of diversion - 2 weeks

⑰ Run exp. for 7 days: Unit of diversion - 2 weeks

⑱ Run exp. for 7 days: Unit of diversion - 2 weeks

⑲ Run exp. for 7 days: Unit of diversion - 2 weeks

⑳ Run exp. for 7 days: Unit of diversion - 2 weeks

㉑ Run exp. for 7 days: Unit of diversion - 2 weeks

㉒ Run exp. for 7 days: Unit of diversion - 2 weeks

㉓ Run exp. for 7 days: Unit of diversion - 2 weeks

㉔ Run exp. for 7 days: Unit of diversion - 2 weeks

㉕ Run exp. for 7 days: Unit of diversion - 2 weeks

㉖ Run exp. for 7 days: Unit of diversion - 2 weeks

㉗ Run exp. for 7 days: Unit of diversion - 2 weeks

㉘ Run exp. for 7 days: Unit of diversion - 2 weeks

㉙ Run exp. for 7 days: Unit of diversion - 2 weeks

㉚ Run exp. for 7 days: Unit of diversion - 2 weeks

㉛ Run exp. for 7 days: Unit of diversion - 2 weeks

㉜ Run exp. for 7 days: Unit of diversion - 2 weeks

㉝ Run exp. for 7 days: Unit of diversion - 2 weeks

㉞ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days: Unit of diversion - 2 weeks

㉟ Run exp. for 7 days