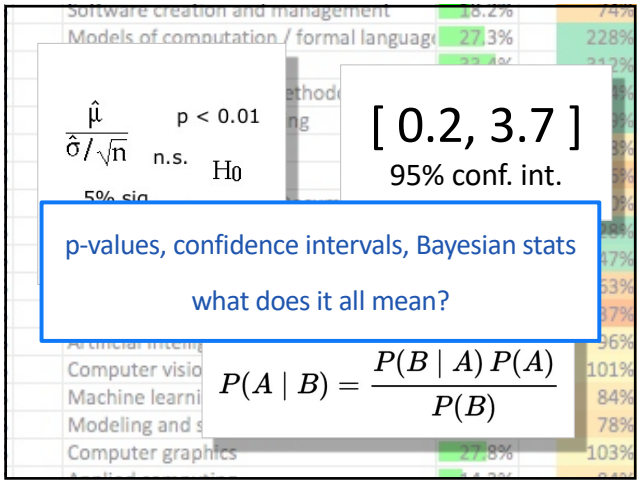
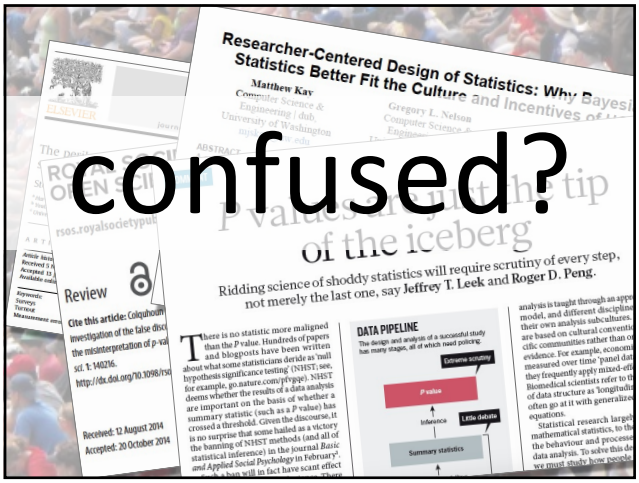
p < 0.01
n.s.
H0
5% sig.
CHI22
New Orleans, LA
Statistics for HCI
Alan Dix
https://alandix.com/statistics/chi2022/
Statistics for HCI – CHI 2022 – Alan Dix

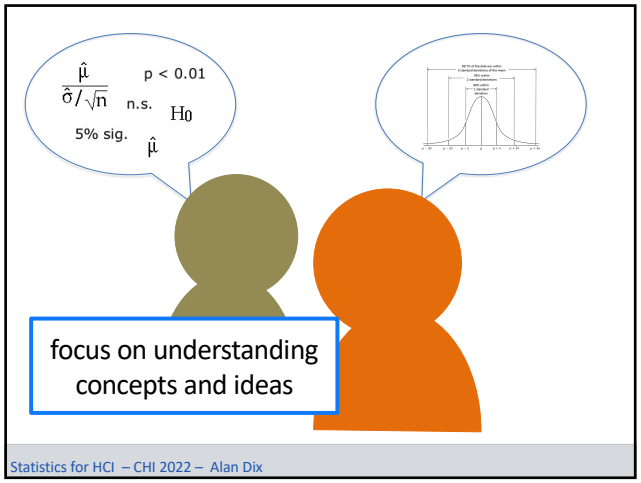
1


p < 0.01
n.s.
H0
[0.2, 3.7]
95% conf. int.
p-values, confidence intervals, Bayesian stats
what does it all mean?
$P(A \mid B) = \frac{P(B \mid A)\,P(A)}{P(B)}$

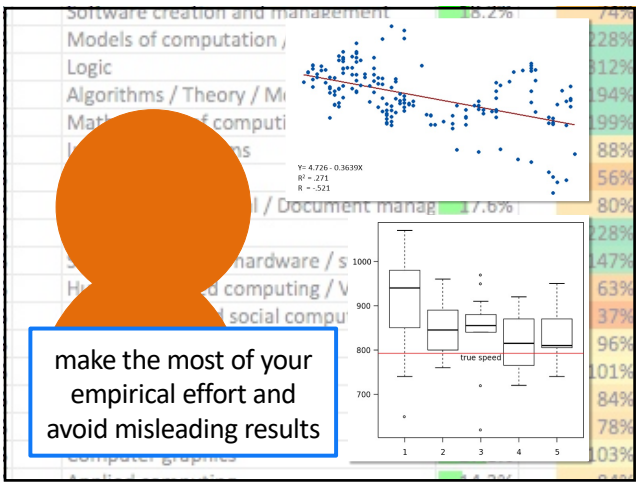
2


confused?
Researcher-Centered Design of Statistics: Why Bayesian Statistics Better Fit the Culture and Incentives of HCI
Matthew Kay
Gregory L. Nelson
P values are just the tip of the iceberg
Review
Ridding science of shoddy statistics will require scrutiny of every step, not merely the last one, say Jeffrey T. Leek and Roger D. Peng.
DATA PIPELINE

3


p < 0.01
n.s.
H0
5% sig.
focus on understanding concepts and ideas
Statistics for HCI – CHI 2022 – Alan Dix

4



make the most of your empirical effort and avoid misleading results

5

overview – three parts

wild and wide

- **understanding randomness**, bias and variability
- **distributions and long tails**

doing it

- alternative statistical analyses: the ubiquitous 'p' to Bayesian
- **common issues and critical differences**

design and interpretation

- **gaining power** – avoid the dreaded 'too few participants'
- making sense of your data and avoiding the pitfalls

Statistics for HCI – CHI 2022 – Alan Dix

6

overview – three parts

wild and wide

- **understanding ra**
- **distrib**

doing

- altern … sian
- **comm**

design a

- **gaining p** … eaded 'too few participants'
- making se … your data and avoiding the pitfalls

… but more important to make sense then 'get through' the material

videos, book, notes

Statistics for HCI – CHI 2022 – Alan Dix

7

the book …

Statistics for HCI: Making Sense of Quantitative Data

Morgan and Claypool

https://alandix.com/statistics/book/

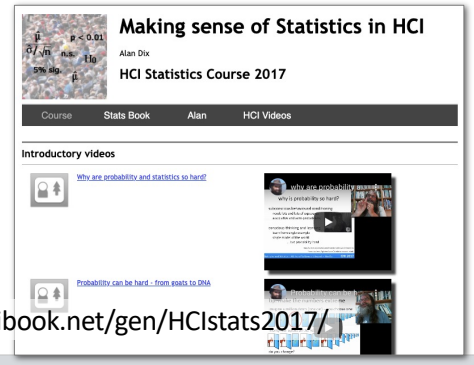
Statistics for HCI – CHI 2022 – Alan Dix

8

the website

https://alandix.com/statistics/

the videos



https://hcibook.net/gen/HCIstats2017/

overview – three parts

wild and wide

- **understanding randomness**, bias and
- **distributions and long tail**

doing it

- altern… …bayesian
- **comm…**

design a… …ation

- **gaining power** – avoid the dreaded 'too few participants'

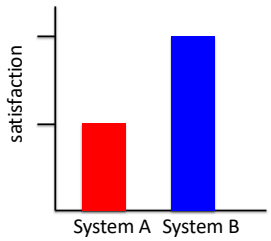
making sense of your data and avoiding the pitfalls

… but first …
who needs
statistics anyway?

do I need statistics?

just eyeball the data …

is system B better?



how many participants?

- 3000 – sure thing
- 3 – maybe just chance
- 6, 15, 30, 90, 300, …?

you do stats all the time!

but how can you know that you are right?

13

14

why are you doing it?

exploration vs. validation
process vs. product

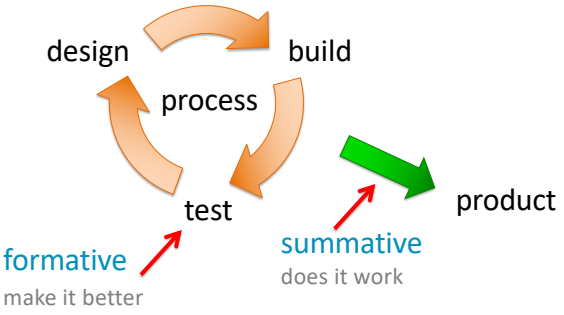
15

research

exploration	validation	explanation
?	✓	
finding questions	answering them	finding why and how
ethnography in-depth interviews detailed observation big data	experiments large-scale survey quantitative data	qualitative data theoretical models mechanism

16

development



17

exploration / formative

- find *any* interesting issues
- stats about deciding priorities

validation / summative

- exhaustive: find all problems/issues
- verifying: is hypothesis true, does system work
- mensuration: how good, how prevalent

explanation

- matching qualitative/quantitative, small/large samples

18

are five users enough?

original work

Nielsen & Landauer (1993) about *iterative* process

not summative – not for stats!

N.B. later work on saturation

how many?

to find enough to do in next development cycle

depends on size of project and complexity

now-a-days with cheap development maybe n=1

but always more in next cycle

19

20

why are statistics so hard?

holding on to opposities

21

Why are statistics so hard?

in two minds
- – conscious vs. sub-conscious

mathematics and materiality

accepting uncertainty
- – probability and the unknown

counterfactual reasoning

22

Why are statistics so hard?

in two minds
- – conscious vs. sub-conscious

23

Why are statistics so hard?

in two minds – conscious vs unconscious

mathematics and materiality
- – a foot in the world

3 apples + 2 oranges = ?

24

Why are statistics so hard?

in two minds – conscious vs unconscious

mathematics and materiality – a foot in the world

accepting uncertainty

– probability and the unknown

– probability helps to tame the unknown

… but it never goes away!

Diacritica, CC BY-SA 3.0. https://commons.wikimedia.org/wiki/File:6sided_dice.jpg

25

Why are statistics so hard?

in two minds – conscious vs unconscious

mathematics and materiality – a foot in the world

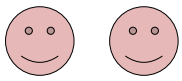
accepting uncertainty – probability vs unknown

counterfactual reasoning

– from observation to cause

– lots of "what if?"

26



27