

$\hat{\mu}$ $p < 0.01$
 $\hat{\sigma}/\sqrt{n}$ n.s. H_0

5% sig.

Understanding Statistics for Human-Computer Interaction and Related Disciplines

Alan Dix

<http://alandix.com/statistics/>

$$P(A | B) = \frac{P(A \cap B)}{P(B)}$$

Understanding Statistics for HCI and Related Disciplines – Alan Dix

$\hat{\mu}$ $p < 0.01$
 $\hat{\sigma}/\sqrt{n}$ n.s. H_0

5% sig.

[0.2, 3.7]
95% conf. int.

p-values, confidence intervals, Bayesian stats
what does it all mean?

$$P(A | B) = \frac{P(B | A) P(A)}{P(B)}$$

Understanding Statistics for HCI and Related Disciplines – Alan Dix

confused?

Researcher-Centered Design of Statistics: Why Bayesian Statistics Better Fit the Culture and Incentives
 Matthew Kay
 Gregory L. Nibwa

Tip of the iceberg
 P-values are the tip of the iceberg
 Ridging 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
 The design and analysis of a scientific study has many steps, all of which need paying attention.

These are the statistics more maligned than the P value. Hundreds of papers and blogposts have been written and doctored to make significance tests as 'null' as possible. NIST uses hypothesis significance testing (NHST) to determine whether the results of a data analysis are important on the basis of whether a summary statistic (such as a P value) has crossed a threshold. Given the disclosure, it is no surprise that some hailed as a victory in the journal *Basic and Applied Social Psychology* in 'Discovery' and 'Applied Social Psychology' in 'Discovery' and 'Applied Social Psychology' in 'Discovery'.

Cite this article: Colebatch, R. (2014) The investigation of the false discovery rate: a review. *Journal of the Royal Society Open Science*, 11, 140204. <http://dx.doi.org/10.1098/rsos.140204>
 Received: 12 August 2014
 Accepted: 20 October 2014

Understanding Statistics for HCI and Related Disciplines – Alan Dix

$\hat{\mu}$ $p < 0.01$
 $\hat{\sigma}/\sqrt{n}$ n.s. H_0

5% sig.

focus on understanding concepts and ideas

Understanding Statistics for HCI and Related Disciplines – Alan Dix

Software creation and management 18.2%
 Models of computation / formal language 27.3%
 Logic 12.2%
 Algorithms / Theory / Methods 19.4%
 Mathematics of computation 9.9%
 Hardware / software 17.6%
 Human-computer interaction / user interface 17.6%
 Hardware / software 17.6%
 Human-computer interaction / user interface 17.6%
 Hardware / software 17.6%
 Human-computer interaction / user interface 17.6%

$r = 0.36398$
 $R^2 = 0.1325$
 $R = -0.36398$

make the most of your empirical effort and avoid misleading results

Understanding Statistics for HCI and Related Disciplines – Alan Dix

overview – four parts

- wild and wide
 - exploring randomness, uncertainty and 'distributions'
- doing it
 - alternative statistical analyses: the ubiquitous 'p' to Bayesian
- gaining power
 - avoid the dreaded 'too few participants'
- so what?
 - making sense of your data and avoiding the pitfalls

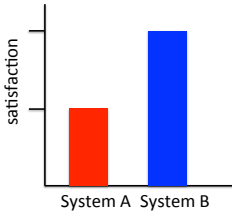
Understanding Statistics for HCI and Related Disciplines – Alan Dix

do I need statistics?

just eyeball the data ...

Understanding Statistics for HCI and Related Disciplines – Alan Dix


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?

Understanding Statistics for HCI and Related Disciplines – Alan Dix






Understanding Statistics for HCI and Related Disciplines – Alan Dix

why are you doing it?

exploration vs. validation
process vs. product

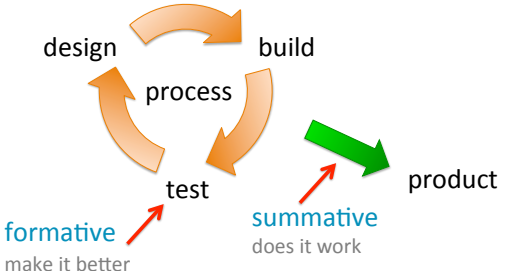
Understanding Statistics for HCI and Related Disciplines – Alan Dix

research

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

Understanding Statistics for HCI and Related Disciplines – Alan Dix

development



design process build test product

formative make it better

summative does it work

Understanding Statistics for HCI and Related Disciplines – Alan Dix

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

Understanding Statistics for HCI and Related Disciplines – Alan Dix

are five users enough?

original work

Nielsen & Landauer (1993) about *iterative* process

not summative – not for stats!

how many?

to find enough to do in next development cycle

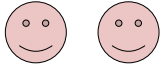
depends on size of project and complexity

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

but always more in next cycle

N.B. later work on saturation

Understanding Statistics for HCI and Related Disciplines – Alan Dix



Understanding Statistics for HCI and Related Disciplines – Alan Dix