Statistics for HCI

Alan Dix
https://alandix.com/statistics/chi2022/

What does it all mean?

Confused?

Focus on understanding concepts and ideas.
make the most of your empirical effort and avoid misleading results

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

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

the book ...
Statistics for HCI: Making Sense of Quantitative Data
Morgan and Claypool
overview – three parts

wild and wide
understanding randomness, bias and variability

… but first …
who needs statistics anyway?

bias and variability
distributions and long tails

doing it
alternative statistical analyses: the ubiquitous ‘p’ to Bayesian

design and interpretation

common issues and critical differences

gaining power – avoid the dreaded ‘too few participants’
making sense of your data and avoiding the pitfalls

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?

why are you doing it?
exploration vs. validation
process vs. product

research

exploration
finding
in-depth interviews
detailed observation
big data

validation
answering
experiments
large-scale survey
quantitative data

explanation
finding
why and how
qualitative data
theoretical models
mechanism
development

<table>
<thead>
<tr>
<th>development</th>
</tr>
</thead>
<tbody>
<tr>
<td>design</td>
</tr>
<tr>
<td>build</td>
</tr>
<tr>
<td>process</td>
</tr>
<tr>
<td>test</td>
</tr>
<tr>
<td>product</td>
</tr>
<tr>
<td>formative</td>
</tr>
<tr>
<td>make it better</td>
</tr>
<tr>
<td>summative</td>
</tr>
<tr>
<td>does it work</td>
</tr>
</tbody>
</table>

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

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
Why are statistics so hard?

- In two minds
  - Conscious vs. sub-conscious
- Mathematics and materiality
- Accepting uncertainty
  - Probability and the unknown
- Counterfactual reasoning