Experimental design

How would you evaluate fertilizer effect?

Discuss with partner (5’)

Experimental design principles

Replication

Replication!

Replication

- Replication is key: we need several samples.

Replication

- Replication is key: we need several samples.
- How many? The more the better! See Gelman & Carlin 2014.

Replication

- Replication is key: we need several samples.
- How many? The more the better! See Gelman & Carlin 2014.
- But there are trade-offs. See Lakens' sample size justification

The most important aspect of a statistical analysis is not what you do with the data, it's what data you use

H. Stern / A. Gelman

The importance of sample size

- Many studies have **too low sample sizes**.

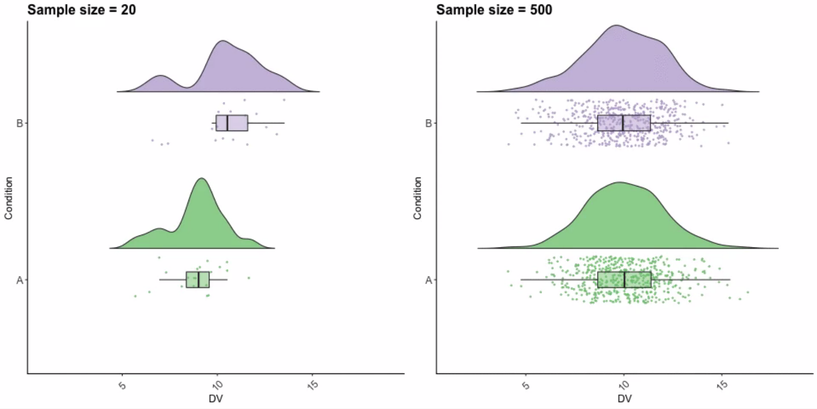
The importance of sample size

- Many studies have **too low sample sizes**.
- Low sample sizes **miss subtle effects**, but also **prone to bias**.

The importance of sample size

- Many studies have **too low sample sizes**.
- Low sample sizes **miss subtle effects**, but also **prone to bias**.
- The **fallacy** of assuming that which does not kill statistical significance makes it stronger (Loken & Gelman 2017).

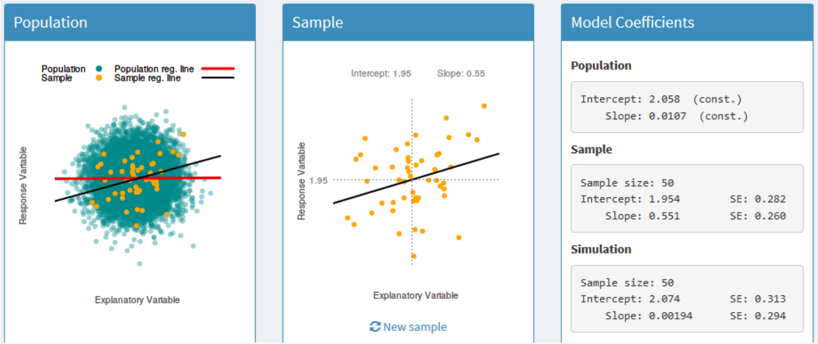
Low sample sizes are very sensitive to random noise



https://twitter.com/ajstewart_lang/status/1020038488278945797

Play yourself here

Low sample sizes may bias inferences about population



Source: statisticalgate.com

Low sample sizes may bias inferences

See *The evolution of correlations*

Stopping rules

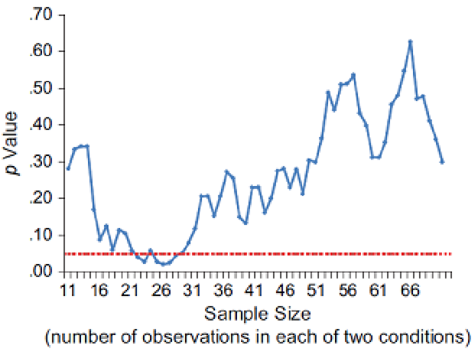


Fig. 2. Illustrative simulation of *p* values obtained by a researcher who continuously adds an observation to each of two conditions, conducting a *t* test after each addition. The dotted line highlights the conventional significance criterion of $p \leq .05$.

Sample size estimation

- **Plan** model/statistical analysis **before** data collection.

Sample size estimation

- **Plan** model/statistical analysis **before** data collection.
- **Do simulations**. Power/Sample size/Precision analyses (e.g. see papers like this & this, or software like this & this).

Sample size estimation

- **Plan** model/statistical analysis **before** data collection.
- **Do simulations**. Power/Sample size/Precision analyses (e.g. see papers like this & this, or software like this & this).
- **Justify** your sample size (precision, minimal effect size…). See Lakens' book.

Sample size estimation

- **Plan** model/statistical analysis **before** data collection.
- **Do simulations**. Power/Sample size/Precision analyses (e.g. see papers like this & this, or software like this & this).
- **Justify** your sample size (precision, minimal effect size...). See Lakens' book.
- Plan to have at least **10-30 observations per predictor**.

Sample size estimation

- **Plan** model/statistical analysis **before** data collection.
- **Do simulations**. Power/Sample size/Precision analyses (e.g. see papers like this & this, or software like this & this).
- **Justify** your sample size (precision, minimal effect size...). See Lakens' book.
- Plan to have at least **10-30 observations per predictor**.
- Complex models (w/ many predictors, interactions etc) require **high** sample sizes.

Sample size estimation

Calculating sample size for Gaussian (Normal) response model:

- expected mean: 30
- expected sd: 10
- 10 parameters (predictors)
- expected R-squared: 0.2

```
library(pmsampsize)
pmsampsize(type = "c", parameters = 10, intercept = 30, sd = 10, rsquared = 0.2)
```

```
NB: Assuming 0.05 acceptable difference in apparent & adjusted R-squared
NB: Assuming MMOE <= 1.1 in estimation of intercept & residual standard deviation
SPP - Subjects per Predictor Parameter

            Samp_size Shrinkage Parameter Rsq  SPP
Criteria 1        313     0.900        10 0.2 31.3
Criteria 2        161     0.827        10 0.2 16.1
Criteria 3        244     0.876        10 0.2 24.4
Criteria 4*       313     0.900        10 0.2 31.3
Final             313     0.900        10 0.2 31.3

 Minimum sample size required for new model development based on user inputs = 313

 * 95% CI for intercept = (29.01, 30.99), for sample size n = 313
```

Sample size estimation

Calculating sample size for binary response model:

- expected prevalence: 0.1
- 20 parameters (predictors)
- expected R-squared: 0.2

```
library(pmsampsize)
pmsampsize(type = "b", parameters = 20, prevalence = 0.1, nagrsquared = 0.2)
```

```
NB: Assuming 0.05 acceptable difference in apparent & adjusted R-squared
NB: Assuming 0.05 margin of error in estimation of intercept
NB: Events per Predictor Parameter (EPP) assumes prevalence = 0.1

           Samp_size Shrinkage Parameter CS_Rsq Max_Rsq Nag_Rsq  EPP
Criteria 1      1774     0.900        20  0.096   0.478   0.201 8.87
Criteria 2       786     0.801        20  0.096   0.478   0.201 3.93
Criteria 3       139     0.900        20  0.096   0.478   0.201 0.70
Final           1774     0.900        20  0.096   0.478   0.201 8.87

 Minimum sample size required for new model development based on user inputs = 1774,
 with 178 events (assuming an outcome prevalence = 0.1) and an EPP = 8.87
```

Randomization

Randomization

Randomization

- Haphazard $\neq$ Random

Randomization

- Haphazard $\neq$ Random
- Stratify: randomize within groups (e.g. species, soil types)

Controls

Have controls

- Untreated individuals, plots… (assigned randomly, of course).

Have controls

- Untreated individuals, plots… (assigned randomly, of course).
- Must differ only in treatment (i.e. homogeneous environment).

Have controls

- Untreated individuals, plots… (assigned randomly, of course).
- **Must differ only in treatment** (i.e. homogeneous environment).
- Measure **before & after** treatment.

Have controls

- Untreated individuals, plots... (assigned randomly, of course).
- **Must differ only in treatment** (i.e. homogeneous environment).
- Measure **before & after** treatment.
- Consider **blind designs** to avoid observer bias.

Experimental design principles

1. Replication

Experimental design principles

1. Replication
2. Randomization

Experimental design principles

1. Replication
2. Randomization
3. Controls

To read more

- Ruxton & Colegrave: Experimental Design for the Life Sciences. OUP

To read more

- Ruxton & Colegrave: Experimental Design for the Life Sciences. OUP
- Lakens: Improving your statistical inferences

To read more

- Ruxton & Colegrave: Experimental Design for the Life Sciences. OUP
- Lakens: Improving your statistical inferences
- DeclareDesign