Statistical Thinking in Biology Research

Understanding Statistical Inference through Simulation

Terry Neeman

Australian National University

12th August 2021

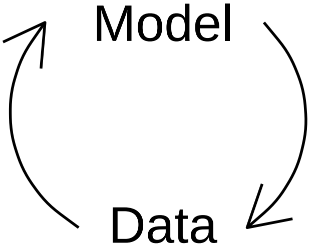
Key questions

- How do we decide whether a treatment effect is real?
- How do we decide if a "pattern" in our data is real or imagined?
- Given the experimental data, what can we *infer* about the effect of treatments?
- What counts as evidence?

Let's use simulation to explore these questions.

Simulation and Inference

Model $\rightarrow$ *Data* $\rightarrow$ *Model*



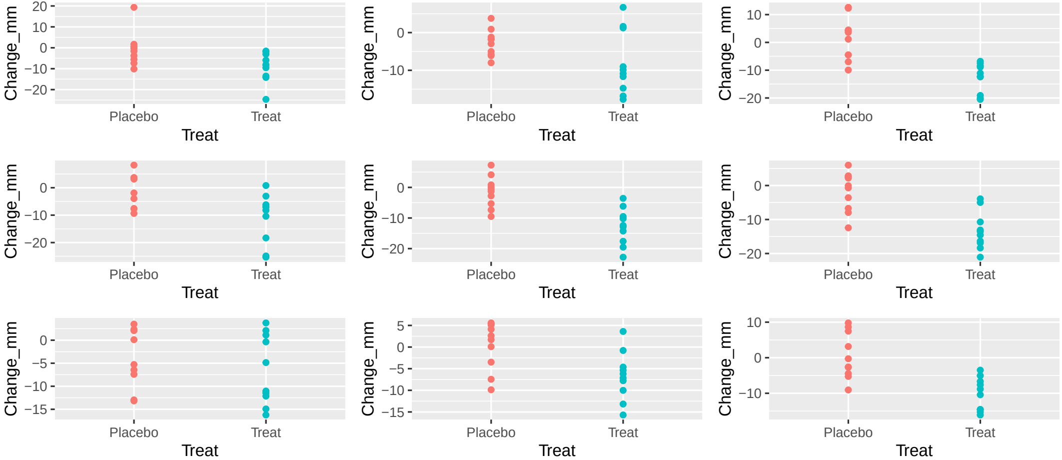
inference

simulation

Let's simulate some data!

- Made-up scenario: testing new blood pressure lowering medication
- 9 clinical centres
- 20 patients per centre randomised 1:1 to treatment or placebo
- Primary outcome: change in blood pressure (SBP) from baseline (mm)
- Model: variation between patients normally distributed

Simulated data: evidence of a treatment effect?



Model for simulation experiment

$$(\Delta SBP | treatment) = -10 + rnorm(mean = 0, sd = 6)$$

$$(\Delta SBP | placebo) = rnorm(mean = 0, sd = 6)$$

“Signal” = mean difference = 10

“Noise” = variation around mean = random “normal” variation

A few observations

- Model for each simulation exactly the same
- Some centres - more convincing evidence
- Variation - may interfere with seeing “signal”
- Can we combine data across sites?
- Will combining data add to the “signal” or “noise”?

Fit a model to the data

```
library(lmerTest)
model1 <- lmer(Change_mm~Treat + (1|Centre), data = sim2)
anova(model1)
```

```
## Type III Analysis of Variance Table with Satterthwaite's method
##       Sum Sq Mean Sq NumDF DenDF F value    Pr(>F)
## Treat 3688.4  3688.4     1   178  87.441 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Can the model recover the “truth”?

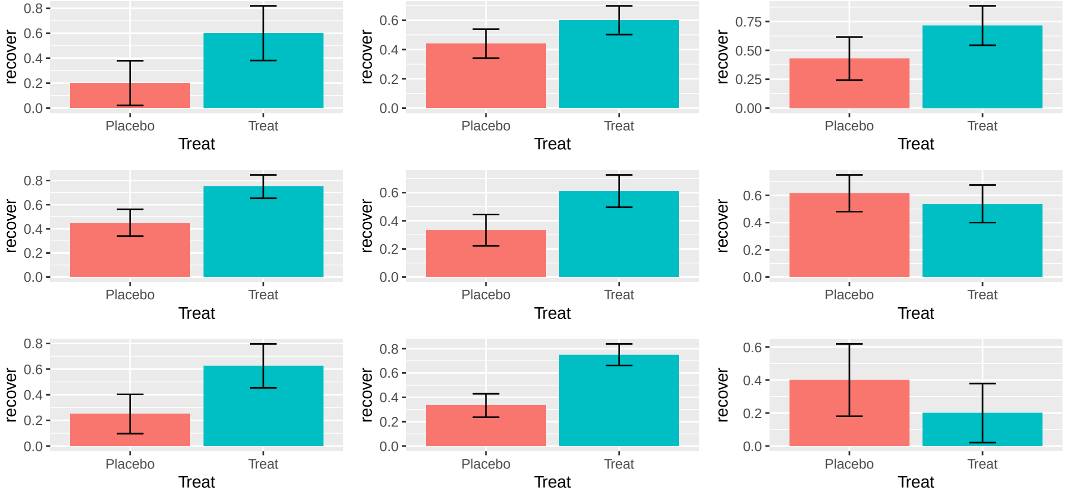
```
library(emmeans)
emmeans(model1, ~Treat)
```

```
##  Treat    emmean    SE   df lower.CL upper.CL
##  Placebo  -0.999 0.685 30.6     -2.4    0.398
##  Treat   -10.052 0.685 30.6    -11.4   -8.655
##
## Degrees-of-freedom method: kenward-roger
## Confidence level used: 0.95
```

Let's simulate some more data

- Testing a new treatment for COVID-19
- 9 centres
- Randomise between 10-50 patients to treatment or placebo
- Primary outcome: full recovery within 7 days
- Model: Binomial model, probability of recovery $= p$

Simulated data: what patterns do you see?



Model for simulation experiment - “biased coin” flipping

$$Prob(recovery|treatment) = 0.70$$

$$Prob(recovery|placebo) = 0.40$$

A few observations

- Model for each simulation same
- Number of patients per centre vary
- Role of chance - may interfere with signal
- How does number of patients affect “signal”?
- Will combining data add to the “signal” or “noise”?

Fit a model to the data

```
library(car)
model2 <- glmer(Recovered ~ Treat + (1|Centre), family = binomial,
                data = sim2)
Anova(model2)
```

```
## Analysis of Deviance Table (Type II Wald chisquare tests)
##
## Response: Recovered
##        Chisq Df Pr(>Chisq)
## Treat 14.136  1  0.0001701 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

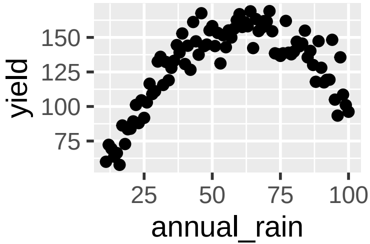
Can the model recover the "truth"?

```
library(emmeans)
emmeans(model2, ~Treat, type = "response")
```

```
##  Treat   prob     SE  df asymp.LCL asymp.UCL
##  Placebo 0.40 0.0438 Inf     0.318     0.488
##  Treat   0.64 0.0429 Inf     0.552     0.719
##
## Confidence level used: 0.95
## Intervals are back-transformed from the logit scale
```

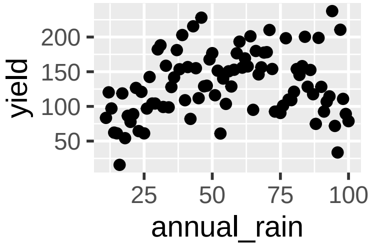
Re-visit precipitation vs yield data

```
set.seed(202073)
annual_rain<-seq(11,100, 1)
yield <- 2 + 5*annual_rain - 0.04* annual_rain^2 + rnorm(90,0,10)
yield_dat<-tibble(annual_rain = annual_rain, yield=yield)
ggplot(yield_dat, aes(annual_rain, yield))+geom_point()
```



What happens when we increase the “noise”?

```
set.seed(202073)
annual_rain<-seq(11,100, 1)
yield <- 2 + 5*annual_rain - 0.04* annual_rain^2 + rnorm(90,0,40)
yield_dat<-tibble(annual_rain = annual_rain, yield=yield)
ggplot(yield_dat, aes(annual_rain, yield))+geom_point()
```



Inference and Evidence

- Inference: deciding observed "signal" is REAL
- Fail to INFER signal $\neq$ "no signal"
- Evidence of signal: depends on signal:noise ratio
- Weak evidence: LOW signal:noise ratio
- STRONG evidence: HIGH signal:noise ratio
- INFER signal is "real" when there is STRONG evidence
- INFERENCE $\neq$ PROOF

Inference with Noisy Data

- more noise means harder to INFER signal is real
- More data = more information, higher signal:noise ratio
- Replication important for inference
- Combining experiments: combine information about signal

Summary - let's answer our key questions

- How do we decide whether a treatment effect is real?
  - **strong EVIDENCE that effect is real**
- How do we decide if a "pattern" in our data is real or imagined?
  - **Model data, model fit includes measures of evidence**
- Given the experimental data, what can we *infer* about the effect of treatments?
  - **INFERENCE = strong evidence of treatment effect**
- What counts as evidence?
  - **evidence measured by signal:noise ratio**