ADAPTIVE HEALTH INTELLIGENCE

EVIDENCE IN ACTION

Advice to Medical Students: A Statistical Sermon

Michael Dymock

29th May 2024

About me…

- 2014-2020:

  BSc (Hons) Mathematics & Statistics

  UWA Centre for Applied Statistics

- 2020 onwards:

  Biostatistician at Telethon Kids, et. al.

- 2023 onwards

  PhD @ UWA

  Council member of SSA WA Branch

Statistical Society of Australia

THE UNIVERSITY OF WESTERN AUSTRALIA

ADAPTIVE HEALTH INTELLIGENCE

EVIDENCE IN ACTION

A Statistical Sermon

STRIVE WA
Student Research Initiative

Overview

- Probability, statistics and models
- Embracing uncertainty
- Hypothesis testing
- Sample size
- The ten commandments
- Where to find more help

…to improve health outcomes, we need to improve clinical decision-making…

…which requires the efficient pursuit of truth (to useful questions)…

…we simplify the complex world with biological/clinical/statistical models…

…and estimate the parameters of these models using knowledge and data…

Probability vs Statistics

If only I knew the **parameters**, then I could predict the **observations**!

If only I knew the **observations**, then I could infer the **parameters**!

Probability vs Statistics

ADAPTIVE HEALTH INTELLIGENCE
EVIDENCE IN ACTION

More formally

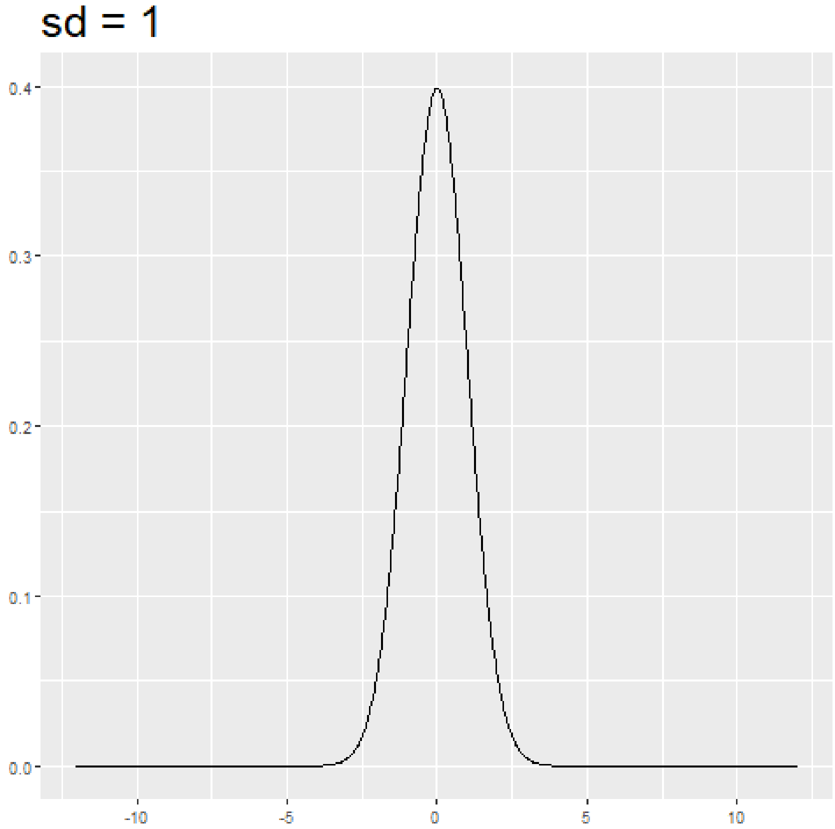
- We can use **probability distributions** to understand the behaviour of the world around us
  - E.g., a clinician can make an **informed** decision on prescribing a treatment if they understand its behaviour
    (e.g., **mean** and **variance**)
- We can use **statistical methods** to infer the probability distributions of interest
  - E.g., by collecting and analysing data, we can estimate
    parameters (e.g., **mean** and **variance**)

The normal (Gaussian) distribution

- A **probability distribution** (something that allocates probabilities over a set of possible outcomes)
- Has both a **mean** parameter and a **variance** parameter and is **symmetric**

$$N(\mu, \sigma^2)$$

mean — variance



Statistical models

- We
- Der
- E.g.
- If w
- Do
- Our
- We could use **linear regression** (or another statistical method)

“All models are wrong, but some are useful” – George Box

Many different models with many different parameters

Usually want to estimate a “parameter of interest”

But we need to understand the uncertainty in our estimation

What do we mean by uncertainty?

- In **frequentist** statistics, parameters have **unknown** but **fixed** values
- Because they are **unknown**, we cannot be sure how close our guess/estimate is to the true **fixed** value
- But we can estimate our **uncertainty** in the estimation itself
- We usually do this using **confidence intervals**
- E.g., our **point estimate** may be 4 but our 95% confidence interval may be (2,6), this describes our uncertainty in the point estimate

*Bayesians have a different take on this – speak to me later if you want to join the dark side

What is a hypothesis test?

- We **assess** the claim of a **hypothesis** against the evidence
- Specifically, we assess the evidence that a **model parameter** takes on a certain value or lies within a certain range
- E.g., one may **hypothesise** that $\mu > 0$ (i.e., that the mean response is positive)
- We can test this claim using the two **hypotheses:**
  - **Null hypothesis**: $H_0: \mu = 0$
  - **Alternative hypothesis**: $H_1: \mu > 0$

The philosophical argument…

- Proof by **contradiction**
  - Suggest Theory X
  - Find a contradiction (or counter example) to Theory X
  - Therefore, Theory X is false
- Scientific arguments or theories (rarely) can ever be proven
- Instead, we gather evidence to **support** or **counter** a theory
- With a hypothesis test, we aim to assess evidence that **counters** the claim of the null hypothesis, thus **supporting** the alternative hypothesis
- **BUT** the failure to find counter evidence **does not** prove the null hypothesis
- We do this with p-values!

P-values: holy grail or poisoned apple?

- A p-value is the *a priori* probability of observing the data (or more extreme) under the **assumption** that the null hypothesis is **true**
- A small p-value is therefore evidence that the data were unlikely to be observed **if** the **null hypothesis was true** (i.e., $\mu = 0$)
- This is the **counter evidence** against the null hypothesis, and so we **reject** it
- We need to *a priori* set a **threshold** or **significance level**
- How small does the p-value need to be to convince me that the null hypothesis is false
- Historically, and preferably in the eyes of grant review panels, this is set at the magical value of 5% (p-values under 5% are “good” otherwise we just try again or file it away and pretend it never happened)

Is this an issue?

- The significance level is also called **type one error**

This is called multiple testing

There are methods to deal with this issue

But it can creep into science when we are not transparent

We must report out insignificant results!

Why do we care about sample size?

Decreasing the sample size

Increasing the sample size

- Save resources!
- Ethics??

- Increase precision!
- Ethics??

From a (purely) statistical point of view…

- Large sample sizes are always preferable **with caution**
- At **study design**, we compute the required sample size to achieve the desirable **type one error** and **power**
    - Although, this is usually done backwards!
- **Power** is chance we **correctly** reject the null hypothesis
- **Before** seeing data, the sample size can help us understand the possible behaviour of the trial and guide our interpretation of the results
- **After** seeing data, the sample size no longer matters!

The ten (statistical) commandments for medical students

1) Pursue the truth with integrity and enthusiasm;
2) Respect tradition, the scientists and their methods, for they paved the path you walk;
3) Challenge dogma, sometimes they were wrong, and we can do better;
4) Be humble, you are probably also wrong, but the journey is worthwhile;
5) Embrace uncertainty: an uncertain answer to the right question is better than a certain answer to the wrong question;
6) Think carefully about what you are trying to estimate and why;
7) Beware of biases: there will always be a snake in the garden;
8) Sacrifice the project of your dreams for the supervisors you love;
9) Befriend the internet: a problem you face now was likely solved long ago;
10) Employ a statistician for they too have families to feed (and they may also be useful);

Where to find more help

Perth Children's Hospital (seminars):

CAHS Research Education Program Research Skills Seminars

Telethon Kids Institute (consultancy service):

biometrics@telethonkids.org.au

UWA Centre for Applied Statistics (short courses and consultancy service):

consulting-cas@uwa.edu.au

Joint clinical and statistical supervision (unlimited access to knowledge)

ADAPTIVE HEALTH INTELLIGENCE
EVIDENCE IN ACTION

Checklist for talking to a statistician

- Clear hypothesis
- Proposed study design
- Primary endpoint & estimate of variability
- Clinically relevant effect size
- Estimate of feasible sample size
- Important confounders and source of bias
- Similar publications or systematic reviews

How can I learn more about statistics?

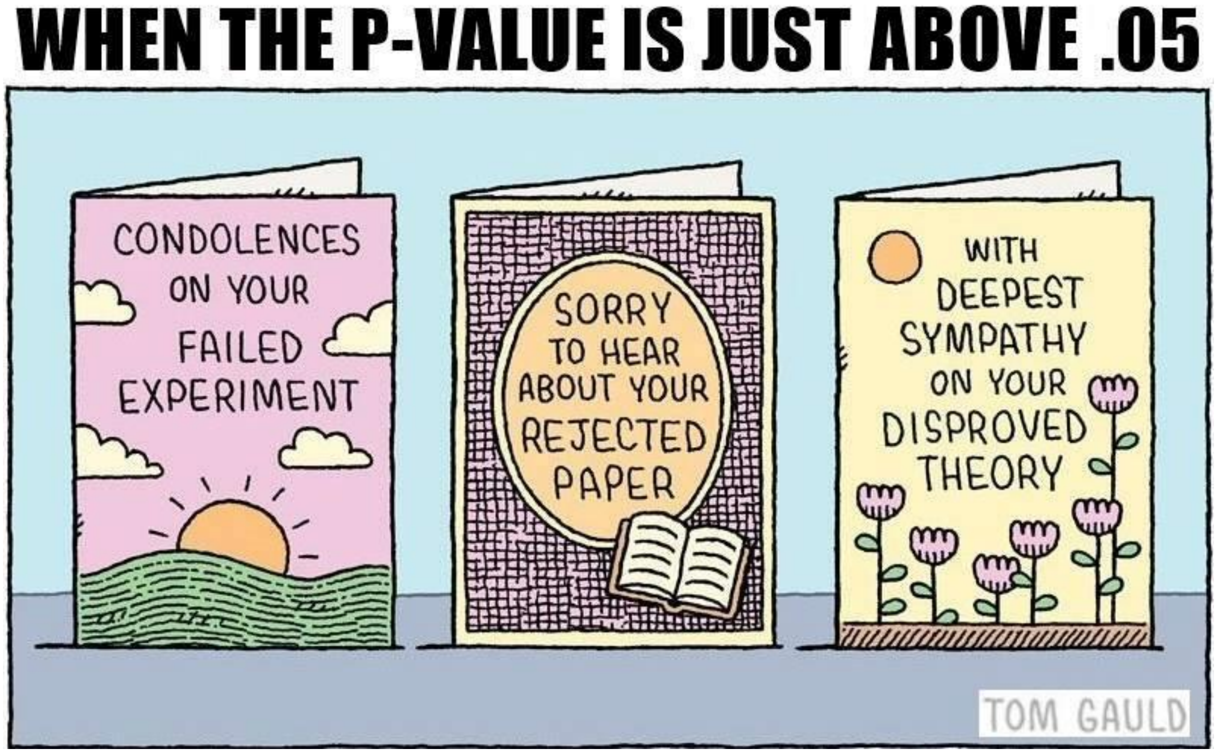
- In the absence of large, randomised, well-controlled clinical trials to address every research question we all need to increase our **statistical literacy**

- Explore online and in-person courses

- Ask questions

- Be brave!

Thank you


WHEN THE P-VALUE IS JUST ABOVE .05
CONDOLENCES ON YOUR FAILED EXPERIMENT
SORRY TO HEAR ABOUT YOUR REJECTED PAPER
WITH DEEPEST SYMPATHY ON YOUR DISPROVED THEORY
TOM GAULD