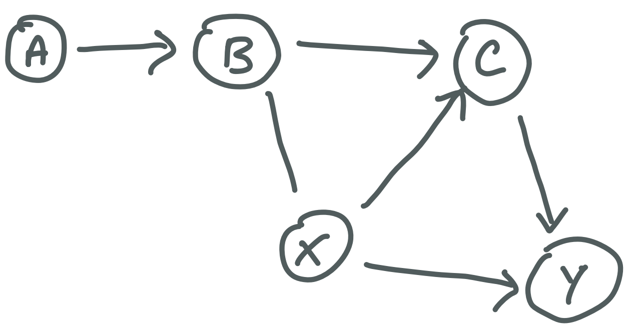
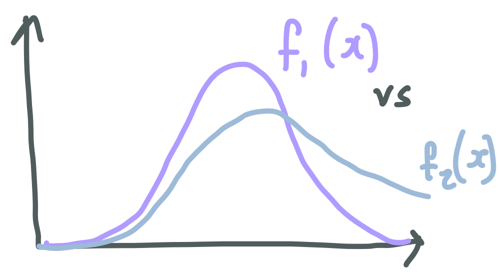
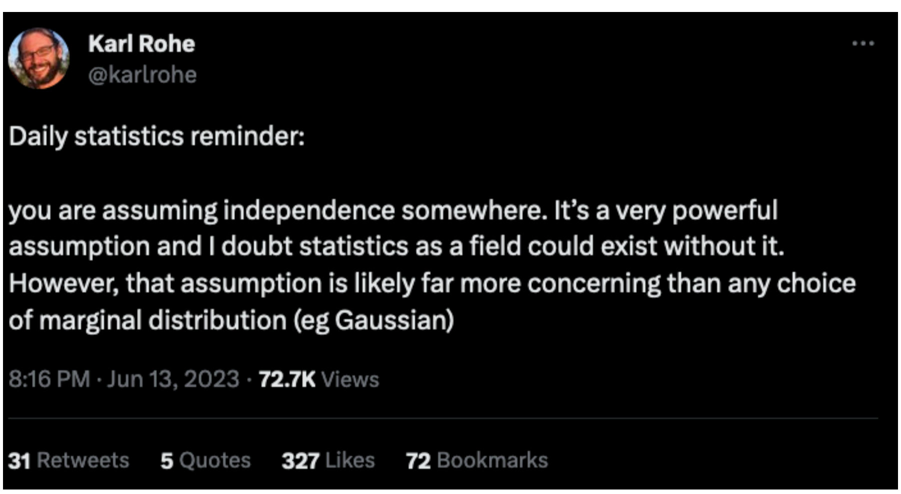
ETHICS II LIVE SESSION 4

SAMPLE SIZE CALCULATIONS FOR RCTs & A/B TESTING.

last week in tweets:







Karandeep Singh
@kdpsinghlab

Some textbooks say that meteors *caused* dinosaurs to go extinct.

This is wrong.

Meteors were a collider.

3:16 PM · Jun 19, 2023 · **35.9K** Views

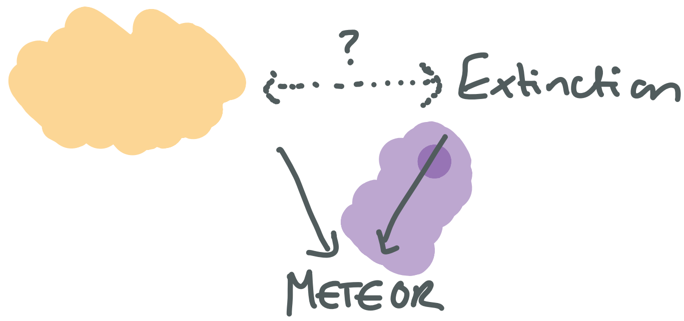
METEOR

↓

Climate Change

↓

Extinction



REMINDER

Assessment 1 is due this coming **Thursday 29th June** at **11:59pm (BST)** (or earlier!)

Dr Meming
@Dr_Meming

Me proudly standing by my poster at conferences waiting for somebody to come and talk to me

Review of hypothesis testing

- Within a hypothesis test we have a null & an alternative hypothesis.

- Construct a test statistic T and it's sampling dist$^{\underline{n}}$ under the null hypothesis

- Reject H_0 if t is very unlikely under H_0.

Review of Error types

- Within a hypothesis test we have a null & an alternative hypothesis.

- Construct a test statistic T and it's sampling dist$^{\underline{n}}$ under the null hypothesis

- Reject H_0 if t is very unlikely under H_0.

	fail to reject H_0	reject H_0
H_0 true	✓	type I error α
H_0 false	type II error β	✓

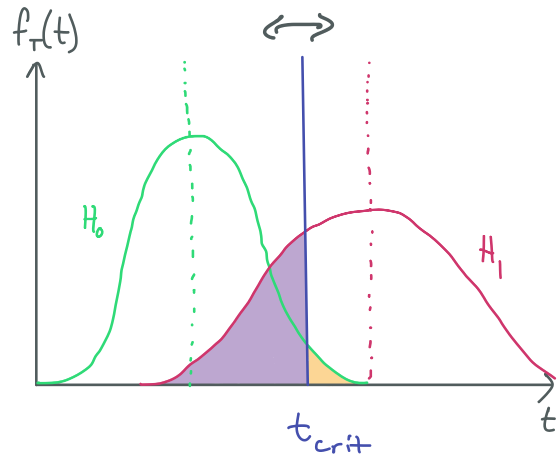
Review of Error types

- Within a hypothesis test we have a null & an alternative hypothesis.

- Construct a test statistic T and it's sampling dist^n under the null hypothesis

- Reject H_0 if t is very unlikely under H_0.

	fail to reject H_0	reject H_0
H_0 true	✓	type I error α
H_0 false	type II error β	✓

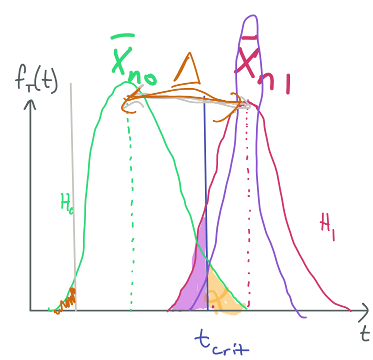


Power of a test

Power = Ability to detect a real difference.

$$= \Pr(\text{Reject } H_0 \mid H_0 \text{ false})$$

$$= 1 - \beta$$



Will depend on:

- Δ true effect size
- n sample size
- α size of the test
- T test statistic used
- F_X dist^n of data
- H_1 nature of alternative hypothesis

Sample Size Calculations

Aim: Pick smallest n such that for a given α, Δ, T we get at least $1-\beta$ in power

Question:

Why is this an ethical issue?

Sample Size Calculations

1. Establish dist^{n} of T under H_0 and H_1.

2. Agree on a minimum relevent difference Δ.

3.
$$
\begin{aligned}
1-\beta &= Pr(\text{reject } H_0 \mid H_1 \text{ true}) \\
&= Pr(T > t_{crit} \mid H_1 \text{ true}) \\
&\geq Pr(T > t_{crit} \mid \text{worst } H_1) \\
&= h(n, \theta)
\end{aligned}
$$

4. Solve for n:

$$h(n, \theta) = 1-\beta.$$

EXAMPLE: Drug for blood pressure

$X_1, \dots, X_n$ Change in BP following medication

Hypotheses: $H_0 : \mu_x = 0$ $H_1 : \mu_x < 0$.

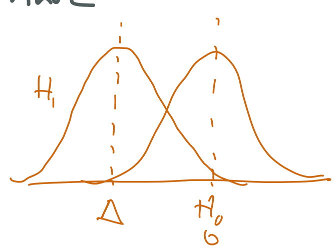
* for simplicity, suppose BP variance σ_x^2 is known & unchanged by the medication * (1-sample z-test)

Test Statistic

Estimating μ_x by $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, we have

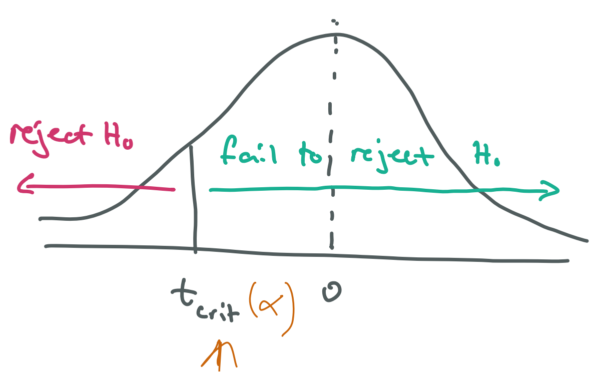
$$H_0 : \bar{X} \overset{\text{approx.}}{\sim} N\left(0, \frac{\sigma_x^2}{n}\right)$$

$$H_1 : \bar{X} \overset{\text{approx}}{\sim} N\left(\Delta, \frac{\sigma_x^2}{n}\right)$$



(where $\Delta < 0$)

test statistic $T = \frac{\bar{X} - 0}{\sigma_x / n} \sim N(0,1)$ under the null.

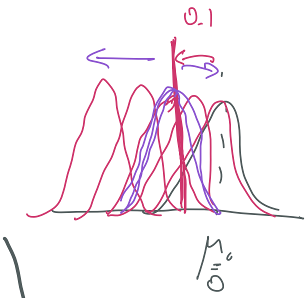


z_α

Power Calculation

$$\text{Power} = 1 - \beta$$

$$= \Pr(\text{reject } H_0 \mid H_1 \text{ true})$$



$$= \Pr\left(\frac{\bar{X}\sqrt{n}}{\sigma_x} < z_\alpha \,\middle|\, H_1 \text{ true} \right)$$

Pessimism

$$\geq \Pr\left(\frac{\bar{X}\sqrt{n}}{\sigma_x} < z_\alpha \,\middle|\, \bar{X} \sim N\left(\Delta, \frac{\sigma_x^2}{n}\right) \right)$$

$$= \Pr\left(\bar{X} < \frac{\Phi^{-1}(\alpha)\,\sigma_x}{\sqrt{n}} \,\middle|\, \bar{X} \sim N\left(\Delta, \frac{\sigma_x^2}{n}\right) \right)$$

$$= \Pr\left(Z < \frac{\left[\frac{\Phi^{-1}(\alpha)\,\sigma_x}{\sqrt{n}} - \Delta\right]}{\sigma_x / n} \right)$$

$$= \Pr\left(Z < \Phi^{-1}(\alpha) - \frac{\Delta\sqrt{n}}{\sigma_x} \right)$$

$$= \Phi\left(\Phi^{-1}(\alpha) - \frac{\Delta\sqrt{n}}{\sigma_x} \right) = h(n, \alpha, \Delta, \sigma_x)$$

Setting $h(n, \theta) = 1 - \beta$ and rearranging for n:

$$\text{Power} \geq h(n,\theta) = \Phi\left(\Phi^{-1}(\alpha) - \frac{\Delta\sqrt{n}}{\sigma_x}\right) = 1-\beta$$

rearranging:

$$n \geq \left[\frac{\sigma_x}{\Delta}\left(\Phi^{-1}(\beta) + \Phi^{-1}(\alpha)\right)\right]^2 .$$

note: $\Phi^{-1}(1-a) = -\Phi^{-1}(a)$

We now have an expression to calculate the required sample size for given α, β, Δ & σ_x .

Challenges

- Properly defining H_0 and H_1.
- Agreeing on Δ
- Establishing dist^{n} of test statistic
- Inverting $h(n,\theta) = 1-\beta$

  ↳ 2-sided tests have absolute values

  ↳ σ_x^2 unknown $\Rightarrow$ t-test $\Rightarrow$ df depend on n

  ↳ 2 sample tests: dealing with group size

  ↳ Estimating something other than a mean.