

# F6 del II - Decision Errors

## Statistisk översikt kurs

Ulf Högnäs

Statistiska institutionen  
Stockholms universitet

April 5, 2025



Stockholm  
University

- 1 Type I and Type II Errors
- 2 Systematic Reviews och metastudier

## Type I and Type II Errors

# Type I and Type II Errors

**Table 14.1:** Four different scenarios for hypothesis tests.

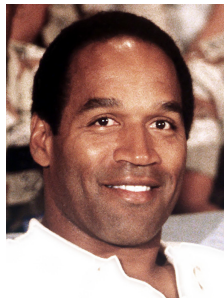
Truth	Test conclusion	
	Reject null hypothesis	Fail to reject null hypothesis
Null hypothesis is true	Type I error	Good decision
Alternative hypothesis is true	Good decision	Type II error

# Exempel: Feltyper i en domstol

I en domstol är den tilltalade antingen oskyldig ( $H_0$ ) eller skyldig ( $H_A$ ).

Vad representerar felen i detta sammanhang?

- **Typ I-fel** Domstolen fäller en oskyldig person.  
( $H_0$  är sann, men förkastas – oskyldig men döms)
- **Typ II-fel** Domstolen friar en skyldig person.  
( $H_A$  är sann, men  $H_0$  förkastas inte – skyldig men frias)



**Figure:** O.J. Simpson 1990. Public Domain.

## Type I error

- Att förkasta nollhypotesen när nollhypotesen är sann
- “Convicting the innocent”
- Exempel:
  - Power Posing - vetenskaplig studie and TED talk
  - Upprepade försök att replikera studien har misslyckats



Fig. 1. The two high-power poses used in the study. Participants in the high-power-poses condition were posed in expansive positions with open limbs.

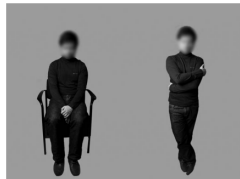


Fig. 2. The two low-power poses used in the study. Participants in the low-power-poses condition were posed in contractive positions with closed limbs.

Power Posing : Brief Nonverbal Displays Affect Neuroendocrine Levels and Risk Tolerance Carney, Cullen, Yap. Psychological Science (2010)

# Typ I-fel – Replikationskris

## Experiment och typ I-fel

- Antag att tjugo forskarlag har samma hypotes, men hypotesen är falsk –  $H_0$  är sann
- Alla forskarlagen använder en signifikansnivå på 5% och har ingen kommunikation med varandra
- Eftersom nollhypotesen är sann har varje lag 5% risk att begå ett typ I-fel
- Det finns 64% risk att minst ett typ I-fel inträffar och 26% risk att minst två typ I-fel inträffar
- Vi kan få flera studier som visar stöd för en falsk hypotes, medan nollresultat glöms bort!

## Replikationskris

- Forskare ledda av Brian Nosek försökte replikera 100 resultat från prestigefyllda psykologitidskrifter
- Endast 36% gav signifikanta resultat. Effektstorlekarna var också mindre
- Ett problem inom många discipliner
- Några orsaker ([Wikipedia: Replication Crisis](#))
  - Publikationsbias – endast positiva resultat publiceras
  - Felaktig matematik eller statistiska metoder
  - “Publish or perish”-kultur
  - Rena bedrägerier



## Typ II-fel

- Att felaktigt underlåta att förkasta nollhypotesen
- "Att fria den skyldige"
- Mycket vanligt, särskilt när det är svårt att få stora urval
- Exempel (påkittat)
  - Hypotes: en majoritet av studenterna vid SU har en iPhone
  - Jag frågar 25 studenter – 14 svarar "ja, iPhone", 11 svarar "nej"
  - Använder 5% signifikansnivå
  - `prop.test(14, 25, 0.5, alternative = "greater")`

```
> prop.test(14, 25, 0.5, alternative = "greater")
```

```
1-sample proportions test with continuity correction
```

```
data: 14 out of 25  
X-squared = 0.2, df = 1, p-value = 0.3
```

# Att inte förkasta är inte (bra) bevis för nollhypotesen

Viktigt! När vi säger att vi “inte förkastar nollhypotesen” säger vi inte att “nollhypotesen är sann”.

- Vi kan misslyckas för att stickprovet är för litet (iPhone-exemplet)
- Vi kan misslyckas för att sanningen ligger nära nollhypotesen, även om den är falsk
- Sätt alltid den hypotes som kräver bevis som alternativhypotes
- Om du vill studera effekten av ett läkemedel – vad bör då vara alternativhypotesen?

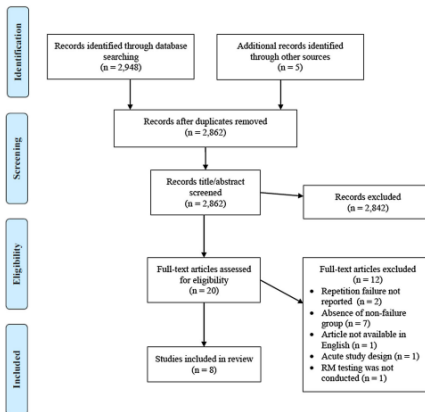
## Systematic Reviews och metastudier

En **systematic review** sammanfattar metodiskt flera studier inom samma ämnesområde

- Man sätter upp regler för vilka typer av studier som ska inkluderas (kvalitet, relevans)
- Söker i databaser efter relevanta studier
- Slutar med en samling av de bästa och mest relevanta studierna inom ämnet
- Ibland kan mycket liknande studier kombineras till en **metastudie** eller till och med en **meta-regression**

# Systematic Reviews and Meta Studies

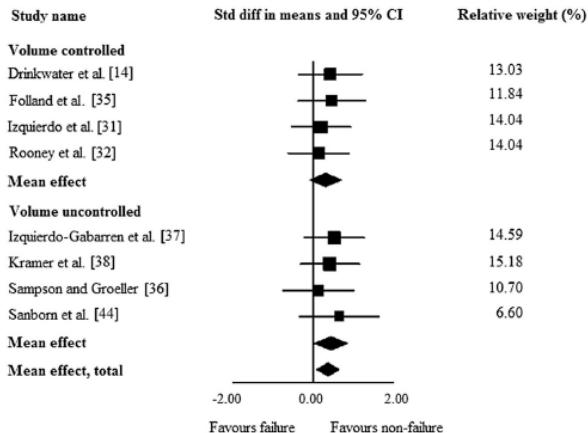
Fig. 1 Flow diagram.  
RM repetition maximum



Effect of Training Leading to Repetition Failure on Muscular Strength:  
A Systematic Review and Meta-Analysis. [Davies et al. Sports. Med. 2016 Apr](#)

# A Systematic Review - Training to Muscular Failure

Youtube: from Pumping Iron



Effect of Training Leading to Repetition Failure on Muscular Strength:  
A Systematic Review and Meta-Analysis. Davies et al. Sports. Med. 2016 Apr