

# Kausal inferens av Sebastian, Henrik og Gard

## 1. Definisjon av en kausal effekt

En kausal effekt er hvor mye et utfall endrer seg på grunn av en bestemt påvirkning. For eksempel, hvis en intervensjon øker Lean body mass med 2kg sammenlignet med en kontrollgruppe som ikke trener, er enhetene økning i intervensjonsgruppen den kausale effekten av treningsprogrammet på lean bodymass.

## 2. Hvordan kan vi bruke en eksperimentell studie for å estimere en kausal effekt?

I en eksperimentell studie, som i et randomisert kontrollert forsøk, deles deltakerne tilfeldig inn i to eller flere grupper: en gruppe som får intervensjonen, og en gruppe som forblir kontroll.

Randomiseringen gjør at gruppene i snitt blir like med hensyn til alle faktorer som kan påvirke utfallet, både faktorer vi kan måle og faktorer vi ikke kan måle. Dette betyr at vi kan sikre eventuelle forskjeller i bakgrunns karakteristikk mellom gruppene blir fordelt tilfeldig.

Når vi sammenligner resultatene mellom gruppene, vet vi derfor at forskjellen i utfall ikke skyldes andre faktorer, men skyldes intervensjonen. Den eventuelt observerte forskjellen mellom gruppene kan gi et direkte estimat på den kausale effekten av tiltaket.

## 3. Hvordan kan vi bruke en observasjonsstudier for å estimere en kausal effekt?

I observasjonsstudier styrer vi ikke hvem som får tiltaket eller eksponeringen – vi registrerer det som skjer naturlig. For å isolere effekten av eksponeringen må vi kontrollere for faktorer som påvirker både eksponering og utfall, altså konfunderende variabler.

Målet er å gjøre de som har ulik eksponering så like som mulig på andre relevante faktorer, slik at forskjellen i utfall i større grad kan tilskrives eksponeringen selv, omtrent som randomisering i et eksperiment.

For eksempel, om vi studerer styrketrening og lean body mass, må vi kontrollere for kosthold, alder eller kjønn som kan påvirke både hvor mye folk trener og lean body mass. Når vi har justert for dette, kan vi estimere den kausale effekten av trening.

#### **4. Hvilke farer finnes i (særlig) observasjonsstudier når vi ønsker å finne en kausal effekt?**

I observasjonsstudier kan vi ikke direkte estimere kausal effekt, fordi deltakerne ikke følger en intervensjon, de randomiseres ikke og grupper kan skille seg systematisk. Vi kan derfor kun estimere effekten ved å justere for kjente og observerbare forskjeller, men usikkerhet og skjevheter kan fortsatt påvirke resultatet.

- Konfundering: Variabler som påvirker både eksponering og utfall kan gi falsk sammenheng.
- Seleksjonsbias: Systematiske forskjeller mellom personer som inngår i studien og de som ikke gjør det.
- Omvendt kausalitet: Utfallet kan påvirke eksponeringen, ikke omvendt.
- Collider bias: Justering for en variabel som påvirkes av både eksponering og utfall kan skape en falsk sammenheng.
- Målefeil: Unøyaktige målinger av eksponering eller utfall gir skjeve estimater.

#### **5. Beskriv en konfunderende effekt.**

En konfunder er en variabel som påvirker både eksponering og utfall. Hvis vi ikke kontrollerer for den, kan vi feilaktig tro at eksponeringen har effekt på utfallet.

#### **6. Beskriv effekten av å kontrollere for en variabel som påvirkes av en intervensjon når vi ønsker å estimere effekten av intervensjonen på en annen variabel.**

En mediator er en variabel som ligger mellom intervensjonen og utfallet, og forklarer hvordan eller hvorfor intervensjonen påvirker utfallet.

Når man kontrollerer for en mediator, fjerner man den indirekte effekten som går gjennom denne variabelen. Resultatet blir at man kun estimerer den direkte effekten av intervensjonen, ikke den totale effekten. Dette kan gi et skjevt estimat hvis målet er å finne den totale effekten.

For eksempel, en treningsintervensjon kan øke lean body mass både direkte og indirekte via økt matinntak. Hvis vi kontrollerer mediators matinntak, vil den målte effekten av treningen bli lavere enn den totale effekten, fordi den indirekte effekten via matinntaket fjernes.

**7. Beskriv effekten av å kontrollere for en variabel som påvirkes av utfall og avhengig variabel i en regresjonsanalyse (en collider).**

En collider er en variabel som påvirkes både av den avhengige og uavhengige variabelen. Eksempelvis:

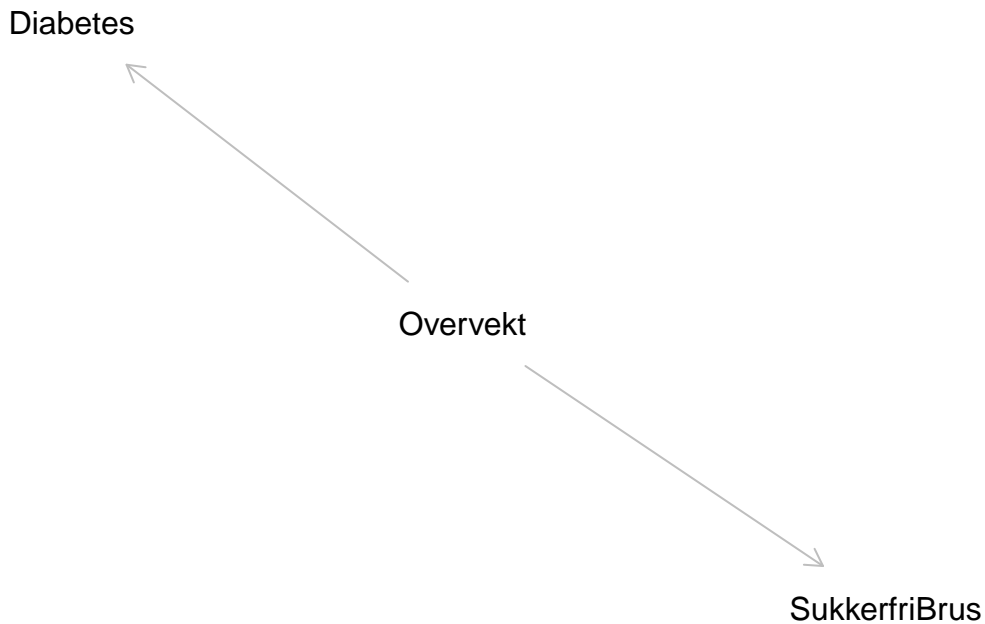
- Treningsintensitet = X (Uavhengig variabel)
- Romtemperatur = Y (Avhengig variabel)
- Mengde svette = C (Collider)

Både romtemperatur og treningsintensitet påvirker mengde svette, men treningsintensitet påvirker ikke romtemperatur. Visuelt eksempel:

- Romtemperatur -> Mengde svette <- Treningsintensitet

Hvis man kontrollerer for mengde svette i en regresjonsmodell vil det skape en falsk korrelasjon mellom Treningsintensitet og Romtemperatur. Derfor er det ugunstig å kontrollere for mengde svette, som er en konsekvens av både treningsintensitet og romtemperatur i modellen.

**8. Beskriv ved hjelp av en grafisk modell (Directed Acyclic Graph) en positiv sammenheng mellom konsumpsjon av sukkerfri brus og diabetes. Begrense antall variabler til < 4 og begrunn din modell**



Forklaringen på den positive sammenhengen mellom sukkerfri brus og diabetes er at det fins en confounder: Overvekt. Overvekt øker sjansen for utvikling av diabetes, i tillegg til et høyere inntak av sukkerfri brus. Derfor kan man si at det er korrelasjon mellom sukkerfri brus og diabetes, ikke en kausal effekt.