

Konfidensintervall

Stein Atle Lie, statistiker, professor. Institutt for klinisk odontologi, Det medisinske fakultet, Universitetet i Bergen

I statistikk betyr begrepet konfidens i hvilken grad man kan ha tillit til en verdi. Konfidensintervall kunne man tenke på som et tillitsintervall eller sikkerhetsintervall.

Ved presentasjon av oppsummerende tall og statistikk for en studie må man ta forbehold om usikkerhet for de innsamlede dataene. I en studie finner man for eksempel at gjennomsnittlig lommedybde hos røykere er 3,8 mm. Da vet man at dette gjennomsnittet representerer røykerne i studien og altså er et (presist) gjennomsnitt for disse. Hvis røykerne i studien representerer røykere generelt kan man også anta at 3,8 mm vil være «den beste gjetning» for alle røykere, for eksempel i populasjonen av røykere i Norge.

Ønsker man å trekke slutninger om en populasjon basert på et utvalg i en studie oppstår et problem. Hvis studien gjentas flere ganger vil gjennomsnittet variere mellom de ulike studiene. Har man et stort utvalg, for eksempel 1 000, vil man også være mer sikker på at det observerte gjennomsnittet er nærmere «det riktige» gjennomsnittet for hele populasjonen, enn om studien er liten, for eksempel 30.

I odontologi, medisin og biologi er det nærmest en fast regel å benytte 95 % konfidensintervall (unntaksvis 90 % og 99 %). Det vil si at man beregner et intervall som med 95 % sikkerhet omfatter det «virkelige» gjennomsnittet i populasjonen. Beregningen tar hensyn til antall deltakere i studien og variasjonen mellom studiedeltakere. Konfidensintervall kan beregnes for mange statistiske størrelser, slik som andeler, prevalenser, korrelasjoner og gjennomsnitt.

Eksempel

I en studie fra 2015 med 40 røykere (Bunæs DF& al. Site-specific treatment outcome in smokers following non-surgical and surgical

periodontal therapy. J Clin Periodontol. 2015; 42: 933-42.) ble det funnet at lommedybde for røykerne var 3,8 mm. Denne lommedybden vil da representere gjennomsnittet for de 40 røykerne i studien, men vil (antagelig) ikke være det samme tallet om alle røykere ble inkludert. Man beregner derfor et konfidensintervall rundt det observerte gjennomsnittet der det er en viss (95 %) sannsynlighet for at det «riktige» gjennomsnittet finnes i intervallet.

For røykere var 95 % konfidensintervall (95 % KI) for lommedybde fra 3,6 mm til 4,0 mm. Det vil si at vi med 95 % sikkerhet kan anta at lommedybde for røykere i populasjonen ligger i dette intervallet. Det skrives ofte slik: «3,8 mm (95 % KI: 3,6-4,0)». For ikke-røykere var gjennomsnitt og konfidensintervall 3,4 mm (95 % KI: 3,2-3,5).

Vi ser i vårt eksempel at gjennomsnittlig lommedybde for røykere var høyere enn for ikke-røykere. For denne forskjellen kan vi også beregne et 95 % KI. Vi finner at gjennomsnittlig differanse er 0,4 (95 % KI: 0,2-0,7). Siden konfidensintervallet ikke omslutter verdien null, altså ingen forskjell, er det her rimelig å anta at røykere og ikke-røykere har ulik lommedybde.

Konfidensintervall og p-verdi

Konfidensintervall kan benyttes som et alternativ til p-verdi for å bedømme om grupper er ulike. Ut fra eksempelet ovenfor er verdien null ikke innenfor 95 % konfidensintervallet. Hvis vi hadde beregnet p-verdien for forskjellen i eksempelet vet vi at denne ville vært mindre enn 0,05 (5 %). Når man sammenligner to grupper, som i eksempelet, kan konfidensintervall og p-verdi betraktes som to sider av samme sak og de vil ha direkte sammenheng.

stein.lie@uib.no