



Når målingen går i taket

MEDISIN OG TALL

ARE HUGO PRIPP

E-post: apripp@ous-hf.no

Are Hugo Pripp er forsker og biostatistiker ved Oslo senter for biostatistikk og epidemiologi, Forskningsstøtteavdelingen, Oslo universitetssykehus, og professor II ved Fakultet for helsevitenskap, OsloMet – storbyuniversitetet.

Forfatter har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

Noen målinger gir bare resultater mellom bestemte maksimums- og minimumsverdier som ofte kalles målemetodens tak og gulv. Når målingen går i taket (eller i gulvet), blir både tolkningen og den statistiske analysen påvirket.

En del målemetoder, som tekniske måleinstrumenter, laboratorieanalyser og spørreskjemaer, har bestemte maksimums- og minimumsverdier. For eksempel består livskvalitetsskjemaet SF-36 av åtte dimensjoner, der man på hver dimensjon skårer 0–100 poeng. Uansett hvor sprek man er, er det ikke mulig å oppnå mer enn 100 poeng for dimensjoner som fysisk funksjon, vitalitet eller mental helse. Noen ganger oppnår en stor andel av studiedeltagerne målemetodens maksimums- eller minimumsverdi, hvilket kompliserer tolkningen og påvirker de statistiske analysene. Vi kaller dette tak- eller gulveffekten. Det statistiske prinsippet for tak- og gulveffekten er likt, så her omtaler jeg i hovedsak takeffekten.

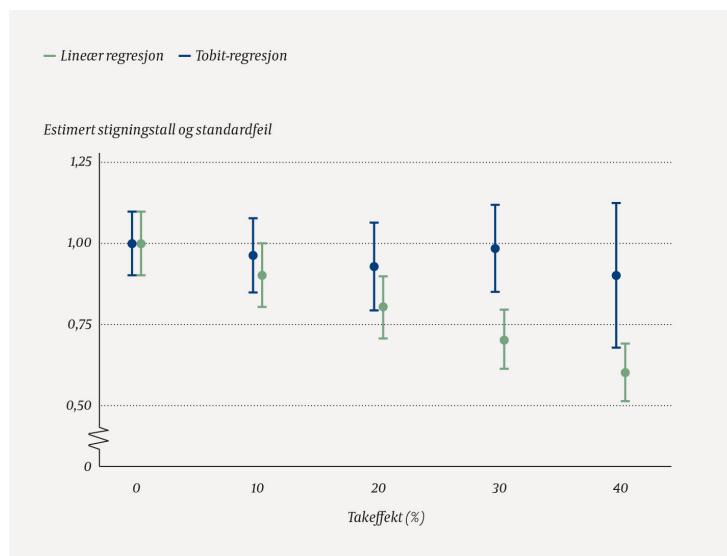
Årsaker til en takeffekt

En maksimumsverdi kan uttrykke en perfekt tilstand, men i mange tilfeller kunne verdien blitt høyere med en modifisert målemetode. Hvis vi ikke måler en egenskap ut over en gitt grense, kan vi miste verdifull informasjon. La oss tenke oss en eksamensoppgave i statistikk der halvparten av studentene oppnår maksimal skår. Hvis eksamensoppgaven hadde vært mer krevende, kunne vi sannsynligvis differensiert disse studentene. På samme vis kan noen helserelevante målinger og tester ikke differensiere de med lavest eller høyest funksjon i utvalget (1). Hvis dette skjer for en betydelig andel i utvalget, tyder det på at måleinstrumentet er utviklet og tilpasset en annen studiepopulasjon. Det er derfor hensiktsmessig å velge målemetoder tilpasset studieutvalget.

Effekt på statistiske resultater

Statistiske resultater, for eksempel stigningstallet fra en regresjonsanalyse, påvirkes av takeffekten. I figur 1 er dette illustrert med simulerte data. Forklaringsvariabelen er i alle eksemplene normalfordelt og uten takeffekt. Derimot uttrykker utfallsvariabelen målemetoder med økende grad av takeffekt. Med en tilpasset målemetode uten takeffekt er det forventede stigningstallet på 1,0. For målemetoder med takeffekt vil deltagere med de

høyeste verdiene få et resultat tilnærmet likt målemetodens maksimumsverdi. Da blir det estimerte stigningstallet fra regresjonsanalysen systematisk redusert. Er takeffekten 20 % eller mer, blir effekten betydelig, og gir et prosentvis redusert estimat tilnærmet lik andel deltagere med målemetodens maksimumsverdi.



Figur 1 Det estimerte stigningstallet og standardfeilen (angitt med hhv. punktet og intervallet) fra lineær regresjon eller tobit-regresjon av simulerte data med økende andel studiedeltagere med maksimumsverdi for utfallsvariabelen (økende prosentvis takeffekt). Forklaringsvariabelen i alle regresjonsanalysene var normalfordelt og uten takeffekt.

Statistisk analyse

Mange statistiske metoder, deriblant t-test, korrelasjon og regresjonsanalyse, forutsetter ingen takeffekt i dataene. Tar vi ikke hensyn til dette, vil de statistiske resultatene bli påvirket og systematisk forskjøvet. Ideelt sett bør vi unngå takeffekter ved å velge målemetoder slik at alle måles og differensieres uansett grad av den aktuelle tilstanden. Hvis det ikke er mulig, bør vi bruke statistiske metoder som kan håndtere variabler med takeffekter. Ikke-parametriske metoder har vært brukt, men resultatene er avhengige av fordelingen av dataene (2). En annen mulighet kan være å kategorisere utfallsvariabelen og endre regresjonsmodellen fra lineær regresjon til mer avanserte metoder som logistisk, ordinal eller multinomial regresjon. Selv om kategorisering er en opplagt metode for å omgå problemet med takeffekt, er det av mange årsaker en dårlig løsning med blant annet tap av statistisk informasjon og styrke (3).

En bedre tilnærming er å anta at målingen teoretisk sett kunne overskredet maksimumsverdien. Hadde vi for eksempel lagt inn mer krevende øvelser i en fysisk funksjonstest, kunne vi differensiert mellom alle dem som fikk maksimumsverdien i den opprinnelige testen. I så fall antar vi at studiedeltagere med målemetodens maksimumsverdi egentlig skårer denne eller en høyere verdi. Vi kaller dette sensurerte data. Egne regresjonsmetoder er utviklet for dette, deriblant metoden tobit-regresjon (4, 5). Selv med en betydelig takeffekt for dataene i figur 1, ga tobit-regresjon estimerte stigningstall tilnærmet like den forventede verdien på 1,0. At standardfeil i tobit-regresjonsanalyse øker med økende takeffekt, uttrykker større usikkerhet til målinger med takeffekt, noe som er forventet.

Takeffekten kan gi systematiske feil

Forskjeller i bakgrunnsvariabler mellom grupper som sammenlignes, manglende data, konfundering og endring i målemetoder underveis i en studie kan gi systematiske feil. På samme måte som man kartlegger og kontrollerer for slike feilkilder i studiedesignet og de statistiske analysene, må man også undersøke effekten av målinger med tak- og

gulveffekter. De kan påvirke de statistiske resultatene i like stor grad. Metoder for sensurerte data, som for eksempel tobit-regresjon, er tilgjengelig i blant annet programpakken Stata og R. Vær oppmerksom på takeffekter. Da unngår man at studien går i gulvet når målingen går i taket.

LITTERATUR:

1. Bruce B, Fries J, Lingala B et al. Development and assessment of floor and ceiling items for the PROMIS physical function item bank. *Arthritis Res Ther* 2013; 15: R144. [PubMed][CrossRef]
 2. Rorden C, Bonilha L, Nichols TE. Rank-order versus mean based statistics for neuroimaging. *Neuroimage* 2007; 35: 1531-7. [PubMed][CrossRef]
 3. Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ* 2006; 332: 1080. [PubMed][CrossRef]
 4. Austin PC, Escobar M, Kopec JA. The use of the Tobit model for analyzing measures of health status. *Qual Life Res* 2000; 9: 901-10. [PubMed][CrossRef]
 5. McBee M. Modeling outcomes with floor or ceiling effects: an introduction to the Tobit model. *Gift Child Q* 2010; 54: 314-20. [CrossRef]
-

Publisert: 18. februar 2019. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.18.0880
© Tidsskrift for Den norske legeforening 2020. Lastet ned fra tidsskriftet.no