



Pearsons khikvadrattest

MEDISIN OG TALL

STIAN LYDERSEN

E-post: stian.lydersen@ntnu.no

Stian Lydersen er dr.ing. og professor i medisinsk statistikk ved Regionalt kunnskapssenter for barn og unge – psykisk helse og barnevern (RKBU Midt-Norge) ved Institutt for psykisk helse, NTNU. Forfatteren har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

MORTEN WANG FAGERLAND

Morten Wang Fagerland er ph.d. og leder for Seksjon for biostatistikk og epidemiologi ved Oslo universitetssykehus.

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

PETTER LAAKE

Petter Laake er professor emeritus ved Avdeling for biostatistikk ved Institutt for medisinske basalfag, Universitetet i Oslo, og professor II ved Avdeling for helse- og sosialfag ved Høgskolen i Molde.

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

En vanlig statistisk problemstilling er å undersøke om det er en sammenheng mellom to variabler i en krysstabell (2x2-tabell). I den første av tre artikler om dette temaet viser vi at Pearsons khikvadrattest er intuitivt forståelig og lett å beregne, og at den trygt kan brukes i store utvalg.

Resultatene av et randomisert kontrollert forsøk med to behandlingsgrupper og to utfall kan oppsummeres som i tabell 1 (1). Tilsvarende kan resultatene av en kasus-kontroll-studie presenteres som i tabell 2 (2). Vi vil undersøke om det er en sammenheng mellom de to variablene trening med fysioterapeut og kontinens i det første eksempelet, og røyking og lungekreft i det andre.

Tabell 1

Trening av bekkenbunnsmuskulatur etter radikal prostatektomi. Resultater fra en randomisert kontrollert studie med ukentlig trening med fysioterapeut i ett år (gruppe A) versus veiledning om egentrening (gruppe B) (1).

Kontinent etter 1 år			
Gruppe	Ja	Nei	Sum
A	33	3	36
B	28	11	39
Sum	61	14	75

Tabell 2

Pasienter med lungekreft versus andre pasienter. Fra en kasus-kontroll-studie (2).

Eksposering	Lungekreft	Andre sykdommer	Sum
Røyker	688	650	1 338
Ikke-røyker	21	59	80
Sum	709	709	1 418

Vi skal altså sammenlikne to sannsynligheter: I det første eksempelet vil vi sammenlikne sannsynlighetene for kontinens ved de to behandlingene. I det andre eksempelet vil vi sammenlikne sannsynlighetene for å røyke hos personer med og uten diagnosen lungekreft. Under nullhypotesen er det ingen sammenheng, det vil si at de to sannsynlighetene er like. For tallene i tabell 1 vil estimert sannsynlighet for kontinens da bli $61 / 75 = 0,813$, uansett behandlingsgruppe. Forventet antall kontinente i den første behandlingsgruppen ville bli $36 \cdot 0,813 = 29,28$. Hvor mye avviker det vi har observert, fra det vi ville forventet under nullhypotesen? Pearsons khikvadrattest bygger på følgende avstandsmål: Vi beregner avstanden mellom det observerte og det forventede som $(33 - 29,28)^2 / 29,28$. Ved å summere dette over de fire cellene i tabellen fås et samlet mål på avvik fra nullhypotesen, også kalt Pearsons khikvadratobservator, χ^2 . Denne blir i de to eksemplene lik henholdsvis 4,87 og 19,13. P-verdien er sannsynligheten for å observere det vi har observert eller noe mer ekstremt, gitt at nullhypotesen er sann. Pearsons khikvadratobservator er tilnærmet khikvadratfordelt med én frihetsgrad, og dette gir tilnærmede p-verdier på $p = 0,027$ og $p < 0,001$ for henholdsvis tabell 1 og tabell 2. Ifølge Cochrans kriterium bør denne tilnærmingen bare brukes hvis alle forventede antall er over 5. I tabell 1 blir det minste forventede antallet $36 \cdot 14 / 75 = 6,72$, så kriteriet er oppfylt. Merk at kriteriet bygger på det forventede antallet som her er 6,72, og ikke det observerte antallet, som er 3. I tabell 2 blir det tilsvarende forventede antallet $80 \cdot 709 / 1418 = 40$, som oppfyller kriteriet med stor margin, og Pearsons khikvadrattest kan trygt brukes.

Unngå Yates' kontinuitetskorreksjon

Noen benytter Pearsons khikvadrattest med Yates' kontinuitetskorreksjon. Hensikten med denne korreksjonen er å beregne p-verdien med større nøyaktighet enn den tilnærmede p-verdien. Men Yates' korreksjon har vært omdiskutert. Denne korreksjonen fører til redusert statistisk styrke, samtidig som p-verdien fremdeles er tilnærmet og ikke eksakt. Karim F. Hirji skriver at korreksjonen ikke bør brukes og mest er å betrakte som en historisk kuriositet (3, s. 149). Vi har tidligere evaluert alternative tester for 2x2-tabeller og konkluderer også med at denne korreksjonen ikke er å anbefale (4, s. 105–13 og 175).

Alternativer ved små utvalg

Pearsons khikvadrattest som beskrevet her bør bare brukes dersom forventet antall observasjoner i cellene er større enn 5. Tradisjonelt har mange anbefalt å bruke Fishers eksakte test i små utvalg. Dette vil vi komme tilbake til i den andre og tredje artikkelen i denne serien om 2x2-tabeller i Medisin og tall.

LITTERATUR:

1. Overgård M, Angelsen A, Lydersen S et al. Does physiotherapist-guided pelvic floor muscle training reduce urinary incontinence after radical prostatectomy? A randomised controlled trial. *Eur Urol* 2008; 54: 438–48. [PubMed][CrossRef]
2. Doll R, Hill AB. Smoking and carcinoma of the lung; preliminary report. *BMJ* 1950; 2: 739–48. [PubMed][CrossRef]

3. Hirji KF. Exact analysis of discrete data. Boca Raton, FL: Chapman and Hall/CRC, 2006.
 4. Fagerland MW, Lydersen S, Laake P. Statistical Analysis of Contingency Tables. Chapman and Hall/CRC, 2017.
-

Publisert: 9. september 2019. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.18.0125

© Tidsskrift for Den norske legeforening 2020. Lastet ned fra tidsskriftet.no