
Alternativer til Fishers eksakte test

MEDISIN OG TALL

STIAN LYDERSEN

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.

Fishers eksakte test er det tradisjonelle alternativet for å undersøke sammenheng mellom to variabler i en krysstabell dersom utvalget er lite. Vi anser imidlertid eksakte ubetingede tester som gullstandarden for 2×2 -tabeller.

Dersom man vil undersøke om det er en sammenheng mellom to variabler i en krysstabell, kan Pearsons khikvadrattest brukes for store utvalg (1), mens Fishers eksakte test er det tradisjonelle alternativet for små utvalg (2). Tabell 1 viser resultatet av et randomisert kontrollert forsøk (3). Nullhypotesen er at sannsynligheten for suksess (24 timers overlevelse) er den samme for de to behandlingene, og alternativhypotesen er at det er forskjell i overlevelse

mellom behandlingene. Utvalget er for lite til at Pearsons khikvadrattest kan anbefales, og den tradisjonelle løsningen er å bruke Fishers eksakte test, som gir en tosidig p-verdi på 0,054 (2). Denne testen betinger på kolonnesummene, dvs. at p-verdien beregnes som om det totale antallet suksesser var forutbestemt hvis nullhypotesen er riktig.

Tabell 1

Behandling av barn med hjertestans. Høy dose versus standard dose adrenalin (3).

Behandling	24 timers overlevelse		
	Ja	Nei	Sum
Standard dose	7	27	34
Høy dose	1	33	34
Sum	8	60	68

En eksakt ubetinget test

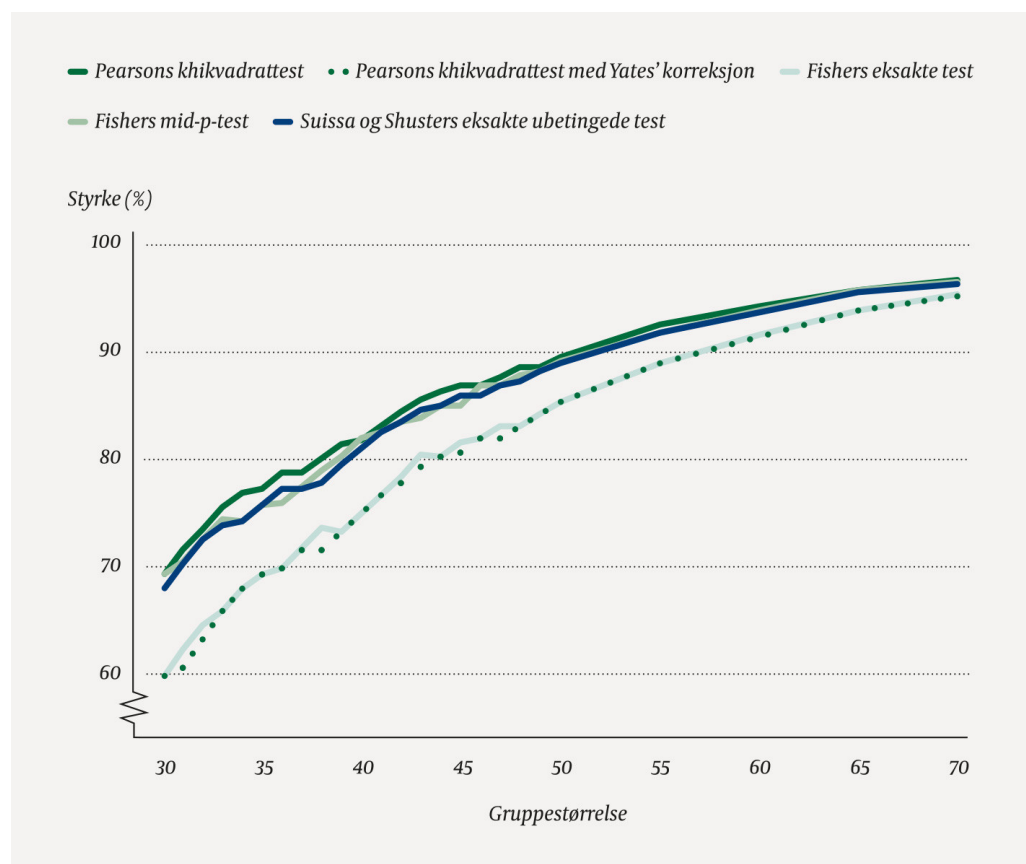
En betinget hypotesetest som Fishers eksakte test bygger på anerkjente statistiske prinsipper som garanterer at testen holder signifikansnivået. Det betyr at dersom man setter signifikansnivået lik 5 %, så er man garantert at sannsynligheten for feilaktig å forkaste nullhypotesen ikke overstiger 5 %. Men det er mulig å oppnå dette i en såkalt ubetinget test uten å betinge på kolonnesummen. En ubetinget test basert på Pearsons khikvadratobservator (1) kalles Suissa og Shusters test eller z-pooled test ((4), s. 98–112). De underliggende beregningene er komplekse, men kan enkelt gjennomføres med programvarene StatXact, Matlab og R med filer fra www.contingencytables.com, eller på nettstedet www.stat.ncsu.edu/exact/. For tabell 1 fås $p = 0,029$.

Mid-p-test

En mid-p-verdi er basert på en betinget eksakt p-verdi, f.eks. fra Fishers eksakte test. Den er enkel å beregne og gir tilnærmet samme resultat som en eksakt ubetinget p-verdi. Mid-p-verdien inkluderer bare halve sannsynlighetene for det observerte utfallet (punktsannsynligheten), og kan beregnes ved å trekke halve punktsannsynligheten fra den betingede p-verdien ((4), s. 97). Den ensidige Fishers eksakte p-verdi for tabell 1 er 0,027. Punktsannsynligheten for det observerte utfallet er 0,025 (dette oppgis ofte i statistiske programvarer), og den ensidige Fisher-mid-p-verdien blir $0,027 - 0,025 / 2 = 0,015$, som gir en tosidig Fisher-mid-p-verdi på $2 \cdot 0,015 = 0,030$.

Anbefalinger

For tallene i tabell 1 så vi at p-verdien ble forskjellig avhengig av om vi brukte Fishers eksakte test ($p = 0,054$), en eksakt ubetinget test ($p = 0,029$), eller Fishers mid-p-test (mid $p = 0,030$). Valg av metode skal selvfølgelig ikke gjøres i ettertid ut fra hvilken p-verdi som er lavest, men bestemmes på forhånd for gitt studiedesign og utvalgsstørrelse. Figur 1 viser et representativt eksempel på statistisk styrke for fem alternative tester i en 2×2 -tabell. Pearsons khikvadrattest, Fishers mid-p-test og en eksakt ubetinget test har tilnærmet lik styrke. Fishers eksakte test og Pearsons khikvadrattest med Yates' korreksjon har lavere styrke. Den sistnevnte har i tillegg den ulempen at den ikke alltid holder signifikansnivået, og den bør ikke brukes.



Figur 1 Statistisk styrke for fem tester, beregnet for sannsynligheter på 0,2 og 0,5 i to like store grupper. Basert på (4, s. 111).

Vi anser eksakte ubetingede tester som gullstandarden for 2×2 -tabeller fordi de har høy statistisk styrke og holder signifikansnivået. Men mangel på tilgjengelig programvare kan sette begrensninger, og i store utvalg er det enklest å benytte Pearsons khikvadrattest. I små utvalg er Fishers mid-p-test enkel å beregne ((4), s. 110–3), og den gir høyere styrke enn Fishers eksakte test.

LITTERATUR

1. Lydersen S, Fagerland MW, Laake P. Pearsons khikvadrattest. Tidsskr Nor Legeforen 2019; 139. doi: 10.4045/tidsskr.18.0125. [CrossRef]

2. Lydersen S, Fagerland MW, Laake P. Fishers eksakte test – hvordan smaker teen? Tidsskr Nor Legeforen 2019; 139. doi: 10.4045/tidsskr.19.0237. [CrossRef]
 3. Perondi MBM, Reis AG, Paiva EF et al. A comparison of high-dose and standard-dose epinephrine in children with cardiac arrest. N Engl J Med 2004; 350: 1722–30. [PubMed][CrossRef]
 4. Fagerland M, Lydersen S, Laake P. Statistical analysis of contingency tables. Boca Raton, FL: Chapman and Hall/CRC, 2017.
-

Publisert: 7. oktober 2019. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.19.0238
Opphavsrett: © Tidsskriftet 2026 Lastet ned fra tidsskriftet.no 14. februar 2026.