

Hypotesetest til skriftlig eksamen på stx

JAN B. SØRENSEN, Aalborg Studenterkursus

Kære fagkonsulent, opgavekommission og øvrige matematikkolleger. Vi har nu haft χ^2 -hypotesetest til skriftlig eksamen på stx B og stx A i nogle år, så måske det er tid til nogle overvejelser om de stillede opgaver og mulige forbedringer heraf. Jeg vil tillade mig at anbefale min artikel fra LMFK bladet nr. 4 fra 2012, da jeg bygger videre på indholdet derfra, lmfk.dk/artikler/data/artikler/1204/1204_18.pdf.

Som jeg ser det, har vi formået at lære eleverne i rimeligt omfang at anvende et CAS værktøj og at nå frem til enten at forkaste eller at acceptere en hypotese på mekanisk korrekt vis, men så desværre heller ikke så meget mere end det. Efter min mening har vi ikke reelt med eksamensopgaverne påvirket eleverne (eller hinanden for den sags skyld) til at få en faktisk forståelse af, hvad et hypotesetest er, hvornår det kan bruges, og at der faktisk er en del gråzoner, som man i forbindelse med hypotesetest bør være opmærksom på. Desuden har der i nogle af eksamensopgaverne været fejl eller i hvert fald uheldigheder i forhold til at fremme en klar forståelse af det jo ret abstrakte begreb hypotesetest.

Det helt store problem ved dette er efter min mening, at vi alt for let risikerer, at eleverne (og måske også kolleger i andre fag eller endda i matematik) i anvendelser i SRP, AT eller andre fag kommer til at bruge hypotesetest forkert eller uheldigt, fordi de mangler forståelse, men kun har lært mekanikken. Mht. fejlene i de skriftlige eksamenssæt kan det i værste fald også betyde, at en SRP- eller AT-censor får svært ved at kritisere noget klart forkert, hvis samme fejl er set til skriftlig eksamen i matematik uden at blive anfægtet.

Konkret mener jeg, at vi med fordel kan arbejde hen mod, at opgaverne til skriftlig eksamen ud over selve udregningerne fremmer en større forståelse, bl.a. vha. nedenstående 4 punkter. Jeg eksemplificerer efterfølgende ud fra allerede stillede eksamensopgaver.

- A. Fokus på, at hypotesen altid handler om den population, som stikprøven er taget fra – ikke om stikprøven selv! I hver opgave bør der være enten en eksplicit angivelse af, hvilken population stikprøven er taget fra, eller en så tydelig case, at der findes en naturlig, måske næsten åbenlys, population, som det kunne være. Hvor det sidstnævnte er tilfældet, kan der så i opgaven passende være et spørgsmål om, hvad populationen er, så det altid er tydeligt, at der er tale om en stikprøve fra en bestemt (veldefineret) population, og at hypotesen handler om populationen, ikke om stikprøven.
- B. Opgaver, hvor de forventede værdier bliver så små, at testen ikke umiddelbart er anvendelig. Et ofte anvendt kriterie, som Susanne Christensen også anfører i sin note, math.aau.dk/fileadmin/user_upload/www.math.aau.dk/Samarbejder/MatLaerer/chi2AUG09.pdf, er, at de forventede værdier alle skal være større end 5. Andre forfattere benytter dog andre, men dog lignende kriterier. Når der ikke er et entydigt kriterie, skyldes det, at fordelingen af stikprøven asymptotisk i stikprøvens størrelse er χ^2 og problemet handler om, hvornår stikprøven er for lille til, at dette holder rimeligt stik. I en korrekt besvarelse bør vi så forvente, at eleven hver gang undersøger, om der skulle være et problem med for små forventede værdier, og hvis det er tilfældet beskriver, at man så enten kan indsamle en større stikprøve, eller at man kan sammenlægge beslægtede kategorier, hvor dette giver mening. Hvis det sidste er rimeligt, vil eleven så kunne gøre dette, for derefter at lade sit CAS værktøj regne på tallene.
- C. Opgaver, hvor stikprøven er taget fra en endelig population, hvoraf stikprøven udgør for stor en andel til, at det er rimeligt at foretaget testen. Disse kan varieres mellem, at stikprøven næsten er hele populationen, så der bliver tale om deskriptiv statistik (som let kan

regnes på) eller, at stikprøven er mindre end det, men dog betænkelig stor, så man enten må afvise at kunne regne på tallene, eller i hvert fald kun gør det med forbehold.

- D. Opgaver, hvori der lægges op til, at man *efter* at have set data beslutter sig for en hypotese, man vil teste, fordi man ser noget suspekt i data. Yderligere skal der lægges op til, at eleven gør opmærksom på, at dette er problematisk, og at man ideelt set burde indsamle en ny stikprøve, som så kunne benyttes til testen, men alligevel gennemfører testen med de givne data, blot med forbehold.

Eksempel 1

Til Mat A december 2013 stillede man opgaven, der er vist vist i Opgave – Eksempel 1.

Der er tale om en stikprøve på 165 personer, taget fra en population som er de 20–70 årige – men lige efter får vi eksplicit en opgjort fordeling for netop denne population, så tingene hænger ikke så godt sammen. Endvidere står der reelt, at vi skal undersøge *stikprøvens* aldersfordeling, ikke den bagvedliggende *populations*, hvilket efter min mening reelt vil være deskriptiv statistik, da aldersfordelingen i stikprøven jo er fuldstændigt kendt i opgaven, og der derfor ikke er nogen usikkerhed at lave inferentiell statistik på.

Efter min mening kunne opgaven fint være lavet glimrende med ganske små ændringer. Fx kunne man i første linje i stedet have skrevet: *Tabellen viser aldersfordelingen blandt 165 tilfældigt udvalgte 20-70 årige tyskere*, så opgaven blev at undersøge om aldersfordelingen kan være den samme for tyskere som for danskere. I opgaven kunne man udover de allerede stillede spørgsmål eksplicit så have bedt om en forklaring på, hvad der er stikprøve og hvilken population denne stammer fra, for at fremme forståelsen af disse begreber og deres betydning. Opgave 11 fra Mat B 29. maj 2013 er et fint eksempel på dette.

Alternativt kunne opgaven laves, så den er tænkt som en deskriptiv statistikopgave, ved at undlade ord som hypotese-test i spørgsmålet, men ganske enkelt bede om en undersøgelse af, om aldersfordelingen blandt de 165 tilnærmelsesvis er den samme som i hele befolkningen. Dette kunne så fx besvares ved at eleven tegnede pindediagrammer for de to datasæt.

Endeligt kunne man have lavet opgaven åben: *Kommenter ovenstående undersøgelse med inddragelse af statistiske begreber, og gennemfør eventuelle statistiske test eller andre beregninger, som undersøgelsen giver anledning til.*

Eksempel 2

Til Mat B august 2013 stillede man opgaven, der er vist i Opgave – Eksempel 2.

Hvis man regner på tallene, viser det sig, at det forventede antal for karakteren -3 er så lavt som 1,4. Jeg har set i flere lærebøger i statistik og uanset hvilket kriterie de benytter, så er denne værdi for lav til, at man vil tillade brugen af χ^2 -testen. Hvis vi blot er enige om, at opgaven er stillet bl.a. med det formål at teste, om eleverne er klar over dette forhold, og kan agere fornuftigt i forhold hertil, er opgaven efter min mening perfekt. Så vil eleven skulle skrive, at man enten skal indsamle en større stikprøve, eller at man i dette tilfælde med nogen rimelighed kan slå karaktererne -3 og 00 sammen, hvorefter testen kan gennemføres.

Alternativt kunne man lave en tilsvarende opgave, hvor det blot ikke giver mening at slå kategorier sammen, og eleven derfor i sin besvarelse skal konstatere dette. Det kunne fx gøres ved at lave kategorier uden naturlig rangordning, så sammenlægning vil være noget vrøvl, eller fx ved at man i denne opgave skulle lægge 00 og 02 sammen, hvilket kunne anses for betænkeligt, da det er en dumpet og en bestået karakter, så de hører ikke så heldigt sammen.

Opgave – Eksempel 1

Tabellen viser aldersfordelingen af 165 tilfældigt udvalgte 20-70-årige danskere.

Alder	20-30	30-40	40-50	50-60	60-70
Antal	21	39	51	30	24

Aldersfordelingen af alle 20-70-årige danskere fremgår af nedenstående tabel.

Alder	20-30	30-40	40-50	50-60	60-70
Antal i %	18,1	20,3	22,6	20,0	19,0

Man ønsker at undersøge, om aldersfordelingen blandt de tilfældigt udvalgte danskere er den samme som aldersfordelingen blandt alle 20-70-årige danskere.

- a) Opstil en nulhypotese, og undersøg, om nulhypotesen kan forkastes på et 5% signifikansniveau.

Opgave – Eksempel 2

Karakterfordelingen ved folkeskolens 9. klasses afgangsprøve i matematik sommeren 2011 var som vist i tabellen.

Karakter	-3	00	02	4	7	10	12
Andel	0,14%	7,43%	10,96%	22,04%	31,23%	19,83%	8,37%

I en stikprøve fra 2012 blandt landets 9. klasses elever var karakterfordelingen således:

Karakter	-3	00	02	4	7	10	12
Antal elever	30	68	124	198	279	225	76

Man ønsker at undersøge nulhypotesen:

Karakterfordelingen blandt 9. klasses elever i 2012 er den samme som karakterfordelingen fra 2011.

- a) Beregn med udgangspunkt i nulhypotesen den forventede karakterfordeling for eleverne i stikprøven.
 b) Undersøg på et 5% signifikansniveau, om nulhypotesen må forkastes.

Opgave – Eksempel 3

Nogle elever på et gymnasium ønsker at undersøge om, der er en sammenhæng mellem forældres og børns rygevaner. De vælger at spørge 137 tilfældigt udvalgte elever på skolen. Resultatet af undersøgelsen fremgår af tabellen nedenfor.

Rygevaner	Forældre ryger	Forældre ryger ikke
Barn ryger	35	33
Barn ryger ikke	23	46

- a) Opstil en nulhypotese, og undersøg på et 5% signifikansniveau, om undersøgelsens resultat giver belæg for at hævde, at børns rygevaner er uafhængige af forældres rygevaner.

Eksempel 3

Til Mat B august 2012 stillede man opgaven Opgave – Eksempel 3.

Der er flere indvendinger mod opgaven. Hvis man faktisk ønsker at undersøge om børns rygevaner er uafhængige af forældrenes rygevaner, må det skarpt kritiseres kun at tage en stikprøve blandt gymnasieelever og det endda kun på en enkelt skole. Hvis man i stedet blot ønsker at undersøge hypotesen for den population, som udgøres af skolens elever og deres forældre, er stikprøven bedre, men så er den pludseligt potentielt for stor til at benytte χ^2 -test, især hvis der er tale om et mindre gymnasium. Endelig kan man indvende, at forældre ryger kontra forældre ryger ikke er uklart mht. den situation, at den ene, men ikke begge forældre ryger.

Faktisk mener jeg, at opgaven som stillet, er rigtig interessant, hvis man lave-

de den mere åben som tidligere nævnt: a) *Kommenter ovenstående undersøgelse med inddragelse af statistiske begreber, og gennemfør eventuelle statistiske test eller andre beregninger, som undersøgelsen giver anledning til.*

Så har vi pludseligt en særdeles interessant, men også langt mere åben opgave – altså ikke længere bare en ren type-opgave. Det kunne jeg godt tænke mig kom til at ske, da det for alvor ville sætte fokus på forståelsen, fremfor blot beregningerne, men derved også ville stille større krav til os som undervisere i vores egen forståelse og følgende formidling. Dette vil dog helt sikkert kræve en nærmere overvejelse.

Eksempel 4

I opgave 11 i de vejledende eksamensopgaver i χ^2 ses opgaven Opgave – Eksempel 4.

I delopgave c ses et tydeligt eksempel på, at man ud fra data i stikprøven ser noget *suspekt*, som man derfor får lyst til at teste. Hvis vi bliver enige om, at en fyldestgørende besvarelse her skal indeholde et forbehold overfor dette og en opfordring til at indsamle og teste på et nyt datasæt, så er opgaven efter min mening glimrende, da den så har et pædagogisk formål, men hvis vi accepterer løsninger uden sådan forbehold, er det efter min mening skidt.

Man kunne give flere eksempler og udvikle flere forskellige typer af opgaver, men det arbejde vil jeg gerne opfordre til sker i en gruppe under fagkonsulenten og/eller opgavekommissionen med henblik på at gøre eksamensopgaverne i statistik mere interessante.

Opgave – Eksempel 4

Ved folketingsvalget i 2007 fordelte stemmerne sig således:

Parti	S	Rad	Kons	SF	DF	V	EL	Lib. All.	Kr. Dem.
Stemmer i %	25,5	5,1	10,4	13,0	13,8	26,3	2,2	2,8	0,9

Ved en meningsmåling med 968 respondenter foretaget af Greens Analyseinstitut 1/2 - 3/2 fordelte stemmerne sig således

Parti	S	Rad	Kons	SF	DF	V	EL	Lib. All.	Kr. Dem.
Stemmer i %	26,3	5,4	9,6	16,9	14,2	24,0	2,1	1,1	0,5

Vi vil anvende meningsmålingen til at undersøge følgende nulhypotese: stemmefordelingen er uændret in forhold til valget.

- Hvad der er populationen og hvad der er stikprøven. Lav en tabel over de observerede antal stemmer og de forventede antal stemmer under nulhypotesen.
- Bestem teststørrelsen, der tester om hypotesen skal accepteres eller forkastes, og afgør på et 5% signifikansniveau om man kan sige, at stemmefordelingen har ændret sig.
- Det ser ud som om, partiet SF har haft betydelig fremgang. Opstil en 2x2 tabel inddelt efter SF og øvrige partier med observerede og forventede værdier og afgør med statistiske metoder om SF's fremgang er signifikant på 5% niveau.