

## Eksamensopgaver vedr. statistik og stikprøver

### **Skrivelse fra fagkonsulenten**

<http://www.emu.dk/gym/fag/ma/undervisningsministeriet/nyt/variabelbegrebet-stikproever.pdf>

### **Stikprøver**

Statistik er svært, fordi vi her ofte er helt ude og røre ved virkeligheden. Men netop derfor er statistik et centralt felt, når det drejer sig om samarbejde med andre fag.

Statistik drejer sig om metoder til at uddrage en indsigt om tendenser og sammenhænge ud fra kendskab til et givet talmateriale. Det givne kan være både meget begrænset og ret diffust. Statistik søger også at svare på, med hvilken grad af sikkerhed vi kan udtale os om disse tendenser, sammenhænge osv.

Hovedparten af det faglige stof er placeret under supplerende stof i læreplanerne. Det er gjort for at give bedre plads til det faglige samarbejde med andre fag. Eksempelvis er alt hvad der vedrører bearbejdning af et talmateriale til belysning af en opstillet hypotese, placeret under supplerende stof (på B- og A-niveau).

Kernestoffet, der er genstand for den skriftlige eksamen, omfatter deskriptiv statistik på hf C og på stx B og stx A yderligere stikprøvers repræsentativitet. I undervisningsvejledningen, samt i de vejledende eksamensopgaver er anført eksempler på sådanne opgavetyper.

Lad os se på opgave 4.003 i hæftet til stx B:

Vil indtagelse af urtete styrke helbredet hos de ældre? Dette ønsker en gruppe studerende at undersøge. Over en periode på 6 måneder besøger de nogle tilfældigt udvalgte beboere på et plejehjem og serverer urtete for dem. Efter 6 måneder viser det sig, at de beboere, der fik serveret urtete, faktisk har færre sygedage, end de som ikke fik serveret noget. De studerende publicerer resultatet af deres undersøgelse under overskriften: "Urtete styrker helbredet hos de ældre".

a) Kommenter denne påstand ved at stille mindst tre kritiske spørgsmål til undersøgelsen.

Hvordan går man til denne opgave, når man ikke ved mere, end der står? Sådan er det jo ofte med avisnotitser. Kan man som elev sige, at man antager undersøgelsen er gennemført efter alle forskrifter, og at den er offentliggjort i The Lancet, og at der derfor ikke er noget at komme efter? Nej det kan man ikke – hverken her eller andre steder får man point for at digte en anden opgave i stedet for at løse den stillede.

*Det første* der springer i øjnene er, at vi hverken får oplyst noget om *populationen eller stikprøven*. Er populationen alle beboere på plejehjemmet? Hvor stort er dette og hvor mange af beboerne er besøgt?

*Det næste* er, at vi ikke af teksten kan vide, hvorledes stikprøven er udvalgt. Det kan naturligvis ikke udelukkes, at stikprøven er repræsentativ efter alle kunstens regler og en række parametre, men på det oplyste grundlag spørger vi: *Er der begået systematiske fejl* (i statistikbøger oftest betegnet 'bias'). Går jeg ud på gaden og spørger om holdningen til et eller andet, er det nok helt tilfældigt, hvem jeg spørger. Men naturligvis er der begået systematiske fejl. Tilfældighed garanterer ikke, at det er en repræsentativ stikprøve. Serverer de te for de syge, demente, sengeliggende, for de med særligt dårligt helbred lige så vel som for de oppegående? Er der taget hensyn til alder og køn? Er der taget hensyn til bestemte sygdomme, der kan have indflydelse? Hvad med den gruppe, der absolut ikke drikker te? osv. osv. Disse spørgsmål drejer sig om en række specifikke parametre eller variable – og listen er svær at udtømme! – som vi siger, der skal tages hensyn til, for at undgå systematiske fejl (bias).

*Det tredje* er, at selv om det var en repræsentativ undersøgelse, hvor stikprøven var udvalgt efter alle kunstens regler, så kunne der alligevel være nogle *skjulte variable* på spil. Man tror man måler holdningen til eller virkningen af det som er i fokus i spørgsmålet – men den, man spørger, reagerer på noget helt andet eller tredje. Det kan ofte være svært at isolere effekten af sådanne skjulte variable – i det givne tilfælde, hvad skyldes urtete og hvad skyldes de unge menneskers besøg? (Sådanne skjulte variable, hvor det er svært at isolere den enes virkning frem for den anden kaldes i statistisk litteratur konfunderede variable.)

*Det fjerde*, der springer i øjnene, er selve konklusionen: "Styrker helbredet". Helbredet er her åbenbart en variabel, men *hvordan mon man måler denne styrkelse?* Det står der ikke noget om. Og det kan der sikkert være flere bud på – blodtryk, bestemte stoffer i blodet, kolesteroltal, evnen til at optage vitaminer og mineraler, søvnproblemer, appetit, hukommelse osv. Listen er udtømmelig, men allerede de nævnte viser, hvor vidt forskellige bud, der kan være. Og det er da ikke ligegyldigt, hvad der styrkes – har de målt på alt? Er der noget der måske svækkes samtidig? Og videre: Styrkelse i forhold til hvad? En fremgang fra 27 til 28 på en eller anden skala eller hvad?

*Det femte*, man i forlængelse heraf må spørge om, er, om der ikke har været en kontrolgruppe? Det tyder det ikke på – og det er også svært at drikke urtete uden at lægge mærke til det. Det er jo problemet i mange forsøg, men ikke desto mindre – uden kontrolgruppe bliver det betydeligt vanskeligere at drage pålidelige slutninger. Ideelt set skulle det naturligvis være et dobbelt-blind-forsøg.

Eleverne kan godt lære at håndtere sådanne opgaver, hvis der i undervisningen arbejdes med en række gode eksempler på det. Flere lærebøger har et ret righoldigt udvalg af sådanne eksempler, også formuleret som opgaver. Jeg skal ikke udtale mig om alle lærebøger er lige godt polstret på dette område, men hver skole bør jo have mindst et eksemplar af hvert lærebogssystem på markedet. Så kan man orientere sig og hente inspiration. Avisklip kan også give et materiale til undervisningen.

Kravet til en elevbesvarelse er selvfølgelig ikke, at de skal skrive noget så omfattende som de nævnte 5 punkter. Men af ovenstående skulle forhåbentlig fremgå, at der kan være et væld af forskellige elevbesvarelser til en sådan opgave, der alle giver topkarakter.

*Fagkonsulent Bjørn Grøn*

### **Stikord til kommentering af stikprøve**

- 1. Populationen?**
  - 2. Systematiske fejl ('bias')?**
  - 3. Skjulte variable ('konfunderede variable')?**
  - 4. Konklusionen?**
  - 5. Kontrolgruppe?**
- 

### **Vejledende eksamensopgaver STX A og B**

4.002 På en skole med 700 elever ønsker en af de politiske ungdomsorganisationer at få mulighed for at stille et bord op, hvor eleverne i spisebureauet kan hente materialer og få information. Da skolens ledelse siger nej, opfordrer organisationen alle elever til at tilkendegive, om de er for eller imod dette. 127 afgav deres stemme og heraf støttede 92 forslaget. Organisationen omdeler derefter løbesedler, hvor de skriver: "Elevundersøgelse viser, at over 70 % støtter de politiske organisationers ret til at uddele materialer på skolen".

a) Kommentér denne påstand med brug af statistiske begreber som stikprøve, population, systematiske fejl og skjulte variable.

4.003 Vil indtagelse af urtete styrke helbredet hos de ældre? Dette ønsker en gruppe studerende at undersøge. Over en periode på 6 måneder besøger de nogle tilfældigt udvalgte beboere på et plejehjem og serverer urtete for dem. Efter 6 måneder viser det sig, at de beboere, der fik serveret urtete, faktisk har færre sygedage, end de som ikke fik serveret noget. De studerende publicerer resultatet af deres undersøgelse under overskriften: "Urtete styrker helbredet hos de ældre".

a) Kommentér denne påstand ved at stille mindst tre kritiske spørgsmål til undersøgelsen.

4.004 I Hite-rapporten, der omhandler amerikanske kvinders seksuelle adfærd, er en af konklusionerne: "Amerikanske kvinder er langt mere frigjorte, end hidtil antaget". Undersøgelsen kom til veje gennem udsendelse af spørgeskemaer til 100.000 kvinder, hvoraf 4.500 svarede.

a) Kommentér påstanden med brug af statistiske begreber som stikprøve, population, systematiske fejl og skjulte variable.

**STX083-MAA og STX083-MAB**

**Opgave 15** Et firma, der sælger kosttilskud via deres hjemmeside, oplyser, at 75% af de personer, der besøger hjemmesiden, indtager kosttilskud dagligt. På hjemmesiden opfordres de besøgende til at deltage i en undersøgelse vedrørende kosttilskuds betydning for helbredet. Det viser sig, at 80% af dem, der deltager i undersøgelsen, oplyser, at de mener, at indtagelse af kosttilskud har en positiv virkning på helbredet, mens de resterende 20% mener, at kosttilskud ikke har nogen positiv virkning på helbredet. I en pressemeddelelse hævder firmaet på grundlag af undersøgelsen, at 80% af befolkningen tydeligt mærker en positiv virkning af kosttilskud.

- a) Kommentér firmaets påstand ved brug af statistiske begreber som population, stikprøve, systematiske fejl og skjulte variable.

**STX071-MAB**

**Opgave 15** Forskere i Kræftens Bekæmpelse har fulgt 770 danske patienter, som fik konstateret tarmkræft i perioden 1985-1990, med henblik på at afdække, om det sociale netværk har betydning for patienternes evne til at overleve sygdommen. Undersøgelsen viser ifølge forskerne, at tarmkræftpatienter, der er blevet enlige på grund af skilsmisse eller dødsfald, før sygdommen bliver konstateret, har hele 40 % større risiko for at dø sammenlignet med patienter, der har en partner. Undersøgelsen præsenteres i en avis, hvor der fremsættes følgende påstand: "Det er en rigtig god idé at finde en at leve sammen med. Ægteskaber redder kræftpatienter".

- a) Kommentér denne påstand ved at stille 3 kritiske spørgsmål til påstanden eller undersøgelsen.

*Kilde: Urban, 10.01.07.*