

STATSVIDENSKABELIG KANDIDATEKSAMEN 1997 II

UDVALGT FAG

STIKPRØVETEORI

Alle hjælpemidler tilladt

Gallup har undersøgt, hvor mange danskere, der tager på kortvarige (under 36 timer) indkøbsture til Tyskland, samt hvor meget spiritus, de bringer med tilbage, hvilket egentlig er smugling. Analysen er gennemført telefonisk. Fra denne analyse betragtes i denne opgave kun data for respondenter med bopæl vest for Storebælt og i visse spørgsmål kun respondenter i Sønderjylland.

- 1) Antal indbyggere vest for Storebælt over 14 år er 2390000, hvoraf 1763 er udtaget til stikprøven. Af disse har 615 været på indkøbstur til Tyskland inden for det sidste år. Estimer andelen, der har været på indkøbstur til Tyskland indenfor det seneste år. Angiv 95% sikkerhedsgrænser for denne andel under antagelsen om, at stikprøven er udvalgt simpelt tilfældigt.
- 2) Af stikprøvens 1763 personer har 9 bragt mere end 4 liter spiritus med hjem på den seneste indkøbstur. Angiv 95% sikkerhedsgrænser for det antal i den samlede befolkning vest for Storebælt, der smugler i dette omfang.
- 3) Af geografiske grunde er det naturligt at betragte en stratifikation efter geografiske kriterier, som i tabel 1. Antag at de andele, tabellen angiver for at have været på indkøbstur, er korrekte. Angiv den proportionale og den optimale allokering af en stikprøve på 1763 personer.
- 4) Vurder, gerne uden beregning af varianserne, om der er væsentlige præcisionsgevinster ved stratifikationen med allokeringerne i 3).
- 5) De fleste, der tager på indkøbstur, svarer, at de ikke har hjembragt spiritus, og de, der ikke har været af sted, har naturligvis ikke hjembragt noget. Men nogle opgiver at have hjembragt større eller mindre mængder. I gennemsnit over alle 1763 personer i stikprøven, dvs. også de, der ikke har medbragt noget, er der medbragt 0.0698 liter fra den seneste tur og variansen er $s^2 = 0.4230^2$. Beregn 95% sikkerhedsgrænser for den gennemsnitlige mængde spiritus fra den seneste tur.

6) Andelen, der har hjembragt spiritus, afhænger meget af svarpersonens køn, idet den er 0.0697 for mænd og 0.0196 kvinder. Beregn hvor meget præcisionen kan forbedres ved at anvende vægtning efter køn. Betragt både det konkrete tilfælde, hvor 845 mænd og 918 kvinder har besvaret samt en planlægning af en kommende undersøgelse ligeledes af 1763 personer. I beregningerne forudsættes simpel tilfældig udvælgelse, og det antages, at der vest for Storebælt er lige mange af hvert køn.

7) I Sønderjylland foreligger der svar fra 24 personer, der har medbragt spiritus. Antag nu at der er taget en stikprøve på 24 personer fra populationen af sønderjyder, der har smuglet spiritus indenfor det sidste år. Den medbragte mængdes afhængighed af antal personer i svarpersonens husstand er vist i figur 1. Hvorfor er ratioestimation med antal personer i svarpersonens husstand ikke en god idé?

8) Korrelationskoefficienten for de 24 observationer i figur 1 er - 0.48. Angiv hvor meget præcisionen ville forbedres i forhold til estimation med simpel tilfældig udvælgelse ved anvendelse af regressionsmetoden, hvis man vel at mærke kendte den sande gennemsnitlige familiestørrelse for de personer, der smugler spiritus.

9) I undersøgelsen er der spurgt én person i hver husstand, idet undersøgelsen er gennemført telefonisk med valg af en tilfældig beboer i de udvalgte husstande. Antag nu, at alle i husstanden sammen foretager eventuelle indkøbsture til Tyskland, altså at alle i husstanden svarer ens. Den varians, der blev angivet i spørgsmål 5) som en varians mellem individer, er derfor egentlig variansen mellem husstandenes totale indkøb. I stikprøvens 1763 husstande vest for Storebælt bor i alt 3664 personer over 14 år, og tilsammen har de 1763 husstande i stikprøven hjembragt 123 liter spiritus på den seneste indkøbstur. Det samlede antal husstande vest for Storebælt er 1160000. Angiv de to "klynge"-estimer for den totale mængde hjembragt spiritus.

10) Beregn varianserne af estimatorerne i spørgsmål 9). Udover variansen på den medbragte mængde spiritus, der jf. spørgsmål 5) er 0.4230^2 , kan det benyttes at variansen på antal voksne pr. husstand er 0.85^2 og korrelationen mellem den hjembragte mængde spiritus og antal voksne i husstanden for hele området vest for Storebælt er -0.02.

11) Sammenhold resutaterne af spørgsmålene 9) og 10) med spørgsmål 5), der jo omhandler samme problemstilling under den egentlig forkerte antagelse om simpel tilfældig udvælgelse på personniveau. Giv en samlet konklussion.

Tabel 1. Andel, der har været på indkøbstur til Tyskland indenfor det sidste, år for fem delområder af Danmark vest for Storebælt.

	Andel	N_k
Fyns amt	0.41	390000
Nordjyllands amt	0.22	405000
Århus amt	0.29	515000
Vestjylland (Ringkøbing og Viborg amter)	0.27	410000
Sønderjylland (Sønderjyllands, Ribe og Vejle amter)	0.48	670000

Figur 1. Sammenhængen mellem husstandsstørrelse og hjembragt mængde spiritus. Enheden på anden akse er centiliter, altså hundrededele liter.