

Skriftlig tentamen i Undersökningsmetodik inom fortsättningskursen i statistik, måndagen den 23 maj 2005

- För studenter som påbörjat kursen HT04 eller tidigare är tentamen på 3 poäng. Uppgifternas poängantal gäller enbart för dessa studenter.
- För studenter som påbörjat kursen VT05 utgör tentamen en del av examinationsportföljen, som i sin helhet är på 5 poäng. "G" resp. "VG" vid en fråga betecknar om frågan anses kräva huvudsakligen faktakunskaper eller även en betydande mängd av egna insikter.
- De sistnämnda studenterna har möjlighet att lämna in en reflektion kring salskrivningen (vilken då utgör en del av examinationsportföljen); inlämning **senast kl. 10 den 26 maj 2005** genom institutionens expedition.

Skrivtid: 4 timmar

Hjälpmedel: miniräknare, bifogat formelblad.

Ange Ditt fullständiga personnummer och inte Ditt namn på varje inlämnat pappersark.

Kom ihåg att en motiverad lösning/lösning med angivna formella uttryck kan ge poäng även om beräkningarna inte är helt korrekta.

Lös vänligen inte mer än en uppgift per inlämnat pappersark.

Skrivningsgenomgång: tisdagen den 7 juni 2005, kl. 10 i sal B705.

Uppgift 1

I. Vad ska vi tänka på, då vi formulerar frågor för en enkät eller intervju, för att göra frågorna lätta att förstå. (G/5p)

II. Diskutera anledningar till - och följderna av - att ställa demografiska- och bakgrundsfrågor först respektive sist i en frågeblankett. (VG/5p)

Uppgift 2

I. Ange kvotestimatoren i någon av dess former, och diskutera estimatorns för- och nackdelar. (G/5p)

II. Visa att

$$MSE(\hat{\theta}) = V(\hat{\theta}) + Bias^2(\hat{\theta}). \quad (VG/5p)$$

Uppgift 3

Vi har framför oss en populationsram som vi finner lämplig att dela in i 3 strata. Situationen kan vi sammanställa på följande vis (där vi bl a har anledningar att tro att vi känner till undersökningvariabelns varians i de olika strata):

Stratum	I	II	III
N_ℓ	1500	500	1000
σ_ℓ^2	9	25	16

I. Om vi ämnar dra ett stickprov av storleken $n = 100$, hur ska vi fördela denna stickprovsstorlek på strata om vi tar hänsyn till informationen i ovanstående tabell? (G/5p)

II. Förutsatt att: vi ämnar skatta en populationstotal, att det är ett OSU och π -estimatoren vi använder inom varje stratum, samt att våra förhandsgissningar om variansen i de olika strata. har varit helt överensstämmande med verkligheten, vad blir standardfel (SE), dvs

$$SE(\hat{\tau}_{\pi,STOSU}) = 1.96 \times \sqrt{\hat{V}(\hat{\tau}_{\pi,STOSU})}$$

av vår punktskattning i denna undersökning (VG/5p).

Uppgift 4

En population av korvar består av två grupper, grupp 1 och grupp 2. I grupp nummer 1 finns det 1000 korvar, i grupp nummer 2 finns det 10 korvar. Vi har för avsikt att skatta medelvärde för vår undersökningsvariabel *antal korvänder på en korv*.

Vi använder oss av OSU av grupper, med $n = 1$. I den valda gruppen antecknar vi totalen för vår undersökningsvariabel; vi utgår ifrån att vi redan känner till det totala antalet element i varje grupp i populationen (vilket angivits ovan).

I. Med förutsättning att grupp 1 valdes in i stickprovet:

(a) skatta medelvärdet med kvotestimatoren (vilket i allmänhet rekommenderas för föreliggande situation),

(b) skatta medelvärdet med π -estimatoren (vilket i allmänhet *inte* rekommenderas för föreliggande situation).

II. Nu med förutsättning att grupp 2 valdes in i stickprovet, upprepa (a) och (b) ovan.

III. Visa numeriskt, m hj a data från I och II, att både kvot- och π -estimatoren är väntevärdesriktiga i den här situationen¹. (VG/10p)

Uppgift 5

I. En skola har ett antal gymnasieelever som tyvärr är okänt (inget register finns). Eleverna går dessutom i klasser på ett flexibelt sätt vars sammansättning anpassa ofta. Varje elev tillhör en och endast en klass; det finns just för tillfället 34 klasser. På grund av flexibiliteten vet man inte heller hur stora klasserna är.

För att skatta totalt antal timmar som gymnasieeleverna går på per vecka väljer man då ut med hjälp av systematiskt urval 8 klasser ur en slumpmässigt ordnad populationsram innehållande de 34 klasserna. För varje utvald klass antecknar man hur många gymnasieelever går i den klassen samt hur många timmar sammanlagt dessa elever har deltagit i undervisningen under referensveckan. Därpå kunde man konstatera följande:

¹Denna uppgift är en förenklad version av en övningsuppgift i Särndal, Swenson och Wretman (1992), som i sin tur sägs komma från Tore Dalenius, professor vid institutionen under en större del av förra seklets andra halvan.

Klass	Antal elever	Summa besökta undervisningstimmar
1	31	995
2	14	380
3	20	600
4	16	580
5	42	1280
6	29	770
7	27	807
8	13	455

Skatta totalen för besökta undervisningstimmar i gymnasiet under referensveckan i föreliggande situationen och ge tillhörande 95%-igt konfidensintervall. (G/5p)

II. Diskutera automatisering av kodnig, granskning och andra bearbetningsprocesser. (VG/5p)