

## 732G26

## SURVEYMETODIK MED UPPSATS

### TENTAMEN

### TISDAGEN 8 JUNI 2010, KL. 08.00–12.00

*Hjälpmedel:*

Läroboken "Lohr, S: Sampling-Design and analysis. 1:a el. 2:a uppl. Brooks/Cole"  
(anteckningar får finnas). Räknedosa.

*Jourhavande lärare:*

Anders Nordgaard (endast via telefon)

*Poänggränser:*

Skrivningen ger maximalt 20 poäng. För betyget Godkänd krävs normalt 12 poäng.

*Lycka till!*

1. Antag att vi har gjort ett OSU om 200 individer från en population bestående av 55560 individer. Olika frågor ställs till individerna och bl.a. frågas om det totala antalet timmar man sett på TV den senaste veckan. För hela urvalet blir totalsumman av antalet timmar 5540 och standardavvikelsen blir 15 timmar.
  - a) Beräkna en punktskattning och ett 99% konfidensintervall för det *sammanlagda antalet* timmar,  $t$ , alla individer i populationen har sett på TV den senaste veckan. (2p)
  - b) Utgående från informationen i urvalet, hur stort urval borde göras om bredden hos ett 99% konfidensintervall för  $t$  skall bli högst 280 000 (timmar). (2p)
  - c) Antag nu att 10 av de 200 individerna i urvalet visar sig inte tillhöra målpopulationen. För dessa 10 individer är medelvärde 28.5 timmar och standardavvikelsen är 10.9 timmar. Ta hänsyn till detta och beräkna en ny punktskattning och ett nytt 99% konfidensintervall för  $t$ . (*Ledning:* De individer i urvalet som faktiskt tillhör populationen kan betraktas som en redovisningsgrupp (domain of study).) (1p)
2. Betrakta undersökningen i uppgift 1 igen. Någon föreslår att den aktuella urvalsdesignen inte passar för att mäta TV-tittandet. I en parallell undersökning har ett stratifierat urval gjorts om 150 individer bland de i populationen som har en anställning och 50 individer bland de som saknar anställning. De 150 har totalt tittat 3350 timmar på TV den senaste veckan och standardavvikelsen är 10.3 timmar. De 50 har totalt tittat 1800 timmar och standardavvikelsen är 9.5 timmar. I populationen är det c:a 7.1% av individerna som saknar anställning.
  - a) Beräkna utgående från dessa data en punktskattning och ett 99% konfidensintervall för  $t$ , dvs. det sammanlagda antalet timmar alla individer i populationen har sett på TV den senaste veckan. (3p)
  - b) I såväl den parallella undersökningen som i den som analyserades i uppgift 1 har totalt 200 individer valts ut. Vad blir en optimal allokering av 200 individer över de två strata "individer med anställning" och "individer utan anställning"? (2p)

3. Antag att det i en stor stad finns totalt 55 skolor med elever i åldrarna 13–16 år. I varje skola är eleverna indelade i klasser utifrån vilket skolår de tillhör och utifrån vilken skola de gick i innan de kom till denna skola. Om vi specifikt undersöker elever som tillhör skolår 8 så är det totala antalet elever i alla 55 skolorna 5954 stycken. Vidare vet man att det totala antalet klasser i de 55 skolorna (alla skolår inräknade) är 631.

Ett OSU om 3 skolor valdes och i varje skola gjordes ett OSU om 10% av alla elever i skolår 8. I varje sådant urval tillfrågades eleverna om de hade druckit alkohol den senaste månaden. Resultatet blev

<i>Skola</i>	<i>Totalt antal klasser (alla skolår)</i>	<i>Antal elever i skolår 8</i>	<i>Antal elever i urvalet som druckit alkohol den senaste månaden</i>
1	11	92	2
2	9	83	1
3	13	149	3

- a) Beräkna en väntevärdesriktig punktskattning och ett 95% konfidensintervall för andelen elever,  $p$ , i studieår 8 (i de 55 skolorna) som druckit alkohol den senaste månaden. (3p)
- b) Antag hypotetiskt att undersökningen skulle ha ett bortfall om c:a 50% i var och en av skolorna, men att vi ändå skulle ha fått de svar vi fick i tabellen ovan. Om man ignorerar bortfallet, vilken punktskattning och vilket konfidensintervall motsvarande det i a)–uppgiften får man? Kommentera resultatet. (Avrunda 50% av respektive urval till närmaste heltal.) (3p)
4. Ett OSU om 700 individer gjordes från ett nyligen uppdaterat register av medlemmar i ett politiskt parti omfattande totalt 381 960 individer. Till individerna i urvalet ställdes en fråga om hur mycket pengar de hade spenderat på sin semester närmast föregående år. Vidare finns i registret en 5 år gammal uppgift om taxerad inkomst för varje individ. Man vill använda denna information (även om den inte är uppdaterad) för att uppskatta det genomsnittliga beloppet en medlem spenderar på semestern. Följande har beräknats från de svar man fick (inget bortfall förekom osannolikt nog) och från information i registret:

$$\begin{aligned}\sum(\text{belopp spenderat på semestern}) &= 3655.89 (\times \text{SEK } 1000) \\ \sum(\text{belopp spenderat på semestern})^2 &= 23881.3 (\times (\text{SEK } 1000)^2) \\ \sum(\text{årlig taxerad inkomst}) &= 248721 (\times \text{SEK } 1000) \\ \sum(\text{årlig taxerad inkomst})^2 &= 92032277 (\times (\text{SEK } 1000)^2) \\ \sum(\text{belopp spenderat på semestern}) \times (\text{årlig taxerad inkomst}) &= 1400947 (\times (\text{SEK } 1000)^2)\end{aligned}$$

Totalsumman av taxerade inkomster i hela registret är ungefär 135730000 ( $\times \text{SEK } 1000$ )

- a) Beräkna en kvotskattning tillsammans med ett 95% konfidensintervall för det genomsnittliga belopp en medlem spenderade på sin semester föregående år. (2p)
- b) Beräkna en regressionsskattning tillsammans med ett 95% konfidensintervall för det genomsnittliga belopp en medlem spenderade på sin semester föregående år. (2p)