

Lösningförslag till tentamen

Institutionen för datavetenskap, STIMA



| | |
|----------------------|--|
| Kurskod och namn: | 732G38 , Surveymetodik med teori |
| Datum och tid: | 2018-06-05, 8-12 |
| Jourhavande lärare: | Annika Tillander |
| Tillåtna hjälpmedel: | Miniräknare av valfri modell Kursbok Sharon L. Lohr SAMPLING: Design and Analysis (alla upplagor tillåtna - får innehålla markeringar, understrykningar och flärpar, men inte anteckningar) A4-blad med vad som helst skrivet på fram- och baksida (för hand/skrivet på dator/kopierat etc. - inga begränsningar) |
| Betygsgränser: | Tentamen omfattar totalt 20 poäng, G från 12p, VG från 16p |
| Övrigt: | Siffrorna i uppgifterna är påhittade. Saknas någon siffra för att kunna lösa uppgiften? Skriv då tydligt ut att du saknar denna information, anta ett godtyckligt värde och lös uppgiften med detta antagande. Lösningförslag till denna tentamen läggs upp på kursens LISAM-sida den 7/6. |

Redovisa, tolka och motivera tydligt alla dina lösningar!

Lycka till!

Uppgift 1 (6p)

Linköping stad med 6900 barn i förskola önskar skatta genomsnittligt antal sjukdagar per månad för förskolebarn. Eftersom antal sjukdagar (y) varierar med ålder görs ett stratifierat urval baserat på åldersgrupper.

| Strata | N_h | n_h | \bar{y}_h | s_h |
|------------|-------|-------|-------------|-------|
| < 2 år | 978 | 200 | 14.38 | 5.51 |
| 2 – 3 år | 1251 | 200 | 8.48 | 3.64 |
| > 3 – 4 år | 1374 | 200 | 4.01 | 1.91 |
| > 4 år | 3297 | 200 | 3.37 | 0.61 |
| Samtliga | 6900 | 800 | 8.23 | 4.27 |

Tabell 1: Genomsnittligt antal sjukdagar per månad och stratum.

- Skatta genomsnittligt antal sjukdagar per månad för förskolebarn med 95% konfidensintervall baserat på det stratifierade urvalet. 2p
- Vilken typ av allokering har använts och beräkna två alternativa typer av allokering? 2p
- Om skattningen istället skulle baseras på OSU, hur stort urval skulle krävas för att uppnå samma precision (e) som i a)? 2p

Uppgift 2 (8p)

Genomsnittlig sockerhalt per liter jordgubbar ska skattas där medelvikt per liter jordgubbar är 500.9 gram beräknat för totalt 20 liter. Bland dessa 20 liter gjordes ett slumpmässigt urval på 4 liter och sockerhalten mättes i gram, se tabell 1.

| | vikt | socker |
|---|-------|--------|
| 1 | 472.9 | 18.5 |
| 2 | 534.1 | 32.5 |
| 3 | 519.9 | 27.6 |
| 4 | 446.4 | 16.4 |

Tabell 2: Vikt i gram och sockerhalt i gram per liter jordgubbar

- Skatta den genomsnittliga sockerhalten per liter i urvalet med 95% konfidensintervall baserat på OSU. 2p
- Skatta den genomsnittliga sockerhalten per liter med 95% konfidensintervall med kvotskattning. 2p
- Skatta den genomsnittliga sockerhalten per liter med 95% konfidensintervall med regressionskattning givet modellen: 2p

| | Estimate | Std. Error | t value | $\Pr(> t)$ |
|-------------|----------|------------|---------|-------------|
| (Intercept) | -29.85 | 9.51 | -3.1 | 0.01 |
| vikt | 0.11 | 0.02 | 5.7 | 0.00 |

- Beräkna designeffekten för kvotskattning och regressionskattning. Den sanna genomsnittliga sockerhalten är 24.1 gram, utifrån detta och designeffekterna vilken skattning är att föredra? 2p

Uppgift 3 (4p)

För att undersöka studenters inställning till elektronisk kurslitteratur ska en studie genomföras vid Linköpings universitet med 18120 studenter vid Campus Valla och 4880 vid Campus Norrköping. Urvalet önskas uppgå till totalt 1394 och ett stratifierat urval gjordes med 906 studenter från Campus Valla och 488 från Campus Norrköping.

- a) Tidigare undersökningar har visat ett bortfall på ca. 74% för studenter Campus Valla och ca. 61% för studenter Campus Norrköping. Om urvalet skulle gjorts med oversampling för få in 1394 svar, hur stort urval borde gjorts i respektive strata? 1p
- b) Undersökningen genomförs och beräkna studiens oviktade och viktade bortfall givet följande resultat 2p

| Resultat | Valla | Norrköping | Totalt |
|------------------|-------|------------|--------|
| Svar (S) | 371 | 327 | 698 |
| Bortfall (B) | 371 | 117 | 488 |
| Okänd status (O) | 164 | 44 | 208 |

- c) Av de 698 svarande var 52 stycken 35 år eller äldre. Av dessa var det endast 15 stycken som var positivt inställda till elektronisk kurslitteratur. Skatta totala antalet positivt inställda för gruppen 35 år eller äldre. 1p

Uppgift 4 (2p)

En skola med 15 klasser är intresserade av totala tiden eleverna lägger ned på läsläsning per vecka (y) för att eventuellt kunna erbjuda läxhjälp. Först görs ett slumpmässigt urval på 3 klasser och sedan ytterligare ett slumpmässigt urval av elever. Skatta totala tiden läsläsning per vecka med 95% konfidensintervall 2p

| Klass | M_i | m_i | \bar{y}_i | s_i^2 |
|-------|-------|-------|-------------|---------|
| A | 25 | 14 | 4.36 | 8.09 |
| B | 18 | 9 | 6.78 | 39.19 |
| C | 23 | 16 | 4.88 | 0.78 |

Tabell 3: Läsläsning i timmar per vecka i urvalet