

Tentamen 732G38

Linköpings universitet
Institutionen för datavetenskap, IDA
Avdelningen för Statistik och maskininläring, STIMA

Kurskod och namn:	732G38, Surveymetodik med teori
Datum och tid:	2019-06-03, 14-18
Jourhavande lärare:	Isak Hietala
Tillåtna hjälpmedel:	- Räknedosa av valfri modell; - kursboken <i>Sampling: Design and Analysis</i> av Lohr som får innehålla markeringar, överstrykningar samt små flärpar med anteckningar; - ett dubbelsidigt A4 med egna anteckningar
Betygsgränser:	Tentamen omfattar totalt 20 poäng, betyg G fås vid minst 12 poäng; betyg VG vid minst 16 poäng
Annan information:	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.

Redovisa, tolka och motivera tydligt alla dina lösningar!

Uppgift 1 (4p)

Bortfall har en stor påverkan på alla undersökningar som genomförs och kan hanteras på olika sätt.

- Motivera vilken av de tre bortfallsantaganden som gäller för följande tillfällen: (0.5p per deluppgift)
 - I en undersökning av boendesituationen i en kommun svarar inte personer som inte kan språket.
 - I en på-stan undersökning deltar inte personer som inte känner att de har tid.
 - I en partisympatiundersökning deltar inte personer som inte kommer rösta i nästa val.
 - I en pop-up-webbundersökning om nöjdhet deltar inte personer som har en s.k. Adblocker som förhindrar pop-up rutor från hemsidan.
- Nämn ett sätt att hantera det bortfall som uppkommit för vardera av de tre antaganden. (1.5p)
- I ett OSU har det visat sig att det oviktade bortfallsandelen är 0.325. Vad är den viktade bortfallsandelen? (0.5p)

Uppgift 2 (3.5p)

På en viss campingplats vill man undersöka hur nöjda camparna är med storleken på platsen man kan hyra. Därför valde man slumpmässigt 50 campare bland de första 250 campinggästerna som kom till campingplatsen. Bland dessa 50 campinggäster var 38 nöjda med storleken på platsen.

- Beräkna ett 95-procentigt konfidensintervall för andelen nöjda bland de 250 campinggästerna. (2p)
- Om man vid ett senare tillfälle vill göra om undersökningen, hur stort urval bör man då dra för att få en felmarginal på högst 5 procent? Använd data från a)-uppgiften. (1.5p)

Uppgift 3 (4p)

Ett företag har liknande kontor på 24 olika platser runtom i landet. Ledningen är intresserade av att undersöka antalet timmar per vecka som spenderas på administration för vardera anställd. Information från 6 slumpmässigt utvalda kontor sammanställs i tabellen nedan:

Tabell 1: Antalet timmar per vecka som spenderats på administration

	M_i	m_i	\bar{y}_i	s_i^2
1	10	10	4.98	25.10
2	22	22	8.01	4.29
3	21	21	4.33	15.50
4	76	76	6.03	21.16
5	59	59	7.88	8.68
6	73	73	4.36	16.31

- Skatta den totala tiden som spenderas på administration i alla företag med ett 95-procentigt konfidensintervall. (2p)
- Beräkna punktskattningen \hat{y}_U . (1p)
- Beskriv två skillnader mellan urvalsmetoden som användes i denna uppgift och ett stratifierat urval. (1p)

Uppgift 4 (4.5p)

En viss dag finns det på marknadsplatsen Blocket 102148 bilar till salu. 72873 av dessa bilar säljs av privatpersoner och resten av bilfirmor. Slumpmässigt väljer man 2500 bilar bland privatpersonerna och 2500 bland bilfirmorna, och följande beräknas:

Tabell 2: Beskrivande statistik över urvalets försäljningar

	\bar{x}	s^2
Privatperson	99200	23920
Bilfirma	105200	29460

- Beräkna ett 95-procentigt konfidensintervall för snittpriset bland de bilar som är till salu på Blocket den aktuella dagen. (2.5p)
- Vad heter allokeringmetoden som har använts? (0.5p)
- Beräkna nya urvalsstorlekar om Neymann-allokering skulle ha använts. (1.5p)

Uppgift 5 (4p)

Under en sommar mäter en försäljare slumpmässigt 8 dagars medeltemperatur och antalet glassar som säljs. Den genomsnittliga temperaturen för hela sommaren (bestående av 90 dagar) är 22.4 grader.

Tabell 3: Beskrivande statistik över urvalets försäljningar

Temperatur	Antal sålda glassar
26.8	403
20.7	417
31.8	452
21.4	382
25.3	458
29.2	461
27.0	445
17.1	396

- Beräkna det genomsnittliga antalet sålda glassar per dag med ett 95-procentigt konfidensintervall baserat på kvotskattning. (2p)
- Beräkna designeffekten för denna kvotskattning. Är denna metod att föredra framför ett OSU? Motivera. (2p)