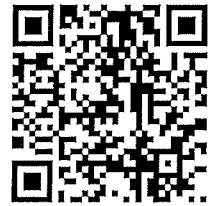


Försättsblad till skriftlig tentamen vid Linköpings universitet



Datum för tentamen	2019-08-16
Sal (2)	TER1(21) TERE(1)
Tid	8-12
Utb. kod	732G38
Modul	TENA
Utb. kodnamn/benämning Modulnamn/benämning	Surveymetodik med teori Tentamen
Institution	IDA
Antal uppgifter som ingår i tentamen	4
Jour/Kursansvarig Ange vem som besöker salen	Isak Hietala
Telefon under skrivtiden	013-281970
Besöker salen ca klockan	Endast telefonjour
Kursadministratör/kontaktperson (namn + tfnr + mailaddress)	Erika Larsson 013-28 18 68 erika.larsson@liu.se
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)
Övrigt	
Antal exemplar i påsen	

Försättsblad till skriftlig tentamen vid Linköpings universitet



Datum för tentamen	2019-08-16
Sal (2)	TER1(21) <u>TERE(1)</u>
Tid	8-12
Utb. kod	732G38
Modul	TENA
Utb. kodnamn/benämning Modulnamn/benämning	Surveymetodik med teori Tentamen
Institution	IDA
Antal uppgifter som ingår i tentamen	4
Jour/Kursansvarig Ange vem som besöker salen	Isak Hietala
Telefon under skrivtiden	013-281970
Besöker salen ca klockan	Endast telefonjour
Kursadministratör/kontaktperson (namn + tfnr + mailaddress)	Erika Larsson 013-28 18 68 erika.larsson@liu.se
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)
Övrigt	
Antal exemplar i påsen	

Tentamen 732G38

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

Kurskod och namn:	732G38, Surveymetodik med teori
Datum och tid:	2019-08-16, 8-12
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
Antalet uppgifter:	4 stycken
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.

- a) Motivera vilken av de tre bortfallsantaganden som gäller för följande tillfällen: (0.5p per deluppgift)
- I en undersökning om vikt av personer som deltar i en studie tar batterierna slut i vågen och vikterna kan inte registreras.
 - En kundundersökning har svårt att få in svar från äldre kunder som inte aktivt använder sin mail-adress.
 - Personer som tagit droger dagen innan ett drogtest dyker inte upp för sin undersökning.

	Yngre	Äldre
Svar (S)	316	407
Bortfall (B)	249	235
Okänd status (O)	201	114
Övertäckning (\emptyset)	80	22

Tabell 1: Resultat från en studie.

- b) Beräkna den viktade samt oviktade bortfallsandelen av studien i tabell 1. Vi vet att populationsstorleken av de yngre var 19458 och de äldre var 12448.

Uppgift 2 (8p)

Storleken av åkrar i olika områden av USA mäts vart femte år. Tabell 2 är ett urval av 8 områden och dess areal från två olika mätperioder. Medelarealen av alla 3044 områden från ACRES82 är 323.536.

Tabell 2: Åkerareal från två mätperioder (angivna i tusental).

ACRES87	ACRES82
732	774
280	293
158	179
87	95
186	195
1439	1446
377	379
11	11

- Skatta den genomsnittliga åkerarealen år 1987 med ett 95-procentigt konfidensintervall baserat på OSU. (2p)
- Skatta den genomsnittliga åkerarealen år 1987 med ett 95-procentigt konfidensintervall baserat på kvotskattning. (2p)
- Beräkna designeffekten för denna kvotskattning. Är denna metod att föredra framför ett OSU? Motivera. (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 100? Använd data från a)-uppgiften. (2p)

Uppgift 3 (4p)

En skola med 12 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 4 klasser och sedan tillfrågas alla elever hur mycket tid de spenderar på läxor i veckan.

Tabell 3: Antalet timmar per vecka som spenderats på läxor

	M_i	m_i	\bar{y}_i	s_i^2
A	36	36	9.79	17.47
B	45	45	4.22	24.77
C	47	47	5.06	21.90
D	29	29	9.32	31.02

- Skatta totala tiden läsläsning per vecka med ett 95-procentigt konfidensintervall. (2p)
- Beskriv två skillnader mellan en- och två-stegs klusterurval. (2p)

Uppgift 4 (6p)

I ett företag vill man undersöka antalet sjukdagar per år bland administrativ personal och chefer. Slumpmässigt väljer man 25 personer för båda grupperna, och följande beräknas:

Tabell 4: Beskrivande statistik över urvalets sjukdagar

	N_h	n_h	\bar{y}	s^2
Administrativ	208	25	30.9	8.1
Chefer	199	25	37.4	9.1

- Beräkna ett 95-procentigt konfidensintervall för det genomsnittliga antalet sjukdagar som företaget har. (3p)
- Beräkna nya urvalsstorlekar om Neymann samt optimal allokering skulle ha använts. Antag att kostnaderna för grupperna är 1.4, 0.8. (3p)