

Tentamen i Surveymetodik

Skrivtid:	8.00-12.00
Hjälpmedel:	Miniräknare, boken "Sampling: Design and Analysis" (1999) av Sharon L. Lohr och häftet "Standard för bortfallsberäkning" (2005-01-22) utgivet av Sektionen för Surveystatistik och Svenska Statistikersamfundet. Både bok och häfte ska vara utan anteckningar, men färgöverstrykningar, understrykningar och flikar är tillåtna.
Jourhavande lärare:	Annica Isaksson.
Betygsgränser:	Tentamen omfattar totalt 18 poäng. Godkänt från 10 poäng, väl godkänt från 14 poäng.

Redovisa och motivera kort alla dina lösningar. Skriv namn på varje papper du lämnar in.

Uppgift 1 (7 poäng)

Familjen Eriksson vill flytta från landet in till staden, och letar nu hus i Sveriges tre storstadsregioner; Stor-Stockholm, Stor-Malmö och Stor-Göteborg. Under en tremånadersperiod bevakar familjen bostadsmarknaden i dessa regioner. Av tabellen nedan framgår hur många hus som såldes. Där finns även, för varje region, familjens observationer av försäljningspriser (i 1000-tals kronor) på nio slumpmässigt valda hus.

Region	Antal sålda	1	2	3	4	5	6	7	8	9
Stor-Stockholm	2091	3400	2900	3100	4100	3700	1800	3400	5600	3300
Stor-Göteborg	1257	2400	2700	2800	1900	2700	2900	3100	2200	4400
Stor-Malmö	827	2300	4200	3000	2100	2600	2700	2800	2500	3200

Kanske har du hjälp av följande uppgifter (beräknade utifrån data i tabellen) för stickprovet från varje region: summan av försäljningspriserna, betecknad $\sum y_i$, och summan av de kvadrerade försäljningspriserna, betecknad $\sum y_i^2$.

Stor-Stockholm: $\sum y_i = 31300$, $\sum y_i^2 = 117130000$

Stor-Göteborg: $\sum y_i = 25100$, $\sum y_i^2 = 74010000$

Stor-Malmö: $\sum y_i = 25400$, $\sum y_i^2 = 74720000$

Betrakta endast den utvalda tidsperioden.

- Beräkna ett 95-procentigt konfidensintervall för den totala försäljningssumman för de hus som såldes i storstadsregionerna.
- Beräkna ett 95-procentigt konfidensintervall för andelen av de hus som såldes i storstadsregionerna som såldes för över tre miljoner kronor.

En journalist får syn på familjens lilla undersökning och vill göra om den i större skala. Framförallt är hon intresserad av genomsnittligt försäljningspris på de hus som såldes i storstadsregionerna, och hon vill skatta detta med en felmarginal på högst 100 000 kronor.

- c) Bestäm hur stort urval journalisten ska göra, och hur hon ska fördela det på ett optimalt sätt över de tre storstadsregionerna. Ta hjälp av data från familjen Erikssons undersökning.

Uppgift 2 (6 poäng)

Hästar utfodras i allmänhet med en kombination av hö eller hösilage, kraftfoder, saftfoder samt eventuellt extra vitaminer eller mineraler. Kraftfoder kan utgöras av exempelvis havre, korn eller vetekli. Det är mera energirikt än hö, och huvudskälet till att ge en häst kraftfoder är att den ska hålla (eller öka) sin vikt. Antag att ett företag har utvecklat ett nytt kraftfoder speciellt avpassat för travhästar, och att man vill studera hur det påverkar hästarnas vikt. I samarbete med ett antal travstall låter man 350 travhästar delta i en studie där man ersätter deras vanliga kraftfoder med rekommenderad dos av det nya fodret. I samband med att studien startar väger man alla deltagande hästar. Efter en månad gör man en mindre avstämning där man kontrollväger tio slumpmässigt valda hästar. Dessa hästars vikt efter en månad, y , samt deras utgångsvikt, x , återges i tabellen nedan (i 100-tals kilo). Viktförändringen mellan de två tidpunkterna, $e=y-x$, anges också. Den samlade utgångsvikten för alla hästar som ingick i studien var 180 ton.

Häst nr	1	2	3	4	5	6	7	8	9	10
y_i	5,5	6,4	4,6	5,1	5,0	5,0	5,1	4,9	5,7	5,8
x_i	5,3	6,1	4,5	5,1	4,9	4,8	5,0	4,7	5,5	5,6
e_i	0,2	0,3	0,1	0	0,1	0,2	0,1	0,2	0,2	0,2

Kanske har du hjälp av följande uppgifter, beräknade utifrån tabellen: $\sum y_i = 53,1$, $\sum y_i^2 = 371,91$, $\sum x_i = 51,5$, $\sum x_i^2 = 267,31$, $\sum y_i x_i = 275,76$, $\sum e_i = 1,6$ och $\sum e_i^2 = 0,32$.

- a) Beräkna en differensskattning av medelvikten hos alla hästar i studien efter en månad. Beräkna även ett 95-procentigt konfidensintervall.
- b) Beräkna en kvotskattning av medelvikten hos alla hästar i studien efter en månad. Beräkna även ett 95-procentigt konfidensintervall.

Efter sex månader gör företaget en lite större avstämning. Denna gång drar man ett OSU av tre av de totalt 12 travstall som medverkar i studien. Bland hästarna i dessa stall gör man ett urval (återigen med OSU) av hästar som kontrollvägs. Av tabellen nedan framgår, för varje utvalt stall, hur många hästar som finns i stallet (och härmed ingår i studien), hur många av dessa som valts ut till sexmånadersavstämningen, samt

medelvärde och varians för utvalda hästars vikt efter sex månader med det nya fodret (i 100-tals kilo).

Stall	Totalt antal	Antal utvalda	Medelvärde	Varians
1	10	5	5,2	205,0
2	110	10	5,3	150,3
3	40	10	5,4	250,4

- c) Punktskatta medelvikten hos alla hästar i studien efter sex månader. Beräkna även ett 95-procentigt konfidensintervall.

Uppgift 3 (3 poäng)

I en urvalsundersökning har man, för ett urval av åtta individer, observerat värden på variablerna y och x i enlighet med tabellen nedan. Totalt innehåller populationen 300 individer. Populationstotalen för x är 3600.

Individ nr	1	2	3	4	5	6	7	8
y_i	10	12	15	5	6	12	18	18
x_i	8	9	11	4	2	9	16	17

- a) Antag att urvalet gjordes utan återläggning, och att sannolikheterna för individerna att inkluderas i urvalet var följande:

Individ 1, 3 och 5: 0,03

Individ 2, 4 och 6: 0,05

Individ 7 och 8: 0,10

Beräkna två olika punktskattningar av populationstotalen för y : en kvotskattning och en väntevärdesriktig skattning. (Glöm inte att ange vilken som är vilken.)

- b) Vilken av skattningarna i a-uppgiften skulle du välja? Motivera ditt svar.

Uppgift 4 (2 poäng)

- a) I en undersökning av hushållens konsumtion visar det sig att en stor del av de undersökta hushållen kraftigt har underrapporterat sina inköp av en viss grupp av varor. Är detta ett exempel på täckningsfel, mätfel eller partiellt bortfall?
- b) Vad ser du för huvudsakliga problem med bortfall i en undersökning?