

Tentamen i Surveymetodik med uppsats. 2011-03-29

Skrivtid: kl: 8-12

Hjälpmedel: Kursboken: Lohr, 'Sampling: Design and Analysis som ej får innehålla anteckningar men får ha över/understrykningar och flärpar med anteckning, Räknedosa

Jourhavande lärare: Lotta Hallberg

Redovisa och motivera kort alla dina lösningar

1

En minipopulation består av 6 personer som alla får betygsätta kvaliteten på automatkaffet som serveras gratis på deras jobb. Betyg 1 för uselt kaffe och betyg 5 för utmärkt kaffe.

De 6 personerna har följande betyg:

Person nr	1	2	3	4	5	6
Betyg	5	2	4	4	4	2

- Beräkna populationsmedelvärdet och populationsvariansen. 1p
- Dra ett OSU om tre personer och skatta populationsmedelvärdet och beräkna urvalsvariansen. 1.5p
- Totala antalet OSU utan återläggning man kan dra är 20 urval. Hur beräknas detta antal? 0.5p

Här ges urvalsmedelvärdet för de 20 urvalen dragna med OSU uå.

3.667 3.667 3.667 3.000 4.333 4.333 3.667 4.333 3.667 3.667 3.333
3.333 2.667 3.333 2.667 2.667 4.000 3.333 3.333 3.333

- Ge samplingfördelningen för urvalsmedelvärdet. 1p

Vidare ges urvalsvarianserna för de 20 urvalen

2.333 2.333 2.333 3.000 0.333 0.333 2.333 0.333 2.333 2.333
1.333 1.333 1.333 1.333 1.333 1.333 0.000 1.333 1.333 1.333

- Visa att urvalsmedelvärdet är en väntevärdesriktig skattning av populationsmedelvärdet samt visa att urvalsvariansen är en väntevärdesriktig skattning av populationsvariansen. 2p

2

En undersökning ska planeras. Man vill då skatta andelen invånare p som ställer sig positiva till ett namnbyte av en galleria. Om man vill att felmarginalen i ett 95% konfidensintervall av p inte får överstiga 5%. Hur stort urval ska man då dra vid ett OSU utan återläggning? Totala antalet invånare är 25 000. 2p

3

En skogsägare vill bilda sig en uppfattning om den skada en storm har orsakat på beståndet. Skogarna är sedan tidigare indelade i till arealen lika stora mindre områden som är tydligt utmarkerade. Hon väljer slumpmässigt ut 5 områden och beräknar den totala skadan i tkr (tusentals kr) i varje område. Dessutom tas ett flygfoto och andelen stormfälld skog kan uppskattas, både totalt och i de 5 delområdena. Låt y vara skadan i tkr och låt x vara andelen stormfälld skog. I tabellen nedan visas (det fiktiva) resultatet.

OMRÅDE NR	1	2	3	4	5
y i tkr	104	120	102	144	142
x i %	33	55	28	75	80

Totala antalet delområden är 100 stycken. Andelen stormfälld skog uppskattades från flygfotot till 62%.

Skatta den totala skadan på skogarna som drabbat skogsägaren med ett 95% konfidensintervall. Använd kvotskattning. 3p

4

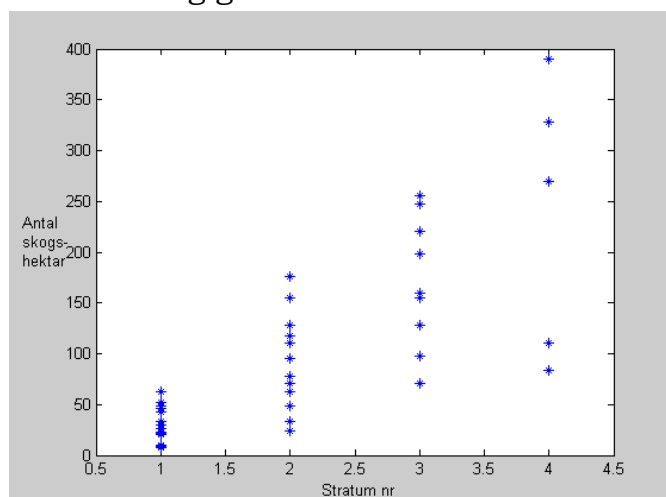
Förklara hur regressionimputation går till vid partiellt bortfall. 1p

5

En jägmästare vill uppskatta antalet skogshektar i ett område. Eftersom antalet skogshektar varierar väldigt med gårdsstorlek så stratifieras populationen av gårdar för dess storlek. Ett stratifierat urval om 40 gårdar görs med proportionell allokering. Resultatet visas i nedanstående tabell. Låt y vara antalet skogshektar. Kostnad c för urvalsdragning är också given i tabellen.

	Stratum 1 1-100 hektar	Stratum 2 100-200 hektar	Stratum 3 200-300 hektar	Stratum 4 >300 hektar
N_i	86	72	52	30
n_i	14	12	9	5
\bar{y}_i	30.96	91.50	170.28	236.20
s_i	17.32	47.58	64.80	134.51
C_i	c	$1.10c$	$1.20c$	$1.40c$

- Skatta det totala antalet skogshektar på gårdarna i området. Beräkna även ett 95% konfidensintervall. 3p
- Om vi på nytt skulle dra ett stratifierat urval om 40 gårdar hur skulle vi då allokera med optimal allokering. (Använd s^2 ur tabellen). 2p
- Förklara varför optimal allokering är att föredra framför proportionell allokering genom att studera tabellen och figuren nedan. 1p



6

Vid en sälinventering på Svalbard var man bl.a. intresserad av att skatta hur många andningshål det fanns i isen. Man delade in området i tre strata, där man trodde sig kunna anta att stratum 1 hade få sälar medan stratum 3 hade flest sälar. Hela området delades sedan in i 200 delområden (grids). Man gjorde sedan ett urval av grids, åkte ut och inventerade dessa och räknade då antalet andningshål.

Det visade sig att stratum 1 innehöll totalt 68 grids, stratum 2 hade 84 grids medan stratum 3 hade 48 stycken grids. Ur de tre stratumen valdes med OSU 17, 12 resp. 11 grids, dvs den totala urvalsstorleken var 40 grids.

Data lades in i en fil som i SAS-utskriften nedan har namnet seals. Filen har tre variabler, stratum (med värdena 1, 2 och 3), holes (antalet andningshål) samt samplingweight (urvalsvikt). Studera SAS-kommandona och resultatutskriften för att svara på följande:

```
data seals;
infile 'h:seals.DAT';
input stratum holes samplingweight;
run;
data stratumtot;
input stratum _total_;
cards;
1 68
2 84
3 48
;
run;
proc surveymeans data=seals total=stratumtot mean clm;
strata stratum;
var holes;
weight samplingweight;
run;
```

The SAS The SURVEYMEANS Procedure

Data Summary

Number of Strata	3
Number of Observations	40
Sum of Weights	200.00004

Variable	Mean	Statistics		
		Std Error of Mean	Lower 95% CL for Mean	Upper 95% CL for Mean
holes	4.985910	0.590132	3.790188	6.181632

- I utskriften finns ett 95 % konfidensintervall för genomsnittliga antalet andningshål per grid. Detta är inte så intressant, utan beräkna i stället ett 95% konfidensintervall för totala antalet andningshål på Svalbard. 1p
- Man kan vara litet tveksam till allokeringen av urvalet. Vilka urvalsstorlekar erhålls vid proportionell allokering av en total urvalsstorlek på 40 grids? 1p