

Övningsuppgifter

Rekommenderad läsning:

Uppgift 1-8 : Kapitel 1, 2 och 3

Uppgift 1

En undersökning ska planeras. Man vill då skatta andelen invånare p som ställer sig positiva till en utvidgning av ett köpcentrum. 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 2500.

2p

Uppgift 2

En mindre kommun är intresserad av att undersöka kostnaderna för fallolyckor som skett under en viss period. En av faktorerna de är intresserade av är hur många olyckor som resulterat i att en person fortfarande har men från olyckan 12 månader efter det att olyckan inträffade. Totalt har 13417 olyckor inträffat och från dessa väljs 200 olyckor ut slumpmässigt för att studeras vidare. I undersökningen framgår att i 38 olyckor har personer fortfarande men 12 månader efter olyckan.

- Beräkna en skattning av andelen olyckor där en person fortfarande har men 12 månader efter olyckstillfället med tillhörande konfidensintervall (95 %).
- Beräkna en skattning av det totala antalet olyckor med tillhörande konfidensintervall (95 %) där en person fortfarande har men 12 månader efter olyckstillfället.
- Baserat på denna undersökning vill kommunen löpande börja studera denna aspekt av olyckorna. För att kunna identifiera skillnader över tid har de bestämt sig för att konfidensintervallet (95 %) för skattningen av andelen maximalt får vara ± 0.05 , oavsett andelen olyckor i urvalet. Baserat på att totala antalet olyckor är samma som i den första undersökningen, beräkna vilken urvalsstorlek som skulle krävas för att uppnå denna precision.

Uppgift 3

Visa att populationsvariansen S^2 är lika med $\frac{Np(1-p)}{N-1}$ då variabeln y endast kan anta värdena 1 och 0. 2p

Uppgift 4

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 urvals varianserna 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. 1p

Uppgift 5

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). Det visade sig att stratum 1 innehöll totalt 68 grids, stratum 2 hade 84 grids medan stratum 3 hade 48 stycken grids.

Man gjorde sedan ett urval av grids, åkte ut och inventerade dessa och räknade då antalet andningshål. 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:

- Bestäm med hjälp av utskriften ett 95% konfidensintervall för totala antalet andningshål. 1p
- Ange urvalsvikterna som används i programmet. Med andra ord, hur ser urvalsvikterna ut vid stratifierat urval och i detta ex? 2p

```
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 sum clsum;
strata stratum;
var holes;
weight samplingweight;
run;
```

The SAS System 18:16 Thursday, March 17, 2005 3

The SURVEYMEANS Procedure

Data Summary

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

Statistics

Variable	Mean	Std Error of Mean	Lower 95% CL for Mean	Upper 95% CL for Mean	Sum	Std Dev
holes	4.985910	0.590132	3.790188	6.181632	997.182240	118.026499

Statistics

Variable	Lower 95% CL for Sum	Upper 95% CL for Sum
holes	758.037837	1236.326643

Uppgift 6

Visa att $E[\bar{y}_s] = \bar{y}_U$. Förklara noggrant alla steg.

2p

Uppgift 7

Under denna vinter har snöröjningen varit ett hett ämne i pressen. Anta nu att kommunen vill undersöka attityden till snöröjningen i innerstaden. Eftersom man tror att attityden skiljer sig beroende på var man bor så har tre stadsdelar valts ut, Lambohov, Ryd och innerstaden. Låt p vara andelen invånare i de tre stadsdelarna som tycker att snöröjningen har skötts på ett tillfredställande sätt i innerstaden. Resultatet av denna fiktiva undersökning visas i tabellen nedan.

Stadsdel	Totala antalet invånare	Urvalsstorlek	Antal positiva i urvalet
Lambohov	3500	20	11
Ryd	4000	20	10
innerstaden	3000	20	5

- a) Se de tre stadsdelarna som tre strata. Skatta p med ett 95% konfidensintervall. 3p
- b) Vilka tydliga brister har denna undersökning? Hur skulle den kunna ha förbättrats? 1p
- c) Skatta med ett 95% konfidensintervall skillnaden i andelen positiva mellan Lambohov och innerstaden. Tolka. 1p
- d) Man har här använt sig av lika allokering. Hur skulle allokeringen bli om man vid ett senare tillfälle vill göra om undersökningen och då använda Neyman-allokering med totalt 60 observationer. 2p

Uppgift 8

Systemebolagets dotterbolag IQ är intresserade av hur mycket pengar studenter vid Linköpings universitet spenderat på alkohol den senaste terminen. De är intresserade av de sammanlagda kostnaderna (y) och skickar därför ut en postenkät till ett stratifierat slumpmässigt urval av storlek 1000 från de 17866 studenter som studerar vid Linköpings universitet. Stratifiering gjordes baserat på antalet terminer studenterna studerat. Kostnaderna i kr per respondent (c) är lika för alla (postenkät). Totalt deltog $n_r = 635$ studenter i undersökningen.

Det fanns också intresserade av hur många som spenderat mer än 2000 kronor på alkohol (p).

Undersökningen gav följande resultat:

	N_h	n_h	n_{rh}	\bar{y}_h	s_h	p_h	c_h
6 -	2902	162	98	1502.67	306.18	0.06	12
2 - 6	10730	601	378	2015.26	459.84	0.51	12
< 2	4234	237	159	2894.03	734.33	0.90	12
Samtliga	17866	1000	635	2156.19	699.18	0.54	

- Baserat på resultatet ovan beräkna en punktskattning för \bar{y}_U med tillhörande konfidensintervall (99%).
- Beräkna designvikterna för respektive strata.
- Visa att proportionerlig allokering användes i undersökningen.
- Beräkna designeffekten.

Tabell

NORMAL CUMULATIVE DISTRIBUTION FUNCTION

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986