



Försättsblad till skriftlig tentamen vid Linköpings Universitet

Datum för tentamen	2010-06-08
Sal (1) Om tentan går i flera salar ska du bifoga ett försättsblad till varje sal och <u>ringa in</u> vilken sal som avses	U11
Tid	8-12
Kurskod	732G26
Provkod	TENA
Kursnamn/benämning Provnamn/benämning	Surveymetodik med uppsats Tentamen
Institution	IDA
Antal uppgifter som ingår i tentamen	4
Jour/Kursansvarig Ange vem som besöker salen	Anders Nordgaard
Telefon under skrivtiden	0761-354599
Besöker salen ca kl.	Endast telefonjour
Kursadministratör/kontaktperson (namn + tfnr + mailaddress)	Carita Lilja, 1463, carli@ida.liu.se
Tillåtna hjälpmedel	Läroboken "Lohr, S.L.: Sampling-Design and Analysis. 1:a el. 2:a uppl. Brooks/Cole" (anteckningar får finnas), Räknedosa (valfri)
Övrigt	
Vilken typ av papper ska användas, rutigt eller linjerat	Rutigt
Antal exemplar i påsen	

732G26

SURVEYMETODIK MED UPPSATS

TENTAMEN

TISDAGEN 8 JUNI 2010, KL. 08.00–12.00

Hjälpmedel:

Läroboken "Lohr, S: Sampling-Design and analysis. 1:a el. 2:a uppl. Brooks/Cole"
(anteckningar får finnas). Räknedosa.

Jourhavande lärare:

Anders Nordgaard (endast via telefon)

Poänggränser:

Skrivningen ger maximalt 20 poäng. För betyget Godkänd krävs normalt 12 poäng.

Lycka till!

1. Antag att vi har gjort ett OSU om 200 individer från en population bestående av 55560 individer. Olika frågor ställs till individerna och bl.a. frågas om det totala antalet timmar man sett på TV den senaste veckan. För hela urvalet blir totalsumman av antalet timmar 5540 och standardavvikelsen blir 15 timmar.
 - a) Beräkna en punktskattning och ett 99% konfidensintervall för det *sammanlagda antalet* timmar, t , alla individer i populationen har sett på TV den senaste veckan. (2p)
 - b) Utgående från informationen i urvalet, hur stort urval borde göras om bredden hos ett 99% konfidensintervall för t skall blir högst 280 000 (timmar). (2p)
 - c) Antag nu att 10 av de 200 individerna i urvalet visar sig inte tillhöra målpopulationen. För dessa 10 individer är medelvärdet 28.5 timmar och standardavvikelsen är 10.9 timmar. Ta hänsyn till detta och beräkna en ny punktskattning och ett nytt 99% konfidensintervall för t . (*Ledning:* De individer i urvalet som faktiskt tillhör populationen kan betraktas som en redovisningsgrupp (domain of study).) (1p)
2. Betrakta undersökningen i uppgift 1 igen. Någon föreslår att den aktuella urvalsdesignen inte passar för att mäta TV-tittandet. I en parallell undersökning har ett stratifierat urval gjorts om 150 individer bland de i populationen som har en anställning och 50 individer bland de som saknar anställning. De 150 har totalt tittat 3350 timmar på TV den senaste veckan och standardavvikelsen är 10.3 timmar. De 50 har totalt tittat 1800 timmar och standardavvikelsen är 9.5 timmar. I populationen är det c:a 7.1% av individerna som saknar anställning.
 - a) Beräkna utgående från dessa data en punktskattning och ett 99% konfidensintervall för t , dvs. det sammanlagda antalet timmar alla individer i populationen har sett på TV den senaste veckan. (3p)
 - b) I såväl den parallella undersökningen som i den som analyserades i uppgift 1 har totalt 200 individer valts ut. Vad blir en optimal allokering av 200 individer över de två strata "individer med anställning" och "individer utan anställning"? (2p)

3. Antag att det i en stor stad finns totalt 55 skolor med elever i åldrarna 13–16 år. I varje skola är eleverna indelade i klasser utifrån vilket skolår de tillhör och utifrån vilken skola de gick i innan de kom till denna skola. Om vi specifikt undersöker elever som tillhör skolår 8 så är det totala antalet elever i alla 55 skolorna 5954 stycken. Vidare vet man att det totala antalet klasser i de 55 skolorna (alla skolår inräknade) är 631.

Ett OSU om 3 skolor valdes och i varje skola gjordes ett OSU om 10% av alla elever i skolår 8. I varje sådant urval tillfrågades eleverna om de hade druckit alkohol den senaste månaden. Resultatet blev

Skola	Totalt antal klasser (alla skolår)	Antal elever i skolår 8	Antal elever i urvalet som druckit alkohol den senaste månaden
1	11	92	2
2	9	83	1
3	13	149	3

- a) Beräkna en väntevärdesriktig punktskattning och ett 95% konfidensintervall för andelen elever, p , i studieår 8 (i de 55 skolorna) som druckit alkohol den senaste månaden. (3p)
- b) Antag hypotetiskt att undersökningen skulle ha ett bortfall om c:a 50% i var och en av skolorna, men att vi ändå skulle ha fått de svar vi fick i tabellen ovan. Om man ignorerar bortfallet, vilken punktskattning och vilket konfidensintervall motsvarande det i a)–uppgiften får man? Kommentera resultatet. (Avrunda 50% av respektive urval till närmaste heltal.) (3p)
4. Ett OSU om 700 individer gjordes från ett nyligen uppdaterat register av medlemmar i ett politiskt parti omfattande totalt 381 960 individer. Till individerna i urvalet ställdes en fråga om hur mycket pengar de hade spenderat på sin semester närmast föregående år. Vidare finns i registret en 5 år gammal uppgift om taxerad inkomst för varje individ. Man vill använda denna information (även om den inte är uppdaterad) för att uppskatta det genomsnittliga beloppet en medlem spenderar på semestern. Följande har beräknats från de svar man fick (inget bortfall förekom osannolikt nog) och från information i registret:

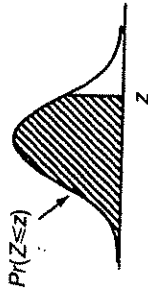
$$\begin{aligned}
\Sigma(\text{belopp spenderat på semestern}) &= 3655.89 (\times \text{ SEK } 1000) \\
\Sigma(\text{belopp spenderat på semestern})^2 &= 23881.3 (\times (\text{SEK } 1000)^2) \\
\Sigma(\text{årlig taxerad inkomst}) &= 248721 (\times \text{ SEK } 1000) \\
\Sigma(\text{årlig taxerad inkomst})^2 &= 92032277 (\times (\text{SEK } 1000)^2) \\
\Sigma(\text{belopp spenderat på semestern}) \times (\text{årlig taxerad inkomst}) &= 1400947 (\times (\text{SEK } 1000)^2)
\end{aligned}$$

Totalsumman av taxerade inkomster i hela registret är ungefär 135730000 (\times SEK 1000)

- a) Beräkna en kvotskattning tillsammans med ett 95% konfidensintervall för det genomsnittliga belopp en medlem spenderade på sin semester föregående år. (2p)
- b) Beräkna en regressionskattning tillsammans med ett 95% konfidensintervall för det genomsnittliga belopp en medlem spenderade på sin semester föregående år. (2p)

Tabell 3a. Normalfördelningen

Om Z är en standardiserad normalfördelad variabel ger tabellen $Pr(Z \leq z)$.



z	0	1	2	3	4	5	6	7	8	9
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	0,7704	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

Kommentar:

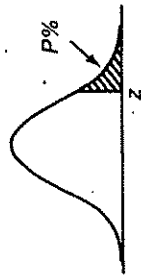
På grund av den standardiserade normalfördelningens symmetri kring punkten noll är sannolikheterna endast tabellerade för positiva z -värden.

z	0	1	2	3	4	5	6	7	8	9
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	0,9983	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9988	0,9989	0,9989	0,9990	0,9990
3,1	0,9990									
3,2	0,9993									
3,3	0,9995									
3,4	0,9997									
3,5	0,9998									
3,6	0,9998									
3,7	0,9999									

För större värden se tabell 3b.

Tabell 3b. Normalfördelningen

Det mot en given sannolikhet svarande z -värdet.



$P\%$	z	$P\%$	z	$P\%$	z	$P\%$	z
50	0,0000	4,8	1,6646	2,4	1,9774	0,9	2,3656
45	0,1257	4,6	1,6849	2,3	1,9954	0,8	2,4089
40	0,2533	4,4	1,7060	2,2	2,0141	0,7	2,4573
35	0,3853	4,2	1,7279	2,1	2,0335	0,6	2,5121
30	0,5244	4,0	1,7507	2,0	2,0537	0,5	2,5758
25	0,6745	3,8	1,7744	1,9	2,0749	0,4	2,6521
20	0,8416	3,6	1,7991	1,8	2,0969	0,3	2,7478
15	1,0364	3,4	1,8250	1,7	2,1201	0,2	2,8782
12	1,1750	3,2	1,8522	1,6	2,1444	0,1	3,0902
10	1,2816	3,0	1,8808	1,5	2,1701	0,05	3,2905
9	1,3408	2,9	1,8957	1,4	2,1973	0,01	3,7190
8	1,4051	2,8	1,9110	1,3	2,2262	0,005	3,8906
7	1,4758	2,7	1,9268	1,2	2,2571	0,001	4,2649
6	1,5548	2,6	1,9431	1,1	2,2904	0,0005	4,4172
5	1,6449	2,5	1,9600	1,0	2,3263	0,00005	4,8916

