

Tentamen

Linköpings Universitet, Institutionen för datavetenskap, Statistik

Kurskod och namn:	732G60 Statistiska metoder
Datum och tid:	2014-03-27 08:00 – 12:00
Jourhavande lärare:	Tommy Schyman
Tillåtna hjälpmedel:	Kursens formelsamling samt häfte med tabeller, båda dessa ska vara utan anteckningar. Valfri räknedosa. Linjal.
Betygsgränser:	Tentamen omfattar totalt 20p. Godkänt från och med 12p, väl godkänt från och med 16p.

Kom ihåg:

- Redovisa och motivera tydligt alla dina lösningar!
- Tolka det du kommer fram till i ord!

Uppgift 1 (5p)

Ett företag bestående av tre större avdelningar har efter många om och men valt att byta operativsystem från Windows till Linux på de anställdas datorer. En månad efter detta skifte gjordes en undersökning, där de anställda fick ange om de tyckte att deras datorer fungerar bättre eller sämre efter skiftet, eller om de inte märkte någon skillnad. Utfallet blev följande:

Avdelning	Åsikt		
	Sämre	Ingen skillnad	Bättre
Avdelning 1	12	26	45
Avdelning 2	32	47	43
Avdelning 3	36	21	7

- Skapa ett lämpligt stapeldiagram för att illustrera skillnader/likheter mellan avdelningarna. (2p)
- Undersök med hjälp av ett test om det finns statistiskt säkerställda skillnader i åsikt mellan avdelningarna. Använd 5 % signifikansnivå. (3p)

Uppgift 2 (5p)

En trafikplanerare gör antagandet att 40 % av invånarna i Linköping cyklar till sin dagliga sysselsättning (arbete, skola etc.). För att undersöka detta ska man fråga 50 slumpmässigt utvalda Linköpingsbor om de cyklar till sin dagliga sysselsättning eller inte.

- a) Vad är sannolikheten att mellan 18 och 20 personer i stickprovet cyklar till sin dagliga sysselsättning om trafikplanerarens antagande är korrekt? (2p)

Undersökningen genomförs, och bland de 50 tillfrågade personerna svarar 23 stycken att de cyklar till sin dagliga sysselsättning.

- b) Beräkna ett 95 % konfidensintervall för andelen Linköpingsbor som cyklar till sin dagliga sysselsättning. Resonera utifrån detta intervall om trafikplanerarens antagande kan stämma. (2p)
- c) Trafikplaneraren vill undersöka denna andel en gång till, och vill att felmarginalen i ett 95 % konfidensintervall ska vara högst 5 %. Använd andelen från b) och beräkna hur många personer som minst bör ingå i denna undersökning. (1p)

Uppgift 3 (6p)

Ett företag vill undersöka sambandet mellan inköpskostnad för datorer (i tusentals kronor) och hur länge de används innan de skrotas (i månader). Närmare bestämt vill man undersöka hur inköpskostnaden kan förklara datorernas livslängd (i månader). Tio datorer som har skrotats ingår i undersökningen, och följande summor sammanställs:

$$\begin{aligned} \sum \text{Inköpskostnad} &= 89,7 & \sum \text{Inköpskostnad}^2 &= 853,25 \\ \sum \text{Månader} &= 176 & \sum \text{Månader}^2 &= 3306 & \sum \text{Inköpskostnad} * \text{Månader} &= 1676,5 \end{aligned}$$

- a) Beräkna a och b i ekvationen $y = a + b * x$ samt tolka dessa (om lämpligt). (3p)
- b) Företagets inköpare planerar att köpa in en dator som kostar tiotusen kronor. Använd ekvationen från a) för att uppskatta dess livslängd. (1p)
- c) Motivera vilket intervall som är lämpligt att beräkna till uppskattningen i b). (1p)
- d) Anta att förklaringsgraden (determinationskoefficienten) i denna modell är 0,943. Tolka denna förklaringsgrad i ord. (1p)

Uppgift 4 (4p)

Företaget från uppgift 3 köper framförallt in datorer av två märken och anar att märke A:s datorer håller längre än datorer från märke B. 10 datorer från märke A undersöks, och från märke B undersöks 9 datorer. I tabellen nedan redovisas datorernas livslängd i månader (vi bortser från eventuellt samband mellan inköpskostnad och livslängd).

Märke A	Märke B
21	20
18	14
17	15
12	13
16	18
18	11
22	12
23	14
17	10
19	

- Beräkna medellivslängd och standardavvikelse för båda märkena. (2p)
- Undersök med hjälp av hypotesprövning om företagets aningar kan stämma, dvs. att märke A:s datorer har längre livslängd än datorer från märke B (5 % signifikansnivå). (2p)