

Tentamen

Linköpings Universitet, Institutionen för datavetenskap, Statistik

Kurskod och namn:	732G60 Statistiska metoder
Datum och tid:	2014-04-26 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 (3p)

En uppdatering för Windows ska enligt uppgift leda till att datorerna startar upp snabbare. För att undersöka detta mättes starttiderna (i sekunder) före och efter uppdateringen på tio stycken utvalda datorer (det är alltså samma tio datorer som undersöks före och efter uppdateringen). Utfallet blev följande:

Dator	Innan uppdatering	Efter uppdatering
1	30	27
2	32	30
3	27	28
4	21	19
5	41	35
6	26	23
7	23	25
8	35	35
9	39	35
10	37	36

Undersök med hjälp av hypotesprövning om datorerna startar upp snabbare efter uppdateringen. Använd 5 % signifikansnivå. (3p)

Uppgift 2 (3p)

Genomsnittspriset (i svenska kronor) för 1 kg svensk nötfärs har under de senaste fem åren haft följande utveckling:

År	Pris nötfärs	KPI (basår = 1980)
2009	66,1	299,01
2010	67,3	302,47
2011	68,2	311,43
2012	68,9	314,20
2013	70,0	314,06

- Beräkna och tolka ett enkelt prisindex för den svenska nötfärsen. (1p)
- Beräkna och tolka ett relativprisindex för den svenska nötfärsen. (2p)

Uppgift 3 (6p)

För att undersöka TV-tittartiden bland gymnasieungdomar tillfrågades tio stycken slumpmässigt valda gymnasieungdomar hur många timmar de kollar på TV i veckan. Svaren blev enligt följande:

TV-tittartid
10
7
9
12
14
16
8
11
7
13

- Beskriv datamaterialet med hjälp av ett lådagram. (2p)
- Beräkna ett 95 % dubbelsidigt konfidensintervall för den genomsnittliga TV-tittartiden bland gymnasieungdomar. (3p)
- I en kommande undersökning gällande TV-tittartid vill man skapa ett 95 % dubbelsidigt konfidensintervall med en intervallbredd på högst två timmar per vecka. Använd information från detta stickprov och beräkna hur många ungdomar som minst bör ingå i denna kommande undersökning. (1p)

Uppgift 4 (5p)

För att utreda hur systemvetares ålder (i år) påverkar deras månadslön (i tusentals kronor) tillfrågades 22 slumpmässigt utvalda systemvetare. Följande summor sammanställdes:

$$\begin{aligned}\sum \text{Ålder} &= 862 & \sum \text{Ålder}^2 &= 36\,096 \\ \sum \text{Lön} &= 832 & \sum \text{Lön}^2 &= 33\,860 & \sum \text{Lön} * \text{Ålder} &= 34\,843\end{aligned}$$

- a) Beräkna och tolka korrelationskoefficienten med hjälp av de ovanstående summorna. (2p)

Nedan visas regressionsmodellen, anpassad med hjälp av Minitab.

Regression Analysis: Lön versus Ålder

The regression equation is
Lön = - 0,05 + 0,967 Ålder

Predictor	Coef	SE Coef	T	P
Constant	-0,055	2,829	-0,02	0,985
Ålder	0,96659	0,06985	13,84	0,000

S = 3,36527 R-Sq = 90,5% R-Sq(adj) = 90,1%

- b) Tolka regressionskoefficienten tillhörande variabeln *Ålder* i utskriften ovan. (1p)
c) Tolka determinationskoefficienten (förklaringsgraden) i utskriften ovan. (1p)
d) När ska konfidensintervall respektive prediktionsintervall för prognoser användas? (1p)

Uppgift 5 (3p)

Ett svenskt universitet har genomfört en mindre undersökning gällande vad studenterna tycker om lunchtubudet på campus. Det ställdes enbart en fråga, och på den fick studenterna helt enkelt ange vad de tycker om lunchtubudet på campus på en fyrgradig skala. Svaren delades upp på kvinnor och män, fördelningen blev följande:

	Mycket dåligt	Dåligt	Bra	Mycket bra
Kvinnor	12	15	28	45
Män	18	10	30	38

Undersök med hjälp av ett test om det finns skillnad mellan könen i åsikt om lunchtubudet på campus. Använd 5 % signifikansnivå. (3p)