



Försättsblad till skriftlig tentamen vid Linköpings Universitet

Datum för tentamen	2011-02-24
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	TER1
Tid	8-12
Kurskod	732G60
Provkod	TEN1
Kursnamn/benämning Provnamn/benämning	Statistiska metoder Tentamen
Institution	IDA
Antal uppgifter som ingår i tentamen	4
Jour/Kursansvarig Ange vem som besöker salen	Oleg Sysoev
Telefon under skrivtiden	0735673482
Besöker salen ca kl.	10
Kursadministratör/kontaktperson (namn + tfnr + mailaddress)	Carita Lilja, 1463, carita.lilja@liu.se
Tillåtna hjälpmedel	Med skrivningen häftad formelsamling, Räknedosa.
Övrigt	
Vilken typ av papper ska användas, rutigt eller linjerat	Rutigt
Antal exemplar i påsen	43

Tentamen

Linköpings Universitet, Institutionen för datavetenskap, Statistik

Kurskod och namn:	732G60 Statistiska Metoder
Datum och tid:	2011-02-24, 8-12
Jourhavande lärare:	Oleg Sysoev
Tillåtna hjälpmedel:	Med skrivningen häftad formelsamling. Räknedosa.
Betygsgränser:	Tentamen omfattar totalt 20p. Godkänt från 12p, väl godkänt från 16p. Siffrorna i uppgifterna är påhittade.

Redovisa och motivera tydligt alla dina lösningar!

Formelsamlingen finns på de sista sidorna

Uppgift 1 (3p)

Under tillverkningskontrollprocessen undersökte man 20 spritflaskor som har volymen 0,7L. Det har visat sig att flaskorna innehöll 0.67L sprit i genomsnitt och standardavvikelsen i stickprovet var 0.064. Genomför en hypotesprövning för att undersöka om flaskorna i genomsnitt innehåller mindre sprit än de skall. Använd signifikansnivån 5% och tolka resultatet. Vilket är P-värde av testet? (3p)

Uppgift 2 (6p)

- Man planerar att genomföra en partisympatiundersökning för parti S och då får man i uppdrag att beräkna hur stort stickprov man behöver för att få säkra resultat. Då vill uppdragsgivaren få ett 99% konfidensintervall för andelen sympatisörer för S och det medges att ha 0.2% som intervallets bredd. Beräkna stickprovets storlek som kan uppfylla dessa krav. (3p)
- Partisympatiundersökningen genomfördes i november och i december 2010 och man valde då 2000 personer slumpmässigt varje månad. Undersökningen visade att 28.5% skulle vilja rösta för S i november och 26.9% i december. Beräkna ett intervall med 95% konfidensgrad som visar skillnaden mellan andelen sympatisörer i november och december 2010. Innebär detta intervall att det blev en minskning av andelen sympatisörer? (3p)

Uppgift 3 (6p)

I en undersökning av huspriser i Albuquerque (USA) har man fått fram information om de 10 senast sålda husen. Datamaterialet visas nedan som en tabell där X är husarean uttryckt i 100-tal kvadratfot och Y är priset uttryckt i 10000-tal USD.

Pris	5,4	7	7,5	8,8	9,5	10,2	11	12,4	13,8	20
Husarea	11,4	15,1	13,4	11,6	15,8	14,8	16,6	18,9	18,4	25,8

- Skatta regressionskoefficienter med minsta-kvadratmetoden (redovisa uträkning) och redogör den skattade regressionsekvationen. Tolka båda regressionskoefficienterna. (3p)
- Vilken area skulle ett hus ha i genomsnitt enligt modellen om priset ligger på 90000 USD? (1p)
- Man har utökat analysen till 117 observationer och har genomfört en ny regressionsanalys (se utskriften nedan). Beräkna ett konfidensintervall för lutningen med konfidensgraden 95 procent. (2p)

Regression Analysis: Pris versus Husarea

The regression equation is
Pris = - 4,54 + 0,933 Husarea

Predictor	Coef	SE Coef	T	P
Constant	-4,542	1,920	-2,37	0,046
Husarea	0,9334	0,1152	8,10	0,000

Uppgift 4 (5p)

Följande tabell visar hur mycket järn produceras i Sveriges gruvor.

Industriproduktionsindex (IPI/PVI), ej kalenderkorrigerad, 2000=100 efter näringsgren SNI 2002 och tid																			
År	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
13.1 järnmalmgruvor	95	98	99	83	89	102	104	107	103	90	100	95	98	105	109	114	115	123	124

- Betrakta samtliga indextalen som ett stickprov och rita ett lådagram. Vad kan man säga om uteliggarnärvaro och fördelningens symmetri? (3p)
- Byt basår till 1990 och visa den nya indexserien som ett tidseriediagram. Hur många procent mer eller mindre är det nya indexet år 2008 i jämförelse med år 1990? (2p)

Medelvärde $\bar{x} = \frac{\sum x}{n}$

Medianen

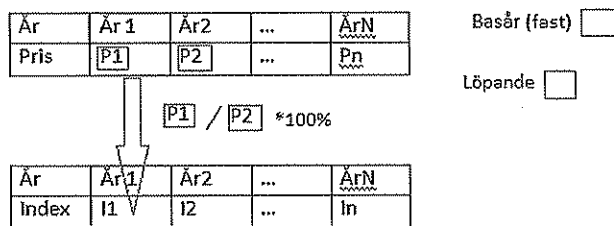
- Udda n $x_m = x_{(n+1)/2}$
- Jämmt n $x_m = \frac{x_{n/2} + x_{n/2+1}}{2}$

Standardavvikelse: $s = \sqrt{\frac{n}{n-1} (\overline{x^2} - \bar{x}^2)}$

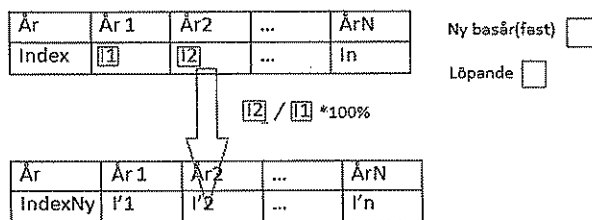
Lådagrammets punkter:

- Utebyggare (punkter $< Q1 - 1.5 * \text{kvartilavstånd}$)
- Max(minvärde, $Q1 - 1.5 * \text{kvartilavstånd}$)
- Q1
- Medianen
- Q3
- Min(maxvärde, $Q3 + 1.5 * \text{kvartilavstånd}$)
- Utebyggare (punkter $> Q3 + 1.5 * \text{kvartilavstånd}$)

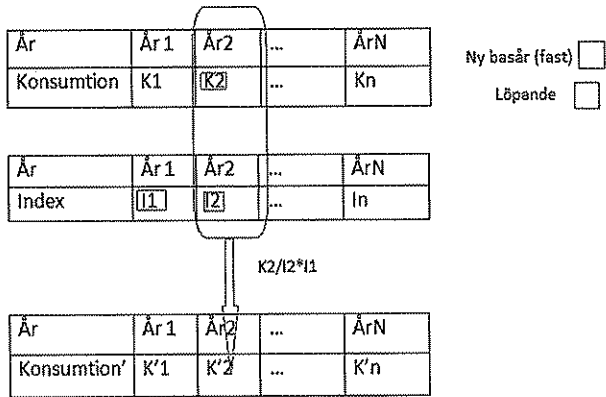
Omräkning till Index



Byte av basår



Deflatering



Regression

$$y = a + bx \quad r = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sqrt{(\overline{x^2} - \bar{x}^2)(\overline{y^2} - \bar{y}^2)}} \quad b = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2} \quad a = \bar{y} - b\bar{x}$$