

STOCKHOLMS UNIVERSITET
Statistiska institutionen
Jörgen Säve-Söderbergh

VT 2009

Skriftlig tentamen på momentet Statistisk dataanalys III (SDA III), 3 högskolepoäng ingående i kursen Undersökningsmetodik och statistisk dataanalys, 15 högskolepoäng, den 25 mars 2009 kl 9.00 – 11.00. Resultatet anslås senast den 8 april på anslagstavlan, plan 3.

Skrivtid: 2 timmar.

Hjälpmedel: godkänd miniräknare utan lagrade formler eller text.

Tentamen består av 18 uppgifter som kan ge totalt 25 poäng, således sju stycken tvåpoängsuppgifter. Examinationen betraktas som avklarad om poäng motsvarande lägst betyget E uppnås. Följande betygsgränser gäller:

Betyg	Poäng
A	24-25
B	22-23
C	19-21
D	17-18
E	15-16
Fx	12-14
F	0-11

Observera att felaktiga svar ej ger minuspoäng. Använd den särskilda svarsbilagan och ringa in det svarsalternativ som du tycker **bäst besvarar frågan**. Fler inringade alternativ samt andra oklarheter gör att frågan anses obesvarad. **Observera att endast den särskilda svarsbilagan skall lämnas in. Beräkningar beaktas ej.** Var noga med att tydligt skriva namn och personnummer på svarsbilagans båda sidor. Skriv dessutom på svarsbilagans båda sidor det platsnummer du har i tentamenssalen.

LYCKA TILL!

1. Vi har ett material om fyra stycken observationspar (x_i, y_i) som är $(1, -2)$, $(2, -3)$, $(3, -2)$ och $(4, -3)$. Om vi med hjälp av minsta kvadratmetoden skattar en regressionsekvation till detta material får den vilken form? (2 poäng)

- a) $y = -x$
- b) $y = -2,5$
- c) $y = -2 - 0,2x$
- d) $y = 5 - 2,3x$

2. En lantbrukare experimenterar på sina ägor med olika mängder konstgödselnivå. Nedan visas en förteckning över veteskörd i kilo samt kvävegödning i kilo per viss ytenhet.

Skörd	Kvävegödning
5300	100
5700	200
6500	300
6800	400
7100	500
7100	600

En minitabkörning gav följande resultat;

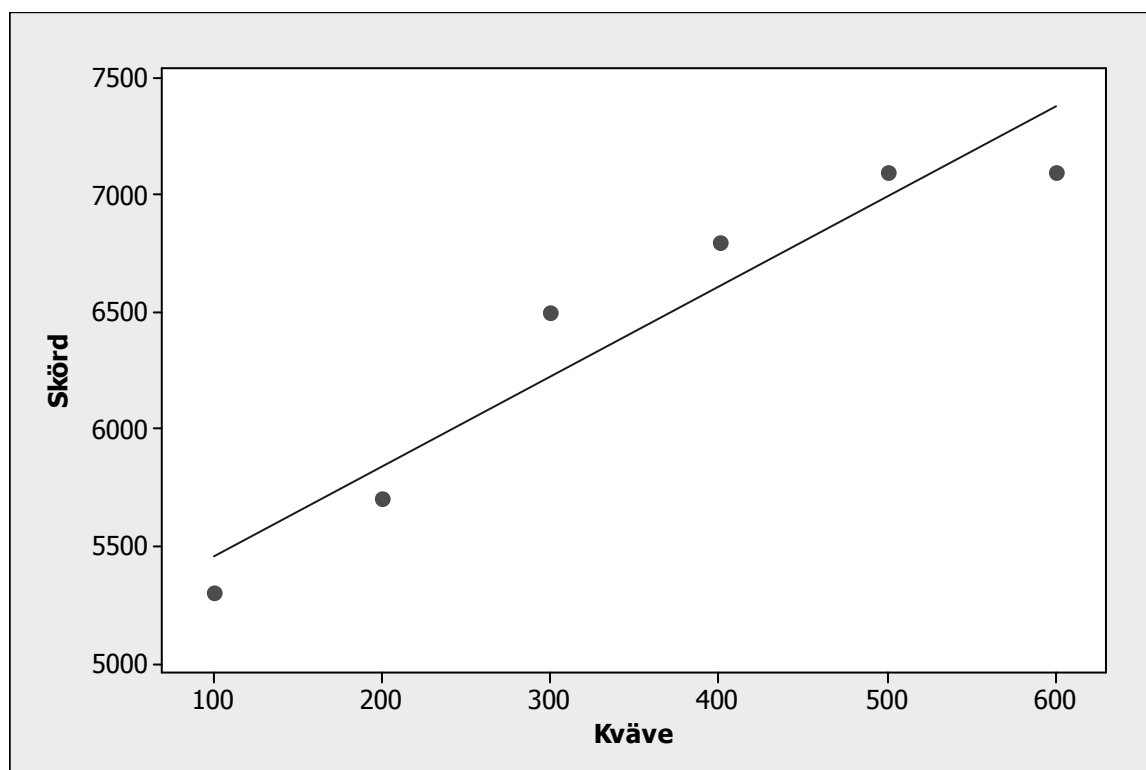
Regression Analysis: Skörd versus Kväve

The regression equation is
Skörd = 5067 + 3.86 Kväve

Predictor	Coef	SE Coef	T	P
Constant	5066.7	230.3	22.00	0.000
Kväve	3.8571	0.5913	6.52	0.003

S = 247.367 R-Sq = 91.4% R-Sq(adj) = 89.3%

Grafiskt ser sambandet ut enligt;



Hur stor andel av variationen i datamaterialet förklaras av den anpassade regressionslinjen?

- a) 95,6 procent.
- b) 91,4 procent.
- c) 3,86 procent.
- d) 22 procent.

3. Hur stor är regressionskoefficienten för modellen i fråga 2?

- a) 0,616
- b) 6,082
- c) 0,785
- d) 3,8571

4. Hur skall interceptet i modellen i fråga 2 tolkas? (2 poäng)

- a) Den genomsnittliga skörden är 5067 kilo.
- b) Skörden ökar 3,86 kilo för varje kilo kväve vi tillför.
- c) Skörden är i genomsnitt 5067 kilo när vi inte tillför något kväve alls.
- d) Utan gödning är skörden 50,67 procent av den genomsnittliga skörden med gödning.

5. Hur stor är residualvariansen i materialet i fråga 2?

- a) 61190,43269
- b) 247,367
- c) 15,728
- d) Denna kvantitet kan inte beräknas med den givna informationen.

6. Beräkna residualen för den enskilda observationen (500,7100).

- a) 105
- b) 205
- c) 305
- d) 405

7. Hur stor är korrelationen mellan variablerna i fråga 2? (2 poäng)

- a) 0,914.
- b) 0,993.
- c) 0,956.
- d) Denna kvantitet kan inte beräknas med den givna informationen.

8. Följande indexserie visar prisutvecklingen för en viss vara;

År	2004	2005	2006	2007	2008
Pris	100	112	119	129	143

Mellan vilka år var den procentuella förändringen i priset som störst?

- a) Mellan 2007 och 2008.
- b) Mellan 2006 och 2007.
- c) Mellan 2005 och 2006.
- d) Mellan 2004 och 2005.

9. Hur stor har den genomsnittliga årliga procentuella förändringen varit för materialet i fråga 8? (2 poäng)

- a) 7,4
- b) 12,7
- c) 6,1
- d) 9,4

10. Vad avser HIPK att möjliggöra?

- a) Underlätta klassifikationer av olika länders bilindustrier.
- b) Jämförelser bland EU-länderna.
- c) Jämföra olika sätt att deflatera en tidsserie.
- d) Värdesäkra EU-medborgarnas sparande.

11. Betrakta följande påståenden:

Påstående 1: NPI visar prisutvecklingen rensad från indirekta skatter och subventioner.

Påstående 2: KPI:s basår är 1914.

- a) Påstående 1 är sant och påstående 2 är falskt.
- b) Påstående 1 är falskt och påstående 2 är sant.
- c) Båda påståendena är sanna.
- d) Inget av påståendena är sant.

12. För fyra olika varor A, B, C och D vill vi studera prisutvecklingen från 2003 till 2008. Vara A kostade 2003 34 kronor och det såldes 350 stycken. 2008 kostade den 39 kronor och försäljningen var tre gånger så stor som 2003. Vara B kostade 2003 95 kronor och det såldes 5000 stycken. 2008 var både pris och kvantitet för vara B fördubblade. Vara C kostade 2003 50 kronor och 50 stycken såldes. År 2008 hade priset på vara C stigit med 50 procent och försäljningen halverats i antal räknat. Slutligen kostade vara D 68 kronor 2003 och 500 stycken såldes. År 2008 hade priset på vara D sänkts med 50 procent och försäljningen halverats i antal räknat.

Vad blir ett Laspeyres fastbasindex för år 2008 för de fyra varorna sammantaget (2003=100)? (2 poäng)

- a) 188,7
- b) 188,1
- c) 190,4
- d) 179,3

13. Beräknat på samma material som i fråga 12; vad blir ett Paasches fastbasindex för år 2008 för de fyra varorna sammantaget (2003=100)?

- a) 193,3
- b) 188,2
- c) 194,4
- d) 190,0

14. Vad menas med *multipl* regressionsanalys?

- a) Vi har fler beroende variabler.
- b) Vi har fler oberoende variabler.
- c) En variant av minsta-kvadratmetoden som inte förutsätter minst intervallskala hos variablerna.
- d) Vi använder fler skattningsmetoder på samma statistiska material.

15. Nedanstående data anger omsättningen i ett företag (i miljoner kronor) under åren 2003-2008

År	Omsättning
2003	2,3
2004	3,5
2005	5,3
2006	7,4
2007	11,7
2008	17,8

Gör först en lämplig transformation av tidsvariabeln där en enhets förändring av t ska motsvara ett halvår. Anpassa sedan med hjälp av minsta-kvadratmetoden en exponentiell trendmodell. Vilket utseende får denna? (2 poäng)

- a) $y = 6,35 \cdot 1,27^t$
- b) $y = 6,35 \cdot 1,22^t$
- c) $y = 6,35 + 1,26t$
- d) $y = 6,35 \cdot 1,28^t$

16. Gör en prognos för 2009 enligt rätt anpassad modell i fråga 15?

- a) 20,9

- b) 25,5
- c) 26,7
- d) 25,3

17. Vi har på tertialdata anpassat en trend med hjälp av ett glidande medelvärde (3-termer) på en viss tidsserie. Vi vill nu också säsongrensa materialet och skattar därför säsongskoefficienter i en multiplikativ modell. Nedan följer en förteckning över de faktiska värdena dividerade med de skattade trendvärdena:

Tertial 1	Tertial 2	Tertial 3
	0,88	1,26
0,90	0,93	1,18
0,94	0,92	

Hur många observationer innehåller den ursprungliga tidsserien på vilken trend och säsongskoefficienter skattas?

- a) 7
- b) 8
- c) 9
- d) 10

18. Beräkna för materialet i fråga 17 en justerad (korrigerad) säsongskoefficient för tertial 3? (2 poäng)

- a) 1,20
- b) 1,21
- c) 1,22
- d) 1,23