

**Del A:** Schema för ifyllande av svar finns på sista sidan

- 1 Följande två modeller har anpassats till data, där man dels använt originalresponsdata (Modell 1) och dels en transform av responsdata (Modell 2).

Modell 1:

$$Y_1 = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \epsilon$$

Modell 2:

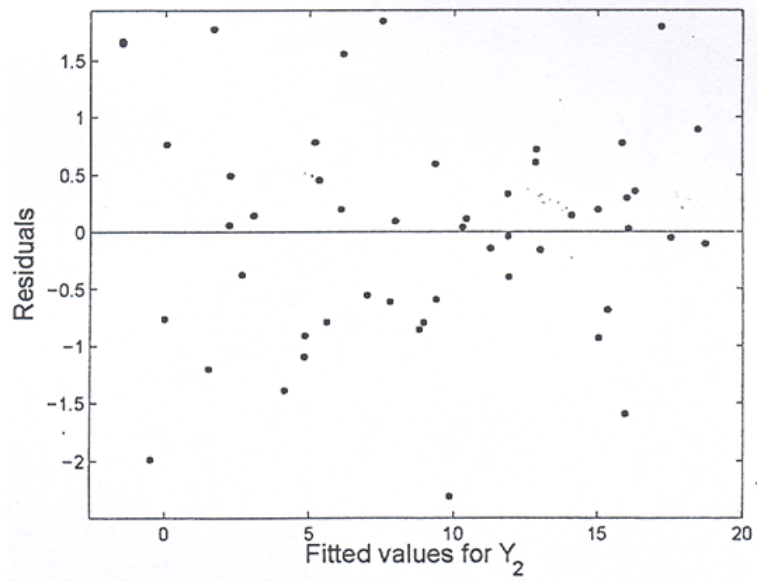
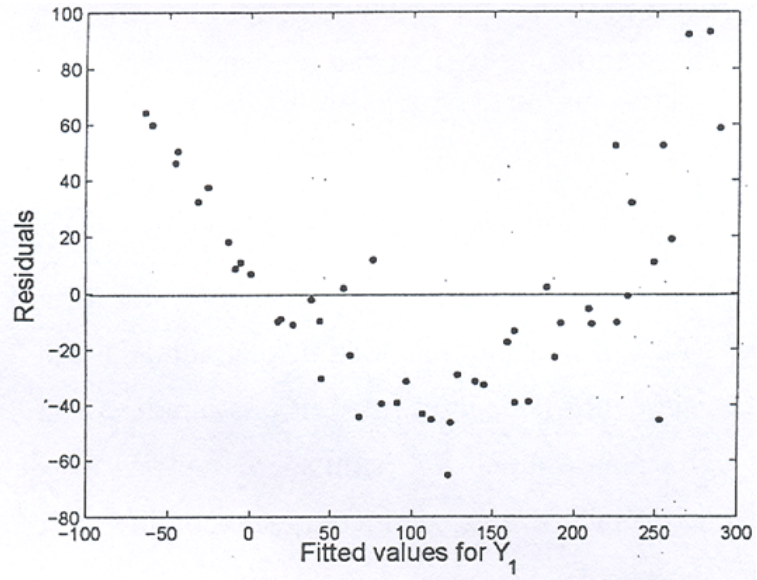
$$Y_2 = \sqrt{Y_1} = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \epsilon$$

Studera figurerna på nästa sida som visar residualplottar för de båda modellerna och läs följande påståenden:

- 1: Baserat på figurerna ser det ut som om transformationen har löst problemet med multikolaritet.
- 2: Baserat på figurerna ser det ut som om transformationen har löst problemet med icke-linearitet mellan responsvariabeln och regressorerna.
- 3: Enbart dessa figurer räcker inte för att avgöra om Modell 2 är en bra modell.

Vilket eller vilka av dessa påståenden är korrekt/korrekta?

- (a)  Påstående 2 är sant, men inte de andra.
- (b)  Påstående 1 och 3 är sanna, men inte 2.
- (c)  Påstående 1 är sant, men inte de andra.
- (d)  Påstående 2 och 3 är sant, men inte 1.
- (e)  Inget påstående är sant.
- (f)  Vet ej.



2 Läs följande påståenden om blockning:

- 1: Syftet med blockning är framför allt att studera blockeffekterna.
- 2: Samspeleplottar ger information om hur väl normalfördelningsantagandet stämmer.
- 3: På grund av hur man randomiserar försöket bör man vara försiktig med att tolka hypotestest på blockeffekterna.

Vilket av följande påståenden är sanna?

- (a)  Påstående 1 och 2 är sanna, medan påstående 3 är falskt.
- (b)  Påstående 2 är sant, medan påstående 1 och 3 är falska.
- (c)  Påstående 1 är sant, medan påstående 2 och 3 är falska.
- (d)  Påstående 2 och 3 är sanna, medan påstående 1 är falskt.
- (e)  Påstående 3 är sant, medan påstående 1 och 2 är falska.
- (f)  Vet ej.

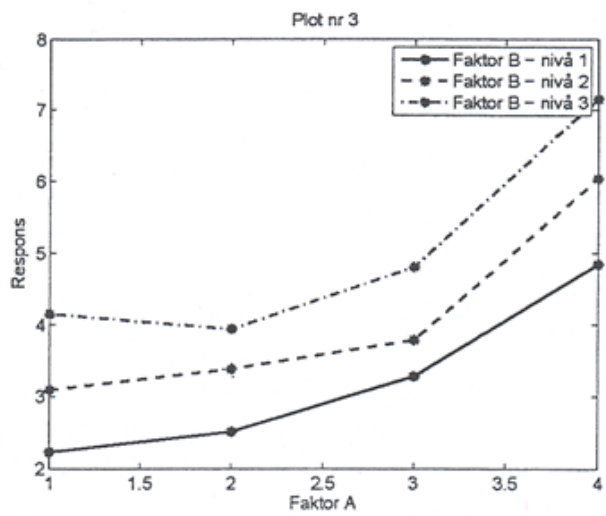
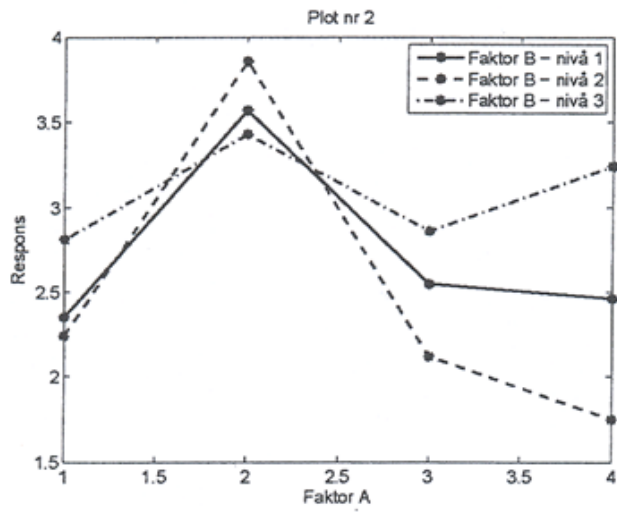
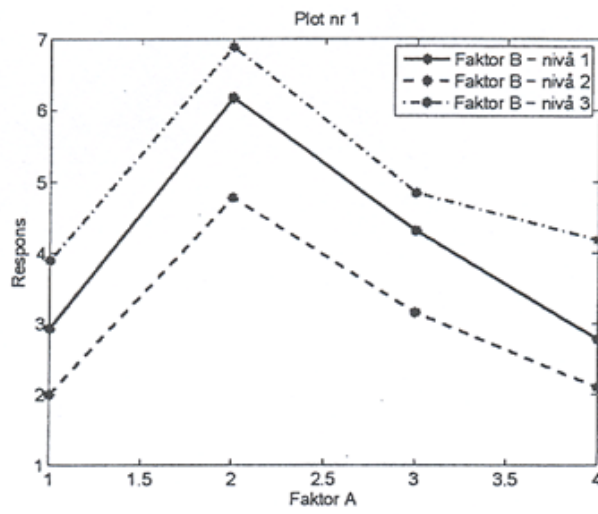
3 Vilket av följande påståenden om multikolaritet är INTE sant?

- (a)  Vid multikolaritet kan man få ett F-test som är signifikant medan inget av de individuella t-testen för de olika regressionskoefficienterna är signifikanta
- (b)  Vid många kontrollerade experiment kan man välja nivån på regressorerna så att man undviker multikolaritetsproblemet.
- (c)  Om multikolaritet förekommer betyder det att om man ändrar en regressor tenderar de andra också att ändra sig.
- (d)  Toleransen och variansinflationsfaktorn (VIF) är två mått på multikolaritet. Om toleransen är låg är också VIF låg.
- (e)   $VIF=2$  indikerar att multikolaritet inte är ett problem.
- (f)  Vet ej.

- 4 Slitaget på 4 olika däcktyper har undersökts genom att 5 olika förare har kört 2000 mil vardera med var och en av däcktyperna. Resultatet har analyserats med hjälp av en variansanalys med förare som block. F-testet för däcktyp ger ett lågt p-värde och alla residualplottar ser bra ut. Man kan inte se att det skulle finnas någon effekt av förare. Varför skulle man vilja göra konfidensintervall enligt Tukeymetoden?
- a  För att kontrollera att variansen är lika stor för alla däcktyper.
  - b  För att undersöka om det finns några samspelseffekter mellan förare och däcktyp.
  - c  För att få konfidensintervall för medeleffekten för varje däcktyp så att alla konfidensintervall har en simultan konfidensgrad.
  - d  För att få konfidensintervall för skillnaden i medeleffekt för varje par av däcktyper så att alla konfidensintervall har en simultan konfidensgrad.
  - e  För att undersöka om det överhuvud taget finns några skillnader i slitage mellan däckstyperna.
  - f  Inget av ovanstående.

5 Nedan ges några exempel på plottar av samspelseffekter. I vilken/vilka kan man tydligt se samspelseffekter mellan faktor A och faktor B?

- a  I plot 1.
- b  I plot 2.
- c  I plot 1 och 2.
- d  I plot 1 och 3.
- e  I plot 3.
- f  Inget av ovanstående.





6 Det justerade förklaringsvärdet,  $R_a^2$ , definieras av

$$R_a^2 = 1 - \frac{SSE/(n - q - 1)}{SSTO/(n - 1)}.$$

Vilket av följande påståenden är korrekt?

- a   $R_a^2$  används för att jämföra två modeller med olika antal regressorer.
- b   $R_a^2$  används för att se hur stor andel av variationen i data som förklaras av modellen.
- c   $R_a^2$  används för att se om modellen passar bra till data.
- d   $R_a^2$  används för att undersöka om det finns multikollinjäritet.
- e   $R_a^2$  används för att undersöka om data är normalfördelade.
- f  Inget av ovanstående.

**Poängberäkning:** Uppgifterna är av flervalstyp, där endast ett alternativ är rätt. Korrekt besvarad uppgift ger 2 poäng, obesvarad uppgift (vet inte eller alternativ f) ger 0 poäng och felaktigt besvarad uppgift ger -0.5 poäng (flera ifyllda alternativ ger automatiskt -1/2 poäng). Inlämnade lösningar kommer ej tas hänsyn till vid rättningen. Fyll i och lämna in denna sida.  
**Svar:** Lägg ut i studieportalen efter tentamens slut.

Uppgift	a	b	c	d	e	f (vet ej)	Poäng
1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
5	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
6	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	