

Samtliga behandlade uppgifter skall förses med utförlig lösning och motivering.

1. Du och en kollega har fått i uppdrag att skatta hur mycket (i kronor) en bils pris påverkas av dess motorstyrka (i hästkrafter) och ni har fått data på en massa olika bilmodeller. Din kollega kör en regression

$$(\text{pris}) = \alpha_0 + (\text{motorstyrka})\alpha_1 + e_1 \quad (1)$$

Du tänker att priset bör ju också bero på bilens storlek, vilken man kan mäta i dess vikt (kilo), så du kör regressionen

$$(\text{pris}) = \beta_0 + (\text{motorstyrka})\beta_1 + (\text{vikt})\beta_2 + e_2 \quad (2)$$

Det visar sig att din kollegas skattning $\hat{\alpha}_1$ och din skattning $\hat{\beta}_1$ skiljer sig kraftigt åt.

- Förklara varför skattningarna är helt olika.
- Vilken olikhet kommer att gälla: $\hat{\alpha}_1 > \hat{\beta}_1$ eller $\hat{\alpha}_1 < \hat{\beta}_1$?

Ni får nu uppdraget att ändra modell-specifikationen (2) så att relativa prisförändringen (procentuella förändringen) beror på skillnaden i motorstyrka i hästkrafter.

- Ange en sådan regressionsmodell. (10p.)

2. Du har skattat en regressionsmodell

$$y = \beta_0 + x_1\beta_1 + x_2\beta_2 + x_3\beta_3 + e$$

och fått resultaten

$$(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3) = (1.253, 4.300, 0.232, 0.894)$$

och kovariansmatrisen

$$\begin{pmatrix} 1.464 & 0.732 & 1.342 & 0.610 \\ 0.732 & 0.427 & 0.739 & 0.488 \\ 1.342 & 0.739 & 1.586 & 1.098 \\ 0.610 & 0.488 & 1.098 & 1.342 \end{pmatrix}$$

Du vill nu testa hypotesen $\beta_1 = 2\beta_3$. Vad blir p-värdet för detta test? Dvs. bestäm den signifikansnivå vid vilken hypotesen precis förkastas. (En tabell över normalfördelningen bifogas.)

(10p.)

3. Enligt en modell för kaffe-priset gäller att priset p bestäms av jämviktsvillkoret $Q^S = Q^D$, dvs. utbud = efterfrågan (i kvantiteter). Utbudet beskrivs av ekvationen

$$\ln Q^S = \alpha_0 + \alpha_1 \ln p + \alpha_2 \ln x + \varepsilon_s$$

där x är storleken på kaffeskörden i Brasilien, och efterfrågan beskrivs av

$$\ln Q^D = \beta_0 - \beta_1 \ln p + \varepsilon_d \quad (\text{D})$$

där ε_* naturligtvis är residualer. Du vill nu skatta efterfråge-kurvan (D). Ange en lämplig skattningsmetod för (D), och motivera varför ditt val är lämpligt. (10p.)

4. En ekonometriker lät sin assistent James köra en OLS-regression

$$y_i = x_i' \beta + e_i, \quad i = 1, \dots, n$$

på ett data set (här är $x_i' = (1, x_{1,i}, \dots, x_{k,i})$ och $\beta' = (\beta_0, \dots, \beta_k)$). Hon fick skattningarna $\hat{\beta}$ av James, men tyckte resultatet såg väldigt egendomligt ut och ville kontrollera att allt gått rätt till. Hon bad därför James att göra om regressionen.

Nu var det tyvärr så att observationerna på x -variablerna var konfidentiella, så dessa hade James förstört, helt enligt överenskommelse, men han hade kvar y_i -na och residualerna \hat{e}_i . Nu bad hon James att beräkna de två summorna

$$\sum_{i=1}^n y_i \hat{e}_i \quad \text{och} \quad \sum_{i=1}^n \hat{e}_i^2$$

När ekonometrikern fick resultatet av dessa summor utbrast hon: "James, din slarver – vad har du ställt till med! Det här är inte residualerna till en OLS!"

- Förklara vad ekonometrikern upptäckt, och hur hon vet att skattningen inte stämmer. (10p.)

5. Du vill skatta sannolikheten p för att man skall lyckas klara förarprovet för körkort för bil beroende på hur många timmar t man övningskört. Du bestämmer dig för en specifikation

$$p = \frac{\beta t}{1 + \beta t}$$

Du har en mängd data för olika personer på antal övningstimmar t_i och om personen klarade eller misslyckades på provet.

En möjlighet är att skatta

$$D = \frac{\beta t}{1 + \beta t} + e$$

med icke-linjär OLS, där D är en dummy-variabel för "klarade provet".

- Visa att detta är en ineffektiv skattning. (Den är konsistent, men det behöver du inte visa.)
- Ange en lämplig skattningsmetod. (10p.)

6. Betrakta “strukturekvationen”

$$y = \beta_0 + x\beta_1 + e$$

där $E[e] = 0$, men $E[e | x] \neq 0$, så vi har inte en regressionsekvation. Vi har dock en stokastisk variabel z som fungerar som instrument för x .

Bestäm ett uttryck för β_1 . Uttrycket får bara innehålla storheterna väntevärdena, varianserna och kovarianserna för variablerna x , y och z . (10p.)