Løsningsforslag INT010 vår 2010

Oppgave 1

a) F-test. Testobservator $F = s_1^2 / s_2^2 = 25^2 / 20^2 =$ **1,56**. Forkastningsområde på 5% signifikansnivå med 29 og 29 frihetsgrader (ikke i tabell – benytter derfor 30 og 30): $F > F_{0,025,\,30,\,30} =$ **2,07** eller $F < 1/F_{0,025,30,30} = 1/2,07 =$ **0,48**. **Aksepter H_0** om lik varians.

b) Testobservator $t = (x_1 - x_2) / \sqrt{[s_p^2\,(1/n_1 + 1/n_2)]} = (420\text{-}408) / \sqrt{[512,5 \cdot (1/30 + 1/30)]} =$ **2,05**, da $s_p^2 = (29 \cdot 25^2 + 29 \cdot 20^2)/58 = 512,5$. Kritisk grense på 5% nivå er med $n_1 + n_2 - 2 = 58$ frihetsgrader (ikke i tabell, benytter 60) $t > t_{0,025,\,60} =$ **2,0** eller $t < -\,t_{0,025,\,60} = -2,0$. **Forkast H_0** (men med liten margin) og påstå at forskjellen i produksjon er signifikant.

c) Hypotesen **H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$ vs. H_1: minst en av forventningene er forskjellig** blir nå testet. Testobservatøren er F = 44,27 og p-verdien er 0,000 (fra Minitab), kan derfor **forkaste H_0**, og påstå at produksjonen av biler er signifikant forskjellig ved fabrikkene. Fra konfidensintervallene for forventningene kan vi påstå at fabrikk 4 produserer flest biler.

Oppgave 2

a) Benytter t-test. Antall frihetsgrader er n-k-1 = 31-2-1 = **28**. Kritisk verdi er $t_{0,025,28} =$ **2,048**. Vi forkaster nullhypotesen for t-verdier større eller lik 2,048 (eller mindre eller lik -2,048). Dvs. alle koeffisienter er signifikante, unntatt real import prices.

b) F-test med frihetsgrader lik 2 og 28 i hhv teller og nevner. Kritisk verdi på 5%-nivå er 3,34. Med en F-verdi på 368,23 kan vi klart **forkaste nullhypotesen**.

c) Fra resultatet i b) og den høye R^2 konkluderer vi med at modellen er god (høy forklaringsgrad). Koeffisienten til real income (inntekt) er positiv, dette virker rimelig. Dess høyere inntekt, dess mer import. Koeffisienten til real import price er også positiv. Dette virker noe mer tvilsomt, dess høyere importpriser (relativt til de innenlandske prisene), dess høyere import. Men dette er nærmest irrelevant, da koeffisienten er nesten lik null og t-verdien lav. Den manglende effekten av pris kan skyldes liten variasjon i variablene eller at modellen ikke beskriver virkeligheten særlig godt og derfor ikke passer med dataene.

d) $\eta_{gdp} = b \cdot \frac{\overline{X}}{\overline{\overline{Y}}} = 0,59 \cdot \frac{536,4}{146,3} = 2,16.$ Dvs. en 1% økning i brutto nasjonalprodukt (inntekt eller GDP) fører til en 2,16% økning i importen. $\eta_{pm} = b \cdot \frac{\overline{X}}{\overline{\overline{Y}}} = 0,05 \cdot \frac{109,4}{146,3} = 0,04.$ Dvs. en 1% økning i relativer priser fører til en 0,04% økning i importen. Som i c) virker dette urimelig.

e) 2004: Predikert real import = -172,61 + 0,585 · 784,9 + 0,046 · 71,7 = **289,8**.
2005: Predikert real import = -172,61 + 0,585 · 799,6 + 0,046 · 72,7 = **298,5**.
De observerte verdier er hhv. 314,2 og 331,1, altså en **mindre god prediksjon**.

f) Alle koeffisienter er signifikante (lav p-verdi), i c) var ikke koeffisienten til de relative priser signifikant. R^2 er høy, så modellen har stor forklaringskraft . Merk at det ikke er faglig tillatelig å sammenligne R^2 på tvers av modeller der Y-variabelen er forskjellig. Siden vi nå har transformert med ln, vil b_1 og b_2 være **elastisiteten** til importen mht. GDP og relativer priser. Koeffisienten til inntekt (GDP) er 1,66, dvs. en en 1% økning i brutto nasjonalprodukt (inntekt eller GDP) fører til en 1,66% økning i importen. Dette er noe lavere enn det vi fant i d). Koeffisienten til de relative priser er -0,41, dvs. en 1% økning i relativer priser fører til en 0,41% **nedgang** i importen. Dette virker mer rimelig enn det vi fant i c) og d).

g) 2004: Predikert ln (real import) = -3,60 + 1,66 ·ln(784,9) - 0,41 ·ln(71,7) = 5,73. Dette gir predikert real import = exp(5,73) = **308,2**.
2005: Predikert ln(real import) = -3,60 + 1,66 · ln(799,6) – 0,41· ln(72,7) = 5,76. Dette gir predikert real import = exp(5,76) = **316,0**. Dette er bedre prediksjoner enn de i e).
[Merk at rett prediksjon fra en ln-transformasjon krever at man egentlig tar hensyn til estimert standardavvik, vi forventer ikke at studentene skal utdype dette.]

h) H_0: Ingen autokorrelasjon vs. H_1: Positiv autokorrelasjon. Durbin-Watsons testobservator er oppgitt i utskrift fra Minitab, d = **0,855**. Forkastningsområde er $d < d_L$ = **1,30** (fra tabell med n = 31, k =2 og signifikansnivå 5%). **Forkast dermed H_0.**

i) Innfør en dummyvariabel D_t som er lik 0 for årene 1973-1979, og er lik 1 fra og med 1980 og utover. Modellen blir da: $\ln(m_t) = b_0 + b_1 \cdot \ln(gdp_t) + b_2 \cdot \ln(pm_t) + \mathbf{b_3 \cdot D_t} + e_t$
For å teste om det var et skift, test nullhypotesen $b_3 = 0$ mot alternativet $b_3 \neq 0$ med en standard t-test.

Oppgave 3

a) Fra sannsynlighetsfordelingen *p(x)* ser vi at $p(x\text{-}1) = \dfrac{\mu^{x-1}}{(x-1)!} e^{-\mu}$

Dette uttrykket kan vi sette inn på høyre side i formelen (1),

$$\frac{\mu}{x} p(x\text{-}1) = \frac{\mu}{x} \frac{\mu^{x-1}}{(x-1)!} e^{-\mu} = \frac{\mu^x}{x!} e^{-\mu} = p(x)$$

Vi ser dermed at formelen (1) gjelder.

b) For μ=1: $p(0) = \dfrac{1^0}{0!} e^{-1} = e^{-1} = \underline{0{,}37}$, benytter så (1) for de resterende sannsynligheter

$$p(1) = \frac{1}{1} e^{-1} = e^{-1} = \underline{0{,}37},\ p(2) = \frac{1}{2} e^{-1} = \frac{1}{2} e^{-1} = \underline{0{,}18},\ p(3) = \frac{1}{3}\frac{1}{2} e^{-1} = \frac{1}{6} e^{-1} = \underline{0{,}06}$$

c) Fra formelen (1) ser vi at når x = μ, er p(x) = p(x-1). Når x > μ, vil p(x) bli mindre, mens når x < μ, vil p(x) gå mot μ·p(0) (da formelen kun gjelder fra x = 1). Fra tabellen og beregningen ser vi at p(x) har to maksima når x = μ og x = μ-1 (når μ er et heltall). Når μ ikke er et heltall, ligger maksima i nærheten av μ. [Teoretisk sett har ikke dette maksimeringsproblemet en eksakt løsning].

d) $E(Y) = q \cdot F(q) - \mu \cdot F(q-1) - (q-\mu) = 12 \cdot F(12) - 10 \cdot F(11) - (12 - 10) = 12 \cdot 0,7916 - 10 \cdot 0,6968 - 2 =$ **0,531**, der tabell er benyttet for den kumulative Poissonfordeling med $\mu = 10$ og $k = 11$ og 12.

e) Steg i en simulering:

   a. Velg en verdi for q og μ (for eksempel $q = 15$ og $\mu = 10$)
   b. Trekk n (for eksempel 1000) verdier av X svarende til Poissonfordeling med $\mu = 10$. Kaller disse verdiene $x_1, \ldots, x_{1000}$.
   c. For hver av de n trekningene fra b., la $y_i = 0$ om $x_i \leq 15$ eller $y_i = x_i - 15$ om $x_i > 15$. Har da 1000 verdier av Y; $y_1, \ldots, y_{1000}$.
   d. Beregn gjennomsnittet til de 1000 verdiene av Y funnet i c. Dette gjennomsnittet kan vi så sammenligne med E(Y) fra formelen. Om det beregnede gjennomsnittet er i samsvar med svaret fra formelen, kan vi slutte at formelen synes korrekt.

[Godkjenn også om besvarelsen består av Minitab-kode]