Viðskipta- og Hagfræðideild
Háskóli Íslands

Tölfræði II, 04.51.09-960
Skriflegt próf

14. desember 2005, 13.30-16.30.

Leyfileg hjálpargögn: Skriffæri, strokleður

Rökstyðjið öll svör og gerið grein fyrir forsendum. Þar sem til teljið vanta forsendur gefið ykkur það sem þið teljið nauðsynlegt.

1. (5%) Hending X hefur þéttifall:

$$f(x) = cx^{\alpha} \quad \alpha > -1, 0 < x < 1$$

   Finnið fastann c. Finnið dreifingu $Y = 1/X$.

2. (15%) Gefið að $X_1, \ldots, X_n$ eru óháðar einsdreifðar með þéttifall:

$$f(x) = cx^{\alpha} \quad \alpha > -1, 0 < x < 1$$

   Leiðið út metil mesta sennileika (MLE=maximum-likelihood-estimator) fyrir α. Leiðið út MM (method-of-moments) metil fyrir α. Leiðið út upplýsingafallið (Information fall) fyrir α. Hversu nákvæmlega er hægt að álykta um α út frá random úrtaki af n mælingum. Giskið á nafn á dreifingu MLE-metilsins. Hver dreifing MLE-metilsins í stórum úrtökum?

3. (10%) Gefið að $X_1, \ldots, X_n$ eru óháðar einsdreifðar með þéttifall:

$$f(x) = cx^{\alpha} \quad \alpha > -1, 0 < x < 1$$

   Reiknið LR (likelihood-ratio) prófstærð fyrir kenninguna $H_0 : \alpha = 0$.

4. (10%) Gefið er að hendingin X er normaldreifð með meðaltal μ og staðalfrávik σ. Runan Z_n er samsett úr óháðum hendingum með líkindadreifingu:

$$P(Z_n = j) = \begin{cases} \frac{n^{-2}}{2} & \text{ef } j = \pm n \\ 1 - \frac{1}{n^2} & \text{ef } j = 0 \\ 0 & \text{annars} \end{cases}$$

   Hendingarunan Y_n er skilgreind:

$$Y_n = X + Z_n \quad X \text{ óháð } Z_n$$

   Er $plim(Y_n)$ til? Er Y_n samleitin í dreifingu?

5. (5%) Skýrið lauslega hugmyndina að baki QQ-plot.

6. (10%) Fyrir hendi eru n óháð pör af mælingum, (x_i, y_i). Meta á halla línu sem fer í gegnum 0, $y_i = \beta x_i + \varepsilon_i$. Stungið er upp á tveim metlum:

$$\hat{\beta}_A = \frac{\sum_{i=1}^{n} y_i x_i}{\sum_{i=1}^{n} x_i^2} \quad \text{og } \tilde{\beta}_B = \frac{\bar{y}}{\bar{x}}$$

   Hvaða rök eru fyrir að velja $\hat{\beta}_A$ og hvaða rök eru fyrir því að velja $\tilde{\beta}_B$?

7. (15%) Gefið er ferli í tíma sem er myndað með:

$$X_t = X_{t-1} + \varepsilon_t \quad E(\varepsilon_t) = 0, E(\varepsilon_t^2) = 1, E(\varepsilon_t \varepsilon_s) = 0 \text{ ef } t \neq s$$

Hvað er algengt nafn á ferli sem hegðar sér eins og ε_t? Hvað er algengt nafn á ferli sem hegðar sér eins og X_t? Reiknið $E(X_t|X_0 = x_0)$ og $V(X_t|X_0 = x_0)$. Er ferlið X_t sístætt (stationary)? Rissið upp fræðilegt correlogram fyrir ε_t. Rissið upp fræðilegt correlogram fyrir X_t. Rissið upp dæmigert úrtakscorrelogram fyrir tímaröð sem er samsett úr T mælingum á X_t.

8. (5%) Hvers vegna er stundum áhugavert að reikna $\sum_{i=1}^{n}(x_i - \bar{x})^3$ og $\sum_{i=1}^{n}(x_i - \bar{x})^4$. Hvaða hugtök er verið að vinna með þegar þetta tvennt er reiknað?

9. (10%) Berið saman general-to-specific og specific-to-general aðferðirnar í haglíkanagerð.

10. (15%) (Þetta er klippt og límd hráútprentun úr GRETL-forritinu. Það á að sjálfsögðu ekki að skila frá sér svona mörgum aukastöfum eins og hér er sýnt). Breytan Q_t táknar eftirspurn eftir matvælum og breytan X_t táknar heildarútgjöld á föstu verði. Breyturnar eru mældar á tímabilinu 1966-1992 og tengslin metin með aðhvarfsgreiningu:

$$\widehat{\mathrm{log(Q_t)}} = \underset{(34.006)}{7.78759} + \underset{(19.325)}{0.308668}\, log(X_t)$$

$T = 27$ (t-statistics in parentheses)

Mean of dependent variable	12.2126
S.D. of dependent variable	0.0708852
Sum of squared residuals	0.00819686
Standard error of residuals ($\hat{\sigma}$)	0.0181073
Unadjusted R^2	0.937257
Degrees of freedom	25
Durbin–Watson statistic	0.881205
First-order autocorrelation coeff.	0.527364

Hér sér rannsakandi ástæðu til að reikna

$$e_t = log(Q_t) - \widehat{log(Q_t)}$$

Og metur eftirfarandi:

$$\Delta e_t = \phi e_{t-1} + u_t$$

sem gefur:

$$\widehat{\Delta \mathrm{e}_t} = \underset{(-2.908)}{-0.472636}\, \mathrm{e}_{t-1}$$

$T = 26$ (t-statistics in parentheses)

Mean of dependent variable	0.000697561
S.D. of dependent variable	0.0169829
Sum of squared residuals	0.00539722
Standard error of residuals ($\hat{\sigma}$)	0.0146932
Unadjusted R^2	0.253695
$F(1, 25)$	8.45755
p-value for $F()$	0.00752004
Durbin–Watson statistic	1.26983
First-order autocorrelation coeff.	0.359777

Næst ákveður rannsakandi að reikna $\hat{u}_t = \Delta e_t - \Delta \hat{e}_t$ og meta

$$\hat{u}_t = \gamma_0 + \gamma_1 e_{t-1} + \gamma_2 \hat{u}_{t-1} + \gamma_3 \hat{u}_{t-2} + v_t$$

sem gefur:

$$\widehat{\mathrm{u}}_t = \underset{(0.073)}{0.000186027} - \underset{(-1.804)}{0.401895}\,\mathrm{e}_{t-1} + \underset{(3.607)}{0.618055}\,\hat{u}_{t-1} + \underset{(1.054)}{0.221375}\,\hat{u}_{t-2}$$

$T = 24$ (t-statistics in parentheses)

Mean of dependent variable	4.59875e-05
S.D. of dependent variable	0.0150312
Sum of squared residuals	0.00309886
Standard error of residuals ($\hat{\sigma}$)	0.0124476
Unadjusted R^2	0.403666
Adjusted $\bar{R}^2$	0.314216
$F(3, 20)$	4.51275
p-value for $F()$	0.0142199
Durbin–Watson statistic	2.14519
First-order autocorrelation coeff.	−0.0798099

Þegar rannsakandi sér þessa útkomu telur hann að það sé skynsamlegt að meta:

$$\Delta e_t = \phi e_{t-1} + \beta \Delta e_{t-1} + v_t$$

sem gefur:

$$\Delta\widehat{\mathrm{e}}_t = \underset{(-4.674)}{-0.713526}\,\mathrm{e}_t + \underset{(3.817)}{0.601254}\,\Delta\ \mathrm{e}_{t-1}$$

$$T = 25 \quad \bar{R}^2 = 0.5057 \quad F(2, 23) = 12.77 \quad \hat{\sigma} = 0.011927$$

(t-statistics in parentheses)

Hann notar þetta til að reikna $u_t^* = \Delta e_t - \Delta\widehat{e}_t$ og metur

$$\hat{u}_t^* = \gamma_0 + \gamma_1 e_{t-1} + \gamma_2 \Delta e_{t-1} + \gamma_3 \hat{u}_{t-1}^* + \gamma_4 \hat{u}_{t-2}^* + v_t$$

sem gefur:

$$\widehat{\mathrm{u}}_t^* = \underset{(0.120)}{0.000317662} - \underset{(-0.271)}{0.0643087}\,\mathrm{e}_{t-1} - \underset{(-0.778)}{0.165948}\,\Delta\mathrm{e}_{t-1} + \underset{(1.722)}{0.463009}\,\mathrm{u}_{t-1}^* - \underset{(-1.537)}{0.570089}\,\mathrm{u}_{t-2}^*$$

$$T = 23 \quad R^2 = 0.166 \quad F(5,17) = 0.67959 \quad \hat{\sigma} = 0.012651$$

(t-statistics in parentheses)

Nú virðir rannsakandinn með velþóknun fyrir sér töluna -4.674. Til öryggis ákveður hann að meta einnig líkanið:

$$\Delta e_t = \phi e_{t-1} + \beta_1 \Delta e_{t-1} + \beta_2 \Delta e_{t-2} + z_t$$

Það gefur þær útkomur sem hann væntir (hversvegna), þ.e. að mat og t-gildi á ϕ breyttist lítið og því ákveður hann að meta líkanið:

$$\Delta log(Q_t) = \alpha_0 + \gamma(log(Q_{t-1}) - 7.79 - 0.309 log(X_{t-1})) + \text{ tafin gildi á } \Delta log(Q_t) \text{ og } \Delta log(X_t) + \varepsilon_t$$

Túlkið þau skref sem rannsakandi tók. Hvernig á að túlka γ í síðustu jöfnunni? Af hverju varð rannsakandinn glaður þegar hann sá töluna -4.674?

Aukaspurning (+): Hvers vegna er óhepplegt að þýða maximum-likelihood-estimator á íslensku sem metil hámarkslíkinda?

Lausn verður sett á heimasíðu námskeiðs að prófi loknu.