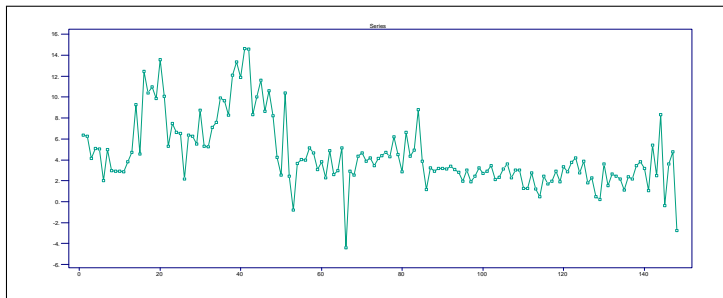


ARMA-mallien rakentaminen

Mallin riittävyyden tarkistaminen



Kuvio 4.1. USAn neljännesvuosittainen inflaatioisarja ajanjaksolta 1970I - 2006IV ($T = 148$). Alkuperäisen kausipuhdistetun kuluttajien hintaindeksin logaritmiset differenssit kerrottuna neljälläsadalla.

ARMA-mallien rakentaminen

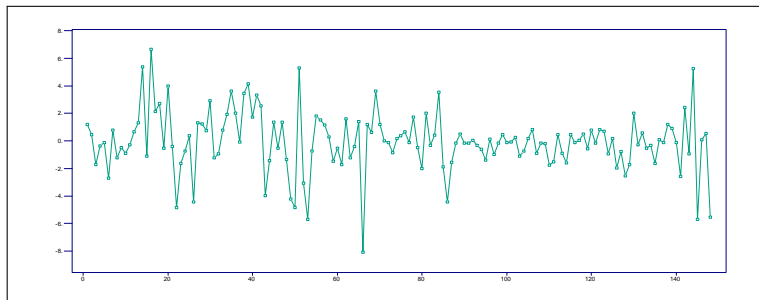
Mallin riittävyyden tarkistaminen

- Estimoidun autokorrelaatio- ja osittaisautokorrelaatiofunktion perusteella AR(3)-malli vaikutti sopivalta inflaationsarjalle.
- Hannanin ja Rissasen menetelmä valinnoilla $m = 10$ ja $p^* = q^* = 4$ johti myös AR(3)-malliin kaikilla kolmella sakkofunktiolla (samoin, kuin valinnoilla $m = 10$ ja $p^* = q^* = 5, \dots, 9$).
- SU-menetelmällä estimoitu AR(3)-malli inflaationsarjalle ($\bar{y}_t = y_t - 4.56$):

$$\bar{y}_t = \underset{(0.080)}{0.275}\bar{y}_{t-1} + \underset{(0.081)}{0.263}\bar{y}_{t-2} + \underset{(0.080)}{0.321}\bar{y}_{t-3} + \hat{\varepsilon}_t, \quad \hat{\sigma}^2 = 4.637,$$

ARMA-mallien rakentaminen

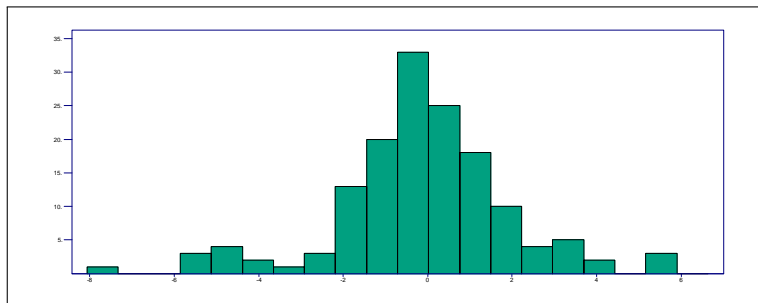
Mallin riittävyyden tarkistaminen



Kuvio 4.3(i). Inflaationsarjalle estimoidun AR(3)-mallin residuaalit ($\hat{\sigma} \approx 2.15$).

ARMA-mallien rakentaminen

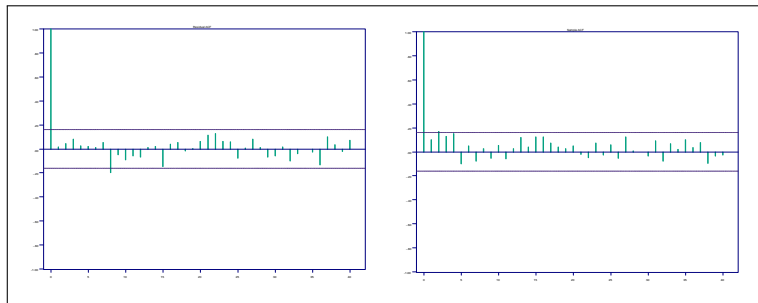
Mallin riittävyyden tarkistaminen



Kuvio 4.3(ii). Inflaationsarjalle estimoidun AR(3)-mallin residuaalien histogrammi.

ARMA-mallien rakentaminen

Mallin riittävyyden tarkistaminen



Kuvio 4.3(iii). Inflaatiotasjalle estimoidun AR(3)-mallin residuaalien otosautokorrelaatiofunktio (vasemmalla) ja neliöityjen residuaalien otosautokorrelaatiofunktio (oikealla) viipymillä $h = 0, \dots, 40$.

ARMA-mallien rakentaminen

Mallin riittävyyden tarkistaminen

- SU-menetelmällä estimoitu AR(3)-malli inflaatioasarjalle ($\bar{y}_t = y_t - 4.56$):

$$\bar{y}_t = \underset{(0.080)}{0.275}\bar{y}_{t-1} + \underset{(0.081)}{0.263}\bar{y}_{t-2} + \underset{(0.080)}{0.321}\bar{y}_{t-3} + \hat{\varepsilon}_t, \quad \hat{\sigma}^2 = 4.637$$

$$\text{AIC} = 656.25$$

- ARMA(4,0): $\hat{\phi}_4 = -0.045$ (0.084), AIC = 657.87
ARMA(3,1): $\hat{\theta}_1 = 0.106$ (0.214), AIC = 658.00.
- Siis, AIC (ja siten myös BIC ja HQ) suosii AR(3)-mallia.
- Ljung-Box: Testisuureen arvo = 14.74, kun $H = 15$ ja χ^2_{12} -jakaumasta saatava p-arvo 0.26.
- McLeod ja Li: Testisuureen arvo = 13.06, kun $H = 6$ ja χ^2_6 -jakaumasta saatava p-arvo 0.042.

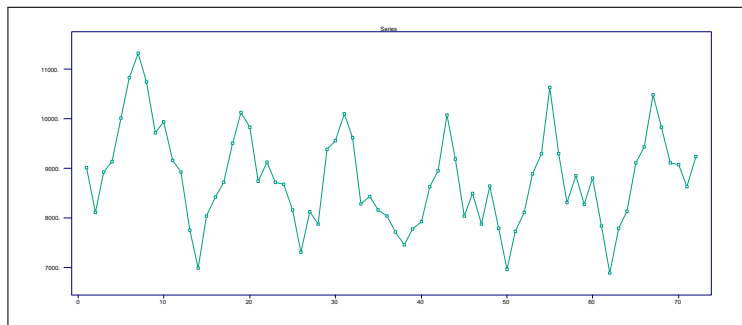
ARMA-mallien rakentaminen

Mallin riittävyyden tarkistaminen

- Johtopäätös residuaalitarkasteluista.
- Inflaatiosarjalle valittu AR(3)-malli pystyy kuvaamaan sarjan autokorrelaatorakenteen, mutta sen virheitä ei voida pitää normaalisina ja riippumattomuuttakin voidaan McLeodin ja Liin testin perusteella epäillä.
- Vaikka valittu malli ei olekaan optimaalinen, voi se kuitenkin tuottaa kohtuullisen hyviä ennusteita.
- Paremman mallin löytäminen on todennäköisesti vaikeaa tarkasteltujen ARMA-mallien ja niille esitettyjen menetelmien puitteissa.

ARMA-mallien rakentaminen

Kausivaihtelumallit

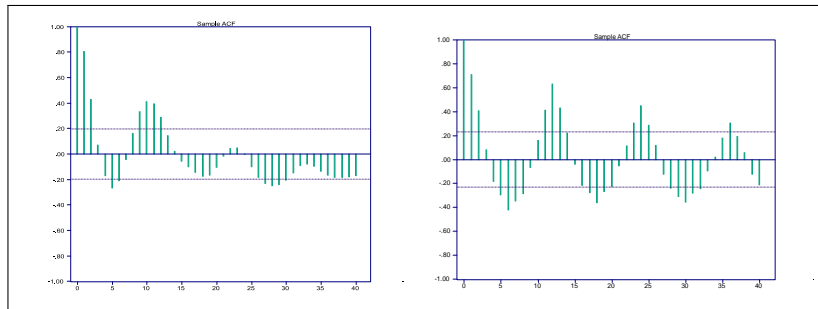


Kuvio 1.2. USA:ssa auto-onnettomuuksissa kuukausittain kuolleiden lukumäärä ajanjaksolta 1973I - 1978XII.

- Ei selvää nousevaa tai laskevaa trendiä.
- Sen sijaan selvä *kausivaihtelu*. Vuotuinen maksimi aina heinäkuussa ja minimi helmikuussa.

ARMA-mallien rakentaminen

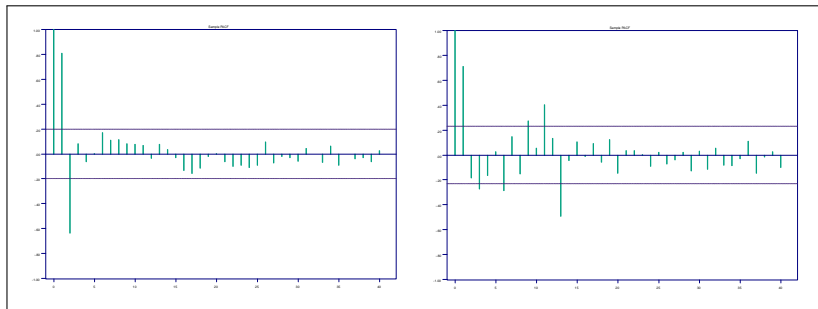
Kausivaihtelumallit



Kuvio 2.1. Kuvion 1.4 auringonpilkkusarjan otosautokorrelaatiofunktio (vas.) ja Kuvion 1.2 onnettomuussarjan otosautokorrelaatiofunktio (oik.) viipymillä $h = 0, \dots, 40$.

ARMA-mallien rakentaminen

Kausivaihtelumallit



Kuvio 3.2. Kuvion 1.4 auringonpilkkusarjan otosittaisautokorrelaatiofunktio (vas.) ja Kuvion 1.2 onnettomuussarjan otosittaisautokorrelaatiofunktio (oik.) viipymillä $h = 0, \dots, 40$.

- Kuvion 1.2 kausivaihtelua sisältävälle onnettomuussarjalle autokorrelaatio- ja osittaisautokorrelaatiofunktio viittaavat AR(13)-malliin, jossa kuitenkin paljon parametreja ($T = 72!$). Osa kertoimien ϕ_1, \dots, ϕ_{13} todellisista arvoista voi kuitenkin olla nollia.
- Estimoidaan ensin kaikki mahdolliset mallit, joista on vuorotellen poistettu yksi viipymistä y_{t-1}, \dots, y_{t-13} ja lasketaan kriteerifunktion (esim. BIC) arvo, jota verrataan AR(13)-mallista saatavaan arvoon.
- Jos AR(13)-mallista saatava kriteerifunktion arvo on pienin, valitaan malliksi AR(13). Muutoin poistetaan viipymistä se, jota vastaava 12 viipymän malli tuottaa suurimman kriteerifunktion arvon.
- Näin jatketaan poistaen aina viipymä kerrallaan, kunnes minkä tahansa viipymän poistaminen kasvattaa kriteerifunktion arvoa.

- Edellä kuvattu menettely & PNS & BIC ($\bar{y}_t = y_t - 879$):

$$\bar{y}_t = \underset{(0.093)}{0.661}\bar{y}_{t-1} + \underset{(0.066)}{0.835}\bar{y}_{t-12} - \underset{(0.099)}{0.576}\bar{y}_{t-13} + \hat{\varepsilon}_t, \quad \hat{\sigma}^2 = 115205.$$

- Residuaalitarkastelujen perusteella tämä malli on riittävä, vaikka AIC ja HQ ehdottavatkin hieman laajempia malleja (virheiden ei-normaalisuudesta varsin selviä viitteitä).
- Huom.: $-\hat{\phi}_1\hat{\phi}_{12} = -0.661 \times 0.835 = -0.552 \approx -0.576 = \hat{\phi}_{13}$.
- Jos rajoite $\phi_{13} = -\phi_1\phi_{12}$ oletetaan, päädytään malliin

$$(1 - \phi_1 B)(1 - \phi_{12} B^{12})(y_t - \mu) = \varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, \sigma^2),$$

- Rajoitettua SU-estimointia käyttäen saadaan (riittävä) malli

$$\left(1 - \underset{(0.076)}{0.763}B\right) \left(1 - \underset{(0.049)}{0.852}B^{12}\right) \bar{y}_t = \hat{\varepsilon}_t, \quad \hat{\sigma}^2 = 128227.$$

- Edellä esitetty malli

$$(1 - \phi_1 B) (1 - \Phi_1 B^{12}) y_t = \varepsilon_t, \quad \varepsilon_t \sim \text{iid} (0, \sigma^2)$$

on autoregressiivinen esimerkki ns. multiplikaatiivisesta ARMA-mallista, jollaisia käytetään paljon kausivaihtelun mallintamisessa.

- Jos kausivaihtelujakson pituus on yleinen s , niin tämän mallin yksinkertainen ARMA-yleistys on

$$(1 - \phi_1 B) (1 - \Phi_1 B^s) y_t = (1 + \theta_1 B) (1 + \Theta_1 B^s) \varepsilon_t$$

tai lyhyesti

$$\phi(B) \Phi(B^s) y_t = \theta(B) \Theta(B^s) \varepsilon_t, \quad \varepsilon_t \sim \text{iid} (0, \sigma^2).$$

- Kun edellä määritellyssä malliyhtälössä

$$\phi(B)\Phi(B^s)y_t = \theta(B)\Theta(B^s)\varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, \sigma^2)$$

asetetaan

$$\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$$

$$\Phi(B^s) = 1 - \Phi_1 B^s - \dots - \Phi_P B^{sP}$$

$$\theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q$$

$$\Theta(B^s) = 1 + \Theta_1 B^s + \dots + \Theta_Q B^{sQ},$$

saadaan ns. (multiplikatiivinen) $SARMA(p,q) \times (P,Q)_s$ -malli.

- Lisäksi oletetaan $p < s$ ja $q < s$. Asteita P ja Q ei rajoiteta, mutta käytännössä yleensä $0 \leq P, Q \leq 2$.

SARMA(p,q) × (P,Q)_s-malli saadaan ARMA(p+sP,q+sQ)-mallista

$$\phi^*(B) y_t = \theta^*(B) \varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, \sigma^2)$$

olettamalla polynomeille $\phi^*(B)$ ja $\theta^*(B)$ rakenne

$$\phi^*(B) = \phi(B) \Phi(B^s) \quad \text{ja} \quad \theta^*(B) = \theta(B) \Theta(B^s).$$

Vaihtoehtoinen ei-multiplikatiivinen kausivaihtelumalli on esimerkiksi ARMA(s+1,s+1)

$$(1 - \phi_1 B - \phi_s B^s - \phi_{s+1} B^{s+1}) y_t = (1 + \theta_1 B + \theta_s B^s + \theta_{s+1} B^{s+1}) \varepsilon_t,$$

jossa saadaan "tavallisesta" ARMA(s+1,s+1)-mallista asettamalla nollarajoitteet $\phi_2 = \dots = \phi_{s-1} = 0$ ja $\theta_2 = \dots = \theta_{s-1} = 0$.

SARMA(p,q) × (P,Q)_s-mallin

$$\phi(B) \Phi(B^s) y_t = \theta(B) \Theta(B^s) \varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, \sigma^2)$$

stationaarisuus- ja käännettävyys- ja identifioituvuusehto saadaan vastaavista ARMA-mallin ehdoista.

Stationaarisuus:

$$\phi(z) \Phi(z^s) \neq 0, \text{ kun } |z| \leq 1 \Leftrightarrow \phi(z) \neq 0 \text{ ja } \Phi(z^s) \neq 0, \text{ kun } |z| \leq 1$$

Käännettävyys:

$$\theta(z) \Theta(z^s) \neq 0, \text{ kun } |z| \leq 1 \Leftrightarrow \theta(z) \neq 0 \text{ ja } \Theta(z^s) \neq 0, \text{ kun } |z| \leq 1.$$

Identifioituvuus: Polynomeilla $\phi(z)$ ja $\theta(z)$ ei yhteisiä juuria eikä myöskään polynomeilla $\Phi(z^s)$ ja $\Theta(z^s)$ sekä $\phi_p \neq 0$ tai $\theta_q \neq 0$ ja $\Phi_P \neq 0$ tai $\Theta_Q \neq 0$.

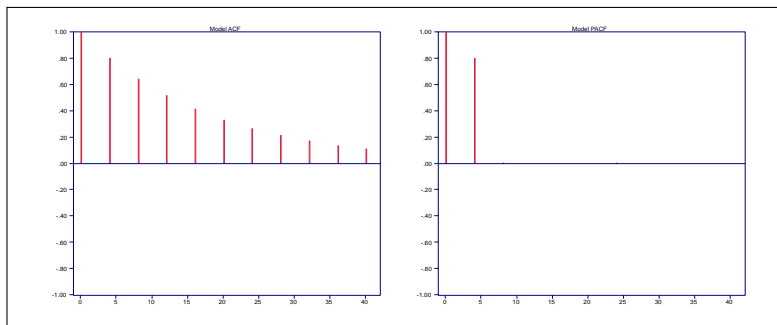
- Tarkastellaan nyt tapausta $s = 12$ ja SARMA(0,0) \times (1,1)₁₂-mallia

$$y_t = \Phi_1 y_{t-12} + \varepsilon_t + \Theta_1 \varepsilon_{t-12}, \quad \varepsilon_t \sim \text{iid}(0, \sigma^2).$$

- Jos tarkastellaan eri vuosien yksittäisiin kuukausiin liittyvien havaintojen mallinnusta, voidaan todeta, että *saman ARMA(1,1)-mallin oletetaan pätevän kaikille vuoden kuukausille.*
- Voi olla epärealistista olettaa, että esim. maaliskuun havainto y_t riippuisi menneisyydestä vain edellisen vuoden maaliskuun havainnon y_{t-12} ja virhetermin ε_{t-12} kautta eikä lainkaan saman vuoden helmikuun havainnosta y_{t-1} tai virhetermistä ε_{t-1} . SARMA(p,q) \times (1,1)₁₂-malli ottaa tällaisen riippuvuuden huomioon olettamalla, että edellä $\varepsilon_t \sim \text{ARMA}(p,q)$ -prosessia.

ARMA-mallien rakentaminen

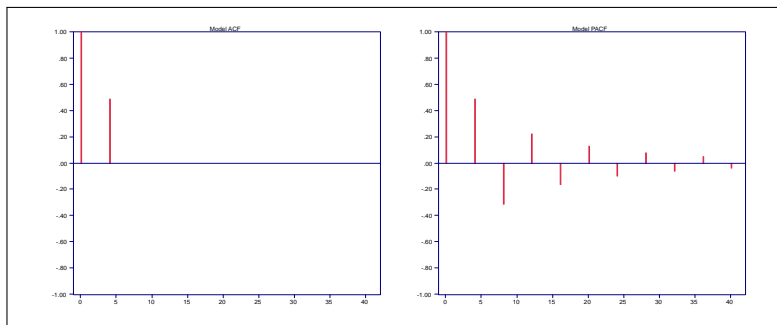
Kausivaihtelumallit



- Mallin $y_t = 0.8y_{t-4} + \varepsilon_t$, $\varepsilon_t \sim \text{nid}(0, 1)$ autokorrelaatiofunktio (vas.) ja osittaisautokorrelaatiofunktio (oik.) viipymillä $h = 0, \dots, 40$.

ARMA-mallien rakentaminen

Kausivaihtelumallit



- Mallin $y_t = \varepsilon_t + 0.8\varepsilon_{t-4}$, $\varepsilon_t \sim \text{nid}(0, 1)$ autokorrelaatiofunktio (vas.) ja osittaisautokorrelaatiofunktio (oik.) viipymillä $h = 0, \dots, 40$.

- Jos tarkasteltava aikasarja on epästationaarinen, mutta stationaarisuus saavutetaan differensoinnilla, määritellään multiplikatiivinen ARMA-malli differensoitua prosessia käyttäen kuten edellä.
 - Kausivaihtelutapauksessa differensointi voidaan suorittaa käyttäen joko operaattoria $\Delta = 1 - B$ tai vastaavaa kausivaihteluoperaattoria $\Delta_s = 1 - B^s$.
- Malli, joka sallii molemmat vaihtoehdot ja myös korkeammat differensoinnit, on SARIMA(p,d,q) × (P,D,Q)_s-malli

$$\phi(B) \Phi(B^s) \Delta^d \Delta_s^D y_t = \theta(B) \Theta(B^s) \varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, \sigma^2).$$

Käytännössä differensointiasteet d ja D ovat yleensä joko nolla tai yksi. Sovelluksia, joissa molemmat ovat ykkösiä, näkee kuitenkin varsin usein.

- Differensointiasteiden valinnan jälkeen SARIMA(p,d,q) × (P,D,Q)_s-mallin rakentaminen sujuu periaateessa samaan tapaan kuin tavallisen ARMA(p,q)-mallin.
 - Mallinvalinta on hankalampaa, koska kahden asteen asemesta täytyy valita neljä astetta.
- Jaksossa 4.4 esitetty SU-estimointi yleistyy niin, että parametrivektoriin β liitetään SARMA(p,q) × (P,Q)_s-mallin multiplikatiiviset rajoitteet.
 - Esimerkiksi,

$$(1 - \phi_1 B - \phi_{12} B^{12} - \phi_{13} B^{13})y_t = \varepsilon_t, \quad \varepsilon_t \sim \text{nid}(0, \sigma^2),$$

$$\boldsymbol{\phi} = (\phi_1, \dots, \phi_{13}) \text{ rajoitteet } \phi_{13} = -\phi_1 \phi_{12} \text{ ja } \phi_2 = \dots = \phi_{11} = 0.$$

Ehdollisen varianssin mallintaminen

- ARMA-malleja voidaan luonnehtia malleiksi stationaarisen prosessin ehdolliselle odotusarvolle.
- Käännettävyys \Rightarrow AR(∞)-esitys

$$y_t = - \sum_{j=1}^{\infty} \pi_j y_{t-j} + \varepsilon_t, \quad \varepsilon_t \sim \text{iid} (0, \sigma^2),$$

joten $[E_{t-1}(\cdot) = E(\cdot | y_{t-1}, y_{t-2}, \dots)]$

$$E_{t-1}(y_t) = - \sum_{j=1}^{\infty} \pi_j y_{t-j} + E_{t-1}(\varepsilon_t) = - \sum_{j=1}^{\infty} \pi_j y_{t-j}.$$

ja

$$\text{Var}_{t-1}(y_t) = \text{Var}_{t-1}(\varepsilon_t) = \text{Var}(\varepsilon_t) = \sigma^2.$$

- Siis, ehdollinen varianssi on vakio eli riippumaton ehtomuuttujista y_{t-1}, y_{t-2}, \dots

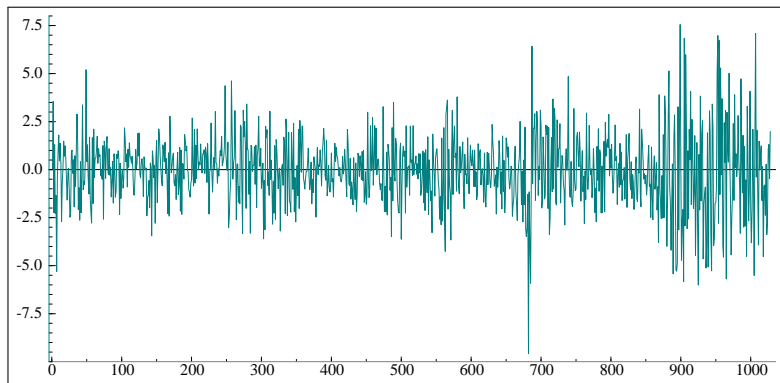
- ARMA-malleilla ei ilmeisestikään ole käyttöä tilanteissa, joissa havaittu aikasarja todetaan autokorreloimattomaksi.
- Tällaisissa aikasarjoissa voi kuitenkin olla riippuvuutta, kuten Esimerkkissä 2.2(ii):

$$y_t = \varepsilon_t \sqrt{\omega + \alpha \varepsilon_{t-1}^2}, \quad \varepsilon_t \sim \text{iid}(0, \sigma^2), \quad \omega > 0, \alpha > 0.$$

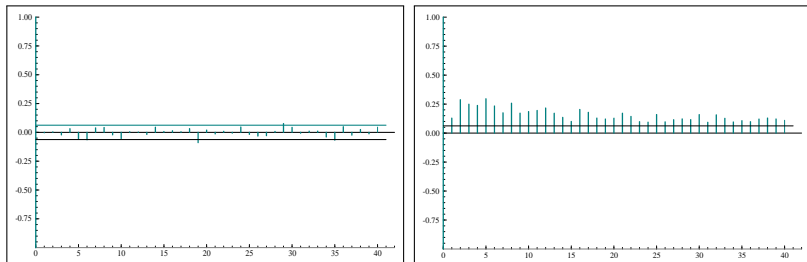
Koska $y_{t-1} = \varepsilon_{t-1} \sqrt{\omega + \alpha \varepsilon_{t-2}^2}$, on selvää, ettei y_t ole ajassa riippumaton.

- Erityisesti ehdollisessa varianssissa voi esiintyä systemaattista vaihtelua, mitä ei voida ottaa huomioon ARMA-mallilla.

Ehdollisen varianssin mallintaminen



Kuvio 5.1(i). Frankfurtin pörssin DAX-indeksin päivätuotot ajalta 1.1.1999 - 31.1.2003. Tuotot on laskettu kaavalla $y_t = 100 (\log P_t - \log P_{t-1})$, jossa P_t on indeksin arvo päivänä t .



Kuvio 5.1(ii). Frankfurtin pörssin DAX-indeksin päivätuotoista ajalta 1.1.1999 - 31.1.2003 laskettu otosautokorrelaatiofunktio (vas.) ja nelöiden otosautokorrelaatiofunktio (oik.) viiptymillä $h = 1, \dots, 40$.

Ehdollisen varianssin mallintaminen

Mallin yleinen rakenne

- Olkoon y_t (vahvasti ja heikosti) stationaarinen prosessi, jolle pätee $E_{t-1}(y_t) = 0$ [$E_{t-1}(y_t) = E(y_t | y_s, s < t)$].
- Yleinen malli y_t :n ehdolliselle varianssille on

$$y_t = h_t^{1/2} \varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, 1),$$

jossa

- h_t on funktio muuttujista $y_{t-j}, j > 0$ ja $\varepsilon_t \perp\!\!\!\perp \{y_{t-j}, j > 0\}$.
- Tällöin erityisesti $h_t \perp\!\!\!\perp \varepsilon_t$.

Ehdollisen varianssin mallintaminen

Mallin yleinen rakenne

- Yleinen malli

$$y_t = h_t^{1/2} \varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, 1),$$

h_t funktio muuttujista $y_{t-j}, j > 0$ ja $\varepsilon_t \perp\!\!\!\perp \{y_{t-j}, j > 0\}$

- Oletuksista seuraa

$$E_{t-1}(y_t) = 0 \quad \text{ja} \quad \text{Var}_{t-1}(y_t) = h_t.$$

- $h_t^{1/2}$ on siten y_t :n ehdollinen hajonta eli ns. *volatiliteetti*.

- Yleinen malli

$$y_t = h_t^{1/2} \varepsilon_t, \quad \varepsilon_t \sim \text{iid}(0, 1).$$

- Kun y_t on (vahvasti) stationaarinen, niin ehdollinen varianssi $h_t = h_t(y_{t-1}, y_{t-2}, \dots)$ myös (vahvasti) stationaarinen.
- Koska $\text{Var}(y_t) = E(y_t^2)$ ja $h_t \perp\!\!\!\perp \varepsilon_t$, saadaan

$$\sigma_y^2 \equiv \text{Var}(y_t) = E(h_t \varepsilon_t^2) = E(h_t)E(\varepsilon_t^2) = E(h_t).$$

- Samanlaiset laskelmat osoittavat, että y_t on autokorreloimaton eli

$$\text{Cov}(y_t, y_{t+k}) = 0, \quad \text{kaikilla } k \neq 0.$$