Stationaariset aikasarjat sl 2016, HT 5, viikko 41

1. Tarkastellaan epästationaarista prosessia $y_t = \alpha + \beta t + x_t$, $t = 1, 2, ...$, jossa x_t noudattaa stationaarista AR(1)-prosessia

$$x_t = \phi x_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \mathsf{iid}\left(0, \sigma^2\right), \quad |\phi| < 1.$$

Johda lausekkeet y_{t+h}:n optimaaliselle ennusteelle $\mathsf{E}_t\left(y_{t+h}\right) = \mathsf{E}\left(y_{t+h}\,|y_1, ..., y_t\right)$, ennustevirheelle $y_{t+h} - \mathsf{E}_t\left(y_{t+h}\right)$ $(h \geq 1)$ ja ennustevirheen varianssille $\mathsf{Var}(y_{t+h} - \mathsf{E}_t(y_{t+h})) \equiv \sigma_h^2$. Selvitä myös mitä tapahtuu ennustevirheen varianssille σ_h^2 ja oletukseen $\varepsilon_t \sim \mathsf{nid}\left(0, \sigma^2\right)$ perustuville ennusteiden luottamusrajoille, kun $h \to \infty$.

2. Tarkastellaan ARIMA(1,1,0)-prosessia $\Delta y_t = \phi \Delta y_{t-1} + \varepsilon_t, \quad t = 1, 2, ..., \varepsilon_t \sim \mathsf{iid}\left(0, \sigma^2\right), |\phi| < 1$.

(i) Osoita, että $\mathsf{E}_t\left(y_{t+h}\right) - \mathsf{E}_t\left(y_{t+h-1}\right) = \phi^h \Delta y_t$

(ii) Osoita, että $\mathsf{E}_t\left(y_{t+h}\right) = y_t + \dfrac{\phi\left(1 - \phi^h\right)}{(1 - \phi)} \Delta y_t$

(iii) Oletetaan nyt, että $\phi = 0$. Laske ennustevirheen $y_{t+h} - \mathsf{E}_t\left(y_{t+h}\right)$ odotusarvo ja varianssi. Mitä tapahtuu ennustevirheen varianssille ja edelleen ennusteen oletukseen $\varepsilon_t \sim \mathsf{nid}\left(0, \sigma^2\right)$ perustuville luottamusrajoille, kun $h \to \infty$?

Vihje: Yksi mahdollisuus kohdassa (ii) on käyttää identiteettiä $y_{t+h} = y_t + \Delta y_{t+1} + \Delta y_{t+2} + \cdots + \Delta y_{t+h}$.

3. Tarkastellaan prosessia

$$(1 - \phi B)\Delta y_t = \nu + \varepsilon_t, \;\; t = 1, 2, ..., \quad \varepsilon_t \sim \mathsf{iid}\left(0, \sigma^2\right), \quad |\phi| < 1.$$

(i) Johda $\mathsf{E}_t\left(y_{t+h}\right)$:n lauseke.

(ii) Oletetaan nyt, että $\phi = 0$. Laske ennustevirheen $y_{t+h} - \mathsf{E}_t\left(y_{t+h}\right)$ odotusarvo ja varianssi. Mitä tapahtuu ennustevirheen varianssille ja edelleen ennusteen oletukseen $\varepsilon_t \sim \mathsf{nid}\left(0, \sigma^2\right)$ perustuville luottamusrajoille, kun $h \to \infty$?

Vihje: Huomaa, että tehtävän ARIMA-prosessin määrittely-yhtälö voidaan lausua myös muodossa $(1 - \phi B)\Delta(y_t - \beta t) = \varepsilon_t$, jossa $\beta = \nu / (1 - \phi)$.

Seuraavassa tehtävässä havainnollistetaan ARIMA-prosessien ennusteita simuloimalla (piste-ennusteiden lisäksi näet 80%:n ja 95%:n luottamusvälit). Koska R:stä ei löytynyt koodia, jolla ennusteet voisi laskea annetuilla parametriarvoilla, käytetään todellisten parametriarvojen paikalla estimaatteja. Tehtävissä annetaan tilanteeseen sopivat R-koodit. Koodit vaativat R:n paketit `forecast` ja `portes` (lisätietoa komennolla `?forecast`).

4. (i) Simuloi stationaarisesta AR(2)-prosessista

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim \mathsf{nid}\left(0, 1\right),$$

(esim.) 200:n havainnon sarjoja ja ennusta (esim.) 50 periodin päähän. Koodi:

```
y <- numeric(200)
y <- arima.sim(list(order = c(2,0,0), ar = c(ϕ1,ϕ2)),n = 200)
fit <- arima(y, order = c(2,0,0), include.mean = FALSE)
fit
plot(forecast(fit,h = 50))
```

Ehdotus parametrivalinnoiksi: $(\phi_1, \phi_2) = (0.6, 0.3)$, $(-0.6, 0.3)$

(ii) Simuloi stationaarisesta MA(4)-prosessista

$$y_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \theta_3\varepsilon_{t-3} + \theta_4\varepsilon_{t-4}, \quad \varepsilon_t \sim \mathsf{nid}\,(0,1)$$

(esim.) 200:n havainnon sarjoja ja ennusta (esim.) 50 periodin päähän. Koodi:

```
y <- numeric(200)
y <- arima.sim(list(order = c(0,0,4),ma = c(θ1,θ2,θ3,θ4)),n = 200)
fit <- arima(y, order = c(0,0,1), include.mean = FALSE)
fit
plot(forecast(fit,h = 50))
```

Ehdotus parametrivalinnoiksi: $(\theta_1, \theta_2, \theta_3, \theta_4) = (0.3, 0.2, 0.2, 0.2)$, $(-0.3, -0.2, -0.2, -0.2)$

(iii) Simuloi stationaarisesta ARMA(2,1)

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t + \theta_1\varepsilon_{t-1}, \quad \varepsilon_t \sim \mathsf{nid}\,(0,1)$$

(esim.) 200:n havainnon sarjoja ja ennusta (esim.) 50 periodin päähän. Koodi:

```
y <- numeric(200)
y <- arima.sim(list(order = c(2,0,1), ar = c(ϕ1,ϕ2), ma = c(θ1)),n = 200)
fit <- Arima(y, order = c(2,0,1), include.mean = FALSE)
fit
plot(forecast(fit,h = 50))
```

Ehdotus parametrivalinnoiksi: $(\phi_1, \phi_2) = (1.6, -0.7)$, $\theta_1 = 0.8$, -0.8

(iv) Simuloi epästationaarisesta ARIMA(1,1,0)-prosessista (ks. tehtävä 3)

$$(1 - \phi_1 B)\Delta y_t = \nu + \varepsilon_t, \;\; \varepsilon_t \sim \mathsf{iid}\,(0,1)$$

(esim.) 200:n havainnon sarjoja ja ennusta (esim.) 50 periodin päähän. Koodi:

```
y <- numeric(200)
y <- varima.sim(phi = c(ϕ1), d = 1, sigma = 1, n = 200, constant = ν,
trend = 0)
fit <- Arima(y, order = c(1,1,0), include.drift = TRUE)
fit
plot(forecast(fit,h = 50))
```

Ehdotus parametrivalinnoiksi: $(\phi_1, \nu) = (0.8, 1)$, $(0.8, -1)$

(v) Simuloi epästationaarisesta tehtävän 1 tyyppisestä (ns. trendistationaarisesta) prosessista

$$y_t = \alpha + \beta t + x_t, \quad x_t = \phi_1 x_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \mathsf{iid}\,(0, 1)$$

(esim.) 200:n havainnon sarjoja ja ennusta (esim.) 50 periodin päähän. Koodi:

```
y <- numeric(200)
y <- varima.sim(phi = c(ϕ1), d = 0, sigma = 10, n = 200, constant = a,
trend = b)
fit <- Arima(y, order = c(1,0,0), include.drift = TRUE)
fit
plot(forecast(fit,h = 50))
```

Ehdotus parametrivalinnoiksi: $\phi_1 = 0.8$, $a = 1$, $b = 0.5$, 0.1

Huom.: Edellä olevat esimerkit on pyritty valitsemaan (jossain mielessä) "edustavasti". Parametriarvoja muuttamalla voit kokeilla vaihtoehtoja. Samoja parametriarvoja kokeilemalla voi myös saada erinäköisiä kuvia.