Aikasarjat ja ennustaminen

Jussi Klemelä

2015, toukokuu

Sisältö

1) Ennustaminen lineaarisen regression avulla
   - Osaketuoton ennustaminen osinkotuoton avulla
   - Onko ennustettavuutta olemassa?

2) Ennustaminen liukuvan keskiarvon avulla
   - Osaketuoton volatiliteetin ennustaminen

Osa I: Ennustaminen lineaarisen regression avulla

Ennustaminen lineaarisen regression avulla

- Havaitaan $Y_1, \ldots, Y_t \in \mathbf{R}$ ja $X_1, \ldots, X_t \in \mathbf{R}$.

- Ennustetaan Y_{t+s}, $s \geq 1$, käyttäen ennustetta

$$\hat{Y}_{t+s} = \hat{\beta}_0 + \hat{\beta}_1 X_t,$$

missä $\hat{\beta}_0$ ja $\hat{\beta}_1$ minimoivat neliöpoikkeamien summan

$$\sum_{i=1}^{t-s} (Y_{i+s} - \beta_0 - \beta_1 X_i)^2.$$

Osaketuoton ennustaminen osinkotuoton avulla

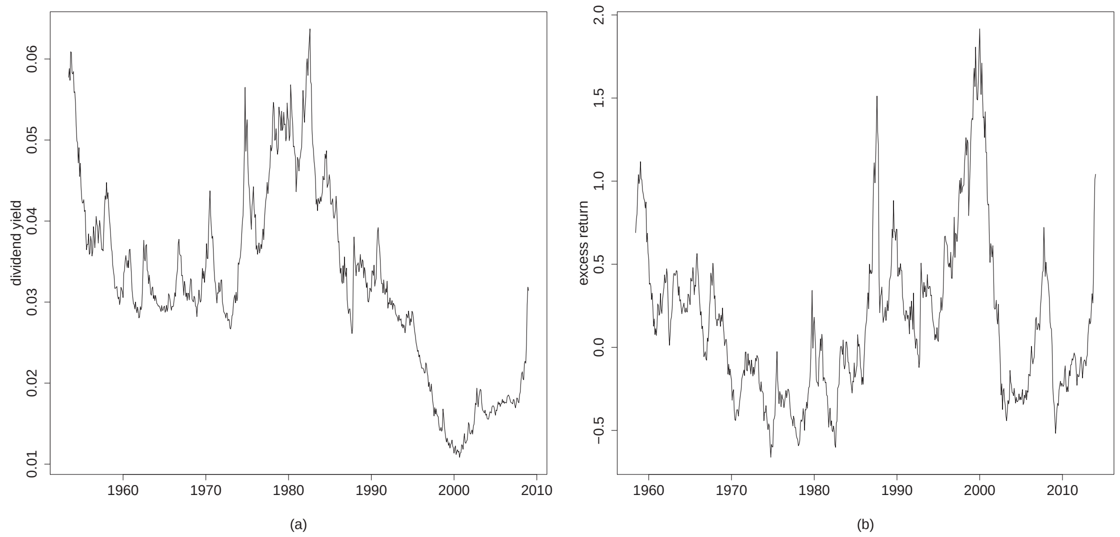

Figure 1: Aikasarjat: (a) Ennustava muuttuja X_t; (b) ennustettava muuttuja Y_{t+1}.

- $X_t = D_t/S_t$ on osinkotuotto, missä S_t on S&P 500 indeksin arvo ja D_t on osinkojen dollarimäärä 12 edellisen kuukauden aikana indeksin osakkeille.

- Y_{t+s} = 5 vuoden ylijäämätuotto S&P 500 indeksille: $Y_{t+s} = S_{t+5v}/S_t - r_t^f$, missä r_t^f on 5 vuoden riskitön tuotto.

Osaketuottojen ennustaminen

- 1960–1970: Osakkeiden hinta on geometrinen satunnaiskävely.

- 1980–1990: Osakkeiden tuottoja voidaan ennustaa 3-5 vuoden horisontilla käyttäen esim. osinkotuottoa ennustavana muuttujana.

- Nobelin muistopalkinto taloustieteissä 2013 (Eugene Fama, Lars Peter Hansen, Robert Shiller).
  "High future returns are viewed as compensation for holding risky assets during unusually risky times."

Lineaarinen regressio: hajontakuvio ja ennustukset

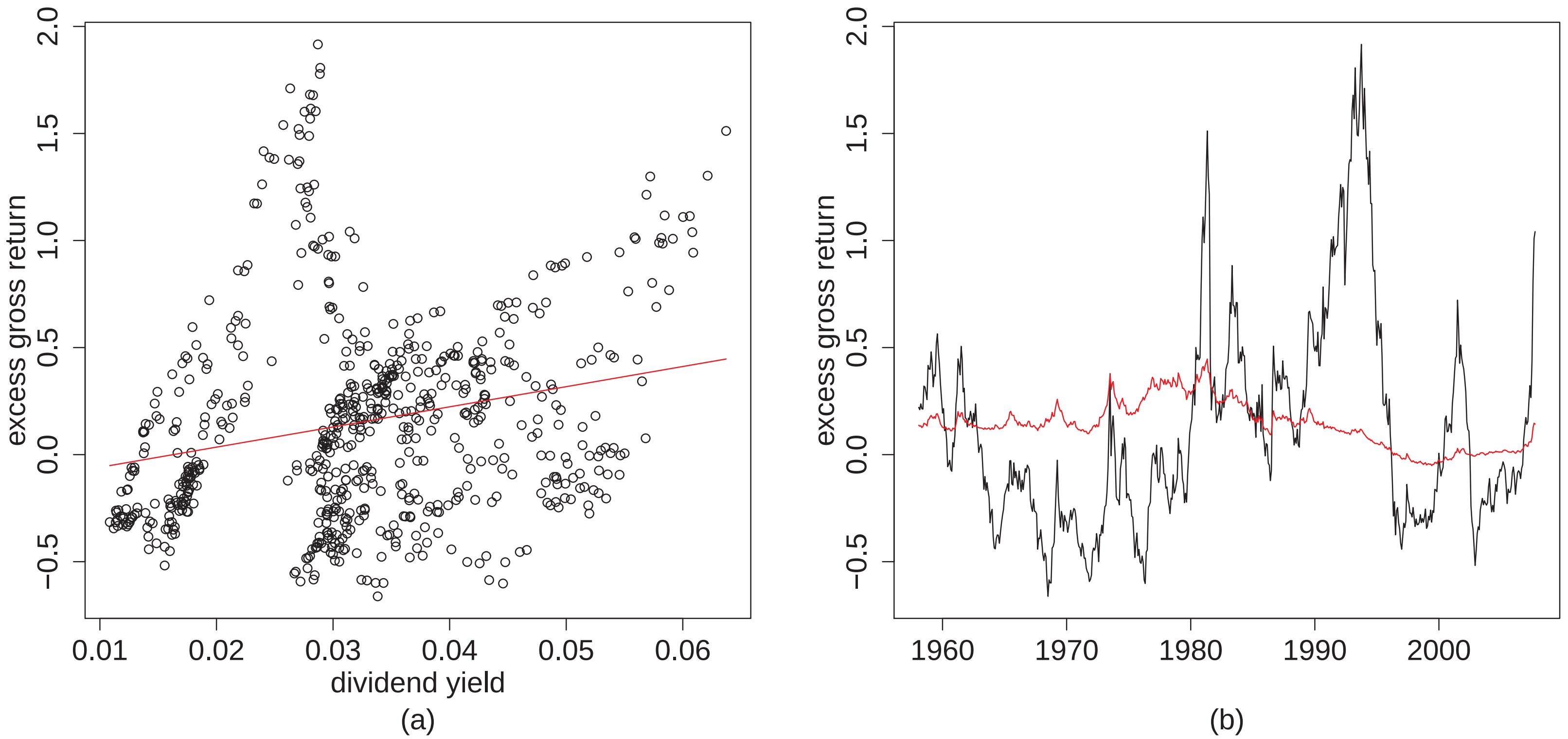

Figure 2: Lineaarinen regressio: (a) Aineiston hajontakuvio ja regressiofunktio; (b) Ennusteet (punainen) ja toteutuneet (musta).

Onko ennustettavuutta olemassa?

- Havaitaan $Y_1, \ldots, Y_t$ ja $X_1, \ldots, X_t$. Olkoon ennuste $\hat{Y}_{i+s} = \hat{\beta}_0 + \hat{\beta}_1 X_i$. Ennustevirheiden neliösumma on

$$\text{SSPE}(\hat{Y}) = \sum_{i=1}^{t-s} \left(Y_{i+s} - \hat{Y}_{i+s} \right)^2 .$$

Ennustevirheiden neliösumma kun ennuste on keskiarvo:

$$\text{SSPE}(\bar{Y}) = \sum_{i=1}^{t-s} \left(Y_{i+s} - \bar{Y} \right)^2 , \qquad \bar{Y} = \frac{1}{t} \sum_{i=1}^{t} Y_i.$$

- Selitysaste on

$$R^2 = 1 - \frac{\text{SSPE}(\hat{Y})}{\text{SSPE}(\bar{Y})}.$$

Testataan hypoteesia $H_0 : R^2 = 0$.

- Huom! Määriteltiin “in-sample”-ennustusvirheiden neliösumma. Onko “out-of-sample”-ennustusvirheiden neliösumma parempi?

Osa II: Ennustaminen liukuvan keskiarvon avulla

Liukuva keskiarvo

- Havaitaan aikasarja $Y_1, \ldots, Y_t$. Ennustetaan Y_{t+1} liukuvalla keskiarvolla

$$\hat{Y}_{t+1} = \frac{1}{h+1} \sum_{i=t-h}^{t} Y_i, \qquad h = 0, 1, 2, \ldots, t-1.$$

- Käytetään eksponentiaalisesti painotettua liukuvaa keskiarvoa

$$\hat{Y}_{t+1} = \frac{1-\gamma}{1-\gamma^t} \sum_{i=1}^{t} \gamma^{t-i} Y_i = \frac{1-\gamma}{1-\gamma^t} \Big(Y_t + \gamma Y_{t-1} + \gamma^2 Y_{t-2} + \cdots + \gamma^{t-1} Y_1 \Big), \qquad 0 \le \gamma < 1.$$

- Yleinen määritelmä:

$$\hat{Y}_{t+1} = \sum_{i=1}^{t} p_i(t)\, Y_i,$$

missä $\sum_{i=1}^{t} p_i(t) = 1$ ja $p_t(t) \ge p_{t-1}(t) \ge \cdots \ge p_1(t)$.

Esimerkki: Volatiliteetin ennustaminen

- Volatiliteetin ennustamisella tarkoitetaan arvon Y_{t+1}^2 ennustamista havaintojen $Y_1, \ldots, Y_t$ perusteella, kun $Y_t = S_t / S_{t-1} - 1$ on osakkeen tuotto, missä S_t on osakkeen hinta.

- Käytetään liukuvaa keskiarvoa:

$$\hat{Y}_{t+1}^2 = \sum_{i=1}^{t} p_i(t)\, Y_i^2,$$

missä $p_i(t) = (1 - \gamma)/(1 - \gamma^t)\, \gamma^{t-i}$ ja $0 \leq \gamma < 1$.

- Ennuste on lähellä GARCH(1,1) mallista saatavaa ennustetta.

S&P 500 volatiliteetin ennustaminen

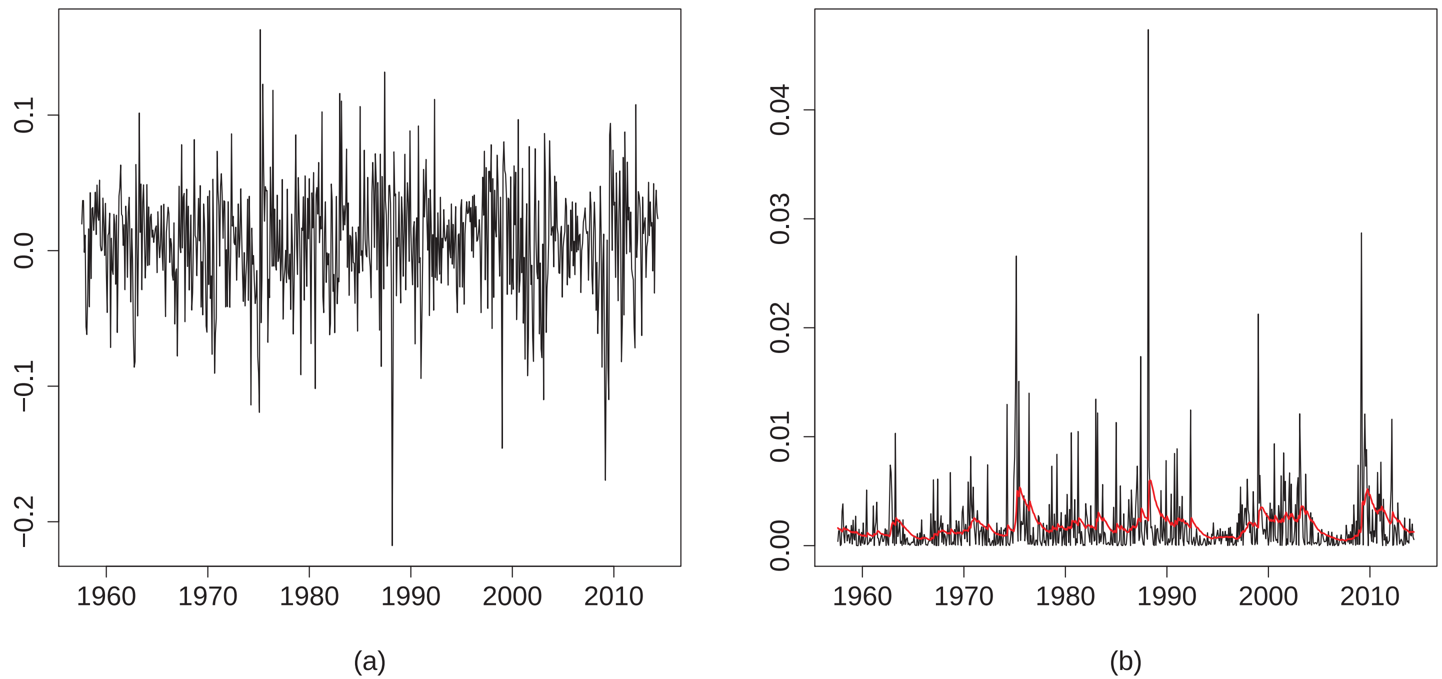

Figure 3: (a) Tuottojen $Y_t = S_t/S_{t-1} - 1$ aikasarja. (b) Ennusteiden aikasarja $\hat{Y}^2_{t+1}$ punaisella ja toteutuneitten aikasarja Y^2_{t+1} mustalla.

Yhteenveto

- Aikasarjan tulevia arvoja voidaan ennustaa esimerkiksi regressiolla ja liukuvalla keskiarvolla. Lisäksi arvoja voidaan ennustaa olettamalla aikasarjamalli (ARMA-malli, GARCH-malli,. . .).

- Ennustamista voidaan tutkia ainakin kahdella tavalla.
  (1) Ensin määritellään ennustusmenetelmä ja sitten tutkitaan mille aineistoille ja missä malleissa menetelmä toimii.
  (2) Ensin määritellään malli ja sitten tutkitaan mille aineistoille malli sopii ja mikä on paras ennustusmenetelmä tässä mallissa.