Egenskaper til estimatorene

To "triks":

Forventningsverdi $\hat{\beta}$

① Viser at $\sum_{i=1}^{n}(x_i-\bar{x}) = 0$

$$\sum_{i=1}^{n}(x_i-\bar{x}) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n}\bar{x}$$

$$= \sum_{i=1}^{n} x_i - n\bar{x}$$

$$\overset{(\mathrm{I})}{=} \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i = 0$$

$(\mathrm{I})\ \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

② Viser likhet mellom $\sum_{i=1}^{n}(x_i-\bar{x})y_i$ og $\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$.

Lettere å forkorte:

$$\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) = \sum_{i=1}^{n}(x_iy_i - x_i\bar{y} - y_i\bar{x} + \bar{y}\bar{x})$$

$$\stackrel{(I)}{=} \sum_{i=1}^{n} x_iy_i - \sum_{i=1}^{n} x_i\bar{y} - \sum_{i=1}^{n} y_i\bar{x} + \sum_{i=1}^{n}\bar{y}\bar{x}$$

$$\stackrel{(II)}{=} \sum_{i=1}^{n} x_iy_i - \bar{y}\sum_{i=1}^{n} x_i - \bar{x}\sum_{i=1}^{n} y_i + n\bar{x}\bar{y}$$

$$= \sum_{i=1}^{n} x_iy_i - \cancel{\bar{y}\cdot n\bar{x}} - \sum_{i=1}^{n} y_i\bar{x} + \cancel{n\bar{x}\bar{y}}$$

$$= \sum_{i=1}^{n} x_iy_i - \sum_{i=1}^{n} y_i\bar{x}$$

$$= \sum_{i=1}^{n}(x_iy_i - \bar{x}y_i) = \sum_{i=1}^{n}(x_i-\bar{x})y_i$$

(I) $\sum_{i=1}^{n}(a_i+b_i+c_i) = (a_1+b_1+c_1)+(a_2+b_2+c_2) + \ldots + (a_n+b_n+c_n)$
$= a_1+a_2+\ldots+a_n$
$+ b_1 + b_2 + \ldots + b_n$
$+ c_1 + \ldots + c_n$
$= \sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i + \sum_{i=1}^{n} c_i$

(II) $\sum_{i=1}^{n} a x_i = ax_1 + ax_2 + \ldots + ax_n = a(x_1+\ldots+x_n)$
$= a\cdot\sum_{i=1}^{n} x_i$

(III) $\sum_{i=1}^{n}\bar{x}x_i = \bar{x}\sum_{i=1}^{n} x_i = \bar{x}\cdot\frac{n}{n}\sum_{i=1}^{n} x_i = n\bar{x}\cdot\frac{1}{n}\sum_{i=1}^{n} x_i$
$= n\bar{x}\cdot\bar{x} = n\bar{x}^2$

Merk: dette gjelder også for $S_{xx} = \sum_{i=1}^{n}\overbrace{(x_i-\bar{x})(x_i-\bar{x})}^{(x_i-\bar{x})^2} = \sum_{i=1}^{n}(x_i-\bar{x})x_i$

og $S_{yy} = \sum_{i=1}^{n}\underbrace{(y_i-\bar{y})(y_i-\bar{y})}_{(y_i-\bar{y})^2} = \sum_{i=1}^{n}(y_i-\bar{y})y_i$

Eksamen H2020, 13B

Ikke "standard situasjon".

$(x_1, y_1), \ldots, (x_n, y_n)$

$Y_i \sim N(\beta \cdot \ln(x_i), \sigma^2 x_i^2)$ uavhengige

Ekvivalente representasjoner av modellen

$\Rightarrow Y_i = \beta \ln(x_i) + x_i \varepsilon_i \qquad \varepsilon_i \sim N(0, \sigma^2)$

$$E[Y_i] = E[\beta \cdot \ln(x_i) + x_i \varepsilon_i] = \beta \cdot \ln(x_i) + x_i E[\varepsilon_i] = \underline{\beta \cdot \ln(x_i)}$$

$$Var[Y_i] = Var[\beta \cdot \ln(x_i) + x_i \varepsilon_i] = 0 + x_i^2 Var[\varepsilon_i] = \underline{x_i^2 \sigma^2}$$

SME: $\hat{\beta}_{SME} = \dfrac{\sum_{i=1}^{n} Y_i \frac{\ln(x_i)}{x_i^2}}{\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}}$

utledning hvis vi har tid (kan vise start evt.)

Bestem $E[\hat{\beta}_{SME}]$ og $Var[\hat{\beta}_{SME}]$.

Hva er fordelingen til $\hat{\beta}_{SME}$?

$$E[\hat{\beta}_{SME}] = E\left[\frac{\sum_{i=1}^{n} Y_i \cdot \frac{\ln(x_i)}{x_i^2}}{\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}}\right]$$

$$\overset{\mathrm{I}}{=} \frac{1}{\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}} E\left[\sum_{i=1}^{n} Y_i \cdot \frac{\ln(x_i)}{x_i^2}\right]$$

$$\overset{\mathrm{II}}{=} \frac{1}{\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}} \sum_{i=1}^{n} E\left[Y_i \cdot \frac{\ln(x_i)}{x_i^2}\right]$$

$$\overset{\mathrm{I}}{=} \frac{1}{\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}} \sum_{i=1}^{n} \frac{\ln(x_i)}{x_i^2} E[Y_i]$$

Modellantakelse

$$= \frac{1}{\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}} \sum_{i=1}^{n} \frac{\ln(x_i)}{x_i^2} \ln(x_i) \cdot \beta$$

$$= \frac{\beta \cdot \cancel{\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}}}{\cancel{\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}}} = \beta$$

Forventningsrett

$$\mathrm{I} \quad E[aX] = aE[X]$$

$$\mathrm{II} \quad E[X+Y] = E[X] + E[Y]$$

$$\mathrm{III} \quad E[b] = b$$

$$\text{Var}[\hat{\beta}_{SME}] = \text{Var}\left[\frac{\sum_{i=1}^{n} Y_i \frac{\ln(x_i)}{x_i^2}}{\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}}\right]$$

$$\overset{I}{=} \left(\frac{1}{\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}}\right)^2 \text{Var}\left[\sum_{i=1}^{n} Y_i \frac{\ln(x_i)}{x_i^2}\right]$$

II + uavh.

$$= \left(\frac{1}{\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}}\right)^2 \sum_{i=1}^{n} \text{Var}\left[Y_i \frac{\ln(x_i)}{x_i^2}\right]$$

$$\overset{I}{=} \frac{1}{\left(\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}\right)^2} \sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^4} \text{Var}[Y_i]$$

Modellantakelse

$$= \frac{1}{\left(\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}\right)^2} \sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^{\cancel{4}\,2}} \sigma^2 \cancel{x_i^2}$$

$$= \frac{\sigma^2 \cancel{\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}}}{\left(\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}\right)^{\cancel{2}}} = \frac{\sigma^2}{\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}}$$

$\text{Var}[aX] = a^2 \text{Var}[X]$

$\text{Var}[X+Y] = \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X,Y]$

$\text{Var}[b] = 0$

$$\hat{\beta}_{SME} = \frac{\sum_{i=1}^{n} Y_i \frac{\ln(x_i)}{x_i^2}}{\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}}$$

$$= \sum_{i=1}^{n} \frac{\frac{\ln(x_i)}{x_i^2} Y_i}{\sum_{j=1}^{n} \frac{\ln(x_j)^2}{x_j^2}}$$

$$= \sum_{i=1}^{n} a_i Y_i \quad , \qquad a_i = \frac{\frac{\ln(x_i)}{x_i^2}}{\sum_{j=1}^{n} \frac{\ln(x_j)^2}{x_j^2}}$$

lineær kombinasjon
av uavhengige,
normalfordelte
Y_i, $i = 1, \dots, n$

$$\Rightarrow \hat{\beta}_{SME} \sim N\left(\beta, \frac{\sigma^2}{\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}}\right)$$

Utledning av SME: $Y_i \sim N(\beta \ln x_i \,, \sigma^2 x_i^2)$

$i = 1, 2, \ldots, n$ uavhengige

$X \sim N(\mu, \sigma^2)$

$\rightarrow f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$

Rimelighetsfunksjon (fra simultanfordeling av $Y_1, \ldots, Y_n$)

$$L(\underbrace{\beta, \sigma^2}_{\text{ukjente parametere}}) = f(y_1, \ldots, y_n ; \beta, \sigma^2)$$

$$\overset{\text{uavh.}}{=} f(y_1; \beta, \sigma^2) \cdot f(y_2; \beta, \sigma^2) \cdot \ldots \cdot f(y_n; \beta, \sigma^2)$$

$$= \prod_{i=1}^{n} f(\overset{\text{"x"}}{y_i} ; \beta, \sigma^2)$$

$$= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi \cdot \underbrace{\sigma^2 x_i^2}_{\text{"}\sigma^2\text{"}}}} \cdot e^{-\frac{1}{2} \cdot \frac{(y_i - \overbrace{(\beta \ln x_i)}^{\text{"}\mu\text{"}})^2}{\underbrace{\sigma^2 x_i^2}_{\text{"}\sigma^2\text{"}}}}$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n \left(\prod_{i=1}^{n} \frac{1}{\sqrt{\sigma^2 x_i^2}}\right) \cdot e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} \frac{(y_i - \beta \ln x_i)^2}{x_i^2}}$$

$$\ell(\beta, \sigma^2) = \ln L(\beta, \sigma^2) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2}\sum_{i=1}^{n} \ln(\sigma^2 x_i^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n} \frac{(y_i - \beta \ln x_i)^2}{x_i^2}$$

$$\frac{\partial \ell}{\partial \beta} = 0 + 0 + \frac{1}{\cancel{2}\sigma^2}\sum_{i=1}^{n} \frac{\cancel{2}(y_i - \beta \ln x_i) \cdot (\cancel{-}\ln x_i)}{x_i^2} = 0$$

$$\frac{1}{\sigma^2}\sum_{i=1}^{n} \frac{(y_i - \beta \ln x_i) \cdot \ln x_i}{x_i^2} = 0 \qquad | \cdot \sigma^2 \text{ på begge sider}$$

$$\Rightarrow \sum_{i=1}^{n} \frac{(y_i - \beta \ln x_i) \ln x_i}{x_i^2} = 0$$

$$\Rightarrow \sum_{i=1}^{n} \left(\frac{\ln x_i}{x_i^2} \cdot y_i - \frac{(\ln x_i)^2}{x_i} \beta\right) = 0$$

$$\Rightarrow \sum_{i=1}^{n} \frac{\ln x_i}{x_i^2} Y_i = \beta \sum_{i=1}^{n} \frac{(\ln x_i)^2}{x_i^2}$$

$$\beta = \frac{\sum_{i=1}^{n} \frac{\ln x_i}{x_i^2} y_i}{\sum_{i=1}^{n} \frac{(\ln(x_i))^2}{x_i^2}}$$

$$\Rightarrow \hat{\beta} = \frac{\sum_{i=1}^{n} \frac{\ln x_i}{x_i^2} Y_i}{\sum_{i=1}^{n} \frac{(\ln x_i)^2}{x_i^2}}$$

Diverse "regneregler" som dukker opp (med "bevis"):

$$\sum_{i=1}^{n} a = \underbrace{a+a+\ldots+a}_{n \text{ ganger}} = n \cdot a$$

$$\sum_{i=1}^{n} ax_i = ax_1 + ax_2 + \ldots + ax_n = a(x_1 + \ldots + x_n) = a\sum_{i=1}^{n} x_i$$

$$E\left[\sum_{i=1}^{n} a_i X_i\right] = E\left[a_1X_1 + a_2X_2 + \ldots + a_nX_n\right] = a_1E[X_1] + a_2E[X_2] + \ldots + a_nE[X_n]$$
$$= \sum_{i=1}^{n} a_i E[X_i]$$

$$Var\left[\sum_{i=1}^{n} a_i X_i\right] = Var\left[a_1X_1 + a_2X_2 + \ldots + a_nX_n\right] \overset{\text{uavh.}}{=} Var[a_1X_1] + Var[a_2X_2] + \ldots + Var[a_nX_n]$$
$$= a_1^2 Var[X_1] + \ldots + a_n^2 Var[X_n]$$
$$= \sum_{i=1}^{n} a_i^2 Var[X_i]$$

$$\prod_{i=1}^{n} a = \underbrace{a \cdot a \cdot a \cdot \ldots \cdot a}_{n \text{ ganger}} = a^n$$

$$\prod_{i=1}^{n} ax_i = ax_1 \cdot ax_2 \cdot \ldots \cdot ax_n = \underbrace{a \cdot a \cdot \ldots \cdot a}_{n \text{ ganger}} \cdot x_1 \cdot \ldots \cdot x_n = a^n \prod_{i=1}^{n} x_i$$

$$\prod_{i=1}^{n} a \cdot b^{x_i} = a \cdot b^{x_1} \cdot a \cdot b^{x_2} \cdot \ldots \cdot ab^{x_n} = a \cdot a \cdot \ldots \cdot a \cdot b^{x_1} \cdot b^{x_2} \cdot \ldots \cdot b^{x_n}$$
$$= a^n \, b^{x_1 + x_2 + \ldots + x_n}$$
$$= a^n \, b^{\sum_{i=1}^{n} x_i}$$

HVIS TID: Stack 13 oppg. 1a: Regn ut S^2 ved størrelsene oppgitt

$$S^2 = \frac{1}{n-2} \sum_{i=1}^{n} \left(y_i - (\hat{\alpha} + \hat{\beta} x_i)\right)^2$$

$$= \frac{1}{n-2} \sum_{i=1}^{n} \left(y_i - (\bar{y} - \hat{\beta}\bar{x} + \hat{\beta} x_i)\right)^2$$

$$= \frac{1}{n-2} \sum_{i=1}^{n} \left(y_i - \bar{y} - (x_i - \bar{x})\hat{\beta}\right)^2$$

$$= \frac{1}{n-2} \sum_{i=1}^{n} \left[(y_i - \bar{y})^2 - 2(y_i - \bar{y})(x_i - \bar{x})\hat{\beta} + (x_i - \bar{x})^2 \hat{\beta}^2\right]$$

$$= \frac{1}{n-2} \left[\sum_{i=1}^{n} (y_i - \bar{y})^2 - 2\hat{\beta} \sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x}) + \hat{\beta}^2 \sum_{i=1}^{n} (x_i - \bar{x})^2\right]$$

$$= \frac{1}{n-2} \left[S_{yy} - 2\hat{\beta} S_{xy} + \hat{\beta}^2 S_{xx}\right] = \frac{1}{n-2}\left[S_{yy} - 2 \cdot \frac{S_{xy}}{S_{xx}} S_{xy} + \left(\frac{S_{xy}}{S_{xx}}\right)^2 S_{xx}\right]$$

$$= \underline{\underline{\frac{1}{n-2}\left[S_{yy} - \frac{S_{xy}^2}{S_{xx}}\right]}}$$