第6章
定式化
練習問題

問 1

Y は年収(万円)、X_1 は女性ダミー(女性なら 1、男性なら 0)、X_2 は大卒ダミー(大卒以上なら 1、それ以外なら 0)とした結果、次の推定式が得られたと仮定しよう。

$$\hat{Y} = 25 - 5X_1 + 15X_2 + 5X_1X_2$$

ただし、$X_1 X_2$ は、X_1 と X_2 との交差項である。大卒男女の所得差(大卒男性の所得−大卒女性の所得)はいくつか。

1. 0 円
2. −5 万円
3. 15 万円
4. 40 万円

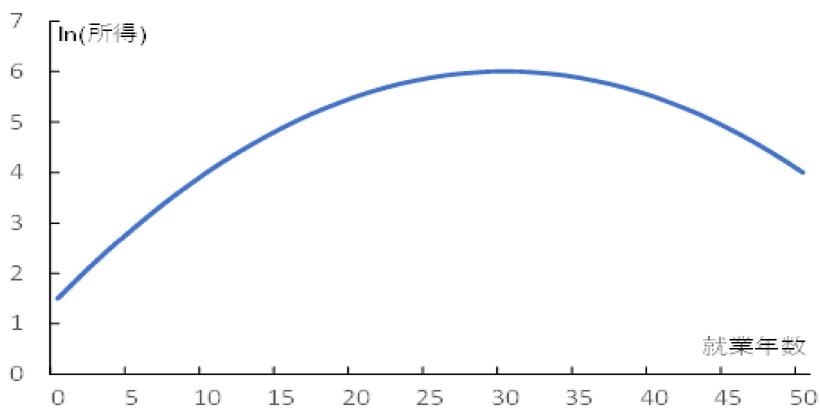
問 2

被説明変数を所得の対数、説明変数を就業年数、就業年数の 2 乗として推定した結果、次の式が得られた。

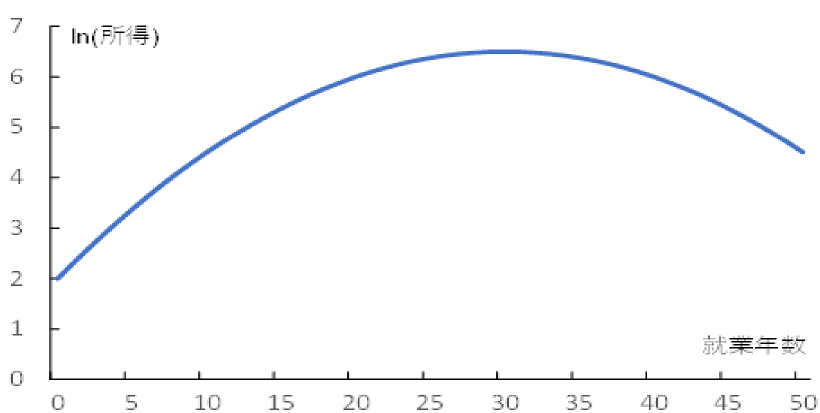
$$\ln(\text{所得}) = 1.5 + 0.3\ \text{就業年数} - 0.005\ \text{就業年数}^2$$

就業年数と ln(所得)の関係を描いた図として正しいものを選択しなさい。

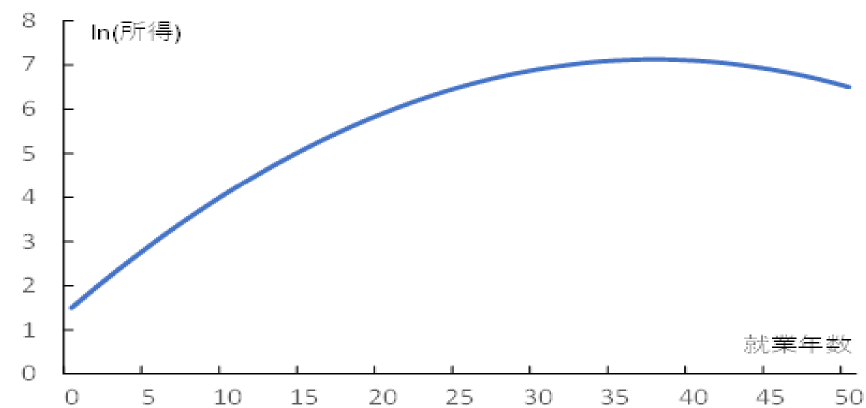
1.



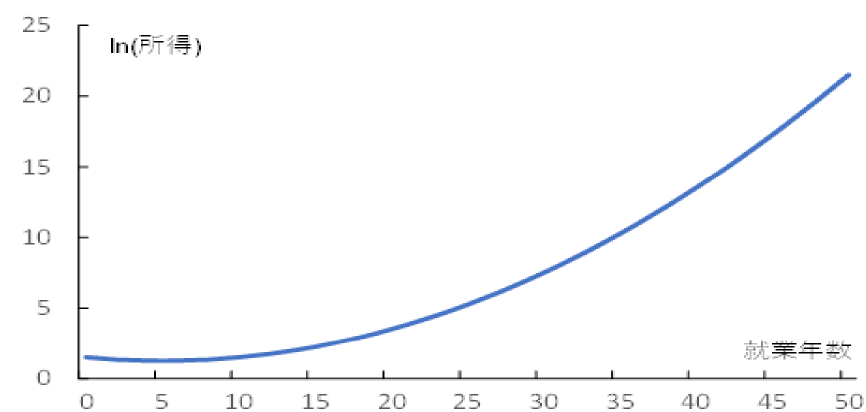
2.



3.



4.



問 3

体重(g)を Y、身長(cm)を X としたとき、$Y = \alpha + \beta X + u$ という関係がある。体重を kg 表示、身長を m 表示にした $Y^* = Y/1000$、$X^* = X/100$ を用いて、$Y^* = \alpha^* + \beta^* X^* + u^*$ という関係を分析したい。回帰係数 β と β^* との関係として正しいものを選択しなさい。

1. $\beta^* = \beta/10$
2. $\beta^* = \beta/1000$
3. $\beta^* = 1000\beta$
4. $\beta^* = \beta$

問 4

空欄に当てはまる記号もしくは数値を選びなさい。

BIC は、次式によって定義される。ただし、SSR は残差 2 乗和、T はサンプルサイズ、K は説明変数の数である。

$$\mathrm{BIC} = \ln\left(\frac{SSR}{T}\right) + (K+1)\frac{(\qquad\qquad)}{T}$$

1. 2
2. T
3. $\ln(T)$
4. $\ln(T-1)$

問 5

分布ラグモデルを推定した結果、BIC の値は p=1 なら 1000、p=2 なら 400、p=3 なら 1200、p=4 なら 1400 となる。p の推定値はいくつか。

1. p は 1 と推定される
2. p は 2 と推定される
3. p は 3 と推定される
4. p は 4 と推定される

問 6

<複数回答>被説明変数を所得の対数、説明変数を教育年数、就業年数、就業年数の 2 乗として推定した。

$$\ln(\text{所得}) = 3.5 + 0.13\ \text{教育年数} + 0.2\ \text{就業年数} - 0.005\ \text{就業年数}^2$$

推定結果の解釈として正しい記述をすべて選択しなさい。

1. 教育年数が 1 年増えると、所得は 13%上昇する。
2. 就業年数が増えると、当初は所得が大きく増えるが、徐々に所得の増加は小さくなる。
3. 大学新卒(教育年数 16 年、就業年数 0 年)なら、所得の対数は 5.5 となる。
4. 教育年数が 1 年増えると、所得は 0.13%上昇する。

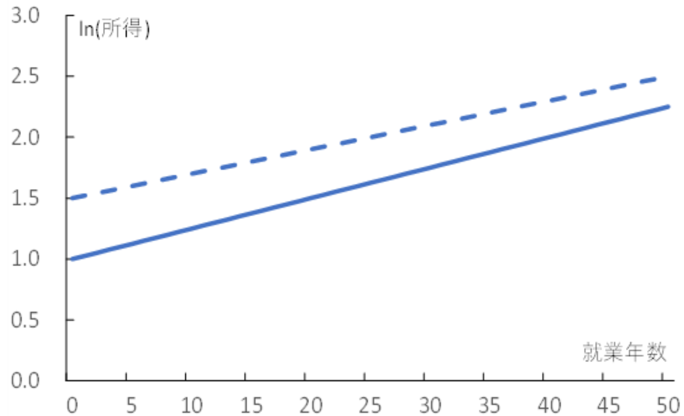
問 7

被説明変数を所得の対数、説明変数を就業年数、女性ダミー(女性なら 1 となるダミー変数)、女性ダミー×就業年数として推定した結果、次の式が得られたとしよう。

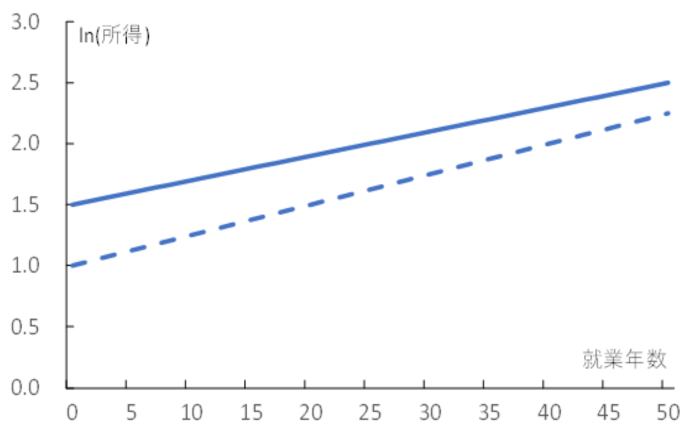
$$\ln(\text{所得}) = 1.5 + 0.02\ \text{就業年数}\ -\ 0.5\ \text{女性ダミー}\ + 0.005\ \text{女性ダミー} \times \text{就業年数}$$

就業年数と ln(所得)の関係を描いた図を選択しなさい。なお、実線が女性、点線が男性の関係式としている。

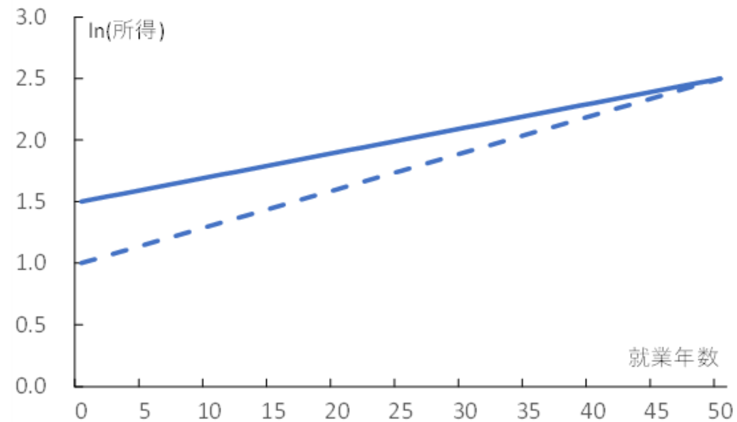
1.



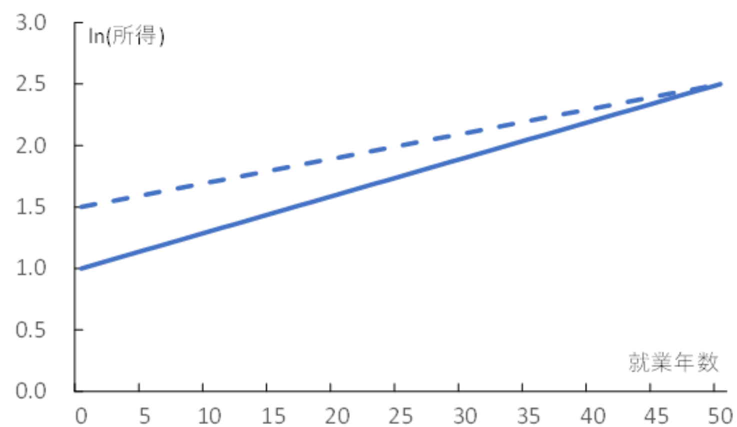
2.



3.



4.



問 8

<複数回答>以下のモデルのうち、線形回帰モデルに変換できるものをすべて選択しなさい。

1. $Y_i = \alpha X_i^{\beta} + u_i$
2. $Y_i = \alpha + \beta_1 X_i + \beta_2 X_i Z_i + u_i$
3. $Y_i = e^{\alpha + \beta_1 X_i + u_i}$
4. $Y_i = e^{\alpha + \beta_1 X_i} + u_i$

問 9

分布ラグモデル $Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + u_t$ において、X_t が 1 単位変化したとき、累積動学乗数は t+2 期にいくつになるか。

1.	$\beta_0 + \beta_1 + \beta_2$
2.	$\beta_0 + \beta_1$
3.	β_0
4.	β_2

問 10

データを推定した結果、$\widehat{\ln(Y)} = 10 + 0.5X$となった。$X$が 1 単位変化したとき、$Y$は何%変化するか。なお、$e^{0.5} = 1.649$とする。

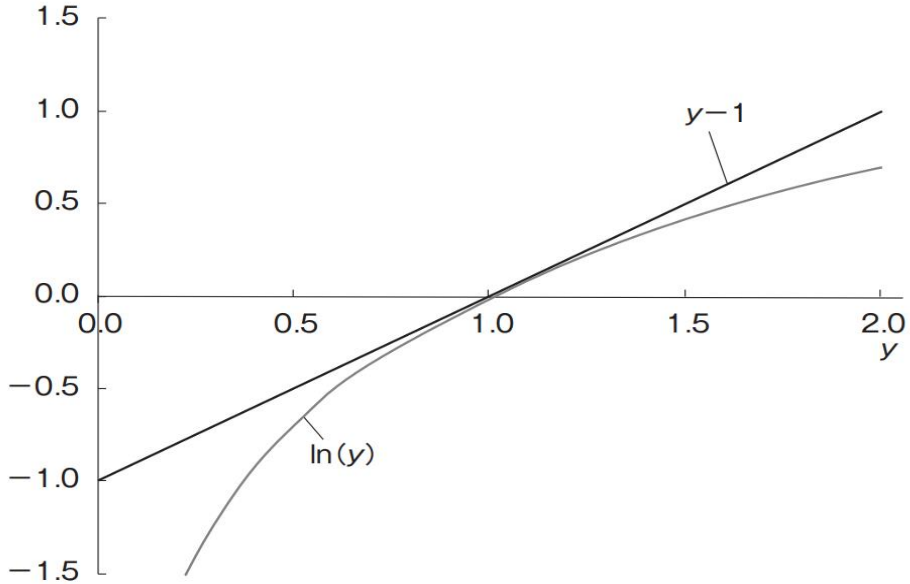
1. 64.9%
2. 50%
3. 0.5%
4. 0.649%

- **対数の差は変化率と解釈される**

$$\ln(S_t) - \ln(S_{t-1}) = \ln\left(\frac{S_t}{S_{t-1}}\right)$$
$$= \ln\left(1 + \frac{S_t - S_{t-1}}{S_{t-1}}\right) \approx \frac{S_t - S_{t-1}}{S_{t-1}}$$

[補足]　εが0に近い値であれば(正でも負でも)、
$$\ln(1+\varepsilon) \approx \varepsilon$$

図 A - 2　ln(*y*) と直線 *y*－1 の関係



$y = 1$付近では、$\ln(y) \approx y - 1$
$y = 1 + \varepsilon$と設定すると、
εが0に近いとき
$$\ln(1+\varepsilon) \approx (1+\varepsilon) - 1 = \varepsilon$$

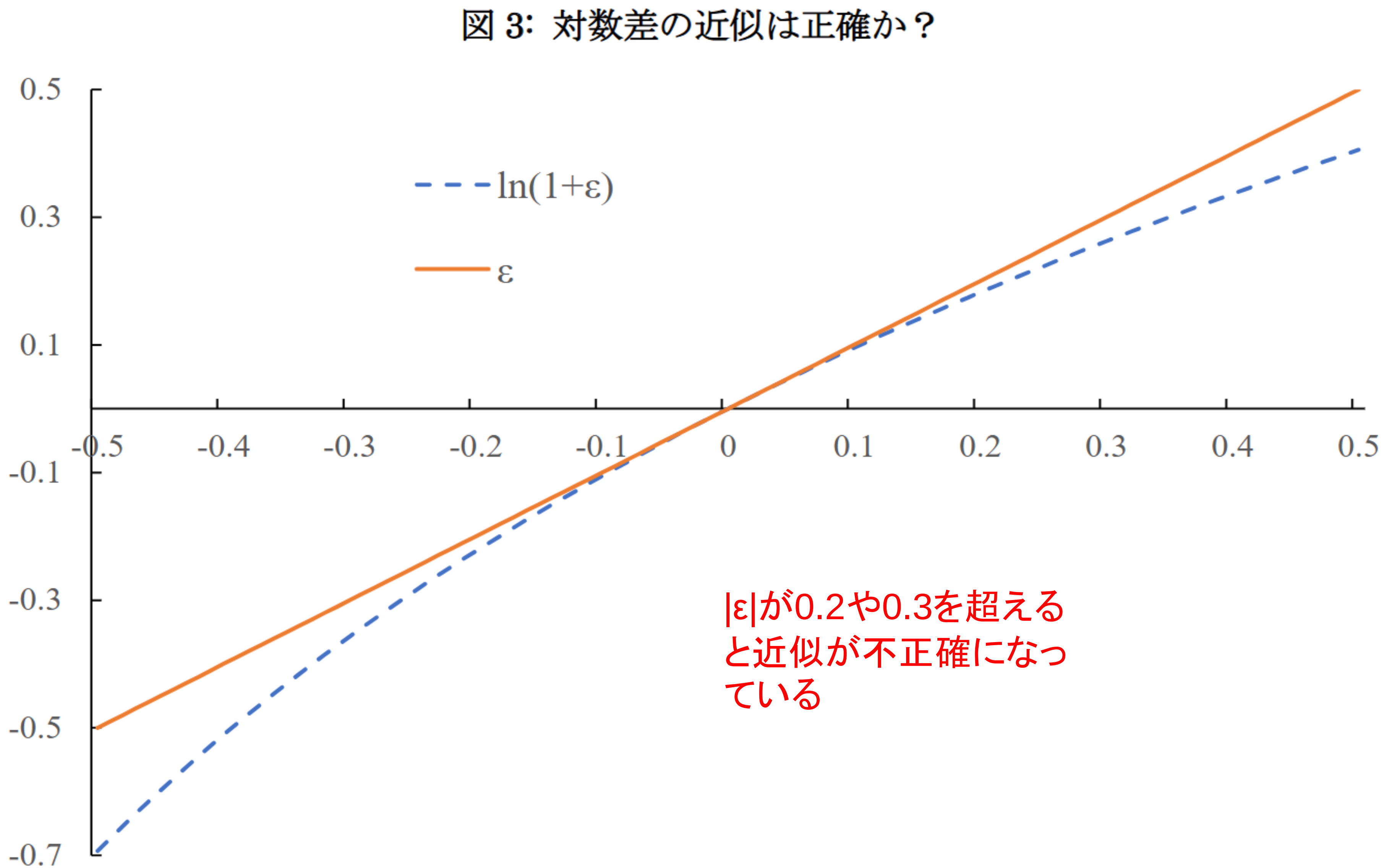


図 3: 対数差の近似は正確か？
0.5
0.3
0.1
-0.1
-0.3
-0.5
-0.7
-0.5
-0.4
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
0.4
0.5
ln(1+ε)
ε
|ε|が0.2や0.3を超える
と近似が不正確になっ
ている

対数線形(log-linear)モデル

$$\ln(Y) = \alpha + \beta X + u$$

β: Xが1単位変化すると、Yは$100 \times \beta$%変化する

[証明] XがX'に変化したとき、$\ln(Y)$から$\ln(Y')$への変化は、

$$\ln(Y') - \ln(Y) = (\underbrace{\alpha + \beta X'}_{\ln(Y')}) - (\underbrace{\alpha + \beta X}_{\ln(Y)})$$

$$= \beta(X' - X)$$

この関係式を、βについて解くと、

$$\beta = \frac{\ln(Y') - \ln(Y)}{X' - X} = \frac{\ln\left(\frac{Y'}{Y}\right)}{X' - X} = \frac{\ln\left(1 + \frac{Y' - Y}{Y}\right)}{X' - X} \approx \frac{\frac{Y' - Y}{Y}}{X' - X}$$

Yの変化が小さいなら、

$$\frac{Y' - Y}{Y} = \beta\ (X' - X)$$

対数近似の注意点

Yの変化が大きいなら、

$$\ln\left(1+\frac{Y'-Y}{Y}\right) \neq \frac{Y'-Y}{Y}$$

この場合、近似を使うことができない。

$$\beta = \frac{\ln(Y') - \ln(Y)}{X'-X} = \frac{\ln\left(\frac{Y'}{Y}\right)}{X'-X} = \frac{\ln\left(1+\frac{Y'-Y}{Y}\right)}{X'-X}$$

ここで、 $X'-X=1$とすると、

$$\beta = \ln\left(1+\frac{Y'-Y}{Y}\right)$$

このため、

$$e^{\beta} = 1 + \frac{Y'-Y}{Y}$$

$$\frac{Y'-Y}{Y} = e^{\beta} - 1$$

例:　重力モデル－－対数対数モデルの例

・万有引力の法則:「2つの物体間に働く力(F)」は、それぞれの質量(M、m)と相互の距離(r)に依存する

$$F = G\frac{Mm}{r^2}$$

--- Gは万有引力定数

・重力モデル:「2国間の貿易量($Trade_{ij}$)」は、それぞれのGDP(GDP_i、GDP_j)と相互の距離(D_{ij})に依存する

$$Trade_{ij} = A\frac{GDP_i GDP_j}{D_{ij}}e^{u_{ij}}$$

--- 両辺の対数をとると、対数対数モデルになる

$$\ln\left(Trade_{ij}\right) = \ln(A) + \beta_1 \ln(GDP_i) + \beta_2 \ln\left(GDP_j\right) + \beta_3 \ln(D_{ij}) + u_{ij}$$

--- 非常に当てはまりの良いモデル

重力モデルの問題

- 本当の重力モデル

$$Trade_{ij} = A\frac{GDP_i GDP_j}{D_{ij}} + u_{ij}$$

--- そのままでは対数をとることができない

--- 対数をとるため、

$$Trade_{ij} = A\frac{GDP_i GDP_j}{D_{ij}}\underbrace{\left(1+\frac{1}{A}\frac{D_{ij}}{GDP_i GDP_j}u_{ij}\right)}_{=\eta_{ij}}$$

--- 両辺の対数をとると、

$$\ln(Trade_{ij}) = \ln(A) + \beta_1\ln(GDP_i) + \beta_2\ln(GDP_j) + \beta_3\ln(D_{ij}) + ln(\eta_{ij})$$

しかし、説明変数とη_{ij}は相関してしまう

--- 非線形最小2乗法もしくはポアソン回帰を用いる

(推定方法は、教科書サポートウェブサイトの「カウントデータ」、「貿易における重力モデル」を参照)

15. ★　所得Y_iの自然対数である$\ln(Y_i)$を分析したい。

(a) 男性所得の対数の平均は 6.2、女性所得の対数の平均は 5.9 とする。この結果から、男性は女性より所得が平均 30%高いといえるか。Hint: Yの幾何平均は$(Y_1 \times Y_2 \times \ldots \times Y_n)^{1/n}$と定義される[1]。

(b) 被説明変数を所得の対数、説明変数を教育年数としたところ、係数は 0.1 と推定された。この結果から、教育年数が 1 年増えると、所得は平均 10%増えるといえるか。

- **幾何平均　$(\boldsymbol{Y_1 \times Y_2 \times \cdots \times Y_n})^{\boldsymbol{1/n}}$**
  平均変化率を求めるための指標

<例> 株価変化率70%、-20%、10%なら、

株価は1.7倍、0.8倍、1.1倍となる(計1.7×0.8×1.1=1.496倍)

幾何平均　$(1.7 \times 0.8 \times 1.1)^{1/3} = 1.144$　(計 1.144^3=1.496倍)

平均　$1.2 = (1.7 + 0.8 + 1.1)/3$　(計1.2^3=1.728倍)

・対数の平均は、幾何平均の対数である

$$\frac{1}{n}\sum_{i=1}^{n} \ln(Y_i) = \ln\left((Y_1 \times Y_2 \times \cdots \times Y_n)^{\frac{1}{n}}\right)$$

(a)X_i：男所得、Z_i：女所得(男n_1人、女n_2人)

$$\frac{1}{n_1}\sum_{i=1}^{n_1}\ln(X_i)-\frac{1}{n_2}\sum_{i=1}^{n_2}\ln(Z_i)$$

$$=\ln\left(\left(X_1\times X_2\times\cdots\times X_{n_1}\right)^{\frac{1}{n_1}}\right)-\ln\left(\left(Z_1\times Z_2\times\cdots\times Z_{n_2}\right)^{\frac{1}{n_2}}\right)$$

$$=\ln\left(\frac{\left(X_1\times X_2\times\cdots\times X_{n_1}\right)^{\frac{1}{n_1}}}{\left(Z_1\times Z_2\times\cdots\times Z_{n_2}\right)^{\frac{1}{n_2}}}\right)$$

$$=\ln\left(1+\frac{\left(X_1\times X_2\times\cdots\times X_{n_1}\right)^{\frac{1}{n_1}}-\left(Z_1\times Z_2\times\cdots\times Z_{n_2}\right)^{\frac{1}{n_2}}}{\left(Z_1\times Z_2\times\cdots\times Z_{n_2}\right)^{\frac{1}{n_2}}}\right)$$

$$\approx\frac{\left(X_1\times X_2\times\cdots\times X_{n_1}\right)^{\frac{1}{n_1}}-\left(Z_1\times Z_2\times\cdots\times Z_{n_2}\right)^{\frac{1}{n_2}}}{\left(Z_1\times Z_2\times\cdots\times Z_{n_2}\right)^{\frac{1}{n_2}}}$$

対数平均の差が0.3なら、男性は女性より平均で所得が30%高いといって誤りではないが、厳密には、平均は幾何平均である

(b) Yを所得、Xを教育年数としたところβは0.1と推定された

$$\ln(Y) = \alpha + \beta X + u$$

このとき、$\hat{\alpha}$の公式から、

$$\overline{\ln(Y)} = \frac{1}{n}\sum_{i=1}^{n} \ln(Y_i) = \hat{\alpha} + \hat{\beta}\bar{X}$$

$\bar{X}$が1年増えると、$\overline{\ln(Y)}$から$\overline{\ln(Y)}'$に変化する

$$\overline{\ln(Y)}' = \hat{\alpha} + \hat{\beta}(\bar{X} + 1)$$

よって、

$$\overline{\ln(Y)}' - \overline{\ln(Y)} = \left(\hat{\alpha} + \hat{\beta}\left(\bar{X} + 1\right)\right) - \left(\hat{\alpha} + \hat{\beta}\bar{X}\right)$$
$$= \hat{\beta}$$

教育年数が1年増えると、所得は平均して10%増える(幾何平均であることに注意)

対数線形モデルからY_iを予測する

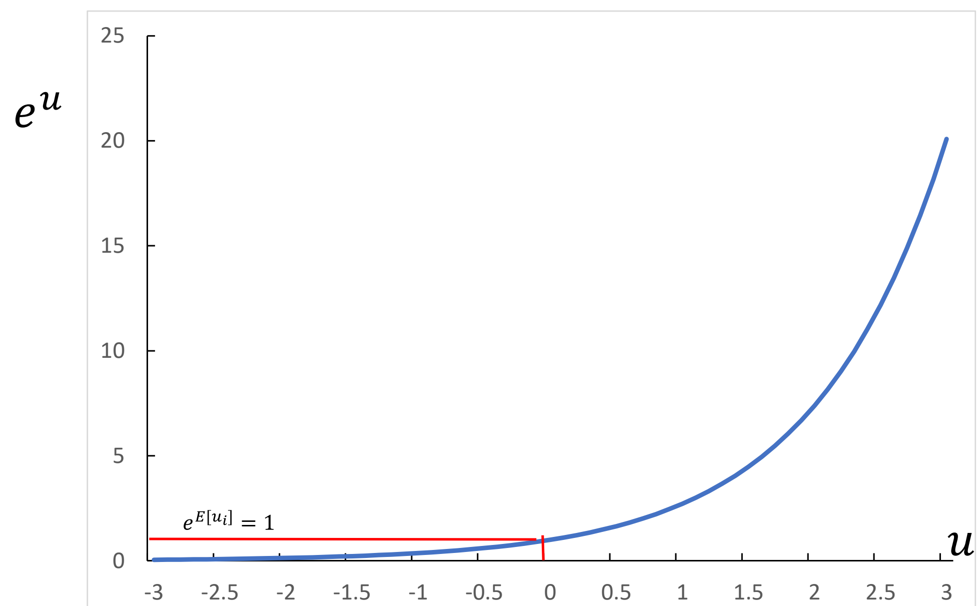
- $\mathrm{E}[\ln(Y_i)]$ ではなく、$\mathrm{E}[Y_i]$に関心がある
- 対数線形モデル$\ln(Y_i) = \alpha + \beta X_i + u_i$なら

$$Y_i = e^{\alpha+\beta X_i} e^{u_i}$$

$$E[Y_i] = e^{\alpha+\beta X_i} E[e^{u_i}]$$

- $E[e^{u_i}] \neq e^{E[u_i]} = 1$　　(一般には、$E[e^{u_i}] > 1$)

$$E[Y_i] > e^{\alpha+\beta X_i}$$



スミアリング推定量(smearing estimate)

① $\ln(Y_i) = \alpha + \beta X_i + u_i$をOLS推定し残差を計算

$$\hat{u}_i = \ln(Y_i) - \hat{\alpha} - \hat{\beta} X_i$$

② $E[e^{u_i}]$の推定量を計算

$$\widehat{E[e^{u_i}]} = \frac{1}{n}\sum_{i=1}^{n} e^{\hat{u}_i}$$

③ $\widehat{E[e^{u_i}]}$を用いて$E[Y_i] = e^{\alpha+\beta X_i} E[e^{u_i}]$を推定

$$e^{\alpha+\beta X_i}\widehat{E[e^{u_i}]}$$