2024/04/12

藪友良

第 6 章の答え

練習問題 1

ここで推定式は次のようになる。

$$\hat{Y}_i = 20 - 5X_{1i} + 10X_{2i} + 5X_{1i}X_{2i}$$

ここで、X_{1i}は女性ダミー、X_{2i}は大卒ダミーである。

(a) 高卒男性なら、$X_{1i} = X_{2i} = 0$となるため、交差項$X_{1i}X_{2i}$も 0 となる。よって、高卒男性の所得は$\hat{Y}_i = 20$万円である。高卒女性なら、$X_{1i} = 1$、$X_{2i} = 0$となるため、交差項$X_{1i}X_{2i}$は 0 となる。よって、高卒女性の所得は$\hat{Y}_i = 20 - 5 = 15$である。以上より、

$$\text{高卒男性の所得} - \text{高卒女性の所得} = 20 - 15 = 5$$

(b) 大卒男性なら、$X_{1i} = 0$、$X_{2i} = 1$となるため、交差項$X_{1i}X_{2i}$は 0 となる。よって、大卒男性の所得は$\hat{Y}_i = 20 + 10 = 30$である。大卒女性なら、$X_{1i} = 1$、$X_{2i} = 1$となるため、交差項$X_{1i}X_{2i}$は 1 となる。よって、大卒女性の所得は$\hat{Y}_i = 20 - 5 + 10 + 5 = 30$である。

$$\text{大卒男性の所得} - \text{大卒女性の所得} = 30 - 30 = 0$$

練習問題 2

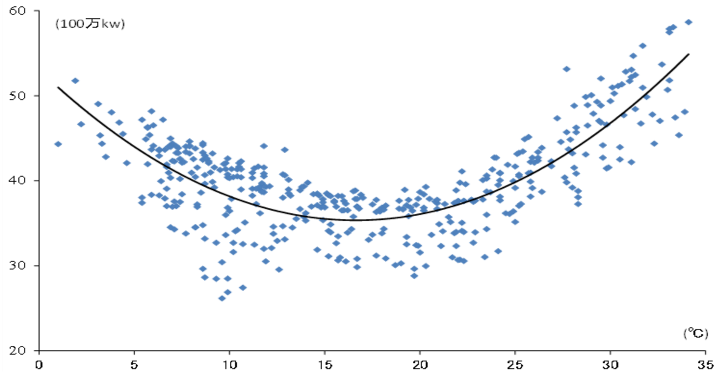
5.7.3 節では、決定係数は被説明変数を変えると異なる意味を持つため、その値の相互比較には意味がないことを学習した。本問題では、2 つのモデルは異なる被説明変数であるため、決定係数による相互比較はできない。この場合、経済理論や t 値などを参考にしながら、定式化を決めることになる。

練習問題 3

ここで、図 6-1 を再掲載した。実線は

$$\hat{Y}_t = 53.04 - 2.127X_{1t} + 0.064X_{1t}^2$$

であり、これをみると、約 17 度で電力需要が最小になっていることがわかる。



正確な気温を求めるため、

$$\widehat{Y}_t = 53.04 - 2.127X_{1t} + 0.064X_{1t}^2$$

を気温X_{1t}で偏微分して 0 と置く。

$$\frac{\partial(53.04 - 2.127X_{1t} + 0.064X_{1t}^2)}{\partial X_{1t}} = -2.127 + 2 \times 0.064X_{1t} = 0$$

これを気温X_{1t}について解くと、電力需要を最も低くする気温が得られる。

$$X_{1t} = \frac{2.127}{2 \times 0.064} = 16.617$$

練習問題 4

線形回帰モデルに変換できるのは(b)のみである。(b)では、$X_{2i} = X_i^2$、$X_{3i} = X_i^3$と定義すれば、線形回帰モデルになる。

(a)式はパラメータに関して線形ではなく、また、対数をとっても線形に変換できない。

(c)式は、誤差項が和の形で含まれており、対数をとっても線形にならない。仮にモデルを、

$$Q_i = AK_i^{\beta_1}L_i^{\beta_2}u_i^*$$

と変更すれば、対数をとることで、線形回帰モデルに変換できる。

$$\underbrace{\ln(Q_i)}_{=Y_i} = \underbrace{\ln(A)}_{=\alpha} + \beta_1 \underbrace{\ln(K_i)}_{=X_{1i}} + \beta_2 \underbrace{\ln(L_i)}_{=X_{2i}} + \underbrace{\ln(u_i^*)}_{=u_i}$$

練習問題 5

(a) 制約$\beta_1 = 2\beta_2$を書き換えた$\beta_2 = \beta_1/2$を代入すると、

$$\begin{aligned} Y_i &= \alpha + \beta_1 X_{1i} + \frac{\beta_1}{2}X_{2i} + u_i \\ &= \alpha + \beta_1 \underbrace{\left(X_{1i} + \frac{X_{2i}}{2}\right)}_{=X_i} + u_i \end{aligned}$$

となる。ここで、説明変数X_iを

$$X_i = X_{1i} + \frac{X_{2i}}{2}$$

と定義して、新しいモデルを

$$Y_i = \alpha + \beta_1 X_i + u_i$$

とすれば、制約($\beta_2 = \beta_1/2$)を課した推定ができる。

(b) 制約$\alpha + \beta_1 + \beta_2 = 1$を書き換えた$\beta_2 = 1 - \alpha - \beta_1$をモデルに代入すると、

$$Y_i = \alpha + \beta_1 X_{1i} + (1 - \alpha - \beta_1)X_{2i} + u_i$$

となる。これを書き換えると、

$$Y_i - X_{2i} = \alpha(1 - X_{2i}) + \beta_1(X_{1i} - X_{2i}) + u_i$$

となる。ここで、被説明変数を$Y_i - X_{2i}$とし、説明変数を$(1 - X_{2i})$と$(X_{1i} - X_{2i})$とすれば、αとβ_1を推定できる[1]。

練習問題 6

(a) Xが 1%変化すると、Yは$\beta \times 0.01$単位分変化する

(b) Xが 1%変化すると、Yはβ%分変化する

(c) Xが 1 単位変化すると、Yは$100 \times \beta$%分変化する

詳細は 6.3 節と補足を参照されたい。

練習問題 7

AIC の値は$p = 3$で最小となるため、$\hat{p} = 3$となる。

練習問題 8

教科書の式展開では、

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

という関係を用いたが、これは変化率$(X' - X)/X$が小さいときのみ成立する。ここで、変化率$(X' - X)/X$は 50%と大きいため、

$$\ln\left(1 + \frac{X' - X}{X}\right) = \ln(1.5) = 0.4054$$

となり、上記は 0.5 と大きく異なる。

近似を使わないで、変化量$Y' - Y$を評価しよう。単純化のため、$u = 0$とすると、

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

となる。ここで、$(X' - X)/X = 0.5$とすると、

$$Y' - Y = \beta \ln(1.5) = \beta \times 0.4054$$

となり、Yは$\beta \times 0.4054$だけ変化するといえる($\ln(1.5) = 0.4054$)。この例からも明らかなとおり、変化率が大きいときは、近似関係が使えないことがわかる。

練習問題 9

単純化のため、$u = 0$とする。このとき、モデルは$\ln(Y) = \alpha + \beta X$となり、

$$Y = e^{\alpha + \beta X}$$

を意味している。つまり、XがX'に変化すると、Yの変化率は、

---

1 モデル$Y_i - X_{2i} = \alpha(1 - X_{2i}) + \beta_1(X_{1i} - X_{2i}) + u_i$では、説明変数は$(1 - X_{2i})$と$(X_{1i} - X_{2i})$)であり、これらの係数は$\alpha$と$\beta_1$となる。このモデルに定数項は存在しないこと。つまり、モデルを推定する際、定数項がない回帰モデルとして、推定する必要がある(3 章練習問題 12 では、定数項がない回帰モデルの推定方法が書かれているので参考にしてほしい)。

$$\frac{Y'-Y}{Y}=\frac{e^{\alpha+\beta X'}-e^{\alpha+\beta X}}{e^{\alpha+\beta X}}=e^{\beta(X'-X)}-1$$

となる。Xの 1 単位の変化なら、Yの変化率は次のようになる。

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

練習問題 10

ラグの長さを$p=3$とすると、分布ラグモデルは、

$$Y_t=\alpha+\beta_0X_t+\beta_1X_{t-1}+\beta_2X_{t-2}+\beta_3X_{t-3}+u_t$$

となる。ここで、右辺に$\beta_0(X_{t-1}-X_{t-1})$を足すと、

$$Y_t=\alpha+\beta_0(X_t-X_{t-1})+(\beta_0+\beta_1)X_{t-1}+\beta_2X_{t-2}+\beta_3X_{t-3}+u_t$$

となる(ここで$\beta_0(X_{t-1}-X_{t-1})=0$であり、0 を足しても等号関係は変わらない)。さらに、右辺に$(\beta_0+\beta_1)(X_{t-2}-X_{t-2})$を足すと、

$$Y_t=\alpha+\beta_0(X_t-X_{t-1})+(\beta_0+\beta_1)(X_{t-1}-X_{t-2})+(\beta_0+\beta_1+\beta_2)X_{t-2}+\beta_3X_{t-3}+u_t$$

となる。最後に、右辺に$(\beta_0+\beta_1+\beta_2)(X_{t-3}-X_{t-3})$を足すと、

$$\begin{aligned}Y_t=\alpha&+\beta_0(X_t-X_{t-1})+(\beta_0+\beta_1)(X_{t-1}-X_{t-2})+(\beta_0+\beta_1+\beta_2)(X_{t-2}-X_{t-3})\\&+(\beta_0+\beta_1+\beta_2+\beta_3)X_{t-3}+u_t\end{aligned}$$

となる。ここで、

$$\Delta X_t=X_t-X_{t-1}、\Delta X_{t-1}=X_{t-1}-X_{t-2}、\Delta X_{t-2}=X_{t-2}-X_{t-3}$$

$$\theta_0=\beta_0,\ \theta_1=\beta_0+\beta_1,\ \theta_2=\beta_0+\beta_1+\beta_2,\ \theta_3=\beta_0+\beta_1+\beta_2+\beta_3$$

と定義すれば、上式は次のようになる。

$$Y_t=\alpha+\theta_0\Delta X_t+\theta_1\Delta X_{t-1}+\theta_2\Delta X_{t-2}+\theta_3X_{t-3}+u_t$$

変形したモデルを OLS 推定し、θ_hの推定値と標準誤差を求めれば、累積動学乗数の仮説検定や 95%信頼区間の計算が可能となる。横軸を h とし、縦軸を θ_hとして図示すれば、h が変わると累積動学乗数がどのように変化したのかを視覚的に示すことができる。

練習問題 11

被説明変数は$Y_i^*=c_YY_i$、説明変数は$X_i^*=c_XX_i$に変換する。元モデルは次のようになる。

$$Y_i=\alpha+\beta X_i+u_i$$

6.6 節で確認した通り、上式の両辺にc_Yを掛けると、

$$c_YY_i=c_Y\alpha+\left(\frac{c_Y}{c_X}\beta\right)c_XX_i+c_Yu_i$$

となるため、$\alpha^*=c_Y\alpha$、$\beta^*=\frac{c_Y}{c_X}\beta$、$u_i^*=c_Yu_i$と定義すると、

$$Y_i^*=\alpha^*+\beta^*X_i^*+u_i^*$$

と表現できる。ここで、u_i^*の分散をσ^{*2}と表記すると、$u_i^*=c_Yu_i$より、次のようになる。

$$\sigma^{*2}=E[u_i^{*2}]=E[(c_Y u_i)^2]=c_Y^2\,E[u_i^2]=c_Y^2\sigma^2$$

以下では、スケール変更しても、t 統計量の値は全く影響を受けないことを示す。まず、OLS 推定量$\hat{\beta}^*$は、次のように書き換えられる。

$$\hat{\beta}^*=\frac{\sum_{i=1}^n(X_i^*-\bar{X}^*)(Y_i^*-\bar{Y}^*)}{\sum_{i=1}^n(X_i^*-\bar{X}^*)^2}$$

$$=\frac{\sum_{i=1}^n(c_X X_i-c_X\bar{X})(c_Y Y_i-c_Y\bar{Y})}{\sum_{i=1}^n(c_X X_i-c_X\bar{X})^2}$$

$$=\frac{c_X c_Y}{c_X^2}\underbrace{\frac{\sum_{i=1}^n(X_i-\bar{X})(Y_i-\bar{Y})}{\sum_{i=1}^n(X_i-\bar{X})^2}}_{=\hat{\beta}}=\frac{c_Y}{c_X}\hat{\beta}$$

次に、OLS 推定量$\hat{\beta}^*$の分散は

$$s_{\hat{\beta}^*}^2=\frac{\sigma^{*2}}{\sum_{i=1}^n(X_i^*-\bar{X}^*)^2}$$

となる。上式に$\sigma^{*2}=c_Y^2\sigma^2$、$X_i^*=c_X X_i$、また、

$$\bar{X}^*=\frac{\sum_{i=1}^n X_i^*}{n}=\frac{\sum_{i=1}^n c_X X_i}{n}=c_X\frac{\sum_{i=1}^n X_i}{n}=c_X\bar{X}$$

を代入すると、

$$s_{\hat{\beta}^*}^2=\frac{\sigma^{*2}}{\sum_{i=1}^n(X_i^*-\bar{X}^*)^2}=\frac{c_Y^2\sigma^2}{\sum_{i=1}^n(c_X X_i-c_X\bar{X})^2}=\frac{c_Y^2}{c_X^2}\frac{\sigma^2}{\sum_{i=1}^n(X_i-\bar{X})^2}$$

となる。以上から、t 統計量は次のようになる。

$$\frac{\hat{\beta}^*-\beta^*}{\sqrt{s_{\hat{\beta}^*}^2}}=\frac{\frac{c_Y}{c_X}\hat{\beta}-\frac{c_Y}{c_X}\beta}{\sqrt{\frac{c_Y^2}{c_X^2}\frac{\sigma^2}{\sum_{i=1}^n(X_i-\bar{X})^2}}}=\frac{\frac{c_Y}{c_X}\left(\hat{\beta}-\beta\right)}{\frac{c_Y}{c_X}\sqrt{\frac{\sigma^2}{\sum_{i=1}^n(X_i-\bar{X})^2}}}=\frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^n(X_i-\bar{X})^2}}}=\frac{\hat{\beta}-\beta}{\sqrt{s_{\hat{\beta}}^2}}$$

スケール変更をしても、t 統計量は影響を受けないことが確認できた。

練習問題 12

(a) 対数の性質から、$Y_i^*=c_Y Y_i$と$X_i^*=c_X X_i$の対数は、

$$\ln(Y_i^*)=\ln(c_Y)+\ln(Y_i)、\ln(X_i^*)=\ln(c_X)+\ln(X_i)$$

となる。対数対数モデルである

$$\ln(Y_i)=\alpha+\beta\ln(X_i)+u_i$$

の両辺に $\ln(c_Y)$を足して、右辺に$\beta(\ln(c_X)-\ln(c_X))$を足すと、

$$\ln(c_Y)+\ln(Y_i)=\alpha+\ln(c_Y)+\beta(\ln(c_X)-\ln(c_X))+\beta\ln(X_i)+u_i$$

$$=(\alpha+\ln(c_Y)-\beta\ln(c_X))+\beta(\ln(c_X)+\ln(X_i))+u_i$$

となる($\beta(\ln(c_X)-\ln(c_X))=0$であるため、右辺に足しても等号関係は変わらない)。ここで、

$$\ln(Y_i^*)=\ln(c_Y)+\ln(Y_i)、\ln(X_i^*)=\ln(c_X)+\ln(X_i)$$

という関係に注意すると、上式は次のように書き換えることができる。

$$\ln(Y_i^*) = (\alpha + \ln(c_Y) - \beta \ln(c_X)) + \beta \ln(X_i^*) + u_i$$

また、$\alpha^* = \alpha + \ln(c_Y) - \beta \ln(c_X)$と定義すると、次のようになる。

$$\ln(Y_i^*) = \alpha^* + \beta \ln(X_i^*) + u_i$$

　以上から、こうしたスケール変更では、説明変数の係数は影響を受けていないが、定数項は変化したことがわかる。

(b) 線形対数モデルである

$$Y_i = \alpha + \beta \ln(X_i) + u_i$$

の右辺に$\beta(\ln(c_X) - \ln(c_X))$を足すと、

$$\begin{aligned} Y_i &= \alpha + \beta(\ln(c_X) - \ln(c_X)) + \beta \ln(X_i) + u_i \\ &= (\alpha - \beta \ln(c_X)) + \beta(\ln(c_X) + \ln(X_i)) + u_i \end{aligned}$$

となり、さらにc_Yを両辺に掛けると、

$$c_Y Y_i = c_y(\alpha - \beta \ln(c_x)) + c_y \beta(\ln(c_x) + \ln(X_i)) + c_y u_i$$

となる。ここで、$X_i^* = c_X X_i$の対数は、$\ln(X_i^*) = \ln(c_X) + \ln(X_i)$となることに注意すると、上式は、次のよう表現できる。

$$Y_i^* = \alpha^* + \beta^* \ln(X_i^*) + u_i^*$$

ただし、$\alpha^* = c_Y(\alpha - \beta \ln(c_X))$、$\beta^* = c_Y \beta$、$u_i^* = c_Y u_i$とした。

　以上から、こうしたスケール変更によって、定数項と係数はともに変化したことがわかる。

(c) 対数線形モデルである

$$\ln(Y_i) = \alpha + \beta X_i + u_i$$

の両辺に $\ln(c_y)$を足してから展開すると、

$$\begin{aligned} \ln(c_Y) + \ln(Y_i) &= \alpha + \ln(c_Y) + \beta X_i + u_i \\ &= \alpha + \ln(c_Y) + \frac{\beta}{c_X} c_X X_i + u_i \end{aligned}$$

となる。ここで、$Y_i^* = c_Y Y_i$、$X_i^* = c_X X_i$であることに注意すると、上式は次のように表現できる($\ln(Y_i^*) = \ln(c_Y) + \ln(Y_i)$に注意)。

$$\ln(Y_i^*) = \alpha^* + \beta^* X_i^* + u_i$$

ただし、$\alpha^* = \alpha + \ln(c_Y)$、$\beta^* = \frac{\beta}{c_X}$とした。

　以上より、こうしたスケール変更によって、定数項と係数はともに変化したことになる。

練習問題 13、14

ウェブサイトから、再現に必要なデータと再現コードをダウンロードできる。

＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊＊

こちらは初版(第 1 刷)には掲載されていない新しい問題になります。

練習問題 15

15. ★　所得Y_iの対数である$\ln(Y_i)$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。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[2]。

(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。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。Hint: εが小さいとき、$\ln(1+\varepsilon) \approx \varepsilon$ となる(巻末付録 A.3.2 節参照)。

練習問題 15 の答え

(a) これは大まかには正しい記述だが、厳密には、正確ではない。この点を厳密に議論していく。対数の平均は、次のように表現できる。

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

つまり、対数の平均は、Y_iの幾何平均の対数となる。

男性の所得をX_iとし、女性の所得をZ_iとしよう(男性は計n_1人、女性は計n_2人いるとする)。このとき、男性の対数平均から女性の対数平均を引くと、次のように展開できる。

$$\frac{1}{n_1}\sum_{i=1}^{n_1} \ln(X_i) - \frac{1}{n_2}\sum_{i=1}^{n_2} \ln(Z_i) = \ln\left(\left(X_1 \times X_2 \times \ldots \times X_{n_1}\right)^{\frac{1}{n_1}}\right) - \ln\left(\left(Z_1 \times Z_2 \times \ldots \times Z_{n_2}\right)^{\frac{1}{n_2}}\right)$$

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

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

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

---

[2] 幾何平均は、変化率の平均を求めるのに便利な指標である。たとえば、株価の変化率が 70%、-20%、10%としよう。このとき、株価は 1.7 倍、0.8 倍、1.1 倍となるので、幾何平均は$(1.7 \times 0.8 \times 1.1)^{1/3} = 1.14$となる。これに対して、通常の平均は$1.2 (= (1.7 + 0.8 + 1.1)/3)$となる。幾何平均は$1.7 \times 0.8 \times 1.1 = 1.14^3$となり、平均変化率として適切であることがわかる。一方、平均は$1.7 \times 0.8 \times 1.1 < 1.2^3$となり、平均変化率を過大に評価し不適節である。

最後の近似は、εが小さいとき、$\ln(1+\varepsilon) \approx \varepsilon$となることを用いた(巻末付録 A.3.2 節参照)

この例では、対数平均の差は 0.3(=6.2－5.9)であることから、男性所得は女性所得より幾何平均で 30%高いといえる。男性の所得は女性より平均で 30%と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ことを覚えていてほしい。

(b) 対数の平均を次のように定義する。

$$\overline{\ln(Y)} = \frac{1}{n}\sum_{i=1}^{n} \ln(Y_i)$$

OLS 推定量$\hat{\alpha}$の公式から(P32 参照)、

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

という関係が成立する。ここで、$\bar{X}$が 1 年増えると、$\overline{\ln(Y)}$は$\overline{\ln(Y)}'$に変化する。

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

よって、$\overline{\ln(Y)}$から$\overline{\ln(Y)}'$ の変化は$\hat{\beta}$となる。

$$\overline{\ln(Y)}' - \overline{\ln(Y)} = \hat{\beta}$$

この結果から、教育年数が 1 年増えると、所得は幾何平均でみて 10%増えるといえる。教育年数が 1 年増えると、所得は平均 10%増えるといっても間違いではないが、厳密には、平均は幾何平均であることを覚えておいてほしい。