2017/12/29

藪友良

第 11 章の答え

1. β は真の値であり、$\hat{\beta}$はデータからβの推定値です。u は誤差項であり、$\hat{u}$は残差（誤差項の推定値）です。

2. (1)標準誤差は、以下となります。

$$\sqrt{s_{\hat{\beta}}^2} = \sqrt{\frac{\sum \hat{u}_i^2/(n-2)}{\sum (X_i - \bar{X})^2}} = \sqrt{\frac{120/3}{250}} = 0.4$$

2)ここで n=5 ですから、βの t 統計量は自由度 3 の t 分布に従います。よって、β の信頼区間は、

$$0.95 = \mathrm{P}\{-t_{3,0.05} < t_{\hat{\beta}} < t_{3,0.05}\}$$

$$= \mathrm{P}\{-t_{3,0.05} < \frac{\hat{\beta}-\beta}{\sqrt{s_{\hat{\beta}}^2}} < t_{3,0.05}\}$$

$$= \mathrm{P}\{\hat{\beta} - t_{3,0.05}\sqrt{s_{\hat{\beta}}^2} < \beta < \hat{\beta} + t_{3,0.05}\sqrt{s_{\hat{\beta}}^2}\}$$

として求めることができます。ここで $t_{3,0.05}$=3.182 から、信頼区間は

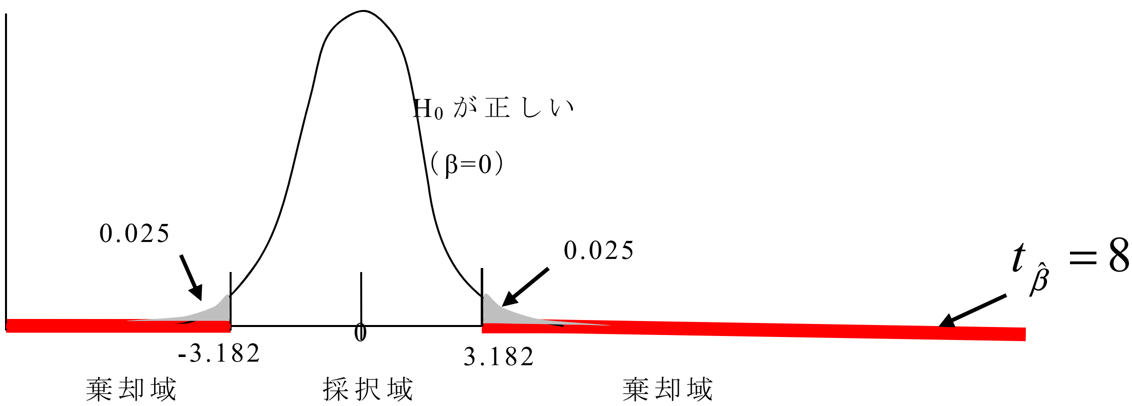
$$3.2 - 3.182 \times 0.4 < \beta < 3.2 + 3.182 \times 0.4$$

となります（計算すると 1.93<β<4.47 です）。

(3) ここで t 値は

$$t_{\hat{\beta}} = \frac{\hat{\beta} - 0}{\sqrt{s_{\hat{\beta}}^2}} = \frac{3.2}{0.4} = 8$$

であり、これは $t_{3,0.05}$=3.182 より大きいため H_0 を棄却します（下図参照）。



3. (1)　n=10、 $t_{8,0.05}$=2.306 から、95%の信頼区間は、

$$53.5-2.306\times5.3<\beta<53.5+2.306\times5.3$$

となります（計算すると 41.28<β<65.72）。

(2) t 値は、

$$t_{\hat{\beta}}=\frac{\hat{\beta}-\beta}{\sqrt{s_{\hat{\beta}}^{2}}}=\frac{53.5-55}{5.3}=-0.283$$

となり、$|t_{\hat{\beta}}|$=0.283<2.306 ですから H_0 を採択します。

(3) t 分布は左右対称ですから、片側検定を有意水準 5%で行うためには、t 分布表の α=0.10 の列から数値を見つけます。この場合、自由度 8 ですから、$t_{\hat{\beta}}$が-1.860 より小さいかで有意性を判断します[1]。t 値は$t_{\hat{\beta}}$=-0.283 であり、これは-1.860 より大きいため、H_0が採択されます。

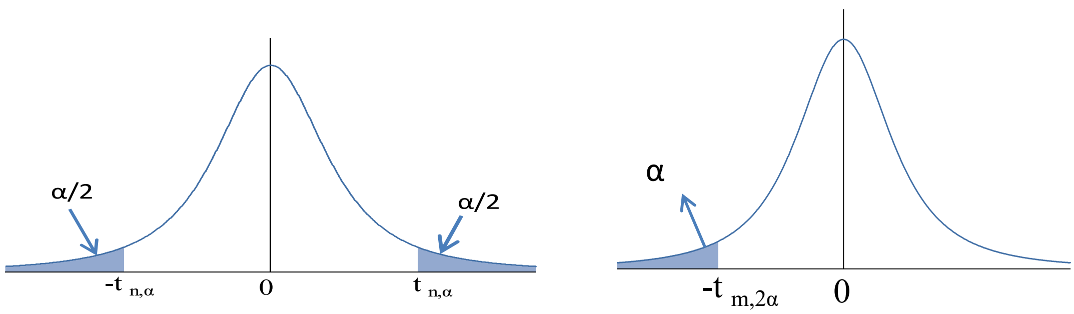
4. 1）喫煙をしない場合（X=0）体重は 3kg、喫煙量が 1 日 20 本の場合（X=20）体重は 3-0.012×20=2.76kg と予想されます。(2)サンプルサイズ n は 500 もあるので、t 分布と正規分布はほとんど同じです。よって、β の 95%信頼区間は、-0.012-1.96×0.005<β<-0.012+1.96×0.005 です(-0.0218<β<-0.0022)。　(3)

$$t_{\hat{\beta}}=\frac{\hat{\beta}}{s_{\hat{\beta}}}=\frac{-0.012}{0.005}=-2.4$$

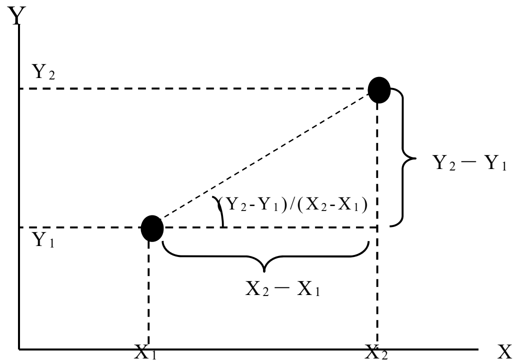
であり、$|t_{\hat{\beta}}|$=2.4>1.96 ですから有意水準 5%で H_0を棄却します。

---

[1] 9 章で学習した通り、t 分布表には、$P\{t_{n,\alpha}<|U|\}=\alpha$ が成立する $t_{n,\alpha}$ の値が表されています。t 分布は左右対称なので、$P\{t_{n,\alpha}<U\}=P\{U<-t_{n,\alpha}\}=\alpha/2$ です（左下図参照）。この問題では、t 統計量が十分に小さいかで H_0 を採択するかが判断されます。よって、有意水準 α ならば、t 統計量が $-t_{m,\alpha}$ より小さいかではなく（こうすると有意水準は $\alpha/2$ になる）、$-t_{m,2\alpha}$ より小さいかで判断すればよいのです（右下図参照）。



5.　この推定量は何を意味しているかを考えてみましょう。いま 2 個のデータ（$(X_1、Y_1)$と$(X_2、Y_2)$）があり、下図の●で表されています。



ここで、底辺 X_2-X_1、高さ Y_2-Y_1 となる三角形を考えましょう。この三角形の角度は$(Y_2-Y_1)/(X_2-X_1)$で表されます。つまり、この推定量は、2 点から作られる三角形の角度に当たるわけです。この角度が、X が 1 単位増えたとき、Y が何単位変化するかを表すのは直観的でしょう。

この推定量の統計的性質をみてみましょう。ここでモデルは標準的仮定を満たしているとして、この推定量の期待値と分散を求めます。まず、この推定量の確率的表現を求めます。

$$\frac{Y_2-Y_1}{X_2-X_1}=\frac{(\alpha+\beta X_2+u_2)-(\alpha+\beta X_1+u_1)}{X_2-X_1}$$

$$=\frac{\beta(X_2-X_1)+(u_2-u_1)}{X_2-X_1}=\beta+\frac{u_2-u_1}{X_2-X_1}$$

この期待値をとると、

$$E\left[\frac{Y_2-Y_1}{X_2-X_1}\right]=E\left[\beta+\frac{u_2-u_1}{X_2-X_1}\right]$$

$$=\beta+\frac{E[u_2]-E[u_1]}{X_2-X_1}=\beta+\frac{0-0}{X_2-X_1}=\beta$$

となり、不偏性を満たしていることが確認できます。

次に、分散は、

$$E\left[\left(\frac{Y_2-Y_1}{X_2-X_1}-\beta\right)^2\right]=E\left[\left(\frac{u_2-u_1}{X_2-X_1}\right)^2\right]=\frac{E[(u_2-u_1)^2]}{(X_2-X_1)^2}$$

$$= \frac{E[u_1^2] + E[u_2^2] - 2E[u_1 u_2]}{(X_2 - X_1)^2} = \frac{2\sigma^2}{(X_2 - X_1)^2}$$

となります（標準的仮定から、$E[u_i^2]=\sigma^2$、$E[u_1u_2]=0$ としました）。

この推定量は 2 つの観測値しか使っていないため、最小 2 乗推定量より分散は大きくなっています。分母は$(X_2-X_1)^2$ですから、X_2 と X_1 が互いに離れているほど、その分散は小さくなり、より望ましい推定量となります(説明変数のばらつきが大きいほど正確に推定できる点については、p273 の図 11-4 を参照してください)。

6. この問題を通じて、変数のスケール変更によって推定結果に何が生じるかを理解しましょう。もともと X_i と Y_i には $Y_i = \alpha + \beta X_i + u_i$ という関係があります（u の分散は σ^2）。この式の両辺に c_y を掛けると

$$c_y Y_i = c_y \alpha + \left(\frac{c_y}{c_x} \beta \right) c_x X_i + c_y u_i$$

と書き換えられます。ここで Y_i*=$c_y Y_i$、X_i*=$c_x X_i$、α*=$c_y \alpha$、$\beta^* = (c_y / c_x)\beta$、$u_i$*=$c_y u_i$ と定義すると、上式は

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

と表記できます（u_i*の分散を σ^{2*}と表記すれば $\sigma^{2*}=E[(c_y u_i)^2]=c_y^2\sigma^2$）。

この式から、スケール変更した変数 Y*、X*同士にどのような関係があるかを理解できるでしょう。たとえば、説明変数 X だけを 100 倍したとしましょう（c_y=1、c_x=100）。このとき、新しい説明変数 X*の係数 β*は、β を 1/100 倍したものとなります。これに対し、被説明変数だけを 100 倍したら（c_y=100、c_x=1）、説明変数 X*（この場合、X*=X）の係数 β*は、β を 100 倍したものとなります。以上から、スケール変更によって、推定されるパラメータの値が影響を受けることが分かりました。

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します。β*の最小 2 乗推定量は、

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

$$=\frac{\sum_{i=1}^{n}(c_xX_i-c_x\overline{X})(c_yY_i-c_y\overline{Y})}{\sum_{i=1}^{n}(c_xX_i-c_x\overline{X})^2}$$

$$=\frac{c_xc_y}{c_x^2}\frac{\sum_{i=1}^{n}(X_i-\overline{X})(Y_i-\overline{Y})}{\sum_{i=1}^{n}(X_i-\overline{X})^2}=\frac{c_y}{c_x}\hat{\beta}$$

と書き換えることができます。当然ですが、この推定量は不偏性を満たしています（$E[\hat{\beta}^*]=E[(c_y/c_x)\hat{\beta}]=(c_y/c_x)E[\hat{\beta}]=(c_y/c_x)\beta=\beta^*$）。次に、$\hat{\beta}^*$の分散は

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

となります。上式に、$\sigma^{2*}=c_y^2\sigma^2$、$X_i^*=c_xX_i$、また

$$\overline{X}^*=\frac{\sum_{i=1}^{n}X_i^*}{n}=\frac{\sum_{i=1}^{n}c_xX_i}{n}=c_x\frac{\sum_{i=1}^{n}X_i}{n}=c_x\overline{X}$$

を代入すると、

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

と書き換えることができます。以上から、t 統計量は

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

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

となります（式展開では、$\hat{\beta}^*=(c_y/c_x)\hat{\beta}$、$\beta^*=(c_y/c_x)\beta$ に注意）。この結果から、スケール変更により、t 統計量は影響を受けないことが確認できました。

7. 以下では、最小 2 乗推定量$\hat{\alpha}$と$\hat{\beta}$の確率的表現を用いるので再掲載します。

$$\hat{\alpha}=\alpha-(\hat{\beta}-\beta)\overline{X}+\overline{u}$$

$$\hat{\beta}=\beta+\frac{\sum_{i=1}^{n}(X_i-\overline{X})u_i}{\sum_{i=1}^{n}(X_i-\overline{X})^2}$$

ここで$\hat{\alpha}$の分散は

$$E[(\hat{\alpha}-\alpha)^2]=E[(-(\hat{\beta}-\beta)\overline{X}+\overline{u})^2]$$

$$=\overline{X}^2E[(\hat{\beta}-\beta)^2]-2\overline{X}E[(\hat{\beta}-\beta)\overline{u}]+E[\overline{u}^2]$$

と 3 つに分解できます。第 1 項は、$\hat{\beta}$の分散式である$\sigma_{\hat{\beta}}^2 = \sigma^2 / \sum_{i=1}^{n}(x_i - \bar{x})^2$を用いて計算できます。第 2 項は、

$$E[(\hat{\beta}-\beta)\bar{u}] = E\left[\left(\frac{\sum_{i=1}^{n}(X_i - \bar{X})u_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right)\left(\frac{\sum_{i=1}^{n} u_i}{n}\right)\right]$$

$$= \frac{1}{n\sum_{i=1}^{n}(X_i - \bar{X})^2} E\left[\left(\sum_{i=1}^{n}(X_i - \bar{X})u_i\right)\left(\sum_{i=1}^{n} u_i\right)\right] = 0$$

を用いると 0 となることを示せます[2]。第 3 項は

$$E[\bar{u}^2] = E\left[\left(\frac{\sum_{i=1}^{n} u_i}{n}\right)^2\right]$$

$$= \frac{1}{n^2} E\left[\sum_{i=1}^{n} u_i^2 + \sum_{i \neq j}\sum u_i u_j\right] = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

となります（i≠j のとき $E[u_i u_j]=0$ を用いました）。以上から、

$$E[(\hat{\alpha} - \alpha)^2] = \sigma^2\left(\frac{\bar{X}^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} + \frac{1}{n}\right)$$

と書けます。この式から、n が大きくなるにつれて、（両項とも分母が無限大に大きくなるため）分散が 0 に近づくのは明らかでしょう。

分散は次のようにも書き換えられます。

$$E[(\hat{\alpha} - \alpha)^2] = \frac{\sigma^2 \bar{X}^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} + \frac{\sigma^2}{n}$$

$$= \frac{\sigma^2 n \bar{X}^2}{n\sum_{i=1}^{n}(X_i - \bar{X})^2} + \frac{\sigma^2}{n}\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

$$= \frac{\sigma^2 n \bar{X}^2}{n\sum_{i=1}^{n}(X_i - \bar{X})^2} + \sigma^2 \frac{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2}{n\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{\sigma^2 \sum_{i=1}^{n} X_i^2}{n\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

[2] 期待値の部分が 0 となることを n=2 として証明しましょう。

$$E\left[\left(\sum_{i=1}^{2}(X_i - \bar{X})u_i\right)\left(\sum_{i=1}^{2} u_i\right)\right] = E\left[\left((X_1 - \bar{X})u_1 + (X_2 - \bar{X})u_2\right)(u_1 + u_2)\right]$$
$$= (X_1 - \bar{X})E[u_1^2] + (X_2 - \bar{X})E[u_2^2] + (X_1 - \bar{X})E[u_1 u_2] + (X_2 - \bar{X})E[u_1 u_2]$$
$$= (X_1 - \bar{X})\sigma^2 + (X_2 - \bar{X})\sigma^2 = \sigma^2 \sum_{i=1}^{2}(X_i - \bar{X}) = 0$$

式展開において、$E[u_1^2]=\sigma^2$、$E[u_2^2]=\sigma^2$、$E[u_1 u_2]=0$、また偏差の和は 0 を用いました。