第3章
最小2乗推定量の統計的性質

藪友良

『入門　実践する計量経済学』

(東洋経済新報社)

PPT

・確率的モデル

・回帰分析における標準的仮定

・最小2乗推定量の確率的性質

・最小2乗推定量の分散の推定方法

・区間推定

確率的モデル

- **単回帰モデル**

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

--- Y :被説明変数、X :説明変数

u :誤差項は確率変数

- **なぜ誤差項が含まれるのか？**

---Yを家賃、Xを面積としたとき、他要因が存在する

--- これらをすべて足し合わせものを誤差項とする

- **なぜ誤差項を確率変数とするのか？**

--- 築年数、階数、駅からの距離など全て同じ条件でも家賃は異なる

--- 人間は神のように全ての構造を知らない

回帰分析における標準的仮定

標準的仮定

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

仮定1:説明変数X_iは確率変数ではなく、固定した値を持つ

仮定2:n が大きくなると、$\sum_{i=1}^{n}(X_i - \bar{X})^2$は$\infty$に近づく

仮定3:$E[u_i] = 0$

$E\left[(u_i - E[u_i])^2\right]$

仮定4:$V(u_i) = E[u_i^2] = \sigma^2$

$E\left[(u_i - E[u_i])(u_j - E[u_j])\right]$

仮定5:$Cov\left(u_i, u_j\right) = E[u_i u_j] = 0$

仮定6: u_iは正規分布に従う

$u_i \sim i.i.d.\, N(0, \sigma^2)$

仮定2の意味

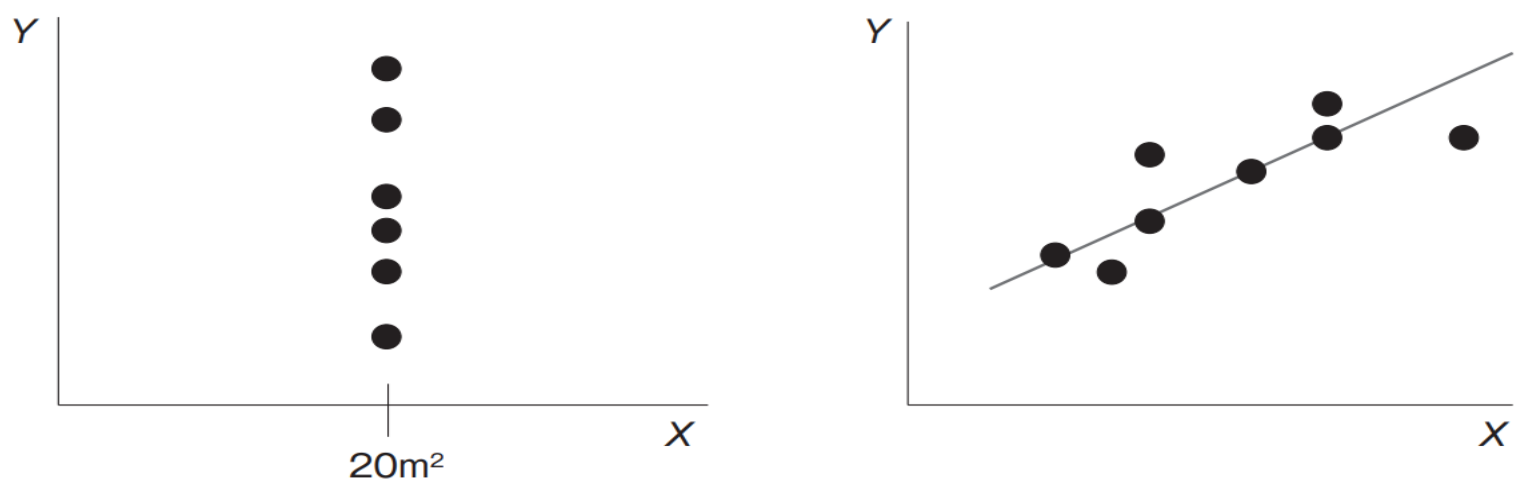
n が大きくなると、$\sum_{i=1}^{n}(X_i - \bar{X})^2$ は∞に近づく

・2乗和であるため、Xに多少の変動があれば必ず満たされる

・Xに変動がまったくなければ、仮定は満たされない。

図３-１　説明変数の変動と係数 β の推定との関係

（a）すべてのXが20m²の場合　（b）Xにばらつきがある場合



Xに変動がない限り、XとYの安定的な関係を測ることはできない

仮定3（$E[u_i]=0$）の意味

・X以外の要因を足し合わせたものが、なぜ期待値0なのか？

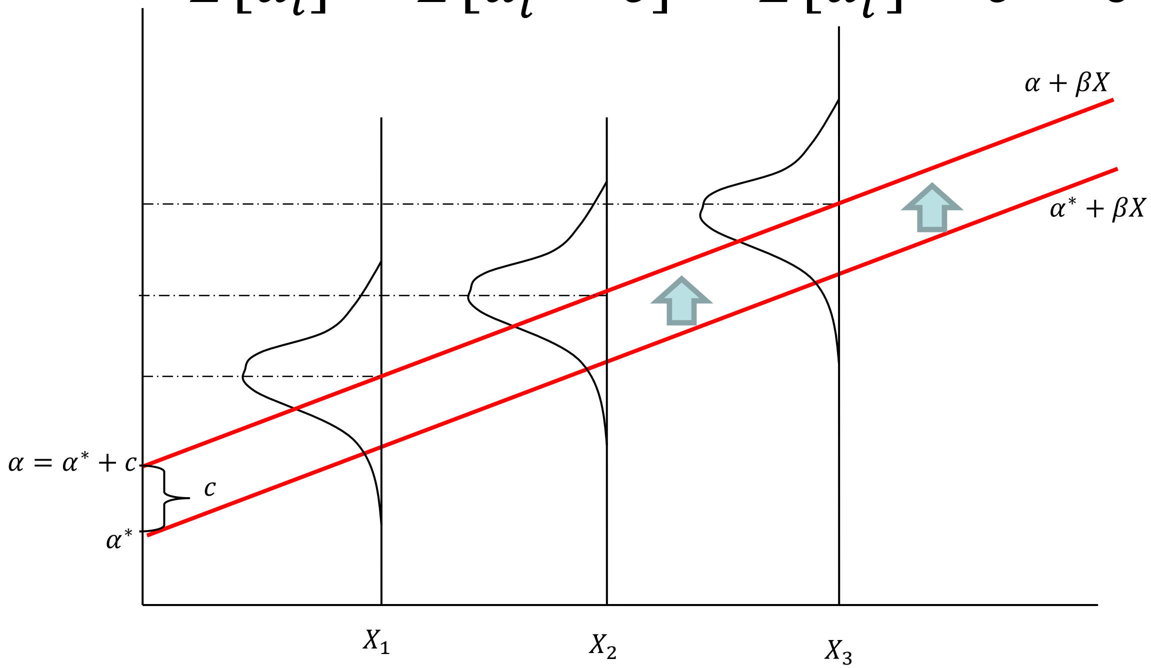
答え：　定数項αは$E[u_i]=0$が満たされるように定義されているから

[証明]　$Y_i=\alpha^*+\beta X_i+u_i^*$　ただし、$E[u_i^*]=c\neq 0$

$$\Rightarrow\ Y_i=(\alpha^*+c)+\beta X_i+(u_i^*-c)$$
$$=\alpha+\beta X_i+u_i$$

ここで$\alpha=\alpha^*+c$、$u_i=u_i^*-c$と定義した。

$$E[u_i]=E[u_i^*-c]=E[u_i^*]-c=0$$



仮定4の意味

$$V(u_i) = E[u_i^2] = \sigma^2 \text{ for any } i$$

・全てのiについてu_iの分散はσ^2で一定（均一分散）

---無作為抽出で選ばれたデータなら成立する

---通常、$V(u_i) = \sigma_i^2$となる（不均一分散）

仮定5の意味

$$Cov(u_i, u_j) = E[u_i u_j] = 0 \text{ for any } i \neq j$$

・無作為抽出した横断面データなら問題ない

時系列データでは成立しない

---無作為抽出ならi番目とj番目のデータは無関係になる

---時系列データでは、u_tとu_sは相関している（自己相関）

仮定6の意味
u_iは正規分布に従う

・X以外の要因を足し合わせたものが、なぜ正規分布に従うのか

----X以外の要因を$e_1, e_2, e_3, \ldots, e_m$と定義しよう

$$Y_i = \alpha + \beta X_i + (e_1 + e_2 + e_3 + \cdots + e_m)$$

$$\Rightarrow \quad u = e_1 + e_2 + e_3 + \cdots + e_m$$

mが十分に大きいとき、中心極限定理によって

正規分布に従う

最小2乗推定量の確率的性質

最小2乗推定量の別表現

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$$

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

$\Rightarrow$

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

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

--- 最小2乗推定量$\hat{\alpha}$, $\hat{\beta}$は確率変数uの関数であり、それ自体も確率変数である

$\hat{\alpha}$の別表現の導出

[証明]

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}$$
$$= \frac{\sum_{i=1}^{n}(\alpha + \beta X_i + u_i)}{n}$$
$$= \frac{n\alpha + \beta \sum_{i=1}^{n} X_i + \sum_{i=1}^{n} u_i}{n}$$
$$= \alpha + \beta \bar{X} + \bar{u}$$

次に、$\bar{Y} = \alpha + \beta \bar{X} + \bar{u}$を、$\hat{\alpha}$の公式に代入する

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$$
$$= \left(\alpha + \beta \bar{X} + \bar{u}\right) - \hat{\beta}\bar{X}$$
$$= \alpha - (\hat{\beta} - \beta)\bar{X} + \bar{u}$$

$\hat{\beta}$の別表現の導出

[証明]

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

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

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

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

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

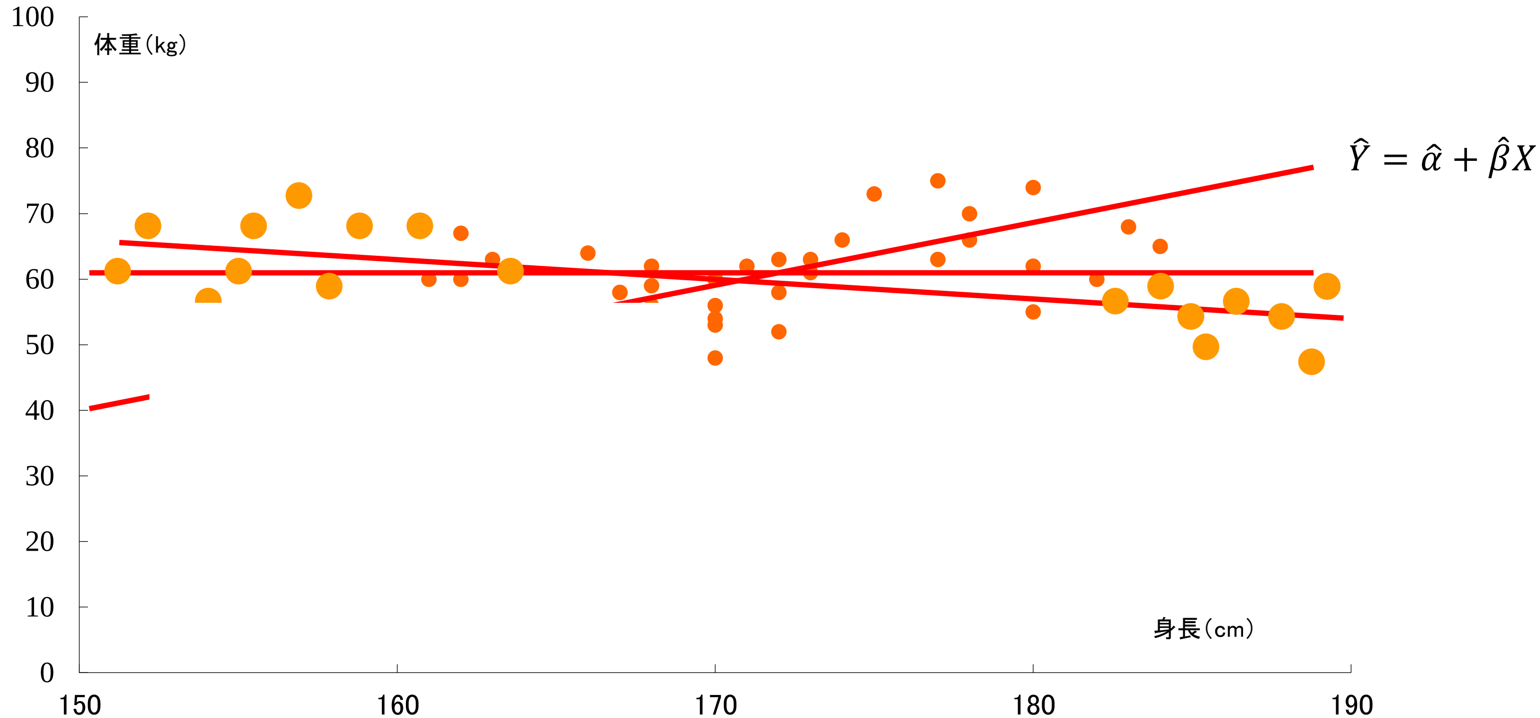
$$\sum_{i=1}^{n}(X_i - \bar{X})(X_i - \bar{X}) = \sum_{i=1}^{n}(X_i - \bar{X})X_i - \bar{X}\sum_{i=1}^{n}(X_i - \bar{X})$$

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

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

例） 身長と体重の関係

- Yを体重、Xを身長とする



- 得られたデータによって、最小2乗推定量の値が確率的に変動する

最小2乗推定量の期待値と分散

$$\boxed{\begin{gathered}\hat{\alpha} = \alpha - (\hat{\beta} - \beta)\bar{X} + \bar{u} \\ \hat{\beta} = \beta + \frac{\sum_{i=1}^{n}(X_i - \bar{X})\,u_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\end{gathered}}$$

・**期待値**

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

$$E[\bar{u}] = E\left[\frac{\sum_{i=1}^{n} u_i}{n}\right] = \frac{\sum_{i=1}^{n} E[u_i]}{n} = 0$$

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

--- 最小2乗推定量は不偏性をもっている

・**分散**

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

① σ^2が小さいほど、最小2乗推定量の分散も小さい

② Xのばらつきが大きいほど、最小2乗推定量の分散は小さい

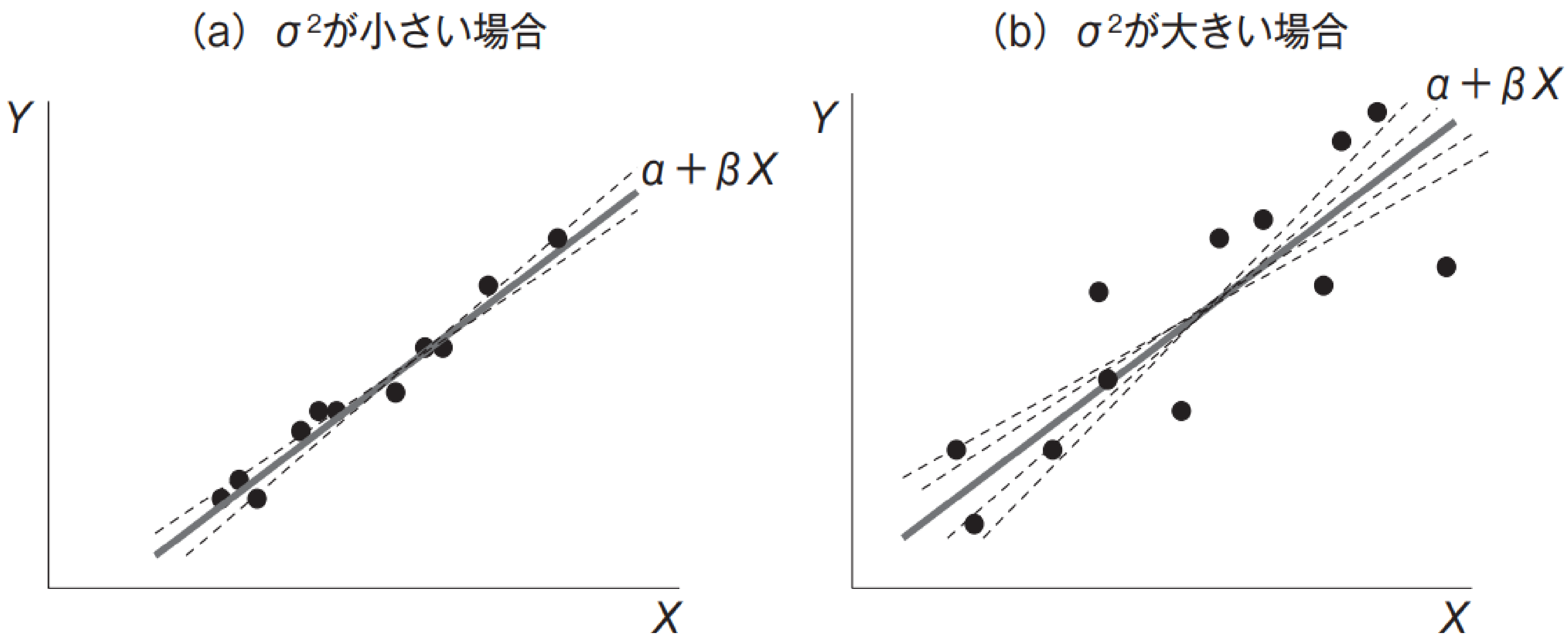
③ nが増えると、最小2乗推定量の分散は0に収束する

---最小2乗推定量は一致性をもっている

分散の性質①

① σ^2が小さいほど、最小2乗推定量の分散も小さい

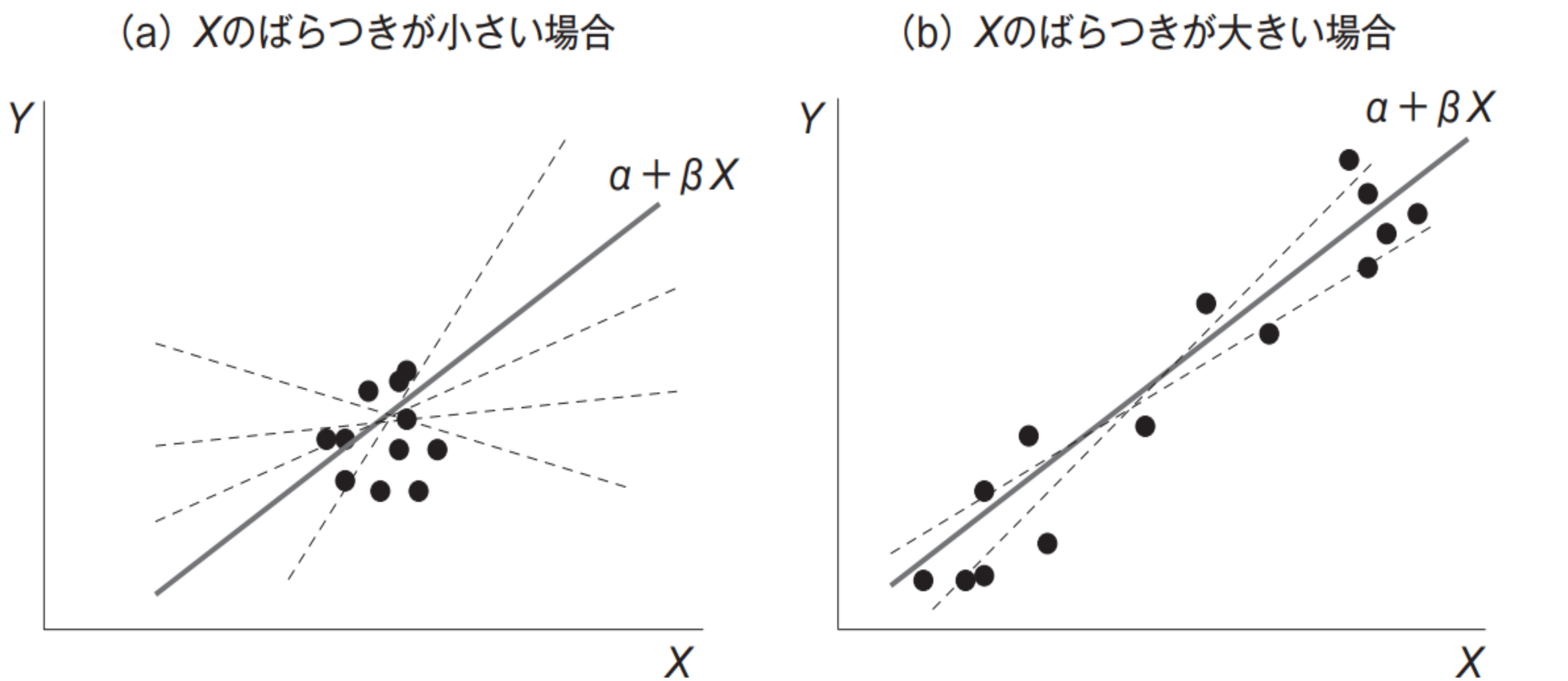
図 3 - 3　誤差項の分散σ^2 が OLS 推定量の分散に与える影響



分散の性質②

②Xのばらつきが大きいほど、最小2乗推定量の分散は小さい

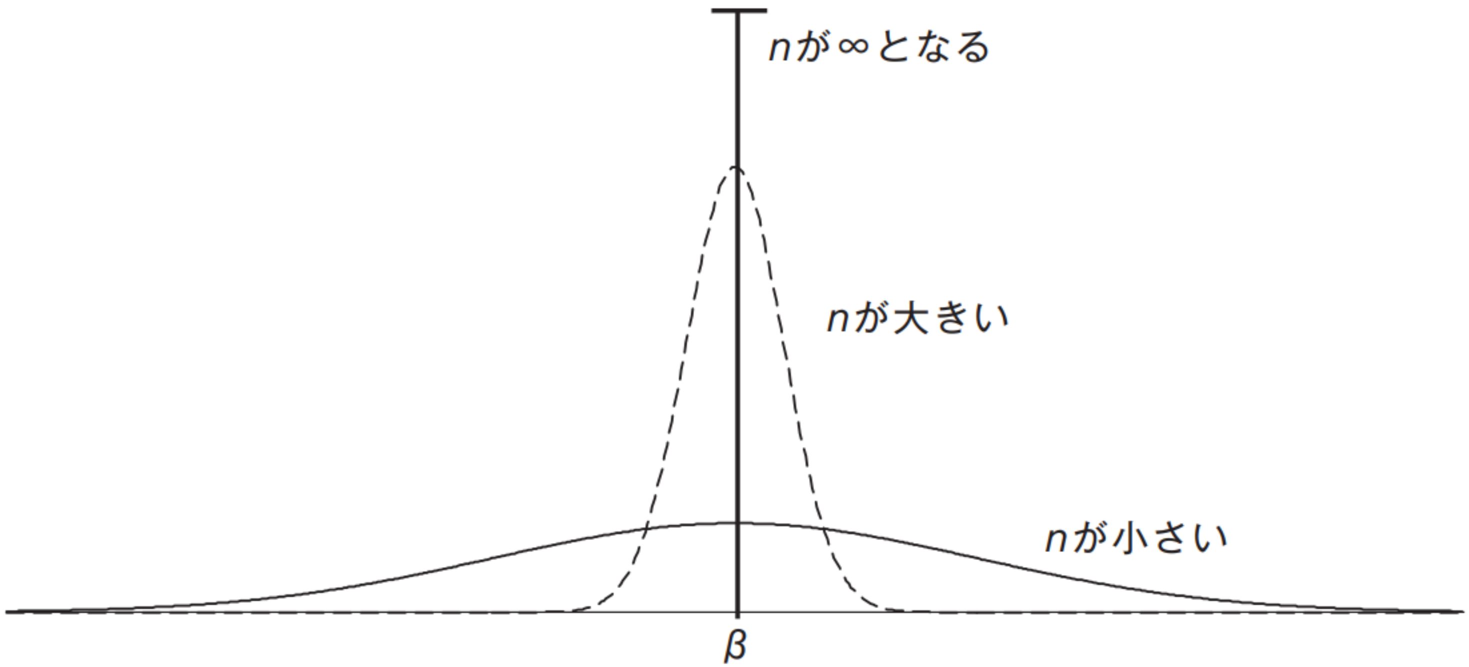
図 3 - 4　説明変数 X のばらつきが推定量の分散に与える影響



分散の性質③

③nが増えると最小2乗推定量の分散は0に

図 3 - 5　サンプルサイズ n と OLS 推定量の分布の関係



--- 不偏性と分散が0になることから、最小2乗推定量は一致性をもつ

$\hat{\boldsymbol{\beta}}$の分散$\mathrm{V}(\hat{\beta})$の導出

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

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

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

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

$$\boxed{E\left[u_i^2\right] = \sigma^2}$$

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

式展開では、以下に注意($n=2$の場合)

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

$$= \left(X_1-\bar{X}\right)^2E\left[u_1^2\right] + \left(X_2-\bar{X}\right)^2E\left[u_2^2\right] + 2\left(X_1-\bar{X}\right)\left(X_2-\bar{X}\right)\underbrace{E[u_1u_2]}_{=0}$$

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

最小2乗推定量の分布

$$\boxed{E\left[\hat{\beta}\right] = \beta, \quad \mathrm{V}(\hat{\beta}) = \frac{\sigma^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}}$$

・最小2乗推定量は**正規確率変数の線形関数であるため、全体が正規分布に従う(正規分布の再生性)**

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

$$= \beta + \frac{(X_1 - \bar{X})u_1 + (X_2 - \bar{X})u_2 + \dots + (X_n - \bar{X})u_n}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

$$= \beta + \underbrace{\left(\frac{X_1 - \bar{X}}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right)}_{c_1} u_1 + \underbrace{\left(\frac{X_2 - \bar{X}}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right)}_{c_2} u_2 + \dots + \underbrace{\left(\frac{X_n - \bar{X}}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right)}_{c_n} u_n$$

まとめると

$$\boxed{\hat{\beta} \sim N(\beta, V(\hat{\beta}))}$$

誤差項u_iと残差$\hat{u}_i$の違いとは何か

- モデルは$Y_i = \alpha + \beta X_i + u_i$から、誤差項$u_i$は

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

--- 誤差項u_iは、実現値Y_iと真の回帰式からの予測値$\alpha + \beta X_i$との差

- 残差$\hat{u}_i$は、

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

--- 残差$\hat{u}_i$は、実現値Y_iと推定された回帰式からの予測値$\hat{\alpha} + \hat{\beta} X_i$との差

- 残差$\hat{u}_i$は誤差項u_iの推定量

---誤差項u_iは観察できないが、残差$\hat{u}_i$は観察できる

最小2乗推定量の分散の推定方法

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

▪ σ^2の望ましい推定方法

$$\sigma^2 = E[(u_i - 0)^2] = E[u_i^2]$$

$$\Rightarrow \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^{n} u_i^2}{n}$$

▪誤差項u_iは観察できないため、残差$\hat{u}_i$で代用する

$$s^2 = \frac{\sum_{i=1}^{n} \hat{u}_i^2}{n-2}$$

▪最小2乗推定量の分散の推定量

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

s^2の性質

- $\boldsymbol{u_i \sim i.i.d.N(0,\sigma^2)}$、よって $\boldsymbol{\frac{u_i}{\sigma} \sim i.i.d.N(0,1)}$

$$\Rightarrow \ \sum_{i=1}^{n}\left(\frac{u_i}{\sigma}\right)^2 \sim \chi^2(n)$$

自由度(自由に動ける確率変数の数)はn

- $\boldsymbol{u_i}$は観察できないため、残差で置き換える

$$\sum_{i=1}^{n}\left(\frac{\hat{u}_i}{\sigma}\right)^2 \sim \chi^2(n-2)$$

自由度は$n-2$

--- 残差はn個あるが、2つの式(正規方程式)を満たさなければいけない

① $\sum_{i=1}^{n}(Y_i - \hat{\alpha} - \hat{\beta}X_i) = \sum_{i=1}^{n}\hat{u}_i = 0$

② $\sum_{i=1}^{n}(Y_i - \hat{\alpha} - \hat{\beta}X_i)X_i = \sum_{i=1}^{n}X_i\hat{u}_i = 0$

--- **自由に動ける残差は$\boldsymbol{n-2}$個しかない！！**

たとえば、$\hat{u}_1$、$\hat{u}_2$、…、$\hat{u}_{n-3}$、$\hat{u}_{n-2}$の残差がわかれば、

① $\sum_{i=1}^{n} \hat{u}_i = 0$

② $\sum_{i=1}^{n} X_i \hat{u}_i = 0$

①②の式を解くことで、$\hat{u}_{n-1}$と$\hat{u}_n$が得られる。つまり、

自由度は$n-2$

$s^2 = \frac{\sum_{i=1}^{n} \hat{u}_i^2}{n-2}$ は不偏性を満たす

[証明]

$$E[s^2] = E\left[\frac{\sum_{i=1}^{n} \hat{u}_i^2}{n-2}\right]$$

$$= E\left[\frac{\sigma^2}{n-2}\sum_{i=1}^{n}\left(\frac{\hat{u}_i}{\sigma}\right)^2\right]$$

$$= \frac{\sigma^2}{n-2} E\left[\sum_{i=1}^{n}\left(\frac{\hat{u}_i}{\sigma}\right)^2\right]$$

$\chi^2(n-2)$

$$= \frac{\sigma^2}{n-2}(n-2) = \sigma^2$$

最小2乗推定量の標準誤差

・標準誤差：推定量の標準偏差を推定したもの

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

例) 賃料Yと面積Xの関係を分析する

① 2.2節の表2のデータ(8物件：$n_1 = 8$)

$$\hat{Y} = 2.62 + 0.163X$$
$$(1.05) \quad (0.029)$$

② すべてのデータ(724物件： $n_2 = 724$)

$$\hat{Y} = 2.69 + 0.160X$$
$$(0.101) \quad (0.003)$$

サンプルサイズが約100倍に増えたことで標準誤差は約1/10に減少した

区間推定

t統計量

- $\hat{\beta} \sim N(\beta, V(\hat{\beta}))$を標準化すると

$$\Rightarrow \quad \frac{\hat{\beta}-\beta}{\sqrt{V(\hat{\beta})}} = \frac{\hat{\beta}-\beta}{\sqrt{\frac{\sigma^2}{\sum (X_i-\bar{X})^2}}} \sim N(0,1)$$

- $V(\hat{\beta}) = \frac{\sigma^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2}$ を$s_{\hat{\beta}}^2 = \frac{s^2}{\sum_{i=1}^{n}(X_i-\bar{X})^2}$で置き換えると

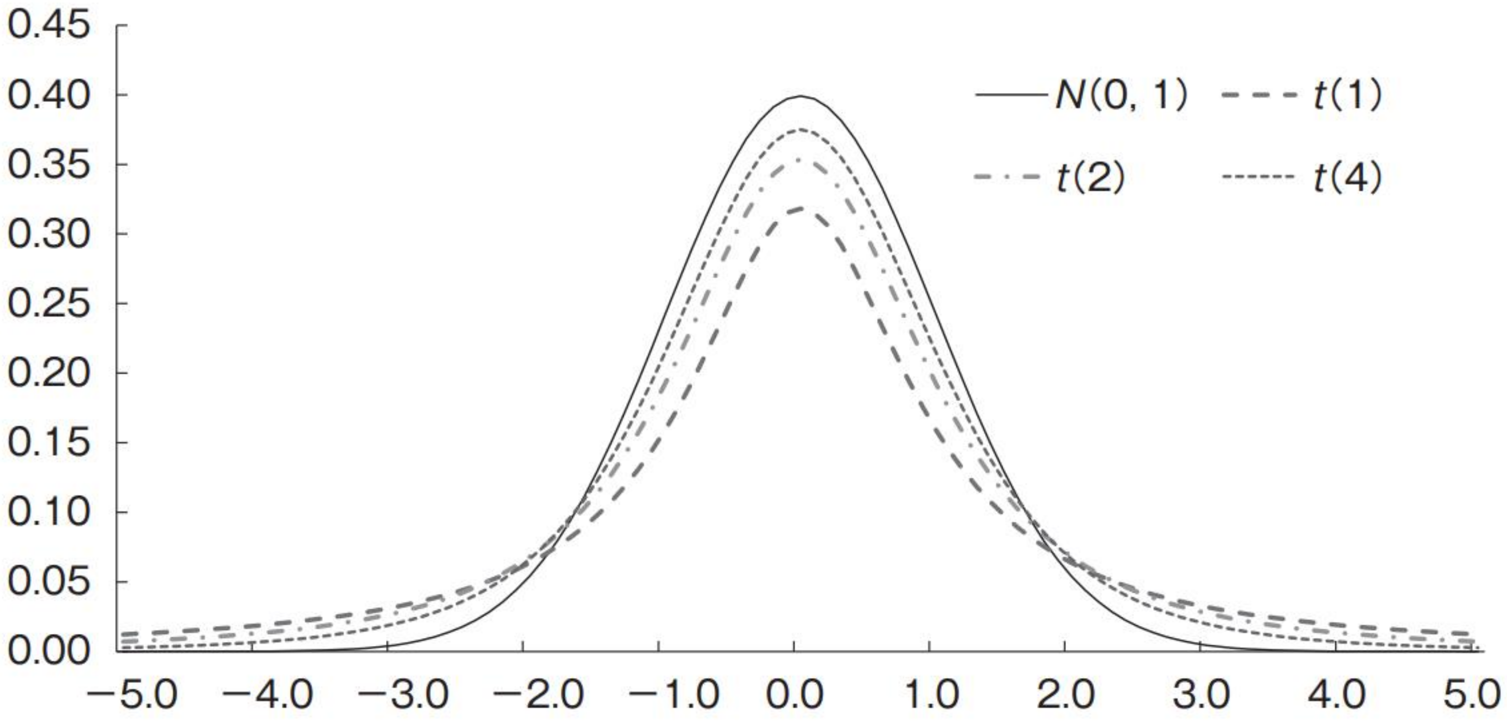
$$\Rightarrow \quad t_{\hat{\beta}} = \frac{\hat{\beta}-\beta}{\sqrt{s_{\hat{\beta}}^2}} = \frac{\hat{\beta}-\beta}{\sqrt{\frac{\sum \hat{u}_i^2/(n-2)}{\sum (X_i-\bar{X})^2}}}$$

$$= \frac{\frac{\hat{\beta}-\beta}{\sqrt{\sigma^2/\sum (X_i-\bar{X})^2}}}{\sqrt{\frac{1}{n-2}\sum\left(\frac{\hat{u}_i}{\sigma}\right)^2}} \sim t(n-2)$$

N(0,1)

$\chi^2(n-2)$

図３-７　自由度と *t* 分布の形状



サンプルサイズが大きくなると、自由度n-2も大きくなり、t分布は標準正規分布に近づく(nが30以上なら両者はほとんどかわらない)

信頼区間

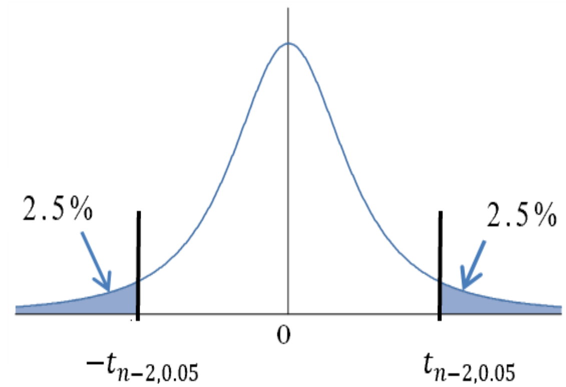
$$t_{\hat{\beta}} = \frac{\hat{\beta}-\beta}{\sqrt{s_{\hat{\beta}}^2}} \sim t(n-2)$$

$$\Rightarrow \ 0.95 = P\left\{-t_{n-2,0.05} < \frac{\hat{\beta}-\beta}{\sqrt{s_{\hat{\beta}}^2}} < t_{n-2,0.05}\right\}$$

$$= P\left\{\underline{\hat{\beta} - t_{n-2,0.05}\sqrt{s_{\hat{\beta}}^2} < \beta < \hat{\beta} + t_{n-2,0.05}\sqrt{s_{\hat{\beta}}^2}}\right\}$$

95%信頼区間

図 3-8　t 統計量の分布



95%信頼区間　$\hat{\beta} - t_{n-2,0.05}\sqrt{s_{\hat{\beta}}^2} < \beta < \hat{\beta} + t_{n-2,0.05}\sqrt{s_{\hat{\beta}}^2}$

90%信頼区間　$\hat{\beta} - t_{n-2,0.1}\sqrt{s_{\hat{\beta}}^2} < \beta < \hat{\beta} + t_{n-2,0.1}\sqrt{s_{\hat{\beta}}^2}$

99%信頼区間　$\hat{\beta} - t_{n-2,0.01}\sqrt{s_{\hat{\beta}}^2} < \beta < \hat{\beta} + t_{n-2,0.01}\sqrt{s_{\hat{\beta}}^2}$

例）同性愛者の割合

- 米国州別データ($n = 50$)
- X_i : 同性婚を支持する割合
- Y_i : 同性愛を自己申告する割合
- OLS推定すると、

$$\hat{Y}_i = 1.34 + 0.045 X_i$$
$$(0.509) \quad (0.011)$$

- $t_{n-2,0.05} = t_{48,0.05} = 2.01$を用いると、95%信頼区間は

$$\underbrace{0.045 - 2.01 \times 0.011}_{=0.023} < \beta < \underbrace{0.045 + 2.01 \times 0.011}_{=0.066}$$

---簡便法として、2.01の代わりに2を用いても、ほぼ同じ結果

$$\underbrace{0.045 - 2 \times 0.011}_{=0.023} < \beta < \underbrace{0.045 + 2 \times 0.011}_{=0.067}$$

まとめ

- **確率的モデル**

  ---誤差項の導入

- **回帰分析における標準的仮定**

- **最小2乗推定量の確率的性質**

  --- 最小2乗推定量は不偏性と一致性を満たす

- **最小2乗推定量の標準誤差**

  --- 最小2乗推定量のばらつきを表す指標

- **区間推定**

・ サンプルサイズを100倍すると標準誤差は1/10になる

$$s_{\widehat{\beta}} = \sqrt{\frac{s^2}{\sum_{i=1}^{n}(X_i - X)^2}} = \sqrt{\frac{\sum_{i=1}^{n} \hat{u}_i^2 /(n-2)}{n\sum_{i=1}^{n}(X_i - X)^2 /n}}$$

[証明] 標準誤差の構成要素のうち、

$$\frac{1}{n-2}\sum_{i=1}^{n} \hat{u}_i^2, \quad \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

はサンプルサイズを増やしても値はあまりかわらない。よって、$n_2 = 100n_1$として、標準誤差の比をとると、

$$\frac{\sqrt{\frac{\sum_{i=1}^{n_2} \hat{u}_i^2 /(n_2-2)}{n_2\sum_{i=1}^{n_2}(X_i - \bar{X})^2/n_2}}}{\sqrt{\frac{\sum_{i=1}^{n_1} \hat{u}_i^2 /(n_1-2)}{n_1\sum_{i=1}^{n_1}(X_i - \bar{X})^2/n_1}}} \approx \frac{\sqrt{\frac{1}{n_2}}}{\sqrt{\frac{1}{n_1}}} = \sqrt{\frac{n_1}{n_2}} = \sqrt{\frac{1}{100}} = \frac{1}{10}$$