統計学ノートNo.10
単回帰分析

担当教員：蛭川雅之
研究室：紫英館3階345研究室
メール： hirukawa@econ.ryukoku.ac.jp
オフィスアワー：毎週金曜日3講時

1　回帰分析の例

2つの変数間の関係

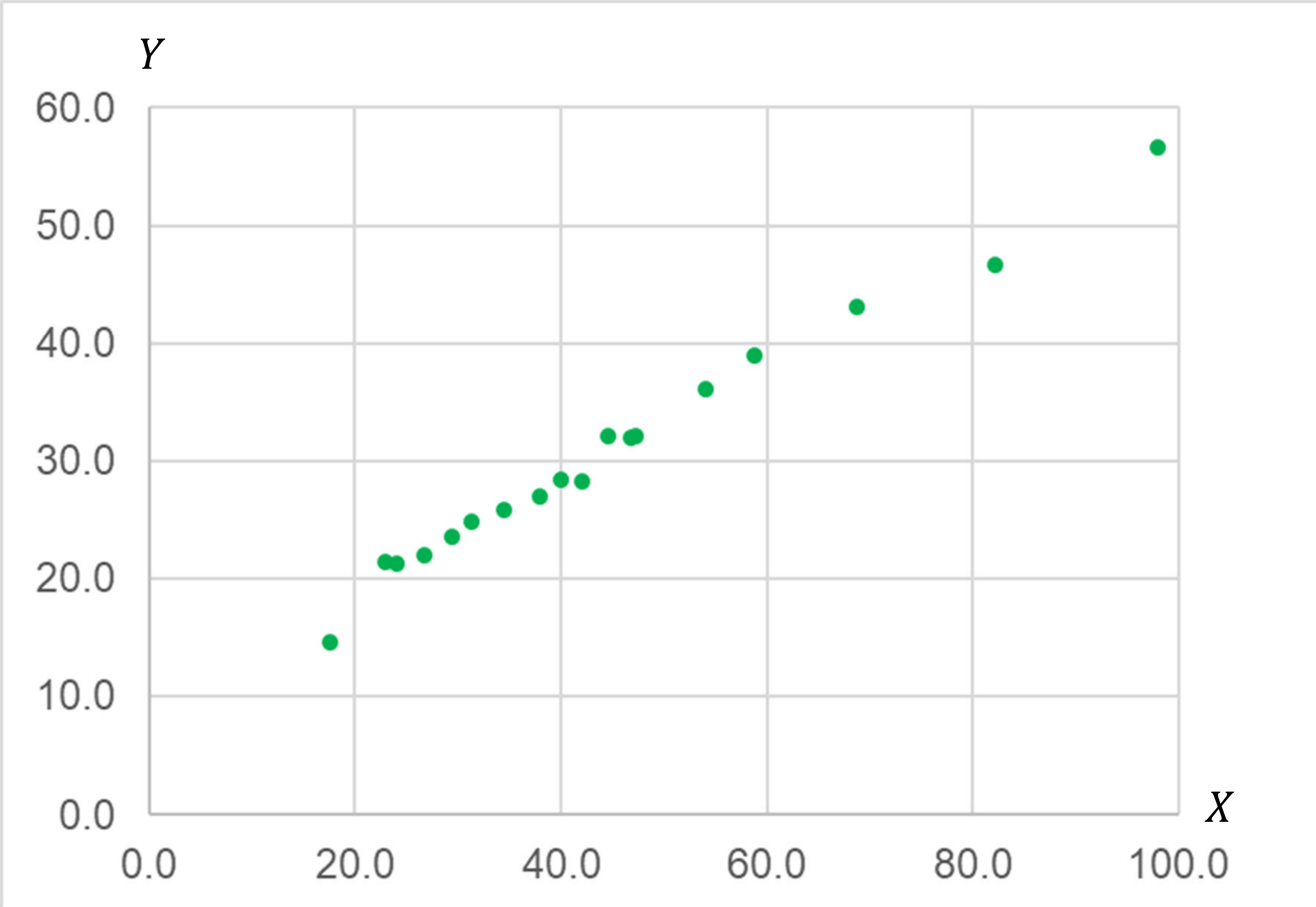
- 様々な分野で、2変数間の関係に関心を持つ場合が多い。
  - ある科目の試験の得点と学習時間
  - 所得と消費支出
  - 価格と需要量
  - 物価上昇率と失業率（フィリップス曲線）…
- 相関係数を見るだけでこの関係を十分把握できるか？

所得と消費の関係【データ】

- 2019年家計調査から以下のデータを抽出した。
  - $X=1$月当たり可処分所得（万円）
  - $Y=1$月当たり消費支出（万円）
  - $n=18$

X	17.5	23.0	24.1	26.7	29.5	31.3	34.5	38.0	40.0	42.1	44.5	47.3	46.8	54.0	58.8	68.8	82.3	97.9
Y	14.6	21.4	21.3	22.0	23.5	24.8	25.8	27.0	28.4	28.3	32.1	32.1	32.0	36.1	39.0	43.1	46.6	56.6

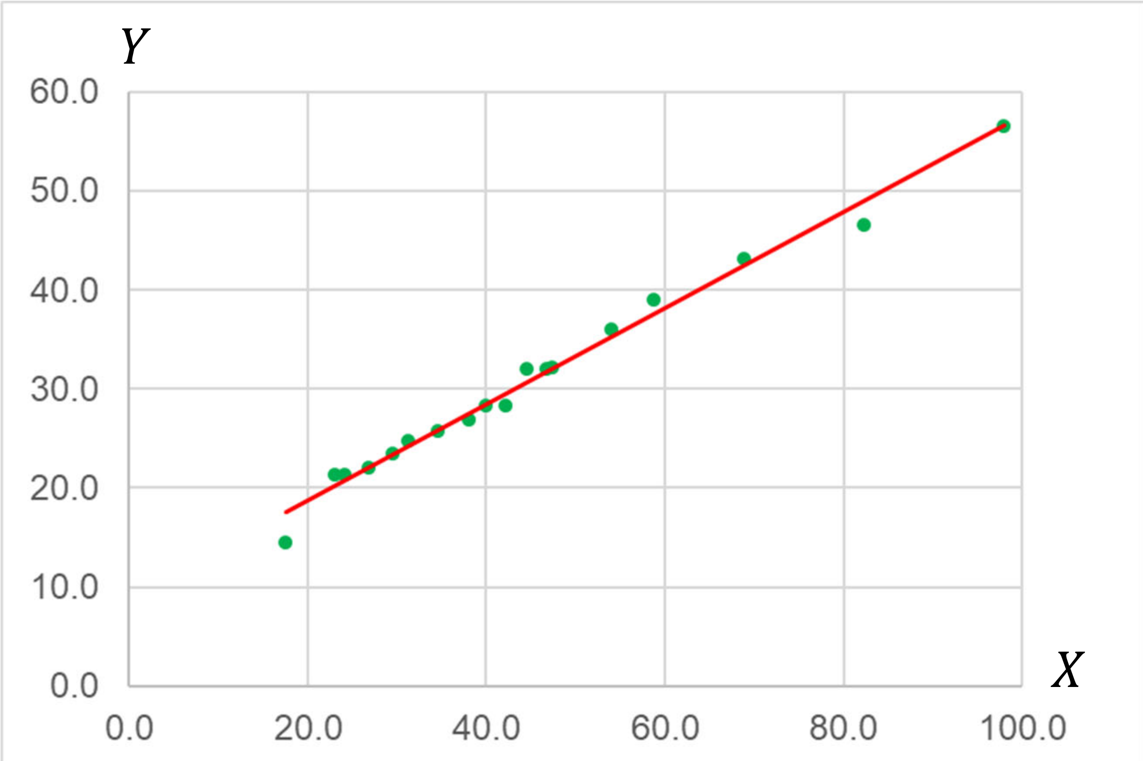
所得と消費の関係【散布図】



所得と消費の関係【疑問1】

- データに対して直線$Y = \beta_0 + \beta_1 X$をどのようにあてはめればよいか？

■これは**推定**の問題である。



所得と消費の関係【疑問2】

- **散布図を見る限り（あるいは、直感的に）直線の傾きは正であると予想される。**

■**仮説$\beta_1 > 0$はデータから支持されるか？**

■**これは検定の問題である。**

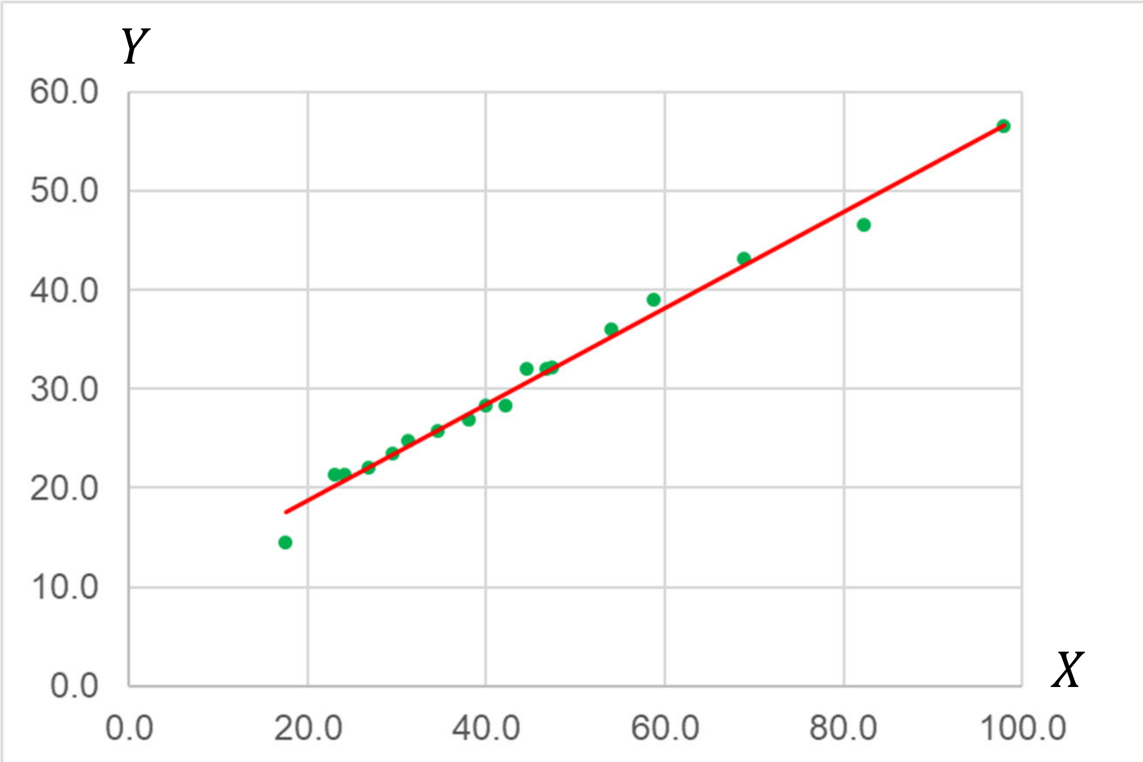
2　最小2乗法

直線の当てはめ

- **この散布図のデータに当てはめた直線は**

$$Y = 9.0639 + 0.4853X$$

である。



単回帰モデル

- n個のデータ$\{(X_i, Y_i)\}_{i=1}^{n}$の散布図を見る限り、(X, Y)には直線の関係が予想される。
  - 直線は$Y = \beta_0 + \beta_1 X$と表現される
  - ただし、切片β_0、傾きβ_1とも未知である。

単回帰モデル（つづき）

- n個のデータ$\{(X_i, Y_i)\}_{i=1}^{n}$全てが同じ直線上に乗っているわけではない。
  - 個々のデータには直線$Y = \beta_0 + \beta_1 X$では説明しきれない部分が存在する。
  - 誤差項ϵ_iを含め、各データを統一的に

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

と表現する。

単回帰モデル（つづき）

- **関係式**

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

では、2つの変数(X, Y)間に直線の関係を想定している。

- **この式を線形回帰モデルという。**

単回帰モデル（つづき）

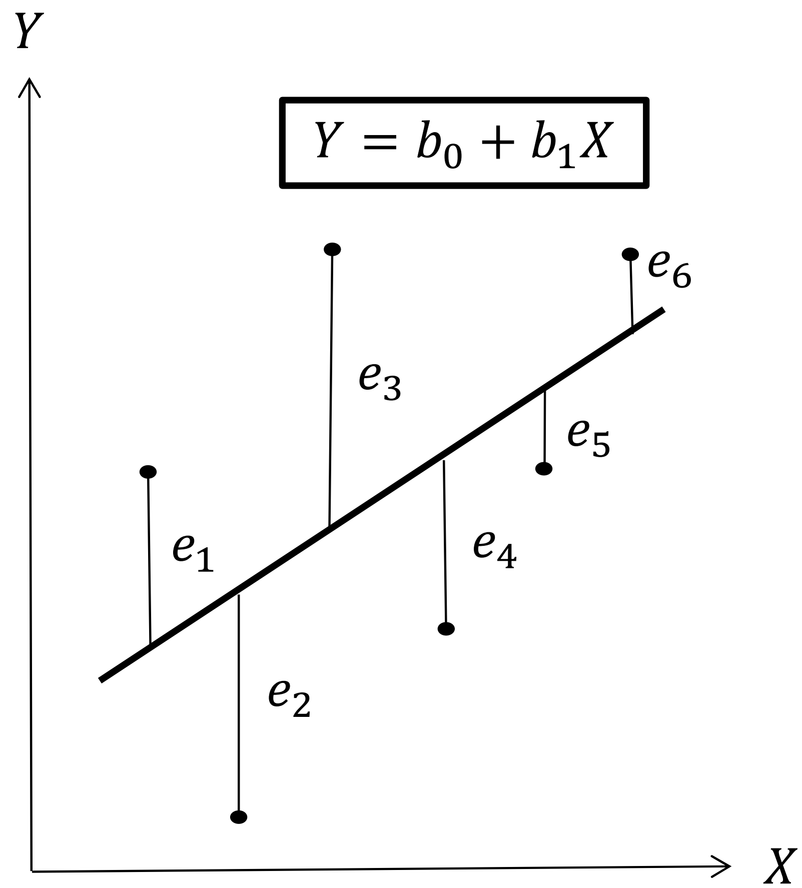
- Xは**説明変数（独立変数）**、Yは**被説明変数（従属変数）**と呼ばれる。
  - 説明変数が1つだけの線形回帰モデルを**単回帰モデル**という。

当てはまりの基準

- n個のデータ$\{(X_i, Y_i)\}_{i=1}^{n}$を使って、未知の回帰係数(β_0, β_1)をどのように推定するか？
  - そのためには、当てはまりの基準を決める必要がある。

残差の導入

- 右の散布図では、回帰係数(β_0, β_1)を予め(b_0, b_1)と決めてある。
- 全6点と直線との距離もしくは「外れ具合」$e_1, \ldots, e_6$が決まる。
  - これらを**残差**という。



最小2乗法

- データがn個の場合に話を戻す。
- 残差2乗和

$$SSR(b_0, b_1) = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \{Y_i - (b_0 + b_1 X_i)\}^2$$

を最小にするような値(b_0, b_1)を(β_0, β_1)の推定値とすればよいのではないか？

- このような回帰係数推定法を**最小2乗法**という。

単回帰モデルの最小2乗推定量

- **回帰係数**(β_0, β_1)**の最小**2**乗推定量を**$(\hat{\beta}_0, \hat{\beta}_1)$**と表記する。**
- **これらは簡単な公式で与えられる。**

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

単回帰モデルの最小2乗推定量（つづき）

- **傾きβ_1の最小2乗推定量は、以下のように表現することができる。**

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

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

$(\hat{\beta}_0, \hat{\beta}_1)$をどのように導出するのか？

- **偏微分を利用して**

$$SSR(b_0, b_1) = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - b_0 - b_1 X_i)^2$$

を最小にする(b_0, b_1)を求めてもよいが、ここでは原始的（直感的？）な方法（＝平方完成）を用いる。

$(\hat{\beta}_0, \hat{\beta}_1)$をどのように導出するのか？（つづき）

- **具体的には、次の特殊ケースを考える。**

1. **切片のみの回帰式**
2. **原点を通る回帰式**

切片のみの回帰式

- **まず回帰式**

$$Y_i = \beta_0 + \epsilon_i, i = 1, \ldots, n$$

の切片β_0**を最小**2**乗推定する。**

- **残差**2**乗和**

$$SSR(b_0) = \sum_{i=1}^{n} (Y_i - b_0)^2$$

を以下のようにb_0**について平方完成する。**

切片のみの回帰式（つづき）

$$SSR(b_0) = \sum_{i=1}^{n} {Y_i}^2 - 2n\bar{Y}b_0 + nb_0^2$$

$$= n(b_0 - \bar{Y})^2 + \sum_{i=1}^{n}(Y_i - \bar{Y})^2$$

- $SSR(b_0)$**は**$b_0 = \bar{Y}$**のとき最小となるから、**β_0**の最小**2**乗推定量は**$\hat{\beta}_0 = \bar{Y}$**である。**

原点を通る回帰式

- **次に、回帰式**

$$Y_i = \beta_1 X_i + \epsilon_i, i = 1, \dots, n$$

の傾きβ_1を最小2乗推定する。

- **残差2乗和**

$$SSR(b_1) = \sum_{i=1}^{n} (Y_i - b_1 X_i)^2$$

を以下のようにb_1について平方完成する。

原点を通る回帰式（つづき）

$$SSR(b_1)$$
$$= \sum_{i=1}^{n} {Y_i}^2 - 2\left(\sum_{i=1}^{n} X_i Y_i\right) b_0 + \left(\sum_{i=1}^{n} {X_i}^2\right) b_1^2$$
$$= \left(\sum_{i=1}^{n} {X_i}^2\right)\left(b_1 - \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} {X_i}^2}\right)^2$$
$$+ \frac{\left(\sum_{i=1}^{n} {X_i}^2\right)\left(\sum_{i=1}^{n} {Y_i}^2\right) - \left(\sum_{i=1}^{n} X_i Y_i\right)^2}{\sum_{i=1}^{n} {X_i}^2}$$

原点を通る回帰式（つづき）

- **先程と同様の手順から、β_1の最小2乗推定量は**

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} {X_i}^2}$$

である。

最小2乗推定量の導出

STEP 1

- **単回帰モデル**

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, i = 1, \dots, n$$

において、傾きβ_1をある値b_1に固定する。

最小2乗推定量の導出（つづき）

STEP 1

- **右辺の$\beta_1 X_i = b_1 X_i$を左辺に移項し、**

$$\tilde{Y}_i = Y_i - b_1 X_i$$

と書き改めることにより、切片のみの回帰式

$$\tilde{Y}_i = \beta_0 + \epsilon_i$$

が得られる。

最小2乗推定量の導出（つづき）

<u>STEP 1</u>

- **この場合のβ_0の最小2乗推定量は次の通りである。**

$$\tilde{\beta}_0 = \tilde{\beta}_0(b_1) = \frac{1}{n}\sum_{i=1}^{n} \tilde{Y}_i = \bar{Y} - b_1\bar{X}$$

最小2乗推定量の導出（つづき）

STEP 2

- **残差**2**乗和**$SSR(b_0, b_1)$**に**

$$b_0 = \tilde{\beta}_0 = \tilde{\beta}_0(b_1) = \bar{Y} - b_1\bar{X}$$

を代入し、

$$(X_i^*, Y_i^*) = (X_i - \bar{X}, Y_i - \bar{Y})$$

と書き改めることにより以下の結果を得る。

最小2乗推定量の導出（つづき）

STEP 2

$$SSR\left(\tilde{\beta}_0, b_1\right) = \sum_{i=1}^{n} \{Y_i - (\bar{Y} - b_1\bar{X}) - b_1X_i\}^2$$

$$= \sum_{i=1}^{n} \{(Y_i - \bar{Y}) - b_1(X_i - \bar{X})\}^2$$

$$= \sum_{i=1}^{n} (Y_i^* - b_1X_i^*)^2$$

最小2乗推定量の導出（つづき）

STEP 2

- **右辺は原点を通る回帰式に対する残差2乗和であるから、β_1の最小2乗推定量は次の通りである。**

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} X_i^* Y_i^*}{\sum_{i=1}^{n} {X_i^*}^2} = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

最小2乗推定量の導出（つづき）

STEP 3

- $\tilde{\beta}_0$にSTEP　2の結果を代入することにより、最終的にβ_0の最小2乗推定量は次のようになる。

$$\hat{\beta}_0 = \tilde{\beta}_0(\hat{\beta}_1) = \bar{Y} - \hat{\beta}_1 \bar{X}$$

例題1

本ノートのスライドNo.4にある1月当たり可処分所得Xと同消費支出Yのデータから以下の結果を得た。

$$\bar{X} = 44.8360, \bar{Y} = 30.8234,$$

$$\hat{\sigma}_{XY} = 206.7856, \hat{\sigma}_X^2 = 426.0867$$

このデータから単回帰モデル$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$, $i = 1, \ldots, 18$ を最小2乗推定すると、回帰係数(β_0, β_1)の推定値$(\hat{\beta}_0, \hat{\beta}_1)$はどのようになるか。

例題1（つづき）

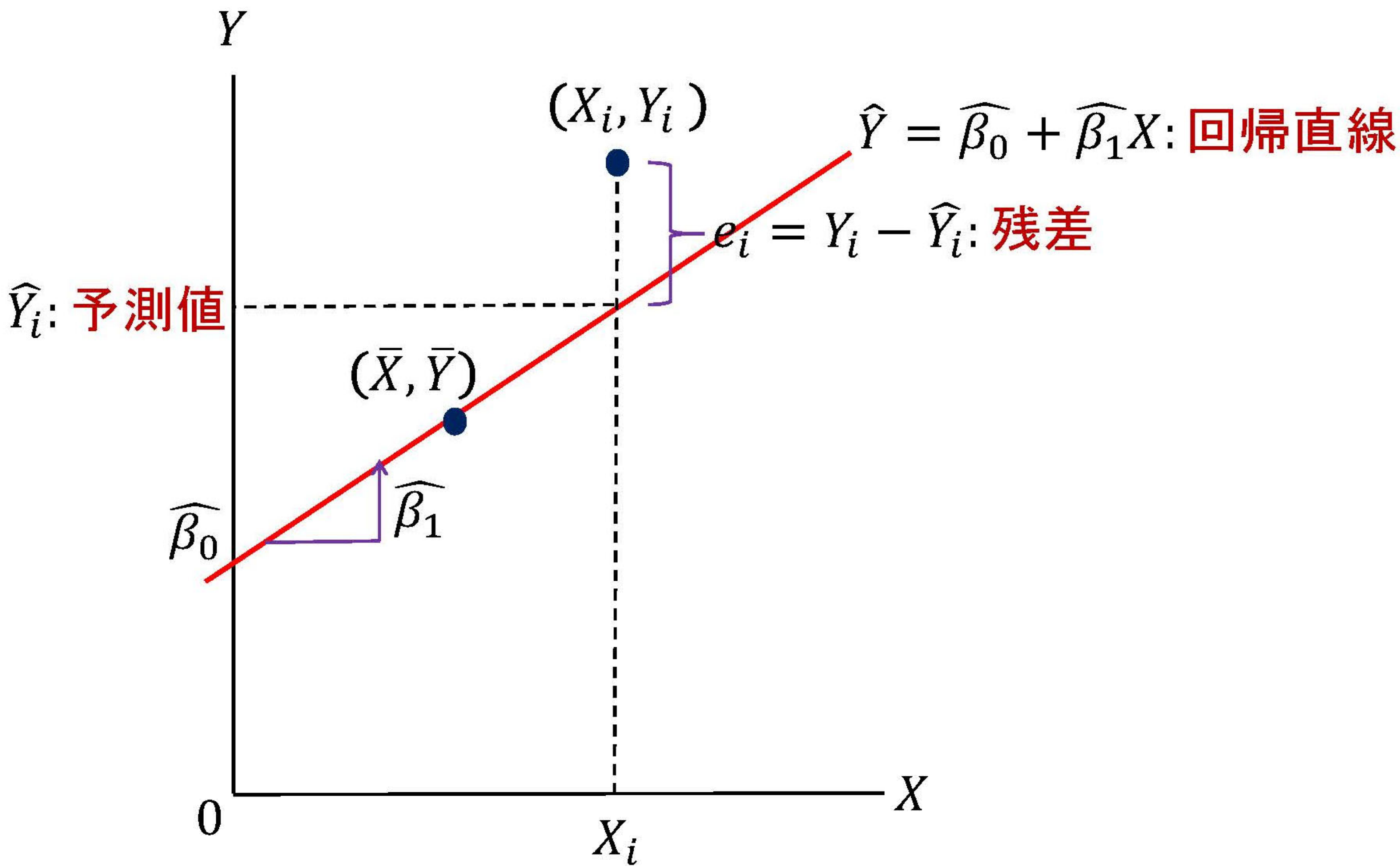
- **まず、傾きβ_1の最小2乗推定値を計算する。**

$$\hat{\beta}_1 = \frac{\hat{\sigma}_{XY}}{\hat{\sigma}_X^2} = \frac{206.7856}{426.0867} \approx 0.4853$$

- **次に、切片β_0の最小2乗推定値を計算する。**

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = 30.8234 - 0.4853 \times 44.8360 \approx 9.0639$$

回帰直線・予測値・残差


Y
(X_i, Y_i)
$\hat{Y} = \widehat{\beta_0} + \widehat{\beta_1}X$: 回帰直線
$e_i = Y_i - \widehat{Y_i}$: 残差
$\widehat{Y_i}$: 予測値
$(\bar{X}, \bar{Y})$
$\widehat{\beta_1}$
$\widehat{\beta_0}$
0
X
X_i

3　最小2乗推定量の性質

予測値・残差の性質

1. **予測値** $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i,\ i = 1, \dots, n$**の平均**

$$\bar{\hat{Y}} = \frac{1}{n}\sum_{i=1}^{n} \hat{Y}_i$$

は被説明変数の平均$\bar{Y}$**に等しい。**

2. **回帰直線**$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$**は点**$(\bar{X}, \bar{Y})$**を通る。**

予測値・残差の性質（つづき）

3. **残差** $e_i = Y_i - \hat{Y}_i,\ i = 1, \ldots, n$ **の和はゼロである。**

$$\sum_{i=1}^{n} e_i = 0$$

4. **説明変数と残差との積の和はゼロである。**

$$\sum_{i=1}^{n} X_i e_i = 0$$

分散分解

- **平均回りの被説明変数Yの偏差2乗和を全変動（Total Sum of Squares；略称TSS）と呼ぶ。**

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

分散分解（つづき）

- **平均回りの予測値$\hat{Y}$の偏差2乗和を回帰変動（Explained Sum of Squares；略称ESS）と呼ぶ。**

$$ESS = \sum_{i=1}^{n}\left(\hat{Y}_i - \bar{\hat{Y}}\right)^2 = \sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2$$

分散分解（つづき）

- **残差eの2乗和を残差変動（Residual Sum of Squares；略称RSS）と呼ぶ。**

$$RSS = \sum_{i=1}^{n} e_i^2$$

分散分解（つづき）

- **これら3つの2乗和には次のような関係がある。**

$$\underbrace{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}_{TSS} = \underbrace{\sum_{i=1}^{n}\left(\hat{Y}_i - \bar{Y}\right)^2}_{ESS} + \underbrace{\sum_{i=1}^{n} e_i^2}_{RSS}$$

$$\boxed{TSS = ESS + RSS}$$

- **この関係を分散分解という。**

４　推定モデルのあてはまりの尺度

重相関係数

- **重相関係数**R**は被説明変数**Y**とその予測値**$\hat{Y}$**に関する相関係数である。**

$$R = r_{Y\hat{Y}} = \frac{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})(\hat{Y}_i - \bar{Y})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}}$$

決定係数

- **決定係数**R^2**はは回帰変動（**ESS**）の全変動（**TSS**）に対する比率である。**

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

重相関係数と決定係数の性質

- $0 \le R, R^2 \le 1$である。
- **決定係数**R^2**は重相関係数**R**の**2**乗に等しい。**
- (X, Y)**の相関係数**r_{XY}**に対し、**

$$R^2 = r_{XY}^2$$

が成り立つ。

自由度修正済み決定係数

- 説明変数を追加すると、一般にR^2は増加する（少なくとも減少はしない）。
  - 無意味な説明変数を回帰モデルに追加してもR^2は改善する（＝1に近づく）。
- この問題の解決策として、**自由度修正済み決定係数**$\bar{R}^2$が提案されている。

$$\bar{R}^2 = 1 - \frac{RSS/(n-k-1)}{TSS/(n-1)}$$

k: 定数項を除く説明変数の数

自由度修正済み決定係数の性質

- $\bar{R}^2$の上限は1であるが、下限はゼロでなく負値も取りうる。
- 同じデータを使って同じ回帰モデルを推定する場合、$\bar{R}^2 \leq R^2$が成り立つ。

例題2

本ノートのスライドNo.4にある1月当たり可処分所得Xと同消費支出Yのデータから単回帰モデル$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i,\ i = 1, \ldots, 18$を最小2乗推定し、$ESS = 1806.4051$および$RSS = 23.7085$を得た。このとき、重相関係数$R$、決定係数$R^2$、自由度修正済み決定係数$\bar{R}^2$を計算せよ。また、$(X, Y)$の相関係数が$r_{XY} = 0.9935$であることを考慮して、結果についてコメントせよ。

例題2（つづき）

- **分散分解から**

$$TSS = ESS + RSS$$

$$= 1806.4051 + 23.7085 = 1830.1136$$

- **従って、決定係数R^2は**

$$R^2 = \frac{ESS}{TSS} = \frac{1806.4051}{1830.1136} \approx 0.9870$$

例題2（つづき）

- **また、重相関係数Rは**

$$R = \sqrt{R^2} = \sqrt{0.9870} \approx 0.9935$$

- **さらに、自由度修正済み決定係数$\bar{R}^2$は**

$$\bar{R}^2 = 1 - \frac{RSS/(n-2)}{TSS/(n-1)} = 1 - \frac{23.7085/16}{1830.1136/17}$$
$$\approx 0.9862$$

例題2（つづき）

- **計算結果から、以下の点が確認できる。**

$$\bar{R}^2 = 0.9862 \leq 0.9870 \ = R^2$$

$$R = r_{XY} = 0.9935$$

$$R^2 = r_{XY}^2 = 0.9935^2 = 0.9870$$

5　最小2乗推定量の統計的特性

古典的諸仮定

1. **誤差項ϵの平均はゼロである。**
2. **説明変数Xは誤差項ϵと無相関である。**
3. **各観測値に付随する誤差項ϵは相互に無相関である。**
4. **誤差項ϵの分散は均一である。**
5. **誤差項ϵは正規分布に従う。**

古典的諸仮定：コメント

- 仮定1-4（仮定1-5）を満たす誤差項を**古典的誤差項**（古典的正規誤差項）と呼ぶ。
- 仮定2が満たされない場合、最小2乗推定量は一致性を持たない。
- 仮定3または4が満たされない場合でも、最小2乗推定量は一致性を持つ。
  - ただし、最小2乗推定量は有効性を失う。

最小2乗推定量の統計的特性

- **仮定**1-4**の下で、**

1. **不偏性**：$E(\hat{\beta}_0) = \beta_0,\ E(\hat{\beta}_1) = \beta_1$
2. **一致性**：$\hat{\beta}_0 \xrightarrow{p} \beta_0, \hat{\beta}_1 \xrightarrow{p} \beta_1$

の両方が成り立つ。

最小2乗推定量の統計的特性（つづき）

- **さらに仮定5（誤差項の正規性）が加わると、誤差分散$\sigma^2 = Var(\epsilon)$に対し、以下が成り立つ。**

$$\hat{\beta}_0 \sim N\left(\beta_0, Var(\hat{\beta}_0)\right)$$
$$= N\left(\beta_0, \sigma^2\left\{\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right\}\right)$$

$$\hat{\beta}_1 \sim N\left(\beta_1, Var(\hat{\beta}_1)\right) = N\left(\beta_1, \frac{\sigma^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right)$$

ガウス＝マルコフ定理

- **仮定1-4の下で、最小2乗推定量$(\hat{\beta}_0, \hat{\beta}_1)$は、$(\beta_0, \beta_1)$の全ての線形不偏推定量の中で最小の分散を持つ。**
  - **線形推定量とは、ある定数$c_i,\ i = 1, \dots, n$を用いて$\sum_{i=1}^{n} c_i Y_i$の形に表現できる推定量を指す。**
  - **仮定5（誤差項の正規性）が加わると、最小2乗推定量は最小分散不偏推定量となることが知られている。**

6　標準誤差

標準誤差

- **最小2乗推定量$(\hat{\beta}_0, \hat{\beta}_1)$の分散**

$$Var(\hat{\beta}_0) = \sigma^2 \left\{ \frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} \right\}$$

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

は未知の定数$\sigma^2 (= Var(\epsilon))$に依存する。

- σ^2**をどのように推定するか？**

標準誤差（つづき）

- **まず、σ^2を$s^2 = \frac{RSS}{n-2}$で推定する。**
  - **この推定量の正の平方根**

$$\boxed{s = \sqrt{s^2} = \sqrt{\frac{RSS}{n-2}}}$$

は、しばしば回帰式の標準誤差（Standard Error of Regression；略称SER）と呼ばれる。

標準誤差（つづき）

- σ^2 を s^2 に置き換えることにより、$Var(\hat{\beta}_0)$ および $Var(\hat{\beta}_1)$ の推定量が得られる。

$$\widehat{Var}(\hat{\beta}_0) = s^2 \left\{ \frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} \right\}$$

$$\widehat{Var}(\hat{\beta}_1) = \frac{s^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

標準誤差（つづき）

- これらの正の平方根

$$SE(\hat{\beta}_0) = \sqrt{\widehat{Var}(\hat{\beta}_0)}$$

$$SE(\hat{\beta}_1) = \sqrt{\widehat{Var}(\hat{\beta}_1)}$$

は$\hat{\beta}_0$および$\hat{\beta}_1$の**標準誤差（Standard Error；略称SE）**と呼ばれる。

例題3

本ノート例題1・2の情報を利用して、回帰式の標準誤差、最小2乗推定量$\hat{\beta}_0$および$\hat{\beta}_1$の標準誤差をそれぞれ計算せよ。

例題3（つづき）

- **まず、誤差分散の推定量は**

$$s^2 = \frac{RSS}{n-2} = \frac{23.7085}{18-2} \approx 1.4818$$

- **従って、回帰式の標準誤差は**

$$s = \sqrt{1.4818} \approx 1.2173$$

例題3（つづき）

- $\hat{\beta}_0$および$\hat{\beta}_1$の標準誤差は以下のようになる。

$$SE(\hat{\beta}_0) = \sqrt{\widehat{Var}(\hat{\beta}_0)} = \sqrt{s^2\left\{\frac{1}{n} + \frac{\bar{X}^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right\}}$$

$$= \sqrt{s^2\left(\frac{1}{n} + \frac{\frac{1}{n}\bar{X}^2}{\hat{\sigma}_X^2}\right)} = \sqrt{\frac{s^2}{n}\left(1 + \frac{\bar{X}^2}{\hat{\sigma}_X^2}\right)}$$

$$= \sqrt{\frac{1.4818}{18}\left(1 + \frac{44.8360^2}{426.0867}\right)} \approx 0.6861$$

例題3（つづき）

$$SE(\hat{\beta}_1) = \sqrt{\widehat{Var}(\hat{\beta}_1)} = \sqrt{\frac{s^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}}$$

$$= \sqrt{\frac{s^2}{n\hat{\sigma}_X^2}} = \sqrt{\frac{1.4818}{18 \times 426.0867}} \approx 0.0139$$

推定結果の報告

- **回帰モデルの推定結果を報告する際には、係数推定値・標準誤差の両方を明示する。**
  - **係数推定値は点推定値、一方、標準誤差は推定の精度を示す。**
- **例えば、例題1-3の推定結果は次のように報告すればよい。**

$$\hat{Y} = \underset{(0.6861)}{9.0639} \quad \underset{(0.0139)}{+0.4853X}$$

$$n = 18, R^2 = 0.9870, \bar{R}^2 = 0.9862$$

7　回帰係数の仮説検定

回帰係数に関するt検定

- **傾きβ_1が特定の値bであるかどうかを検定する問題を考える。**
  - **帰無仮説は**

$$H_0: \beta_1 = b$$

である。

回帰係数に関するt検定（つづき）

- **多くの場合$b = 0$、即ち帰無仮説は**

$$H_0 : \beta_1 = 0$$

である。

  - **これは「説明変数Xは効いている（＝被説明変数Yに対して影響を及ぼしている）かどうか」に関する検定である。**

回帰係数に関するt検定（つづき）

- **しばしばゼロ以外の数値も用いられる。**
  - **需要の価格弾力性（両対数モデル）**$\Rightarrow b = -1$
  - CAPM $\Rightarrow b = 1$

回帰係数に関するt検定（つづき）

- β_1**の最小**2**乗推定値**$\hat{\beta}_1$**およびその標準誤差** $SE(\hat{\beta}_1)$**から、検定統計量として*t*統計量**

$$t_1 = \frac{\hat{\beta}_1 - b}{SE(\hat{\beta}_1)}$$

が定義される。

回帰係数に関するt検定（つづき）

- **標準的な統計パッケージでは、$b = 0$に対応する*t値***

$$t_1 = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)}$$

が自動的に計算される。

回帰係数に関するt検定（つづき）

- **帰無仮説H_0が正しい場合、$t_1 \sim t(n-2)$（＝自由度$(n-2)$のt分布）に従う。**
  - **標本数nが十分大きい場合は、漸近的に$t_1 \sim N(0,1)$として差し支えない。**
- **帰無仮説H_0と対立仮説H_1の組がどのように設定されているかにより、棄却域は異なる。**

回帰係数に関するt検定（つづき）

- **有意水準を**$\alpha(=0.05, 0.01 \ldots)$**とする。**

1. $H_1: \beta_1 \neq b$**の場合、棄却域は**$t(n-2)$**の両側**$100\alpha\%$**点より外側の領域になる（両側検定）。**

2. $H_1: \beta_1 < b$**の場合、棄却域は**$t(n-2)$**の左側**$100\alpha\%$**点より小さい部分の領域になる（左側検定）。**

3. $H_1: \beta_1 > b$**の場合、棄却域は**$t(n-2)$**の右側**$100\alpha\%$**点より大きい部分の領域になる（右側検定）。**

P値

- **標準的な統計パッケージでは、t値以外にP値も報告される。**

- **P値は帰無仮説H_0: $\beta_1 = 0$が真である場合に絶対値がt値以上の値が観測される確率を表す。**

P値（つづき）

- P値はこのH_0を**両側検定で棄却できる最小の有意水準**と解釈することもできる。
  - P値が予め定めた有意水準を下回っていれば、両側検定で帰無仮説$H_0: \beta_1 = 0$を棄却できる。

例題4

本ノート例題1-3の情報を利用して、帰無仮説 H_0: $\beta_1 = 0$ を H_1: $\beta_1 > 0$ に対し有意水準1%で検定せよ。

例題4（つづき）

- **検定統計量の値は**

$$t_1 = \frac{\hat{\beta}_1}{SE\left(\hat{\beta}_1\right)} = \frac{0.4853}{0.0139} \approx 34.91$$

である。

- **自由度**$18 - 2 = 16$**の**t**分布の右側**1%**点は** $t_{0.01} = 2.583$**である。**

例題4（つづき）

- $t_{0.01} = 2.583 < 34.91 = t_1$であるから、有意水準1%で$H_0$は棄却される。
  - β_1が正の値である（可処分所得が大きいと消費支出も大きくなる）強い証拠があると解釈される。