回帰分析

モデルの評価

村田 昇

講義の内容

- 第 1 回 : 回帰モデルの考え方と推定
- **第 2 回 : モデルの評価**
- 第 3 回 : モデルによる予測と発展的なモデル

回帰分析の復習

線形回帰モデル

- **目的変数** を **説明変数** で説明する関係式を構成
  - 説明変数 : $x_1, \dots, x_p$ (p 次元)
  - 目的変数 : y (1 次元)
- **回帰係数** $\beta_0, \beta_1, \dots, \beta_p$ を用いた一次式

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p$$

- **誤差項** を含む確率モデルで観測データを表現

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \epsilon_i \quad (i = 1, \dots, n)$$

簡潔な表現のための行列

- デザイン行列 (説明変数)

$$X = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}$$

簡潔な表現のためのベクトル

- ベクトル (目的変数・誤差・回帰係数)

$$\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}$$

問題の記述

- 確率モデル

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \text{確率分布}$$

- 回帰式の推定 : **残差平方和** の最小化

$$S(\boldsymbol{\beta}) = (\boldsymbol{y} - X\boldsymbol{\beta})^\top (\boldsymbol{y} - X\boldsymbol{\beta})$$

解の表現

- 解の条件 : **正規方程式**

$$X^\top X\boldsymbol{\beta} = X^\top \boldsymbol{y}$$

- 解の一意性 : **Gram 行列** $X^\top X$ が正則

$$\hat{\boldsymbol{\beta}} = (X^\top X)^{-1} X^\top \boldsymbol{y}$$

最小二乗推定量の性質

- **あてはめ値** $\hat{\boldsymbol{y}} = X\hat{\boldsymbol{\beta}}$ は X の列ベクトルの線形結合
- **残差** $\hat{\boldsymbol{\epsilon}} = \boldsymbol{y} - \hat{\boldsymbol{y}}$ はあてはめ値 $\hat{\boldsymbol{y}}$ と直交

$$\hat{\boldsymbol{\epsilon}}^\top \hat{\boldsymbol{y}} = 0$$

- 回帰式は説明変数と目的変数の **標本平均** を通過

$$\bar{y} = (1, \bar{\boldsymbol{x}}^\top)\hat{\boldsymbol{\beta}}, \quad \bar{\boldsymbol{x}} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{x}_i, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i,$$

寄与率

- **決定係数** (R-squared)

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \hat{\epsilon}_i^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

- **自由度調整済み決定係数** (adjusted R-squared)

$$\bar{R}^2 = 1 - \frac{\frac{1}{n-p-1}\sum_{i=1}^{n} \hat{\epsilon}_i^2}{\frac{1}{n-1}\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

  - 不偏分散で補正

日付	気温	降雨	日射	降雪	風向	風速	気圧	湿度	雲量
2024-10-01	23.3	0.5	11.45	0	NNW	2.6	1006.0	81	5.8
2024-10-02	26.5	0.0	18.32	0	S	2.9	1007.9	77	6.0
2024-10-03	23.1	11.0	5.88	0	E	2.7	1015.9	87	10.0
2024-10-04	25.9	2.0	12.60	0	S	3.5	1015.4	87	10.0
2024-10-05	21.3	9.5	1.88	0	NNE	2.5	1018.4	94	10.0
2024-10-06	21.3	0.0	5.01	0	NNW	1.7	1017.1	93	10.0
2024-10-07	25.0	0.0	14.99	0	S	2.9	1008.9	83	8.0
2024-10-08	18.8	33.5	1.98	0	NE	3.0	1008.9	97	10.0
2024-10-09	16.0	53.5	3.58	0	NNW	2.9	1009.3	93	10.0
2024-10-10	17.8	0.0	7.52	0	NNW	2.6	1009.8	75	6.0
2024-10-11	19.0	0.0	16.14	0	SSE	1.9	1013.1	69	7.5
2024-10-12	20.6	0.0	16.44	0	N	1.9	1019.0	73	2.5
2024-10-13	20.9	0.0	16.27	0	NNW	2.2	1021.1	70	0.8
2024-10-14	20.8	0.0	16.02	0	NNW	2.3	1022.6	71	4.0
2024-10-15	22.1	0.0	16.53	0	SSW	2.2	1020.3	72	3.8
2024-10-16	22.6	0.0	8.50	0	NNE	1.5	1017.3	76	7.5
2024-10-17	22.8	0.0	8.10	0	ENE	2.3	1020.0	79	9.3
2024-10-18	21.6	2.0	3.27	0	N	1.8	1019.5	92	10.0
2024-10-19	24.2	1.5	11.29	0	S	2.7	1009.2	84	10.0
2024-10-20	17.4	0.0	13.59	0	ENE	3.6	1023.6	55	5.8
2024-10-21	16.2	0.0	12.31	0	NW	2.7	1029.2	61	7.0
2024-10-22	19.7	0.0	12.02	0	NNW	1.9	1022.1	69	8.3
2024-10-23	21.9	6.5	4.24	0	NW	2.3	1012.3	90	10.0
2024-10-24	22.6	0.0	9.18	0	NE	2.2	1013.5	79	8.0
2024-10-25	20.2	0.5	3.61	0	NNE	2.4	1021.1	77	9.8
2024-10-26	19.0	0.0	3.90	0	NNW	1.7	1019.1	80	10.0
2024-10-27	19.7	4.0	8.46	0	NW	1.5	1011.4	87	9.5
2024-10-28	18.8	8.0	3.54	0	NE	1.9	1006.4	87	9.8
2024-10-29	15.4	24.5	2.79	0	NE	2.9	1017.4	79	10.0
2024-10-30	16.8	17.5	9.07	0	NW	3.2	1012.4	79	7.5
2024-10-31	16.2	0.0	11.86	0	NNE	2.0	1021.9	65	7.5

解析の事例

気温に影響を与える要因の分析

- データの概要

- 気温を説明する 5 種類の線形回帰モデルを検討
  - モデル 1 : 気温 = F(気圧)
  - モデル 2 : 気温 = F(日射)
  - モデル 3 : 気温 = F(気圧, 日射)
  - モデル 4 : 気温 = F(気圧, 日射, 湿度)
  - モデル 5 : 気温 = F(気圧, 日射, 雲量)

分析の視覚化

- 関連するデータの散布図

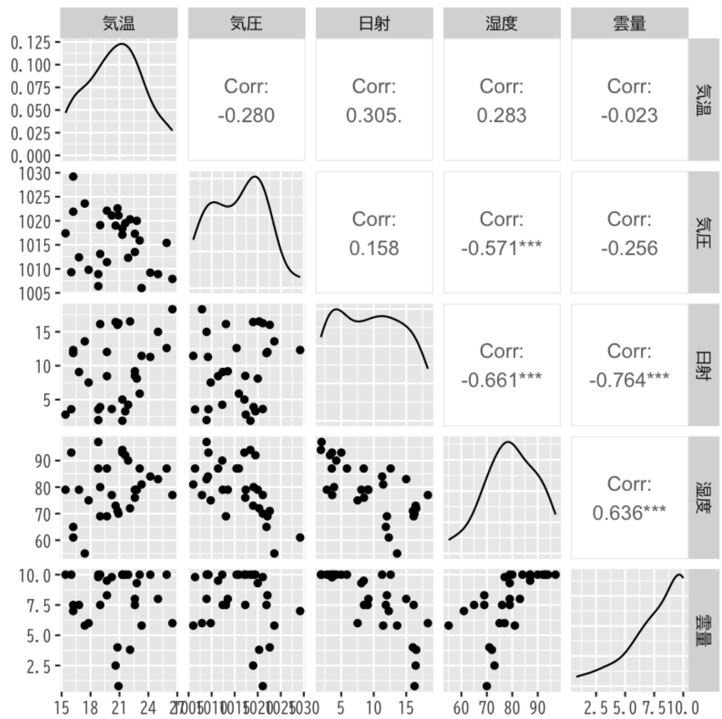

Figure 1: 散布図

	モデル 1		モデル 2		モデル 3		モデル 4		モデル 5	
変数	**係数**	**標準誤差**	**係数**	**標準誤差**	**係数**	**標準誤差**	**係数**	**標準誤差**	**係数**	**標準誤差**
気圧	-0.14	0.090			-0.17	0.086	0.06	0.088	-0.14	0.086
日射			0.17	0.101	0.21	0.098	0.53	0.109	0.38	0.146
湿度							0.28	0.067		
雲量									0.49	0.306
R^2	0.078		0.093		0.204		0.519		0.272	
Adjusted R^2	0.047		0.062		0.147		0.466		0.191	

Abbreviations: CI = Confidence Interval, SE = Standard Error

- モデル 1 の推定結果
- モデル 2 の推定結果
- モデル 3 の推定結果
- 観測値とあてはめ値の比較

モデルの比較

- 決定係数 (R^2, Adjusted R^2)

あてはめ値の性質

あてはめ値

- さまざまな表現

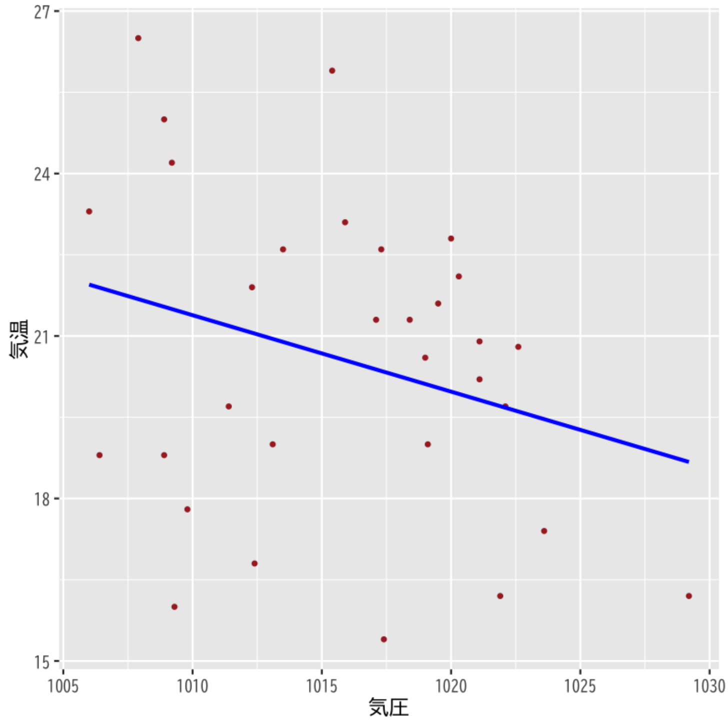

Figure 2: モデル 1

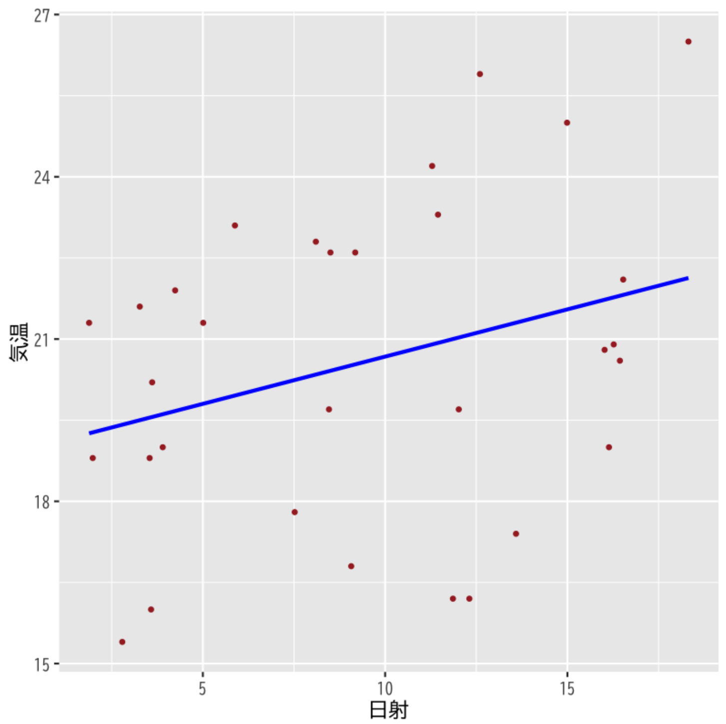

Figure 3: モデル 2

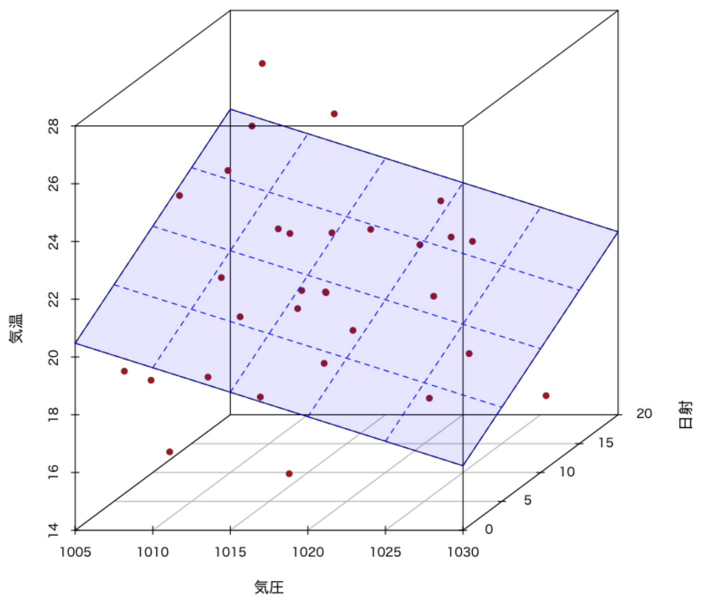

Figure 4: モデル 3

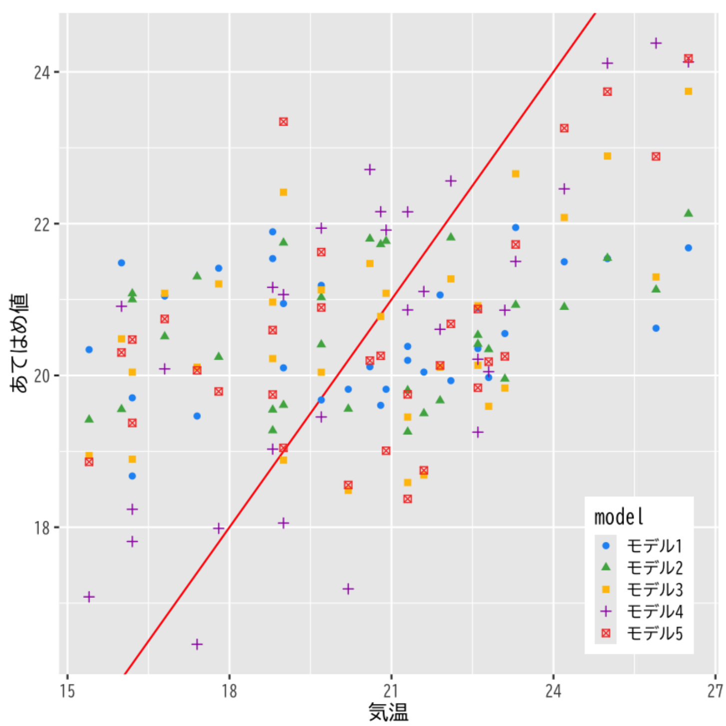

Figure 5: モデルの比較

$$
\begin{aligned}
\hat{\boldsymbol{y}} &= X\hat{\boldsymbol{\beta}} \\
&\qquad (\hat{\boldsymbol{\beta}} = (X^\top X)^{-1}X^\top \boldsymbol{y}\text{を代入}) \\
&= X(X^\top X)^{-1}X^\top \boldsymbol{y} && (A) \\
&\qquad (\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}\text{を代入}) \\
&= X(X^\top X)^{-1}X^\top X\boldsymbol{\beta} + X(X^\top X)^{-1}X^\top \boldsymbol{\epsilon} \\
&= X\boldsymbol{\beta} + X(X^\top X)^{-1}X^\top \boldsymbol{\epsilon} && (B)
\end{aligned}
$$

  - (A) あてはめ値は **観測値の重み付けの和** で表される
  - (B) あてはめ値と観測値は **誤差項** の寄与のみ異なる

あてはめ値と誤差

- 残差と誤差の関係

$$
\begin{aligned}
\hat{\boldsymbol{\epsilon}} &= \boldsymbol{y} - \hat{\boldsymbol{y}} \\
&= \boldsymbol{\epsilon} - X(X^\top X)^{-1}X^\top \boldsymbol{\epsilon} \\
&= \left(I - X(X^\top X)^{-1}X^\top\right)\boldsymbol{\epsilon} && (C)
\end{aligned}
$$

  - (C) 残差は **誤差の重み付けの和** で表される

ハット行列

- 定義

$$
H = X(X^\top X)^{-1}X^\top
$$

- ハット行列 H による表現

$$
\begin{aligned}
\hat{\boldsymbol{y}} &= H\boldsymbol{y} \\
\hat{\boldsymbol{\epsilon}} &= (I - H)\boldsymbol{\epsilon}
\end{aligned}
$$

  - あてはめ値や残差は H を用いて簡潔に表現される

ハット行列の性質

- 観測データ (デザイン行列) のみで計算される
- 観測データと説明変数の関係を表す
- 対角成分 (**テコ比**; leverage) は観測データが自身の予測に及ぼす影響の度合を表す

$$
\hat{y}_j = (H)_{jj}y_j + (\text{それ以外のデータの寄与})
$$

  - $(A)_{ij}$ は行列 A の (i, j) 成分
  - テコ比が小さい : 他のデータでも予測が可能
  - テコ比が大きい : 他のデータでは予測が困難

演習

問題

- ハット行列 H について以下を示しなさい
  - H は対称行列であること
  - H は冪等であること

$$H^2 = H, \quad (I - H)^2 = I - H$$

  - 以下の等式が成り立つこと

$$HX = X, \quad X^\top H = X^\top$$

ヒント

- いずれも H の定義にもとづいて計算すればよい

$$\begin{aligned}
H^\top &= (X(X^\top X)^{-1} X^\top)^\top \\
H^2 &= (X(X^\top X)^{-1} X^\top)(X(X^\top X)^{-1} X^\top) \\
(I - H)^2 &= I - 2H + H^2 \\
HX &= (X(X^\top X)^{-1} X^\top)X \\
X^\top H &= (HX)^\top
\end{aligned}$$

推定量の統計的性質

最小二乗推定量の性質

- 推定量と誤差の関係

$$\begin{aligned}
\hat{\boldsymbol{\beta}} &= (X^\top X)^{-1} X^\top \boldsymbol{y} \\
&= (X^\top X)^{-1} X^\top (X\boldsymbol{\beta} + \boldsymbol{\epsilon}) \\
&= (X^\top X)^{-1} X^\top X\boldsymbol{\beta} + (X^\top X)^{-1} X^\top \boldsymbol{\epsilon} \\
&= \boldsymbol{\beta} + (X^\top X)^{-1} X^\top \boldsymbol{\epsilon}
\end{aligned}$$

- 正規分布の重要な性質 (**再生性**)

  正規分布に従う独立な確率変数の和は正規分布に従う

推定量の分布

- **誤差の仮定** : 独立，平均 0 分散 σ^2 の **正規分布**
- 推定量は以下の多変量正規分布に従う

$$\begin{aligned}
\mathbb{E}[\hat{\boldsymbol{\beta}}] &= \boldsymbol{\beta} \\
\mathrm{Cov}(\hat{\boldsymbol{\beta}}) &= \sigma^2 (X^\top X)^{-1}
\end{aligned}$$

$$\hat{\boldsymbol{\beta}} \sim \mathcal{N}(\boldsymbol{\beta}, \sigma^2 (X^\top X)^{-1})$$

演習

問題

- 誤差が独立で，平均 0 分散 σ^2 の正規分布に従うとき，最小二乗推定量 $\hat{\boldsymbol{\beta}}$ について以下を示しなさい
  - 平均は $\boldsymbol{\beta}$ (真の母数) となること
  - 共分散行列は $\sigma^2(X^\top X)^{-1}$ となること

解答例

- 定義にもとづいて計算する

$$\begin{aligned}\mathbb{E}[\hat{\boldsymbol{\beta}}] &= \mathbb{E}[\boldsymbol{\beta} + (X^\top X)^{-1}X^\top\boldsymbol{\epsilon}] \\ &= \boldsymbol{\beta} + (X^\top X)^{-1}X^\top\mathbb{E}[\boldsymbol{\epsilon}] \\ &= \boldsymbol{\beta}\end{aligned}$$

  - 定義にもとづいて計算する

$$\begin{aligned}\mathrm{Cov}(\hat{\boldsymbol{\beta}}) &= \mathbb{E}[(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})^\top] \\ &= \mathbb{E}[(X^\top X)^{-1}X^\top\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top X(X^\top X)^{-1}] \\ &= (X^\top X)^{-1}X^\top\mathbb{E}[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top]X(X^\top X)^{-1} \\ &= (X^\top X)^{-1}X^\top(\sigma^2 I)X(X^\top X)^{-1} \\ &= \sigma^2(X^\top X)^{-1}\end{aligned}$$

誤差の評価

寄与率 (再掲)

- **決定係数 (R-squared)**
  - 回帰式で説明できるばらつきの比率

$$R^2 = 1 - \frac{\sum_{i=1}^n \hat{\epsilon}_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

- **自由度調整済み決定係数 (adjusted R-squared)**
  - 決定係数を不偏分散で補正

$$\bar{R}^2 = 1 - \frac{\frac{1}{n-p-1}\sum_{i=1}^n \hat{\epsilon}_i^2}{\frac{1}{n-1}\sum_{i=1}^n (y_i - \bar{y})^2}$$

各係数の推定量の分布

- 推定された回帰係数の精度を評価
  - 誤差 ϵ の分布は平均 0 分散 σ^2 の正規分布
  - $\hat{\boldsymbol{\beta}}$ の分布 : p+1 変量正規分布

$$\hat{\boldsymbol{\beta}} \sim \mathcal{N}(\boldsymbol{\beta}, \sigma^2(X^\top X)^{-1})$$

- $\hat{\beta}_j$ の分布 : 1 変量正規分布

$$\hat{\beta}_j \sim \mathcal{N}(\beta_j, \sigma^2((X^\top X)^{-1})_{jj}) = \mathcal{N}(\beta_j, \sigma^2 \zeta_j^2)$$

  - $(A)_{jj}$ は行列 A の (j, j) (対角) 成分

標準誤差

- **標準誤差** (standard error)
  - $\hat{\beta}_j$ の標準偏差の推定量

$$\text{s.e.}(\hat{\beta}_j) = \hat{\sigma}\zeta_j = \sqrt{\frac{1}{n-p-1}\sum_{i=1}^{n}\hat{\epsilon}_i^2} \cdot \sqrt{((X^\top X)^{-1})_{jj}}$$

  - 未知母数 σ^2 は不偏分散 $\hat{\sigma}^2$ で推定
  - $\hat{\beta}_j$ の精度の評価指標

演習

問題

- 以下の問に答えなさい
  - 不偏分散 $\hat{\sigma}^2$ が母数 σ^2 の不偏な推定量となることを示せ
    以下が成り立つことを示せばよい

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

  - 回帰のばらつき S_r をハット行列 H と目的変数 $\boldsymbol{y}$ で表せ

解答例

- ハット行列 H を用いた表現を利用する

$$\begin{aligned}
\hat{\boldsymbol{\epsilon}} &= (I_n - H)\boldsymbol{\epsilon} \\
\mathbb{E}\left[\sum_{i=1}^{n}\hat{\epsilon}_i^2\right] &= \mathbb{E}[\hat{\boldsymbol{\epsilon}}^\top\hat{\boldsymbol{\epsilon}}] \\
&= \mathbb{E}[\text{tr}(\hat{\boldsymbol{\epsilon}}\hat{\boldsymbol{\epsilon}}^\top)] \\
&= \mathbb{E}[\text{tr}(I_n - H)\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top(I_n - H)] \\
&= \text{tr}(I_n - H)\mathbb{E}[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top](I_n - H) \\
&= \text{tr}(I_n - H)(\sigma^2 I_n)(I_n - H) \\
&= \sigma^2\text{tr}(I_n - H)
\end{aligned}$$

  - I_n は $n \times n$ 単位行列

- さらに以下が成立する

$$
\begin{aligned}
\mathrm{tr}H &= \mathrm{tr}X(X^\top X)^{-1}X^\top \\
&= \mathrm{tr}(X^\top X)^{-1}X^\top X \\
&= \mathrm{tr}I_{p+1} \\
&= p+1
\end{aligned}
$$

  - 行列のサイズに注意

- H の性質より以下が成り立つ

$$
\begin{aligned}
H\mathbf{1} &= \mathbf{1} \\
\hat{\boldsymbol{y}} &= H\boldsymbol{y}
\end{aligned}
$$

- 行列 M を以下で定める

$$
M = \frac{1}{n}\mathbf{1}\mathbf{1}^\top
$$

  このとき以下が成り立つ

$$
\begin{aligned}
&M^2 = M \\
&HM = MH = M
\end{aligned}
$$

- 行列 M を用いると目的変数の平均は以下で表される

$$
\bar{\boldsymbol{y}} = M\boldsymbol{y}
$$

- S_r の定義から

$$
\begin{aligned}
S_r &= (\hat{\boldsymbol{y}} - \bar{\boldsymbol{y}})^\top(\hat{\boldsymbol{y}} - \bar{\boldsymbol{y}}) \\
&= ((H-M)\boldsymbol{y})^\top((H-M)\boldsymbol{y}) \\
&= \boldsymbol{y}^\top(H-M)^\top(H-M)\boldsymbol{y} \\
&= \boldsymbol{y}^\top(H-M)\boldsymbol{y} \\
&= \mathrm{tr}(H-M)\boldsymbol{y}\boldsymbol{y}^\top
\end{aligned}
$$

- 後に $\beta_1 = \cdots = \beta_p = 0$ ならば

$$
\begin{aligned}
\mathbb{E}[\boldsymbol{y}\boldsymbol{y}^\top] &= \beta_0^2\mathbf{1}\mathbf{1}^\top + \mathbb{E}[\boldsymbol{\epsilon}\boldsymbol{\epsilon}^\top] \\
&= n\beta_0^2 M + \sigma^2 I_n
\end{aligned}
$$

  となることを利用して，モデルの評価のための統計量を導く

係数の評価

t 統計量

- **回帰係数の分布** に関する定理

  t 統計量 (t-statistic)

  $$t = \frac{\hat{\beta}_j - \beta_j}{\mathrm{s.e.}(\hat{\beta}_j)} = \frac{\hat{\beta}_j - \beta_j}{\hat{\sigma}\zeta_j}$$

  は自由度 $n-p-1$ の t 分布に従う

  - 証明には以下の性質を用いる
    - $\hat{\sigma}^2$ と $\hat{\boldsymbol{\beta}}$ は独立となる
    - $(\hat{\beta}_j - \beta_j)/(\sigma\zeta_j)$ は標準正規分布に従う
    - $(n-p-1)\hat{\sigma}^2/\sigma^2 = S(\hat{\boldsymbol{\beta}})/\sigma^2$ は自由度 $n-p-1$ の χ^2 分布に従う

t 統計量による検定

- 回帰係数 β_j が回帰式に寄与するか否かを検定
  - 帰無仮説 $H_0 : \beta_j = 0$ (t 統計量が計算できる)
  - 対立仮説 $H_1 : \beta_j \neq 0$
- p 値 : 確率変数の絶対値が $|t|$ を超える確率
  - $f(x)$ は自由度 $n-p-1$ の t 分布の確率密度関数

  $$(p\ 値) = 2\int_{|t|}^{\infty} f(x)dx \quad (両側検定)$$

  帰無仮説 H_0 が正しければ p 値は小さくならない

モデルの評価

平方和の分解 (再掲)

- いろいろなばらつき
  - $S_y = (\boldsymbol{y} - \bar{\boldsymbol{y}})^\top(\boldsymbol{y} - \bar{\boldsymbol{y}})$: 目的変数のばらつき
  - $S\ \ = (\boldsymbol{y} - \hat{\boldsymbol{y}})^\top(\boldsymbol{y} - \hat{\boldsymbol{y}})$: 残差のばらつき ($\hat{\boldsymbol{\epsilon}}^\top\hat{\boldsymbol{\epsilon}}$)
  - $S_r = (\hat{\boldsymbol{y}} - \bar{\boldsymbol{y}})^\top(\hat{\boldsymbol{y}} - \bar{\boldsymbol{y}})$: あてはめ値 (回帰) のばらつき
- 3 つのばらつき (平方和) の関係

$$(\boldsymbol{y} - \bar{\boldsymbol{y}})^\top(\boldsymbol{y} - \bar{\boldsymbol{y}}) = (\boldsymbol{y} - \hat{\boldsymbol{y}})^\top(\boldsymbol{y} - \hat{\boldsymbol{y}}) + (\hat{\boldsymbol{y}} - \bar{\boldsymbol{y}})^\top(\hat{\boldsymbol{y}} - \bar{\boldsymbol{y}})$$

$$S_y = S + S_r$$

F 統計量

- **ばらつきの比** に関する定理

  $\beta_1 = \cdots = \beta_p = 0$ ならば F **統計量** (F-statistic)

  $$F = \frac{\frac{1}{p}S_r}{\frac{1}{n-p-1}S} = \frac{n-p-1}{p}\frac{R^2}{1-R^2}$$

  は自由度 $p, n-p-1$ の F 分布に従う

  - 証明には以下の性質を用いる
    - S_r と S は独立となる
    - S_r/σ^2 は自由度 p の χ^2 分布に従う
    - S/σ^2 は自由度 $n-p-1$ の χ^2 分布に従う

F 統計量を用いた検定

- 説明変数のうち 1 つでも役に立つか否かを検定
  - 帰無仮説 $H_0 : \beta_1 = \cdots = \beta_p = 0$ (S_r が χ^2 分布になる)
  - 対立仮説 $H_1 : \exists j\ \beta_j \neq 0$
- p 値 : 確率変数の値が F を超える確率
  - $f(x)$ は自由度 $p, n-p-1$ の F 分布の確率密度関数

  $$(p \text{ 値}) = \int_F^\infty f(x)dx \quad (\text{片側検定})$$

  帰無仮説 H_0 が正しければ p 値は小さくならない

解析の事例

気温に影響を与える要因の分析 (再掲)

- データの概要
- 気温を説明する 5 種類の線形回帰モデルを検討
  - モデル 1 : 気温 = F(気圧)
  - モデル 2 : 気温 = F(日射)
  - モデル 3 : 気温 = F(気圧, 日射)
  - モデル 4 : 気温 = F(気圧, 日射, 湿度)
  - モデル 5 : 気温 = F(気圧, 日射, 雲量)

分析の視覚化 (再掲)

- 観測値とあてはめ値の比較

モデルの比較

- t 統計量・F 統計量
- 診断プロット
- 診断プロット
- 診断プロット

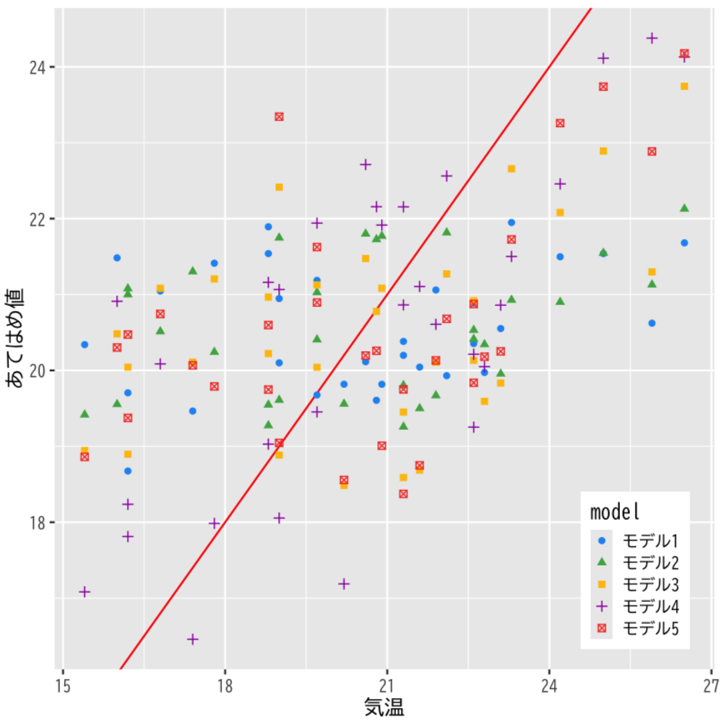

Figure 6: モデルの比較

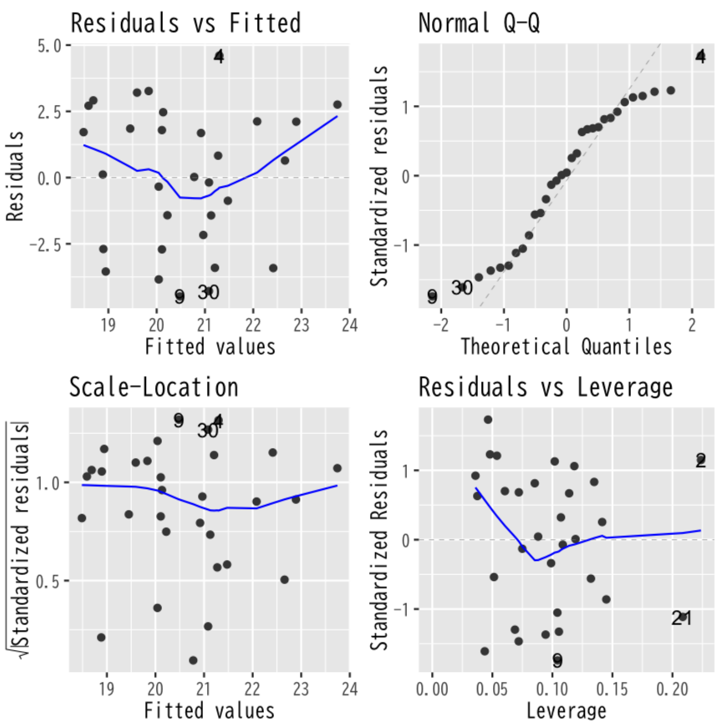

Figure 7: モデル 3

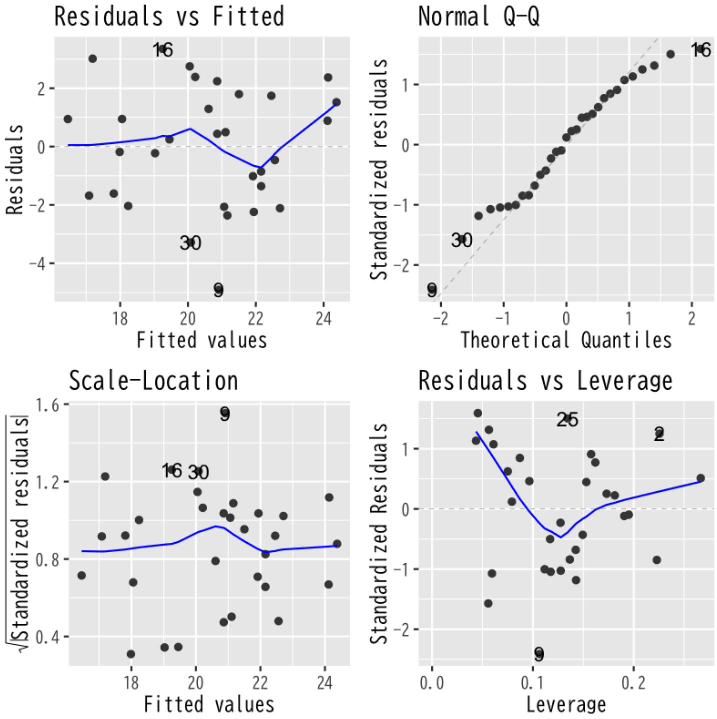

Figure 8: モデル 4

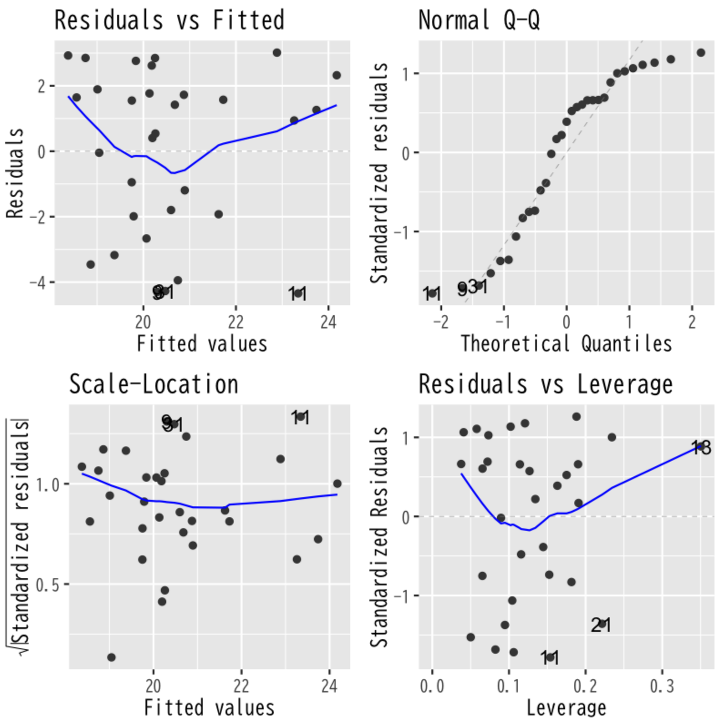

Figure 9: モデル 5

日付	気温	降雨	日射	降雪	風向	風速	気圧	湿度	雲量
2024-10-01	23.3	0.5	11.45	0	NNW	2.6	1006.0	81	5.8
2024-10-02	26.5	0.0	18.32	0	S	2.9	1007.9	77	6.0
2024-10-03	23.1	11.0	5.88	0	E	2.7	1015.9	87	10.0
2024-10-04	25.9	2.0	12.60	0	S	3.5	1015.4	87	10.0
2024-10-05	21.3	9.5	1.88	0	NNE	2.5	1018.4	94	10.0
2024-10-06	21.3	0.0	5.01	0	NNW	1.7	1017.1	93	10.0
2024-10-07	25.0	0.0	14.99	0	S	2.9	1008.9	83	8.0
2024-10-08	18.8	33.5	1.98	0	NE	3.0	1008.9	97	10.0
2024-10-09	16.0	53.5	3.58	0	NNW	2.9	1009.3	93	10.0
2024-10-10	17.8	0.0	7.52	0	NNW	2.6	1009.8	75	6.0
2024-10-11	19.0	0.0	16.14	0	SSE	1.9	1013.1	69	7.5
2024-10-12	20.6	0.0	16.44	0	N	1.9	1019.0	73	2.5
2024-10-13	20.9	0.0	16.27	0	NNW	2.2	1021.1	70	0.8
2024-10-14	20.8	0.0	16.02	0	NNW	2.3	1022.6	71	4.0
2024-10-15	22.1	0.0	16.53	0	SSW	2.2	1020.3	72	3.8
2024-10-16	22.6	0.0	8.50	0	NNE	1.5	1017.3	76	7.5
2024-10-17	22.8	0.0	8.10	0	ENE	2.3	1020.0	79	9.3
2024-10-18	21.6	2.0	3.27	0	N	1.8	1019.5	92	10.0
2024-10-19	24.2	1.5	11.29	0	S	2.7	1009.2	84	10.0
2024-10-20	17.4	0.0	13.59	0	ENE	3.6	1023.6	55	5.8
2024-10-21	16.2	0.0	12.31	0	NW	2.7	1029.2	61	7.0
2024-10-22	19.7	0.0	12.02	0	NNW	1.9	1022.1	69	8.3
2024-10-23	21.9	6.5	4.24	0	NW	2.3	1012.3	90	10.0
2024-10-24	22.6	0.0	9.18	0	NE	2.2	1013.5	79	8.0
2024-10-25	20.2	0.5	3.61	0	NNE	2.4	1021.1	77	9.8
2024-10-26	19.0	0.0	3.90	0	NNW	1.7	1019.1	80	10.0
2024-10-27	19.7	4.0	8.46	0	NW	1.5	1011.4	87	9.5
2024-10-28	18.8	8.0	3.54	0	NE	1.9	1006.4	87	9.8
2024-10-29	15.4	24.5	2.79	0	NE	2.9	1017.4	79	10.0
2024-10-30	16.8	17.5	9.07	0	NW	3.2	1012.4	79	7.5
2024-10-31	16.2	0.0	11.86	0	NNE	2.0	1021.9	65	7.5

次回の予定

- 第 1 回 : 回帰モデルの考え方と推定
- 第 2 回 : モデルの評価
- **第 3 回 : モデルによる予測と発展的なモデル**

	モデル 1			モデル 2			モデル 3			モデル 4			モ	
変数	**係数**	**t 統計量**	**p 値**	**係数**	**t 統計量**	**p 値**	**係数**	**t 統計量**	**p 値**	**係数**	**t 統計量**	**p 値**	**係数**	**t**
気圧	-0.14	-1.57	0.13				-0.17	-1.97	0.059	0.06	0.715	0.5	-0.14	
日射				0.17	1.73	0.10	0.21	2.10	0.045	0.53	4.85	<0.001	0.38	
湿度										0.28	4.21	<0.001		
雲量													0.49	
Adjusted R^2	0.047			0.062			0.147			0.466			0.191	
F 統計量	2.47			2.98			3.58			9.72			3.36	
p 値	0.13			0.10			0.041			<0.001			0.033	

Abbreviation: CI = Confidence Interval