

データ解析 2003 年度試験問題

静岡大学工学部システム工学科

安藤 和敏

2003 年 12 月 10 日

注意事項

- 教科書, プリント, ノート, 他人のノートのコピー等の 持ち込みは許可する. ただし, 持ち込んだものの貸し借りは禁止する.
- ノートパソコン, (関数) 電卓の使用を許可する. ただし, これらの機器の貸し借り, 及び, ネットワーク等を利用した通信を禁ずる.
- 試験の時間は 10:20-11:50 である.
- 数値計算の結果は, 小数点以下 (第 4 位を四捨五入して), 第 3 位まで記すこと.
- 証明問題の解答はなるべく詳しく記述せよ.
- 11:20 以降, 退室を許可する. 他の人の迷惑にならないように退室すること.
- 問題用紙は持ち帰ってよい.

データ解析

問題 1. (配点 20)

表 1 に示したデータは, 正規分布 $N(\mu, \sigma^2)$ にしたがう母集団から採取された $n = 15$ 個のデータである. ここで, μ と σ^2 は未知のパラメータである.

表 1:

No	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x	30	35	28	35	30	34	29	28	29	29	31	28	32	36	32

設問 (1) x の平均 \bar{x} と分散 V_x を数値で求めよ.

設問 (2) $\frac{\bar{x}-\mu}{\sqrt{V_x/n}}$ はどのような分布にしたがう確率変数か? ここで, $n = 15$ である.

設問 (3) 帰無仮説 $H_0: \mu = 33$, 対立仮説 $H_1: \mu \neq 33$ として, 有意水準 5% で検定を行なえ.

設問 (4) μ の 95% 信頼区間を (数値で) 求めよ.

問題 2. (配点 30)

表 3 に示すような 2 変数 x, y についてのデータが得られているとする. このデータに対して, 単回帰モ

表 2:

No	x	y
1	x_1	y_1
2	x_2	y_2
\vdots	\vdots	\vdots
i	x_i	y_i
\vdots	\vdots	\vdots
n	x_n	y_n

デル

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (i = 1, \dots, n)$$

を仮定する. ここで, ε_i ($i = 1, \dots, n$) は独立に同一の分布 $N(0, \sigma^2)$ に従う.

以下の設問に答えよ. その際に, 以下の事実を必要ならば用いよ.

- 各 $i = 1, \dots, n$ に対して, y_i は期待値 $\beta_0 + \beta_1 x_i$, 分散 σ^2 の正規分布に従う確率変数である.
- $i \neq j$ ならば, y_i と y_j は独立である.

設問 (1) β_0 と β_1 の推定量 $\hat{\beta}_0$ と $\hat{\beta}_1$ を求める方法として, 最小 2 乗法がある. 最小 2 乗法がどのような方法なのかを, 70 字程度で述べよ (全角 1 文字を 1 字と数える). もちろん数式を含んでもよい.

設問 (2) 最小 2 乗法を用いて求められる β_0 と β_1 の推定量 $\hat{\beta}_0$ と $\hat{\beta}_1$ を $\bar{x}, \bar{y}, S_{xx}, S_{xy}$ を用いて表せ.

設問 (3) $\hat{\beta}_1$ を y_1, \dots, y_n の一次結合 (線型結合) の形で表現しなさい.

設問 (4) $E(\hat{\beta}_1) = \beta_1$ を証明せよ.

設問 (5) $\hat{\beta}_1$ はどのような分布にしたがう確率変数か?

設問 (6) $\hat{\beta}_1$ の 95% 信頼区間を式で表せ.

問題 3. (配点 30)

表 3 に示すような 2 変数 x, y についてのデータが得られているとする.

表 3:

No	x	y
1	52	68
2	56	69
3	48	63
4	62	86
5	52	64
6	66	89
7	30	36
8	60	80
9	55	72
10	64	89
11	69	94
12	47	60

このデータに対して, 単回帰モデル

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (i = 1, \dots, n)$$

を仮定する. ここで, ε_i ($i = 1, \dots, n$) は独立に同一の分布 $N(0, \sigma^2)$ に従う.

以下の設問に答えよ.

設問 (1) β_0, β_1 の最小 2 乗推定量 $\hat{\beta}_0, \hat{\beta}_1$ を数値で求めよ.

設問 (2) y の全変動 S_{yy} , 残差平方和 S_e , 回帰による変動 S_R は, それぞれいくつ? また, これら 3 つの量はどのような関係にあるか?

設問 (3) 寄与率 R^2 を S_{yy}, S_R を用いた式で表現し, その数値を与えよ. また, その値から, なにが言えるか?

設問 (4) β_1 の 95% 信頼区間を数値で求めよ.

設問 (5) $x_0 = 60$ に対する母回帰 $\beta_0 + \beta_1 x_0$ の 95% 信頼区間を数値で求めよ.

設問 (6) $x_0 = 60$ に対する予測値 $\beta_0 + \beta_1 x_0 + \varepsilon$ の 95% 信頼区間 (予測区間) を数値で求めよ.

問題 4. (問題 2 の続き, 配点 20)

以下の設問に答えよ. 必要ならば, 問題 2 の結果を用いても良い.

設問 (1) 次式を証明せよ.

$$E(\hat{\beta}_0) = \beta_0$$

設問 (2) 次式を証明せよ.

$$\hat{\beta}_0 = \sum_{i=1}^n \left(\frac{1}{n} - \frac{x_i - \bar{x}}{S_{xx}} \bar{x} \right) y_i$$

設問 (3) 上式を用いて, 次式を証明せよ.

$$V(\hat{\beta}_0) = \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}} \right) \sigma^2$$

設問 (4) $\hat{\beta}_0$ はどのような分布にしたがう確率変数か?

学籍 番号		氏 名	
----------	--	--------	--

問題 1(1) の解答欄

問題 1(2) の解答欄

問題 1(3) の解答欄

問題 1(4) の解答欄

問題 2(1) の解答欄

.....

.....

.....

.....

.....

学籍 番号		氏 名	
----------	--	--------	--

問題 2(2) の解答欄

問題 2(3) の解答欄

問題 2(4) の解答欄

問題 2(5) の解答欄

問題 2(6) の解答欄

学籍 番号		氏 名	
----------	--	--------	--

問題 3(1) の解答欄

問題 3(2) の解答欄

問題 3(3) の解答欄

問題 3(4) の解答欄

問題 3(5) の解答欄

問題 3(6) の解答欄

学籍 番号		氏 名	
----------	--	--------	--

問題 4(1) の解答欄

問題 4(2) の解答欄

学籍 番号		氏 名	
----------	--	--------	--

問題 4(3) の解答欄

問題 4(4) の解答欄