統計検定

Japan Statistical Society Certificate

準1級

2019 年 6 月 16 日

【注意事項】

1 試験開始の合図があるまで，この問題冊子および部分記述問題・論述問題解答用紙の中を見てはいけません。

2 この問題冊子は，40 ページあります。**3〜23** ページは選択問題及び部分記述問題，**25〜32** ページは論述問題です。

3 論述問題は，3 問から 1 問のみを選択して，解答しなさい。

4 試験時間は 120 分です。

5 試験中に問題冊子の印刷不鮮明，ページの落丁・乱丁およびマークシートの汚れ等に気付いた場合は，手を挙げて監督者に知らせなさい。

6 **解答用紙・マークシートの A 面には次の項目があるので，それぞれの指示に従い記入あるいは確認しなさい。**項目の内容に誤りがある場合は，手を挙げて監督者に知らせなさい。

① **マークシートの氏名**
氏名を記入しなさい。

② **マークシートの検定種別と受験番号**
受験する検定種別と受験番号を確認しなさい。

③ **マークシートの Web 合格発表**
Web 合格発表について，希望の有無をマークしなさい。

④ **部分記述問題・論述問題解答用紙**の表紙上部に**受験番号を記入**しなさい。

⑤ 論述問題の解答面には，最終的な解答だけでなくそれに至る過程も記しなさい。最終的な解答が正しくないときにも点が与えられることがあります。

7 選択問題の解答は，マークシートの B 面の解答にマークしなさい。
（この冊子裏面の記入例参照）

8 部分記述問題の解答は，**部分記述問題解答面**に解答しなさい。 **記述 2** のように表示のある問は部分記述問題です。

9 **選択問題**（問 3 〜 問 10）の**解答番号**は **24** まで，**部分記述問題**（問 1，2，11，12）の**解答番号**は **10** まであります。

10 33 ページ以降に付表を掲載しています。必要に応じて利用しなさい。

11 問題冊子の余白等は適宜利用してよいが，どのページも切り離してはいけません。

12 試験終了後，問題冊子は持ち帰りなさい。

（冊子裏面につづく）

選択問題及び部分記述問題

問 1　あるサッカーの試合において，チーム T1 があげた得点 X およびチーム T2 があげた得点 Y がそれぞれ独立に平均 3 および 2 のポアソン分布に従うと仮定する。次の空欄に当てはまる数値または用語を答えよ。

〔1〕2 チームの合計得点 $X+Y$ の従う分布は，平均が $\boxed{\textbf{記述 1}}$，分散が $\boxed{\textbf{記述 2}}$ のポアソン分布である。

〔2〕2 チームの合計得点 $X+Y$ が 4 であるという条件のもとで，チーム T1 の得点 X は平均が $\boxed{\textbf{記述 3}}$ の $\boxed{\textbf{記述 4}}$ 分布に従う。ただし，必要であれば平均が λ のポアソン分布の確率関数は

$$P(X=x)=e^{-\lambda}\frac{\lambda^x}{x!} \qquad (x=0,1,\ldots)$$

であることを用いよ。

問 2　あるお菓子を買うと，3 種類のアニメキャラクターのカードのうちの 1 つが等確率でおまけとして付いてくる。

〔1〕無作為復元抽出を仮定できるとき，3 種類すべてのカードを揃えるまでに必要な購入回数の期待値を求めよ。ただし，必要であればパラメータ p $(0 < p < 1)$ をもつ幾何分布の平均は p^{-1} となること，つまり

$$\sum_{k=1}^{\infty} k(1-p)^{k-1}p = \frac{1}{p}$$

となることを用いよ。$\boxed{\textbf{記述 5}}$

〔2〕3 種類のカードをすべて集めた後，お菓子を買うのをやめていたが，新しい種類のカード 1 枚が追加されたため再び購入を始めた。この場合に，はじめの 3 種類と追加の 1 種類の，4 種類すべてを揃えるのに必要な購入回数の期待値を x とする。一方，はじめから 4 種類が発売されていた場合に，4 種類すべてを揃えるまでに必要な購入回数の期待値を y とする。このとき，購入回数の期待値の差 $x - y$ の値を求めよ。ただし，いずれの購入時期においても等確率の無作為復元抽出を仮定してよい。$\boxed{\textbf{記述 6}}$

注：記述 7〜10 は問 11，問 12 にあります。

問 3　医薬品の開発段階において観察された有害事象は，審査時に臨床的重要性について検討がなされ，製造販売後の調査において適切に監視される。

〔1〕開発段階の臨床試験において，ある有害事象の発現割合（母比率）を p とする。また，症例数は十分大きく，発現割合の推定量 $\hat{p}$ は近似的に正規分布に従うと仮定する。

(1) 症例数 475 で，帰無仮説 $H_0 : p = 0.05$，対立仮説 $H_1 : p > 0.05$ の片側検定をするとき，帰無仮説の下で $\hat{p}$ が 0.0733 以上になる確率はいくらか。次の ① 〜 ⑤ のうちから最も適切なものを一つ選べ。 **[1]**

① 0.01　② 0.025　③ 0.05　④ 0.1　⑤ 0.2

(2) 帰無仮説 $H_0 : p = 0.05$，対立仮説 $H_1 : p = 0.1$ に対し，有意水準 2.5 %の片側検定を行うとき，検出力を 90 %とする製造販売後調査の必要症例数は何例になるか。次の ① 〜 ⑤ のうちから最も適切なものを一つ選べ。 **[2]**

① 114　② 164　③ 214　④ 264　⑤ 314

〔2〕発現が懸念されるある有害事象が，開発段階の臨床試験では観察されなかった。該当の有害事象は，発現割合が 0.001 未満であれば，安全性の観点からは許容可能であると考えられている。

(1) 発現割合が 0.05 の事象について，独立に 8 症例を調べた。このとき，少なくとも 1 例の有害事象が観測される確率はいくらか。次の ① 〜 ⑤ のうちから最も適切なものを一つ選べ。 **[3]**

① 0.05　② 0.24　③ 0.34　④ 0.40　⑤ 0.66

(2) 発現割合が 0.001 の独立な事象について，95 %の確率で少なくとも 1 例の有害事象が観察されるような症例数を n とする。この症例数 n で独立に観察を行ったときに，1 例も有害事象が観察されなければ，その事象の発現割合は 0.001 未満であると判断する。この場合の製造販売後調査の必要症例数は何例になるか。次の ① 〜 ⑤ のうちから最も適切なものを一つ選べ。ただし，必要に応じて付表 5 を用い，また $\varepsilon(>0)$ が十分小さいときに $\log(1-\varepsilon) \simeq -\varepsilon$ であることを用いてよい。ここで，log は自然対数である。 **[4]**

① 1000　② 1500　③ 2000　④ 2500　⑤ 3000

問 4　ある商品について，CM の影響の有無と購入の有無について調査した結果，次の分割表が得られた。

	購入あり	購入なし	計
CM の影響あり	93	42	135
CM の影響なし	97	68	165
計	190	110	300

〔1〕CM の影響の有無と購入の有無に関連がないと仮定して確率を推定する。このとき，CM の影響ありかつ購入ありの頻度の期待値として，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **5**

① 34.3　② 48.7　③ 58.3　④ 85.5　⑤ 106.7

〔2〕CM の影響の有無と購入の有無の関連性に関するピアソンの χ^2 統計量はいくつか。次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **6**

① -4.32　② 0.33　③ 3.26　④ 10.49　⑤ 22.93

〔3〕「CM の影響の有無と購入の有無の関連性がない」という帰無仮説に対する片側検定の結果について，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **7**

① 有意水準 10 %では帰無仮説は棄却されない。
② 有意水準 10 %では帰無仮説は棄却されるが，有意水準 5 %では棄却されない。
③ 有意水準 5 %では帰無仮説は棄却されるが，有意水準 2.5 %では棄却されない。
④ 有意水準 2.5 %では帰無仮説は棄却されるが，有意水準 1 %では棄却されない。
⑤ 有意水準 1 %では帰無仮説は棄却される。

問 5　ある冠動脈疾患の治療施設では，心筋梗塞と喫煙の関係を調べるため，以下のような調査を行った。まず，急性心筋梗塞を発症してこの施設に入院した患者（86 名）について，喫煙歴の有無を調査した。次に，86 名のそれぞれに対して，同じ期間に別の急性疾患を発症してこの施設に入院した患者の中から，年齢，性別，身長，体重が比較的近い者を 3 名ずつ選んだ。選ばれた 258 名をコントロール群とよび，コントロール群に対しても喫煙歴の有無を調査した。調査結果をまとめたのが，次の表である。

	心筋梗塞患者	コントロール群	合計
喫煙歴あり	65	66	131
喫煙歴なし	21	192	213
合計	86	258	344

この調査から読み取れる心筋梗塞に罹る確率（罹患率）の解釈について，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 $\boxed{\textbf{8}}$

① 心筋梗塞患者に関する喫煙歴ありのオッズ（$65/21 = 3.10$）はコントロール群に関する喫煙歴ありのオッズ（$66/192 = 0.344$）の約 9 倍である。心筋梗塞の罹患率は小さい値であると知られているので，喫煙歴がある場合とない場合のそれぞれに対する心筋梗塞の罹患率の比（相対リスク）は，およそ 9 であると推定できる。

② 心筋梗塞患者に関する喫煙歴ありのオッズ（$65/21 = 3.10$）はコントロール群に関する喫煙歴ありのオッズ（$66/192 = 0.344$）の約 9 倍である。コントロール群として 3 倍の人数を選んでいるので，喫煙歴がある場合とない場合のそれぞれに対する心筋梗塞の罹患率の比（相対リスク）は，およそ 3 であると推定できる。

③ 喫煙歴のある患者に関する心筋梗塞患者の割合は $65/131 = 0.496$ であり，喫煙歴のない患者に関する心筋梗塞患者の割合は $21/213 = 0.0986$ である。これらはそれぞれ，喫煙歴がある場合とない場合のそれぞれに対して，心筋梗塞の罹患率の妥当な推定値である。

④ 喫煙歴のある患者に関する心筋梗塞患者の割合は $65/131 = 0.496$ であり，喫煙歴のない患者に関する心筋梗塞患者の割合は $21/213 = 0.0986$ であり，この差はおよそ 0.40 である。このことから，喫煙歴があると，喫煙歴がない場合にくらべて，心筋梗塞の罹患率が 40 %増えると推定できる。

⑤ この調査では，調査対象の 4 分の 1 は心筋梗塞患者となる。これは，実際の心筋梗塞の罹患率とはかけ離れた値であるから，この調査から読み取れるものはない。

問 6　ある時点で生成したコンクリートの圧縮強度 (y) を調べるため，セメント量 (x_1)，高炉スラグ量 (x_2)，飛散灰量 (x_3)，水分量 (x_4)，高性能 AE 減水剤量 (x_5)，粗骨材量 (x_6)，細骨材量 (x_7)，およびこれらを観測した時点 (x_8) を記録した標本サイズ 50 のデータが観測されている。説明変数 $x_1, \ldots, x_8$ について相関係数行列に基づく主成分分析を行ったところ，第 1 主成分 (PC1) から第 8 主成分 (PC8) までの固有値，寄与率，長さ 1 の固有ベクトルは以下の通りであった。

主成分		PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
固有値		2.334	1.540	1.372	1.012	0.936	0.659	0.113	0.035
寄与率		0.292	0.193	0.172	0.127	0.117	0.082	0.014	0.004
固有ベクトル	x_1	0.288	−0.349	0.454	−0.553	0.034	−0.158	0.057	−0.504
	x_2	−0.416	0.319	0.374	0.212	0.191	0.485	0.224	−0.468
	x_3	0.250	0.578	−0.353	0.023	0.095	−0.469	0.277	−0.415
	x_4	−0.593	−0.014	−0.018	−0.006	−0.041	−0.448	−0.610	−0.272
	x_5	0.461	0.474	0.206	−0.049	0.014	0.266	−0.668	0.016
	x_6	−0.021	−0.202	−0.693	−0.299	0.192	0.471	−0.167	−0.321
	x_7	0.301	−0.344	−0.058	0.662	−0.401	0.041	−0.105	−0.420
	x_8	−0.161	0.242	−0.060	−0.345	−0.868	0.160	0.111	−0.032

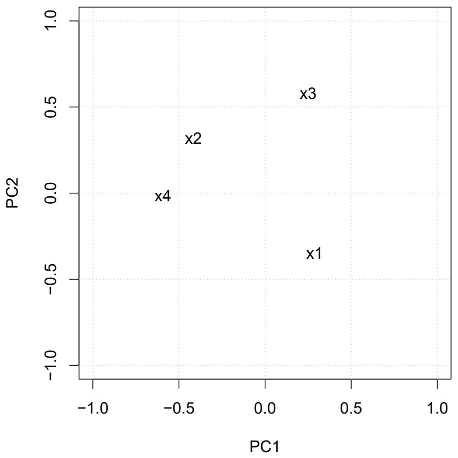
資料: UCI Machine Learning Repository,
Concrete Compressive Strength Data Set

〔1〕第何主成分まででではじめて累積寄与率が 80 %以上になるか。次の ① ～ ⑤ のうちから最も適切なものを選べ。 **9**

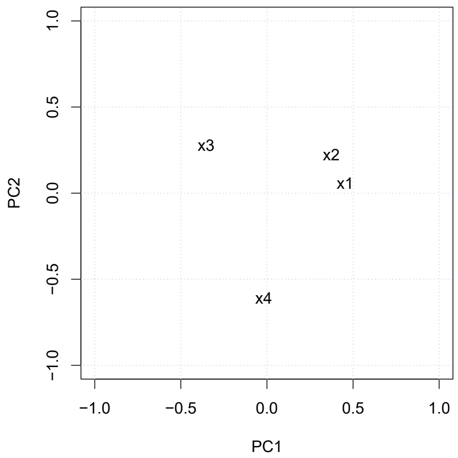
① 第 3 主成分　② 第 4 主成分　③ 第 5 主成分
④ 第 6 主成分　⑤ 第 7 主成分

〔2〕横軸を第 1 主成分 (PC1)，縦軸を第 2 主成分 (PC2) とした場合の $x_1, \ldots, x_4$ の固有ベクトルのプロットはどれか。次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 $\boxed{\textbf{10}}$

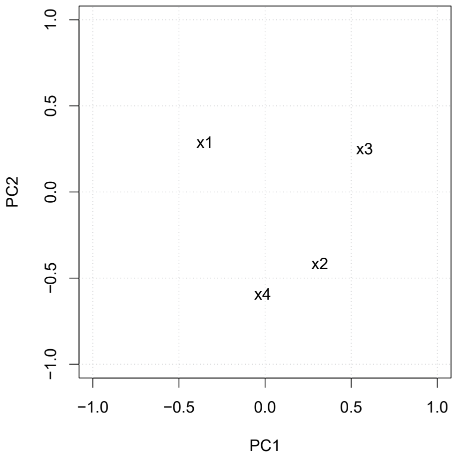
①



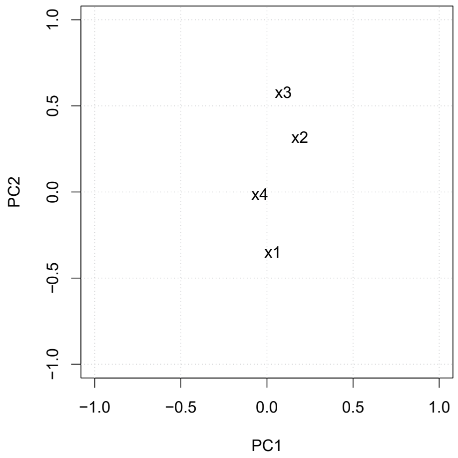
②



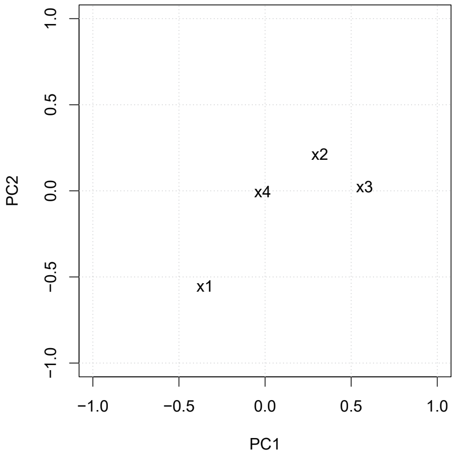
③



④



⑤



〔3〕次に，主成分スコア $z_1, \ldots, z_8$ を説明変数とした線形回帰モデルを考える。説明変数として $\{z_1\}, \{z_1, z_2\}, \ldots, \{z_1, z_2, \ldots, z_8\}$ を用いた階層型モデルを順にモデル 1, モデル 2, ..., モデル 8 とする。このとき，各モデルの AIC を計算したところ，図 1 の結果が得られた。

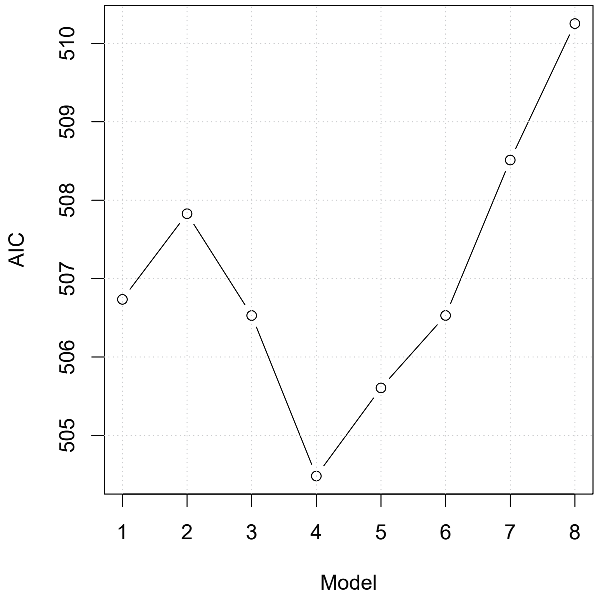


図 1　各モデルの AIC の値

ここで，縦軸は AIC の値，横軸は AIC を計算するために用いたモデルを示している。図 1 から，予測の観点から最適なモデルは何か。次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **11**

① モデル 2　② モデル 4　③ モデル 6
④ モデル 7　⑤ モデル 8

〔4〕次の ① ～ ⑤ の文章のうちから最も適切なものを一つ選べ。 **12**

① 主成分分析を行う際には，前処理としてデータを標準化することが不可欠である。

② 相関行列に対する主成分分析では，各主成分の主成分負荷量（因子負荷量）はその主成分ともとの変量との相関係数と一致する。

③ AIC を用いて比較できるのはモデルのパラメータ集合間に包含関係がある場合のみである。

④ AIC の特徴として，一般にモデル同定の一致性を持つことがあげられる。

⑤ AIC によるモデル選択は，交差検証法に比べて一般に計算量が大きくなるという欠点がある。

問 7　平均 μ が未知，分散 σ^2 が既知の正規分布に従うサイズ 1 の標本 $X \sim N(\mu, \sigma^2)$ が観測されたとする。このとき，μ に対する事前分布として正規分布 $N(\mu_0, \sigma_0^2)$ を仮定すると，事後分布も正規分布となるが，これを $N(\tilde{\mu}, \tilde{\sigma}^2)$ と表すことにする。例えば，$\mu_0 = 0, \sigma_0 = 1, \sigma = 2$ のときに観測値 $X = 2$ が得られた場合の事前分布と事後分布の密度関数のグラフは，それぞれ図 1 の破線と実線のようになる。

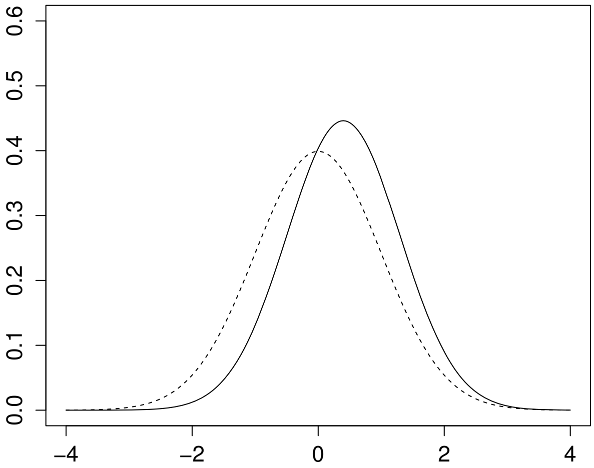


図 1　事前分布 (破線) と事後分布 (実線) の密度関数

〔1〕事後平均と事後分散の組合せ $(\tilde{\mu}, \tilde{\sigma}^2)$ として，次の ① ～ ⑤ のうちから適切なものを一つ選べ。 **13**

① $\tilde{\mu} = \dfrac{\sigma^2 X + \sigma_0^2 \mu_0}{\sigma^2 + \sigma_0^2}, \quad \tilde{\sigma}^2 = \sigma\sigma_0$

② $\tilde{\mu} = \dfrac{\sigma^2 X + \sigma_0^2 \mu_0}{\sigma^2 + \sigma_0^2}, \quad \tilde{\sigma}^2 = \dfrac{\sigma^2 + \sigma_0^2}{2}$

③ $\tilde{\mu} = \dfrac{\sigma^2 X + \sigma_0^2 \mu_0}{\sigma^2 + \sigma_0^2}, \quad \tilde{\sigma}^2 = \left(\dfrac{1}{\sigma^2} + \dfrac{1}{\sigma_0^2}\right)^{-1}$

④ $\tilde{\mu} = \dfrac{\sigma_0^2 X + \sigma^2 \mu_0}{\sigma^2 + \sigma_0^2}, \quad \tilde{\sigma}^2 = \dfrac{\sigma^2 + \sigma_0^2}{2}$

⑤ $\tilde{\mu} = \dfrac{\sigma_0^2 X + \sigma^2 \mu_0}{\sigma^2 + \sigma_0^2}, \quad \tilde{\sigma}^2 = \left(\dfrac{1}{\sigma^2} + \dfrac{1}{\sigma_0^2}\right)^{-1}$

〔2〕ある施設で養殖されている蟹の重さ（グラム）の測定データを正規分布で近似したうえで，平均の事後分布を求める。ただし，蟹の重さの平均の事前分布については，測定日の前日までのデータをもとにして得られた $N(13, 2.7^2)$ を用い，分散 σ^2 としては当日のデータの標本分散を用いることにする。当日の測定で次の 10 個体のデータが新たに得られたとする。

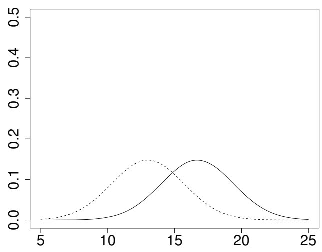
重さ (g)	14.05	19.25	23.00	16.00	13.90	14.70	20.35	15.05	15.30	15.50

平均： 16.71，　標準偏差： 3.07

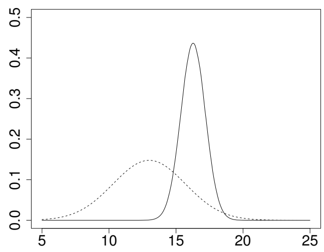
資料: CRAN Package ‘isdals’

このとき，事前分布（破線）と事後分布（実線）の密度関数のグラフとして，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **14**

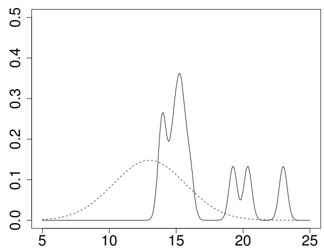
①



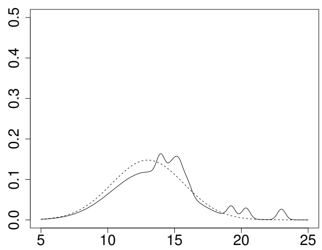
②



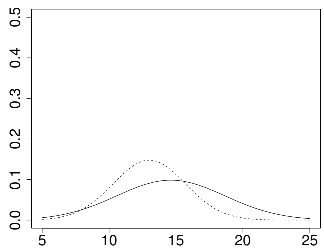
③



④



⑤



〔3〕上の正規分布の例のように，特定の確率分布のパラメータに対して，事前分布と事後分布が同じ分布族に属するような性質を持つ事前分布は共役事前分布とよばれる。これに関して述べた次の (A) ～ (C) の文章の正誤について，下の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **15**

(A) 共役事前分布を用いる利点の一つは，事後分布の計算がハイパーパラメータの更新として表現できる点である。

(B) 正規分布の平均が既知で分散が未知のとき，分散に対する共役事前分布としては例えばベータ分布を用いることができる。

(C) 共役事前分布を用いることができない場合には，一般にモンテカルロ法等の数値計算を用いて事後分布を近似計算する。

① (A)，(B)，(C) はすべて正しい。
② (A)，(B) のみが正しい。
③ (A)，(C) のみが正しい。
④ (B)，(C) のみが正しい。
⑤ (A)，(B)，(C) はすべて誤り。

問 8　中性子モニターは宇宙から地球の大気に当たる高エネルギー荷電粒子の数を測定するために設計された地上ベースの検出器である。 図 1 は，2000 年 10 月から 2018 年 10 月までの，フィンランド Oulu 大学における中性子モニターによって計測された中性子のカウントデータ (月次）である。

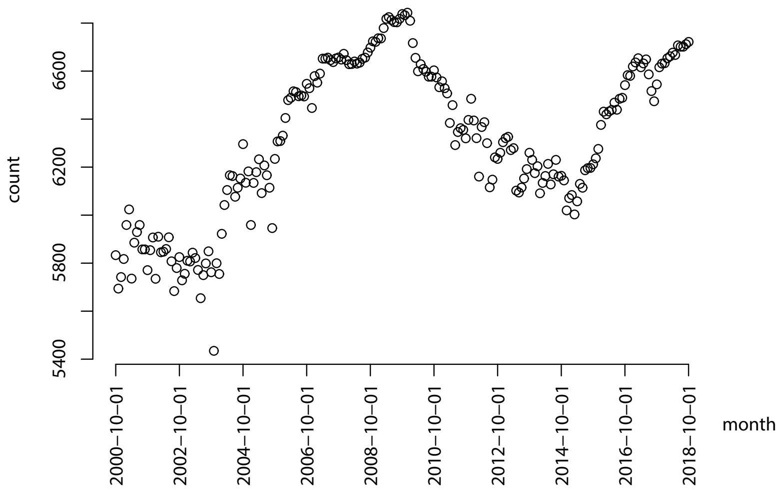


図 1: 中性子のカウントデータ

資料: Cosmic Ray Station of the University of Oulu
(`http://cosmicrays.oulu.fi/#solar`)

217 点の観測データを $y_i(i = 1, 2, \ldots, 217)$ とし，このデータに対して平滑化を行うことで，中性子数の変動に関する特徴を捉えたい。特に Fused Lasso は，次の式により (y_i) を平滑化した実数列 (β_i) を生成する手法である。

$$\hat{\boldsymbol{\beta}} = \mathop{\arg\min}_{\boldsymbol{\beta}\in\mathbb{R}^{217}} \frac{1}{2}\sum_{i=1}^{217}(y_i - \beta_i)^2 + \lambda\sum_{i=1}^{216}|\beta_{i+1} - \beta_i|.$$

ここで λ は非負の平滑化パラメータであり，$\mathop{\arg\min}_{\boldsymbol{\beta}\in\mathbb{R}^{217}} f(\boldsymbol{\beta})$ は関数 $f(\boldsymbol{\beta})$ を $\boldsymbol{\beta} \in \mathbb{R}^{217}$（217 次元ベクトルの集合）の範囲で最小化するような $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_{217})$ の値を意味する。

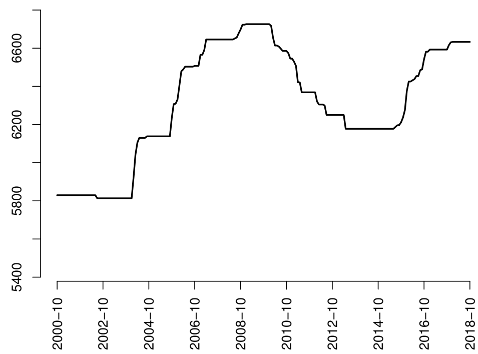
〔1〕$\lambda = 500$ としたFused Lassoで平滑化を行った結果の図として，次の ① ～ ④ から最も適切なものを一つ選べ。 **16**

①

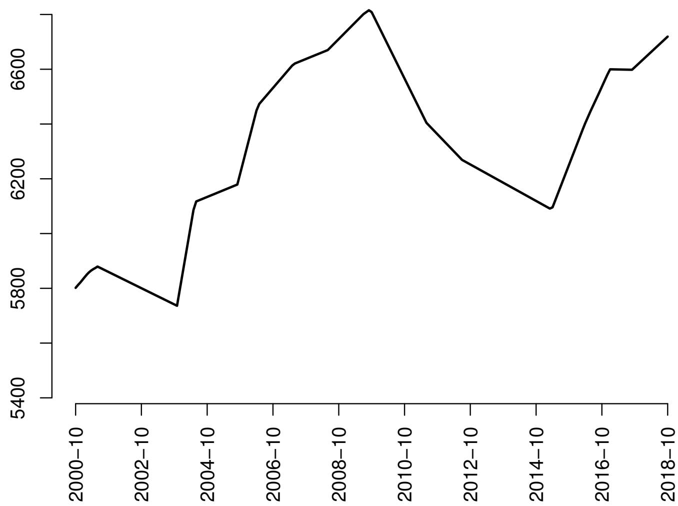
②

③

④



〔2〕次に，同じデータに対して別の平滑化手法を適用したところ次の図のようになった。



この結果に対応する平滑化手法として，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 $\boxed{\textbf{17}}$

① $\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}\in\mathbb{R}^{217}}{\arg\min} \frac{1}{2}\sum_{i=1}^{217}|y_i - \beta_i| + 100\sum_{i=1}^{217}\beta_i^2$

② $\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}\in\mathbb{R}^{217}}{\arg\min} \frac{1}{2}\sum_{i=1}^{217}|y_i - \beta_i| + 500\sum_{i=1}^{217}|\beta_i|$

③ $\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}\in\mathbb{R}^{217}}{\arg\min} \frac{1}{2}\sum_{i=1}^{217}(y_i - \beta_i)^2 + 500\sum_{i=1}^{217}|\beta_i|$

④ $\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}\in\mathbb{R}^{217}}{\arg\min} \frac{1}{2}\sum_{i=1}^{217}(y_i - \beta_i)^2 + 500\sum_{i=1}^{215}|\beta_{i+2} - 2\beta_{i+1} + \beta_i|$

⑤ $\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}\in\mathbb{R}^{217}}{\arg\min} \frac{1}{2}\sum_{i=1}^{217}(y_i - \beta_i)^2 + 500\sum_{i=1}^{214}|\beta_{i+3} - 3\beta_{i+2} + 3\beta_{i+1} - \beta_i|$

問９は次のページにあります。

問 9　図 1 および図 2 は，ある一週間における米ドル／円とユーロ／円の為替レートの一例である。

図 1: 米ドル／円 為替レート

図 2: ユーロ／円 為替レート

〔1〕一週間における取引市場の開始時刻を $t = 0$，終了時刻を $t = 100$ として，時刻 $t \in [0, 100]$ において 1 米ドル $= x_t$ 円が以下の式に従うとする：

$$x_t = x_0 + \sigma B_t.$$

ただし，$x_0, \sigma > 0$ は定数で，$(B_t)_{0 \le t \le 100}$ は標準ブラウン運動とする。観測データ x_k $(k = 0, 1, \ldots, 100)$ を用いて x_t の増分の二乗の平均を計算したところ，

$$V = \frac{1}{100} \sum_{k=1}^{100} (x_k - x_{k-1})^2 = 0.001224$$

となった。このとき，観測データを用いて σ をモーメント法により推定したときの推定値 $\hat{\sigma}$ として，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **18**

① 0.019　② 0.035　③ 0.11　④ 0.19　⑤ 0.35

〔2〕〔1〕とは別の一週間の取引市場に対して得られた，より高頻度の観測データ $x_{k/10}$ $(k = 0, 1, \ldots, 1000)$ を用いて x_t の増分の二乗の平均を計算したところ，

$$V_1 = \frac{1}{1000} \sum_{k=1}^{1000} (x_{\frac{k}{10}} - x_{\frac{k-1}{10}})^2 = 0.000595$$

となった。〔1〕と同じ確率過程モデルを仮定したとき，この観測データを用いて σ をモーメント法により推定したときの推定値 $\hat{\sigma}$ として，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **19**

① 0.0077　② 0.024　③ 0.077　④ 0.24　⑤ 0.77

〔3〕次に，1 米ドル $= x_t$ 円，1 ユーロ $= y_t$ 円としたとき，x_t と y_t が以下の式に従うとする：

$$\begin{aligned} x_t &= x_0 + \sigma_1\sqrt{\rho}B_t^{(1)} + \sigma_1\sqrt{1-\rho}B_t^{(2)}, \\ y_t &= y_0 + \sigma_2\sqrt{\rho}B_t^{(1)} + \sigma_2\sqrt{1-\rho}B_t^{(3)}. \end{aligned}$$

ただし，$x_0, y_0, \sigma_1 > 0, \sigma_2 > 0, \rho \in (0,1)$ は定数で，$(B_t^{(1)}, B_t^{(2)}, B_t^{(3)})_{0 \le t \le 100}$ は 3 個の独立な標準ブラウン運動とする。観測データ $\{x_{k/10}, y_{k/10}\}$ $(k = 0, 1, \ldots, 1000)$ を用いて，x_t, y_t の増分の二乗の平均と積和の平均を計算したところ，

$$V_1 = \frac{1}{1000}\sum_{k=1}^{1000}(x_{\frac{k}{10}} - x_{\frac{k-1}{10}})^2 = 0.000595, \quad V_2 = \frac{1}{1000}\sum_{k=1}^{1000}(y_{\frac{k}{10}} - y_{\frac{k-1}{10}})^2 = 0.001008,$$

$$V_{1,2} = \frac{1}{1000}\sum_{k=1}^{1000}(x_{\frac{k}{10}} - x_{\frac{k-1}{10}})(y_{\frac{k}{10}} - y_{\frac{k-1}{10}}) = 0.000292$$

と計算された。このとき，観測データを用いて ρ をモーメント法により推定したときの推定値 $\hat{\rho}$ として，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。

20

① 0.12　② 0.25　③ 0.38　④ 0.51　⑤ 0.64

問 10　1986 年に起きたスペースシャトル「チャレンジャー号」の爆発事故は，右側固体燃料補助ロケットの密閉用 O リングの破損が原因であったと考えられる。S. R. Dalal, E. B. Fowlkes and B. Hoadley (1989) では，チャレンジャー号爆発事故より以前の 23 回のスペースシャトルの打ち上げにおける，外気温と O リングの破損の有無のデータを分析している。このデータについて，統計ソフトウェアを利用してロジスティック回帰モデルを推定したところ，以下のような出力結果が得られた。ただし，「外気温（華氏）」と「破損の有無（1 が破損あり，0 が破損なし）」に対応する変数名をそれぞれ `Temperature`，`TD` としている。また，出力結果の `(Intercept)` は回帰モデルの定数項を意味している。

出力結果

```
Deviance Residuals:
    Min        1Q    Median       3Q      Max
-1.0611   -0.7613  -0.3783   0.4524   2.2175

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)  15.0429     7.3786   2.039   0.0415
Temperature  -0.2322     0.1082  -2.145   0.0320
```

$i = 1, \ldots, 23$ について，x_i は i 番目の打ち上げ時の外気温（華氏），y_i は i 番目の打ち上げでの O リングの破損の有無とする。次の文章は，このデータに対するロジスティック回帰モデルの説明文である。

> y_i は互いに独立な確率変数 Y_i の実現値であり，Y_i は（ア）に従う。
> π_i $(0 < \pi_i < 1)$ について構造式（イ）を仮定する。

出典：S. R. Dalal, E. B. Fowlkes and B. Hoadley (1989). Risk analysis of the space shuttle: Pre-Challenger prediction of failure. (*Journal of the American Statistical Association*, **84**, 945–957)

〔1〕（ア）にあてはまる分布はどれか。次の ① ～ ④ のうちから適切なものを一つ選べ。 **21**

① ベルヌーイ分布 $Bin(1, \pi_i)$　　② 二項分布 $Bin(23, \pi_i)$

③ ポアソン分布 $Po(\pi_i)$　　④ 正規分布 $N(\pi_i, 1)$

〔2〕（イ）にあてはまる式はどれか。次の ① ～ ④ のうちから適切なものを一つ選べ。 **22**

① $\log \dfrac{\pi_i}{1 - \pi_i} = \alpha + \beta x_i$　　② $\dfrac{\exp(\pi_i)}{1 + \exp(\pi_i)} = \alpha + \beta x_i$

③ $\log \pi_i = \alpha + \beta x_i$　　④ $\pi_i = \alpha + \beta x_i$

〔3〕ロジスティック回帰モデルの推定の出力結果によると，O リングの破損確率が 0.5 となるのは，外気温が何度のときか。次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。 **23**

① 14.9 °F　② 24.3 °F　③ 54.9 °F　④ 64.8 °F　⑤ 71.7 °F

〔4〕チャレンジャー号の事故当日は，異常寒波の影響で外気温は 31°F であった。ロジスティック回帰モデルが正しいと仮定したときの，事故当日の O リングの破損確率の推定値として，次の ① ～ ⑤ のうちから最も適切なものを一つ選べ。必要に応じて付表 5 を用いよ。 **24**

① 0.0028　② 0.1589　③ 0.5820　④ 0.8979　⑤ 0.9996

問 11　企業の信用状態に対する評価（格付）が，ある格付会社によって A（優良），B（投資適格），C（投資不適格・債務不履行）の三つに分類されている。格付は一年毎に更新され，各企業の格付の推移がマルコフ連鎖で表される。各企業の格付推移はそれぞれ独立であるとする。状態 (A,B,C) を $(1,2,3)$ に対応させ，$1 \leq i,j \leq 3$ に対して，一年後に状態 i から状態 j へ推移する確率を p_{ij} と書く。推移確率行列 $M=(p_{ij})_{1\leq i,j\leq 3}$ が定数 $\theta,\phi \in [0,1], \phi+\theta \leq 1$ を用いて以下のように表される：

$$M = \begin{pmatrix} 1-\theta & \theta & 0 \\ \theta & 1-\theta-\phi & \phi \\ 0 & \phi & 1-\phi \end{pmatrix}.$$

〔1〕ある年に格付 A の企業が 100 社，格付 B の企業が 20 社，格付 C の企業が 0 社あったとして，次の年において，A → B に推移した企業の数が 5 社，B → A に推移した企業の数が 1 社で他の企業に格付の変化はなかったとする。$\phi = 0.01$ である時，θ の最尤推定値はいくらか。小数点第 3 位を四捨五入して答えよ。 $\boxed{\textbf{記述 7}}$

〔2〕n を正の整数とする。M の固有値を λ_j $(j=1,2,3)$ とし，直交行列 $U=(u_{ij})_{1\leq i,j\leq 3}$ を

$$U^\top M U = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

をみたすようにとる。ただし，$U^\top$ は行列 U の転置を表す。このとき，t 年において格付 A の企業が $t+n$ 年において格付 C である確率を n, λ_j $(j=1,2,3), u_{ij}$ $(1 \leq i,j \leq 3)$ の式で表せ。 $\boxed{\textbf{記述 8}}$

問 12　時系列モデルに関する以下の問いに答えよ。ただし，ε_t $(t = \dots, -1, 0, 1, \dots)$ は互いに独立に $N(0, \sigma^2)$ に従う確率変数列とする。

〔1〕自己回帰モデル AR(p) とは以下の式で表される時系列モデルである。

$$X_t = c + \sum_{i=1}^{p} a_i X_{t-i} + \varepsilon_t \quad (t = \dots, -1, 0, 1, \dots).$$

ただし，c と各 a_i は実数である。AR(p) モデルが定常であることの必要十分条件は方程式

$$1 - a_1 z - \cdots - a_p z^p = 0$$

のすべての解の絶対値が 1 より大きくなることである。AR(2) モデルが $a_1 = a_2 = a$ $(0 < a)$ のとき，定常であるための a に関する必要十分条件を求めよ。**記述 9**

〔2〕移動平均モデル MA(q) とは以下の式で表される時系列モデルである。

$$X_t = c + \varepsilon_t + \sum_{i=1}^{q} b_i \varepsilon_{t-i} \quad (t = \dots, -1, 0, 1, \dots).$$

ただし c と各 b_i は実数であるとする。

すべての b_i の値を 0.5 に設定した MA(q) モデルにより生成された時系列データ x_t $(t = 1, \dots, 3000)$ のコレログラムを作成したところ，図 1 のようになった。（図中の破線は時系列が無相関であるという帰無仮説のもとでの有意水準 5 %の棄却限界値を表す。）モデルの次数 q の値はいくつと推測できるか。また，そのように考えられる理由を数式を用いて説明せよ。**記述 10**

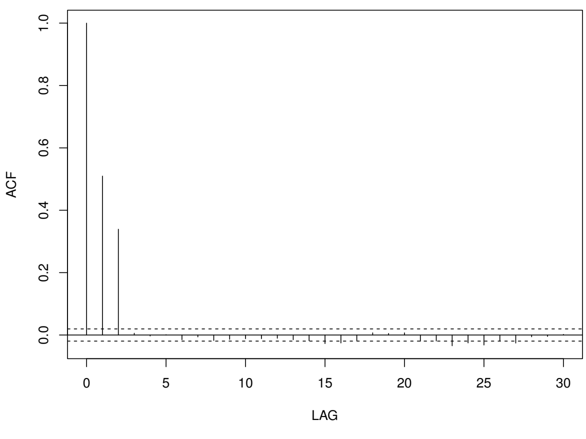


図 1　MA(q) のコレログラム

論述問題

（3 問中 1 問選択）

問 1　白葉枯病（しらはがれ病）は水稲の感染病のひとつで，菌に感染した水稲の葉は縁部分から内側に向かい徐々に枯れてしまう。表は，水稲 4 品種について，温室内のポットを実験単位とし，白葉枯病に感染した株の割合を観測したデータである。各水準の繰り返し数 n_i $(i = 1, 2, 3, 4)$ は一定ではない。

品種	n_i	割合（%）					平均
A_1	4	34	31	29	28		30.5
A_2	5	25	28	30	28	27	27.6
A_3	5	32	31	30	31	28	30.4
A_4	4	33	32	29	34		32.0
計	18						30.0

水準 A_i の第 j 番目の観測データを y_{ij} とする（$i = 1, \ldots, 4;\ j = 1, \ldots, n_i$）。$y_{ij}$ を確率変数 Y_{ij} の実現値とみなし，その期待値を $E[Y_{ij}] = \mu_i$ とする。さらに，以下の，一元配置分散分析モデルを仮定した。

$$Y_{ij} = \mu_i + \varepsilon_{ij} = \mu + \alpha_i + \varepsilon_{ij},\quad i = 1, \ldots, 4,\ j = 1, \ldots, n_i$$

ただし ε_{ij} は互いに独立に正規分布 $N(0, \sigma^2)$ に従う確率変数であると仮定する。

〔1〕通常，一元配置分散分析では，母数 $\alpha_1, \ldots, \alpha_4$ について

$$\sum_{i=1}^{4} \alpha_i = 0 \quad \text{あるいは} \quad \sum_{i=1}^{4} n_i \alpha_i = 0$$

のような制約を仮定する。その理由を説明せよ。また，それぞれの制約の下で，母数 μ, α_i と母平均 μ_i の関係を説明せよ。

〔2〕表のデータに対して，次の帰無仮説および対立仮説

H_0 ： $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$
H_1 ： $\alpha_i \neq \alpha_j$ となる i, j が存在する

を考える。分散分析表の空欄を埋め，検定を行い，その結果の解釈を述べよ。また，母分散 σ^2 の不偏推定量の値を答えよ。

要因	平方和	自由度	分散	F 値
品種				
誤差		$(=\nu)$		
合計				

ただし，誤差の自由度（$=\nu$ とおいた）は次の設問で使う。

〔3〕4 つの品種のうち，A_1, A_2 と A_3, A_4 は，それぞれ異なる母本（品種の親）からの品種であり，白葉枯病に対する抵抗力が異なっている可能性がある。そこで，次の帰無仮説および対立仮説

$$H_0 \ :\ \frac{\mu_1 + \mu_2}{2} = \frac{\mu_3 + \mu_4}{2}$$

$$H_1 \ :\ \frac{\mu_1 + \mu_2}{2} \neq \frac{\mu_3 + \mu_4}{2}$$

の検定を考える。各水準の観測データの標本平均を $\bar{Y}_i = \dfrac{1}{n_i}\displaystyle\sum_{j=1}^{n_i} Y_{ij}$ とする（$i = 1, \ldots, 4$）。検定統計量

$$T = \frac{1}{c}\left(\frac{\bar{Y}_1 + \bar{Y}_2}{2} - \frac{\bar{Y}_3 + \bar{Y}_4}{2}\right)$$

による有意水準 α の両側検定の棄却域が

$$|T| > t_{\alpha/2}(\nu)$$

で与えられるとするとき，c を求めよ。ただし，ν は分散分析表における誤差の自由度の値であり，$t_{\alpha/2}(\nu)$ は自由度 ν の t 分布の上側確率 $\alpha/2$ に対する t の値である。さらに，表のデータについてこの検定を実行し，結論を述べよ。

〔4〕新人データアナリストのＮ君は，表のデータを見て,「品種 A_2 の平均値が最も小さく，品種 A_4 の平均値が最も大きい」ことに注目し，次のことを主張した。

> 品種 A_2 のデータと品種 A_4 のデータから，帰無仮説および対立仮説
>
> $$H_0:\ \mu_2 = \mu_4 \quad \text{vs} \quad H_1:\ \mu_2 < \mu_4$$
>
> の検定を，母分散が等しいと仮定した二標本 t 検定により行うと，t 統計量の値は
>
> $$t = \frac{27.6 - 32.0}{\sqrt{(\frac{1}{5} + \frac{1}{4})\hat{\sigma}^2}} = -3.327 < -2.998 = -t_{0.01}(7)$$
>
> となる（母分散の不偏推定値は $\hat{\sigma}^2 = 3.886$ である）。従って，H_0 は有意水準１％で棄却される。つまり，品種 A_2 は品種 A_4 よりも白葉枯病に対する抵抗力が強いと，有意水準１％で主張できる。

このＮ君の主張に対し，上司のＫ氏は,「検定の多重性が考慮されていない」と指摘した。Ｋ氏の指摘する『検定の多重性』とは何を意味するか説明し，Ｎ君の主張のどこが不適切であるかを説明せよ。

問 2　ある学校の同級生の A 君，B 君，C 君が互いに課題レポートを写しあっているのではないかと T 教員は疑っている。そこで，T 教員は各生徒の課題レポートの点数が正規分布に従うと仮定したうえで，グラフィカルモデルを用いて解析することにした。ここでいうグラフィカルモデルとは，以下のような性質をもつ統計モデルである。

多変量正規分布 $N(\mathbf{0}, \Sigma)$ に従う確率ベクトル $(X_1, \ldots, X_d)$ に対して，頂点集合 $V = \{v_1, \ldots, v_d\}$ をもつ無向グラフ (V, E) の辺集合 E を以下のように定義する。「もし，2 つの頂点 v_i と v_j が辺で結ばれていなければ，X_i と X_j はそれ以外の確率変数で条件付けたときに条件付き独立である。」このようにグラフによって条件付き独立性が表現される統計モデルをグラフィカルモデルという。

〔1〕確率ベクトル (X_1, X_2, X_3) は $N_3(\mathbf{0}, \Sigma)$ に従い，かつその条件付き独立性が図 1 のグラフで表されているとする。

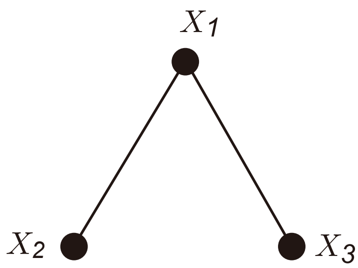


図 1　グラフィカルモデルの例

このグラフの構造からわかる共分散行列 Σ の性質を述べよ。ただし共分散行列の正則性は仮定してよい。また，理由の説明は必要ない。

〔2〕A 君，B 君，C 君の課題レポートの点数 (X_A, X_B, X_C) の標本相関行列 $\hat{C}$ およびその逆行列 $\hat{C}^{-1}$ の成分は以下のようになった。

$$\hat{C} = \begin{pmatrix} 1.00 & 0.80 & 0.52 \\ 0.80 & 1.00 & 0.65 \\ 0.52 & 0.65 & 1.00 \end{pmatrix}, \quad \hat{C}^{-1} = \begin{pmatrix} 2.78 & -2.22 & 0.00 \\ -2.22 & 3.51 & -1.13 \\ 0.00 & -1.13 & 1.73 \end{pmatrix}.$$

この結果をふまえて 3 名の学生にヒアリングを行ったところ，以下のような発言を得た。

A 君：「私は B 君にレポートの解答のヒントをあげていますが，B 君はそれを C 君に伝えているみたいです。」

B 君：「僕はいつも A 君と C 君の両方にレポートのヒントをあげています。」

C 君：「B 君は僕と A 君の両方のレポートをうまく合成して自分のレポートを作成しています。」

A 君，B 君，C 君の発言のうち，標本相関行列 $\hat{C}$ を用いたグラフィカルモデル解析結果と最も適合しないのは誰の発言か。そう考える理由とともに述べよ。

〔3〕調査を進めるうえで，新たにD君の点数と上記3名の点数との相関が高いことが分かった。そこで，D君の点数 X_D を目的変数，3名の点数 X_A, X_B, X_C を説明変数として線形回帰を行い，AICで変数選択を行ったところ次のようなモデルが選ばれた。

$$X_D = \alpha X_A + \gamma X_C + \epsilon$$

ここで，回帰係数 α, γ の推定値は有意に正の値を持っていた。さらに残差 ϵ は正規分布に従い，X_A, X_B, X_C との相関が十分に小さいことも確認できた。このとき，(X_A, X_B, X_C, X_D) の条件付き独立性を最も良く表すような，X_A, X_B, X_C, X_D の4頂点を持つグラフを描け。

〔4〕さらに調査を進めると，他の学級，学年も含めた巨大なレポート情報シンジケートがあることがわかってきた。今後の調査はグラフィカルモデルの頂点数が非常に多くなり，グラフ構造の推定に必要な行列演算の計算量が心配される。ただし，各学級から他の学級への情報の伝達が，各クラスの代表者一人のみを介して行われるとわかっているときには，推定に必要な計算は比較的容易になる。この理由を，確率密度関数の性質とグラフ構造の観点から説明せよ。

問 3　図 1 のような 2 次元データがあり，それぞれ平均および分散共分散行列が

$$\boldsymbol{\mu}_1 = \begin{pmatrix} 1.5 \\ 1.5 \end{pmatrix}, \quad \boldsymbol{\mu}_2 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \quad \Sigma_1 = \begin{pmatrix} 0.8 & 0 \\ 0 & 0.8 \end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix} 0.4 & 0 \\ 0 & 0.4 \end{pmatrix}$$

である 2 変量正規分布 $N_2(\boldsymbol{\mu}_1, \Sigma_1)$ および $N_2(\boldsymbol{\mu}_2, \Sigma_2)$ に従う 2 つのクラスに属するものとする。なお，2 変量正規分布 $N_2(\boldsymbol{\mu}, \Sigma)$ の同時確率密度関数は

$$f(\boldsymbol{x}; \boldsymbol{\mu}, \Sigma) = (2\pi)^{-1}(\det \Sigma)^{-1/2} \exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)$$

である。ただし，$\boldsymbol{x}^\top$ は列ベクトル $\boldsymbol{x}$ の転置を表す。

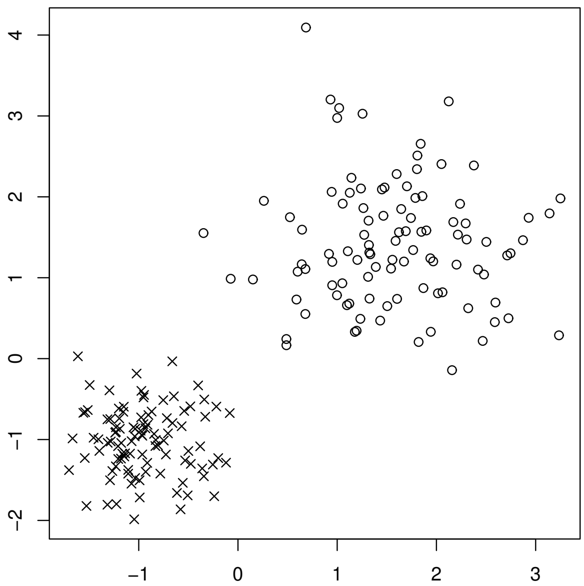


図 1: 判別するデータの集合（○がクラス 1，×がクラス −1 のデータを表す）

〔1〕 二次判別分析を用いて，2 次元の列ベクトルの入力 $\boldsymbol{x}$ から，1 もしくは −1 のクラスラベル y を予測する。$\boldsymbol{x}$ で条件付けたクラスラベル y の確率の比を用いて

$$\hat{y} = \mathrm{sign}\left(\log \frac{\Pr(y=1|\boldsymbol{x})}{\Pr(y=-1|\boldsymbol{x})}\right)$$

で 2 クラス判別器を構成する。$\Pr(\boldsymbol{x}|y=1)$ 及び $\Pr(\boldsymbol{x}|y=-1)$ に対応する条件付き分布をそれぞれ $N(\boldsymbol{\mu}_1, \Sigma_1)$，　$N(\boldsymbol{\mu}_2, \Sigma_2)$ としたとき，この判別関数は $\boldsymbol{x}$ の二次関数

$$f_q(\boldsymbol{x}) = \boldsymbol{x}^\top A\boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{x} + c, \quad A \in \mathbb{R}^{2\times 2}, \boldsymbol{b} \in \mathbb{R}^2$$

となり，判別器はその符号として表すことができる。ただし，A は 2×2 の対称行列，$\boldsymbol{b} \in \mathbb{R}^2$ は 2 次元の列ベクトルである。このとき，$A, \boldsymbol{b}$ を $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \Sigma_1, \Sigma_2$ を用いて表わせ。また，この判別関数が $\boldsymbol{x}$ の一次関数になるための条件を述べよ。

〔2〕サンプルデータから $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \Sigma_1, \Sigma_2$ を経験平均，経験分散共分散行列として推定する。$f_q(\boldsymbol{x})$ における $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \Sigma_1, \Sigma_2$ を推定値で置き換えたものを $\hat{f}_q(\boldsymbol{x})$ として $\hat{f}_q(\boldsymbol{x})$ の等高線を示したものが図 2 である。

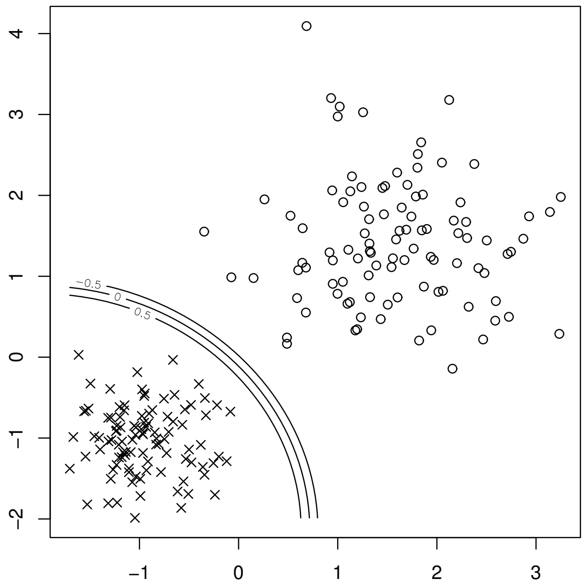


図 2: 二次判別分析による $\hat{f}_q(\boldsymbol{x})$ の等高線

経験平均，経験分散共分散行列がそれぞれ

$$\hat{\boldsymbol{\mu}}_1 = \begin{pmatrix} 1.57 \\ 1.41 \end{pmatrix}, \quad \hat{\boldsymbol{\mu}}_2 = \begin{pmatrix} -0.95 \\ -1.00 \end{pmatrix}, \quad \hat{\Sigma}_1 = \begin{pmatrix} 0.53 & 0 \\ 0 & 0.60 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 0.14 & 0 \\ 0 & 0.17 \end{pmatrix}$$

であったとき,〔1〕で求めた判別器における $A, \boldsymbol{b}$ の値を小数点以下第 2 位まで求めよ。

〔3〕学習データ $\{\boldsymbol{x}_i, y_i\}$ $(i = 1, 2, \ldots, n)$ に対して，サポートベクトルマシン（SVM）の判別関数は，カーネル関数 $k(\boldsymbol{x}, \tilde{\boldsymbol{x}})$ を用いて

$$f_s(\boldsymbol{x}) = \sum_{i \in SV} \alpha_i y_i k(\boldsymbol{x}_i, \boldsymbol{x}) + \beta$$

で表すことができる。ただし SV はサポートベクトル集合を表し，α_i $(i = 1, 2, \ldots, n), \beta$ は実数パラメータである。多項式カーネル

$$k(\boldsymbol{x}, \tilde{\boldsymbol{x}}) = (\boldsymbol{x}^\top \tilde{\boldsymbol{x}} + 1)^2$$

を採用した SVM の判別関数 $f_s(\boldsymbol{x})$ は，入力 $\boldsymbol{x} = (x_1, x_2)^\top \in \mathbb{R}^2$ の二次関数となり，これを

$$f_s(\boldsymbol{x}) = \boldsymbol{x}^\top \tilde{A} \boldsymbol{x} + \tilde{\boldsymbol{b}}^\top \boldsymbol{x} + \tilde{c}$$

とする。このときの行列 $\tilde{A}$ の式を， 学習データ $\{\boldsymbol{x}_i, y_i\}$ $(i = 1, 2, \ldots, n)$ および最適化されたパラメータ値 α_i $(i = 1, 2, \ldots, n), \beta$ を用いて表せ。

〔4〕上記の多項式カーネルを用いたSVMを用いて判別器を学習したところ，判別関数 $f_s(\boldsymbol{x})$ は図3のようになった。

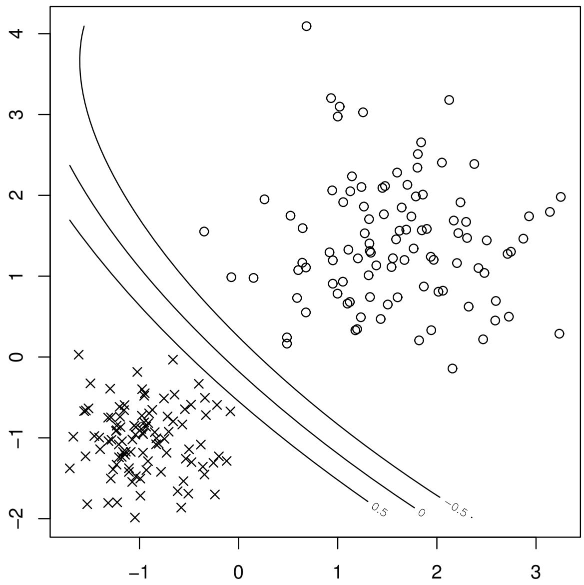
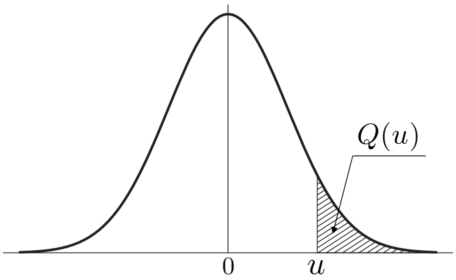


図3: サポートベクトルマシンによる判別関数の等高線

このように，二次判別分析における判別器とサポートベクトルマシンにおける判別器は，どちらも二次関数により記述されても形状が大きく異なることがある。本問題のデータに対してはどちらを用いる方が適当かを，その理由とともに説明せよ。

付　表

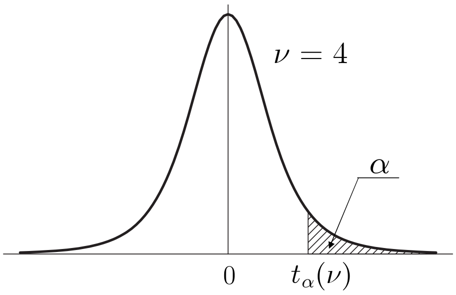
付表 1. 標準正規分布の上側確率



u	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
3.6	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
3.7	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
3.8	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
3.9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

$u = 0.00 \sim 3.99$ に対する，正規分布の上側確率 $Q(u)$ を与える。
例：$u = 1.96$ に対しては，左の見出し 1.9 と上の見出し .06 との交差点で，$Q(u) = .0250$ と読む。表にない u に対しては適宜補間すること。

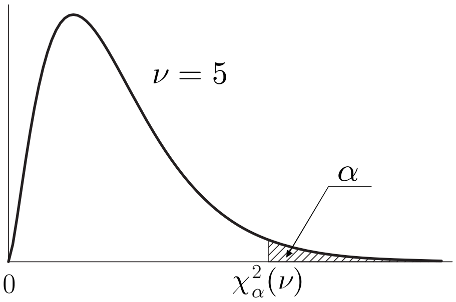
付表 2. t 分布のパーセント点



ν	α = 0.10	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.821	63.656
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
240	1.285	1.651	1.970	2.342	2.596
∞	1.282	1.645	1.960	2.326	2.576

自由度 ν の t 分布の上側確率 α に対する t の値を $t_\alpha(\nu)$ で表す。
例：自由度 $\nu = 20$ の上側 5% 点 ($\alpha = 0.05$) は，$t_{0.05}(20) = 1.725$ である。
表にない自由度に対しては適宜補間すること。

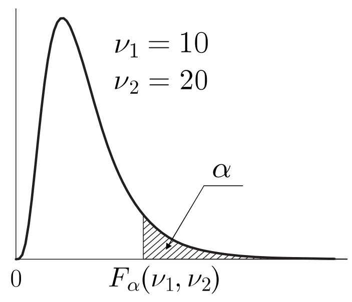
付表 3. カイ二乗分布のパーセント点



	α							
ν	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01
1	0.00	0.00	0.00	0.02	2.71	3.84	5.02	6.63
2	0.02	0.05	0.10	0.21	4.61	5.99	7.38	9.21
3	0.11	0.22	0.35	0.58	6.25	7.81	9.35	11.34
4	0.30	0.48	0.71	1.06	7.78	9.49	11.14	13.28
5	0.55	0.83	1.15	1.61	9.24	11.07	12.83	15.09
6	0.87	1.24	1.64	2.20	10.64	12.59	14.45	16.81
7	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48
8	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09
9	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67
10	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21
11	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.72
12	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22
13	4.11	5.01	5.89	7.04	19.81	22.36	24.74	27.69
14	4.66	5.63	6.57	7.79	21.06	23.68	26.12	29.14
15	5.23	6.26	7.26	8.55	22.31	25.00	27.49	30.58
16	5.81	6.91	7.96	9.31	23.54	26.30	28.85	32.00
17	6.41	7.56	8.67	10.09	24.77	27.59	30.19	33.41
18	7.01	8.23	9.39	10.86	25.99	28.87	31.53	34.81
19	7.63	8.91	10.12	11.65	27.20	30.14	32.85	36.19
20	8.26	9.59	10.85	12.44	28.41	31.41	34.17	37.57
25	11.52	13.12	14.61	16.47	34.38	37.65	40.65	44.31
30	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89
35	18.51	20.57	22.47	24.80	46.06	49.80	53.20	57.34
40	22.16	24.43	26.51	29.05	51.81	55.76	59.34	63.69
50	29.71	32.36	34.76	37.69	63.17	67.50	71.42	76.15
60	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38
70	45.44	48.76	51.74	55.33	85.53	90.53	95.02	100.43
80	53.54	57.15	60.39	64.28	96.58	101.88	106.63	112.33
90	61.75	65.65	69.13	73.29	107.57	113.15	118.14	124.12
100	70.06	74.22	77.93	82.36	118.50	124.34	129.56	135.81
120	86.92	91.57	95.70	100.62	140.23	146.57	152.21	158.95
140	104.03	109.14	113.66	119.03	161.83	168.61	174.65	181.84
160	121.35	126.87	131.76	137.55	183.31	190.52	196.92	204.53
180	138.82	144.74	149.97	156.15	204.70	212.30	219.04	227.06
200	156.43	162.73	168.28	174.84	226.02	233.99	241.06	249.45
240	191.99	198.98	205.14	212.39	268.47	277.14	284.80	293.89

自由度 ν のカイ二乗分布の上側確率 α に対する χ^2 の値を $\chi^2_\alpha(\nu)$ で表す。
例：自由度 $\nu = 20$ の上側 5% 点 $(\alpha = 0.05)$ は，$\chi^2_{0.05}(20) = 31.41$ である。
表にない自由度に対しては適宜補間すること。

付表 4. F 分布のパーセント点



$\alpha = 0.05$

$\nu_2 \backslash \nu_1$	1	2	3	4	5	6	7	8	9	10	15	20	40	60	120	∞
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735	4.619	4.558	4.464	4.431	4.398	4.365
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978	2.845	2.774	2.661	2.621	2.580	2.538
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544	2.403	2.328	2.204	2.160	2.114	2.066
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348	2.203	2.124	1.994	1.946	1.896	1.843
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337	2.282	2.236	2.089	2.007	1.872	1.822	1.768	1.711
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266	2.211	2.165	2.015	1.932	1.792	1.740	1.683	1.622
40	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180	2.124	2.077	1.924	1.839	1.693	1.637	1.577	1.509
60	4.001	3.150	2.758	2.525	2.368	2.254	2.167	2.097	2.040	1.993	1.836	1.748	1.594	1.534	1.467	1.389
120	3.920	3.072	2.680	2.447	2.290	2.175	2.087	2.016	1.959	1.910	1.750	1.659	1.495	1.429	1.352	1.254

$\alpha = 0.025$

$\nu_2 \backslash \nu_1$	1	2	3	4	5	6	7	8	9	10	15	20	40	60	120	∞
5	10.007	8.434	7.764	7.388	7.146	6.978	6.853	6.757	6.681	6.619	6.428	6.329	6.175	6.123	6.069	6.015
10	6.937	5.456	4.826	4.468	4.236	4.072	3.950	3.855	3.779	3.717	3.522	3.419	3.255	3.198	3.140	3.080
15	6.200	4.765	4.153	3.804	3.576	3.415	3.293	3.199	3.123	3.060	2.862	2.756	2.585	2.524	2.461	2.395
20	5.871	4.461	3.859	3.515	3.289	3.128	3.007	2.913	2.837	2.774	2.573	2.464	2.287	2.223	2.156	2.085
25	5.686	4.291	3.694	3.353	3.129	2.969	2.848	2.753	2.677	2.613	2.411	2.300	2.118	2.052	1.981	1.906
30	5.568	4.182	3.589	3.250	3.026	2.867	2.746	2.651	2.575	2.511	2.307	2.195	2.009	1.940	1.866	1.787
40	5.424	4.051	3.463	3.126	2.904	2.744	2.624	2.529	2.452	2.388	2.182	2.068	1.875	1.803	1.724	1.637
60	5.286	3.925	3.343	3.008	2.786	2.627	2.507	2.412	2.334	2.270	2.061	1.944	1.744	1.667	1.581	1.482
120	5.152	3.805	3.227	2.894	2.674	2.515	2.395	2.299	2.222	2.157	1.945	1.825	1.614	1.530	1.433	1.310

自由度 (ν_1, ν_2) の F 分布の上側確率 α に対する F の値を $F_\alpha(\nu_1, \nu_2)$ で表す。
例：自由度 $\nu_1 = 5, \nu_2 = 20$ の上側 5% 点 $(\alpha = 0.05)$ は，$F_{0.05}(5, 20) = 2.711$ である。
表にない自由度に対しては適宜補間すること。

付表 5. 指数関数と常用対数

指数関数			
x	e^x	x	e^x
0.01	1.0101	0.51	1.6653
0.02	1.0202	0.52	1.6820
0.03	1.0305	0.53	1.6989
0.04	1.0408	0.54	1.7160
0.05	1.0513	0.55	1.7333
0.06	1.0618	0.56	1.7507
0.07	1.0725	0.57	1.7683
0.08	1.0833	0.58	1.7860
0.09	1.0942	0.59	1.8040
0.10	1.1052	0.60	1.8221
0.11	1.1163	0.61	1.8404
0.12	1.1275	0.62	1.8589
0.13	1.1388	0.63	1.8776
0.14	1.1503	0.64	1.8965
0.15	1.1618	0.65	1.9155
0.16	1.1735	0.66	1.9348
0.17	1.1853	0.67	1.9542
0.18	1.1972	0.68	1.9739
0.19	1.2092	0.69	1.9937
0.20	1.2214	0.70	2.0138
0.21	1.2337	0.71	2.0340
0.22	1.2461	0.72	2.0544
0.23	1.2586	0.73	2.0751
0.24	1.2712	0.74	2.0959
0.25	1.2840	0.75	2.1170
0.26	1.2969	0.76	2.1383
0.27	1.3100	0.77	2.1598
0.28	1.3231	0.78	2.1815
0.29	1.3364	0.79	2.2034
0.30	1.3499	0.80	2.2255
0.31	1.3634	0.81	2.2479
0.32	1.3771	0.82	2.2705
0.33	1.3910	0.83	2.2933
0.34	1.4049	0.84	2.3164
0.35	1.4191	0.85	2.3396
0.36	1.4333	0.86	2.3632
0.37	1.4477	0.87	2.3869
0.38	1.4623	0.88	2.4109
0.39	1.4770	0.89	2.4351
0.40	1.4918	0.90	2.4596
0.41	1.5068	0.91	2.4843
0.42	1.5220	0.92	2.5093
0.43	1.5373	0.93	2.5345
0.44	1.5527	0.94	2.5600
0.45	1.5683	0.95	2.5857
0.46	1.5841	0.96	2.6117
0.47	1.6000	0.97	2.6379
0.48	1.6161	0.98	2.6645
0.49	1.6323	0.99	2.6912
0.50	1.6487	1.00	2.7183

常用対数			
x	$\log_{10} x$	x	$\log_{10} x$
0.1	-1.0000	5.1	0.7076
0.2	-0.6990	5.2	0.7160
0.3	-0.5229	5.3	0.7243
0.4	-0.3979	5.4	0.7324
0.5	-0.3010	5.5	0.7404
0.6	-0.2218	5.6	0.7482
0.7	-0.1549	5.7	0.7559
0.8	-0.0969	5.8	0.7634
0.9	-0.0458	5.9	0.7709
1.0	0.0000	6.0	0.7782
1.1	0.0414	6.1	0.7853
1.2	0.0792	6.2	0.7924
1.3	0.1139	6.3	0.7993
1.4	0.1461	6.4	0.8062
1.5	0.1761	6.5	0.8129
1.6	0.2041	6.6	0.8195
1.7	0.2304	6.7	0.8261
1.8	0.2553	6.8	0.8325
1.9	0.2788	6.9	0.8388
2.0	0.3010	7.0	0.8451
2.1	0.3222	7.1	0.8513
2.2	0.3424	7.2	0.8573
2.3	0.3617	7.3	0.8633
2.4	0.3802	7.4	0.8692
2.5	0.3979	7.5	0.8751
2.6	0.4150	7.6	0.8808
2.7	0.4314	7.7	0.8865
2.8	0.4472	7.8	0.8921
2.9	0.4624	7.9	0.8976
3.0	0.4771	8.0	0.9031
3.1	0.4914	8.1	0.9085
3.2	0.5051	8.2	0.9138
3.3	0.5185	8.3	0.9191
3.4	0.5315	8.4	0.9243
3.5	0.5441	8.5	0.9294
3.6	0.5563	8.6	0.9345
3.7	0.5682	8.7	0.9395
3.8	0.5798	8.8	0.9445
3.9	0.5911	8.9	0.9494
4.0	0.6021	9.0	0.9542
4.1	0.6128	9.1	0.9590
4.2	0.6232	9.2	0.9638
4.3	0.6335	9.3	0.9685
4.4	0.6435	9.4	0.9731
4.5	0.6532	9.5	0.9777
4.6	0.6628	9.6	0.9823
4.7	0.6721	9.7	0.9868
4.8	0.6812	9.8	0.9912
4.9	0.6902	9.9	0.9956
5.0	0.6990	10.0	1.0000

注: 常用対数を自然対数に直すには 2.3026 をかければよい。

【解答用紙記入例】

- マークシートの記入例

（例） **10** と表示のある問に対して ③ と解答する場合

次のように解答番号 **10** の解答の ③ にマークすること。

解答番号	解 答
10	① ② ● ④ ⑤

- 論述問題解答面のページ先頭の記入例

（例） 論述問題 **問 1** を解答する場合

次のように問題番号に解答する問題番号を記入し，得点の欄には何も書かないこと。

統計検定 準１級 論述問題 解答面

問題番号		得点３
1	※選択した問題番号を記入すること ※２問以上解答した場合は採点対象としない ※以下は自由に使ってよい（裏面も可）	

一般財団法人　統計質保証推進協会

統計検定センター

〒101-0051 東京都千代田区神田神保町３丁目６番

URL **http://www.toukei-kentei.jp**

2019.6