統計検定
Japan Statistical Society Certificate

1級
統計応用

2018 年 11 月 25 日

【注意事項】

1 試験開始の合図があるまで，この問題冊子の中を見てはいけません。

2 この問題冊子は，60 ページあります。**3〜13** ページは人文科学，**15〜25** ページは社会科学，**27〜37** ページは理工学，**39〜49** ページは医薬生物学の問題です。

3 選択した分野の問題 5 問から 3 問を選択して，解答しなさい。

4 試験時間は 90 分です。

5 試験中に問題冊子の印刷不鮮明，ページの落丁・乱丁および解答冊子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 解答冊子・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 解答冊子表紙の左上に受験番号を記入しなさい。

⑤ 問ごと（問 **1**，問 **2**，...）に解答のページを改めなさい。

⑥ 各ページの先頭に受験番号と問題番号を書きなさい。（この冊子裏面の記入例参照）
解答冊子には，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 解答冊子の表紙の選択分野欄の選択した分野と，問題番号欄の選択した問題番号をそれぞれ ◯ で囲みなさい（得点欄には何も書かないこと）。（この冊子裏面の記入例参照）

7 51 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。

8 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。

9 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）

人文科学

問 1　以下の各問に答えよ。

〔1〕　赤球が 6 個，白球が 4 個入っている箱から非復元無作為抽出により n 個の球を取り出す ($1 \leq n \leq 10$)。n 個の中に赤球が x 個ある確率を $f_n(x)$ とするとき，x の値の取り得る範囲，および確率 $f_n(x)$ の式を示せ。また，具体的に $n = 5$ の $f_5(1)$ および $f_5(2)$ を求めよ。

〔2〕　ある社会問題について教育した 6 人の群 A と教育しなかった 4 人の群 B がある。これら 10 人に対して，この社会問題に関する意識調査を点数で聞いたところ，次のような点数になった。

A: 5，6，8，10，12，15

B: 4，6，7，14

この結果を全体の中央値 7.5 より小さい値の観測度数と，大きい値の観測度数に分け次のように 2×2 分割表を作成した。この表について各問に答えよ。

群	中央値 7.5 より小さい	中央値 7.5 より大きい
A	2	4
B	3	1

(1)　オッズ比およびファイ係数を求めよ。これらから群 A と群 B の関連性について述べよ。

(2)　両群の中央値は同じであるという帰無仮説の下での期待度数の分割表を示せ。

(3)　次の帰無仮説および対立仮説について有意水準 10 %の仮説検定を行う。

帰無仮説: 両群の中央値は同じである。

対立仮説: 両群の中央値は異なる。

検定統計量として，ピアソン・カイ二乗統計量

$$Y = \sum \frac{(\text{観測度数} - \text{期待度数})^2}{\text{期待度数}}$$

を用いたとき，およびイェーツの補正を施したカイ二乗統計量 Y' を用いたときの検定統計量の値と仮説検定の結果をそれぞれ示せ。

(4)　上問 (3) の仮説検定を，超幾何分布を用いたフィッシャー検定（直接確率計算法）にて実行し，そのときの P 値および仮説検定の結果を示せ。

問 2　ある高校では，最終学年の生徒に 2 種類のテスト（テスト 1，テスト 2）を行い，テストの成績 $\boldsymbol{x} = (x_1, x_2)$ から，進路 A と進路 B のどちらが好ましいかをアドバイスする第 1 段階目の進路指導を行っている。ここで，進路 A と進路 B のそれぞれにおける 2 種類のテストの点数は，母分散共分散行列が等しい 2 変量正規分布に従うとする。

表 1 は，最終的に進路 A または進路 B に進学した過去の生徒 100 名（各進路 50 名）の成績の標本平均であり，表 2 は，群内平方和をプールした標本分散共分散行列である。また，表 3 は，今年の 3 人の生徒 a，b，c のテストの点数である。

表 1: 標本平均

	テスト 1	テスト 2
進路 A	30	50
進路 B	40	50

表 2: 群内平方和をプールした標本分散共分散行列

	テスト 1	テスト 2
テスト 1	20	14
テスト 2	14	20

表 3: 3 人の点数

生徒	テスト 1	テスト 2
a	30	60
b	40	60
c	50	50

〔1〕　表 2 から，テスト 1 とテスト 2 の相関係数を求めよ。

〔2〕　生徒の成績 $\boldsymbol{x} = (x_1, x_2)$ に対し，各進路のテスト 1 およびテスト 2 の標本平均からのユークリッド 2 乗距離を

進路 A: $U_A^2 = (x_1 - 30)^2 + (x_2 - 50)^2$

進路 B: $U_B^2 = (x_1 - 40)^2 + (x_2 - 50)^2$

とし，距離の値が小さい方にその生徒の進路をアドバイスする。このとき，$U_A^2 = U_B^2$ となる直線を判別境界線 U と呼ぶ。この境界線 U の方程式を x_1，x_2 を用いて求めよ。また，距離 U_A^2，U_B^2 を用いるとき，生徒 a，b，c はそれぞれどちらの進路とアドバイスされるか示し，その理由を述べよ。

〔3〕　上問〔2〕の距離はベクトルを用いて次のように書くことができる。

$$進路 A: U_A^2 = \begin{pmatrix} x_1 - 30, & x_2 - 50 \end{pmatrix} \begin{pmatrix} x_1 - 30 \\ x_2 - 50 \end{pmatrix}$$

$$進路 B: U_B^2 = \begin{pmatrix} x_1 - 40, & x_2 - 50 \end{pmatrix} \begin{pmatrix} x_1 - 40 \\ x_2 - 50 \end{pmatrix}$$

それに対し，群内平方和をプールした標本分散共分散行列 S の逆行列 S^{-1} を用いた距離（マハラノビスの 2 乗距離）を

$$進路 A: M_A^2 = \begin{pmatrix} x_1 - 30, & x_2 - 50 \end{pmatrix} S^{-1} \begin{pmatrix} x_1 - 30 \\ x_2 - 50 \end{pmatrix}$$

$$進路 B: M_B^2 = \begin{pmatrix} x_1 - 40, & x_2 - 50 \end{pmatrix} S^{-1} \begin{pmatrix} x_1 - 40 \\ x_2 - 50 \end{pmatrix}$$

と定義し，この距離の値が小さい方にその生徒の進路をアドバイスする。$M_A^2 = M_B^2$ となる直線を判別境界線 M と呼ぶ。境界線 M の方程式を x_1，x_2 を用いて求めると

$$x_2 = \frac{10}{7} x_1$$

が得られる。境界線 U と境界線 M の違いについて述べよ（イメージを図示してもよい)。また，距離 M_A^2，M_B^2 を用いるとき，生徒 a，b，c はそれぞれどちらの進路とアドバイスされるかを示し，その理由を述べよ。

〔4〕　判別分析には誤判別が生じる。R_A と R_B を母平均ベクトルと母分散共分散行列から計算されるマハラノビスの 2 乗距離 $(M_A^*)^2$，$(M_B^*)^2$ を用いて進路 A および進路 B にアドバイスする領域であるとする。このとき，R_A と R_B は誤判別率最小という意味で最適な判別ルールを与え，そのときの誤判別率は，各進路でのテストの点数の確率密度関数をそれぞれ $f_A(\boldsymbol{x})$，$f_B(\boldsymbol{x})$ としたとき，

$$T(R_A, R_B) = \frac{1}{2} \int_{R_B} f_A(\boldsymbol{x}) d\boldsymbol{x} + \frac{1}{2} \int_{R_A} f_B(\boldsymbol{x}) d\boldsymbol{x}$$

と表せる。ただし，各進路には同数が進学するとする。実際は分布の母数を標本より推定して，それぞれの領域 $\hat{R}_A$，$\hat{R}_B$ を推定するので，実際の誤判別率は

$$T(\hat{R}_A, \hat{R}_B) = \frac{1}{2} \int_{\hat{R}_B} f_A(\boldsymbol{x}) d\boldsymbol{x} + \frac{1}{2} \int_{\hat{R}_A} f_B(\boldsymbol{x}) d\boldsymbol{x}$$

となる。実際の誤判別率は最適誤判別率より大きくなることをその理由とともに示せ。

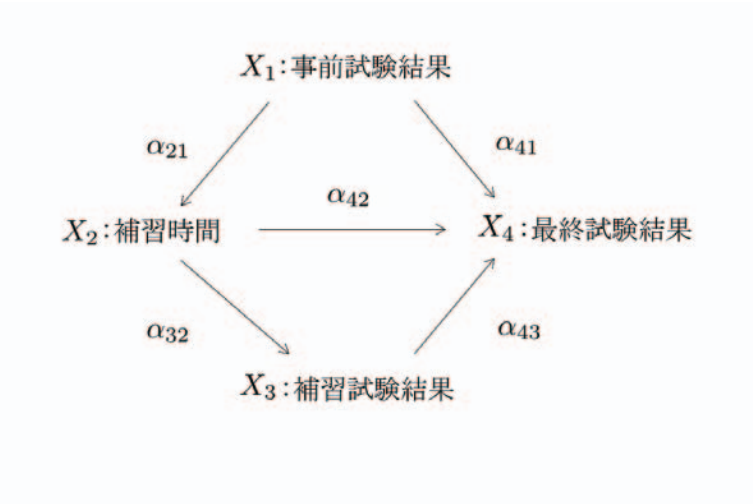
問 3　ある大学では補習時間を設けて学習効果を上げるよう工夫をしている。はじめに事前試験を行い，試験結果に応じて補習時間を決めてよいと学生に指導している。補習が終わった後に簡単な試験（補習試験）を行い，その後，最終試験を受ける。それぞれの量的変数を次のように置く。

X_1: 事前試験結果，X_2: 補習時間，X_3: 補習試験結果，X_4: 最終試験結果

これらの変数間に次のような母相関行列 $R = \{\rho_{ij}\}, i, j = 1, 2, 3, 4$ を仮定する。

	X_1	X_2	X_3	X_4
X_1	1.00	-0.40	-0.20	0.34
X_2	-0.40	1.00	0.50	0.20
X_3	-0.20	0.50	1.00	0.40
X_4	0.34	0.20	0.40	1.00

また，X_1，X_2，X_3，X_4 について次のようなパス図の因果モデルを仮定する。ここで α_{21}，α_{32}，α_{41}，α_{42}，α_{43} は矢線に対応するパス係数である。ただし，これらのパス係数はすべての変数を平均 0，分散 1 に標準化したうえで得たものである。



〔1〕　このパス図には X_1 から X_3 への矢線がない。このことの意味を述べよ。

〔2〕　このパス図には擬相関が生じている箇所がある。それらすべてを示せ。

〔3〕　この因果モデルに対する線形構造方程式を誤差項を付記して示せ。

〔4〕　上問〔3〕で求めた構造方程式から，母相関係数 ρ_{12}，ρ_{13}，ρ_{23}，ρ_{24} を α_{21}，α_{32}，α_{41}，α_{42}，α_{43} を用いて示せ。また，α_{21}，α_{32} の値を具体的に求めよ。

〔5〕　上問〔4〕の結果から X_2 から X_4 への総合効果を求めよ。ただし，$\alpha_{41} = 0.5$ であることを用いてよい。

問 4　都道府県別の酒類販売（消費）数量表を基に，各都道府県を特徴あるグループでまとめた上で解析したい。以下の各問に答えよ。

〔1〕　表 1 は，国税庁による平成 27 年度成人 1 人当たりの酒類販売（消費）数量表（都道府県別）の一部である（販売単位:リットル L）。すべての酒類合計の 1 人当たりの消費量の上位 6 都府県のうち，消費上位であるいくつかの酒類について掲載している。表 2 は，表 1 のデータを基に都府県間の距離をユークリッド距離として求めたものである。図 1 は，あるソフトウェアを用いて，表 2 の距離を元に，最長距離法とウォード法によりデンドログラムを描いたものである（都府県×距離）。

表 1: 平成 27 年度成人 1 人当たりの酒類販売（消費）数量表（都道府県別）より一部抜粋（単位: L）(国税庁)

	東京	高知	大阪	新潟	青森	宮崎
ビール・発泡酒	51.0	45.5	40.7	35.9	34.3	33.7
日本酒	6.7	6.3	5.0	13.3	7.0	2.5
焼酎	9.8	8.2	6.6	7.8	10.7	20.3
ワイン	9.8	1.8	3.8	2.5	2.5	2.5
ウイスキー	2.4	0.9	1.4	1.4	1.7	0.8

表 2: 表 1 より算出した各都府県間のユークリッド距離

距離	東京	高知	大阪	新潟	青森	宮崎
東京	0.00	9.96	12.50	18.16	18.26	21.98
高知	9.96	0.00	5.62	11.92	11.55	17.34
大阪	12.50	5.62	0.00	9.75	7.97	15.65
新潟	18.16	11.92	9.75	0.00	7.12	16.68
青森	18.26	11.55	7.97	7.12	0.00	10.66
宮崎	21.98	17.34	15.65	16.68	10.66	0.00

(1)　表 2 を基に最短距離法を用いてクラスターを作成し，図 1 のようなデンドログラムを描け。

(2)　表 1 および上問 (1) で作成したデンドログラムを参考に，表 1 の 6 都府県ではいくつのクラスターを想定したらよいかを示し，クラスター数を記した上で，クラスターに入る都府県名と各クラスターの特徴を簡潔に記せ。

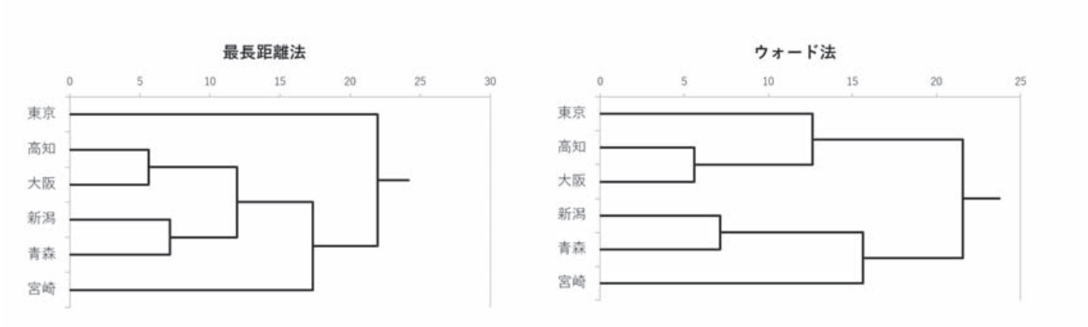


図 1: 最長距離法及びウォード法によるデンドログラム

〔2〕　沖縄を除く 46 都道府県のデータについて，上問〔1〕とは別のソフトウェアを用い，最短距離法とウォード法により算出した結果のデンドログラムは図 2 のようになった。ただし，都道府県名は除いてある。図 2 では，最短距離法に鎖効果が見られる。

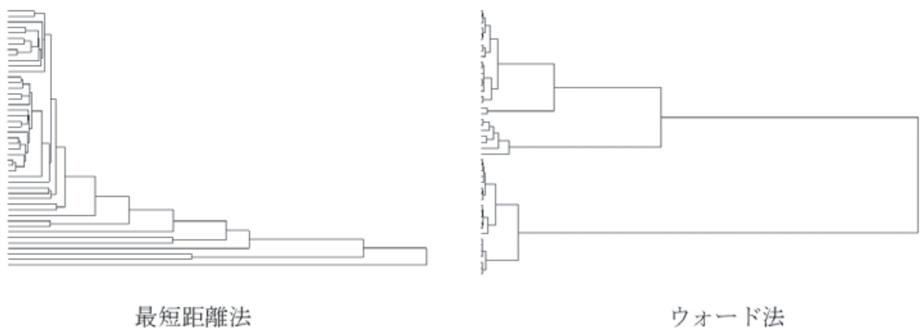


図 2: 最短距離法及びウォード法による 46 都道府県のデンドログラム

(1)　鎖効果とは何かについて 46 都道府県のデンドログラムを利用して説明せよ。また，なぜ最短距離法では鎖効果が生じやすいのかについて述べよ。

(2)　ウォード法のクラスター基準方法を簡潔に記せ。

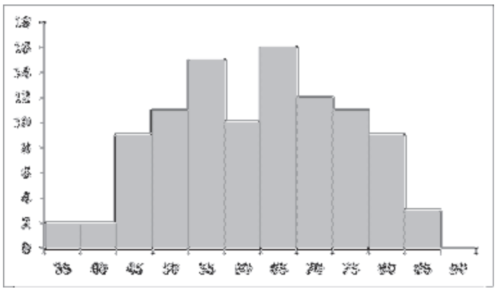
(3)　46 都道府県のデータについて，最短距離法とウォード法のいずれを用いる方が解釈しやすいかについて述べよ。

問 5　ある大学のある学部では，入試科目の数学が選択制となっている。その学部の M 教授は，統計学のクラスを A，B の 2 つ受け持っていて，両クラスとも 100 名ずつの受講者がいる。両クラスとも入試で数学を選択した学生と非選択の学生は 50 名ずつであった。両クラスで統計学の試験を行ったところ入試での数学の選択別の平均と標準偏差は表 1 のようであった。ここで，表の標準偏差は，両クラスとも除数を 50 とした標本分散の正の平方根である。また，それぞれのクラスでヒストグラムを描いたところ図 1 のようであった。

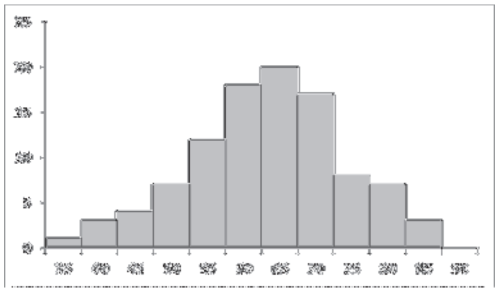
両クラスとも入試での数学の選択の有無で明らかに平均が異なるので，ヒストグラムはふた山型となると予想されたが，A クラスではそうであるものの，B クラスではひと山型であった。M 教授はどのような場合に，混合分布がふた山型を示すのかに興味を持ち，理論的に考察することにした。

表 1: 試験結果の基本統計量

	A クラス		B クラス	
統計量	数学選択	数学非選択	数学選択	数学非選択
平均	69.7	49.6	67.6	53.8
標準偏差	6.8	7.8	7.7	8.7



A クラス



B クラス

図 1: クラス全体の試験結果のヒストグラム

母集団全体の分布 F は，分散は等しいが平均は異なる正規分布 $N(\mu_1, \sigma^2)$ と $N(\mu_2, \sigma^2)$ の混合率 1/2 ずつの混合であるとする。すなわち，$N(\mu_j, \sigma^2)$ の確率密度関数を $f_j(x) = \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left[-\dfrac{(x-\mu_j)^2}{2\sigma^2}\right]$ としたとき $(j = 1, 2)$，分布 F の確率密度関数 $f(x)$ は

$$f(x) = \frac{1}{2}\{f_1(x) + f_2(x)\}$$

で与えられる。図 2 はいくつかの μ_1，μ_2，σ の組み合わせに対応した $f(x)$ の形状である。

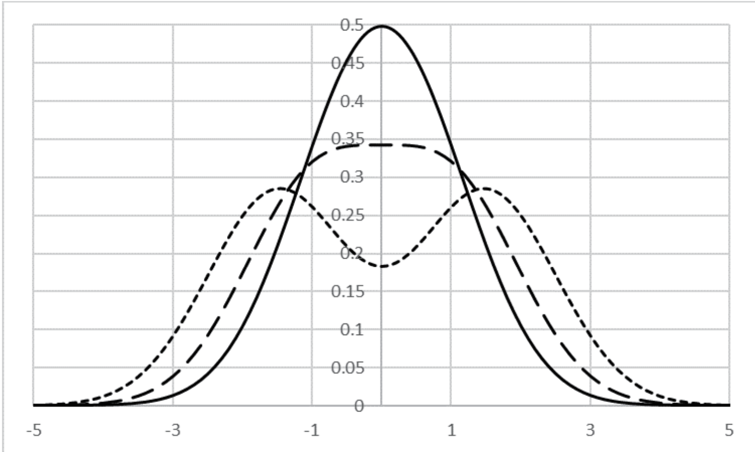


図 2: いくつかのパラメータに関する $f(x)$ の形状

分布 F に従う確率変数を X とするとき，以下の各問に答えよ。

〔1〕　X の期待値と分散はそれぞれ

$$\xi = E[X] = \frac{\mu_1 + \mu_2}{2},\ \tau^2 = V[X] = \sigma^2 + \left(\frac{\mu_1 - \mu_2}{2}\right)^2$$

となることを示せ。

ヒント: $(x - \xi)^2 = \left(x - \dfrac{\mu_1 + \mu_2}{2}\right)^2 = \left(x - \mu_1 + \dfrac{\mu_1 - \mu_2}{2}\right)^2 = \left(x - \mu_2 - \dfrac{\mu_1 - \mu_2}{2}\right)^2$

である。

〔2〕　表 1 の結果から，A クラス全体のデータを $x_i (i = 1, \ldots, 100)$ としたときの平均 $\bar{x} = \dfrac{1}{100}\displaystyle\sum_{i=1}^{100} x_i$ と標準偏差 $s = \sqrt{\dfrac{1}{100}\displaystyle\sum_{i=1}^{100}(x_i - \bar{x})^2}$ を求めよ。

〔3〕　確率密度関数 $f(x)$ の 1 次導関数 $f'(x)$ および 2 次導関数 $f''(x)$ を求めよ。また，$x = \xi$ は $f(x)$ の極値を与えることを示せ。

〔4〕　分布 F がふた山型（二峰性）を示すための μ_1，μ_2，σ の条件を求めよ。

社会科学

問 1　次の表 1 と表 2 は，ある調査における女性の職種と不払い労働時間の有無に関するクロス集計表である。表 1 は女性全体の表であり，表 2 は女性の 50 歳代に限定した表である。以下の各問に答えよ。

表 1: 女性の職種と不払い労働時間の有無別人数

女性		不払い労働時間の有無		
		ある	ない	
職業	専門職	75	58	133
	製造・生産関連	6	19	25
		81	77	158

表 2: 50 歳代の女性の職種と不払い労働時間の有無別人数

女性，50歳代		不払い労働時間の有無		
		ある	ない	
職業	専門職	5	5	10
	製造・生産関連	0	5	5
		5	10	15

注: 労働政策研究・研修機構「働き方の現状と意識に関するアンケート調査結果（2005 年 8 月調査実施）」に基づき，職種は専門職と製造・生産関連に限定して，職種と不払い労働時間に関する無回答を除いて作成した。

〔1〕　表 1 に基づき，職種と不払い労働時間の有無との間に関連性があるか否かを調べるために仮説検定を行う。クロス集計表の第 (i, j) セル確率を p_{ij}，各周辺確率を p_{i+}，p_{+j} としたとき，職種と不払い労働時間の有無との間に関連性がないという帰無仮説を示せ。また，その帰無仮説の下で，表 1 の第 $(1, 1)$ セル（職種: 専門職，不払い労働時間の有無: ある）の期待度数 e_{11} を求めよ。

〔2〕　表 1 に基づき，イェーツの補正を施したカイ二乗分布を用いた独立性の仮説検定を有意水準 0.05 で実行せよ。その際，検定統計量の値や棄却域を明確に示し，検定結果を説明せよ。

〔3〕　表 2 に基づき，フィッシャー検定（直接確率計算法）を用いた独立性の仮説検定を有意水準 0.05 で実行し，検定法の詳細と検定結果を説明せよ。

〔4〕　カイ二乗分布を用いた検定ではなく，上問〔3〕のようなフィッシャー検定を行うのが望ましい状況について簡潔に説明せよ。

問 2　表 1 は，ある地方の小売業の事業所を対象にした過去の全数調査の結果を層別に表したものである。記号は，母集団全体における大きさ（総事業所数）を N とし，層 i での部分母集団の大きさを N_i，母平均を μ_i，母標準偏差を σ_i とする $(i = 1, 2, 3)$。全体の母平均を $\mu = \sum_{i=1}^{3} \frac{N_i}{N}\mu_i$ とする。また，調査における全体での標本の大きさを n，標本平均を $\bar{x}$ とし，層 i での標本の大きさを n_i，標本平均を $\bar{x}_i$ とする。$N_1 + N_2 + N_3 = N$，$n_1 + n_2 + n_3 = n$ である。表 1 の調査結果の情報に基づいて，現在の μ_i を推定するための標本調査を企画する。N，N_i，σ_i は過去の調査と同じとし，以下の各問に答えよ。

表 1: 従業者規模別事業所数，年間商品販売額の平均値と標準偏差

層 i	従業者規模	事業所数 N_i	年間商品販売額 平均値（百万円） μ_i	年間商品販売額 標準偏差（百万円） σ_i
層 1	1〜9 人	4,000	100	10
層 2	10〜99 人	900	1,000	100
層 3	100 人以上	100	25,000	4,000
計		5,000		

〔1〕　層 1，層 2，層 3 における年間商品販売額の層内での散らばりの大きさを比較したい。経済学では，散らばりの大きさの指標として層内の変動係数が用いられることが多い。変動係数が用いられる理由を述べ，表 1 の各層での散らばりの大小について考察せよ。

〔2〕　層 1 において，信頼係数 0.95 で許容誤差は d と設定して母平均 μ_1 を推定する場合，必要な標本の大きさの最小値 n_1^* は，標準正規分布の上側 100α ％点を z_α として

$$n_1^* \geq \left(\frac{z_{0.025}}{d}\right)^2 \sigma_1^2$$

により与えられることを示し，$d = 5$ の場合の n_1^* を求めよ。なおここでは，有限母集団修正は無視している。

〔3〕　層別抽出法を用いた標本抽出を行う。層における標本抽出は，他の層の標本抽出とは無関係であるとする。第 i 層の標本平均 $\bar{x}_i$ の分散を $V[\bar{x}_i] = \frac{N_i - n_i}{N_i - 1} \cdot \frac{\sigma_i^2}{n_i}$ として，μ の推定量 $\hat{\mu} = \sum_{i=1}^{3} \frac{N_i}{N}\bar{x}_i$ の分散 $V[\hat{\mu}]$ を N，N_i，n_i，σ_i^2 を用いて表せ。

〔4〕　$n_1+n_2+n_3=n$ の制約条件の下で，$V[\hat{\mu}]$ を最小にする $n_i^{\#}(i=1,2,3)$ を導出せよ。なおその際，$N_i-1 \approx N_i$ として求めよ。

〔5〕　調査費用に制限があり，標本調査における標本の大きさ n は 120 と決定された。上問〔3〕で求めた割当てにしたがって，層 1，層 2，層 3 に割当てる標本の大きさを求めよ。

問 3　日別の株価分析を考える。ある株式の日付 t における終値を p_t としたとき，$t-1$ 日から t 日にかけての収益率を $r_t = 100(\log p_t - \log p_{t-1})$ とし，この r_t に，m_t をトレンド，ε_t を平均 0，分散一定の撹乱項とした時系列モデル

$$r_t = m_t + \varepsilon_t \tag{1}$$

を当てはめた。図 1 はある銘柄のデータへのモデル (1) の当てはめでの残差の時系列プロット，図 2 は残差の QQ プロットで直線は正規分布のときに想定されるものである。

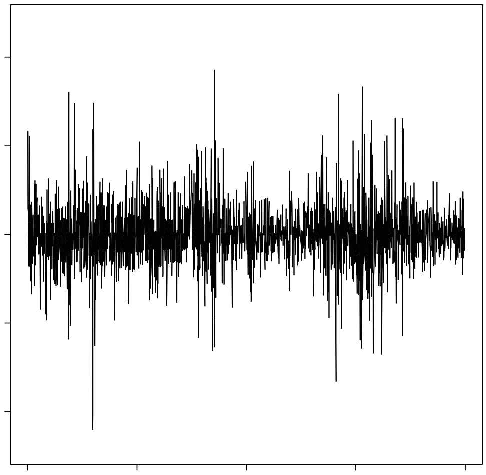


図 1: 残差の時系列プロット

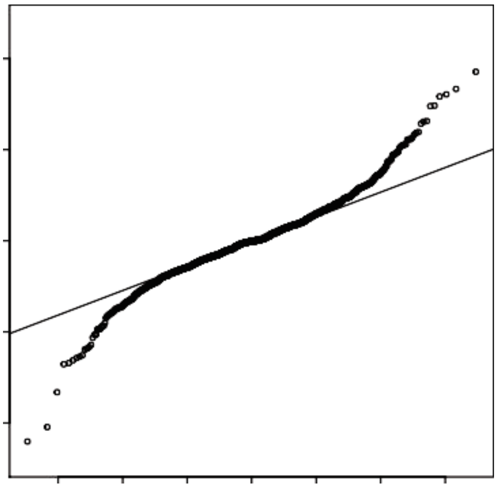


図 2: 残差の QQ プロット

〔1〕　図 1 および図 2 から，撹乱項の分布は正規分布と比較してどのようになっているのかを述べよ。

撹乱項に自己回帰条件付き分散変動型モデル（ARCH 型モデル）を当てはめる。ARCH(1) モデルは，(1) の ε_t を以下のように想定する。

$$\varepsilon_t = \sqrt{h_t}\eta_t,\ \eta_t \overset{iid}{\sim} N(0,1) \tag{2}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \tag{3}$$

ここで，(1) の r_t と (3) の h_t は弱定常性を満たすものと仮定する。また，$\alpha_0 > 0$，$\alpha_1 > 0$ とし，η_t はトレンドや過去の撹乱項とは独立とする。このとき，以下の各問に答えよ。なお解答では，無条件の期待値と，t 期までの情報を所与とした条件付き期待値との区別を明確にすること。

〔2〕　$t-1$期までの情報を所与としたとき，ε_tの条件付き分散はh_tとなることを示せ。

〔3〕　ε_tの無条件分散をα_0とα_1を用いて表せ。

〔4〕　$E[\eta_t^4]=3$となることを利用して，ε_tの無条件4次モーメントをα_0とα_1を用いて表せ。ただし$E[\varepsilon_t^4]=E[\varepsilon_{t-1}^4]$が成り立つものとする。

〔5〕　ε_tの尖度は正になることを示せ。なお，尖度は$\beta_2=\dfrac{E[\varepsilon_t^4]}{(E[\varepsilon_t^2])^2}-3$で定義される。

問 4　世帯所得などの正の値のみを取るデータの分布のモデルとして対数正規分布がある。正の値を取る確率変数 X を自然対数で変換した $\log X$ が $N(\mu, \sigma^2)$ に従うとき，X は対数正規分布に従うといい，$X \sim LN(\mu, \sigma^2)$ と書く。$LN(\mu, \sigma^2)$ の確率密度関数は，$x > 0$ として

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \frac{1}{x} \exp\left[-\frac{(\log x - \mu)^2}{2\sigma^2}\right]$$

である。以下の各問に答えよ。

〔1〕　$X \sim LN(\mu, \sigma^2)$ のとき，X の期待値 $E[X]$，中央値 (median)，最頻値 (mode) をそれぞれ求め，それらの大小関係を示せ。

〔2〕　$LN(\mu, \sigma^2)$ からの n 個の互いに独立な観測値を $x_1, \ldots, x_n$ としたとき，μ の最尤推定量を導出せよ。

〔3〕　ベイズ流の推測において，σ^2 は既知とし μ の事前分布を $N(\nu, \tau^2)$ とする。$LN(\mu, \sigma^2)$ からの n 個の独立な観測値を $x_1, \ldots, x_n$ としたときの μ の事後分布を導出せよ。

　ある会社の非正規職員の年間所得を無作為に 10 人調査したところ，表 1 の所得データ (x) を得た。表 1 には所得データに加え，その自然対数 $(\log x)$ の値，および x と $\log x$ の平均も示している。以下の各問に答えよ。なお，$\sigma^2 = (0.25)^2$ は既知とする。

表 1: 10 名の所得データ（単位は千円）

ID	1	2	3	4	5	6	7	8	9	10	平均
所得 (x)	1635	1772	1781	1899	2217	2241	2460	2745	3000	3525	2327.5
$\log x$	7.40	7.48	7.48	7.55	7.70	7.71	7.81	7.92	8.01	8.17	7.72

〔4〕　上問〔2〕で導出した μ の最尤推定量 $\hat{\mu}$ を元に，上問〔1〕の結果を用いて，表 1 のデータから X の期待値 $E[X]$，中央値 (median)，最頻値 (mode) の推定値をそれぞれ求めよ。巻末付表 5 も参照のこと。

〔5〕　ベイズ流の推測において，μ の事前分布を $N(8, (0.5)^2)$ とする。上問〔3〕の結果を用いて，表 1 のデータから μ の事後期待値を求めよ。

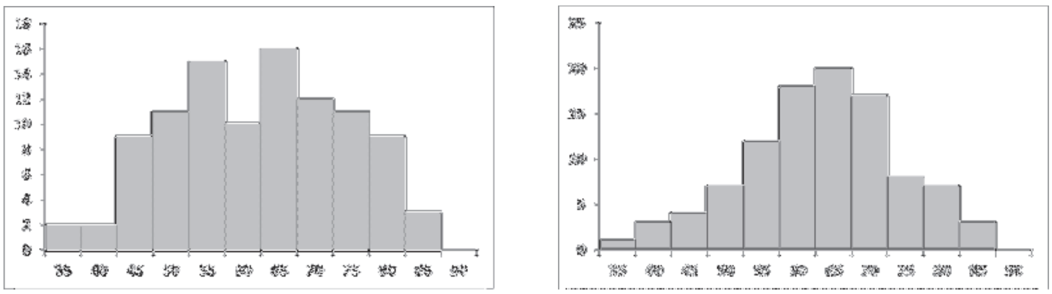
〔6〕　X の期待値 $E[X]$ の推定値としては $x_1, \ldots, x_n$ の標本平均 $\bar{x} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i$ が考えられる。上問〔4〕で求めた期待値 $E[X]$ の推定値と標本平均 $\bar{x}$ をそれぞれ推定量（確率変数）と見たときの，それらの推定量の統計的な特徴について述べよ。

問 5　ある大学のある学部では，入試科目の数学が選択制となっている。その学部のM教授は，統計学のクラスをA，Bの2つ受け持っていて，両クラスとも100名ずつの受講者がいる。両クラスとも入試で数学を選択した学生と非選択の学生は50名ずつであった。両クラスで統計学の試験を行ったところ入試での数学の選択別の平均と標準偏差は表1のようであった。ここで，表の標準偏差は，両クラスとも除数を50とした標本分散の正の平方根である。また，それぞれのクラスでヒストグラムを描いたところ図1のようであった。

両クラスとも入試での数学の選択の有無で明らかに平均が異なるので，ヒストグラムはふた山型となると予想されたが，Aクラスではそうであるものの，Bクラスではひと山型であった。M教授はどのような場合に，混合分布がふた山型を示すのかに興味を持ち，理論的に考察することにした。

表 1: 試験結果の基本統計量

	Aクラス		Bクラス	
統計量	数学選択	数学非選択	数学選択	数学非選択
平均	69.7	49.6	67.6	53.8
標準偏差	6.8	7.8	7.7	8.7



Aクラス　　Bクラス

図 1: クラス全体の試験結果のヒストグラム

母集団全体の分布 F は，分散は等しいが平均は異なる正規分布 $N(\mu_1, \sigma^2)$ と $N(\mu_2, \sigma^2)$ の混合率 $1/2$ ずつの混合であるとする。すなわち，$N(\mu_j, \sigma^2)$ の確率密度関数を $f_j(x) = \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left[-\dfrac{(x-\mu_j)^2}{2\sigma^2}\right]$ としたとき $(j = 1, 2)$，分布 F の確率密度関数 $f(x)$ は

$$f(x) = \frac{1}{2}\{f_1(x) + f_2(x)\}$$

で与えられる。図2はいくつかの μ_1，μ_2，σ の組み合わせに対応した $f(x)$ の形状である。

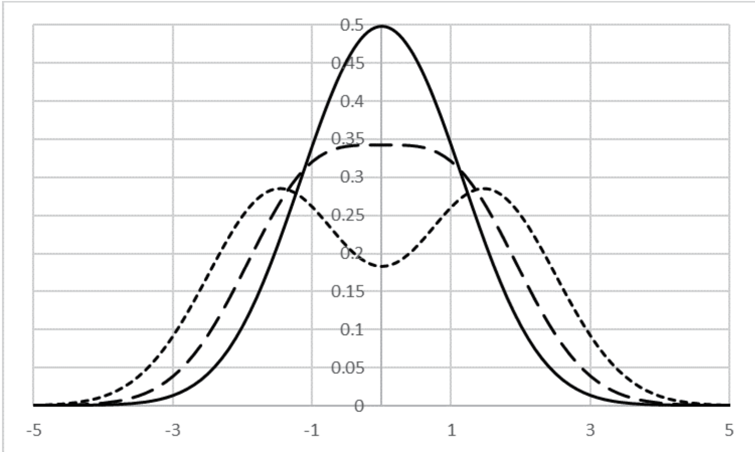


図 2: いくつかのパラメータに関する $f(x)$ の形状

分布 F に従う確率変数を X とするとき，以下の各問に答えよ。

〔1〕　X の期待値と分散はそれぞれ

$$\xi = E[X] = \frac{\mu_1+\mu_2}{2},\ \tau^2 = V[X] = \sigma^2 + \left(\frac{\mu_1-\mu_2}{2}\right)^2$$

となることを示せ。

ヒント: $(x-\xi)^2 = \left(x - \frac{\mu_1+\mu_2}{2}\right)^2 = \left(x-\mu_1+\frac{\mu_1-\mu_2}{2}\right)^2 = \left(x-\mu_2-\frac{\mu_1-\mu_2}{2}\right)^2$

である。

〔2〕　表 1 の結果から，A クラス全体のデータを $x_i(i=1,\ldots,100)$ としたときの平均 $\bar{x} = \frac{1}{100}\sum_{i=1}^{100} x_i$ と標準偏差 $s = \sqrt{\frac{1}{100}\sum_{i=1}^{100}(x_i-\bar{x})^2}$ を求めよ。

〔3〕　確率密度関数 $f(x)$ の 1 次導関数 $f'(x)$ および 2 次導関数 $f''(x)$ を求めよ。また，$x=\xi$ は $f(x)$ の極値を与えることを示せ。

〔4〕　分布 F がふた山型（二峰性）を示すための μ_1，μ_2，σ の条件を求めよ。

理工学

問 1　ある工業製品の生産ラインでは稀に不良品が生じ，相続く 2 つの不良品の生じる時間間隔 X はパラメータ λ の指数分布に従うとする。X の確率密度関数は

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

である。以下の各問に答えよ。なお，対数は自然対数である。

〔1〕　パラメータ λ の指数分布の累積分布関数 $F(x)$ およびモーメント母関数 $M_X(\theta) = E[e^{\theta X}]$ を求めよ。

〔2〕　生産ラインの稼働開始から n 個の不良品が生じるまでの時間を $W_n = X_1 + \cdots + X_n$ とする。$X_1, \ldots, X_n$ が互いに独立にパラメータ λ の指数分布に従うとき，W_n の確率密度関数は

$$g_n(w) = \begin{cases} \dfrac{\lambda^n w^{n-1} e^{-\lambda w}}{(n-1)!} & (w \geq 0) \\ 0 & (w < 0) \end{cases}$$

となることを示せ。また，この分布のモーメント母関数 $M_W(\theta) = E[e^{\theta W}]$ を求めよ。

〔3〕　U を区間 $(0,1)$ 上の一様分布に従う確率変数としたとき，$X = -\lambda^{-1} \log U$ はパラメータ λ の指数分布に従うことを示せ。

〔4〕　不良品の生じる確率を p とする。良品・不良品の生起が独立であるとき，生産ラインの稼働開始から初めて不良品が生じるまでの良品の個数 Y はパラメータ p の幾何分布に従い，その確率関数は

$$g(y) = P(Y = y) = p(1-p)^y \quad (y = 0, 1, 2, \ldots)$$

である。X をパラメータ $\lambda = -\log(1-p)$ の指数分布に従う確率変数としたとき，その整数部分，すなわち X を超えない最大の整数を Y とすると，Y はパラメータ p の幾何分布に従うことを示せ。

〔5〕　互いに独立に区間 $(0,1)$ 上の一様分布に従う確率変数の列を $U_1, U_2, \ldots$ としたとき，それらを順にかけ合わせて初めて $e^{-\lambda}$ より小さくなったときの一つ手前の個数を M とする。すなわち，

$$U_1 \times \cdots \times U_m > e^{-\lambda} > U_1 \times \cdots \times U_m \times U_{m+1} \text{ ならば } M = m$$

である。このとき，M はパラメータ λ のポアソン分布に従うことを示せ。

ただし，パラメータ λ のポアソン分布の確率関数は

$$p(m) = \frac{\lambda^m}{m!} e^{-\lambda} \quad (m = 0, 1, 2, \ldots)$$

であり，ある事象の生じる時間間隔が互いに独立にパラメータ λ の指数分布に従うとき，時刻 0 から T までの間に生じた当該事象の生起回数 N はパラメータ λT のポアソン分布に従うことを証明なしに用いてもよい。

問 2　形状パラメータ $m>0$ と尺度パラメータ $\eta>0$ を持つワイブル分布 $W(m,\eta)$ の累積分布関数 $F(x)$ は

$$F(x)=\begin{cases} 1-\exp\left[-(x/\eta)^m\right] & (x\geq 0) \\ 0 & (x<0) \end{cases}$$

で与えられる。以下の各問に答えよ。

〔1〕　$W(m,\eta)$ の確率密度関数 $f(x)$ および最頻値（mode）を求め，$W(2,10)$ および $W(2,5)$ の確率密度関数の概形をそれぞれ図示せよ。

〔2〕　$W(m,\eta)$ のハザード関数 $h(x)=\dfrac{f(x)}{1-F(x)}$ を求め，m の値による関数 $h(x)$ の特徴について論ぜよ。

〔3〕　k 個の部品からなる直列システムにおいて，各部品の寿命は互いに独立に $W(m,\eta)$ に従うとする。このシステムは k 個の部品の一つでも壊れると稼働を停止する。このとき，このシステムの稼働停止までの寿命がどのような分布に従うか示せ。

〔4〕　ある部品 A の強度を表す確率変数を X とし，部品 A が受けるストレスを表す確率変数を Y とする。X と Y は互いに独立で，それぞれ $W(2,10)$ および $W(2,5)$ に従うとする。部品 A は $X<Y$ のときに壊れるとしたとき，部品 A が壊れる確率を求めよ。

〔5〕　寿命が互いに独立に $W(m,\eta)$ に従うような n 個の部品につき，それぞれ独立に寿命試験を行い，観測値 $x_1,\ldots,x_n$ を得たとする。これらの観測値に基づく $W(m,\eta)$ のパラメータ m および η の対数尤度関数を示せ。また，$\log\log\dfrac{1}{1-F(x)}$ を求め，これを用いてパラメータを推定する方法について論ぜよ。

問 3　機械部品の寿命などの正の値のみを取るデータの分布のモデルとして対数正規分布がある。正の値を取る確率変数 X を自然対数で変換した $\log X$ が $N(\mu, \sigma^2)$ に従うとき，X は対数正規分布に従うといい，$X \sim LN(\mu, \sigma^2)$ と書く。$LN(\mu, \sigma^2)$ の確率密度関数は，$x > 0$ として

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\frac{1}{x}\exp\left[-\frac{(\log x - \mu)^2}{2\sigma^2}\right]$$

である。以下の各問に答えよ。

〔1〕　$X \sim LN(\mu, \sigma^2)$ のとき，X の期待値 $E[X]$，中央値 (median)，最頻値 (mode) をそれぞれ求め，それらの大小関係を示せ。

〔2〕　$LN(\mu, \sigma^2)$ からの n 個の互いに独立な観測値を $x_1, \ldots, x_n$ としたとき，μ の最尤推定量を導出せよ。

〔3〕　ベイズ流の推測において，σ^2 は既知とし μ の事前分布を $N(\nu, \tau^2)$ とする。$LN(\mu, \sigma^2)$ からの n 個の独立な観測値を $x_1, \ldots, x_n$ としたときの μ の事後分布を導出せよ。

ある電子部品 D について，それぞれ独立に寿命試験を行い，表 1 の寿命データ (x) を得た。表 1 には寿命データに加え，その自然対数 $(\log x)$ の値，および x と $\log x$ の平均も示している。以下の各問に答えよ。なお，$\sigma^2 = (0.25)^2$ は既知とする。

表 1: 10 個の寿命データ（単位は時間）

ID	1	2	3	4	5	6	7	8	9	10	平均
寿命 (x)	1635	1772	1781	1899	2217	2241	2460	2745	3000	3525	2327.5
$\log x$	7.40	7.48	7.48	7.55	7.70	7.71	7.81	7.92	8.01	8.17	7.72

〔4〕　上問〔2〕で導出した μ の最尤推定量 $\hat{\mu}$ を元に，上問〔1〕の結果を用いて，表 1 のデータから X の期待値 $E[X]$，中央値 (median)，最頻値 (mode) の推定値をそれぞれ求めよ。巻末付表 5 も参照のこと。

〔5〕　ベイズ流の推測において，μ の事前分布を $N(8, (0.5)^2)$ とする。上問〔3〕の結果を用いて，表 1 のデータから μ の事後期待値を求めよ。

〔6〕　X の期待値 $E[X]$ の推定値としては $x_1, \ldots, x_n$ の標本平均 $\bar{x} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i$ が考えられる。上問〔4〕で求めた期待値 $E[X]$ の推定値と標本平均 $\bar{x}$ をそれぞれ推定量（確率変数）と見たときの，それらの推定量の統計的な特徴について述べよ。

問 4　ある化学物質の合成において，その合成物の収量 Y を最大化する要因を探索し，それらの最適水準を決定したい。以下の各問に答えよ。

〔1〕　化学的知見から収量に影響を与えるであろう要因候補はA，B，C，D，Eの5種類に絞られ，スクリーニング実験を行うこととした。技術者のSさんは5つの要因のうち，E以外の4つの要因につき実験すればいいと考えたが，別の技術者のTさんは5つすべてを取り入れた実験をすべきと主張した。各要因はそれぞれ2水準ずつ（第1水準 $=-1$，第2水準 $=1$）とし，SさんTさんとも8回の実験を行うとする。統計家のFさんはSさんおよびTさんに対し，適当と思われる実験計画を提案した。統計家のFさんが提案したであろうSさん用の4因子実験およびTさん用の5因子実験の実験計画表の空欄部分の水準を示せ。解答用紙に表1のような表を各自書いて解答すること。

表 1: 実験計画表

(a) Sさん用の4因子実験

	要因			
実験	A	B	C	D
1	1	1	1	1
2	1	1	−1	−1
3				
4				
5				
6				
7				
8				

(b) Tさん用の5因子実験

	要因				
実験	A	B	C	D	E
1	1	1	1	1	1
2	1	1	−1	−1	1
3					
4					
5					
6					
7					
8					

〔2〕　上問〔1〕で作成した4因子実験の実験計画につき，実験結果を $y_1, \ldots, y_8$ とするとき，要因Aの主効果の計算式，および要因Aと要因Bの2因子交互作用の計算式を示せ。

〔3〕　上問〔1〕で作成した5因子実験の実験計画につき，各要因の主効果と各要因間の2因子交互作用との関係を論ぜよ。具体的に，要因Aの主効果と交絡する2因子交互作用はどれかを示せ。

〔4〕　技術者のＳさんは統計家のＦさんの提案した計画に基づいて実験を行った。観測結果に基づいて各要因の主効果の推定を行ったところ，以下の結果を得た。各要因の主効果に関する検定を行い，5％有意となる要因をすべて示せ。

要因	係数	標準誤差
切片	9.25	0.78
A	2.75	0.78
B	4.00	0.78
C	1.00	0.78
D	−0.75	0.78

〔5〕　上問〔4〕の実験結果と専門的な知見から，共に連続的な「反応温度」と「反応時間」の影響が大きいことが分かった。それらの水準の最適な組合せを探索するため，x_1 を反応温度，x_2 を反応時間とした2次多項式 $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2$ の当てはめを行う。採用する実験計画は複合中心計画とし，実験点は $(0, 0)$, $(\pm 1, \pm 1)$, $(0, \pm\sqrt{2})$, $(\pm\sqrt{2}, 0)$ とする。中心点で4回実験を行い，その他の点では1回ずつの計12回の実験を行うとしたとき，パラメータ $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_{11}, \beta_{22}, \beta_{12})'$ に関する 12×6 のデザイン行列 X に対し，X' を X の転置行列として

$$(X'X)^{-1} = \frac{1}{32} \begin{pmatrix} 8 & 0 & 0 & -4 & -4 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 \\ -4 & 0 & 0 & 5 & 1 & 0 \\ -4 & 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 8 \end{pmatrix}$$

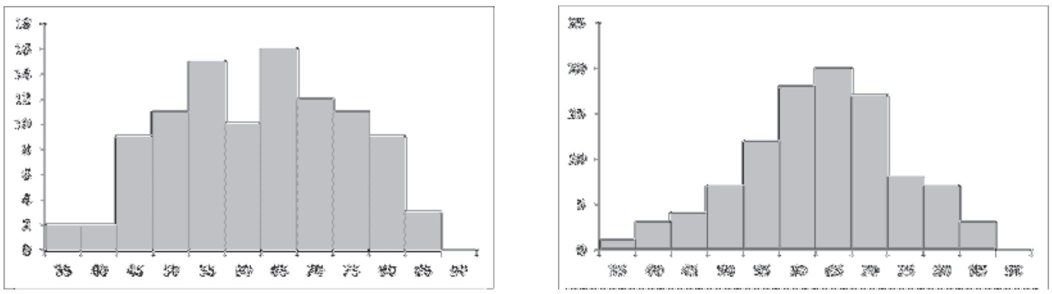
であることが示される。この計画は回転可能性 (rotatability) を満たすこと，すなわち，実験領域内のある点 $\boldsymbol{x} = (x_1, x_2)'$ での予測分散が $x_1^2 + x_2^2$ の関数となることを示せ。

問 5　ある大学のある学部では，入試科目の数学が選択制となっている。その学部の M 教授は，統計学のクラスを A，B の 2 つ受け持っていて，両クラスとも 100 名ずつの受講者がいる。両クラスとも入試で数学を選択した学生と非選択の学生は 50 名ずつであった。両クラスで統計学の試験を行ったところ入試での数学の選択別の平均と標準偏差は表 1 のようであった。ここで，表の標準偏差は，両クラスとも除数を 50 とした標本分散の正の平方根である。また，それぞれのクラスでヒストグラムを描いたところ図 1 のようであった。

両クラスとも入試での数学の選択の有無で明らかに平均が異なるので，ヒストグラムはふた山型となると予想されたが，A クラスではそうであるものの，B クラスではひと山型であった。M 教授はどのような場合に，混合分布がふた山型を示すのかに興味を持ち，理論的に考察することにした。

表 1: 試験結果の基本統計量

	A クラス		B クラス	
統計量	数学選択	数学非選択	数学選択	数学非選択
平均	69.7	49.6	67.6	53.8
標準偏差	6.8	7.8	7.7	8.7



A クラス　　　　B クラス

図 1: クラス全体の試験結果のヒストグラム

母集団全体の分布 F は，分散は等しいが平均は異なる正規分布 $N(\mu_1, \sigma^2)$ と $N(\mu_2, \sigma^2)$ の混合率 1/2 ずつの混合であるとする。すなわち，$N(\mu_j, \sigma^2)$ の確率密度関数を $f_j(x) = \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left[-\dfrac{(x-\mu_j)^2}{2\sigma^2}\right]$ としたとき $(j = 1, 2)$，分布 F の確率密度関数 $f(x)$ は

$$f(x) = \frac{1}{2}\{f_1(x) + f_2(x)\}$$

で与えられる。図 2 はいくつかの μ_1，μ_2，σ の組み合わせに対応した $f(x)$ の形状である。

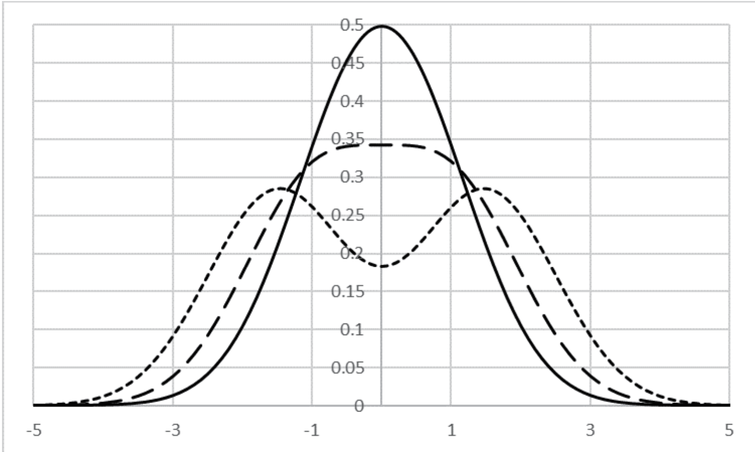


図 2: いくつかのパラメータに関する $f(x)$ の形状

分布 F に従う確率変数を X とするとき，以下の各問に答えよ。

〔1〕　X の期待値と分散はそれぞれ

$$\xi = E[X] = \frac{\mu_1 + \mu_2}{2},\ \tau^2 = V[X] = \sigma^2 + \left(\frac{\mu_1 - \mu_2}{2}\right)^2$$

となることを示せ。

ヒント: $(x - \xi)^2 = \left(x - \dfrac{\mu_1 + \mu_2}{2}\right)^2 = \left(x - \mu_1 + \dfrac{\mu_1 - \mu_2}{2}\right)^2 = \left(x - \mu_2 - \dfrac{\mu_1 - \mu_2}{2}\right)^2$

である。

〔2〕　表 1 の結果から，A クラス全体のデータを $x_i (i = 1, \ldots, 100)$ としたときの平均 $\bar{x} = \dfrac{1}{100}\displaystyle\sum_{i=1}^{100} x_i$ と標準偏差 $s = \sqrt{\dfrac{1}{100}\displaystyle\sum_{i=1}^{100}(x_i - \bar{x})^2}$ を求めよ。

〔3〕　確率密度関数 $f(x)$ の 1 次導関数 $f'(x)$ および 2 次導関数 $f''(x)$ を求めよ。また，$x = \xi$ は $f(x)$ の極値を与えることを示せ。

〔4〕　分布 F がふた山型（二峰性）を示すための μ_1，μ_2，σ の条件を求めよ。

医薬生物学

問 1　生存時間解析において，T を生存時間を表す確率変数，T の確率密度関数を $f(t)$，累積分布関数を $F(t)$，生存関数を $S(t)$，ハザード関数を $h(t)$，累積ハザード関数を $H(t)$ とする。ただし，$f(t)$ はパラメータ λ の指数分布の確率密度関数

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & (t \geq 0) \\ 0 & (t < 0) \end{cases}$$

とする。このとき，以下の各問に答えよ。

〔1〕　互いに独立にパラメータ λ の指数分布に従う n 人の生存時間 $(t_1, t_2, \ldots, t_n)$ が観測されたとする。ただし，打ち切りはないと仮定する。このとき，パラメータ λ に関する尤度関数を示せ。

〔2〕　i 番目の被験者のデータを (t_i, δ_i) $(i = 1, 2, \ldots, n)$ とする。δ_i は生存時間 t_i が打ち切りを受けた場合 0，打ち切りを受けなかった場合 1 とする。ただし，打ち切りはランダムな右側打ち切りであるとする。このとき，パラメータ λ に関する尤度関数を示せ。

〔3〕　上問〔2〕の尤度関数を用いて，λ の最尤推定量とフィッシャー情報量 $I(\lambda)$ を求めよ。ただし，イベント数は $\delta = \sum_{i=1}^{n} \delta_i$ とする。

〔4〕　試験群と対照群の比較をするために，ランダム化比較試験を実施することを考える。試験群と対照群の生存時間は，それぞれハザードが λ_1，λ_2 である指数分布に従うとする。試験群と対照群の期待されるイベント数をそれぞれ d_1，d_2 とする。また，試験群と対照群の割付比は 1 対 1 とする。帰無仮説を $\log \lambda_1 = \log \lambda_2$，対立仮説を $\log \lambda_1 \neq \log \lambda_2$，$\lambda_1$，$\lambda_2$ の最尤推定量をそれぞれ $\hat{\lambda}_1$，$\hat{\lambda}_2$ とする。このとき，デルタ法を用いて $\log \hat{\lambda}_1 - \log \hat{\lambda}_2$ の漸近分散を求め，Z 検定統計量を求めよ。ただし，$\hat{\lambda}_1$，$\hat{\lambda}_2$ の漸近分散がそれぞれ $1/I(\lambda_1)$，$1/I(\lambda_2)$ であることを証明なしに用いてよい。

〔5〕　検出力 $1 - \beta$ を満たす試験群と対照群の必要イベント数を $r_1 = r_2 = r$ としたとき，次式が成り立つことを示せ。

$$r = \frac{2(Z_{\alpha/2} + Z_\beta)^2}{(\log(\lambda_1) - \log(\lambda_2))^2}$$

ただし $Z_{\alpha/2}$ は標準正規分布の上側 $100\alpha/2$ % 点，Z_β は標準正規分布の上側 100β% 点である。

問 2　試験治療と対照治療の有効率を比較するための臨床試験を計画した。背景因子が類似している 2 人の被験者をペアとし，そのペアが n 組あるとする。各ペアに対して，片方には試験治療，もう片方には対照治療をランダムに割り付けた。治療の結果は，有効と無効で評価されるものとする。試験の観察結果を表 1 のように整理した。表 2 は表 1 に対応する母集団確率とする。

表 1: 観察結果

試験治療	対照治療 有効	対照治療 無効	合計
有効	n_{11}	n_{12}	$n_{1\cdot}$
無効	n_{21}	n_{22}	$n_{2\cdot}$
合計	$n_{\cdot 1}$	$n_{\cdot 2}$	n

表 2: 母集団確率

試験治療	対照治療 有効	対照治療 無効	合計
有効	π_{11}	π_{12}	$\pi_{1\cdot}$
無効	π_{21}	π_{22}	$\pi_{2\cdot}$
合計	$\pi_{\cdot 1}$	$\pi_{\cdot 2}$	1

n_{ij} $(i = 1, 2; j = 1, 2)$ は多項分布に従うとして，以下の各問に答えよ。

〔1〕　試験治療と対照治療の有効率を比較するパラメータを $\delta = \pi_{12} - \pi_{21}$ とする。また，δ の推定量を $\hat{\delta} = \dfrac{n_{12} - n_{21}}{n}$ とする。このとき，$\hat{\delta}$ の分散を求めよ。

〔2〕　多項分布に対する正規近似を用いて，$\delta = \pi_{12} - \pi_{21}$ の $100(1 - \alpha)$ % 両側信頼区間を導出せよ。ただし，$Z_{\alpha/2}$ を標準正規分布の上側 $100\alpha/2$ % 点とする。

〔3〕　試験治療と対照治療の有効率が等しいかどうかを検定することとする。帰無仮説を $\delta = 0$，対立仮説を $\delta \neq 0$ とし，帰無仮説と対立仮説の下での π_{ij} の最尤推定量 $\hat{\pi}_{ij}$ をそれぞれ求めよ。

〔4〕　上問〔3〕の検定をするために，帰無仮説と対立仮説の下での尤度から，尤度比検定統計量とその自由度を求めよ。

問 3　ある 1 つのバイオマーカー X と疾患 Y の関連を評価したい。ここで，D を疾患 Y に罹患しているか否かを表す確率変数と定義する。

$$D = \begin{cases} 1 & \text{疾患 } Y \text{ に罹患している} \\ 0 & \text{疾患 } Y \text{ に罹患していない} \end{cases}$$

このときバイオマーカー X と疾患 Y の関連をバイオマーカー X のみを説明変数とするロジスティック回帰モデルによって評価することを考える。16 名の被験者から表 1 のデータが得られたとき，以下の各問に答えよ。ただし，表中の「予測値」は上記モデルをこの 16 名のデータに適用した際に得られる各被験者に対する「疾患 Y に罹患している確率」の推定値である。

表 1: 16 名の被験者のデータ

被験者番号	X	D	予測値	被験者番号	X	D	予測値
1	81	1	0.87	9	47	1	0.43
2	72	1	0.79	10	44	1	0.38
3	67	0	0.73	11	42	1	0.35
4	63	0	0.68	12	39	0	0.31
5	62	1	0.66	13	38	0	0.30
6	58	1	0.61	14	37	0	0.28
7	56	0	0.57	15	36	0	0.27
8	52	1	0.51	16	33	0	0.24

〔1〕　上の表において，予測値が 0.5 以上であれば当該被験者は疾患 Y に罹患していると判断するとしたとき，この集団における真陽性率（感度）と偽陽性率（$1-$特異度）をそれぞれ求めよ。

〔2〕　上の表から得られるこの集団における ROC 曲線の一部を図 1 に示した。偽陽性率が 0.3 以上の部分を補完した ROC 曲線を描け。

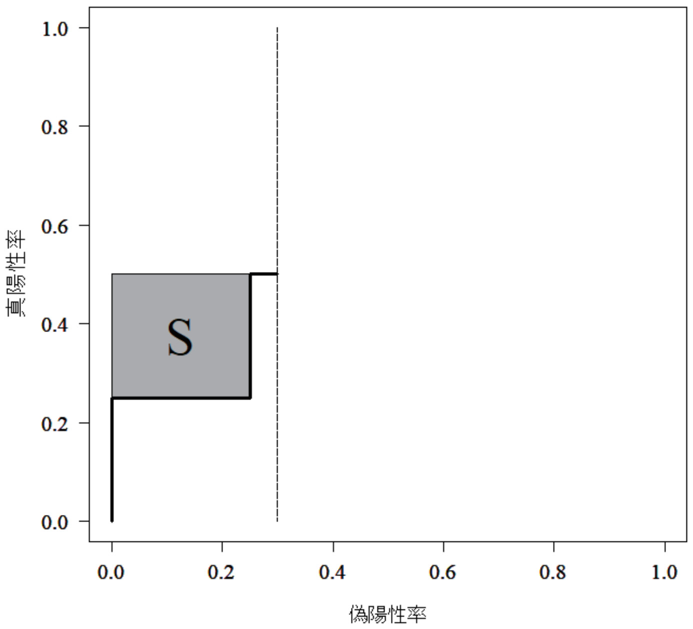


図 1: ROC 曲線の一部

〔3〕　上問〔2〕の ROC 曲線において,「真陽性率 − 偽陽性率」を最大にするには，上の表中の予測値がいくら以上の被験者を疾患 Y に罹患していると判断すべきであるかを答えよ。ただし，回答する予測値は表中に記載されているものとする。

〔4〕　表中の $D = 1$ の被験者のバイオマーカーの値を，予測値が高い方から順に x_{1i} $(i = 1, 2, \ldots, 8)$ とし，$D = 0$ の被験者のバイオマーカーの値を，予測値の高い方から順に x_{0j} $(j = 1, 2, \ldots, 8)$ とする。このとき，図 1 の「S」（背景がグレーとなっている部分）の面積が

$$\frac{1}{8 \times 8} \sum_{i=1}^{4} \sum_{j=1}^{2} I\left(x_{1i} < x_{0j}\right)$$

と表されることを示せ。ただし，

$$I(Z) = \begin{cases} 1 & Z \text{ が正しい} \\ 0 & Z \text{ が正しくない} \end{cases}$$

である。

〔5〕　上問〔2〕で描かれる ROC 曲線の曲線下面積が次式で表されることを示せ。

$$\frac{1}{8 \times 8} \sum_{i=1}^{8} \sum_{j=1}^{8} I\left(x_{1i} > x_{0j}\right)$$

問 4　ある疾病の発症の有無を予測する臨床的リスクスコアを導出するため，ロジスティック回帰モデルを用いて，ステップワイズ法により変数を選択したところ，表1の結果が得られたとする。表中のP値は，カイ二乗検定の結果とする。

表 1: ロジスティック回帰分析の結果

変数	オッズ比	95 %信頼区間	P 値
年齢 (歳)			
[8-12]/[5-7]	1.63	(1.07–2.33)	0.008
[13-18]/[5-7]	2.42	(1.87–3.12)	
性別			
女性/男性	2.42	(1.92–2.88)	0.001
頭痛の既往			
あり/なし	1.81	(1.32–2.65)	0.002
感音性			
あり/なし	1.77	(1.02–1.92)	0.010

得られた回帰モデルから，各変数の各水準にスコアを与えて，検証コホートのもとでの，リスクスコアによる予測のROC曲線と医師による予測のROC曲線を描いたところ，図1が得られたする。リスクスコアによる予測のROC曲線の面積下面積は0.70（95 %信頼区間: 0.67–0.72）とする。

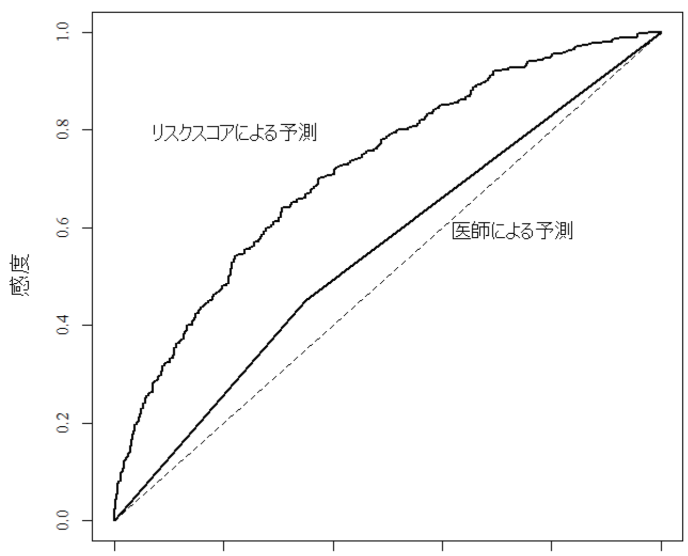


図 1: リスクスコアによる予測のROC曲線と医師による予測のROC曲線

〔1〕　目的変数 y と p 個の説明変数 $x_1, \ldots, x_p$ について n 組のデータ

$$\{(y_i, x_{i1}, \ldots, x_{ip}); i = 1, 2, \ldots, n\}, \quad y_i = \begin{cases} 1 & \text{疾病あり} \\ 0 & \text{疾病なし} \end{cases}$$

が観測されたとする。確率変数 Y を用いて反応確率と非反応確率を

$$\Pr(Y = 1 | x_1, \ldots, x_p) = \theta, \quad \Pr(Y = 0 | x_1, \ldots, x_p) = 1 - \theta$$

と表すことにする。このとき，反応確率と説明変数の関係は次のロジスティック回帰モデルで表されるとする。

$$\theta = \frac{\exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p)}{1 + \exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p)} \tag{1}$$

最尤法を用いて未知パラメータを推定することとする。モデル (1) の対数尤度関数を示せ。

〔2〕　x_1 が 0 または 1 の値をとる二値変数のとき，ロジスティック回帰モデルの偏回帰係数 β_1 は Y の x_1 に対する調整オッズ比の対数であることを示せ。

〔3〕　$\beta_0, \beta_1, \ldots, \beta_p$ の最尤推定量を $\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_p$ とする。モデル (1) に対する AIC を式で示せ。

〔4〕　AIC ではモデルのよさを Kullback-Leibler 情報量を用いて評価する。真のモデル $g(x)$ をベルヌーイ分布 $\mathrm{Bern}(\theta)$，想定したモデル $f(x)$ を $\mathrm{Bern}(\pi)$ とする。このとき $g(x)$ に対する $f(x)$ の Kullback-Leibler 情報量を示せ。

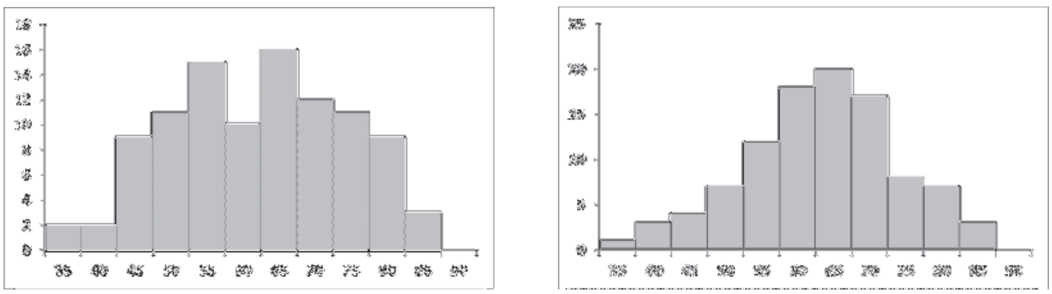
〔5〕　表 1 と図 1 の結果をもとに，導出したリスクスコアから解釈できることを 150 字以内で述べよ。

問 5　ある大学のある学部では，入試科目の数学が選択制となっている。その学部の M 教授は，統計学のクラスを A，B の 2 つ受け持っていて，両クラスとも 100 名ずつの受講者がいる。両クラスとも入試で数学を選択した学生と非選択の学生は 50 名ずつであった。両クラスで統計学の試験を行ったところ入試での数学の選択別の平均と標準偏差は表 1 のようであった。ここで，表の標準偏差は，両クラスとも除数を 50 とした標本分散の正の平方根である。また，それぞれのクラスでヒストグラムを描いたところ図 1 のようであった。

両クラスとも入試での数学の選択の有無で明らかに平均が異なるので，ヒストグラムはふた山型となると予想されたが，A クラスではそうであるものの，B クラスではひと山型であった。M 教授はどのような場合に，混合分布がふた山型を示すのかに興味を持ち，理論的に考察することにした。

表 1: 試験結果の基本統計量

	A クラス		B クラス	
統計量	数学選択	数学非選択	数学選択	数学非選択
平均	69.7	49.6	67.6	53.8
標準偏差	6.8	7.8	7.7	8.7



A クラス　　　　　B クラス

図 1: クラス全体の試験結果のヒストグラム

母集団全体の分布 F は，分散は等しいが平均は異なる正規分布 $N(\mu_1, \sigma^2)$ と $N(\mu_2, \sigma^2)$ の混合率 1/2 ずつの混合であるとする。すなわち，$N(\mu_j, \sigma^2)$ の確率密度関数を $f_j(x) = \dfrac{1}{\sqrt{2\pi}\sigma} \exp\left[-\dfrac{(x-\mu_j)^2}{2\sigma^2}\right]$ としたとき $(j = 1, 2)$，分布 F の確率密度関数 $f(x)$ は

$$f(x) = \frac{1}{2}\{f_1(x) + f_2(x)\}$$

で与えられる。図 2 はいくつかの μ_1，μ_2，σ の組み合わせに対応した $f(x)$ の形状である。

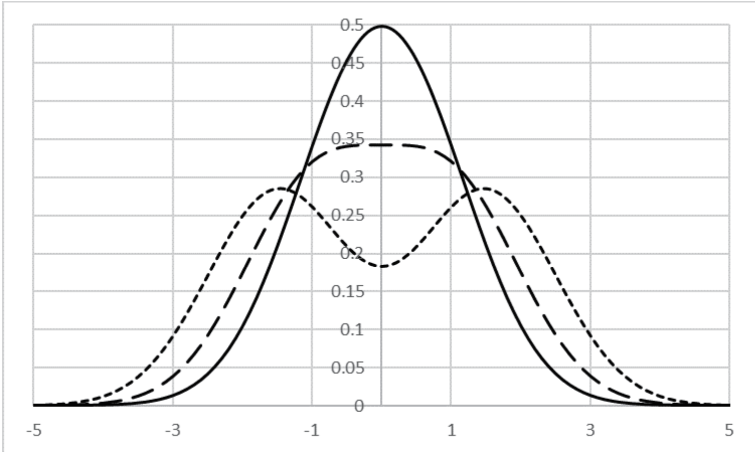


図 2: いくつかのパラメータに関する $f(x)$ の形状

分布 F に従う確率変数を X とするとき，以下の各問に答えよ。

〔1〕 X の期待値と分散はそれぞれ

$$\xi = E[X] = \frac{\mu_1 + \mu_2}{2},\ \tau^2 = V[X] = \sigma^2 + \left(\frac{\mu_1 - \mu_2}{2}\right)^2$$

となることを示せ。

ヒント: $(x - \xi)^2 = \left(x - \frac{\mu_1 + \mu_2}{2}\right)^2 = \left(x - \mu_1 + \frac{\mu_1 - \mu_2}{2}\right)^2 = \left(x - \mu_2 - \frac{\mu_1 - \mu_2}{2}\right)^2$

である。

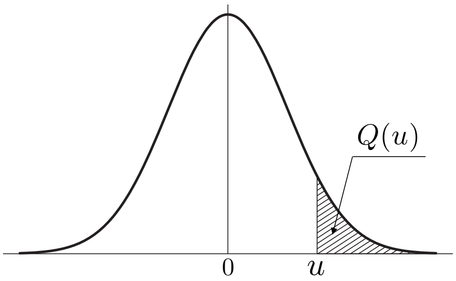
〔2〕 表 1 の結果から，A クラス全体のデータを $x_i (i = 1, \ldots, 100)$ としたときの平均 $\bar{x} = \frac{1}{100}\sum_{i=1}^{100} x_i$ と標準偏差 $s = \sqrt{\frac{1}{100}\sum_{i=1}^{100}(x_i - \bar{x})^2}$ を求めよ。

〔3〕 確率密度関数 $f(x)$ の 1 次導関数 $f'(x)$ および 2 次導関数 $f''(x)$ を求めよ。また，$x = \xi$ は $f(x)$ の極値を与えることを示せ。

〔4〕 分布 F がふた山型（二峰性）を示すための μ_1，μ_2，σ の条件を求めよ。

付　表

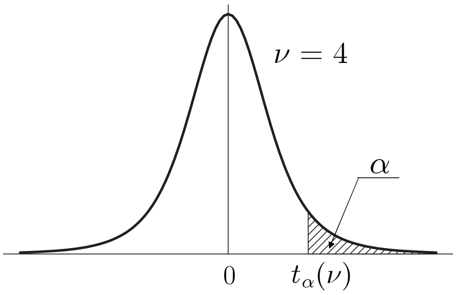
付表 1. 標準正規分布の上側確率



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。

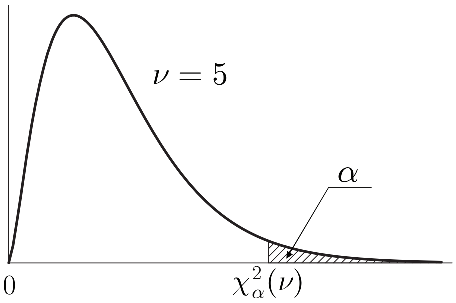
付表 2. t 分布のパーセント点



	α				
ν	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。

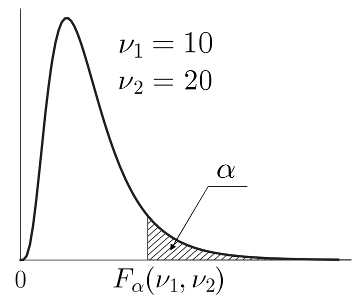
付表 3. カイ二乗分布のパーセント点



	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。

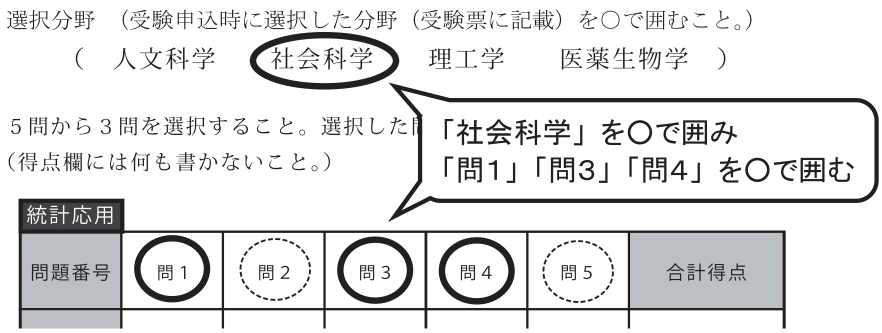
【解答冊子記入例】

- 注意事項 6 ⑥ :〔解答冊子各ページ先頭の記入例〕

（例）問 1 を解答する場合

- 注意事項 6 ⑦ :〔解答冊子表紙選択分野・選択問題の記入例〕

（例）**社会科学** 分野の**問 1**，**問 3**，**問 4** を選択し，解答する場合



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**

2018.11