統計検定

Japan Statistical Society Certificate

1級

統計応用

2016 年 11 月 27 日

人文科学

問 1　ある組織に属する構成員に対して，6 項目 $(X_1, \ldots, X_6)$ からなる心理テストを実施した。このテストの各項目のスコアはそれぞれ連続量で表される。テスト結果から計算された相関係数行列に対して探索的因子分析モデルを当てはめ，斜交回転であるプロマックス回転を行ったところ，表 1 の因子負荷量行列を得た。なお，各因子は平均 0，分散 1 に標準化され，因子間相関係数は 0.5 であった。以下の各問に答えよ。

表 1：得られた因子負荷量行列

変数	因子 1	因子 2
X_1	0.8	0.2
X_2	0.8	−0.2
X_3	0.6	0.2
X_4	0.2	0.5
X_5	0.4	0.6
X_6	−0.2	0.7

〔1〕　X_1 の共通性はいくらか。

〔2〕　X_3 と因子 1 の間の相関係数はいくらか。

〔3〕　因子分析モデルに基づいて，因子負荷量と因子間相関係数から得られる X_3 と X_5 の間の相関係数を求めよ。

〔4〕　直交回転と斜交回転の違いについて論ぜよ。

問 2　ある大学では，英語の授業を「上級コース」と「一般コース」に分けて行っている。コースの選択は学生の自己判断で，学生はいずれかのコースを履修しなくてはならない。また，この大学では，新入生全員に対し，入学時に全国規模の英語能力試験を実施している。その試験の，2011 年度入学者の各コースの履修者の平均点，2014 年度入学者の各コースの履修者の平均点，および両年度での平均点の差は表 1 のようであり，それぞれの年度の各コースの履修割合は表 2 のようであった。以下の各問に答えよ。

表 1：英語能力試験の平均点

コース	2011 年度	2014 年度	差
上級コース	530	536	6
一般コース	445	446	1

表 2：各コースの履修割合

コース	2011 年度	2014 年度
上級コース	0.3	0.25
一般コース	0.7	0.75

〔1〕　この大学の新入生全体の英語能力試験の 2011 年度および 2014 年度での平均点を求めよ。大学の新入生全体の平均点の 2011 年度から 2014 年度への変化を，表 1 の各コースでの平均点の変化と比較してその理由を論ぜよ。

〔2〕　2017 年度は，コース分けを学生の自己判断ではなく，英語能力試験の点数によって行うとする。新入生全体の英語能力試験の点数の分布が平均 468 点，標準偏差 80 点の正規分布 $N(468, 80^2)$ であるとし，英語能力試験の点数が 500 点以上の学生を上級コース，500 点未満の学生を一般コースに振り分けた場合，上級コースおよび一般コースの学生の割合はそれぞれいくらとなるか。

〔3〕　上問〔2〕と同じ状況で，2017 年度の上級コースおよび一般コースでの英語能力試験の点数の平均はそれぞれいくらとなるか。ただし，標準正規分布 $N(0, 1)$ の確率密度関数は $\phi(z) = \dfrac{1}{\sqrt{2\pi}} \exp[-z^2/2]$ である。

問 3　くりかえし受験できる標準化された試験（以下，単に試験と呼ぶ）において，各受験者の得点 x は独立な確率変数と考えられる。ある地域で，中学生の学習達成度を測定するために，この試験を利用することとした。

学校 s $(s = 1, \ldots, M)$ の生徒数を N_s とし，学校 s の生徒 i $(i = 1, \ldots, N_s)$ の得点を x_{si} と表す。x_{si} は各生徒で独立であり，期待値 $\mathrm{E}[x_{si}] = \mu_s$，分散 $\mathrm{V}[x_{si}] = \sigma_s^2$ とする。以下では，学校 s の生徒の得点を表す変数を代表的に x_s と表す。

地域全体の平均 μ を推定するために，地域内にある M 校の中学校から m 校を抽出し，抽出された m 校それぞれから n 人ずつの生徒を無作為抽出して試験を実施するという 2 段抽出の計画を立てた。ただし $n \leq N_s$ $(s = 1, \ldots, M)$ とする。また，地域の生徒数合計を $N = \sum_s N_s$，割合を $W_s = N_s/N$ と表す。この 2 段抽出 (TS, Two Stage sampling) と，この地域全体の生徒から mn 人を等しい確率で無作為に抽出して試験を実施する単純無作為抽出 (SI, SImple sampling) を比較したい。

ここで，学校の抽出方法を記述するために，s が抽出されたとき $R_s = 1$，抽出されないとき $R_s = 0$ となる確率変数 R_s を導入する。この確率変数 R_s は x_s とは独立，$\sum_{s=1}^{M} R_s = m$ であり，s が抽出される確率を $\mathrm{P}(R_s = 1) = W_s$ とする。以下の各問に答えよ。

〔1〕　地域全体の生徒から無作為に 1 人を抽出したときの得点を x とするとき，$x = \displaystyle\sum_{s=1}^{M} R_s x_s$ と表せることを確かめ，この表現を用いて，期待値 $\mathrm{E}[x]$ と分散 $\mathrm{V}[x]$ が，それぞれ次の式で与えられることを示せ。これらが地域全体における得点の平均 μ と分散 σ^2 である。

$$\mu = \mathrm{E}[x] = \sum_s W_s \mu_s, \quad \sigma^2 = \mathrm{V}[x] = \sum_s W_s \sigma_s^2 + \sum_s W_s (\mu_s - \mu)^2$$

〔2〕　学校 s で無作為に抽出された生徒 n 人の得点の平均を $\bar{x}_s$ と表す。1 段目の大きさを $m = 1$ とする 2 段抽出では，学校 s を確率 W_s で選び，その学校の $\bar{x}_s$ を得点 x_{TS} とする。期待値 $\mathrm{E}[x_{TS}]$ と分散 $\mathrm{V}[x_{TS}]$ が，それぞれ次の式で与えられることを示せ。

$$\mathrm{E}[x_{TS}] = \mu, \quad \mathrm{V}[x_{TS}] = \frac{1}{n}\sum_s W_s \sigma_s^2 + \sum_s W_s (\mu_s - \mu)^2$$

〔3〕　地域全体の得点の平均 μ の推定量としてふたつの調査方法を比較したい。全部で mn 人を調査する単純無作為抽出による推定量を $\hat{\mu}_{SI}$，その分散を V_{SI} とする。また第 1 段で m 校を抽出し第 2 段で n 人を抽出する 2 段抽出による推定量を $\hat{\mu}_{TS}$，その分散を V_{TS} とする。ここでは単純化のため，単純無作為抽出および 2 段抽出のいずれも復元抽出を仮定する。

$\hat{\mu}_{SI}$ と $\hat{\mu}_{TS}$ はいずれも不偏推定量であることを示し，それらの分散 V_{TS}，V_{SI} を求め，$\mathrm{V}_{TS} = \mathrm{V}_{SI}$ が成立するための条件を記せ。

〔4〕　調査にかかる費用，手間と推定量の精度を考慮して，単純無作為抽出 (SI) と 2 段抽出 (TS) の得失を論ぜよ。

問 4　ある高校において，数学と理科の間の相関を調べるため，それぞれの科目について 2 つずつのテスト得点をデータとして得た。これら 4 つのテスト得点を使って，図 1 のような 2 因子の確認的因子分析モデルを構成し，共分散構造分析の枠組みで推定を行いたい。以後，テスト得点を観測変数と呼ぶ。図 1 の数学と理科が 2 つの因子（f_1 と f_2，共に平均は 0）であり，4 つの四角（x_1 から x_4）が観測変数（平均からの偏差で表されているとする），その下の 4 つの丸（e_1 から e_4）が誤差を表している。また，r は因子間の共分散, a_1 から a_4 は因子負荷量，ve_1 から ve_4 は誤差分散を表す母数である。また，図 1 では，各因子の分散を 1（各因子の右上の数値），誤差からの影響をすべて 1（各矢印の右の数値）としている。以下の各問に答えよ。

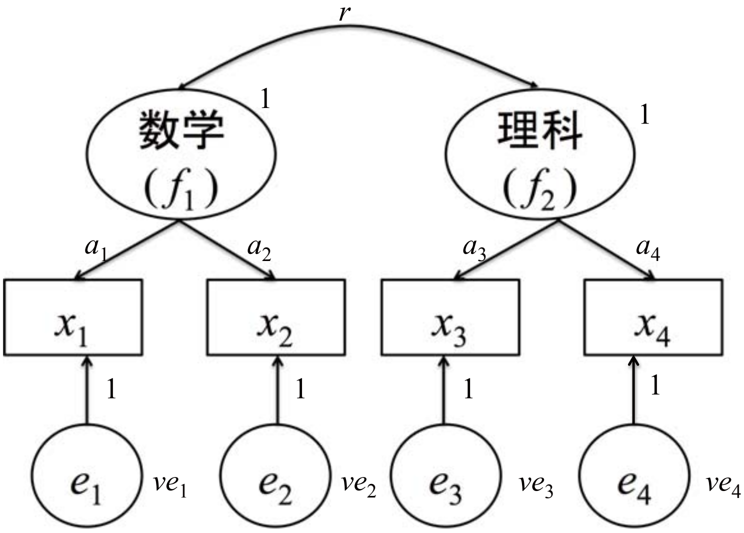


図 1: 2 因子の確認的因子分析モデル

〔1〕　このモデルで推定すべき母数はいくつあるか。

〔2〕　このモデルの自由度はいくらか。ただし，共分散構造分析における自由度とは，観測変数の数を k とするとき，$k(k+1)/2$ から推定すべき母数の数を引いた値である。

〔3〕　x_1 に関する測定方程式を記述せよ。

〔4〕　図 1 で示された 2 因子の確認的因子分析モデルについて，共分散行列はどのようになるか。以下の分散共分散行列の対角要素を含む下三角部分の空欄部分を求めよ。

分散共分散行列

	x_1	x_2	x_3	x_4
x_1		-	-	-
x_2			-	-
x_3				-
x_4				

〔5〕　因子間の共分散 r が 0 であるとの制約を置いた場合，何が問題となるかを答えよ。また，その制約はここでのデータにとって適切であると考えられるかを理由と共に答えよ。

問 5　2 種類の処置 A，B の効果の差を調べるため，被験者 30 人をランダムに 15 人ずつの 2 群に分け，片方には処置 A を施し，もう片方には処置 B を施して経過を観察した。実験結果は正規分布に従う連続量のスコアで表され，処置 A では $x_1, \ldots, x_{15}$，処置 B では $y_1, \ldots, y_{15}$ が観測された。各処置での標本平均

$$\bar{x} = \frac{1}{15}\sum_{i=1}^{15} x_i, \quad \bar{y} = \frac{1}{15}\sum_{i=1}^{15} y_i$$

と標本標準偏差

$$s_x = \sqrt{\frac{1}{14}\sum_{i=1}^{15}(x_i - \bar{x})^2}, \quad s_y = \sqrt{\frac{1}{14}\sum_{i=1}^{15}(y_i - \bar{y})^2}$$

および，t 分布に基づいた処置 A でのスコアの母平均 μ_A の信頼係数 95% の信頼区間の下限と上限が表 1 に示されている。以下の各問に答えよ。

表 1：実験結果の要約

				95% 信頼区間	
	n	平均	標準偏差	下限	上限
処置 A	15	27.0	8.0	22.57	31.43
処置 B	15	20.0	10.0	（ア）	（イ）

〔1〕　処置 B でのスコアの母平均 μ_B の信頼係数 95% の信頼区間の下限（ア）及び上限（イ）を求めよ。

〔2〕　両処置群のスコアの母分散は等しいが未知であるとして，帰無仮説 $H_0 : \mu_A = \mu_B$，対立仮説 $H_1 : \mu_A \neq \mu_B$ の有意水準 5% の 2 標本両側 t 検定を実行し，その結果を述べよ。また，母平均の差 $\mu_A - \mu_B$ の信頼係数 95% の信頼区間を求めよ。

〔3〕　上問〔1〕で求めたような処置群ごとの母平均の信頼係数 95% の信頼区間の重なりと，上問〔2〕の 2 標本両側 t 検定の結果との関係を調べる。簡単のため，両群でのサンプルサイズは等しくともに n とし，標本標準偏差は等しい，$s_x = s_y (= s)$，とする。このとき，信頼区間の重なりがなければ 2 標本両側 t 検定は有意になるかを答えよ。また，信頼区間の重なりがあった場合，その重なりがいくら以下のとき 2 標本両側 t 検定は有意になるかを答えよ。

社会科学

問 1　経済分析で用いられる古典的な Cobb-Douglas 型の生産関数は，生産額を Y，資本ストックを K，労働投入量を L として $Y = A\,K^{\beta_1}\,L^{\beta_2}$ で表される。ここで $\beta_1, \beta_2 > 0$ である。表 1 は各変数を対数に変換した生産額 ($y = \log Y$)，資本ストック ($x_1 = \log K$)，労働投入量 ($x_2 = \log L$) の値であり，平均，標準偏差（不偏分散の平方根）および相関係数は表 2 に示すとおりである。

表 1：生産関数のデータ（標本の大きさは $n = 9$ 期間)

y	.831	.863	.908	.908	.987	1.017	1.045	1.093	1.114
x_1	.423	.447	.471	.503	.535	.564	.593	.631	.669
x_2	.242	.300	.345	.387	.443	.491	.568	.628	.669

表 2：平均，標準偏差，相関係数

	y	x1	x2
平均	0.9740000	0.5373333	0.4525556
sd	0.1014286	0.0842140	0.1483232

（sd は不偏分散の平方根）

y	1.0000000	0.9902736	0.9918601
x1	0.9902736	1.0000000	0.9971798
x2	0.9918601	0.9971798	1.0000000

（相関係数）

係数 β_1, β_2 を推定するために，対数変換した次の回帰式を想定する。ここで u は誤差項である。

$$y = \log A + \beta_1 x_1 + \beta_2 x_2 + u \tag{1}$$

この式を最小二乗法で推定した際，x_1 の 4 期目の観測値 0.503 を誤って 0.530 と入力したところ，表 3 の結果を得た。表 3 の変数名 x1a は誤りを含むデータである。表 3 の結果を疑問に思ってデータを点検し，誤りを修正して実行した回帰分析の結果が表 4 である。

表 3：回帰分析の出力（誤ったデータ）

```
Coefficients:
       Estimate Std. Error t value P (>|t|)
(Int.)   0.7979     0.1193   6.690 0.000541 ***
x1a     -0.4534     0.4098  -1.106 0.310929
x2       0.9304     0.2302   4.042 0.006787 **
---
Residual standard error: 0.01359
Multiple R-squared: 0.9865
F-statistic: 219.8 on 2 and 6 DF
```

表 4：回帰分析の出力（正しいデータ）

```
Coefficients:
       Estimate Std. Error t value  P(>|t|)
(Int.)   0.5943     0.2332   2.548   0.0436 *
x1       0.2589     0.8275   0.313   0.7650
x2       0.5317     0.4698   1.132   0.3010
---
Residual standard error: 0.01479
Multiple R-squared: 0.9840
F-statistic:   185 on 2 and 6 DF
```

以下の各問に答えよ。回帰分析における誤差項は標準的な仮定をおき，特に正規分布を想定してよい。なお，信頼係数は 95%，仮説検定の有意水準は 5%とする。

〔1〕　表 4 の結果にもとづいて，β_1 の信頼区間を求めよ。

〔2〕　表 3 の結果が不適当とみなされる理由を記せ。

〔3〕　表4の結果および以下のW, W^{-1}の数値を用いて，$\hat{\beta}_1+\hat{\beta}_2$の標準誤差を求め，仮説$H_0 : \beta_1+\beta_2 = 1$を検定せよ。ここで$2\times 2$行列$W$の成分は$w_{11} = \sum_{i=1}^{n}(x_{1i}-\bar{x}_1)^2$, $w_{12} = w_{21} = \sum_{i=1}^{n}(x_{1i}-\bar{x}_1)(x_{2i}-\bar{x}_2)$, $w_{22} = \sum_{i=1}^{n}(x_{2i}-\bar{x}_2)^2$である。

偏差の積和・平方和W

```
         x1       x2
x1 0.056736 0.099645
x2 0.099645 0.175998
```

逆行列W^{-1}

```
          x1        x2
x1  3129.264 -1771.703
x2 -1771.703  1008.771
```

〔4〕　生産要素K，Lを$\lambda(>0)$倍したときに生産量もλ倍となる「一次同次性」の条件は$\beta_1+\beta_2=1$と表され，この性質は近似的に成立することが多い。そこで$\beta_1+\beta_2=1$を用いてモデルを$(Y/L) = A\,(K/L)^{\beta_1}$と書き変えて，次の(2)式を想定した。

$$\log(Y/L) = \log A + \beta_1 \log(K/L) + u \tag{2}$$

変数名を y.2 $=\log(Y/L)$, x1.2 $=\log(K/L)$として，(2)式を最小二乗法で推定した結果が表5である。この出力にもとづいて，β_1の信頼区間を求めよ。

表5：回帰分析の出力（1次同次性の仮定）

```
Coefficients:
            Estimate Std. Error t value P(>|t|)
(Intercept) 0.458963   0.008037  57.110 1.32e-10 ***
x1.2        0.737000   0.076965   9.576 2.84e-05 ***
---
Signif.codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Multiple R-squared: 0.9291, Adjusted R-squared:  0.9189
F-statistic:  91.7 on 1 and 7 DF,  p-value: 2.845e-05
```

〔5〕　表3, 4のように，入力データがわずかに変更された場合に推定結果が大きく異なる状況について論評せよ。

問 2　基準時点を s 時点，比較時点を t 時点とする指数を一般に $I_{s,t}$ と表す。多くの国で消費者物価指数 (CPI) に利用されている Laspeyres (ラスパイレス) 指数は「上位レベル」の指数算式と呼ばれるもので,「下位レベル」の指数をさらに合成する方法である。下位レベルの指数算式としては Carli (カルリ) 指数 $C_{s,t}$, Dutot (デュト) 指数 $D_{s,t}$, Jevons (ジェヴォンズ) 指数 $J_{s,t}$ が代表的である。ある下位レベル品目の時点 t における第 k 店舗の価格を p_{tk} と表すとき，各指数の定義は以下のとおりである。なお品目によって店舗数 m は異なる。

$$C_{s,t} = \frac{1}{m}\sum_{k=1}^{m}\left(\frac{p_{tk}}{p_{sk}}\right), \quad D_{s,t} = \frac{\sum_k p_{tk}}{\sum_k p_{sk}}, \quad J_{s,t} = \prod_{k=1}^{m}\left(\frac{p_{tk}}{p_{sk}}\right)^{1/m}$$

〔1〕　一般に $J_{s,t} \leq C_{s,t}$ となることを示せ。

〔2〕　価格指数 $I_{s,t}$ が任意の時点 s，t に対して $I_{s,t}\,I_{t,s} = 1$ を満たすことを「時点の対称性」と呼ぶ。Carli 指数, Dutot 指数, Jevons 指数のそれぞれについて,「時点の対称性」が成立するかどうかを判定せよ。

〔3〕　$s = 0$ を基準時点として，ある下位レベル算式を用いて作成した t 時点の第 i 品目の価格指数を改めて p_{ti} と記し，これらを合成する上位指数を作成する。このとき，基準時点の価格は $p_{0i} = 1$ となるが，t 時点の第 i 品目の支出金額 v_{ti} が $v_{ti} = p_{ti}q_{ti}$ $(t = 0, 1, \ldots; i = 1, \ldots, n)$ となるように消費者が購入する数量 q_{ti} の単位を定める。これらの記号を用いると，t 時点の Laspeyres 価格指数は $P_L = \dfrac{\sum_i p_{ti}q_{0i}}{\sum_i p_{0i}q_{0i}}$ と定義される。

同様にして Paasche (パーシェ) 価格指数 P_P，Laspeyres 数量指数 Q_L，Paasche 数量指数 Q_P の定義式を記せ。

〔4〕　以下では基準時点を $s = 0$ と固定して，簡単のために $x_i = p_{ti}/p_{0i}$, $y_i = q_{ti}/q_{0i}$ と表す。また，t 時点の支出金額合計を $\mathrm{V}_t = \displaystyle\sum_i p_{ti}q_{ti}$，支出構成比を $w_{ti} = v_{ti}/\mathrm{V}_t$ と表す。したがって $\displaystyle\sum_i w_{ti} = 1$ である。基準時点の w_{0i} を単に w_i と書く。

(4-1) w_i をウェイトとする x, y の加重平均 $\bar{x} = \displaystyle\sum_i w_i x_i$, $\bar{y} = \displaystyle\sum_i w_i y_i$ は，それぞれ，Laspeyres 価格指数 P_L および Laspeyres 数量指数 Q_L と一致することを示せ。

(4-2) 次の関係式を示せ。

$$\sum_i w_i x_i y_i = P_P \cdot Q_L = P_L \cdot Q_P$$

(4-3) w_i をウェイトとする x，y の分散を $s_x^2 = \sum_i w_i(x_i - \bar{x})^2$, $s_y^2 = \sum_i w_i(y_i - \bar{y})^2$，共分散を $s_{xy} = \sum_i w_i(x_i - \bar{x})(y_i - \bar{y})$，相関係数を $r = s_{xy}/(s_x\, s_y)$ とする。このとき，次の関係を導け。

$$\frac{P_P - P_L}{P_L} = r\,\frac{s_x}{P_L}\,\frac{s_y}{Q_L}$$

(4-4) 消費者の行動と P_L と P_P の大小関係について，どのような解釈ができるかを記せ。

問 3　ある業種に属する n 企業について，第 i 企業の第 t 年における投資額を y_{it}，企業価値や資本ストックなど p 個の説明変数の列ベクトルを $\boldsymbol{x}_{it}$ とする。以下で $\boldsymbol{x}'$ はベクトル $\boldsymbol{x}$ の転置を表す。なお，説明変数に定数項は含まない。

投資額を説明するモデルとして，定数項を μ，個別効果を α_i，係数の p 次元列ベクトルを $\boldsymbol{\beta}$ とする次式を想定する。

$$y_{it} = \mu + \alpha_i + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \epsilon_{it} \qquad (i = 1, \ldots, n; t = 1, \ldots, T) \tag{1}$$

誤差項 ϵ_{it} については $\mathrm{E}[\epsilon_{it}] = 0$, $\mathrm{E}[\epsilon_{it}^2] = \sigma_\epsilon^2$，$\mathrm{E}[\epsilon_{it}\epsilon_{i't}] = 0$ $(i \neq i')$，かつ説明変数 $\{\boldsymbol{x}_{it}\}$ と独立とする。また，異なる時点のすべての確率変数は互いに独立と仮定する。以下では，次の記号を用いる。

$$\boldsymbol{y}_i = \begin{pmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{pmatrix}, \quad \boldsymbol{X}_i = \begin{pmatrix} \boldsymbol{x}'_{i1} \\ \vdots \\ \boldsymbol{x}'_{iT} \end{pmatrix}, \quad \boldsymbol{\epsilon}_i = \begin{pmatrix} \epsilon_{i1} \\ \vdots \\ \epsilon_{iT} \end{pmatrix}$$

また，これらの変数の時間に関する平均を次のように定義する。

$$\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}, \quad \bar{\boldsymbol{x}}_i = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{x}_{it}, \quad \bar{\epsilon}_i = \frac{1}{T}\sum_{t=1}^{T} \epsilon_{it} \tag{2}$$

1 を並べた T 次元列ベクトルを $\boldsymbol{\iota}_T$ で表すと，(1) 式，(2) 式は次のように書くこともできる。

$$\boldsymbol{y}_i = (\mu + \alpha_i)\boldsymbol{\iota}_T + \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{\epsilon}_i; \quad \bar{y}_i = \frac{1}{T}\boldsymbol{\iota}'_T\boldsymbol{y}_i, \quad \bar{\boldsymbol{x}}'_i = \frac{1}{T}\boldsymbol{\iota}'_T\boldsymbol{X}_i, \quad \bar{\epsilon}_i = \frac{1}{T}\boldsymbol{\iota}'_T\boldsymbol{\epsilon}_i$$

以下の各問に答えよ。

〔1〕　個別効果 α_i を $\sum_{i=1}^{n}\alpha_i = 0$ を満たす未知の定数，または $\mathrm{E}[\alpha_i] = 0, \mathrm{V}[\alpha_i] = \sigma_\alpha^2$ かつ ϵ_{it} と独立な確率変数（ただし，α_i と説明変数 $\boldsymbol{x}_{it}$ の相関は許容する）と想定するモデルを固定効果モデルと呼ぶ。回帰係数 $\boldsymbol{\beta}$ は (1) 式と時間平均 $\bar{y}_i = \mu + \alpha_i + \bar{\boldsymbol{x}}'_i\boldsymbol{\beta} + \bar{\epsilon}_i$ との差を取って μ と α_i を消去した式から推定することができる。この推定量 $\hat{\boldsymbol{\beta}}_0$ を以上の記号を用いて表せ。

〔2〕　前問の固定効果モデルにおいて，$p = 2$，$\hat{\boldsymbol{\beta}}'_0 = (3, 2)$ となった。時間に関する平均 $\bar{y}_i$ と $\bar{\boldsymbol{x}}_i$ の企業平均が $(1/n)\sum_{i=1}^{n}\bar{y}_i = 10.0$，$(1/n)\sum_{i=1}^{n}\bar{\boldsymbol{x}}'_i = (2.5, 1.0)$ であり，特に第 j 企業の値が $\bar{y}_j = 12.0$，$\bar{\boldsymbol{x}}'_j = (2.0, 1.5)$ とするとき，第 j 企業の個別効果 α_j の推定値を求めよ。

〔3〕　個別効果 $\alpha_1, \ldots, \alpha_n$ を確率変数とし，次の仮定をおくモデルを変量効果モデルと呼ぶ。

$\alpha_1, \ldots, \alpha_n$ は平均 0，分散 σ_α^2 を持ち互いに独立，かつ $\boldsymbol{\epsilon}_i$ および説明変数 $\boldsymbol{x}_{it}$ と独立

このとき $\boldsymbol{u}_i = \alpha_i \boldsymbol{\iota}_T + \boldsymbol{\epsilon}_i$ の期待値（T 次元ベクトル）$\mathrm{E}[\boldsymbol{u}_i]$ と $(T \times T)$ 分散共分散行列 $\mathrm{V}[\boldsymbol{u}_i]$ を求め，$\Omega = \mathrm{V}[\boldsymbol{u}_i]$ は i によらず一定であることを示せ。

〔4〕　変量効果モデル $\boldsymbol{y}_i = \boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{u}_i$ における回帰係数 $\boldsymbol{\beta}$ の一般化最小二乗 (GLS) 推定量 $\hat{\boldsymbol{\beta}}_1$ を明示せよ。ただし σ_α^2 と σ_ϵ^2 の一致推定量 $\hat{\sigma}_\alpha^2$ と $\hat{\sigma}_\epsilon^2$ が与えられているものとする。

〔5〕　変量効果モデルにおいて，個別効果と説明変数は無相関であるという仮説 $H_0 : \mathrm{E}[\alpha_i \boldsymbol{x}_i] = 0 \ (i = 1, \ldots, n)$ が正しいときは，上問〔4〕で求めた $\hat{\boldsymbol{\beta}}_0, \hat{\boldsymbol{\beta}}_1$ のいずれも $n \to \infty$ のとき，真の $\boldsymbol{\beta}$ に近づく（一致性を持つ）ことが知られている。

仮説 H_0 が正しくないとき，$\hat{\boldsymbol{\beta}}_0, \hat{\boldsymbol{\beta}}_1$ のそれぞれが一致性を持つかどうか論ぜよ。

〔6〕　変量効果モデルにおいて，仮説 H_0 を検定する「ハウスマン検定」の考え方を記せ。

問 4　ある店舗での接客時間は期待値 μ の指数分布（$Exp(1/\mu)$ と表す）に従うという。$Exp(1/\mu)$ に従う確率変数を X としたとき，X の確率密度関数は

$$f(x) = \begin{cases} \dfrac{1}{\mu}e^{-x/\mu} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

である。$Exp(1/\mu)$ からの互いに独立な大きさ n の無作為標本を $X_1, \ldots, X_n$ としたとき，μ の最尤推定量 $\hat{\mu}$ は標本平均 $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ で与えられる。その店舗である日の午前中に接客された10名の客の接客時間（単位：分）は表1のようであった。以下の各問に答えよ。

表1：接客時間

客	1	2	3	4	5	6	7	8	9	10
接客時間	10	18	12	9	15	4	14	6	5	27

〔1〕　表1のデータから μ の最尤推定値とその標準誤差を求めよ。

〔2〕　この店舗のマネージャーに聞いたところ，接客時間が30分を超える客は別室で対応するとのことで，表1のデータはすべて30分以内で対応した客のみであった。一般に，全部で n 人の来訪者中で，接客時間が30分以内であった人数は m 人で各接客時間は $x_1, \ldots, x_m$ であり，30分を超えて別室で対応した客の数は $n-m$ 人で彼らの具体的な接客時間は分からないとする。このときの μ の最尤推定値 $\tilde{\mu}$ を導出せよ。そして，表1のデータに加えて別室対応者が2名いたときの $\tilde{\mu}$ を求めよ。

〔3〕　上問〔2〕の設定において，別室で対応した客の人数が分からないとしたときの μ の最尤推定値 μ^* を求めるアルゴリズムを示せ（実際に計算する必要はない）。

問 5　2 種類の処置 A，B の効果の差を調べるため，被験者 30 人をランダムに 15 人ずつの 2 群に分け，片方には処置 A を施し，もう片方には処置 B を施して経過を観察した。実験結果は正規分布に従う連続量のスコアで表され，処置 A では $x_1, \ldots, x_{15}$，処置 B では $y_1, \ldots, y_{15}$ が観測された。各処置での標本平均

$$\bar{x} = \frac{1}{15}\sum_{i=1}^{15} x_i, \quad \bar{y} = \frac{1}{15}\sum_{i=1}^{15} y_i$$

と標本標準偏差

$$s_x = \sqrt{\frac{1}{14}\sum_{i=1}^{15}(x_i - \bar{x})^2}, \quad s_y = \sqrt{\frac{1}{14}\sum_{i=1}^{15}(y_i - \bar{y})^2}$$

および，t 分布に基づいた処置 A でのスコアの母平均 μ_A の信頼係数 95% の信頼区間の下限と上限が表 1 に示されている。以下の各問に答えよ。

表 1：実験結果の要約

				95% 信頼区間	
	n	平均	標準偏差	下限	上限
処置 A	15	27.0	8.0	22.57	31.43
処置 B	15	20.0	10.0	（ア）	（イ）

〔1〕　処置 B でのスコアの母平均 μ_B の信頼係数 95% の信頼区間の下限（ア）及び上限（イ）を求めよ。

〔2〕　両処置群のスコアの母分散は等しいが未知であるとして，帰無仮説 $H_0 : \mu_A = \mu_B$，対立仮説 $H_1 : \mu_A \neq \mu_B$ の有意水準 5% の 2 標本両側 t 検定を実行し，その結果を述べよ。また，母平均の差 $\mu_A - \mu_B$ の信頼係数 95% の信頼区間を求めよ。

〔3〕　上問〔1〕で求めたような処置群ごとの母平均の信頼係数 95% の信頼区間の重なりと，上問〔2〕の 2 標本両側 t 検定の結果との関係を調べる。簡単のため，両群でのサンプルサイズは等しくともに n とし，標本標準偏差は等しい，$s_x = s_y (= s)$，とする。このとき，信頼区間の重なりがなければ 2 標本両側 t 検定は有意になるかを答えよ。また，信頼区間の重なりがあった場合，その重なりがいくら以下のとき 2 標本両側 t 検定は有意になるかを答えよ。

理工学

問 1　4 つの処理 A_1，A_2，A_3，A_4 を次のように 5 つのブロック B_1，B_2，B_3，B_4，B_5 に配置して 1 回ずつ実験を行う。

B_1	B_2	B_3	B_4	B_5
A_1	A_1	A_1	A_2	A_1
A_2	A_2	A_3	A_3	A_2
A_3	A_4	A_4	A_4	A_3
				A_4

処理 A_i がブロック B_j で実験されたときの観測値を表す確率変数を Y_{ij} とし，ブロック計画のモデル

$$Y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij} \quad (i = 1, ..., 4; j = 1, ..., 5)$$

を考える。ここで，μ は一般平均，τ_i は処理 A_i の処理効果，β_j はブロック B_j のブロック効果，ϵ_{ij} は観測誤差を表す確率変数である ($\tau_1+...+\tau_4 = 0$, $\beta_1+...+\beta_5 = 0$)。誤差 ϵ_{ij} は互いに独立であるとし, 期待値と分散をそれぞれ $\mathrm{E}[\epsilon_{ij}] = 0$，$\mathrm{V}[\epsilon_{ij}] = \sigma^2$ とする。以下の各問に答えよ。

〔1〕　確率変数 T_1，T_2，T_3 を

$$T_1 = Y_{11} - Y_{21}, \quad T_2 = Y_{12} - Y_{22}, \quad T_3 = Y_{15} - Y_{25}$$

としたとき，これらはいずれも $\tau_1 - \tau_2$ の不偏推定量であることを確かめよ。また，それらの分散 $\mathrm{V}[T_1]$，$\mathrm{V}[T_2]$，$\mathrm{V}[T_3]$ はいくらか。

〔2〕　確率変数 Y_{13}，Y_{33}，Y_{43}，Y_{24}，Y_{34}，Y_{44} の線形結合の形の $\tau_1 - \tau_2$ の不偏推定量のうち分散が最小となる推定量 T_4 を構成し，その分散 $\mathrm{V}[T_4]$ を求めよ。

〔3〕　$w > 0$ として，

$$T = \frac{T_1 + T_2 + T_3 + wT_4}{3 + w}$$

が $\tau_1 - \tau_2$ の不偏推定量であることを確かめ，$\mathrm{V}[T]$ を導出せよ。また，$\mathrm{V}[T]$ が最小となる w を求め，そのときの T と $\mathrm{V}[T]$ を求めよ。

〔4〕　上問〔3〕で求めた推定量 T の最適性について論ぜよ。

問 2　次数 2 の自己回帰過程 (AR(2))

$$Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + \epsilon_t \tag{1}$$

を考える。ここで a_1，a_2 は自己回帰係数であり，自己回帰過程 (1) は定常であるとする。したがって Y_t の確率分布はすべての t について同一である。ϵ_t は，互いに独立に平均 0，分散 σ^2 をもつ白色ガウスノイズである。自己回帰過程 (1) は定常であるので，一般線形過程

$$Y_t = \sum_{j=0}^{\infty} g_j \epsilon_{t-j} \tag{2}$$

と表すこともできる。ただし $g_0 = 1$ である。以下の各問に答えよ。

〔1〕　式 (2) に基づき，式 (1) の自己回帰過程 Y_t の期待値 $\mathrm{E}[Y_t]$ を求めよ。

〔2〕　式 (2) に基づき，式 (1) の自己回帰過程 Y_t の自己共分散 $C(s) = Cov[Y_t, Y_{t-s}]$ を求めよ。

〔3〕　式 (1) の両辺に Y_{t-1} をかけて期待値をとり，自己共分散と自己回帰係数の関係式を導け。

〔4〕　式 (1) の両辺に Y_{t-2} をかけて期待値をとり，自己共分散と自己回帰係数の関係式を導け。

〔5〕　上問〔3〕と〔4〕の結果を行列の式としてまとめることで，ユール・ウォーカー方程式が得られることを示せ。

問 3　製品を製造する工程の能力を測る指標に工程能力指数 (C_P) がある。製品特性の母集団分布を正規分布 $N(\mu, \sigma^2)$ とし，下側規格を S_L，上側規格を S_U とし，製品特性値が区間 (S_L, S_U) の外に出た場合はその製品は不良品とする。母標準偏差 σ が既知のときは

$$C_P = \frac{S_U - S_L}{6\sigma} \tag{1}$$

とする。母標準偏差 σ が未知の場合には，母集団から大きさ n の無作為標本 $X_1, ..., X_n$ を得て，標本平均および標本標準偏差を

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,\ S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

とし，

$$\hat{C}_P = \frac{S_U - S_L}{6S} \tag{2}$$

とする。以下の各問に答えよ。

〔1〕　母標準偏差 σ が既知の一定値であり，S_U, S_L が与えられたとすると，工程平均 μ が $(S_U + S_L)/2$ に一致するときに不良率は最小になることを説明せよ。

〔2〕　上問〔1〕の条件が成立ち，式 (1) で定義される工程能力指数が $C_P = 2/3$ のときの不良率を求めよ。

〔3〕　$\sum_{i=1}^{n}(X_i - \bar{X})^2/\sigma^2$ は自由度 $n-1$ のカイ二乗分布に従うことを用い，C_P の信頼係数 95% の信頼区間が，式 (2) で定義される $\hat{C}_P$ を用いて

$$\left(\hat{C}_P\sqrt{\frac{\chi^2_{0.975}(n-1)}{n-1}},\ \hat{C}_P\sqrt{\frac{\chi^2_{0.025}(n-1)}{n-1}}\right)$$

で与えられることを示せ。ここで $\chi^2_\alpha(\nu)$ は自由度 ν のカイ二乗分布の上側 100α% 点である。

〔4〕　ある製品の下側および上側規格値は $S_L = 12.0$cm，$S_U = 12.6$cm であり，実際に 20 個の製品の特性を計測したところ，標本標準偏差は 0.05cm であった。このときの $\hat{C}_P$ はいくらか。また，C_P の信頼係数 95% の信頼区間を求めよ。これらの数値を基に，C_P は 1.33 以上が望ましいという議論があることを踏まえ，この製造工程の管理状態をどう解釈するかを述べよ。

〔5〕　式 (2) で定義される $\hat{C}_P$ の期待値は，ある定数 a を用いて $\mathrm{E}[\hat{C}_P] = aC_P$ と表される。a の値を求めよ。ただし，自由度 ν のカイ二乗分布に従う確率変数 Y の確率密度関数は，$y > 0$ のとき

$$f(y) = \frac{1}{2\Gamma(\nu/2)}\left(\frac{y}{2}\right)^{(\nu/2)-1} e^{-y/2}$$

で与えられる。ここで $\Gamma(\nu/2)$ はガンマ関数である。

問 4　ある会社の情報部門には，会社で使っている全パーソナル・コンピュータ (PC) のうちで不具合が生じた PC が修理に持ち込まれる。1 日に持ち込まれる PC の数はパラメータ λ のポアソン分布（$Poisson\ (\lambda)$ と表す）に従っているとする。なお，パラメータ λ のポアソン分布に従う確率変数を X としたとき，その確率関数は

$$p(x) = \mathrm{P}(X = x) = \frac{\lambda^x}{x!}e^{-\lambda} \quad (x = 0, 1, 2, ...)$$

である。以下の各問に答えよ。

〔1〕　$Poisson(\lambda)$ のモーメント母関数は

$$M_X(t) = \mathrm{E}[e^{tX}] = \exp[\lambda(e^t - 1)]$$

であることを示し，これを用いて X の期待値 $\mathrm{E}[X]$ と分散 $\mathrm{V}[X]$ を求めよ。

〔2〕　$Poisson(\lambda)$ の最頻値，すなわち $Poisson(\lambda)$ の確率の最大値を与える x を求めよ。

〔3〕　過去 50 日間に修理に持ち込まれた PC の数 (X) とその日数は以下の表のようであった。このデータから λ の最尤推定値を求めよ。

X	0	1	2	3	4	5	6 以上	計
日数	8	15	12	10	3	2	0	50

〔4〕　修理に持ち込まれた PC の数が 0 の日は，何日あったか分からないとする。この場合の確率分布は，正の値のみを取るポアソン分布であり，その分布に従う確率変数を Y とすると確率関数は

$$f(y) = \mathrm{P}(Y = y) = \frac{\lambda^y e^{-\lambda}/y!}{1 - e^{-\lambda}} \quad (y = 1, 2, ...)$$

である。Y のモーメント母関数 $M_Y(t)$ および期待値 $\mathrm{E}[Y]$ を求めよ。また，この場合のパラメータ λ の推定法を論ぜよ。

問 5　2種類の処置A，Bの効果の差を調べるため，被験者30人をランダムに15人ずつの2群に分け，片方には処置Aを施し，もう片方には処置Bを施して経過を観察した。実験結果は正規分布に従う連続量のスコアで表され，処置Aでは $x_1, \ldots, x_{15}$，処置Bでは $y_1, \ldots, y_{15}$ が観測された。各処置での標本平均

$$\bar{x} = \frac{1}{15}\sum_{i=1}^{15} x_i, \quad \bar{y} = \frac{1}{15}\sum_{i=1}^{15} y_i$$

と標本標準偏差

$$s_x = \sqrt{\frac{1}{14}\sum_{i=1}^{15}(x_i - \bar{x})^2}, \quad s_y = \sqrt{\frac{1}{14}\sum_{i=1}^{15}(y_i - \bar{y})^2}$$

および，t 分布に基づいた処置Aでのスコアの母平均 μ_A の信頼係数95%の信頼区間の下限と上限が表1に示されている。以下の各問に答えよ。

表1：実験結果の要約

				95%信頼区間	
	n	平均	標準偏差	下限	上限
処置A	15	27.0	8.0	22.57	31.43
処置B	15	20.0	10.0	（ア）	（イ）

〔1〕　処置Bでのスコアの母平均 μ_B の信頼係数95%の信頼区間の下限（ア）及び上限（イ）を求めよ。

〔2〕　両処置群のスコアの母分散は等しいが未知であるとして，帰無仮説 $H_0 : \mu_A = \mu_B$，対立仮説 $H_1 : \mu_A \neq \mu_B$ の有意水準5%の2標本両側 t 検定を実行し，その結果を述べよ。また，母平均の差 $\mu_A - \mu_B$ の信頼係数95%の信頼区間を求めよ。

〔3〕　上問〔1〕で求めたような処置群ごとの母平均の信頼係数95%の信頼区間の重なりと，上問〔2〕の2標本両側 t 検定の結果との関係を調べる。簡単のため，両群でのサンプルサイズは等しくともに n とし，標本標準偏差は等しい，$s_x = s_y (= s)$，とする。このとき，信頼区間の重なりがなければ2標本両側 t 検定は有意になるかを答えよ。また，信頼区間の重なりがあった場合，その重なりがいくら以下のとき2標本両側 t 検定は有意になるかを答えよ。

医薬生物学

問1　ある臨床研究において，希少疾患の患者7名に対して薬剤Aによる治療を行い，治療の前後に実施したある臨床検査の測定値を表1にまとめた。差 z は臨床検査値の変化量（治療後 − 治療前）である。以下の各問に答えよ。

表1．臨床検査値の測定値と変化量

患者ID	1	2	3	4	5	6	7	平均	標準偏差
治療前	1.1	1.5	1.6	1.8	1.9	2.1	1.4	1.63	0.34
治療後	2.3	7.1	4.5	3.2	2.8	1.4	2.2	3.36	1.91
差 z	1.2	5.6	2.9	1.4	0.9	−0.7	0.8	1.73	2.01

〔1〕　表1の差 z のデータに対して，有意水準を両側5%とした対応のある t 検定を適用し，有意性を判定せよ。この検定における帰無仮説 H_0 は何かを述べ，この検定に通常置かれる仮定を示せ。

〔2〕　表1の差 z のデータに対し，ノンパラメトリック法である符号検定と符号付き順位検定を適用する。これらの検定における帰無仮説 H_0 はそれぞれ何かを述べ，これらの検定に通常置かれる仮定を示せ。

〔3〕　符号検定は，差 z の符号が正であるものの数 T に基づく検定である。総患者数を n とするとき，上問〔2〕で示された帰無仮説 H_0 の下での T の期待値 $\mathrm{E}[T|H_0]$ および分散 $\mathrm{V}[T|H_0]$ を示せ。

〔4〕　符号検定のP値の計算には，検定統計量を標準正規分布で近似して求める方法（正規近似法）と近似を用いずに直接求める方法（exact法）が存在する。上の表のデータにおける符号検定のexactな両側P値を計算せよ。

〔5〕　符号付き順位検定は，差 z の絶対値を元に順位付けした際の，差の符号が正であるものの順位の和 W^+ に基づく検定である。上問〔2〕で示された帰無仮説 H_0 の下での期待値 $\mathrm{E}[W^+|H_0]$ および分散 $\mathrm{V}[W^+|H_0]$ を示せ。実際に，表1のデータに対して，有意水準両側5%の符号付き順位検定（正規近似法）を適用し，有意性を判定せよ。

問 2　ある疾患に対する二重盲検ランダム化並行群間比較試験において，処置が 4 種類あり，治療の結果は有効あるいは無効のいずれかで判定されるとすると，有効および無効の各度数は表 1 の形式に要約される。4 種類の処置群の母有効率 $\pi_i (i = 1, 2, 3, 4)$ に関する帰無仮説

$$H_0 :\ \pi_1 = \pi_2 = \pi_3 = \pi_4\ (= \pi) \tag{1}$$

の検定を考える。ただし，π は未知とする。以下の各問に答えよ。

表 1．データの要約

	有効	無効	合計
群 1	y_1	$n_1 - y_1$	n_1
群 2	y_2	$n_2 - y_2$	n_2
群 3	y_3	$n_3 - y_3$	n_3
群 4	y_4	$n_4 - y_4$	n_4
合計	y	$N - y$	N

〔1〕　第 i 群の母有効率 π_i の推定量を $p_i = y_i/n_i$ とし $(i = 1, 2, 3, 4)$，帰無仮説 (1) における共通の有効率 π の推定量を $p = y/N$ とする。表 1 の分割表の記号を用いて，帰無仮説 (1) の検定のための検定統計量 T を示し，各度数が大きいとき，その検定統計量が帰無仮説の下で近似的に従う確率分布は何かを述べよ。

〔2〕　上問〔1〕の検定の P 値は 0.077 であった。検定の有意水準が 5% であるとき，この結果からいえることを文章で表現せよ。

〔3〕　実は 4 つの処置群はある薬剤の用量群であり，第 i 群の用量を x_i としたとき，$x_1 < x_2 < x_3 < x_4$ という順序があることがわかった。このため，母有効率について

$$\pi_i = \alpha + \beta x_i \quad (i = 1, 2, 3, 4) \tag{2}$$

という単回帰モデルを考えた。モデル (2) における β の最小二乗推定量が $\bar{x} = \sum_{i=1}^{4} n_i x_i / N$ として

$$\hat{\beta} = \frac{\sum_{i=1}^{4} n_i (p_i - p)(x_i - \bar{x})}{\sum_{i=1}^{4} n_i (x_i - \bar{x})^2}$$

となることを示し，その標準誤差 $SE[\hat{\beta}|H_0]$ を求めよ。また，これらを用いて帰無仮説 (1) を検定する検定統計量 Y を示せ。

〔4〕　上問〔3〕の Y を用いた検定の結果，P 値が 0.042 であり，検定の有意性について上問〔2〕と異なる結果を得た。この臨床試験の結果の解釈について述べよ。

問 3　生存時間データの解析に関する以下の各問に答えよ。

〔1〕　2 つの群のそれぞれに対し，表 1 のような生存時間データが得られたとする。このとき群 1 と群 2 の生存曲線をそれぞれカプラン・マイヤー法によって描け。

表 1：生存時間データ

群	観察時間（週）	打ち切りの有無（1:打ち切り, 0:イベント）
1	3	1
1	5	1
1	7	1
1	9	0
1	10	0
2	1	0
2	2	0
2	4	1
2	6	0
2	8	0

〔2〕　表 2 は，イベントが発生した時点ごとの各群のイベントの発生数とリスク集団の人数の推移である。2 群の生存曲線間に差がないとの仮定の下で，イベントの発生数と期待イベント発生数を比較する手法としてログランク検定がある。ログランク検定では，どちらかの群のイベント発生数と期待イベント発生数の差をイベント発生時点ごとに計算し，単純に合計したものをスコアとしている。また，各イベント発生時点でのイベント数が 1 例の場合には，その群の期待イベント発生確率をパラメータとする二項分布の分散をイベント発生時点ごとに計算し，合計したものがスコアの分散となる。表 2 におけるログランク検定のカイ二乗統計量と P 値を求めよ。

表 2：イベントの発生

観察時間（週）	1	2	6	8	9	10
群 1 のイベント発生数	0	0	0	0	1	1
群 2 のイベント発生数	1	1	1	1	0	0
群 1 のリスク集団の人数	5	5	3	2	2	1
群 2 のリスク集団の人数	5	4	2	1	0	0
リスク集団の総数	10	9	5	3	2	1

〔3〕　表 1 のデータに対して，群 1 の場合 0，群 2 の場合 1 となるような共変量 x を用いて Cox 回帰モデルを当てはめることを考える。この場合のハザード関数は以下のようになる。

$$h(x,t) = h_0(t)e^{\beta x}$$

このとき，パラメータ β を推定するための部分尤度 L を求めよ。

問 4　薬物の血中濃度の時間曲線下面積 (AUC) の分布は正の歪みを持ち，経験的に対数正規分布で近似できることが知られている。そこで，被験者 i の AUC を表す確率変数 X_i は対数正規分布 $LN(\mu, \sigma^2)$ に従う，すなわち，$Y_i = \log(X_i)$ が正規分布 $N(\mu, \sigma^2)$ に従うものと仮定する。ここで，log は自然対数を表す。以下の各問に答えよ。

〔1〕　X_i の中央値を求めよ。

〔2〕　X_i の確率密度関数は

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

で与えられることを示せ。

〔3〕　X_i の期待値と分散を求めよ。ただし，正規分布 $N(\mu, \sigma^2)$ に従う確率変数 Y のモーメント母関数が $M_{Y(t)} = \mathrm{E}[\exp(tY)] = \exp[\mu t + \sigma^2 t^2/2]$ となることを用いてよい。

〔4〕　B 薬を被験者 i に投与したときの AUC を X_i^B とし, X_i^B は対数正規分布 $LN(\mu, \sigma^2)$ に従うと仮定する。B 薬の薬物動態試験を計画することを考える。過去の類似薬の結果から，AUC の変動係数が 1 (100%) であると期待されるとき，σ^2 の見積もり値を求めよ。

〔5〕　10 名の被験者に対して 2 剤 2 期のクロスオーバー試験による生物学的同等性試験を行った。A 薬を既存薬とし，B 薬を後発品として，被験者 i のそれぞれの薬剤の AUC を X_i^A および X_i^B とし，これらはそれぞれ対数正規分布に従うとする。すなわち，$Y_i^A = \log(X_i^A) \sim N(\mu_A, \sigma_A^2)$ および $Y_i^B = \log(X_i^B) \sim N(\mu_B, \sigma_B^2)$ である。いま $Y_i = Y_i^A - Y_i^B$ とし，$Y_i (i = 1, ..., 10)$ の標本平均が $\bar{Y} = 0.1$ ，標本標準偏差が $s = 0.5$ と計算されたものとする。このとき，比

$$\frac{A\text{ 薬の AUC の中央値}}{B\text{ 薬の AUC の中央値}}$$

の信頼係数 90% の信頼区間を求め，A 薬と B 薬の生物学的同等性が成立するかを検討せよ。ただし，同等域は (0.8, 1.25) とする。

問 5　2 種類の処置 A，B の効果の差を調べるため，被験者 30 人をランダムに 15 人ずつの 2 群に分け，片方には処置 A を施し，もう片方には処置 B を施して経過を観察した。実験結果は正規分布に従う連続量のスコアで表され，処置 A では $x_1, \dots, x_{15}$，処置 B では $y_1, \dots, y_{15}$ が観測された。各処置での標本平均

$$\bar{x} = \frac{1}{15}\sum_{i=1}^{15} x_i, \quad \bar{y} = \frac{1}{15}\sum_{i=1}^{15} y_i$$

と標本標準偏差

$$s_x = \sqrt{\frac{1}{14}\sum_{i=1}^{15}(x_i - \bar{x})^2}, \quad s_y = \sqrt{\frac{1}{14}\sum_{i=1}^{15}(y_i - \bar{y})^2}$$

および，t 分布に基づいた処置 A でのスコアの母平均 μ_A の信頼係数 95% の信頼区間の下限と上限が表 1 に示されている。以下の各問に答えよ。

表 1：実験結果の要約

				95% 信頼区間	
	n	平均	標準偏差	下限	上限
処置 A	15	27.0	8.0	22.57	31.43
処置 B	15	20.0	10.0	（ア）	（イ）

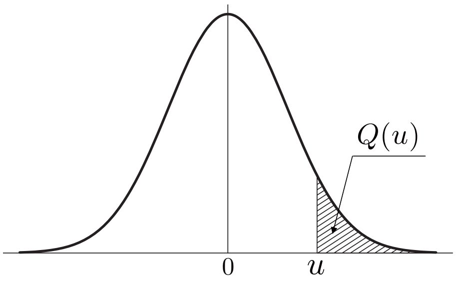
〔1〕　処置 B でのスコアの母平均 μ_B の信頼係数 95% の信頼区間の下限（ア）及び上限（イ）を求めよ。

〔2〕　両処置群のスコアの母分散は等しいが未知であるとして，帰無仮説 $H_0 : \mu_A = \mu_B$，対立仮説 $H_1 : \mu_A \neq \mu_B$ の有意水準 5% の 2 標本両側 t 検定を実行し，その結果を述べよ。また，母平均の差 $\mu_A - \mu_B$ の信頼係数 95% の信頼区間を求めよ。

〔3〕　上問〔1〕で求めたような処置群ごとの母平均の信頼係数 95% の信頼区間の重なりと，上問〔2〕の 2 標本両側 t 検定の結果との関係を調べる。簡単のため，両群でのサンプルサイズは等しくともに n とし，標本標準偏差は等しい，$s_x = s_y (= s)$，とする。このとき，信頼区間の重なりがなければ 2 標本両側 t 検定は有意になるかを答えよ。また，信頼区間の重なりがあった場合，その重なりがいくら以下のとき 2 標本両側 t 検定は有意になるかを答えよ。

付　表

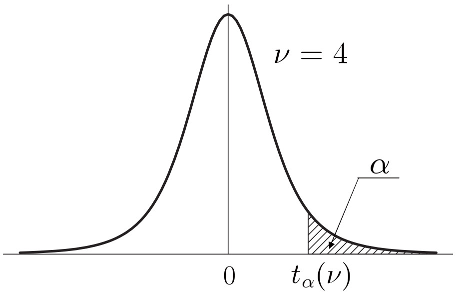
付表 1. 標準正規分布の上側確率



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。

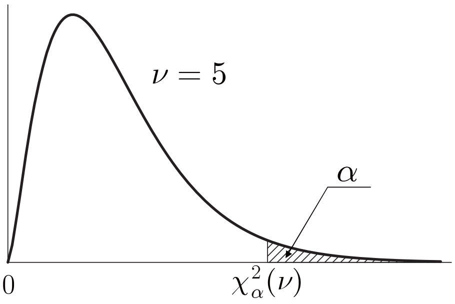
付表 2. t 分布のパーセント点



	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。

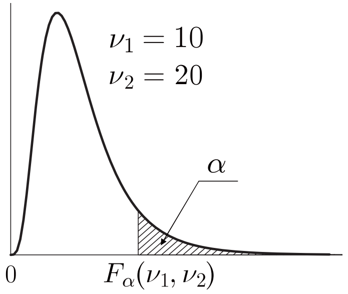
付表 3. カイ二乗分布のパーセント点



	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。

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**

2016.11