2014 年 RSS/JSS 試験（Higher Certificate）

HIGHER CERTIFICATE IN STATISTICS, 2014

モジュール 6：統計学の発展的応用

制限時間： 90 分

4 問中 3 問を選択の上解答のこと．

各問は合計 20 点である．小問の配点は括弧の中に記されている．

グラフ用紙と統計数値表は配布する．

解答にあたっては電卓を使用してよい．
ただし，一般財団法人統計質保証推進協会による「受験要領」に
記された範囲で使用すること．

数学記号 log は **e** を底とする自然対数を表す．
その他の底をもつ対数は，例えば $\log_{10}$ のように底を明示する．

また，$\binom{n}{r}$ は ${}_nC_r$ と同じ意味とする．

問題用紙は 8 頁からなり，**それぞれの頁は片面にのみ印刷されている**．
この表紙が 1 頁目である．
第 1 問は 2 頁目から始まる．

問題は全部で 4 問である．

© JSS 2014

1. ある作業の実行時間について 4 種類の異なる方法（A, B, C, D）の影響を調べるために，作業の技量が異なる 8 人の被験者で乱塊法の実験が行われた． 計測された実行時間（分）は次のようになった．

	被験者								
方法	1	2	3	4	5	6	7	8	総和
A	32	30	32	27	27	32	31	36	247
B	29	38	33	31	36	41	36	35	279
C	30	40	32	30	37	35	32	34	270
D	31	36	30	31	33	33	37	32	263
総和	122	144	127	119	133	141	136	137	1059

$\Sigma y^2 = 35{,}407$

(i) これらのデータに適切な線型モデルを書け．仮定も明確に述べよ．

(3)

(ii) 分散分析表を作成せよ．方法によって平均実行時間が異なるかどうか調べたい．適切な帰無仮説と対立仮説を述べ，有意水準 5% で検定せよ．検定の結論も明確に述べよ．

(9)

(iii) それぞれの方法について平均実行時間の 95% 信頼区間を求めよ．これらの信頼区間が表すことを説明せよ．(ii)で述べた結論に対して，信頼区間が何を補足するかについて説明せよ．

(4)

(iv) この実験で問題になる可能性がある点について述べ，それを改良する方法についてコメントせよ．

(4)

2. (a) いくつかの重回帰モデルをデータ点 16 個，説明変数 x_1 ，x_2 のデータセットに当てはめた．説明変数の次数 2 までのすべての適切な項の組み合わせのモデルが当てはめられて，次のような残差平方和の表が得られた．

説明変数	残差平方和	残差の自由度
$x_1, x_2, x_1^2, x_2^2, x_1x_2$	47.8	10
x_1, x_2, x_1^2, x_2^2	63.1	11
x_1, x_2, x_1^2, x_1x_2	55.2	11
x_1, x_2, x_2^2, x_1x_2	61.4	11
x_1, x_2, x_1^2	64.3	12
x_1, x_2, x_2^2	71.8	12
x_1, x_2, x_1x_2	66.6	12
x_1, x_1^2	81.2	13
x_2, x_2^2	113.9	13
x_1, x_2	89.4	13
x_1	112.4	14
x_2	135.6	14
定数項のみ	168.7	15

後退消去法（変数消去法）を有意水準 5% の検定とともに用いて，最適な回帰モデルを選択せよ．選択結果が得られる過程と結果の正当性について詳細に説明せよ．

(10)

(b) 説明変数 x_1, x_2 が連続変数の場合を考える．それらの目的変数 y への影響を調べるために実験を計画した．説明変数毎に低，中，高の 3 水準を設定し，2 要因の水準のすべて組み合わせで，それぞれ 2 回ずつ観測を行う．

このデータを分析するために，ある研究者は交互作用をもつ 2 元配置分散分析を使うべきだと主張するが，別の研究者は x_1, x_2 の実際の値を用いて変数の 2 乗の項と変数の積の項を含む重回帰分析を行うべきだと主張する．

ダミー変数という概念を説明せよ．ダミー変数を用いることにより，どのようにして 2 元配置分散分析モデルが回帰モデルで表されるかを示せ．

これら 2 つの分析方法を比較対照し，それぞれの方法について用いるモデルと必要な仮定を明確に述べよ．

とくに，柔軟性，自由度，予測への有用性，という観点から 2 つのアプローチの長所と短所を考察せよ．

(10)

3. 異なる種類の薬と運動療法の効果を調べるため，変性状態の患者に対する試験を行う．

試験で効果を調べる薬は 5 種類で，4 種類の新薬と対照群となるプラセボである．試験で効果を調べる運動療法は 4 種類で，3 種類の指定された形式の運動と対照群となる自由な運動である．すべての患者は男女ともに近い年齢である．

この実験を例として次の用語を説明せよ．試験がどのように計画され，どうしてこれらの概念が重要か，一般の場合と今回のケースについて概要を述べよ．とくに，これらの考え方を適用する利点を述べ，利点となる理由の概要を述べよ．実験の考え方を説明できるならば，試験結果の分析についても述べよ．ただし，分析の詳細は不要である．

(i) 要因計画

(5)

(ii) 交互作用

(3)

(iii) 盲検化と二重盲検化

(4)

(iv) 無作為化

(4)

(v) ブロック化

(4)

4. ある会社は機械部品を作っていて，過去の経験から次のことがわかっている．製造上もっとも重要で難しいのは，部品が機械にきちんと取り付けられるように正しい長さで作ることである．

もし長さが 788mm から 792mm の間であれば, 部品は適切に取り付けられる．もしそうでなければ，部品は取り付けられず不良品とみなされる．

(i) その会社の現在の品質管理法は，定期的に 100 個の部品を標本抽出して不良品の割合を計算することである．シューハート管理図の基本的な原理を述べ，今回のケースで不良品の割合が 10% を超えないために管理図をどのように用いるか，どのように警告限界と要処置限界を与えるか説明せよ．

(7)

(ii) その会社は品質管理法の変更を検討しており，4 個の部品を標本抽出して長さの平均値を計算することにした．部品の長さが正規分布に従い，平均 790mm，標準偏差は最大 1.21mm であることが望ましいとき，どのように平均長の管理図を用いるか，どのように警告限界と要処置限界を与えるか説明せよ．

(6)

(iii) 各部品の長さが平均 791mm，標準偏差 1mm の正規分布に従うと仮定する．(i)の場合に標本が要処置限界の外側に出る確率と，(ii)の場合に標本が要処置限界の外側に出る確率を計算せよ．

(4)

(iv) 一般のケースと今回のケースについて，(i)と(ii)で与えた二つの方法の長所と短所を簡潔に比較対照せよ．

(3)

BLANK PAGE

BLANK PAGE

BLANK PAGE