RSS Higher Certificate Statistics, 2010

Module 6 : Further applications of statistics

1. 工業プロセスからの生産量データが集められた．その方法は，標準的な方法 S，または他 4 つの代替方法 A,B,C,D のいずれかによって行われた．25 の実験装置を利用し，実験の順序は完全に無作為に割り付けされ，各方法は反復数が異なっていた．以下の表は，適当な単位での出力 y の結果である．

	Total	*Mean*
S: 24.2, 27.7, 21.8, 30.6, 24.9, 29.3, 28.4, 26.7	213.6	26.7
A: 31.8, 37.7, 32.9, 40.2, 42.4	185.0	37.0
B: 35.7, 40.9, 50.4, 43.8	170.8	42.7
C: 40.2, 39.5, 46.3, 44.8	170.8	42.7
D: 30.6, 33.3, 36.2, 37.5	137.6	34.4
	877.8	

$\Sigma y^2 = 32\,186.08$

(i) このデータの分散分析表を作成し，5 つの生産量平均間に違いがある証拠があるかどうか検定せよ．この分析の基礎となる線形モデルと，その各項が満たすべき仮定を書け．

(ii) 標準的な方法 S と方法 A からの生産量平均の差に対する 95%信頼区間を求めよ．

(iii) 完全無作為実験では，観測された値と，同じ処理を受けた全ユニットの平均値との差が，ユニットに対する残差となる．たとえば，S の平均は 26.7 だから，S を行ったユニットに対する残差は $-2.5, 1.0, -4.9, 3.9, -1.8, 2.6, 1.7, 0.0$ である．

A,B,C,D に対する残差を計算せよ．

(iv) すべての残差をプロットせよ．ただし，5 つの方法を別の行に分けて描け．このプロットを用いて，(i)で述べた仮定についてコメントせよ．

2. 以下の各問に答えよ．

(i) 3 変数 X, Y, Z の観測値の組 (x_i, y_i, z_i) を n 組用意し $(i = 1, 2, 3, \ldots, n)$，これらにモデル $Y = \beta X + \gamma Z$ （定数項なし）をあてはめる．

β と γ の最小二乗推定量を求めよ．

(ii) (a) (i)の結果を用いて，X と Y の間に 2 次関係 $Y = \beta X + \gamma X^2$ を回帰せよ．

(b) いま，(a)での 2 次関係が原点を通るべきか否か，いくらか疑念がある．重回帰のコンピュータプログラムが使えることを仮定する．このプログラムを用いて，(a) でのモデルや定数項αも含んでいるモデルへの回帰を比較せよ．この比較のため，使用するの

に適切な診断方法を提案せよ．

3. 3つのレベル（L1, L2, L3）の栄養素を生育期間中4回（T1, T2, T3, T4）にわたって作物に与える農業実験を行った．実験では，36の単位区画が無作為に3つのブロックにふり分けられた．

下の表は，シーズンの終わりに記録された作物の収穫量yの概要である．各レベル，各時刻の組み合わせにおける，3ブロックの**合計**が与えられている．

3ブロックの合計収穫量

		Level			
		L1	L2	L3	*Time total*
Time	T1	42	54	70	166
	T2	51	62	83	196
	T3	76	84	98	258
	T4	76	89	106	271
Level total		245	289	357	891

各36個の収穫量の平方和は$\sum y^2 = 23649$．ブロックごとの合計は265, 294, 332.

(i) T,Lの組み合わせ12組全てにおける平均(または合計)を表す適当なグラフを作成せよ．

(ii) 以下の分散分析表をうめて完成させよ．

SOURCE OF VARIATION	DEGREES OF FREEDOM	SUM OF SQUARES	MEAN SQUARE	F VALUE
Blocks	2	***	***	***
T	***	***	***	***
L	***	***	***	***
T × L	***	***	***	***
Treatments	***	1375.42		
Residual	***	***	***	
TOTAL	35	1596.75		

(iii) 必要な統計的検定を行って結論を正当化し，分析の結果を報告せよ．

4. 包装食製品の重量をチェックする通常のプロセスにおいて，何バッチ分もの製品を調べる2つの方式が提案されている．経験的に，梱包品の10%が指定された許容限界から外れている（“不良品”である）ことがわかっている．

方法Ｉでは，$n = 24$個のサンプルを一回抜き取り，3個以上が不良品ならばそのバッチは不合

格とする．

方法Ⅱは二回抜取検査方式で，n = 12 個をサンプルとして用いる．1 回目のサンプルに不良品がないならばバッチ全体を合格として，3 個以上あれば不合格とする．1 回目のサンプルに不良品が 1 個か 2 個ある場合は，2 回目のサンプルを 12 個とり，両サンプルの不良品数の合計を考える．この合計が 2 個以下ならばロットは合格, 3 個以上ならばバッチは不合格である．

(i) 方法Ｉを採用したときの不合格となる確率を求めよ．

(ii) 方法Ⅱにおける以下の各確率を求めよ．
(a) 1 回目のサンプルを抜き取るだけでバッチが合格となる確率.
(b) 1 回目のサンプルを抜き取るだけでバッチが不合格となる確率.
(c) サンプルを 2 回抜き取ってから不合格となる確率.

(iii) 方法Ⅱを採用したときにバッチが不合格となる全体の確率を求め，(i)の確率と比較せよ．

(iv) (ii)の結果を用いて，方法Ⅱを採用した際に 2 回目のサンプルが必要となる確率，および方法Ⅱにおけるサンプルサイズの平均を求めよ．
サンプルサイズや不合格の確率に基づいて，どちらの方式が好ましいと考えられるかを説明せよ．