

Microsoft Excelによる分散分析アプリの開発

福森 貢¹⁾, 田平 一行¹⁾

¹⁾ 畿央大学健康科学部理学療法学科 (〒635-0832 奈良県北葛城郡広陵町馬見中4-2-2)

Developing ANOVA Applications for Microsoft Excel

Mitsugu FUKUMORI¹⁾, Kazuyuki TABIRA¹⁾

¹⁾ Department of Physical Therapy, Faculty of Health Sciences, Kio University
(4-2-2 Umami-naka, Koryo-cho, Kitakatsuragi-gun, Nara 635-0832, Japan)

研究の目的

これまでに学部生のための各種教育用統計アプリを開発してきたが、分散分析に関しては下位検定を含んでいなかった。それは単純主効果と多重比較に関する内容である。

Microsoft Excel（以下Excelと表記）で利用できる市販の統計アプリ及びネット上で使用できる統計アプリを調べたが、単純主効果に関して初心者が直観的に使用できるものを確認することができなかった。また、市販の統計アプリでも二元配置分散分析後の単純主効果の分析を行うには追加作業や繰り返し対応ありとなしの一元配置分散分析を再度実施するなど統計の初心者にとっては容易なことではない。

以上のことから本研究では学部生が日常的に使用しているExcelのみで効率よく下位検定までが可能な統計アプリの開発を目標とした。

研究方法

1. 目標

(1) Excelの数式と関数のみを使用

Excel VBA (Visual Basic for Applications) を使用することで汎用的で柔軟な統計アプリを開発できるが、2022年2月よりマイクロソフト社のセキュリティ強化方針により、ネット上からVBAを含んだExcelファイルをダウンロードする場合にマクロ動作がブロックされるようになった。このブロックを解除するには設定を変更する知識が必要となる。そこでパソコン初心者に配慮して特別な設定を必要としないVBAを含まない統計アプリとした。

(2) 直観的なユーザーインターフェース

Excelシートにデータをコピーまたは直接入力することで、下位検定の結果までが出力されるように煩雑

な操作を必要としない直観的なユーザーインターフェースを目指した。

2. 開発環境

- (1) Thinkpad X1 Tablet (gen2)
- (2) Office365 (バージョン2308)

3. 統計アプリの種類

それぞれ個別のExcelファイルとして用意した。

- (1) 一元配置分散分析 (対応なし)
- (2) 一元配置分散分析 (対応あり)
- (3) 二元配置分散分析 (対応なし×対応なし)
- (4) 二元配置分散分析 (対応なし×対応あり)
- (5) 多重比較 (Bonferroni, Holm, Shaffer)

4. 動作の概要

下位検定を必要とする二元配置分散分析 (対応なし×対応あり) の例を示す。

(1) 分散分析表

表1の分散分析表に従って分散比を求め、この分散比をもとにしてp値を算出している。

表1 分散分析表

要因	偏差平方和	自由度	不偏分散	分散比
要因A変動 (行間変動)	S_A	$f_A = I - 1$	$V_A = S_A / f_A$	$F_A = V_A / V_S$
要因B変動 (列間変動)	S_B	$f_B = k - 1$	$V_B = S_B / f_B$	$F_B = V_B / V_E$
個体変動	S_S	$f_S = (r - 1) \times I$	$V_S = S_S / f_S$	
A × B変動 (交互作用)	S_{AB}	$f_{AB} = (I - 1)(k - 1)$	$V_{AB} = S_{AB} / f_{AB}$	$F_{AB} = V_{AB} / V_E$
誤差変動 (残差変動)	S_E	$f_E = I \times k(r - 1) - f_S$	$V_E = S_E / f_E$	
総変動	S	$f = I \times k \times r - 1$		

I : 因子Aの水準数 k : 因子Bの水準数 r : 1要因あたりのデータ数

(2) 単純主効果検定

要因Aの対応ありのデータを抽出して一元配置分散分析（対応あり）を、要因Bの対応なしのデータを抽出して一元配置分散分析（対応なし）を行う。

(3) 多重比較

3種類の「多重比較」の結果を出力する。

結果

二元配置分散分析（対応なし×対応あり）の場合を示す。

1. 例題（その1）

糖尿病患者にブドウ糖負荷試験を行い、経過時間ごとに血糖値を測定した。時間経過により血糖値に変化があるかを調べたい。（交互作用が出るように数値を調整済み、単位は省略）

表2 例題（その1）の測定データ

性別	投与前	30分後	60分後	90分後	120分後
男性	70	120	140	145	156
	92	131	158	164	170
	73	156	179	188	191
女性	103	164	183	160	120
	99	161	178	165	140
	65	137	153	140	120

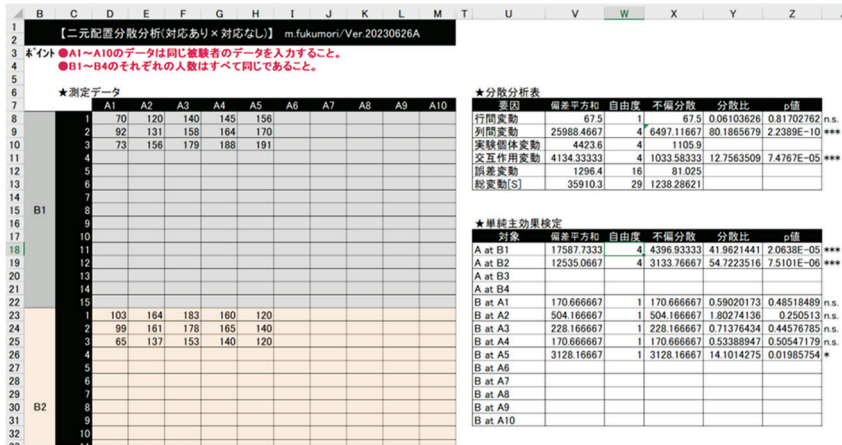


図1 データの入力と結果の出力

(1) データの入出力

図1の左エリアにデータを入力またはコピーする。右エリアに分散分析表と単純主効果検定の結果が出力される。対応ありの要因Aの水準は10、対応なしの要因Bの水準は4、1要因あたりのデータの数は15の制約があるが、データの数300の統計ソフトも用意した。

(2) 分散分析表

図2に分散分析の出力結果を示す。交互作用が認められるため下位検定を行う。

(3) 単純主効果検定の結果

図3に単純主効果検定の出力結果を示す。上部の2行の数値は一元配置分散分析（対応あり）と同等の結果を、その下の5行の数値は一元配置分散分析（対応なし）と同等の結果を出力する。

(4) 多重比較の入出力

最大6群の比較とデータの数200の制約がある。3種類の多重比較の結果が各シートに同時に結果が出力される。図4は男性のデータの入力画面を、図5は多重比較のひとつであるBonferroniの出力画面を示している。例題では5群の比較をしているため5群の表に結果が出力される。

要因	偏差平方和	自由度	不偏分散	分散比	p値	
行間変動	67.5	1	67.5	0.06103626	0.81702762	n.s.
列間変動	25988.4667	4	6497.11667	80.1865679	2.2389E-10	***
実験個体変動	4423.6	4	1105.9			
交互作用変動	4134.33333	4	1033.58333	12.7563509	7.4767E-05	***
誤差変動	1296.4	16	81.025			
総変動[S]	35910.3	29	1238.28621			

図2 分散分析の出力結果

対象	偏差平方和	自由度	不偏分散	分散比	p値	
A at B1	17587.7333	4	4396.93333	41.9621441	2.0638E-05	***
A at B2	12535.0667	4	3133.76667	54.7223516	7.5101E-06	***
A at B3						
A at B4						
B at A1	170.666667	1	170.666667	0.59020173	0.48518489	n.s.
B at A2	504.166667	1	504.166667	1.80274136	0.250513	n.s.
B at A3	228.166667	1	228.166667	0.71376434	0.44576785	n.s.
B at A4	170.666667	1	170.666667	0.53388947	0.50547179	n.s.
B at A5	3128.16667	1	3128.16667	14.1014275	0.01985754	*
B at A6						
B at A7						
B at A8						
B at A9						
B at A10						

図3 単純主効果検定の結果

	A	B	C	D	E	F	G	H	I
1	【多重比較】m.fukumori/Ver.20230620A								
2									
3		n	3	3	3	3	3	0	
4			A1	A2	A3	A4	A5	A6	
5		1	70	120	140	145	156		
6		2	92	131	158	164	170		
7		3	73	156	179	188	191		
8		4							
9		5							
10		6							
11		7							
12		8							
33		29							
34		30							
35		31							
36		32							
37		33							
38		34							
39		35							
40		36							
41		37							
42		38							
43		39							
44		40							
45		41							
46		42							
47		43							
48		44							
49		45							
50		46							
51		47							
52		48							
53		49							
54		50							
55		51							
56		52							
57		53							
58		54							
59		55							
60		56							
61		57							
62		58							
63		59							
64		60							
65		61							
66		62							
67		63							
68		64							
69		65							
70		66							
71		67							
72		68							
73		69							
74		70							
75		71							
76		72							
77		73							
78		74							
79		75							
80		76							
81		77							
82		78							
83		79							
84		80							
85		81							
86		82							
87		83							
88		84							
89		85							
90		86							
91		87							
92		88							
93		89							
94		90							
95		91							
96		92							
97		93							
98		94							
99		95							
100		96							
101		97							
102		98							
103		99							
104		100							
105		101							
106		102							
107		103							
108		104							
109		105							
110		106							
111		107							
112		108							
113		109							
114		110							
115		111							
116		112							
117		113							
118		114							
119		115							
120		116							
121		117							
122		118							
123		119							
124		120							
125		121							
126		122							
127		123							
128		124							
129		125							
130		126							
131		127							
132		128							
133		129							
134		130							
135		131							
136		132							
137		133							
138		134							
139		135							
140		136							
141		137							
142		138							
143		139							
144		140							
145		141							
146		142							
147		143							
148		144							
149		145							
150		146							
151		147							
152		148							
153		149							
154		150							
155		151							
156		152							
157		153							
158		154							
159		155							
160		156							
161		157							
162		158							
163		159							
164		160							
165		161							
166		162							
167		163							
168		164							
169		165							
170		166							
171		167							
172		168							
173		169							
174		170							
175		171							
176		172							
177		173							
178		174							
179		175							
180		176							
181		177							
182		178							
183		179							
184		180							
185		181							
186		182							
187		183							
188		184							
189		185							
190		186							
191		187							
192		188							
193		189							
194		190							
195		191							
196		192							
197		193							
198		194							
199		195							
200		196							
201		197							
202		198							
203		199							
204		200							
205		201							
206		202							
207		203							
208		204							
209		205							
210		206							
211		207							
212		208							
213		209							
214		210							
215		211							
216		212							
217		213							
218		214							
219		215							
220		216							
221		217							
222		218							
223		219							
224		220							
225		221							
226		222							
227		223							
228		224							
229		225							
230		226							
231		227							
232		228							
233		229							
234		230							
235		231							
236		232							
237		233							
238		234							
239		235							
240		236							
241		237							
242		238							
243		239							
244		240							
245		241							
246		242							
247		243							
248		244							
249		245							
250		246							
251		247							
252		248							
253		249							
254		250							
255		251							
256		252							
257		253							
258		254							
259		255							
260		256							

2. 例題（その2）

田平ゼミにて過去の卒業研究で測定したデータから動作確認に使用可能なものを選んだ。健常大学生にダイエットスリッパと通常のスリッパで運動負荷試験を行った。内腹斜筋の筋活動量の計測データをもとに分析してみる。

表3 例題（その2）の測定データ

種類	速度2km	速度3km	速度4km	速度5km	速度6km
ダイエット	5.98	5.39	7.14	7.38	11.13
	15.44	14.29	15.68	19.62	21.27
	9.37	12.05	13.29	14.48	19.69
	15.46	15.32	18.68	22.17	30.15
	9.07	14.84	10.38	19.29	28.23
	11.66	13.85	14.6	15.5	29.78
	12.87	13.99	18.39	28.06	34.99
通常	2.44	2.53	2.21	2.21	3.24
	4.82	5.73	7.38	9.33	9.32
	2.7	3.53	5.14	10.64	19.91
	4.78	7.51	9.98	12.01	13.5
	10	6.91	4.11	3.27	6.17
	3.76	6.58	4.46	6.37	19.97
	6.07	3.09	3.6	5.33	15.1

(1) 分散分析表

① 本統計アプリの出力結果

② SPSSの結果

図7のソースのスリッパの出力結果と図6の本統計アプリの行間変動の出力結果は一致している。

図8のEMG球面性の仮定の出力結果と図6の本統計アプリの列間変動の出力結果は一致している。

要因	偏差平方和	自由度	不偏分散	分散比	p値	
行間変動	1516.18012	1	1516.18012	22.2921401	0.00049574	***
列間変動	1042.12184	4	260.530459	23.8085789	6.8236E-11	***
実験個体変動	816.169349	12	68.0141124			
交互作用変動	86.3981514	4	21.5995379	1.973874	0.11355901	n.s.
誤差変動	525.250251	48	10.9427136			
総変動[S]	3986.11971	69	57.7698508			

図6 分散分析表の出力結果

ソース	タイプ III 平方和	自由度	平均平方	F 値	有意確率	偏イータ 2 乗
切片	9446.596	1	9446.596	138.892	.000	.920
スリッパ	1516.180	1	1516.180	22.292	.000	.650
誤差	816.169	12	68.014			

図7 被験者間効果の検定結果

ソース		タイプ III 平方和	自由度	平均平方	F 値	有意確率	偏イータ 2 乗
EMG	球面性の仮定	1042.122	4	260.530	23.809	.000	.665
	Greenhouse-Geisser	1042.122	1.753	594.484	23.809	.000	.665
	Huynh-Feldt	1042.122	2.200	473.726	23.809	.000	.665
	下限	1042.122	1.000	1042.122	23.809	.000	.665
EMG * スリッパ	球面性の仮定	86.398	4	21.600	1.974	.114	.141
	Greenhouse-Geisser	86.398	1.753	49.286	1.974	.167	.141
	Huynh-Feldt	86.398	2.200	39.275	1.974	.155	.141
	下限	86.398	1.000	86.398	1.974	.185	.141
誤差 (EMG)	球面性の仮定	525.250	48	10.943			
	Greenhouse-Geisser	525.250	21.036	24.969			
	Huynh-Feldt	525.250	26.398	19.897			
	下限	525.250	12.000	43.771			

図8 被験者内効果の検定結果

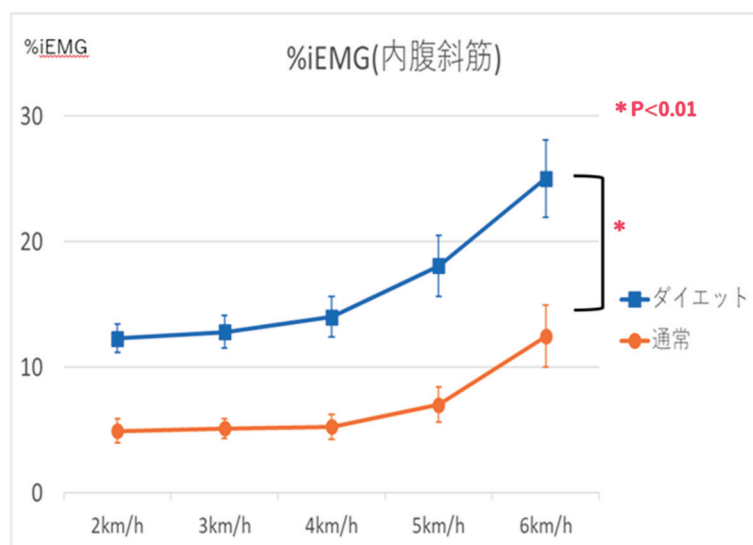


図9 2要因の平均値グラフ

例題（その2）では交互作用が認められなかったので、要因別により一元配置分散分析と多重比較を行うことになる。これらの結果については省略する。

考察

研究の目的に示したとおりExcelの基本操作のみで直観的に使用できる統計アプリを開発できたと考える。これまでの統計教育の経験から統計の初心者は、分散分析で交互作用が認められた場合の下位検定の方法に悩むことが多い。その仕組みを理解してもそれを実現するための環境を用意できないこともある。今回開発した統計アプリには制約はあるが十分実務に耐え

ることができるものとする。

結語

今後の課題は「二元配置分散分析（対応あり×対応あり）」を追加で開発することである。畿央大学の学部生が卒業研究や専門科目で分散分析が必要となったときに活用できるように、担当するゼミやデータサイエンスに関する科目等で案内していきたい。

参考文献

1. 渡辺利夫：フレッシュマンから大学院生までのデータ解析・R言語、ナカニシヤ出版
2. 永田靖：統計的多重比較法の基礎、サイエンティスト社
3. 竹原卓真：SPSSのススメ2要因の分散分析をすべてカバー、北大路書房
4. 石村貞夫：入門はじめての分散分析と多重比較、東京図書
5. 福森貢：看護・医療系データ分析のための基本統計ハンドブック、ピラールプレス
6. 井関龍太のページ、<http://riseki.php.xdomain.jp/>、2023年7月18日アクセス
7. Radi-toko、<https://radi-toko.com/multiple-comparison-test5/>、2023年7月21日アクセス