	患砂眼者	無砂眼者
男	246	1,001
女	210	1,071

	患砂眼者	無砂眼者
4—10歲	91	1,175
11—20歲	365	897

2. 下表爲各專家研究中華民族血屬之總結果(2)，問華北與華南人之血屬分配有差別否？

	血屬				
	O	A	B	AB	總人數
華北	2,244	1,978	2,179	590	6,991
華南	663	711	373	173	1,925

又汪美先氏(19)等研究血屬之結果如下，問能證實性連繫之遺傳(sex-linked inheritance)否？

男女血屬分布

	血屬			
	O	A	B	AB
男	1,075	760	646	219
女	463	353	234	123

3. 下表爲天花患者死亡與曾否種痘之關係(3)：試測驗種痘對於預防天花之實效．計算時可用公式(8-2)求得 χ^2，再查表 8-12B，

	痊愈	死亡
已種痘	20	2
未種痘	89	38

用內插法求得 P 値，最後計算 $\frac{1}{2}P$．在此測驗中，所包含之無效假設爲：在死亡人數方面，已種痘者較未種痘者並不顯著的少；而在痊愈人數方面，則已種痘者較未種痘者並不顯著的多．注意此假設之措詞．

4. 用第8-7節之直接計算法，求第3題之$\frac{1}{2}P$值.

附:

$$\log(109!)=176.159,5252$$

$$\log(127!)=213.478,9503$$

$$\log(149!)=260.580,8023$$

若用Yates校正數後，由表8-12求得之$\frac{1}{2}P$，則與直接計算法所得者較近，試與上題結果一併比較之.

5. 下表爲血屬與梅毒血淸反應之紀錄(16)，問二者有顯著的關係否?

血屬	Wassermann 氏反應 陽性	可疑者	陰性
O	53	26	261
A	75	17	208
B	91	26	238
AB	32	11	95

6. 就下列資料(17)推論三種哺乳方法之嬰兒死亡率有差別否?

	生存	死亡
母乳	1,827	143
奶媽	73	7
人工哺乳	45	20

7. 下表爲國立北平第一助產學校自民國二十四年至二十七年四季接生次數(1). 問嬰兒之出生數有顯著的季節變動(seasonal fluctuation)現象否?(春季係指陽曆三月至五月，餘類推.)

年分	春	夏	秋	冬
二十四年	520	498	520	521
二十五年	426	490	507	543
二十六年	553	459	276	326
二十七年	290	318	338	350

8. 下表為北平傳染病醫院自 1915 至 1923 年猩紅熱患者痊愈與死亡之紀錄(18). 問歷年猩紅熱之死亡率有顯著的差別否?

年分	痊愈者	死亡者	患者人數
1915	6	9	15
1916	152	35	187
1917	46	17	63
1918	6	1	7
1919	4	0	4
1920	12	1	13
1921	59	19	78
1922	109	33	142
1923	77	18	95
總計	471	133	604

表中患者人數最少之四年, 其理論次數將小於 5, 故須設法歸併, 可將連續三年合為一組, 如 1915—17, 1918—20, 1921—23 是.

9. 測驗第七章第 9 題配合常態曲線之適度.

【參考文獻】

(1) 張式溥: 北平學生砂眼之統計, 中華醫學雜誌, 第十八卷第

五期，803–805 頁.

(2) 李振翩：中國民族的血屬，中華醫學雜誌，第十六卷第一期，5–26 頁(1930).

(3) 四川省立第一行政區中心衛生院：溫江衛生工作報告，三十一年.

(4) Liu Bao-Chu, Report of an investigation of infant mortality and its causes in Pishan, The Chinese Medical Journal, 63A, 4 : 178–183 (1945).

(5) Taylor, G. L. and Prior, A. M., The distribution of the *M* and *N* factors in random samples of different races, Annals of Eugenics, 9, 2 : 97–108 (1939).

(6) Mainland, D., The Treatment of Clinical and Laboratory Data, Oliver and Boyd (19[illegible]8).

(7) Lung, Y. Y. & Foster, J. H., Typhoid and paratyphoid fever in Changsha, The National Medical Journal of China, 9, 3 : 185–197 (1923).

(8) Wu, C. and Hsieh, C. K., Roentgenological study of intestinal tuberculosis, The National Medical Journal of China, 16, 5 : 505–523 (1930).

(9) Landauer, E., The temperature factor in the efficacy of sodium cyanide as a larvicide, Trans. Ninth Congree, Far East. Assoc. Trop. Med., Nanking, 2 : 807–812 (1934).

(10) Yule, G. U., The Theory of Statistics, Charles Griffin

and Co. (1924).

(11) Fisher, R. A., Statistical Methods for Research Workers, Chapter IV, Sections 21·01, 21·03, Oliver and Boyd, London, 1936.

(12) Yates, F., Journ. Roy. Stat. Soc., Suppl. i., No. ii, 217 (1934).

(13) Bateson, W. and Punnett, R. C., Journal of Genetics, 1 : 297 (1911).

(14) Snedecor, G. W., Statistical Methods, pp. 372–377, The Iowa State College Press (1940).

(15) Elderton, W. P., Frequency Curves and Correlation, Charles & Edwin Layton (1927).

(16) Chue, C. Y. and Wang, S. H., A system for obtaining Chinese blood donors, Chinese Med. Journ., 46 : 31–42 (1932).

(17) Yang, M. & Yuan, I. C., Report of an investigation on infant mortality and its causes in Peiping, Chinese Medical Journal, 47 : 597–604 (1933).

(18) Yang, T. K. & Shih, W. H., Scarlet fever in China, National Med. Journ. of China, 10, 3 : 153–170 (1924).

(19) Wang, M. S., Chue, C. Y., and Chen, S. P., Blood groups of Chinese, The Chinese Medical Journal, 63A, 1 : 24–25 (1944).

第九章　變異數分析

兩組測量資料相比較時，我們用 t 值測驗其均數相差的顯著性，在本書第四章裏已詳細討論過了．若有兩組以上時，要比較幾個均數之間有無顯著的差別，那麼必須把 t 測驗的方法加以推廣，這個推廣的顯著性測驗法，叫做變異數分析（analysis of variance）．

變異數分析創於 R. A. Fisher 氏(1)，他用 z 值（此 z 值與本書第 6–6 節及第 7–3 節所討論者不同，勿混淆）的 5% 點與 1% 點作爲比較的標準(2, Table VI)，但因需求自然對數，故計算較繁，後經其他統計學家與以簡化(3)，(4)，只須求出兩變異數之商，查 F 值表(5, pp. 184-187)即可，其用法詳後．

9–1．大小相等的三組間之比較——依一個標準分類者　藍天鶴氏(6)在穀類豆類混合蛋白質之生理價值一研究中曾配合三種飼料，計 268 號爲玉米、小米、黃豆、雜合麵；269 號爲玉米、小麵、黃豆、雜合麵；274 號爲豆腐、麵筋、雜合麵．實驗動物，係鼠 18 頭，勻分爲三組，每組給以一種飼料．氏用氮平衡法測定各飼料之生理價值，以不屬於統計學範圍，故從略．此處僅錄其一星期內體重增加的情形，以說明計算的過程；眞正的飼養實驗至少需兩三個月纔能獲得相當結果．又食量與所增體重有密切關係，此點當在下章討論之．表 9–1 所列三組白鼠，在實驗設計時，對於性別、窩別、原始體重等應儘量

【表 9-1】 三組白鼠在一星期內增加之體重(克)

	飼料 268 號 (玉米、小米、黃豆、雜合麪)	飼料 269 號 (玉米、小麪、黃豆、雜合麪)	飼料 274 號 (豆腐、麪筋、雜合麪)	總計
X	13	3	19	
	6	7	20.5	
	15	1	10	
	10	5	16	
	8	8	18	
	23.5	1.5	18	
ΣX	75.5	25.5	101.5	202.5
$\bar{x}$	12.59	4.25	16.92	11.25
ΣX^2	1,146.25	150.25	1,785.25	3,081.75

使之相等. 若無其他因子,如食量多寡等,夾在裏面,則三組均數的差別,或可表示三種飼料營養價值的不同. 我們先看第一組白鼠,吃的飼料相同,但一週內所增的體重卻不相同,這種變異是與飼料的好壞無關的,其變異的大小可以離均差平方和表示之:

$$\Sigma(X_1-\bar{x}_1)^2=1,146.25-\frac{(75.5)^2}{6}=196.21.$$

該組各量數之和爲 75.5, 各量數平方之和爲 1,146.25, 載於表 9-1 的 ΣX 與 ΣX^2 兩行,離均差平方和係用公式(3-5)計算. 同理,第二、三兩組的離均差平方和爲:

$$\Sigma(X_2-\bar{x}_2)^2=150.25-\frac{(25.5)^2}{6}=41.87,$$

$$\Sigma(X_3-\bar{x}_3)^2=1,785.25-\frac{(101.5)^2}{6}=68.21.$$

這些變異也是與食料的好壞無關的. 因此這三數之和, 即

$$\Sigma\Sigma(X_i-\bar{x}_i)^2=196.21+41.87+68.21=306.29,$$

此值爲各組內部變異的總和, 而與食料的優劣無關的. 這個'組內'的變異可目爲機遇的變異, 亦可稱爲誤差.

此外還有一種變異, 即三個均數間的變異, 三組均數爲:

$$12.59 \qquad 4.25 \qquad 16.92$$

其變異可以各組均數與總均數 11.25 (此均數係 18 頭鼠之總平均, 即 202.5÷18) 之離均差平方和表示之, 惟因每組有鼠 6 頭, 故須各以 6 乘之:

$$6(12.59-11.25)^2+6(4.25-11.25)^2+6(16.92-11.25)^2=497.33,$$

此值所表示的變異至少有兩個來源, 一是由於食料的優劣, 一是由於機遇. 由於機遇而來的變異, 我們已在前面求得, 即 306.29. 若各'組間'的變異超過於'組內'的變異很多, 即均數間的變異超過了由於純粹機遇得來的可能, 假定其他因子都已被嚴密控制, 那麼這三個均數間的差別, 大概是飼料的優劣使然.

這裏值得注意的便是: '組內'與'組間'兩個離均差平方和相加恰等於十八頭鼠的總變異.

組內與組間兩離均差平方和之和:

$$306.29+497.33=803.62,$$

總變異: $\Sigma(X-\bar{x})^2=3,081.75-\dfrac{(202.5)^2}{18}=803.62.$

我們就可應用這個特性來簡化計算的過程. (一)先由上式求得所增體重的總變異爲 803.62. (二)再用下列公式求得各'組間'的變異:

$$\Sigma k(\bar{x}_i-\bar{x})^2=\frac{(\Sigma X_1)^2+(\Sigma X_2)^2+\cdots\cdots+(\Sigma X_n)^2}{k}-\frac{(\Sigma X)^2}{nk},$$

公式(9-1)

此處 k 爲每組內的項目數（此例爲動物數），n 爲組數．將例題之數值代入，得

$$\Sigma k(\bar{x}_i-\bar{x})^2=\frac{(75.5)^2+(25.5)^2+(101.5)^2}{6}-\frac{(202.5)^2}{18}$$

$$=2,775.46-2,278.13=497.33.$$

這樣計算'組間'離均差平方和，較前法簡捷而且正確．因爲求各組均數與總均數的相差時往往受小數四捨五入的影響而使結果稍有出入．如上例若各組均數與總均數都用兩位小數，其答案實爲 497.667，較眞正值略大（因此前面不用 497.667，而用 497.33），故計算時以用公式(9-1)爲宜．（三）最後在總變異內減去各'組間'的變異，卽得各'組內'的變異或誤差．其公式爲

$$\Sigma\Sigma(X_i-\bar{x}_i)^2=\Sigma(X-\bar{x})^2-\Sigma k(\bar{x}_i-\bar{x})^2. \qquad \text{公式(9-2)}$$

以數值代入此式，卽得

$$\Sigma\Sigma(X_i-\bar{x}_i)^2=803.62-497.33=306.29.$$

要測驗各組均數間相差的顯著性，我們不能把'組間'與'組內'兩離均差平方和直接比較，而必須顧到它們的自由度，好在自由度與離均差平方和同樣有可加的特性．試看'組間'的變異共有三個均數，求離均差平方和時用去一總均數，故自由度喪失一個，卽爲 $3-1=2$．第一組有鼠六頭，求組內離均差平方和時，用去一個本組均數，其自由度爲 $6-1=5$；同理，第二、三兩組的自由度亦各爲 5；故'組內'離

均差平方和的自由度為 5+5+5=15. 兩種自由度相加, 2+15=17, 恰等於總變異的自由度.

離均差平方和與自由度都求得後, 就可談到變異數分析的本身了. 玆列如下表:

【表 9-2】 變異數分析(三組白鼠所增之體重)

變異來源	自由度	離均差平方和	均方(變異數)	F
總變異	17	803.62		
各組之均數間	2	497.33	248.67	12.18
組內(誤差)	15	306.29	20.42	
$n_1=2$,	$n_2=15$,	1% 點=6.36.		

表內第一橫行是白鼠所增體重的總變異, 其自由度為 17, 離均差平方和為 803.62, 第二、三兩行為'組間'與'組內'的變異, 其自由度與離均差平方和都已在上面解釋過. 將第二、三兩行的自由度, 除同行的離均差平方和, 卽得均方, 亦卽變異數(參閱第 3-4 節),

$$497.33\div 2=248.67, \quad 306.29\div 15=20.42.$$

這時兩變異數可直接比較了. 比較的方法, 卽將'組內'的變異數除'組間'的變異數,

$$F=248.67\div 20.42=12.18.$$

所得的商稱為 F 值, 意卽'三個均數間的變異數'為'誤差變異數'的 12.18 倍. 究竟前者比後者要大多少倍, 纔有顯著的相差, 我們須查 F 值表(表 9-3)纔能知道. 因為 F 值是兩個變異數之比, 故有兩種自由度, 這兩種自由度, 在表 9-3 以縱橫二方向表示之. 該表上端橫行內的數值是較大均方的自由度, 稱為 n_1, 左側直行內的數值是較小

【表 9-3】 F 值表

【相當於機率 .05(上行)與 .01(下行)之 F 值，n_1 爲較大均方之自由度】

	1	2	3	4	5	6	7	8	9	10	11	12
1	161	200	216	225	230	234	237	239	241	242	243	244
	4052	4999	5403	5625	5764	5859	5928	5981	6022	6056	6082	6106
2	18.51	19.00	19.16	19.25	19.30	19.33	19.36	19.37	19.38	19.39	19.40	19.41
	98.49	99.01	99.17	99.25	99.30	99.33	99.34	99.36	99.38	99.40	99.41	99.42
3	10.13	9.55	9.28	9.12	9.01	8.94	8.88	8.84	8.81	8.78	8.76	8.74
	34.12	30.81	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23	27.13	27.05
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.93	5.91
	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.54	14.45	14.37
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.78	4.74	4.70	4.68
	16.26	13.27	12.06	11.39	10.97	10.67	10.45	10.27	10.15	10.05	9.96	9.89
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00
n_2	13.74	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.79	7.72
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.63	3.60	3.57
	12.25	9.55	8.45	7.85	7.46	7.19	7.00	6.84	6.71	6.62	6.54	6.47
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.34	3.31	3.28
	11.26	8.65	7.59	7.01	6.63	6.37	6.19	6.03	5.91	5.82	5.74	5.67
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.13	3.10	3.07
	10.56	8.02	6.99	6.42	6.06	5.80	5.62	5.47	5.35	5.26	5.18	5.11
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97	2.94	2.91
	10.04	7.56	6.55	5.99	5.64	5.39	5.21	5.06	4.95	4.85	4.78	4.71
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.86	2.82	2.79
	9.65	7.20	6.22	5.67	5.32	5.07	4.88	4.74	4.63	4.54	4.46	4.40
12	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.80	2.76	2.72	2.69
	9.33	6.93	5.95	5.41	5.06	4.82	4.65	4.50	4.39	4.30	4.22	4.16
13	4.67	3.80	3.41	3.18	3.02	2.92	2.84	2.77	2.72	2.67	2.63	2.60
	9.07	6.70	5.74	5.20	4.86	4.62	4.44	4.30	4.19	4.10	4.02	3.96

【相當於機率 .05(上行)與 .01(下行)之 F 值，n_1 爲較大均方之自由度】

	14	16	20	24	30	40	50	75	100	200	500	∞
1	245	246	248	249	250	251	252	253	253	254	254	254
	6142	6169	6208	6234	6258	6286	6302	6323	6334	6352	6361	6366
2	19.42	19.43	19.44	19.45	19.46	19.47	19.47	19.48	19.49	19.49	19.50	19.50
	99.43	99.44	99.45	99.46	99.47	99.48	99.48	99.49	99.49	99.49	99.50	99.50
3	8.71	8.69	8.66	8.64	8.62	8.60	8.58	8.57	8.56	8.54	8.54	8.53
	26.92	26.83	26.69	26.60	26.50	26.41	26.35	26.27	26.23	26.18	26.14	26.12
4	5.87	5.84	5.80	5.77	5.74	5.71	5.70	5.68	5.66	5.65	5.64	5.63
	14.24	14.15	14.02	13.93	13.83	13.74	13.69	13.61	13.57	13.52	13.48	13.46
5	4.64	4.60	4.56	4.53	4.50	4.46	4.44	4.42	4.40	4.38	4.37	4.36
	9.77	9.68	9.55	9.47	9.38	9.29	9.24	9.17	9.13	9.07	9.04	9.02
6	3.96	3.92	3.87	3.84	3.81	3.77	3.75	3.72	3.71	3.69	3.68	3.67
	7.60	7.52	7.39	7.31	7.23	7.14	7.09	7.02	6.99	6.94	6.90	6.88
7	3.52	3.49	3.44	3.41	3.38	3.34	3.32	3.29	3.28	3.25	3.24	3.23
	6.35	6.27	6.15	6.07	5.98	5.90	5.85	5.78	5.75	5.70	5.67	5.65
8	3.23	3.20	3.15	3.12	3.08	3.05	3.03	3.00	2.98	2.96	2.94	2.93
	5.56	5.48	5.36	5.28	5.20	5.11	5.06	5.00	4.96	4.91	4.88	4.86
9	3.02	2.98	2.93	2.90	2.86	2.82	2.80	2.77	2.76	2.73	2.72	2.71
	5.00	4.92	4.80	4.73	4.64	4.56	4.51	4.45	4.41	4.36	4.33	4.31
10	2.86	2.82	2.77	2.74	2.70	2.67	2.64	2.61	2.59	2.56	2.55	2.54
	4.60	4.52	4.41	4.33	4.25	4.17	4.12	4.05	4.01	3.96	3.93	3.91
11	2.74	2.70	2.65	2.61	2.57	2.53	2.50	2.47	2.45	2.42	2.41	2.40
	4.29	4.21	4.10	4.02	3.94	3.86	3.80	3.74	3.70	3.66	3.62	3.60
12	2.64	2.60	2.54	2.50	2.46	2.42	2.40	2.36	2.35	2.32	2.31	2.30
	4.05	3.98	3.86	3.78	3.70	3.61	3.56	3.49	3.46	3.41	3.38	3.36
13	2.55	2.51	2.46	2.42	2.38	2.34	2.32	2.28	2.26	2.24	2.22	2.21
	3.85	3.78	3.67	3.59	3.51	3.42	3.37	3.30	3.27	3.21	3.18	3.16

【相當於機率 .05(上行)與 .01(下行)之 F 值，n_1 爲較大均方之自由度】

	1	2	3	4	5	6	7	8	9	10	11	12
14	4.60	3.74	3.34	3.11	2.96	2.85	2.77	2.70	2.65	2.60	2.56	2.53
	8.86	6.51	5.56	5.03	4.69	4.46	4.28	4.14	4.03	3.94	3.86	3.80
15	4.54	3.68	3.29	3.06	2.90	2.79	3.70	2.64	2.59	2.55	2.51	2.48
	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.73	3.67
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.45	2.42
	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.61	3.55
17	4.45	3.59	3.20	2.96	2.81	2.70	2.62	2.55	2.50	2.45	2.41	2.38
	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.52	3.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34
	8.28	6.01	5.09	4.58	4.25	4.01	3.85	3.71	3.60	3.51	3.44	3.37
19	4.38	3.52	3.13	2.90	2.74	2.63	2.55	2.48	2.43	2.38	2.34	2.31
	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.36	3.30
20	4.35	3.49	3.10	2.87	2.71	2.60	2.52	2.45	2.40	2.35	2.31	2.28
	8.10	5.85	4.94	4.43	4.10	3.87	3.71	3.56	3.45	3.37	3.30	3.23
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25
	8.02	5.78	4.87	4.37	4.04	3.81	3.65	3.51	3.40	3.31	3.24	3.17
22	4.30	3.44	3.05	2.82	2.66	2.55	2.47	2.40	2.35	2.30	2.26	2.23
	7.94	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.18	3.12
23	4.28	3.42	3.03	2.80	2.64	2.53	2.45	2.38	2.32	2.28	2.24	2.20
	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.14	3.07
24	4.26	3.40	3.01	2.78	2.62	2.51	2.43	2.36	2.30	2.26	2.22	2.18
	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.25	3.17	3.09	3.03
25	4.24	3.38	2.99	2.76	2.60	2.49	2.41	2.34	2.28	2.24	2.20	2.16
	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.21	3.13	3.05	2.99
26	4.22	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15
	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.17	3.09	3.02	2.96

【相當於機率 .05(上行)與 .01(下行)之 F 值，n_1 爲較大均方之自由度】

	14	16	20	24	30	40	50	75	100	200	500	∞
14	2.48	2.44	2.39	2.35	2.31	2.27	2.24	2.21	2.19	2.16	2.14	2.13
	3.70	3.62	3.51	3.43	3.34	3.26	3.21	3.14	3.11	3.06	3.02	3.00
15	2.43	2.39	2.33	2.29	2.25	2.21	2.18	2.15	2.12	2.10	2.08	2.07
	3.56	3.48	3.36	3.29	3.20	3.12	3.07	3.00	2.97	2.92	2.89	2.87
16	2.37	2.33	2.28	2.24	2.20	2.16	2.13	2.09	2.07	2.04	2.02	2.01
	3.45	3.37	3.25	3.18	3.10	3.01	2.96	2.89	2.86	2.80	2.77	2.75
17	2.33	2.29	2.23	2.19	2.15	2.11	2.08	2.04	2.02	1.99	1.97	1.96
	3.35	3.27	3.16	3.08	3.00	2.92	2.86	2.79	2.76	2.70	2.67	2.65
18	2.29	2.25	2.19	2.15	2.11	2.07	2.04	2.00	1.98	1.95	1.93	1.92
	3.27	3.19	3.07	3.00	2.91	2.83	2.78	2.71	2.68	2.62	2.59	2.57
19	2.26	2.21	2.15	2.11	2.07	2.02	2.00	1.96	1.94	1.91	1.90	1.88
n_2	3.19	3.12	3.00	2.92	2.84	2.76	2.70	2.63	2.60	2.54	2.51	2.49
20	2.23	2.18	2.12	2.08	2.04	1.99	1.96	1.92	1.90	1.87	1.85	1.84
	3.13	3.05	2.94	2.86	2.77	2.69	2.63	2.56	2.53	2.47	2.44	2.42
21	2.20	2.15	2.09	2.05	2.00	1.96	1.93	1.89	1.87	1.84	1.82	1.81
	3.07	2.99	2.88	2.80	2.72	2.63	2.58	2.51	2.47	2.42	2.38	2.36
22	2.18	2.13	2.07	2.03	1.98	1.93	1.91	1.87	1.84	1.81	1.80	1.78
	3.02	2.94	2.83	2.75	2.67	2.58	2.53	2.46	2.42	2.37	2.33	2.31
23	2.14	2.10	2.04	2.00	1.96	1.91	1.88	1.84	1.82	1.79	1.77	1.76
	2.97	2.89	2.78	2.70	2.62	2.53	2.48	2.41	2.37	2.32	2.28	2.26
24	2.13	2.09	2.02	1.98	1.94	1.89	1.86	1.82	1.80	1.76	1.74	1.73
	2.93	2.85	2.74	2.66	2.58	2.49	2.44	2.36	2.33	2.27	2.23	2.21
25	2.11	2.06	2.00	1.96	1.92	1.87	1.84	1.80	1.77	1.74	1.72	1.71
	2.89	2.81	2.70	2.62	2.54	2.45	2.40	2.32	2.29	2.23	2.19	2.17
26	2.10	2.05	1.99	1.95	1.90	1.85	1.82	1.78	1.76	1.72	1.70	1.69
	2.86	2.77	2.66	2.58	2.50	2.41	2.36	2.28	2.25	2.19	2.15	2.13

【相當於機率 .05(上行)與 .01(下行)之 F 值，n_1 爲較大均方之自由度】

	1	2	3	4	5	6	7	8	9	10	11	12
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.30	2.25	2.20	2.16	2.13
	7.68	5.49	4.60	4.11	3.79	3.56	3.39	3.26	3.14	3.06	2.98	2.93
28	4.20	3.34	2.95	2.71	2.56	2.44	2.36	2.29	2.24	2.19	2.15	2.12
	7.64	5.45	4.57	4.07	3.76	3.53	3.36	3.23	3.11	3.03	2.95	2.90
29	4.18	3.33	2.93	2.70	2.54	2.43	2.35	2.28	2.22	2.18	2.14	2.10
	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.08	3.00	2.92	2.87
30	4.17	3.32	2.92	2.69	2.53	2.42	2.34	2.27	2.21	2.16	2.12	2.09
	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.06	2.98	2.90	2.84
32	4.15	3.30	2.90	2.67	2.51	2.40	2.32	2.25	2.19	2.14	2.10	2.07
	7.50	5.34	4.46	3.97	3.66	3.42	3.25	3.12	3.01	2.94	2.86	2.80
34	4.13	3.28	2.88	2.65	2.49	2.38	2.30	2.23	2.17	2.12	2.08	2.05
n_2	7.44	5.29	4.42	3.93	3.61	3.38	3.21	3.08	2.97	2.89	2.82	2.76
36	4.11	3.26	2.86	2.63	2.48	2.36	2.28	2.21	2.15	2.10	2.06	2.03
	7.39	5.25	4.38	3.89	3.58	3.35	3.18	3.04	2.94	2.86	2.78	2.72
38	4.10	3.25	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09	2.05	2.02
	7.35	5.21	4.34	3.86	3.54	3.32	3.15	3.02	2.91	2.82	2.75	2.69
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.07	2.04	2.00
	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.88	2.80	2.73	2.66
42	4.07	3.22	2.83	2.59	2.44	2.32	2.24	2.17	2.11	2.06	2.02	1.99
	7.27	5.15	4.29	3.80	3.49	3.26	3.10	2.96	2.86	2.77	2.70	2.64
44	4.06	3.21	2.82	2.58	2.43	2.31	2.23	2.16	2.10	2.05	2.01	1.98
	7.24	5.12	4.26	3.78	3.46	3.24	3.07	2.94	2.84	2.75	2.68	2.62
46	4.05	3.20	2.81	2.57	2.42	2.30	2.22	2.14	2.09	2.04	2.00	1.97
	7.21	5.10	4.24	3.76	3.44	3.22	3.05	2.92	2.82	2.73	2.66	2.60
48	4.04	3.19	2.80	2.56	2.41	2.30	2.21	2.14	2.08	2.03	1.99	1.96
	7.19	5.08	4.22	3.74	3.42	3.20	3.04	2.90	2.80	2.71	2.64	2.58

【相當於機率 .05(上行)與 .01(下行)之 F 值，n_1 爲較大均方之自由度】

	14	16	20	24	30	40	50	75	100	200	500	∞
27	2.08	2.03	1.97	1.93	1.88	1.84	1.80	1.76	1.74	1.71	1.68	1.67
	2.83	2.74	2.63	2.55	2.47	2.38	2.33	2.25	2.21	2.16	2.12	2.10
28	2.06	2.02	1.96	1.91	1.87	1.81	1.78	1.75	1.72	1.69	1.67	1.65
	2.80	2.71	2.60	2.52	2.44	2.35	2.30	2.22	2.18	2.13	2.09	2.06
29	2.05	2.00	1.94	1.90	1.85	1.80	1.77	1.73	1.71	1.68	1.65	1.64
	2.77	2.68	2.57	2.49	2.41	2.32	2.27	2.19	2.15	2.10	2.06	2.03
30	2.04	1.99	1.93	1.89	1.84	1.79	1.76	1.72	1.69	1.66	1.64	1.62
	2.74	2.66	2.55	2.47	2.38	2.29	2.24	2.16	2.13	2.07	2.03	2.01
32	2.02	1.97	1.91	1.86	1.82	1.76	1.74	1.69	1.67	1.64	1.61	1.59
	2.70	2.62	2.51	2.42	2.34	2.25	2.20	2.12	2.08	2.02	1.98	1.96
34	2.00	1.95	1.89	1.84	1.80	1.74	1.71	1.67	1.64	1.61	1.59	1.57
n_2	2.66	2.58	2.47	2.38	2.30	2.21	2.15	2.08	2.04	1.98	1.94	1.91
36	1.98	1.93	1.87	1.82	1.78	1.72	1.69	1.65	1.62	1.59	1.56	1.55
	2.62	2.54	2.43	2.35	2.26	2.17	2.12	2.04	2.00	1.94	1.90	1.87
38	1.96	1.92	1.85	1.80	1.76	1.71	1.67	1.63	1.60	1.57	1.54	1.53
	2.59	2.51	2.40	2.32	2.22	2.14	2.08	2.00	1.97	1.90	1.86	1.84
40	1.95	1.90	1.84	1.79	1.74	1.69	1.66	1.61	1.59	1.55	1.53	1.51
	2.56	2.49	2.37	2.29	2.20	2.11	2.05	1.97	1.94	1.88	1.84	1.81
42	1.94	1.89	1.82	1.78	1.73	1.68	1.64	1.60	1.57	1.54	1.51	1.49
	2.54	2.46	2.35	2.26	2.17	2.08	2.02	1.94	1.91	1.85	1.80	1.78
44	1.92	1.88	1.81	1.76	1.72	1.66	1.63	1.58	1.56	1.52	1.50	1.48
	2.52	2.44	2.32	2.24	2.15	2.06	2.00	1.92	1.88	1.82	1.78	1.75
46	1.91	1.87	1.80	1.75	1.71	1.65	1.62	1.57	1.54	1.51	1.48	1.46
	2.50	2.42	2.30	2.22	2.13	2.04	1.98	1.90	1.86	1.80	1.76	1.72
48	1.90	1.86	1.79	1.74	1.70	1.64	1.61	1.56	1.53	1.50	1.47	1.45
	2.48	2.40	2.28	2.20	2.11	2.02	1.96	1.88	1.84	1.78	1.73	1.70

【相當於機率 .05(上行)與 .01(下行)之 F 值，n_1 爲較大均方之自由度】

	1	2	3	4	5	6	7	8	9	10	11	12
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.02	1.98	1.95
	7.17	5.06	4.20	3.72	3.41	3.18	3.02	2.88	2.78	2.70	2.62	2.56
55	4.02	3.17	2.78	2.54	2.38	2.27	2.18	2.11	2.05	2.00	1.97	1.93
	7.12	5.01	4.16	3.68	3.37	3.15	2.98	2.85	2.75	2.66	2.59	2.53
60	4.00	3.15	2.76	2.52	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92
	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.56	2.50
65	3.99	3.14	2.75	2.51	2.36	2.24	2.15	2.08	2.02	1.98	1.94	1.90
	7.04	4.95	4.10	3.62	3.31	3.09	2.93	2.79	2.70	2.61	2.54	2.47
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.01	1.97	1.93	1.89
	7.01	4.92	4.08	3.60	3.29	3.07	2.91	2.77	2.67	2.59	2.51	2.45
80	3.96	3.11	2.72	2.48	2.33	2.21	2.12	2.05	1.99	1.95	1.91	1.88
n_2	6.96	4.88	4.04	3.56	3.25	3.04	2.87	2.74	2.64	2.55	2.48	2.41
100	3.94	3.09	2.70	2.46	2.30	2.19	2.10	2.03	1.97	1.92	1.88	1.85
	6.90	4.82	3.98	3.51	3.20	2.99	2.82	2.69	2.59	2.51	2.43	2.36
125	3.92	3.07	2.68	2.44	2.29	2.17	2.08	2.01	1.95	1.90	1.86	1.83
	6.84	4.78	3.94	3.47	3.19	2.95	2.79	2.65	2.56	2.47	2.40	2.33
150	3.91	3.06	2.67	2.43	2.27	2.16	2.07	2.00	1.94	1.89	1.85	1.82
	6.81	4.75	3.91	3.44	3.14	2.92	2.76	2.62	2.53	2.44	2.37	2.30
200	3.89	3.04	2.65	2.41	2.26	2.14	2.05	1.98	1.92	1.87	1.83	1.80
	6.76	4.71	3.88	3.41	3.11	2.90	2.73	2.60	2.50	2.41	2.34	2.28
400	3.86	3.02	2.62	2.39	2.23	2.12	2.03	1.96	1.90	1.85	1.81	1.78
	6.70	4.66	3.83	3.36	3.06	2.85	2.69	2.55	2.46	2.37	2.29	2.23
1,000	3.85	3.00	2.61	2.38	2.22	2.10	2.02	1.95	1.89	1.84	1.80	1.76
	6.66	4.62	3.80	3.34	3.04	2.82	2.66	2.53	2.43	2.34	2.26	2.20
∞	3.84	2.99	2.60	2.37	2.21	2.09	2.01	1.94	1.88	1.83	1.79	1.75
	6.64	4.60	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.24	2.18

【相當於機率 .05(上行)與 .01(下行)之 F 值，n_1 爲較大均方之自由度】

n_2	14	16	20	24	30	40	50	75	100	200	500	∞
50	1.90	1.85	1.78	1.74	1.69	1.63	1.60	1.55	1.52	1.48	1.46	1.44
	2.46	2.39	2.26	2.18	2.10	2.00	1.94	1.86	1.82	1.76	1.71	1.68
55	1.88	1.83	1.76	1.72	1.67	1.61	1.58	1.52	1.50	1.46	1.43	1.41
	2.43	2.35	2.23	2.15	2.06	1.96	1.90	1.82	1.78	1.71	1.66	1.64
60	1.86	1.81	1.75	1.70	1.65	1.59	1.56	1.50	1.48	1.44	1.41	1.39
	2.40	2.32	2.20	2.12	2.03	1.93	1.87	1.79	1.74	1.68	1.63	1.60
65	1.85	1.80	1.73	1.68	1.63	1.57	1.54	1.49	1.46	1.42	1.39	1.37
	2.37	2.30	2.18	2.09	2.00	1.90	1.84	1.76	1.71	1.64	1.60	1.56
70	1.84	1.79	1.72	1.67	1.62	1.56	1.53	1.47	1.45	1.40	1.37	1.35
	2.35	2.28	2.15	2.07	1.98	1.88	1.82	1.74	1.69	1.62	1.56	1.53
80	1.82	1.77	1.70	1.65	1.60	1.54	1.51	1.45	1.42	1.38	1.35	1.32
	2.32	2.24	2.11	2.03	1.94	1.84	1.78	1.70	1.65	1.57	1.52	1.49
100	1.79	1.75	1.68	1.63	1.57	1.51	1.48	1.42	1.39	1.34	1.30	1.28
	2.26	2.19	2.06	1.98	1.89	1.79	1.73	1.64	1.59	1.51	1.46	1.43
125	1.77	1.72	1.65	1.60	1.55	1.49	1.45	1.39	1.36	1.31	1.27	1.25
	2.23	2.15	2.03	1.94	1.85	1.75	1.68	1.59	1.54	1.46	1.40	1.37
150	1.76	1.71	1.64	1.59	1.54	1.47	1.44	1.37	1.34	1.29	1.25	1.22
	2.20	2.12	2.00	1.91	1.83	1.72	1.66	1.56	1.51	1.43	1.37	1.33
200	1.74	1.69	1.62	1.57	1.52	1.45	1.42	1.35	1.32	1.26	1.22	1.19
	2.17	2.09	1.97	1.88	1.79	1.69	1.62	1.53	1.48	1.39	1.33	1.28
400	1.72	1.67	1.60	1.54	1.49	1.42	1.38	1.32	1.28	1.22	1.16	1.13
	2.12	2.04	1.92	1.84	1.74	1.64	1.57	1.47	1.42	1.32	1.24	1.19
1,000	1.70	1.65	1.58	1.53	1.47	1.41	1.36	1.30	1.26	1.19	1.13	1.08
	2.09	2.01	1.89	1.81	1.71	1.61	1.54	1.44	1.38	1.28	1.19	1.11
∞	1.69	1.64	1.57	1.52	1.46	1.40	1.35	1.28	1.24	1.17	1.11	1.00
	2.07	1.99	1.87	1.79	1.69	1.59	1.52	1.41	1.36	1.25	1.15	1.00

均方的自由度，稱爲 n_2；表的本身是 F 值，每一對 n_1 與 n_2 可查得兩個 F 值，上面的一個是 5% 點，下面的一個是 1% 點. 例如表 9–2 的兩個均方以 248.67 較大，故其相當的自由度 2 爲 n_1，而誤差的自由度 15 爲 n_2. 查表 9–3，在 $n_1=2$ 一直行與 $n_2=15$ 一橫行的相交處有兩個數值，上面一數是 3.68，卽 F 的 5% 點，下面一數是 6.36，爲 F 的1% 點. 顯著性的標準還是和以前一樣，卽求得的 F 值若小於 5% 點則爲不顯著，若在 5% 點與 1% 點間者爲顯著，大於1% 點者爲非常顯著. 此例求得的 F 值爲 12.18，遠過於 1% 點，故知三個均數間之差別爲非常顯著. 故僅就一週內所增之體重言，則第三組最多，第一組次之，而第二組最少. 但此處食量問題並未計入，若用第十章方法將三組食量化爲相等後，則發現此一週內三組所增體重之均數並無顯著的差別，或因三種飼料的營養價值本無多大差別，或因此處僅用一週的紀錄，時間過短，不能把差別顯出來的緣故.

9–2. 大小不等的三組間之比較——依一個標準分類者 我們在第四章第 4–7 節，曾以 t 測驗法比較控制組與截除結腸組的犬，其副甲狀腺性搐搦發生的時間，原著者蔡、徐二氏(7)曾另用犬十四頭在摘去其副甲狀腺以前或同時將膽管縛紮，以與控制組之僅摘去副甲狀腺者相比較. 其統計處理方法，當與第 4–7 節相同. 但若欲將三組同時比較，則必須用變異數分析法，而不能用以前的 t 值了. 這裏再引用此例，以說明兩組相互比較與三組（或更多組）同時比較的統計處理法之不同.

表 9–4 爲控制組、截除結腸組與縛紮膽管組（該組因有一犬紀錄不全，故僅錄 13 頭）之犬自摘去副甲狀腺至搐搦發生所需的時間，

【表 9-4】 三組紀錄之比較
（自摘去副甲狀腺至搐搦發生所需之時間，小時）

	控制組	截除結腸組	縛紮膽管組	總計
X	54	48	8	
	48	42	42	
	48	48	96	
	96	48	48	
	45	120	48	
	96	48	60	
	20	53	48	
	68	66	58	
	45		68	
	45		48	
			96	
			96	
			52	
ΣX	565	473	808	1,846
k	10	8	13	31
$\bar{x}$	56.5	59.125	62 15	59.55
ΣX^2	37,055	32,545	55,224	124,824
$(\Sigma X)^2/k$	31,922.5	27,966.12	50,220.31	109,926.32
$\Sigma(X-\bar{x})^2$	5,132.5	4,578.88	5,003.69	14,897.68

此資料與表 9-1 不同的地方，在各組動物數不等，故變異數分析的計算方法微有不同，但原理是一樣的，這裏不再詳述了．該表上半部是原來紀錄，下半部為初步統計結果，其中 ΣX 是各組的總時間，k

爲各組動物數，$\bar{x}$ 爲均數，ΣX^2 爲各量數平方之和，$(\Sigma X)^2/k$ 爲求離均差平方和時所用的校正數，如

$$(565)^2/10=31,922.5.$$

在 ΣX^2 內減去此數，即得離均差平方和

$$37,055-31,922.5=5,132.5.$$

餘類推. 該表右側總計一直行內，除 $\Sigma X, k, \Sigma X^2$ 三數爲同一橫行內三組的總和外，其餘是另行計算的，如

$$\bar{x}=\frac{1,846}{31}=59.55,$$

$$\Sigma(X-\bar{x})^2=124,824-\frac{(1,846)^2}{31}$$

$$=124,824-109,926.32=14,897.68.$$

此兩數値，一爲總均數，一爲總離均差平方和. 至此即可將後者分析爲兩部分:

組內變異　$5,132.5+4,578.88+5,003.69=14,715.07,$

組間變異　$14,897.68-14,715.07=182.61.$

若先算組間變異，則爲

組間變異　$\dfrac{(565)^2}{10}+\dfrac{(473)^2}{8}+\dfrac{(808)^2}{13}-\dfrac{(1,846)^2}{31}$

$$=31,922.50+27,966.12+50,220.31-109,926.32=182.61.$$

其公式爲

$$\Sigma k_i(\bar{x}_i-\bar{x})^2=\frac{(\Sigma X_1)^2}{k_1}+\frac{(\Sigma X_2)^2}{k_2}+\cdots\cdots+\frac{(\Sigma X_n)^2}{k_n}-\frac{(\Sigma X)^2}{\Sigma k_i},$$

公式(9-3)

k_i爲第 i 組的項目數，$\bar{x}_i$ 爲該組均數，$\bar{x}$ 爲總均數. 此處與公式(9-1)不同的地方是在各組項目數不等，故 $(\Sigma X_i)^2$ 須用 k_i 分別除之，換句話說，是要先除後加，卻不能先加後除. 至於組內變異或誤差的平方和，只須在總變異內減去組間變異卽得，故仍可適用公式(9-2).

組內變異　　14,897.68-182.61=14,715.07.

這裏要注意的，卽無論先求組內變異或先求組間變異，兩法所得的結果應相同，若兩結果不符，則計算過程中必有錯誤，讀者可利用這個特性，以資核對.

現在要算各離均差平方和的自由度了. 在總變異方面，動物總數是 31 頭，計算離均差平方和時用去一總均數，故自由度減少一個，卽 31－1＝30. 在組間變異方面，組數爲 3，亦用去一總均數，故自由度爲 3－1＝2. 在誤差方面，求各組離均差平方和時各用去一本組的均數，故自由度爲

$$(10-1)+(8-1)+(13-1)=28.$$

注意組間變異與誤差兩種自由度之和恰等於總變異的自由度，卽

$$2+28=30.$$

最後列成變異數分析表如下:

【表 9-5】 變異數分析(副甲狀腺性搐搦發生之時間)

變異來源	自由度	離均差平方和	均方(變異數)	F
總變異	30	14,897.68		
組間(各均數)	2	182.61	91.31	
組內(誤差)	28	14,715.07	525.54	

表內均方一項，係將自由度除其相當的離均差平方和卽得. 結果，組

間的均方(91.31)反小於誤差的均方(525.54),這表示三組的均數

56.50, 59.13, 62.15

間的變異,反較各組自身的變異為小,實無須再算 F 值.因知三均數間的差別,純由機遇而來,換句話說,三種手術對於動物副甲狀腺性搐搦發生的時間並無眞正的差別.原著者蔡、徐二氏謂:'搐搦症狀如常發現,'又謂:'搐搦症之發現期並不因之而延長,'此項結論與統計處理的結果相符.

讀者至此對於變異數分析的基本原理大概已經明瞭,茲將以上所述,提要錄後.變異數分析,係將所有項目的總變異至少析為兩部分:一為各均數間的變異,簡稱為組間變異;一為各組自身的變異,簡稱組內變異或稱誤差.若前者大於後者若干倍以上,則各均數間有顯著的差別,否則純由機遇而來.計算時先求各離均差平方和及自由度,再求得均方或變異數.最後以誤差之變異數除組間變異數,得 F 值.查表 9-3,得 5% 點與 1% 點,卽可決定其顯著性.

9-3. 依兩個標準分類的三組間之比較 上面兩個例子都是依一個標準分類的,表 9-1 所根據的是飼料,表 9-4 是根據所施的手術.現在我們要討論依兩個標準的分類了.下表(8)為兔的肝臟中脂酸之碘指數(iodine number of the fatty acids).所謂碘指數乃係一百克之脂酸所能吸收碘的克數,碘指數愈大,則脂酸之未飽和度(unsaturation)愈高.實驗動物為白兔十窩,每窩三頭,重量大致相等.在同窩之三兔中,其一注射垂體加壓劑(pitressin),一飼以氯化膽素(choline chloride),另一飼以氯化膽素並注射垂體加壓劑.但因 E 與 I 兩窩僅各有兔二頭,致在第二組紀錄中缺少兩數,括弧內的數值是

【表 9-6】 白兔肝臟脂酸之碘指數

窩別	第一組 注射垂體加壓劑者	第二組 飼食氯化膽素者	第三組 飼食氯化膽素並注射垂體加壓劑者	各窩總和	各窩均數
A	110	105	94	300	103.00
B	104	113	99	316	105.33
C	99	114	116	329	107.67
D	101	129	99	329	107.67
E	100	(105.56)	91	296.56	98.85
F	110	121	115	346	115.33
G	120	114	110	344	114.67
H	112	115	101	328	109.33
I	130	(131.56)	113	373.56	124.52
J	114	130	117	361	120.33
各組總和	1,100	1,178.12	1,055	3,333.12	
各組均數	110.00	117.81	105.50	111.104	
ΣX^2	121,858	139,643.94	112,159	373,660.94	

補進去的. 填補的理論根據係在每一空缺處填入一 X 值，使誤差的離均差平方和爲最小. 所用公式如下:

$$X=\frac{tT+bB-S}{(t-1)(b-1)},\qquad \text{公式(9-4)}$$

式中　　t = 組數或處理數，

b = 窩數，

T = 與缺項同組之項目之和，

B = 與缺項同窩之項目之和，

$$S=\text{觀察所得項目之總和}.$$

若缺少 X_1 與 X_2 兩個項目時，可先用一合理數值（如總均數、所在組或窩的均數）暫代 X_2，由公式求得 X_1 的估計值，於是將 X_2 的暫代值除去，而依公式求之，這樣繼續數次，兩數的估計值就趨於固定了. 上例需要估計的數值，一為 E_2，一為 I_2，茲演算於後.

第一次估計：先令 $I_2=(130+113)\div2=121.5$,

於是 $t=3,\ b=10,\ T=1,062.5,\ B=191,\ S=3,217.5,$

代入公式，得

$$E_2=\frac{3(1,062.5)+10(191)-3,217.5}{(3-1)(10-1)}=104.4.$$

乃以 $E_2=104.4$ 代入，而將暫代的 I_2 除去，此時 t,b 兩值仍相同，惟其餘三數則不同，計為

$$T=1,045.4,\qquad B=243,\qquad S=3,200.4,$$

代入公式求 I_2，得

$$I_2=\frac{3(1,045.4)+10(243)-3,200.4}{(3-1)(10-1)}=131.4.$$

第二次估計：以 $I_2=131.4$ 代入，再求 E_2，此時

$$T=1,072.4,\qquad B=191,\qquad S=3,227.4,$$

得 $$E_2=\frac{3(1,072.4)+10(191)-3,227.4}{(3-1)(10-1)}=105.5.$$

以 $E_2=105.5$ 代入，再求 I_2，此時 $T=1,046.5$, $B=243$, $S=3,201.5$,

得 $$I_2=\frac{3(1,046.5)+10(243)-3,201.5}{(3-1)(10-1)}=131.56.$$

第三次估計： $T=1,072.56,\ B=191,\ S=3,227.56,$

$$E_2=\frac{3(1,072.56)+10(191)-3,227.56}{(3-1)(10-1)}=105.56.$$

$$T=1,046.56,\quad B=243,\quad S=3,201.56,$$

$$I_2=\frac{3(1,046.56)+10(243)-3,201.56}{(3-1)(10-1)}=131.56.$$

若再作第四次估計，將得相同的答案，故待塡補的兩數，事實上已趨於穩定．乃以 $E_2=105.56, I_2=131.56$ 塡入表 9–6 的空缺處，就可進行變異數分析了．凡依兩個標準分類的資料，若遇動物死亡或因他故缺少幾個項目時，都可以用此法補入，但缺項太多，則影響資料之正確性．在更複雜的實驗設計中，遇有缺項需塡補時，可參考本書第十三章第13–8 節，G. W. Snedecor$_{(5,\text{P}.228)}$ 及 C. H. Goulden$_{(9,\text{PP}.263-264)}$ 等著作．

變異數分析時，先求一校正數：

$$\frac{(3,333.12)^2}{30}=370,322.96,$$

再求各離均差平方和：

總變異　　$373,660.94-370,322.96=3,337.98,$

組　間

$$\frac{(1,100)^2+(1,178.12)^2+(1,055)^2}{10}-370,322.96=776.21,$$

窩　間

$$\frac{(309)^2+(316)^2+\cdots\cdots+(361)^2}{3}-370,322.96=1,434.01,$$

誤　差　　$3,337.98-(776.21+1,434.\ 1)=1,127.76.$

所謂組間變異，卽各組均數間的變異，其離均差平方和應與下式結果相等，這點與第9-1節的情形是一樣的.

$$10(110-111.104)^2+10(117.812-111.104)^2$$
$$+10(105.5-111.104)^2=776.21.$$

至於窩間變異，在前面兩節裏是沒有提到的.請看表 9-6 各窩均數，最小的僅 98.85，最大的達 124.52，若窩與窩間有顯著的差別，那麼以前所用的組內變異就不是純由機遇得來的誤差，而不能當做比較的標準了.所以這部分的變異，必須從組內變異中提出來，然後餘下來的纔是眞正的誤差.窩間變異的離均差平方和係各窩均數與總均數相差之平方和，而以各窩項目數乘之，卽

$$3(103-111.104)^2+3(105.33-111.104)^2$$
$$+\cdots\cdots$$
$$+3(124.52-111.104)^2$$
$$+3(120.33-111.104)^2.$$

此式答案應與前面結果 1,434.01 相等，但因小數四捨五入的關係，或微有出入.爲使結果較爲正確起見，以用各窩總和計算爲宜.至於誤差一項，係在總變異內減去組間與窩間兩部分卽得.又窩間與誤差兩部分相加，1,434.01＋1,127.76＝2,561.77，恰等於組內的離均差平方和，讀者可自證之.

關於自由度的計算，要注意補進去的兩數不占自由度，故總變異的自由度爲 28－1＝27(若無缺項時應爲 29)，組間變異的自由度爲3－1＝2,窩間變異的自由度爲10－1＝9,而誤差的自由度則爲 27－(2＋9)＝16 (無缺項時應爲 18).玆將變異數分析列表如下.

【表 9-7】 依兩個標準分數之變異數分析
（白兔肝臟脂酸碘指數之比較）

變異來源	自由度	離均差之平方和	均方	F	5%點	1%點
總變異	27	3,357.98				
組間	2	776.21	388.11	5.51	3.63	6.23
窩間	9	1,434.01	159.33	2.26	2.54	3.78
誤差	16	1,127.76	70.49			

F 值的計算，都以誤差為根據，如 $388.11 \div 70.49 = 5.51$，又 $159.33 \div 70.49 = 2.26$. 由表 9-3，查得 $n_1 = 2, n_2 = 16$ 及 $n_1 = 9, n_2 = 16$ 時的 5%點與 1%點. 這裏可以看到組間的 F 值與 1%點較近，故為顯著. [在補入缺項的資料中，所得 F 值常較其眞值稍大，故 G. W. Snedecor 氏主張(5, p. 224)當求得之 F 值在 5%點與 1%點之間時，以近於 1%點者始認為顯著較妥. 至於精細的校正方法，可參考 F. Yates(13) 及 C. H. Coulden (9, ch. XVI)等著作.]而窩間的 F 值則為不顯著. 因得結論曰：各組平均碘指數有顯著的差別，即第二組動物肝臟脂酸之未飽和度最高，而餘兩組次之. 至於各窩之間，雖均數表面值有大小，實則並無眞正的差別. 故窩間變異可不必提出，即以組內變異作為比較的標準. 茲以第 9-2 節大小不等的三組之比較法重作變異數分析如下表，此時第二組所缺兩數無須補入，所得結論與前相同.

【表 9-8】 依一個標準分類之變異數分析
（白兔肝臟脂酸碘指數之比較）

變異來源	自由度	離均差之平方和	均方	F	5%點	1%點
總變異	27	2,880.86				
組間	2	658.49	329.25	3.704	3.38	5.57
組內(誤差)	25	2,222.37	88.89			

本節開始時，所述白兔肝臟脂酸碘指數的設計，稱爲隨機區組(randomized blocks). 因爲這種設計最初用在農業方面，所用田地包括若干區組(block)，每一區組又分成若干區(plot)，區的多寡與處理的數目相同，某區應施何種處理，以隨機方法決定之(參閱第一章第 1–3 節). 上述碘指數實驗，用兔十窩，每窩三頭，相當於十個區組，每區組又分三個區. 現在隨機區組的實驗已廣爲應用，卽非田間實驗，亦沿用此名稱.

9–4. 交互影響(interaction)**之顯著性**　生物化學告訴我們 pH (氫游子濃度指數)對於滲潤作用(imbibition, 膠體吸收水分的現象)的影響，視膠體(colloid)的性質而異，蛋白質在等電點 (isoelectric point)——兩性電解物粒體，其酸性及鹼性行爲平衡時之 pH 值，稱爲等電點，此點視各種蛋白質而異，如血清蛋白及酪蛋白爲 4.7，蛋白蛋白爲 4.8 等——時滲潤最小，碳水化合物之複糖類則在中性($pH=7$)時滲潤最大. 這種情形，稱爲 pH 與膠體對於滲潤作用的交互影響. 在農業上，如施以化學肥料的作物，其產量視同時所施的腐植土而不同. 又如某作物之各品種，在不同地方之產量各異. 這些都是交互影響的例子. 遇到這種情形，我們在實驗設計中，就應該顧到交互影響的測量及其顯著性測驗法. 下面的例題說明測量交互影響的設計方法及實驗的結果 (5, p. 233). 實驗中所用的公鼠，分爲六組. 三組的飼料中蛋白質的成分高，簡稱高蛋白，餘三組的飼料中蛋白質的成分低，簡稱低蛋白. 至於蛋白質的來源則各爲牛肉、穀類與豬肉. 因爲每組有公鼠10頭，故各組所增體重的均數，只須看它的總和(ΣX)就可知道. 若從各組均數的表面值看，則在高蛋白方面，吃牛

【表 9-9】 六組公鼠增加之體重(克)

	高蛋白			低蛋白		
	牛肉	穀類	豬肉	牛肉	穀類	豬肉
	73	98	94	90	107	49
	102	74	79	76	95	82
	118	56	96	90	97	73
	104	111	98	64	80	86
	81	95	102	86	98	81
	107	88	102	51	74	97
	100	82	108	72	74	106
	87	77	91	90	67	70
	117	86	120	95	89	61
	111	92	105	78	58	82
ΣX	1,000	859	995	792	839	787
ΣX^2	102,062	75,819	100,075	64,462	72,613	64,401

肉與豬肉的兩組動物體重的增加多，而吃穀類的增加少；在低蛋白方面的情形剛好相反. 如果這種現象並非偶然，那麼蛋白質成分的高低，與來源之間就有顯著的交互影響了. 計算交互影響時，我們需要列一個副表:

【表 9-10】 求交互影響用之副表

	牛肉	穀類	豬肉	總計
高	1,000	859	995	2,854
低	792	839	787	2,418
總計	1,792	1,698	1,782	5,272

【表9-12】 自二十五年至二十八年各月初生男女嬰孩之體重(仟克)

	一月		二月		三月		四月		五月		六月		七月		八月		九月		十月		十一月		十二月	
	男	女	男	女	男	女	男	女	男	女	男	女	男	女	男	女	男	女	男	女	男	女	男	女
二十五年	2.61	3.18	2.72	2.72	3.12	3.40	3.18	3.69	2.95	3.18	2.95	2.61	2.27	3.12	2.38	3.57	2.38	3.18	3.06	2.72	3.18	3.18	3.18	2.89
	3.18	3.01	3.03	3.57	3.29	3.52	3.97	2.61	2.72	2.27	3.86	2.95	3.74	2.95	3.52	3.06	3.06	2.83	3.74	3.18	2.72	3.52	3.63	3.18
	2.81	3.18	3.63	3.46	2.95	2.72	3.23	3.06	2.95	2.49	3.63	3.12	3.18	2.95	2.15	2.72	3.52	3.52	3.18	2.72	2.38	3.18	2.72	2.95
	2.83	3.12	3.06	3.29	2.95	2.61	2.38	2.83	3.40	2.95	3.18	3.29	2.83	3.74	3.06	2.61	2.27	2.89	3.63	2.83	2.95	2.83	3.18	3.18
	2.72	3.97	2.61	2.83	3.06	2.95	2.72	2.95	2.83	2.27	3.29	3.40	3.86	2.95	3.57	2.72	4.08	3.06	2.83	3.18	2.72	2.83	3.57	2.72
	14.15	16.46	15.03	15.87	15.37	15.20	15.48	15.14	14.85	13.16	16.91	15.37	15.88	15.71	14.68	14.68	15.31	15.48	16.44	14.63	13.95	15.54	16.28	14.92
二十六年	3.63	2.95	3.40	3.06	3.18	3.06	3.52	2.95	4.03	2.83	2.83	3.40	3.97	3.12	3.29	3.18	2.66	3.06	3.63	3.63	3.40	2.95	3.18	3.06
	2.83	2.49	3.52	3.52	3.18	3.63	3.63	2.49	2.83	2.83	2.72	3.23	2.95	2.72	3.29	3.03	3.06	3.06	3.97	2.83	3.06	3.18	2.95	2.72
	2.72	2.89	3.06	3.52	3.74	2.72	3.40	2.72	3.18	3.23	3.06	2.83	3.06	3.06	3.52	3.40	3.63	3.52	2.61	3.01	3.40	3.06	2.49	2.49
	2.95	2.83	3.12	3.35	2.83	3.06	2.72	3.06	3.40	2.49	2.83	3.18	3.40	2.83	3.18	2.72	3.06	2.95	3.46	3.18	3.06	2.72	3.06	2.61
	3.12	3.23	2.95	3.29	2.95	2.72	3.63	2.72	3.63	2.61	3.63	2.55	3.46	2.49	3.18	2.61	3.57	2.95	3.40	2.89	3.63	3.23	2.38	3.18
	15.25	14.39	16.05	16.74	15.88	15.19	16.90	3.94	17.12	13.99	15.07	15.19	16.84	14.22	16.46	14.97	15.98	15.54	17.07	15.54	16.55	15.14	14.06	14.06
二十七年	3.52	3.18	3.12	3.18	3.12	2.66	2.95	3.63	2.95	3.06	3.74	3.40	3.29	3.86	3.63	2.66	2.95	2.72	4.54	2.49	3.52	3.01	3.18	2.61
	2.27	2.86	3.74	2.44	2.95	2.49	3.46	3.29	3.18	2.72	3.63	3.18	3.18	3.63	3.40	3.29	2.72	2.95	3.63	2.72	2.95	2.83	3.18	3.74
	3.29	3.57	3.35	3.18	3.29	4.03	3.52	2.61	2.83	3.18	3.01	2.83	2.72	3.06	2.89	3.03	2.66	2.49	2.83	2.72	3.18	3.74	3.63	2.61
	3.13	3.18	2.72	2.72	2.61	3.12	3.18	3.06	3.52	3.40	2.27	2.61	3.06	3.52	2.95	2.72	2.61	2.72	2.66	2.95	3.06	3.97	2.61	3.46
	3.52	3.40	2.83	3.18	2.72	2.72	2.83	3.74	3.52	4.69	2.83	3.18	2.72	3.06	2.72	3.63	2.49	3.86	3.06	2.83	3.29	3.18	2.95	3.18
	15.78	16.19	15.76	14.70	14.69	15.02	15.94	16.33	16.00	17.05	15.48	15.20	14.97	17.13	15.59	15.36	13.43	14.74	16.72	13.71	16.00	16.73	15.55	15.60
二十八年	3.86	3.40	3.86	3.06	3.52	2.95	3.09	3.63	3.40	2.55	2.27	3.20	3.86	3.97	2.83	3.12	3.63	2.72	3.23	2.85	3.52	2.49	3.40	2.72
	2.72	2.95	3.63	2.83	3.12	3.35	3.18	4.03	3.40	3.06	2.72	2.95	3.29	2.72	3.01	3.35	3.06	2.72	3.06	2.78	3.03	2.49	3.57	2.38
	3.46	3.18	3.63	2.83	2.49	2.27	3.97	2.64	3.40	2.78	3.77	2.66	2.95	2.49	3.18	2.92	3.18	3.06	3.09	3.18	3.46	3.52	3.06	2.95
	3.01	3.06	2.61	2.72	2.83	3.03	3.18	2.44	3.18	2.72	2.95	2.38	2.72	3.52	4.31	3.29	2.61	3.23	2.78	2.72	2.61	2.83	3.97	2.61
	3.63	3.23	3.52	2.61	2.89	3.40	2.49	2.61	3.10	2.78	3.52	3.52	3.29	3.35	3.18	2.72	2.49	3.06	3.91	3.23	2.72	2.78	3.86	2.95
	16.68	15.82	17.25	14.05	14.85	15.00	15.91	15.35	16.48	13.89	15.23	14.71	16.11	16.05	16.51	15.40	14.97	14.79	16.07	14.76	15.34	14.11	17.86	13.61

由此即可進行下列計算:

總和　$\Sigma X = 5,272,$

$\Sigma X^2 = 479,432.$

校正數　$\dfrac{(5,272)^2}{60} = 463,233.07.$

離均差平方和

總變異 $479,432 - 463,233.07 =$ 　16,198.93

高低　$\dfrac{(2,854)^2 + (2,418)^2}{30} - 463,233.07 =$ 　3,168.26

來源　$\dfrac{(1,792)^2 + (1,698)^2 + (1,782)^2}{20} - 463,233.07 =$ 　266.53

副表總變異

$$\frac{(1,000)^2 + (859)^2 + \cdots\cdots + (787)^2}{10} - 463,233.07$$

$$= 4,612.93$$

交互影響(高低×來源)

$4,612.93 - 3,168.26 - 266.53 =$ 　1,178.14

誤差(餘數) = 　11,586.00

在求'高低'的離均差平方和時,高蛋白方面的總數是 2,854,此數係 30 個項目的總和,低蛋白方面亦然,故其平方須各以 30 除之,減去校正數後,即得離均差平方和(參閱公式 9-1).同理,來源方面,每一總數是 20 個項目的總和,故平方後須各以 20 除之,餘類推.表 9-11 的總變異爲 4,612.93,減去'高低'與'來源'兩方面的變異後,即得交互影響的離均差平方和.最後在總變異 16,198.93 內減去'高低',

'來源',及'交互影響'三者卽得誤差.

關於自由度的計算:總變異方面爲60－1＝59,高低方面爲2－1＝1,來源方面爲3－1＝2,交互影響方面將高低與來源的自由度相乘卽得(2－1)(3－1)＝2,誤差方面爲59－(1＋2＋2)＝54.

【表 9-11】 變異數分析(六組公鼠所增之體重)

變異來源	自由度	離均差平方和	均　　方	F	5%點	1%點
總變異	59	16,198.93				
高　低	1	3,168.26	3,168.26	14.766	——	7.12
來　源	2	266.53	133.27			
交互影響	2	1,178.14	589.07	2.745	3.17	——
誤　差	54	11,586.00	214.56			

變異數分析見表 9-11,比較時皆以誤差的均方爲標準,結果表示吃高蛋白與低蛋白的動物所增體重有非常顯著的差別,而蛋白質的高低與來源之間並無顯著的交互影響,可知前述表面的看法是不正確的. 至於來源方面,其均方小於誤差的均方,便知爲不顯著,實無須再算它的 F 值了.

9-5. 依三種標準分類的多組間之比較　表 9-12 爲民國二十五年至二十八年各月初生男女嬰孩之體重紀錄(14),是項資料係在原來紀錄中,每月隨機抽取男女嬰孩各五人,以視其出生體重有無季節波動的現象(注意第八章練習題第 7 題亦係測驗季節變動現象,但該題爲計數資料,而此處爲測量資料,故統計之處理方法不同).該表是根據年分、月分與性別三種標準分類的,其交互影響就有(年×月)、(月×男女)、(年×男女)與(年×月×男女)多種,末了的一個

稱爲三重交互影響(triple interaction),係表示(年×月)交互影響在男女方面差別的程度.計算這些交互影響時,需先有三個副表,如表 9-13(甲),(乙),(丙).詳細過程如下:

總和 $\Sigma X=1,483.15,$

$\Sigma X^2=4,662.4349.$

校正數 $\dfrac{(1,483.15)^2}{480}=4,582.7790.$

離均差平方和

總變異 $4,662.4349-4,582.7790=\quad 79.6559$

月間 $\dfrac{(124.72)^2+(125.50)^2+\cdots\cdots+(121.94)^2}{40}$

$-4,582.7790=1.0183$

年間 $\dfrac{(366.54)^2+\cdots\cdots+(370.80)^2}{120}-4,582.7790=0.2348$

男女 $\dfrac{(756.78)^2+(726.37)^2}{240}-4,582.7790=\quad 1.9266$

副表(甲)之總變異

$\dfrac{(30.61)^2+(30.95)^2+\cdots\cdots+(31.47)^2}{10}$

$-4,582.7790=6.6435$

交互影響(年×月)

$6.6435-1.0183-0.2348=\quad 5.3904$

交互影響(月×男女)

$\dfrac{(-1.00)^2+(2.78)^2+\cdots\cdots+(5.56)^2}{40}$

$-1.9266=2.2036$

【表 9-13】 計算交互影響所用之副表(甲)

	一月	二月	三月	四月	五月	六月	七月	八月	九月	十月	十一月	十二月	總計
二十五年	30.61	30.95	30.57	30.62	28.01	32.28	31.59	29.36	30.79	31.07	29.49	31.20	366.54
二十六年	29.64	32.79	31.07	30.84	31.11	30.26	31.06	31.43	31.52	32.61	31.69	28.12	372.14
二十七年	31.97	30.46	29.11	32.27	33.05	30.68	32.10	30.95	28.17	30.43	32.73	31.15	373.67
二十八年	32.50	31.30	29.85	31.26	30.37	29.94	32.16	31.91	29.76	30.83	29.4[illegible]	31.47	370.80
總計	124.72	125.50	121.20	124.99	122.54	123.16	126.91	123.65	120.24	124.94	123.36	121.94	1,483.15

【表 9-13】 計算交互影響所用之副表(乙)

	一月	二月	三月	四月	五月	六月	七月	八月	九月	十月	十一月	十二月	總計
男	61.86	64.14	60.79	64.23	64.45	62.69	63.80	63.24	59.69	66.30	61.84	63.75	756.78
女	62.86	61.36	60.41	60.76	58.09	60.47	63.11	60.41	60.55	58.64	61.52	58.19	726.37
總計	124.72	125.50	121.20	124.99	122.54	123.16	126.91	123.65	120.24	124.94	123.36	121.94	1,483.15
相差	-1.00	2.78	.38	3.47	6.36	2.22	.69	2.83	-0.86	7.66	.32	5.56	

【表 9-13】 計算交互影響所用之副表(丙)

	二十五年	二十六年	二十七年	二十八年	總計
男	184.38	193.23	185.91	193.26	756.78
女	182.16	178.91	187.76	177.54	726.37
總計	366.54	372.14	373.67	370.80	1,483.15
相差	2.22	14.32	-1.85	15.72	

交互影響(年×男女)

$$\frac{(2.22)^2+(14.32)^2+\cdots\cdots+(15.72)^2}{120}$$

$$-1.9266=1.9112$$

表 9-12 逐年逐月男女之總變異

$$\frac{(14.15)^2+(16.46)^2+\cdots\cdots+(13.61)^2}{5}$$

$$-4,582.7790=17.9168$$

交互影響(年×月×男女)

$$17.9168-(1.0183+0.2348+1.9266$$
$$+5.3904+2.2036+1.9112)$$
$$=17.9168-12.6849=5.2319$$

誤差(餘數) $79.6559-12.6849-5.2319=$ 61.7391

上列月間變異之離均差平方和,係將副表所載各月之總計值平方後相加,以構成每月總計值之項目數 40 除之,然後減去校正數卽得. 年間及男女兩變異之求法亦同. 計算(年×月)之交互影響時,須先求得副表(甲)之總變異,再減去月間與年間兩變異卽得. 至於(月×男女)之交互影響本可在副表(乙)之總變異內減去月間與男女間之變異,但因性別方面只有男女兩項,故可以簡法計算,卽將該表每月男女之差平方後相加,以每月男女總數 40 除之,再減去男女間之變異,卽爲所求之交互影響. 又交互影響(年×男女)之計算法亦然. 至於三重交互影響,則須在表 9-12 逐年逐月男女之總變異內減去年間、月間、男女間及各二重交互影響始可.

各離均差平方和之自由度計算如下:總變異方面,$480-1=479$;

月間變異 12－1＝11；年間變異 4－1＝3；男女間變異 2－1＝1；交互影響（年×月），(4－1)(12－1)＝33；交互影響（月×男女），(12－1)×(2－1)＝11；交互影響（年×男女），(4－1)(2－1)＝3；交互影響（年×月×男女），(4－1)(12－1)(2－1)＝33；誤差，479－(11＋3＋1＋33＋11＋3＋33)＝384.

變異數分析列於下表：

【表 9-14】依三個標準分類之變異數分析（初生嬰兒之體重）

變異來源	自由度	離均差平方和	均方
總變異	479	79.6559	
月間	11	1.0183	.09257
年間	3	0.2348	.07827
男女間	1	1.9266	1.92660
交互影響（年×月）	33	5.3904	.16335
交互影響（月×男女）	11	2.2036	.20033
交互影響（年×男女）	3	1.9112	.63707
交互影響（年×月×男女）	33	5.2319	.15854
誤差	384	61.7391	.16078

這裏我們可以看到：月間與年間兩均方都小於誤差，故為不顯著，換言之，初生嬰孩之體重並無季節波動的現象；又交互影響（年×月）及（年×月×男女）的兩個均方，幾與誤差相等，這表示逐年逐月之間並無眞正的差別，且交互影響（年×月）在男女方面也沒有什麼不同；又交互影響（月×男女）雖略大於誤差，但為不顯著．上表中值得注意的是男女間與交互影響（年×男女）的兩個均方．後者的 F 值是 .63707÷.16078＝3.962，此值大於 1%點（3.83），故為非常顯著．這是指各年男女嬰孩的體重均數，其趨勢很不一致．試由表 9-13（丙）

計算各年男女的均數，則將如表 9-15. 其中二十八年男女均數相差

【表 9-15】 各年男女嬰孩體重之均數（仟克）

年份	男	女	相差
二十五年	3.073	3.036	+.037
二十六年	3.221	2.982	+.239
二十七年	3.099	3.129	-.030
二十八年	3.221	2.959	+.262
總計	3.153	3.027	+.126

+.262 仟克，而二十七年則相差 -.030 仟克. 正因爲相差符號有正有負，相差數值有大有小，遂使（年×男女）的交互影響非常顯著. 至於男女總均數的相差爲 +.126，其均方爲 1.9266（見表 9-14），這裏我們需要考慮的，在測驗顯著性時應當用那一個均方作爲比較的標準？如果比較男女的相差而對於年月因子不加考慮，那麼祇須把男女的均方與誤差的均方相比較就行了；但若所研究的問題是：在歷年新生嬰兒中，苟以二十五至二十八年爲隨機抽取的樣本，那麼男女嬰孩的體重有顯著的差別否？這裏就要把男女的均方與交互影響（年×男女）的均方相比較，即 $F=1.9266\div .63707=3.024$，其自由度 $n_1=1$，$n_2=3$，5% 點爲 10.13，故知男女均數的差別，若與歷年男女間的變異相比較時實爲不顯著. 此例係說明三個標準的分類及交互影響的用法，在有些問題中交互影響的均方常爲眞正的誤差均方，且爲顯著性測驗時所應當用的. 此點在下節還有更詳細的討論. 至於初生嬰孩體重的研究，因此處僅有很少的資料，不能遽下結論.

9-6. 選擇適當的誤差　顯著性是相對的，而非絕對的. 相差是

否顯著，係與另一變異相比較而言，這種變異的來源是按我們對於實驗結果所用的解釋而任意選擇的.假定現在作一化學測定的實驗.所用的方法係在 A, B 兩物質中各抽取 20 個樣本.每個樣本作 2 次測定.此實驗有兩個誤差，一爲抽樣時所產生的，一爲兩次測定結果的差別所產生的，後者純係實驗技術上的誤差.這兩種誤差彼此獨立，故其大小可以相等，也可以相去懸殊.本實驗的變異數分析有如表 9-16 的形式:

【表 9-16】 化學測定之變異數分析(一)

變異來源	自由度	變異數(均方)
材料(A, B 兩物質間) 组间变异	1	m
A 物各樣本間 组内变异,	19 } 38	a } s
B 物各樣本間	19	b
重複測定間 随机变异	40	d
總計	79	

爲便於討論計，假定 a, b 兩變異數相等，並可認爲合成一個變異數 s.若將 s 與 d 相比較，那麼後者是很小的.現在我們要測驗 A, B 兩物質相差的顯著性，若二者之間並無差別，那麼變異數 m 將與 s 相等，因爲 m 是與樣本間的變異有關的.今 d 值旣很小，若用以測驗 m 的顯著性，即使 A, B 兩物並無差別，而 m 與 d 的比例將很大，這顯然是錯誤的.也許有人要問:若 d 比 s 大得多，將怎樣辦?但仔細一想，這種情形很難得，因爲產生變異數 d 的因子，對於變異數 m 也同樣生作用，若樣本之間並無變異，則就平均言，s 將與 d 相等.故當 d 與 s 大致相等時，即表示 s 大部分由兩次測定的差別而來，於是抽樣誤

差的本身就不顯著了. 此時可用 d 來測驗 m 的顯著性，且因 d 的自由度較 s 的自由度爲多，故可增加測驗之精密度(precision).

還有一個假設的實驗，也用兩種物質，假定兩物質都充分勻淨，故抽樣誤差可略而不計，這點是與前例不同的. 本實驗也可能有兩種誤差：一是實驗室技術方面的誤差，一是工作者方面的誤差，因爲沒有兩個人會得到完全相同的結果的. 工作者共有六人，每人用兩種物質作同樣的試驗，每試驗重複三次，這樣可以測量技術上的誤差. 其變異數分析如表 9-17. 這裏要注意的是變異數 e，其意義可用

【表 9-17】 化學測定之變異數分析(二)

變 異 來 源	自 由 度	變異數(均方)
材料	1	m
工作者	5	o
工作者之誤差 〔即交互影響(材料×工作者)〕	5	e
重複測定之誤差	24	d
總 計	35	

表 9-18 來解釋. 該表中 a_1 代表第 1 位工作者以 A 物質測定三次的

【表 9-18】 各工作者所測定之兩物質之均數

	工作者 1	2	3	4	5	6
材料 A	a_1	a_2	a_3	a_4	a_5	a_6
材料 B	b_1	b_2	b_3	b_4	b_5	b_6

均數，b_1 爲彼測定 B 物質三次的均數，餘類推. 因爲各工作者所測定的兩均數的相差不同，於是就產生變異數 e. 若 $(a_1-b_1)\doteq(a_2-b_2)$

材料

$=\cdots\cdots=(a_6-b_6)$,則 e 等於零.若六對均數之差相去懸殊,那麼 e 就很大了.假定有很多工作者對於 A, B 兩物作同樣的測定,而上述研究祇是這大規模實驗中的一個樣本,那麼產生變異數 e (工作者的誤差)的許多因子對於變異數 m (表示兩物質間的相差者)也會發生作用;故若兩物質間並無差別,則 m 將等於 e. 在這種樣本中,應將 e 當做誤差變異數,以測驗 m 的顯著性.有時用錯了誤差變異數,會引起很不幸的結果.因爲各工作者對於技術的標準化特別注意,故變異數 d 很小,我們假定它比 e 要小得多.今 m 雖僅略大於 e, 但若與 d 比較起來,就很顯著了.因得一結論曰: A 所得的結果比 B 的大得多.某工廠就根據實驗結果用 A, B 兩物質從事製造.製造時需用許多工人,於是工作者的誤差,在實驗室中被忽略的,這時卻表現出來了.這樣經實地製造後,證明兩物質得到同樣的結果.所謂嚴密控制的實驗,因此受人詆毀了.如果研究者能仔細考慮抽樣的情形及其全體(如上述大規模製造中的很多工作者)的正確性質,然後作適當的顯著性測驗,那麼這種錯誤是可以避免的.

也許有人要問:在測驗兩物質相差的顯著性時,既然不用技術上的誤差,那麼每一工作者何必作三次測定,而又把變異數 d 算出來呢.須知在重複測定間若有相當誤差,則變異數 e 亦將受其影響,故工作者之間若並無變異,則就平均言, e 將等於 d. 所以變異數 d 的用途,在用來測驗 e 的顯著性.若 d 值並不太小,足以減低實驗的精密度,這時就需要改良測定的技術了.

關於選擇適當誤差的問題,在 C. H. Goulden (9, pp. 128–134) 氏著作中還有更詳細的例子,讀者可以參考.

【練 習 題】

1. 查下列各對自由度之F之 5%點與1%點；遇表 9-3 未載之自由度，則以最近者代之.

n_1	1	3	4	8	6	30	50	2	3	80
n_2	3	1	4	13	25	16	400	54	384	58

2. 試比較下列四組之均數有差別否？

第一組	第二組	第三組	第四組
8	7	6	7
3	9	12	4
4	6	9	5
3	8	13	3
1	4	10	2
5	7	11	0
4	11	10	9
2	8	8	7
4	7	11	6
6	3	10	7

3. 下表爲不同飼料之三組白鼠，在九週內增加之體重(克)(15)，試比較三種飼料之營養價值：

蠶蛹粉	牛肉粉	豬肉粉
158	155	208
134	132	131
128	132	142
126	110	70
103	85	69
134	106	87

4. 鄭集、戴重光(10)二氏在黃豆蛋白之研究中，將白鼠分爲三

組:甲組飼料爲黃豆蛋白質(glycinin);乙組爲黃豆蛋白質與酪蛋白(casein);丙組爲黃豆蛋白質與卵蛋白(albumin).經三十日後,各組動物所增體重(克)如下,問各組所增體重之均數有顯著的相差否?

甲組	乙組	丙組
23.5	41.3	60.9
15.9	40.7	62.1
6.6	28.8	48.9
9.6	18.8	61.2
8.4	21.7	63.9
11.4	25.1	73.7
20.4	22.2	——

5. 蔡翹、徐豐彥二氏(11)曾研究腸阻塞對於副甲狀腺性搐搦之影響.二氏將犬28頭分爲三組:A組之阻塞在十二指腸(duodenum)或空腸(jejunum)上部;B組之阻塞在迴腸(ileum)下部;C組之阻塞在結腸(colon).下表爲各組動物於施手術(摘除副甲狀腺及施腸阻塞)後之生存日數.問三組有顯著的差別否?

A組	B組	C組
2.0	11.0	17.0
1.8	10.0	8.0
5.0	12.0	8.0
9.0	4.0	8.0
0.5	2.8	8.0
1.3	13.0	4.3
14.0	3.5	1.5
2.0	26.0	20.0
2.5	0.5	
4.8	3.5	

6. 汪敬熙氏(12)等研究貓在麻醉中之皮膚電反射,氏等用貓14

頭，先在大貓完整時實驗之，繼割去其皮質，再割去其視丘，依次測定其皮膚電反射之潛伏期及強度．下表爲一部分紀錄，試測驗其均數相差之顯著性，並解釋之.（注意：此爲依二個標準分類之資料.）

皮膚電反射之潛伏期（秒）

動物號碼	大腦完整時	去大腦皮質後	再去視丘後
31	0.76	0.87	1.28
34	0.86	0.97	1.46
44	0.78	0.90	1.50
66	0.74	0.87	1.64

7．下表(14)爲不同胎次之男女嬰兒初生時之體重（仟克）．試測驗其顯著性，並解釋交互影響（胎次×男女）之意義.（注意：此表亦由隨機抽樣得來，何以各組嬰兒數不等？計算離均差平方和時須用公式 9-3.）

	胎次									
	一		二		三		四		五及以上	
	男	女	男	女	男	女	男	女	男	女
X	2.61	2.72	3.63	3.36	2.95	3.97	3.18	2.83	3.63	3.18
	3.06	3.18	3.01	2.95	2.83	2.61	2.27	3.52	2.95	3.29
	3.06	2.49	3.06	2.83	2.72	3.40	3.54	3.63	3.29	4.20
	2.15	3.06	2.49	3.12	3.74	3.52	3.01	2.95	3.06	2.83
	3.52	3.18	3.18	2.95	2.04	3.12	3.01	3.40	2.95	2.89
	3.40	3.40	3.63	2.21	2.95	3.06			3.23	2.83
	2.61	3.01	2.83	2.72	2.38				2.66	2.95
	3.23	2.61	2.72	2.78					3.18	2.42
	2.72	2.92		2.72					2.55	2.83
	2.61	3.23							3.29	
	2.04	3.18							3.18	
	2.61								3.69	
	3.52								2.63	
	3.06								2.83	
ΣX	40.20	32.98	24.55	25.64	19.61	19.68	15.01	16.33	44.12	27.42
k	14	11	8	9	7	6	5	5	14	9
$\bar{x}$	2.871	2.998	3.069	2.849	2.801	3.280	3.002	3.266	3.151	3.047
ΣX^2	118.2954	99.6903	76.4973	73.8472	56.6259	65.6214	45.9171	53.8387	140.6314	85.5142

【參 考 文 獻】

(1) Fisher, R. A., International Mathematical Conference, Toronto (1924).

(2) Fisher, R. A., Statistical Methods for Research Workers, Oliver and Boyd, Edinburgh (1935).

(3) Mahalanobis, P. C., Indian Journal of Agricultural Science, 2 : 694 (1932).

(4) Snedecor, G. W., Analysis of Variance and Covariance, Collegiate Press, Inc. Ames (1934).

(5) Snedecor, G. W., Statistical Methods, The Iowa State College Press, Ames, Iowa (1940).

(6) Lan, T. H., Biological values of mixed cereal and legume proteins, Chinese Journal of Physiology, 10, 5 : 637–644 (1936).

(7) Tsai, C. and Hsu, F. Y., Studies on the pathogenesis of parathyroid tetany, II. The effect of ligation of the bile duct, Chinese Journal of Physiology, 3, 2 : 197 – 204 (1929).

(8) Mukerji, B. and Van Dyke, H. B., The effect of the pressor principle of the posterior lobe of the pituitary body on the liver-fat after the feeding of choline chloride, Chinese Journal of Physiology, 9, 1 : 69–76 (1935).

(9) Goulden, C. H., Methods of Statistical Analysis, John

Wiley & Sons, Inc. New York (1939)

(10) Tai, T. K. & Cheng, L. T., Studies on soy bean proteins, IV. The supplementary effect of albumin and casein on the nutritive value of glycinin, Journal of Chinese Chemical Society, 11, 1 (1944).

(11) Tsai, C. and Hsu, F. Y., Studies on the pathogenesis of parathyroid tetany, III. Influence of intestinal obstruction, Chinese Journal of Physiology, 3, 4 : 389–398(1929).

(12) Wang, G. H., Pan, J. G., and Lu, T. W., The galvanic skin reflex in normal, thalamic, decerebrated and spinal cats under anaesthesia, Chinese Journal of Physiology, 3, 2 : 109–122 (1929).

(13) Yates, F., The analysis of replicated experiments when the field results are incomplete, The Empire Journal of Experimental Agriculture, 1 : 129–142 (1933).

(14) 成都進益高級助產職業學校附屬產院之紀錄.

(15) 陳朝玉: 蠶蛹營養價值之研究，營養專報，第一號，1–7頁，國立四川大學農學院(三十五年五月).

第十章 共變數分析

本書第五、六兩章所討論的直線迴歸與相關，通常是用來表示同一組內兩變數間的關係的．現在我們要把這些方法加以擴充，俾用於兩組以上的資料，並與前章的變異數分析相聯繫，而成共變數分析(analysis of covariance)．本來共變數可用於幾組含有二個或二個以上變數的測量資料，但本章所討論者以兩個變數爲限．

10-1．修正均數(adjusted means)**之顯著性** 前章表 9-10 六組公鼠增加之體重，若與各鼠食物消費量並列，則如表10-1所示．此處食物消費量係以熱量 10 卡(calorie)爲單位，如 108 卽代表 1,080 卡．若以飼料的實際重量計算，其統計處理方法是一樣的，只是數值不同罷了．該表所列六組公鼠，每組各得一種飼料．因各組所用動物一律是十頭，所以從各組的總和 ΣX 或 ΣY，就可推知其均數．我們可以看到第 1，3 兩組所增體重 (Y) 的均數最高，而第 4，6 兩組則最低．但各鼠食量不同，其體重增加的多少是否食量使然，這點必須考慮的．通常將資料化爲每單位食量所增之體重，或每增加單位體重所需之飼料，這樣雖簡單明瞭，但這些數值實爲體重與食量間之比例，若依前章變異數分析法來測驗平均比例值間的顯著性是不甚妥當的，並且有時會得到不同的結果．最合理的方法，是保存食量與體重的原來紀錄而用共變數處理之．

【表 10-1】 六組公鼠之食物消費量(X, 單位為10卡)
及所增之體重(Y, 克)

	組										別	
	1		2		3		4		5		6	
	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
X, Y	108	73	99	98	194	94	165	90	124	107	140	49
	136	102	117	74	198	79	164	76	95	95	177	82
	138	118	90	56	196	96	161	90	116	97	189	73
	159	104	141	111	198	98	159	64	112	80	142	86
	146	81	106	95	210	102	175	86	123	98	216	81
	141	107	112	88	196	102	135	51	110	74	200	97
	175	100	110	82	230	108	132	72	137	74	255	106
	149	87	117	77	222	91	190	90	105	67	173	70
	174	117	111	86	220	120	145	95	135	89	153	61
	176	111	122	92	228	105	142	78	126	58	160	82
ΣX_i	1,502		1,125		2,092		1,568		1,183		1,805	
ΣY_i		1,000		859		995		792		839		787
ΣX_i^2	229,760		128,245		439,544		148,886		141,525		337,433	
ΣY_i^2		102,062		75,819		100,075		64,462		72,613		64,401
ΣX_iY_i	151,840		97,776		203,892		125,370		99,195		145,872	
總計	$\Sigma X=9,275$, $\Sigma Y=5,272$, $\Sigma XY=828,951$, $\Sigma X^2=1,525,393$, $\Sigma Y^2=479,432$.											

我們先要把所增體重與食量間的關係找出來，最簡單的關係是直線迴歸. 其初步計算附於表10–1的下端. ΣX_i 為各組 X 值的總和，

ΣX_i^2 爲各組 X 值平方後的總和, Y 方面的情形亦然. 至於 $\Sigma X_i Y_i$ 則爲同組各對 X, Y 值乘積的總和, 如

$$(108)(73)+(136)(102)+\cdots\cdots+(176)(111)=151,846.$$

再將各組計算結果相加, 即得該表'總計'欄內的五個總和. 有了這些總和值, 就可作進一步的計算.

校正數: X 方面 $(9,275)^2/60=1,433,760.42,$

Y 方面 $(5,272)^2/60=463,233.07,$

XY 方面 $(9,275)(5,272)/60=814,963.33.$

離均差平方和:

$$\Sigma(X-\bar{x})^2=1,525,393-1,433,760.47=91,632.58,$$

$$\Sigma(Y-\bar{y})^2=479,432-463,233.07=16,198.93.$$

離均差積之和:

$$\Sigma(X-\bar{x})(Y-\bar{y})=828,951-814,963.33=13,987.67.$$

上列結果係用公式(3-5)及(5-1)求得, 這些是六十頭公鼠總計的離均差平方和及積和. 其次求各組均數間的離均差平方和及積和:

$$\Sigma k_i(\bar{x}_i-\bar{x})^2=\frac{(1,502)^2+\cdots\cdots+(1,805)^2}{10}-1,433,760.42$$

$$=67,662.68,$$

$$\Sigma k_i(\bar{y}_i-\bar{y})^2=\frac{(1,000)^2+\cdots\cdots+(787)^2}{10}-463,233.07$$

$$=4,612.93,$$

$$\Sigma k_i(\bar{x}_i-\bar{x})(\bar{y}_i-\bar{y})=\frac{(1,502)(1,000)+\cdots\cdots+(1,805)(787)}{10}$$

$$-814,963.33=5,520.97.$$

此處 $\bar{x}_i, \bar{y}_i$ 爲各組均數，如第 1 組爲 $\bar{x}_1=150.2$, $\bar{y}_1=100.0$; 又 $\bar{x}=154.58$, $\bar{y}=87.87$, 爲六十頭鼠之總均數，k_i 爲各組所有鼠數，此處一律等於 10. 上列兩平方和係用公式 (9-1) 計算，至離均差積之和，則用下列公式計算：

$$\Sigma k_i(\bar{x}_i-\bar{x})(\bar{y}_i-\bar{y})=\frac{(\Sigma X_1)(\Sigma Y_1)}{k_1}+\frac{(\Sigma X_2)(\Sigma Y_2)}{k_2}+\cdots\cdots$$

$$+\frac{(\Sigma X_n)(\Sigma Y_n)}{k_n}-\frac{(\Sigma X)(\Sigma Y)}{\Sigma k_i}.$$

公式(10-1)

此例各組動物數相等，即 $k_1=k_2=\cdots\cdots=k_6$, 故計算式如上. 讀者若將各組均數，對於總均數之離均差，求其乘積而總和之，如

$$10(150.2-154.58)(100.0-87.87)+\cdots\cdots$$

$$+10(180.5-154.58)(78.7-87.87),$$

其結果應等於 5,520.97 (小數點後或稍有出入)，但實際計算時，以用公式(10-1)較爲便利而正確. 再次在總的變異及共變異(參閱第六章第 6-2 節)內，減去組間的變異及共變異，即得組內的離均差平方和及積和：

$$\Sigma\Sigma(X_i-\bar{x}_i)^2=91,632.58-67,662.68=23,969.90,$$

$$\Sigma\Sigma(Y_i-\bar{y}_i)^2=16,198.93-4,612.93=11,586.00,$$

$$\Sigma\Sigma(X_i-\bar{x}_i)(Y_i-\bar{y}_i)=13,987.67-5,520.97=8,466.70.$$

由上列結果，可計算迴歸係數及修正值了. 讓我們先算六十頭鼠總計時，由食量推求所增體重之迴歸. 按公式(5-2)及(5-3)，求得迴歸係數爲

$$13,987.67 \div 91,632.58 = .15265.$$

又 $\bar{x}=154.58, \bar{y}=87.87$，因得迴歸方程式為

$$Y-87.87=.15265(X-154.58),$$

或
$$Y=.15265X+64.28. \qquad (A)$$

若將食量化為相等，卽各鼠所消費之食物皆等於總均數 154.58 時，則按公式(5–7)，求得所增體重的修正值為

$$Y_a=Y-.15265(X-154.58),$$

或
$$Y_a=Y-.15265X+23.59. \qquad (B)$$

將表 10–1 各 X 值代入(B)式，如第 1 組之第一鼠，

$$Y_a=73-.15265(108)+23.59=80.11.$$

全部結果，載於表 10–2. 該表末二行為各組修正值之和及修正值平

【表 10–2】根據總迴歸求得各鼠所增體重之修正值(食量相等)

	組		別			
	1	2	3	4	5	6
Y_a	80.11	106.48	87.98	88.41	111.67	51.22
	104.84	79.74	72.37	74.56	104.09	78.58
	120.53	65.85	89.68	89.02	102.89	67.74
	103.32	113.07	91.37	63.32	86.50	87.92
	82.31	102.41	93.53	82.88	102.82	71.62
	109.07	94.50	95.68	53.99	80.80	90.06
	96.88	88.80	96.48	75.45	76.68	90.66
	87.85	82.74	80.70	84.59	74.57	67.19
	114.03	92.65	110.01	96.46	91.99	61.24
	107.73	96.97	93.79	79.92	62.36	81.17
ΣY_a	1,006.67	923.21	911.59	788.60	894.37	747.40
ΣY_a^2	102,994.69	86,954.11	83,991.04	63,633.43	82,265.55	57,429.81

方之和. 將六組相加，得

$$\Sigma Y_a=5,271.84, \quad \Sigma Y_a^2=477,268.63,$$

由此得修正值之總變異為

$$477,268.63-(5,271.84)^2/60=14,063.68.$$

再將組內的離均差平方和及積和，以同樣步驟求得迴歸係數為

$$8,466.70 \div 23,969.90 = .3532,$$

迴歸方程式為　　$Y = .3532\,X + 33.27,$　　(C)

修正值為　　$Y_b = Y - .3532X + 54.60.$　　(D)

此處修正值的符號用 Y_b，以別於前面的 Y_a. 將表 10–1 各 X 值代入 (D) 式，得表 10–3 的結果.

【表 10.3】根據組內迴歸求得各鼠所增體重之修正值(食量相等)

	組					別
	1	2	3	4	5	6
Y_b	89.45	117.63	80.08	86.32	117.80	54.15
	108.56	87.28	63.67	72.68	116.05	74.08
	123.86	78.81	81.37	87.73	110.63	60.85
	102.44	115.80	82.67	62.44	95.04	90.45
	84.03	112.16	82.43	78.79	109.16	59.31
	111.80	103.04	87.37	57.92	89.75	80.96
	92.79	97.75	81.36	79.98	80.21	70.53
	88.97	90.28	67.19	77.49	84.51	63.50
	110.14	101.39	96.90	98.39	95.92	61.56
	103.44	103.51	79.07	82.45	68.10	80.09
ΣY_b	1,015.48	1,007.65	802.11	784.19	967.17	695.48
ΣY_b^2	104,538.41	102,972.17	65,125.92	62,771.52	96,000.85	49,586.96
校正數	103,119.6	101,535.85	64,338.05	61,495.40	93,541.78	48,369.24
離均差平方和	1,418.45	1,436.32	787.87	1,276.12	2,459.07	1,217.72

該表校正數與離均差平方和之計算法同前. 第 1 組的十個公鼠所食飼料相同，其食量亦經修正值化為相等，但所增體重仍有多有少，這種組內變異，可認為是由機遇使然，換句話說，該組離均差平方和 1,418.45 所表示的變異是與食物的質與量都沒有關係的，其他各組

的組內變異亦然.六組的離均差平方和相加,得

$$1,418.45+1,436.32+\cdots\cdots+1,217.72=8,595.55.$$

構成此總和的各分子既都與飼料的質和量無關,則此總數 8,595.55,當亦為純粹機遇的結果.

至此,我們已求得了體重修正值的兩種變異:由表 10-2 所得的總變異 14,063.68,是把六十頭鼠放在一起求得的,這變異應與飼料優劣及機遇因子都有關係;另由表 10-3 所得的組內變異 8,595.55 卻純由機遇而來.那麼二者的相差 14,063.68－8,595.55＝5,468.13 可認為食量已化為相等後各組飼料的優劣所產生的變異了.因此可用變異數分析法以測驗其顯著性.

【表 10-4】 食量化為相等後,白鼠所增體重之變異數分析

變異來源	自由度	離均差平方和	均方	F
總變異	58	14,063.68		
組內(誤差)	53	8,595,55	162.2	
組間	5	5,468.13	1,093.6	6.74

各離均差平方和的自由度要說明一下.在總變異方面原有鼠 60 頭,在求離均差平方和時用一總均數,又在求修正值時用一迴歸係數(見方程式 B),共用了兩個統計數,故其自由度為 60－2＝58.在誤差方面,各組用了一個本組的均數,又用了一個公共的迴歸係數(見方程式 D),故其自由度為 6(10－1)－1＝53.餘下的 5 個自由度,屬於組間的變異(下面還有更詳細的分析).以自由度除離均差平方和得均方.最後以誤差的均方除組間的均方,得 F 值為 6.74. 查表 9-3,當 $n_1=5$, $n_2=53$ 時, F 的 1% 點約為 3.37,實際求得之 F 值大於 1%

點，故知當食量化為相等後，各組均數之間仍有非常顯著的差別．若設計時沒有其他因子夾雜在內，那麼我們可以說這六種飼料在質的方面有優劣之分了．

10–2. 共變數分析之計算法 上節修正均數顯著性的測驗，是用來說明共變數分析的原理的．實際計算時，我們無須經過這樣冗長的步驟．在第五章第 5–3 節裏，我們已說明，修正值的總和應與觀察值的總和相等，又修正值之標準差，即為標準估計誤差．所以修正值的均數等於觀察值的均數，而修正值的離均差平方和等於估計值誤差的平方和．我們就可利用這些特性，使共變數分析法得一捷徑．

第一步，求表 10–1 各組 X, Y 值的總和、平方之和及乘積之和，結果已載該表末行總計欄內．第二步，求校正數，再求離均差平方和及積和，連其自由度〔白鼠共 60 頭，食量方面用一 $\bar{x}$，故離均差平方和之自由度為 59；體重方面亦然；至離均差之積和，則原有 60 對數值，用去一對總均數 $(\bar{x}, \bar{y})$，故仍為 59.〕寫在表 10–5 左半節第一橫行內．第三步，求各組均數間的離均差平方和及積和，連自由度（6－1

【表 10–5】 共變數分析及修正均數之顯著性測驗

變異來源	自由度	離均差之平方和及積和			估計之誤差		
		$\Sigma(X-\bar{x})^2$	$\Sigma(X-\bar{x})\times(Y-\bar{y})$	$\Sigma(Y-\bar{y})^2$	平方和(1)	自由度	均方
總變異	59	91,632.58	13,987.67	16,198.93	14,063.72	58	
飼料	5	67,662.68	5,520.97	4,612.93			
組內(誤差)	54	23,969.90	8,466.70	11,586.00	8,595.37	53	162.2
修正均數顯著性之測驗					5,468.35	5	1,093.7

(1) $\Sigma(Y-\bar{y})^2-[\Sigma(X-\bar{x})(Y-\bar{y})]^2/\Sigma(X-\bar{x})^2$, $F=1{,}093.7/162.2=6.74$.

＝5)，寫在第二橫行內.第四步，將一、二兩橫行相減，即得組內的變異、共變異及自由度，寫在第三橫行內.這些在上節都已經計算好了.至於右半邊卻是重新計算的，也正是我們所說的捷徑.按公式(5–4)，估計值誤差的平方和爲

$$\Sigma(Y-\bar{y})^2-\frac{[\Sigma(X-\bar{x})(Y-\bar{y})]^2}{\Sigma(X-\bar{x})^2},$$

此式即等於修正值的離均差平方和.如

總變異方面 $16,198.93-\dfrac{(13,987.67)^2}{91,632.58}=14,063.72,$

誤差方面 $11,586.00-\dfrac{(8,466.70)^2}{23,969.90}=8,595.37.$

此兩數與上節由表10–2及表10–3求得的相同，而更爲正確，因上節在計算過程中，小數經過多次的四捨五入，故小數點後的數目與這裏的稍有出入，至於計算時的繁簡卻不可以道里計了.在求估計之誤差時，因另用一迴歸係數，故其自由度較左側少1.以上是第五步的計算.繼將總變異與誤差的兩個平方和相減，得5,468.35，其自由度爲 58－53＝5，是用來測驗修正均數的顯著性的.這是第六步的計算.最後以自由度除平方和，得均方，再以誤差的均方除修正均數的均方，得 F 值，由表9–3，查得 F 的 1%點，藉知修正均數間的相差爲非常顯著.

10–3. 各組人數不等時之共變數分析法 當各組人數不等時，共變數分析的原理，還是一樣的，不過計算方法稍有變化罷了.F. G. Benedict 氏等在燕京大學研究華人之基底代謝時(1)，曾測量該校師

生工友等之身高體重及每分鐘氧消耗量等. 本書著者以是項紀錄按 Stevenson 氏公式(2), 將身高、體重化爲身體表面積, 又依 Du Bois 氏之假定在基底狀況時每一升之氧消耗量相當於 4.8 仟卡之熱量, 而計算每一被試者每小時所產生之熱量. 表 10-6 爲該校男性師生之

【表10-6】 中國學界男子之身體表面積(X, 平方米)
及在基底狀況下每小時之發熱量(Y, 仟卡)

	29歲及以下		30至49歲		50歲及以上	
	X	Y	X	Y	X	Y
X, Y	1.554	69.19	1.848	65.14	1.503	52.40
	1.556	63.40	1.780	64.85	1.824	61.08
	1.453	64.27	1.382	49.22	1.737	55.29
	1.603	65.72	1.501	55.87	1.516	54.43
	1.635	64.85	1.741	68.90	1.457	53.85
	1.427	63.11	1.557	48.06		
	1.567	59.93	1.713	67.45		
	1.564	70.06	1.651	61.66		
	1.664	61.66	1.519	60.80		
	2.018	71.51	1.491	56.16		
	1.574	64.56	1.842	65.14		
	1.644	56.45	1.484	52.11		
	1.530	60.51	1.841	67.74		
			1.936	62.24		
總計	20.789	835.22	23.286	845.34	7.542	277.05
人數	13		14		5	
均數	1.599	64.248	1.663	60.381	1.508	55.410

資料.這裏我們可以看到,在基底狀況下,每小時之平均發熱量隨年齡而降低,不過發熱量與身體表面積是有密切關係的,我們必須把他們的表面積化爲相等,纔能相互比較.一般的習慣是把表面積除發熱量,化成每小時每平方米之發熱量,但所得結果只能憑表面數值作粗約的比較,而不易測驗相差的顯著性,其缺點正與前述動物飼養實驗中化成每單位食量所增的體重一樣.其實這種資料是應該用共變數來處理的.計算的步驟與上節大致相同,若還求各數值之平方及乘積,則用公式(3–5)及(5–1),即可得離均差之平方和及積和.不過這裏都是四位數,平方或相乘後就有七、八位數,相當麻煩,故最好用縮簡法.在 X 方面,我們把每個數值減去 1.5, Y 方面每個數值減去 60(這兩數是任意選擇的,總以在均數附近爲便).縮簡計算的結果,載於下表:

【表 10–7】 前表資料經縮簡後之計算結果

年齡組	人數	$\Sigma(X-1.5)$	$\Sigma(X-1.5)^2$	$\Sigma(Y-60)$	$\Sigma(Y-60)^2$	$\Sigma(X-1.5)\times(Y-60)$
29歲及以下	13	1.289	.373,341	55.22	450.2624	8.21796
30—49 歲	14	2.286	.766,968	5.34	632.2216	13.26632
50歲及以上	5	0.042	.456,314	−22.95	149.9579	2.42318
總計	32	3.617	1.596,623	37.61	1,232.4419	23.90746

根據此項結果,用公式(3–6)及(5–8),即可求得離均差平方和及積和.

總均數:

表面積 $\bar{x}=1.5+(3.617)/32=1.613,$

發熱量 $\bar{y}=60+(37.61)/32=61.175.$

校正數:

表面積 $(3.617)^2/32=.408,834$,

發熱量 $(37.61)^2/32=44.2035$,

乘　積 $(3.617)(37.61)/32=4.25111$.

總離均差平方和:

表面積 $1.596,623-.408,834=1.187,789$,

發熱量 $1,232.4419-44.2035=1,188.2384$.

總離均差積之和:

$$23.90746-4.25111=19.65635.$$

以上是三十二位被試者總計時的情形. 其次計算各年齡組間的變異及共變異, 這裏與上節不同的地方, 在各組人數不等, 故各總和平方或相乘後須分別以該組人數除之, 再將各商數相加, 然後減去校正數, 所用公式爲(9–3)及(10–1).

組間離均差平方和:

表面積
$$\frac{(1.289)^2}{13}+\frac{(2.286)^2}{14}+\frac{(0.042)^2}{5}-.408,834=.092,599,$$

發熱量
$$\frac{(55.22)^2}{13}+\frac{(5.34)^2}{14}+\frac{(-22.95)^2}{5}-44.2035=297.7314.$$

組間離均差積之和:

$$\frac{(1.289)(55.22)}{13}+\frac{(2.286)(5.34)}{14}+\frac{(.042)(-22.95)}{5}$$
$$-4.25111=1.90334.$$

於是以總計及組間各離均差平方和及積和寫在表10-8左半部第一、二兩橫行，將此兩行數值相減，卽得組內(誤差)之變異及共變異.在總變異方面，其自由度爲 $32-1=31$，組間爲 $3-1=2$，組內爲 $31-2=29$，或 $(13-1)+(14-1)+(5-1)=29$. 其次計算右半部'估計之誤差'，我們仍用公式(5-4)以求其平方和：

【表 10-8】 基底代謝資料之共變數分析

變異來源	自由度	離均差之平方和及積和			估計之誤差		
		$\Sigma(X-\bar{x})^2$	$\Sigma(X-\bar{x})(Y-\bar{y})$	$\Sigma(Y-\bar{y})^2$	平方和	自由度	均方
總計	31	1.187,789	19.65635	1,188.2384	862.9516	30	
組間	2	.092,599	1.90334	297.7314			
組內	29	1.095,190	17.75301	890.5070	602.7310	28	21.5261
修正均數顯著性之測驗					260.2206	2	130.1103
$F=130.1103\div21.5261=6.044$, $n_1=2$, $n_2=28$, 1%點$=5.45$.							

總變異 $1,188.2384-(19.65635)^2/1.187,789=862.9516$,

組　內 $890.5070-(17.75301)^2/1.095,190=602.7310$.

在此過程中，因爲各用一迴歸係數(總變異方面爲 16.5487，組內爲 16.2100，參閱第 10-1 節末段)，故其自由度各較左側少 1. 總變異與組內兩估計誤差之平方和相減，得 260.2206，此值表示身體表面積化爲相等後所得發熱量修正均數(見後)間之變異，其自由度爲 $30-28=2$. 以自由度除平方和，得均方，組內之均方爲 21.5261，此值係表面積化爲相等後各組內之機遇變異，故用爲比較之標準. 修正均數之均方爲組內均方之 6.044 倍，此卽爲 F 值. 查表 9-3，得知當自由度 $n_1=2$, $n_2=28$ 時，F 值之 1%點爲 5.45，此數小於觀察值. 故知當

各組表面積化爲相等後，每小時發熱量之均數間仍有非常顯著的差別，易言之，在基底狀況下每小時的發熱量隨年齡而降低，實不能用表面積的大小不等來解釋的.

上述各組修正均數是根據組內迴歸計算的，因組內迴歸是從同組的 X, Y 求得，其與年齡的關係甚少（若每歲爲一組時，則組內迴歸幾與年齡無關），故較能表示發熱量與表面積間的眞正關係. 由表 10–8 組內一橫行的數值，求得迴歸方程式及修正值如下：

迴歸方程式 $$Y-61.175=\frac{17.75301}{1.09519}(X-1.613),$$

或 $$Y=16.210X+35.028. \tag{E}$$

修正值 $$Y_a=Y-16.210(X-1.613),$$

或 $$Y_a=Y-16.210X+26.147. \tag{G}$$

以表 10–6 各組均數分別代入(G)式，得三組發熱量之修正均數爲：

64.475, 59.571, 57.112.

這是指身體表面積一律等於總均數 1.613 時各組發熱量的均數，由前面共變數測驗的結果，知這三均數有非常顯著的差別. 若用通俗的方法，將1.613除各修正均數，則得各組每小時每平方米的發熱量爲：

（已修正者） 39.97, 36.93, 35.41.

至於未經迴歸方法修正，而直接用表 10–6 各組的表面積均數除發熱量均數所得的結果則爲：

（未修正者） 40.18, 36.31, 36.74.

迴歸係數的計算，是根據最小二乘方原理的，故上列兩行數值，當以

已修正者較爲正確．其中以第三組的相差最多，計爲 36.74－35.41 ＝1.33 仟卡，此值不可謂不大．而且用通俗處理法所得的結論有時與用共變數分析法的未必盡同．也許有人認爲共變數分析法太繁，使初學者如墮五里霧中．苟讀者對於本書前面各章能充分了解，則閱讀本章應無困難．至於共變數分析所費的時間，若與全部實驗時間比較起來，眞說不上麻煩呢．

10-4．修正均數與共變數分析之其他用途　共變數分析，除用作測驗修正均數的顯著性外，尙可由此獲得其他的知識．這裏讓我們先把實際均數與修正均數比較一下．由表 10-1，各組公鼠所增體重的總和，卽可得其實際均數，再由表 10-3，則可得其修正均數（因爲組內迴歸與飼料的優劣無關，故能表示食量與所增體重間的眞正關係），玆併列於後：

組　別	1	2	3	4	5	6
實際均數	100.0(1)	85.9(3)	99.5(2)	79.2(5)	83.9(4)	78.7(6)
修正均數	101.5(1)	100.8(2)	80.2(4)	78.4(5)	96.7(3)	69.5(6)

各均數後面括弧裏的數目表示大小的等第，這裏可以看到兩方面的等第大體相同，得到最高與最低等的組別是一樣的，其中等第變動較大的是第三等，在實際均數爲第 2 組，而食量化爲相等後則爲第 5 組，這種等第方面的變動，營養學家從飼料的成分上或可找到解釋，在統計學上只是把這個現象指出來吧了．

關於變異數分析與共變數分析的分野，這裏也順便說一說．若我們只要比較各組公鼠所增體重有無差別，而不管食量的多少，那麼只須把各組的 Y 值用變異數分析法測驗其顯著性就行了．若須

將各組食量化爲相等，然後比較其所增的體重，那就得用共變數分析了．在類似的實驗中是否需要引用自變數 X（如飼養實驗中之食量或原始體重等），要解答這問題須從三方面考慮：（一）比較實際均數與修正均數，若兩方面的大小等第很不相同，就可獲得相當解釋．（二）比較表 10-5 修正均數間的平方和（5,468.35）與實際均數間的平方和（4,612.93）．前者食量相等，而後者則否．這裏前者較後者爲大，但也可能有相反的情形．若這兩個平方和的數值差別很大，就值得注意．（三）檢視誤差方面由迴歸所減少的變異是否顯著．當未修正

【表 10-9】 誤差變異數之分析（參閱表 10-5）

變異來源	自由度	離均差平方和	均方
實際所增體重之誤差	54	11,586	215
修正後之誤差	53	8,595	162
由迴歸所減少之誤差	1	2,991	2,991
$F=2,991\div162=18.$			

時，公鼠所增體重之誤差爲 11,586，其自由度爲 54．內中一個自由度表示由迴歸產生的變異，另 53 個自由度則由機遇而來．由表 10-9，可見誤差的均方，用共變數後不僅從 215 減到 162，而且因迴歸而減少的誤差爲非常顯著，由此可見實驗的精密度大有改進．若上述三方面都得不到更多的知識，如實際與修正均數的等第完全一樣；實際均數間與修正均數間的兩平方和也無甚差別；誤差方面因迴歸而減少的部分也並不顯著．這時自變數 X 似可不必要了．

10-5. 各組組內之共變數 在實驗設計時，若早已準備用共變數分析法來處理的資料，則經本章第 10-2 或 10-3 節的顯著性測驗

後，統計工作卽可告一段落. 但統計人員也常常遇到若干資料，在設計時並未考慮到用共變數的，這時就要把各組分別求其共變數，並找出各組間的關係來，因爲從幾組裏找出的關係，或許比單獨一個顯著性測驗更合適些. 而更重要的，是要明瞭此項測驗的詳細結構. 請分述如下.

先將表 10-1 的資料，分別求各組內的迴歸. 例如第一組，

校正數 X 方面 $(1,502)^2/10=225,600.40$,

Y 方面 $(1,000)^2/10=100,000.00$,

XY 方面 $(1,502)(1,000)/10=150,200.00$.

離均差平方和

$\Sigma(X-\bar{x})^2$ $229,760-225,600.40=4,159.60$,

$\Sigma(Y-\bar{y})^2$ $102,062-100,000.00=2,062.00$.

離均差積之和

$\Sigma(X-\bar{x})(Y-\bar{y})$ $151,846-150,200.00=1,646.00$.

【表 10-10】 各組公鼠所增體重與食量間之迴歸及相關

組別	自由度	離均差之平方和及積和 $\Sigma(X-\bar{x})^2$	$\Sigma(X-\bar{x})(Y-\bar{y})$	$\Sigma(Y-\bar{y})^2$	相關係數[1]	迴歸係數[2]	估計之誤差 平方和[3]	自由度
1	9	4,159.60	1,646.00	2,062.00	0.5620	0.3957	1,410.66	8
2	9	1,682.50	1,138.50	2,030.90	0.6159	0.6767	1,260.51	8
3	9	1,897.60	738.00	1,072.50	0.5173	0.3889	785.48	8
4	9	3,023.60	1,184.40	1,735.60	0.5170	0.3917	1,271.65	8
5	9	1,576.10	−58.70	2,220.90	−0.0314	−0.0372	2,218.71	8
6	9	11,630.50	3,818.50	2,464.10	0.7133	0.3283	1,210.42	8
總計	54	23,969.90	8,466.70	11,586.00	0.5081	0.3532	8,157.43	48

1. 相關係數 $=\Sigma(X-\bar{x})(Y-\bar{y})/\sqrt{\Sigma(X-\bar{x})^2\Sigma(Y-\bar{y})^2}$,
2. 迴歸係數 $=\Sigma(X-\bar{x})(Y-\bar{y})/\Sigma(X-\bar{x})^2$,
3. 估計誤差之平方和 $=\Sigma(Y-\bar{y})^2-[\Sigma(X-\bar{x})(Y-\bar{y})]^2/\Sigma(X-\bar{x})^2$.

其餘各組，依此類推．計算結果詳表 10–10 之左半部．由各組離均差之平方和及積和，即可計算相關係數、迴歸係數及估計誤差之平方和（公式 6–1，5–2 及 5–4 附錄於表末）．這裏可以看到六組的迴歸係數相當參差，相關係數亦然．我們要知道他們的差別是否顯著，還要設法把六組的係數平均起來．爲醒目起見，我們把各組迴歸線繪於圖 10–1．例如第 1 組的迴歸方程式爲

$$Y=100.0+0.3957(X-150.2)=0.3957X+40.57.$$

餘類推．有了迴歸方程式，即可作圖，作圖的方法詳第五章第 5–1 節．圖中迴歸線旁所註數目表示組別，線上'＋'字形爲該組均數 $(\bar{x}_i, \bar{y}_i)$

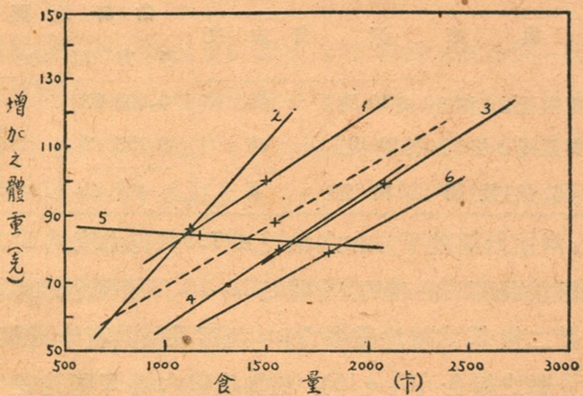


圖 10-1．由各組公鼠食量推算其所增體重之迴歸．

所構成之點．另有一條虛線，則爲用方程式(C)所畫的迴歸線．這裏值得注意的，便是表 10–10 總計行內的離均差平方和及積和與表 10–5 組內一行的數值完全一樣，這證明我們的計算沒有錯，同時說明表 10–5 組內一行的離均差平方和及積和實爲六組相加的結果，其離均

差是根據各該組均數計算的. 所以表 10–5 的組內迴歸(見迴歸方程式 C)是六組迴歸的總平均. 試看圖 10–1 的那條虛線, 在其他六條線的中間, 迴歸係數也有類似的情形. 至此將六個迴歸係數平均的目的已經達到了.

再看表 10–5 組內估計誤差之平方和爲 8,595.37, 此值是用公式(5–4)計算出來的, 而表 10–10 六組估計誤差之平方和相加, 得 8,157.43, 後者是根據各該組均數計算的, 按最小二乘方原理, 此值應爲最小, 這是後者小於前者的原因. 今將此二值並列於表 10–11,

【表 10–11】 由平均組內迴歸所得估計誤差之分析

變異來源	自由度	估計之誤差	
		平方和	均方
與平均組內迴歸之相差(表 10-5)	53	8,595.37	
與各該組迴歸之相差(表 10-10)	48	8,157.43	169.9
各迴歸係數間之相差	5	437.94	87.6

作一比較, 兩估計誤差平方和之相差 437.94(自由度爲 5)乃表示六個迴歸係數彼此間的相差的. 若將表 10–10 每兩個迴歸係數相差之平方乘以該二組 X 之離均差平方和, 此種乘積共有十五個(卽 6 數中每次取 3 數之組合, ${}_6C_2=15$), 相加後以六組 X 離均差平方和之總計除之(3), 應等於 437.94. 卽

$$
\begin{aligned}
&[(4,159.6)(1,682.5)(.3957-.6767)^2\\
&\quad+(4,159.6)(1,897.6)(.3957-.3889)^2+\cdots\cdots\\
&\quad+(1,576.1)(11,630.5)(-.0372-.3283)^2]\div(23,969.9)\\
&\quad=437.94.
\end{aligned}
$$

再以自由度 5 除之，得均方 87.6，此值卽用以測驗迴歸係數間相差的顯著性，測驗時應與均方169.9相比較，因此值與飼料的優劣及各組迴歸無關（參閱第 9–1 節及第 5–2 節）．但此例中前者反較後者爲小，足見六個迴歸係數實非常接近．

各迴歸係數之間旣無顯著的差別，可知它們都從一個均勻的全體中取來，於是前面求得的平均迴歸係數，有了合理的根據（嚴格的說，惟有來自同一全體的各樣本，其平均纔有意義，一般人對於這點往往疏忽的）．現在我們可以把平均迴歸係數（.3532）當作全體迴歸（population regression）的最佳估計值，而各組的迴歸係數，就無須分別考慮了．

在白鼠所增體重與食量的關係方面，各樣本旣由同一全體而來，因此我們可以把各組的相關係數平均起來，平均的方法就是用表 10–10 總計行內的離均差平方和與積和，依公式（6–1）計算，得 $r=0.5081$，此值較由六十頭鼠直接計算的（0.3631）爲大，這是由於分母方面減去了組間變異的結果．

現在我們再翻到表 10–5，用標有‘飼料’一行的數值來計算各組均數間的相關係數、迴歸係數及估計誤差之平方和：

$$r=5,520.97\div\sqrt{(67,662.68)(4,612.93)}=0.3125,$$

$$b=5,520.97\div 67,662.68=0.08160,$$

$$\text{估計誤差之平方和}=4,612.93-(5,520.97)^2\div 67,662.68$$

$$=4,612.93-450.49=4,162.44.$$

最後一數 4,162.44，是各組所增體重均數（100.0，85.9，……，78.7）與組間迴歸線

$$Y=87.87+0.08160(X-154.6)=0.08160X+75.25 \qquad \text{(H)}$$

距離的平方和(參考第 5–2 節及圖 5–1). 又各鼠所增體重與平均組內迴歸線(見第10–1節)

$$Y=0.3532X+33.27 \qquad \text{(C)}$$

距離的平方和爲 8,595.4(見表 10–5). 前者的自由度爲 $6-2=4$, 後者的自由度爲 53. 按表 10–5, 總的估計誤差之平方和爲 14,063.7, 共 58 個自由度. 今上述兩部分占 57 個自由度, 尙餘 1 個自由度, 其平方和爲 1,305.9, 是由(H)與(C)兩迴歸的相差而來的,

$$\frac{(67,662.68)(23,969.90)(.0816-.3532)^2}{67,662.68+23,969.90}=1,305.90.$$

上式分母部分爲表 10–5 所載 X 之兩離均差平方和, 分子的右邊是兩迴歸係數相差之平方, 詳細說明同前 (參閱表 10–11 之說明). 茲將各估計誤差倂列於表 10–12, 以平均組內迴歸之均方爲比較的標

【表 10–12】 三種迴歸的估計誤差之分析

變異來源	自由度	估計誤差之平方和	均方	F
總迴歸(見表 10-5)	58	14,063.7		
各組均數間之迴歸(方程式 H)	4	4,162.4	1,040.6	6.416
平均組內迴歸(方程式 C)	53	8,595.4	162.2	
(H)與(C)兩迴歸之相差	1	1,305.9	1,305.9	8.051

準, 所得兩 F 下值均爲非常顯著. 假如組間迴歸之估計誤差之均方 (1,040.6)爲不顯著, 而兩迴歸係數相差之均方 (1,305.9) 爲顯著, 前者表示各組所增體重均數與迴歸線(H)之距離甚近, 後者表示組間與組內兩迴歸有差別, 易言之, 各組之內, 各鼠食量對於所增體重

之關係與由各組均數(如 150.2, 100.0;112.5,85.9 等,見表 10-1)所構成者不同. 但此例組間迴歸之估計誤差,既非常顯著,卽組均數間並無明顯的趨勢可言〔另法,表 10-5,組均數(飼料)間之總變異爲 4,612.93,此值可分爲兩部分,其由迴歸而來者爲 450.49,估計誤差爲 4,162.44,以其相當之自由度 1 與 4 分別除之,得均方爲 450.49 與 1,040.6,前者小於後者,故迴歸爲不顯著〕,則兩迴歸之差別也就沒有什麼意義了. 要記得只在組內(組內迴歸之顯著性測驗,見表 10-9)與組均數間兩迴歸都有明顯趨勢的時候,兩迴歸之相差,纔有適當的解釋.

10-6. 兩組資料之共變數 當實驗中所用的被試者,只分兩組時,我們仍舊可用共變數分析法,其原理是和前面一樣的,不過計算的過程要簡單些. 這裏可以問兩個問題:(1)兩個修正均數間有無顯著的差別?(2)兩個組內迴歸係數間有無顯著的差別?玆仍以公鼠飼養實驗爲例,表 10-1 前三組的飼料所含蛋白質的成分高,後三組所含蛋白質的成分低. 玆將前後三組各自合併,只比較高低兩大組有無差別. 合併後的總和,見表 10-13.

【表 10-13】 兩組公鼠之食量及所增體重(參閱表 10-1)

蛋白質之成分	ΣX	ΣY	ΣX^2	ΣXY	ΣY^2
高	4,719	2,854	797,549	458,514	277,956
低	4,556	2,418	727,844	370,437	201,476
總計	9,275	5,272	1,525,393	828,951	479,432

由此計算各組(每組有鼠 30 頭)的校正數,再求組內離均差平方和及積和,並仿表 10-10 計算相關係數、迴歸係數及估計誤差等,結果

列於表 10-14.

【表 10-14】 兩組公鼠之迴歸係數及相關係數等

蛋白質之成分	自由度	離均差之平方和及積和 $\Sigma(X-\bar{x})^2$	$\Sigma(X-\bar{x})\times(Y-\bar{y})$	$\Sigma(Y-\bar{y})^2$	相關係數	迴歸係數	計估之誤差 平方和	自由度
高	29	55,250.30	9,579.80	6,445.47	0.5076	0.1734	4,784.44	28
低	29	35,939.47	3,223.40	6,585.20	0.2095	0.0897	6,296.09	28
總　計							11,080.53	56
組　內(平均)	58	91,189.77	12,803.20	13,030.67	0.3714	0.1404	11,233.03	57

高低兩組的修正均數為 94.7 與 81.0,其顯著性測驗見表 10-15, 結果兩修正均數的差別為非常顯著.

【表 10-15】 兩修正均數相差之顯著性測驗

變異來源	估計之誤差 自由度	平方和	均方	F
總迴歸(表 10-5)	58	14,063.72		
高低兩組之平均組內迴歸(表 10-14)	57	11,233.08	197	
修正均數間	1	2,830.64	2,831	14.37

高低兩迴歸係數(見表 10-14)相差的顯著性測驗,見表 10-16, 結果兩迴歸係數頗為相近(均方 152 小於 198).

【表 10-16】 兩迴歸係數相差之顯著性測驗

變異來源	估計之誤差 自由度	平方和	均方
與平均組內迴歸之相差(表 10-15)	57	11,233	
與各組迴歸之相差(表 10-14)	56	11,081	198
兩迴歸係數之相差	1	152	152

表10–16是與表10–11相當的. 我們記得前面六組相比較時，表10–5修正均數方面的5個自由度，在表10–12又析爲兩部分：有4個自由度屬於各組均數間迴歸之變異，另1個自由度屬於組間與組內兩迴歸係數的變異. 今表10–15只有高低兩組，故修正均數間的變異只有1個自由度，這自由度卽屬於組間與組內兩迴歸係數(表10–17)

【表10–17】 組間與組內兩迴歸

變異來源	自由度	離均差之平方和及積和			迴歸	估計之誤差	
		$\Sigma(X-\bar{x})^2$	$\Sigma(X-\bar{x})\times(Y-\bar{y})$	$\Sigma(Y-\bar{y})^2$	係數	平方和	自由度
總變異(表10–5)	59	91,632.58	13,987.67	16,198.93		14,063.72	58
組內(表10–14)	58	91,189.77	12,803.20	13,030.67	0.1404	11,233.08	57
組間	1	442.81	1,184.47	3,168.26	2.6749	0	0
修正均數間(表10–15)						2,830,64	1

的相差的. 這點可用下式說明(參閱上節)：

$$\frac{(442.81)(91,189.77)(2.6749-.1404)^2}{442.81+91,189.77}=2,830.7.$$

還有一點值得注意的，卽表10–17組間迴歸的估計誤差爲0，因爲這裏只有兩組，兩組的均數就決定這根迴歸線，它們都在迴歸線上，故估計誤差等於零. 又因沒有其他的均數點在迴歸線外，所以也沒有誤差的自由度了.

10–7. 隨機區組之共變數 前章第9–3節，曾提到隨機區組的實驗，在這種實驗中，若有 X, Y 兩變數，而要把 X 化做相等，然後測驗各組 Y 均數間相差的顯著性時，就得用共變數分析法. 鄭集與王德寶(5)兩氏曾比較米(R)，麥(W)，米加蔬菜(RV)，及麥加蔬菜

(WV)四種飼料之營養價值.二氏將白鼠 24 頭,先分為六個配偶組(matched groups),屬於同一配偶組之四頭白鼠,其性別相同,體重、窩別亦大致相等.然後以隨機方法〔用 Tippett(6),Fisher and Yates(7)等隨機數目表,或本書第一章、第四章的方法均可〕指定各鼠一種飼料,分派結果,每種飼料各有 6 鼠.惟在實驗過程中發現三鼠有問題,乃拋棄其紀錄,而用前章第 9-3 節的方法,將所缺 X, Y 值分別塡補之.表10-18 有括弧的各數,即為塡補之數值.(塡補缺項,使資料之分析過程益形複雜,精細校正更為費事,故非不得已時,以少用為宜.)

這裏也先算三個校正數:

X 方面 $(16,641.3)^2 \div 24 = 11,538,869.40$,

Y 方面 $(898.2)^2 \div 24 = 33,615.14$,

XY 方面 $(16,641.3)(898.2) \div 24 = 622,800.65$.

【表 10-18】 白鼠之食物消費量(克,X)及所增體重(克,Y)

	R		W		RV		WV		各配偶組總和	
	X	Y	X	Y	X	Y	X	Y	X	Y
	619.0	11.0	722.9	45.7	630.4	42.1	935.2	101.0	2,907.5	199.8
	636.4	16.6	829.3	55.3	683.6	38.5	779.7	62.2	2,929.0	172.6
	562.0	−3.3	859.8	65.5	610.6	19.4	921.4	92.7	2,953.8	174.3
	502.9	−7.7	(723.8)	(33.5)	634.7	20.0	788.0	52.9	2,649.4	98.7
	449.0	−7.5	(697.8)	(42.6)	626.5	31.2	(772.0)	(68.8)	2,545.3	135.1
	603.4	6.4	695.5	35.0	579.4	19.2	778.0	57.1	2,656.3	117.7
ΣX	3,372.7		4,529.1		3,765.2		4,974.3		16,641.3	
ΣY		15.5		277.6		170.4		434.7		898.2
$\bar{x}$	562.12		754.85		627.53		829.05		693.38	
$\bar{y}$		2.58		46.27		28.4		72.45		37.43

$\Sigma X^2 = 11,889,038.67$, $\Sigma Y^2 = 52,991.28$, $\Sigma XY = 701,328.65$.

在表 10–18 末行所載 ΣX^2, ΣY^2 及 ΣXY 內減去相當的校正數，卽得總變異的離均差平方和及積和，載於表 10–19 的第一行．其第二行配偶組的三個數值是這樣計算的:

X 方面

$$\frac{(2,907.5)^2+\cdots\cdots+(2,656.3)^2}{4}-11,538,869.40=38,964.06,$$

Y 方面

$$\frac{(199.8)^2+\cdots\cdots+(117.7)^2}{4}-33,615.14=1,869.43,$$

XY 方面

$$\frac{(2,907.5)(199.8)+\cdots\cdots+(2,656.3)(117.7)}{4}-622,809.65$$

$$=7,030.24.$$

同理，由各飼料組之總和，可求得表 10–19 第三行的各數，因每種飼

【表 10–19】 隨機區組實驗之共變數分析

（白鼠之食物消費量及所增體重）

變異來源	自由度	離均差之平方和及積和			估計誤差		
		$\Sigma(X-\bar{x})^2$	$\Sigma(X-\bar{x})\times(Y-\bar{y})$	$\Sigma(Y-\bar{y})^2$	平方和	自由度	均方
總變異	20	350,169.27	78,528.00	19,376.14			
配偶組	5	38,964.06	7,030.24	1,869.43			
飼料	3	262,504.54	62,778.23	15,601.90			
誤差	12	48,700.67	8,719.53	1,904.81	343.64	11	31.24
飼料+誤差	15	311,205.21	71,497.76	17,506.71	1,080.48	14	
修正均數間之變異					736.84	3	245.61
$F=245.61\div31.24=7.86$, 自由度=3與11, 1%點=6.22.							

料用鼠 6 頭，故各總和平方或乘積之和，應以 6 除之. 最後在總變異內減去配偶組與飼料兩部分，卽得表 10–19 的誤差.

計算自由度時，塡補之三鼠不占自由度，故總變異方面爲 $24-3-1=20$，又配偶組方面爲 $6-1=5$，飼料方面爲 $4-1=3$，而誤差方面則爲 $20-5-3=12$.

表 10–19 第五行是以前沒有講過的. 我們旣然把第一行總變異分成兩個以上的部分，現在就需要一個新的總變異，包含飼料與誤差兩部分的，因此我們把第三、四兩行的自由度、離均差平方和及積和逐項相加，卽得第五行的數值. 於是用公式(5–4)求第四、五兩行的估計誤差平方和(其自由度各較左側少 1). 最後求二者之差，得 736.84，而其相當的自由度爲 3，這是用來測驗四個修正均數〔由表 10–19 第四行求得迴歸係數爲 .17904，修正值爲

$$Y_a = Y - .17904(X - 693.38),$$

以表 10–18 各均數代入，得修正均數爲

26.08, 35.27, 40.19, 及 48.16.〕

的顯著性的. 以誤差的均方除修正均數的均方，得 $F=7.86$，此值大於1%點，故知當食量化爲相等後，此四組白鼠所增之體重仍有非常顯著的差別，卽食'麥與蔬菜'者最優，'米與蔬菜'及'麥'兩組者次之，而純食'米'者最劣.

【練　習　題】

1. 下表爲爲便於計算而假定之資料，試測驗其修正均數之顯著性.

組別	1		2		3		4	
	X	Y	X	Y	X	Y	X	Y
X 或 Y	29	22	15	30	16	12	5	23
	20	22	9	32	31	8	25	25
	14	20	1	26	26	13	16	28
	21	24	6	25	35	25	10	26
	6	12	19	37	12	7	24	23

2. 前章第4題鄭集、戴重光二氏之實驗，各鼠在實驗期間所消費之飼料如下. 若將食量化為相等後，問各組所增體重之均數間仍有顯著的差別否?

白鼠在三十日內之食物消費量(克,X)及所增體重(克,Y)

組別	甲組		乙組		丙組	
飼料	黃豆蛋白質		黃豆蛋白質與酪蛋白		黃豆蛋白質與卵蛋白	
	X	Y	X	Y	X	Y
實驗紀錄	197.1	23.5	225.1	41.3	279.8	60.9
	193.9	15.9	231.2	40.7	275.6	62.1
	213.6	6.6	247.5	28.8	263.9	48.9
	197.6	9.6	202.5	18.8	337.6	61.2
	236.0	8.4	238.5	21.7	326.9	66.9
	239.3	11.4	239.2	25.1	339.4	73.7
	252.6	20.4	255.4	22.2		

3. 下表為實驗心理學(experimental psychology)之四種實驗紀錄(4)，其目的係測量被試者在四種實驗情境下的情緒反應: (1)安坐休息; (2)實驗 2, 令被試者閱讀一短篇故事，這故事突然結束，非常驚人; (3)實驗 3, 呈現很可怕的刺激, 如大蛇等; (4)實驗 16, 放映

各種活動電影。表內 X 爲呼吸速率(一定時間之呼吸次數)，Y 爲同時間內不規則呼吸(atypical breath)之次數。問呼吸速率化爲相等後，各實驗情境下不規則呼吸次數之均數爲何？有顯著的差別否？

四種實驗情境之呼吸速率(X)及不規則呼吸次數(Y)

	休息		實驗2		實驗3		實驗16	
	X	Y	X	Y	X	Y	X	Y
	18	4	18	6	19	10	16	3
	16	9	17	10	15	7	15	10
	21	1	17	1	28	8	22	6
	16	3	17	3	22	9	14	6
	26	2	23	11	22	9	27	1
	16	14	11	8	20	15	8	8
	18	4	16	6	27	10	10	4
	20	4	22	7	29	11	24	2
	16	9	18	9	27	26	18	1
	22	6	19	3	22	7	22	6
	12	7	19	6	14	14	20	1
	18	0	18	2	20	10	19	3
	17	17	16	4	7	7	22	4
	18	0	21	16	16	11	22	0
	21	4	13	5	20	19	20	6
	19	7	18	1	10	7	24	10
	15	6	18	8	20	10	19	0
	16	8	13	1	12	10	15	6
	16	12	16	5	23	13	17	3
	16	11	18	6	25	12	24	0
	18	4	16	6	28	20	19	0
	23	2	14	5	21	10	25	0
	18	18	15	7	21	16	11	11
	20	16	20	7	25	14	23	7
	15	4	20	10	20	9	10	8
	17	0	19	6	16	14	16	10
	16	3	17	2	16	13	17	0
	16	4	16	6	16	3	15	0
	20	0	18	3	28	10	25	0
	15	4	16	3	17	16	10	7
總計	535	183	519	173	606	350	549	123
均數	17.83	6.10	17.30	5.77	20.20	11.67	18.30	4.10

4. 上題呼吸速率與不規則呼吸次數間之關係在四種實驗情境下相同否(參閱表 10-11)？

5. 第3題各組不規則呼吸次數(Y)之均數與組間迴歸線之距離為顯著否(參閱表10-12)?

6. 設第3題資料中同一橫行內之四對數值為同一被試者之紀錄,則此項設計即為隨機區組之實驗,試仿第10-7節,測驗其修正均數之顯著性.所得結論與第3題相同否?

【參考文獻】

(1) Benedict, F.G., Kung, L.C. and Wilson, S.D., The basal metabolism and urinary nitrogen excretion of Chinese, Manchus and others of the Mongolian race, Chinese Journal of Physiology, 12, 1 : 67–100 (1937).

(2) Stevenson, P.H., Height-weight-surface formula for the estimation of surface area in Chinese subjects, Chinese Journal of Physiol., 12, 3 : 327–330 (1937).

(3) Goulden, C.H., Methods of Statistical Analysis, pp. 250–254, John Wiley and Sons, New York (1939).

(4) Harold V. Gaskill and Gertrude M. Cox., Patterns in emotional reactions : I. Respiration; The use of analysis of variance and covariance in psychological data, The Journal of General Psychology, 16, 1st half : 21–38(1937).

(5) Cheng, L.T. and Wang, T.P., Journ. Chinese Chem. Soc., 11, 2.

(6) Tippett, L. H. C., Random Sampling Numbers, Cam-

bridge University Press, London (1927).

(7) Fisher, R. A. and Yates, F., Statistical Tables for Biological, Agricultural and Medical Research, Oliver and Boyd, London (1938).

第十一章　多元迴歸

本書最初討論一組含有一個變數的資料，繼將變數推廣至二個，組數先擴充爲二組，再擴充至二組以上．同組兩變數間的關係，在直線迴歸和相關兩章裏已加討論，多組兩變數間的關係在共變數分析章裏也已經詳細講過了．本章擬將變數數目繼續推廣至二個以上，把一個變數與其他幾個變數間的關係找出來，表示這種關係的方法叫做多元迴歸(multiple regression 或譯作複迴歸)．

11-1．含有三個變數之資料　含有三個變數的資料，常見的如身高、體重與身體表面積；身高、體重與肺活量 (vital capacity)；體重、胸圍與肺活量；飼養實驗中如食量、原始體重與所增體重；原始年齡、原始體重與所增體重等等．如果我們要找出一個方程式，可以從兩個變數來推算第三個變數，例如由身高、體重來推算身體表面積，那麼參加實驗的被試者，必須每一個人都受到這三種測量，有了這三種測量的紀錄，纔能把所要的方程式算出來，若祇有身高、體重的紀錄，而要編造一個由身高、體重推算身體表面積的方程式，是不可能的．

表 11-1 (1) 爲我國成年男女之身高、體重及由 X 光測量所得之心像面積(area of cardiac image)，所謂心像面積，是指在一定距離(二米)處 X 光片上心影(後前位 antero-posterior position 心像)之

【表 11-1】 國人年齡、身高、體重及心像面積之實測紀錄

年齡 (歲)	身高 (厘米) X_1	體重 (仟克) X_2	實測心像面積 (平方厘米) Y	年齡 (歲)	身高 (厘米) X_1	體重 (仟克) X_2	實測心像面積 (平方厘米) Y
19	154	42.0	79.0	25	172	54.3	89.0
19	161	49.5	87.5	26	165	48.0	83.0
19	167	49.5	90.5	26	168	59.4	94.5
20	158	48.5	59.5	26	175	58.4	94.5
20	150	41.0	76.5	27	157	43.2	72.5
20	162	48.5	89.5	28	151	42.0	76.0
20	165	42.8	77.5	28	171	51.0	82.5
20	174	59.1	95.0	28	178	50.3	98.5
20	178	53.2	105.0	29	169	51.0	80.0
20	179	59.7	106.5	29	173	55.5	91.5
21	156	42.2	72.0	30	161	73.0	86.5
21	158	50.5	72.5	30	157	45.2	77.5
21	165	53.3	87.0	30	160	52.0	89.5
21	169	49.8	97.0	30	169	63.4	85.0
22	152	60.0	88.5	31	156	40.0	55.5
22	156	50.3	78.5	31	166	44.0	83.0
22	163	58.8	98.0	31	160	47.2	86.5
22	164	50.4	89.0	31	168	52.3	99.0
22	172	46.7	67.5	31	170	66.0	107.0
22	172	59.9	90.0	32	168	65.5	82.0
22	174	59.0	86.5	32	167	54.3	98.7
23	161	42.5	88.0	33	160	47.2	88.5
23	169	54.4	86.0	34	165	51.1	72.0
23	174	57.0	98.5	35	152	53.5	83.0
23	177	55.9	97.0	36	151	37.6	59.0
24	160	46.8	82.5	37	158	39.0	73.0
24	147	39.0	77.0	38	156	55.5	82.0
24	159	43.5	80.0	38	161	40.0	82.0
24	162	44.3	74.0	39	154	41.0	73.0
24	172	57.0	93.0	39	166	68.0	100.5
24	170	56.5	86.0	42	161	50.0	90.5
25	167	50.0	85.0	42	177	61.5	105.0
25	162	42.0	91.5	46	156	46.0	77.0
25	161	56.0	76.5	50	162	52.7	94.0
25	163	59.5	83.0	50	150	54.5	74.5
25	150	43.2	76.5	52	148	55.7	84.0
25	159	42.6	89.0	56	158	46.1	88.0
25	167	51.4	87.0	57	161	46.3	70.0
25	168	51.8	77.0				

面積，而非心臟之表面或內部的眞實面積. 在臨床方面，要診斷心臟

是否增大，必須有一比較的標準，如西人 Hodges 與 Eyster 二氏(3)曾發表由年齡、身高、體重以推算心像面積之方程式.從這方程式推算出來的數值表示健康人心像的平均面積.若某病人的實測心像面積(即在二米距離處 X 光片上之後前位心像，用求積儀 planimeter 測得之面積)，較估計面積大得很多，就知道他的心臟確實增大，這種方法可以補充普通診斷之不足.不過估計所得面積必須與實際情形非常接近，纔不會得到錯誤的結論.據鄒仲、陳又新二氏(1)的研究，用 Hodges-Eyster 氏方程式以估計國人之心像面積，每較實測面積爲大，因此若以西人標準應用於國人，則國人之多數心臟增大有被忽視之虞.於是二氏建議將 Hodges-Eyster 氏方程式所得的估計值扣除 11%，俾能應用於國人.本書第五章第 5–5 節早已說過：根據外國人的資料所得的迴歸方程式，不能適用於中國人.至於鄒、陳二氏的建議，從統計學上看來也是欠妥的，本書作者已另文詳加討論(2)，這裏不必贅述.二氏既有健康國人之身高、體重及實測心像面積，理應用多元迴歸法以求得推算國人心像面積的方程式.

11–2. 三元迴歸方程式之計算法　在推算心像面積之迴歸方程式中，年齡一項應否列入，爲一先決問題，作者乃求年齡與心像面積之相關係數，得 $r=-.0956$，其自由度爲 75，查表 6–3，知此值爲不顯著，故就此資料言，年齡與心像面積之間並無眞正的關係，因此所求迴歸方程式中，年齡一項可無須列入.茲以 X_1 爲身高，X_2 爲體重，Y 爲心像面積.先計算均數、相關係數、離均差平方和及積和等，此爲求多元迴歸方程式之準備工作.離均差平方和之計算可用公式(3–5)或(3–6)，離均差積之和可用公式(5–1)或(5–8).作者計算時

係用縮簡法，將身高 X_1 各減去 160，體重 X_2 各減去 50，心像面積 Y 各減去 85. 身高簡縮數的總和爲 254.0，其平方之和爲 5,562.00；又身高簡縮數與體重簡縮數乘積之和爲 2,575.1. 詳細結果載於表 11-2. 將此初步結果代入公式 (2-2)、(3-6)、(5-8)及(6-1)，求得均

【表 11-2】 前表資料縮簡後之初步計算結果

$\Sigma(X_1-160)=254.0,$	$\Sigma(X_1-160)^2=5,562.00;$
$\Sigma(X_2-50)=79.8,$	$\Sigma(X_2-50)^2=4,373.64;$
$\Sigma(Y-85)=-16.3,$	$\Sigma(Y-85)^2=8,625.69.$
$\Sigma(X_1-160)(X_2-50)=2,575.1,$	
$\Sigma(X_1-160)(Y-85)=3,897.9,$	
$\Sigma(X_2-50)(Y-85)=3,467.91.$	
$N=77.$	

數、離均差平方和及積和，與相關係數如下(表 11-3). 此表 r_{12} 爲身

【表 11-3】 身高、體重、心像面積之均數及相關係數等

$\bar{x}_1=163.30,$	$\bar{x}_2=51.04,$	$\bar{y}=84.79.$
$\Sigma(X_1-\bar{x}_1)^2=4,724.13,$	$\Sigma(X_1-\bar{x}_1)(X_2-\bar{x}_2)=2,311.86,$	
$\Sigma(X_2-\bar{x}_2)^2=4,290.94,$	$\Sigma(X_1-\bar{x}_1)(Y-\bar{y})=3,951.67,$	
$\Sigma(Y-\bar{y})^2=8,622.24,$	$\Sigma(X_2-\bar{x}_2)(Y-\bar{y})=3,484.80.$	
$r_{12}=.5135,$	$r_{Y1}=.6192,$	$r_{Y2}=.5729.$

高與體重之相關係數，r_{Y1} 爲心像面積與身高之相關係數，r_{Y2} 爲心像面積與體重之相關係數. 將各相關係數之值代入下列公式，得兩標準迴歸係數，其中 $\beta_{Y1\cdot 2}$ 表示 Y 在 X_1 上之迴歸而與 X_2 無關者 (standard regression of Y on X_1 independent of X_2)，即指當 X_2 固

定不變時，由 X_1 推算 Y 的迴歸係數，又因以標準差為單位，故前面加'標準'兩字.

$$\beta_{Y1\cdot2}=\frac{r_{Y1}-r_{Y2}r_{12}}{1-r_{12}^2},\qquad \beta_{Y2\cdot1}=\frac{r_{Y2}-r_{Y1}r_{12}}{1-r_{12}^2}.\qquad \text{公式(11–1)}$$

$$\beta_{Y1\cdot2}=\frac{.6192-(.5729)(.5135)}{1-(.5135)^2}=.4414,$$

$$\beta_{Y2\cdot1}=\frac{.5729-(.6192)(.5135)}{1-(.5135)^2}=.3463.$$

將此迴歸係數化為原來單位，則得兩 b 值:

$$b_{Y1\cdot2}=\beta_{Y1\cdot2}\frac{\sqrt{\Sigma(Y-\bar{y})^2}}{\sqrt{\Sigma(X_1-\bar{x}_1)^2}},\quad b_{Y2\cdot1}=\beta_{Y2\cdot1}\frac{\sqrt{\Sigma(Y-\bar{y})^2}}{\sqrt{\Sigma(X_2-\bar{x}_2)^2}}.$$

公式(11–2)

$$b_{Y1\cdot2}=.4414\frac{\sqrt{8,622.24}}{\sqrt{4,724.13}}=.5963,$$

$$b_{Y2\cdot1}=.3463\frac{\sqrt{8,622.24}}{\sqrt{4,290.94}}=.4909.$$

此為部分迴歸係數(partial regression coefficients)——因為 $b_{Y1\cdot2}$ 是由 X_1 推算 Y 的迴歸係數，而與 X_2 無關，故稱'部分'. 最後將兩 b 值及表 11–3 的三個均數代入迴歸方程式:

$$Y_e=\bar{y}+b_{Y1\cdot2}(X_1-\bar{x}_1)+b_{Y2\cdot1}(X_2-\bar{x}_2),\qquad \text{公式(11–3)}$$

得 $$Y_e=84.79+.5963(X_1-163.30)+.4909(X_2-51.04)$$

$$\left.\begin{aligned}&=.5963X_1+.4909X_2-37.65,\\ &\text{或}\quad \text{心像面積(平方厘米)}=.5963\ \text{身高(厘米)}\\ &\qquad\qquad +.4909\ \text{體重(仟克)}-37.65.\end{aligned}\right\}\text{(K)}$$

此卽爲由身高、體重推算心像面積之多元迴歸方程式. 若仿第五章第5–5節的方法，將此迴歸方程式編成對照表，則在實際應用時當更爲便利.

前面兩 β 的公式是從下列兩方程式解出來的結果:

$$\beta_{Y1\cdot2}+r_{12}\beta_{Y2\cdot1}=r_{Y1},$$

$$r_{12}\beta_{Y1\cdot2}+\beta_{Y2\cdot1}=r_{Y2}.$$

這兩式叫做標準方程式 (normal equations)，是應用最小二乘方的原理求得的，其目的在使由公式 (11–3) 所得的估計值 Y_e 與觀察值 Y 相差之平方和爲最小，因此較由他法（如鄒、陳二氏的建議等）所得的估計值更爲正確.

11–3. 多元相關（multiple correlation，或譯作複相關） 將表11–1各被試者的身高、體重數值代入方程式(K)，則得各人的心像面積估計值. 如

X_1	X_2	Y	Y_e
154	42.0	79.0	74.80
161	49.5	87.5	82.65
……	……	……	……

若用第六章的方法，求觀察值 Y 與估計值 Y_e 間的相關，此值稱爲多元相關係數(multiple correlation coefficient)，通用的符號是 R. 這是第六章簡單相關係數的推廣，讀者試將表5–2的資料，求體重觀察值與估計值間的相關係數，結果應等於食量 X 與體重 Y 間的相關係數，以符號表示之，卽 $r_{YY_e}=r_{XY}$. 所以多元相關 R 與簡單相關 r 的意義是類似的，只是自變數的多寡不同罷了.

至於實際計算R時，我們另有簡捷的公式，正像求r時無須先算估計值一樣．R的公式是：

$$R^2_{Y\cdot 12}=r_{Y1}\beta_{Y1\cdot 2}+r_{Y2}\beta_{Y2\cdot 1} \qquad \text{公式(11–4)}$$

$$=.6192(.4414)+.5729(.3463)=.47171,$$

$$R_{Y\cdot 12}=.687.$$

符號$R_{Y\cdot 12}$表示由X_1及X_2推算Y的多元迴歸所得估計值與觀察值Y間的多元相關係數．若將公式(11–1)兩β代入公式(11–4)，則R的公式更可化簡爲

$$R^2_{Y\cdot 12}=\frac{r^2_{Y1}+r^2_{Y2}-2r_{Y1}r_{Y2}r_{12}}{1-r^2_{12}}. \qquad \text{公式(11–5)}$$

以表 11–3 所列三個相關係數代入公式(11–5)，開方後與前面的結果相同．當變數多於三個時，計算R還有普徧的公式，此點留待後面討論．

R既是觀察值與估計值間的相關係數，而觀察值與估計值通常不會呈相反的關係，所以R最小等於 0，卻不會有負值．這點是與簡單相關係數不同的地方．〔因變數Y與自變數$X_1, X_2, \cdots\cdots$間的關係有正的也有負的，如各迴歸係數所表示者，惟其如此，所以R往往不附符號(4)．〕又通常估計值與觀察值不會完全一樣，故R每小於 1（當然也可以等於 1，祇是很少見吧了）．又R每較r_{Y1}與r_{Y2}爲大，如上例$R=.687$，此值大於 .6192 或 .5729．我們可利用這些特性來核對計算上有無錯誤．

11–4. 估計之誤差 觀察值與估計值之相差稱爲估計之誤差．如上節，已求得第一、二兩位被試者的估計值，第一位的估計誤差是

觀察值	估計值	估計之誤差	估計誤差之平方
Y	Y_e	$Y-Y_e$	$(Y-Y_e)^2$
79.0	74.80	4.20	17.6400
87.5	82.65	4.85	23.5225
……	……	……	………

4.20，第二位的估計誤差是4.85，兩人相差0.65平方厘米，若以觀察值計，則兩人的相差有$87.5-79.0=8.5$平方厘米之多，不過這是不公允的，因爲第一位的身高、體重都比第二位的來得小(參閱上節或表11-1)，其心像面積自然也要小些，而觀察值的相差卻沒有把這點考慮在內．至於估計值，可以說是具有該身高、體重值者之平均心像面積，故第一位之心像面積較平均多4.20，第二位則較平均多4.85．若將身高、體重考慮在內，則此兩人心像面積的相差不能算大．在若干問題中，用估計誤差來比較，要更合理些．

估計誤差有正有負，其代數和爲零．又將各估計誤差平方後相加，得總和爲4,555.04，此值較由其他任何一次方程式(含有三個變數之一次方程式在幾何學上爲一平面)所得者爲小．這兩點特性與直線迴歸是相同的(參閱第五章第5-2節)．計算估計誤差平方和時，實際上不必這樣麻煩，我們另有簡單的公式．其原理與在簡單相關係數中一樣的(參閱第5-2節及第6-8節)，便是將$\Sigma(Y-\bar{y})^2$析爲兩部分：一部分爲$R^2\Sigma(Y-\bar{y})^2$，是由迴歸而來的變異，另一部分爲$(1-R^2)\Sigma(Y-\bar{y})^2$，則與迴歸無關．今計算如下：

$$R^2\Sigma(Y-\bar{y})^2=.47171(8,622.24)=4,067.20,$$

$$(1-R^2)\Sigma(Y-\bar{y})^2=.52829(8,622.24)=4,555.04.$$

後者卽爲估計誤差之平方和. 我們在計算迴歸方程式(K)時,用過三個統計數, 卽 $\bar{y}, b_{Y1\cdot2}$ 及 $b_{Y2\cdot1}$, 故此處自由度爲 $N-3$, 至於 $\bar{x}_1$ 與 $\bar{x}_2$ 已包含在計算過程中,不另占去自由度. 此外還有一種算自由度的簡法,卽將總次數減去變數的個數(包括自變數與因變數). 在方程式(K)中共有三個變數,故自由度爲 $N-3$. 此原則在直線迴歸及其他多元迴歸都可適用.

將估計誤差之平方和以自由度除之,開方卽得標準估計誤差:

$$S_{Y\cdot12}=\sqrt{\frac{\Sigma(Y-Y_e)^2}{N-3}}=\sqrt{\frac{(1-R^2)\Sigma(Y-\bar{y})^2}{N-3}}, \quad \text{公式(11–6)}$$

將各數值代入,得

$$S_{Y\cdot12}=\sqrt{\frac{(1-.47171)(8,622.24)}{77-3}}=\sqrt{\frac{4,555.04}{77-3}}=7.846.$$

又心像面積之標準差爲

$$S_Y=\sqrt{\frac{8,622.24}{77-1}}=10.651.$$

經引用了身高與體重兩變數後, 心像面積的變異從 10.651 減到了 7.846. 可知由身高、體重的不同而使心像面積產生的變異約占標準差的四分之一(參閱第五章第 5–2 節).

11–5. 顯著性測驗 多元迴歸中,所有統計數都受抽樣變異的影響,故須測驗其是否顯著. 含有三個變數的多元迴歸,其兩個 β 值的標準誤是一樣的,公式如下:

$$S_\beta=\sqrt{\frac{1-R^2}{(1-r_{12}^2)(N-3)}}, \quad \text{公式(11–7)}$$

將諸值代入，得

$$S_{\beta}=\sqrt{\frac{1-.47171}{[1-(.5135)^2](77-3)}}=.098.$$

測驗迴歸係數的顯著性，用 t 值，上例 $\beta_{Y1\cdot2}=.4414$，$\beta_{Y2\cdot1}=.3463$，其 t 值如下：

$\beta_{Y1\cdot2}$ $\qquad t=.4414/.098=4.48,$

$\beta_{Y2\cdot1}$ $\qquad t=.3463/.098=3.52.$

查 t 值表（表 4–4），當自由度爲 $77-3=74$ 時，其 1% 點約爲 2.65，此處兩 t 值均大於 1% 點，故知兩 β 值與零之相差爲非常顯著. 至於部分迴歸係數 b 值，只是 β 的倍數，其顯著性與 β 同.

若求 $\beta_{Y1\cdot2}$ 與 $\beta_{Y2\cdot1}$ 兩值相差之顯著性，則其標準誤爲

$$S_{\beta_1-\beta_2}=\sqrt{\frac{2(1-R^2)}{(N-3)(1-r_{12})}}, \qquad \text{公式(11–8)}$$

將上例之各數值代入，得

$$S_{\beta_1-\beta_2}=\sqrt{\frac{2(1-.47171)}{(77-3)(1-.5135)}}=.1713.$$

故 $t=(.4414-.3463)/.1713=.555$. 查 t 值表，當自由度爲 74 時，5% 點約爲 1.994，今求得之 t 值小於 5% 點，故兩 β 值之相差爲不顯著. 易言之，在推算心像面積之迴歸方程式中，身高體重所占分量並無輕重之分. 在此仍須提醒讀者，勿僅憑迴歸係數之表面值，遽下關係大小之結論. 公式 (11–8) 是用於三元迴歸的. 在四元迴歸中兩 β 值相差之顯著性，見第 11–9 節.

關於多元相關係數 R 的顯著性，則需用 F 來測驗，測驗時係將

'由迴歸而來的變異'與'不能由迴歸解釋的變異'相比較，看前者是否比後者大得多. 前面我們已算出這兩部分的離均差平方和，茲列變異數分析表於後:

［表 11-4］ R 之顯著性測驗（多元迴歸之顯著性同）

變異來源	自由度	平方和	均方	F
迴歸, $R^2\Sigma(Y-\bar{y})^2$	2	4,067.20	2,033.60	33.04
估計之誤差, $(1-R^2)\Sigma(Y-\bar{y})^2$	74	4,555.04	61.55	
總計	76	8,622.24		

在迴歸方程式(K)中有 2 個自變數，故自由度為 2. 估計誤差之自由度已述於前. 查 F 值表（表 9-3），當 $n_1=2$, $n_2=74$ 時，1% 點約為 4.92，求得之 F 值遠過於 1%點，故 R 為非常顯著.

測驗 R 的顯著性，還可查表 11-5$_{(5)}$. 該表中 2 個變數的一行數值是與表 6-3 相同的，其他各行是新添的. 上例共有 3 個變數，查該行自由度等於 70（與 74 最近）時，R 值的 5% 點（在上）為 .286，又 1% 點（在下）為 .351，今求得的 R 值為 .687，大於 1% 點，故為非常顯著，此法較變異數分析要省事多了.

11-6. 部分相關（partial correlation） 在多元迴歸中，與部分迴歸係數相當的有部分相關係數，他們的關係與第六章所說 r 和 b 的關係是一樣的. 設有 X_1, X_2 與 X_3 三個變數，我們先求當 X_3 固定時，由 X_2 推算 X_1 的部分迴歸係數（即 $b_{12\cdot3}$）另求當 X_3 固定時，由 X_1 推算 X_2 的部分迴歸係數（即 $b_{21\cdot3}$），此兩部分迴歸係數的幾何均數，即為部分相關係數（參閱第 6-7 節）. 其公式為

$$r_{12\cdot3}=\sqrt{b_{12\cdot3}\,b_{21\cdot3}}. \qquad \text{公式(11-9)}$$

【表 11-5】 r 與 R 之 5% 點與 1% 點

自由度	變數之個數 2	3	4	5	自由度	變數之個數 2	3	4	5
1	.997	.999	.999	.999	24	.338	.470	.523	.562
	1.000	1.000	1.000	1.000		.496	.565	.609	.642
2	.950	.975	.983	.987	25	.381	.462	.514	.553
	.990	.995	.997	.998		.487	.555	.600	.633
3	.878	.930	.950	.961	26	.374	.454	.506	.545
	.959	.976	.983	.987		.478	.546	.590	.624
4	.811	.881	.912	.930	27	.367	.446	.498	.536
	.917	.949	.962	.970		.470	.538	.582	.615
5	.754	.836	.874	.898	28	.361	.439	.490	.529
	.874	.917	.937	.949		.463	.530	.573	.606
6	.707	.795	.839	.867	29	.355	.432	.482	.521
	.834	.886	.911	.927		.456	.522	.565	.598
7	.666	.758	.807	.838	30	.349	.426	.476	.514
	.798	.855	.885	.904		.449	.514	.558	.591
8	.632	.726	.777	.811	35	.325	.397	.445	.482
	.765	.827	.860	.882		.418	.481	.523	.556
9	.602	.697	.750	.786	40	.304	.373	.419	.455
	.735	.800	.836	.861		.393	.454	.494	.526
10	.576	.671	.726	.763	45	.288	.353	.397	.432
	.708	.776	.814	.840		.372	.430	.470	.501
11	.553	.648	.703	.741	50	.273	.336	.379	.412
	.684	.753	.793	.821		.354	.410	.449	.479
12	.532	.627	.683	.722	60	.250	.308	.348	.380
	.661	.732	.773	.802		.325	.377	.414	.442
13	.514	.608	.664	.703	70	.232	.286	.324	.354
	.641	.712	.755	.785		.302	.351	.386	.413
14	.497	.590	.646	.686	80	.217	.269	.304	.332
	.623	.694	.737	.768		.283	.330	.362	.389
15	.482	.574	.630	.670	90	.205	.254	.288	.315
	.606	.677	.721	.752		.267	.312	.343	.368
16	.468	.559	.615	.655	100	.195	.241	.274	.300
	.590	.662	.706	.738		.254	.297	.327	.351
17	.456	.545	.601	.641	125	.174	.216	.246	.269
	.575	.647	.691	.724		.228	.266	.294	.316
18	.444	.532	.587	.628	150	.159	.198	.225	.247
	.561	.633	.678	.710		.208	.244	.270	.290
19	.433	.520	.575	.615	200	.138	.172	.196	.215
	.549	.620	.665	.698		.181	.212	.234	.253
20	.423	.509	.563	.604	300	.113	.141	.160	.176
	.537	.608	.652	.685		.148	.174	.192	.208
21	.413	.498	.552	.592	400	.098	.122	.139	.153
	.526	.596	.641	.674		.128	.151	.167	.180
22	.404	.488	.542	.582	500	.088	.109	.124	.137
	.515	.585	.630	.663		.115	.135	.150	.162
23	.396	.479	.532	.572	1,000	.062	.077	.088	.097
	.505	.574	.619	.652		.081	.096	.106	.115

這公式表示部分相關係數也有雙方面的關係，便是當 X_3 固定時，從 X_2 估計 X_1，或從 X_1 估計 X_2，其間無所選擇，這時便用部分相關係數表示他們的關係（若 X_1 隨着 X_2 與 X_3 而變化，那就得用 $b_{12\cdot3}$ 與 $b_{13\cdot2}$ 了）．前述心像面積的資料，我們已求得 $b_{Y1\cdot2}=.5963$（見公式 11–2 之實例），同理求得 $b_{1Y\cdot2}=.3581$．這裏 .5963 相當於公式(11–9) 裏的 $b_{12\cdot3}$，而 .3581 相當於該公式的 $b_{21\cdot3}$（讀者須注意公式的活用，同時對於符號也要認識清楚），代入上式，得

$$r_{Y1\cdot2}=\sqrt{(.5963)(.3581)}=.462.$$

這是假定有很多體重相等的人，他們的心像面積與身高的相關係數．

將公式(11–9)化爲簡單相關係數的函數，則得

$$r_{12\cdot3}=\frac{r_{12}-r_{13}r_{23}}{\sqrt{(1-r_{13}^2)(1-r_{23}^2)}},\qquad\text{公式(11–10)}$$

如前例　$$r_{Y1\cdot2}=\frac{.6192-(.5729)(.5135)}{\sqrt{[1-(.5729)^2][1-(.5135)^2]}}=.462.$$

測驗部分相關係數的顯著性可用 t 值，其公式爲

$$t=\frac{r}{\sqrt{1-r^2}}\sqrt{N-p-2},\quad \text{自由度}=N-p-2.\qquad\text{公式(11–11)}$$

此處 r 爲部分相關係數，p 爲被固定的變數的個數．在 r_{Y1-2} 中被固定的變數是一個（X_2），故自由度爲 $77-1-2=74$．還有一種算法，是在總次數內減去變數的個數（包括因變數與自變數），其結果相同．

$$t=\frac{.462}{\sqrt{1-(.462)^2}}\sqrt{77-1-2}=5.05,$$

$$\text{自由度}=74,\qquad 1\%\text{點}=2.65.$$

另法可查表 11-5，含有 3 個變數的直行內，自由度等於 70 (表中與 74 最近的一數)時，r 的 1% 點爲 .351，今求得之部分相關係數爲 .462，大於 1%點，故知其爲非常顯著.

11-7. 含有四個以上變數之資料 若要從三個變數推算另外一個變數，那麼每一個被試者必須都受這四種測量，根據這樣的資料纔能求得含有四個變數的迴歸方程式. 過去有些人的研究，因爲幾種測量不在同時舉行，於是參加各種測量的被試者不盡相同，數目也不一致. 從這種參差不齊的紀錄，雖可求得若干簡單相關係數，由此計算多元迴歸方程式，但參加各種測量的被試錯雜，人數不等，終究是個缺點，尤其在顯著性測驗時無從確定其自由度. 這是在搜集材料時必先注意的.

含有四個以上變數的迴歸方程式，其原理雖和以前一樣，但計算步驟卻繁得多了. 初學者不要以爲所用變數愈多，結果就愈爲正確. 相反的，測量的項目增多，則所費時間精力也增加，所包括的問題也更爲複雜. 如資料的殘缺不全、測量技術上的失敗、均勻材料的不易得到、解釋結果時的困難等等，在變數增多時這些情形都是難免的. 多元迴歸並不是一劑萬應靈藥，它並不能使不正確的資料變成可靠的結果. 但若選材謹慎，測量精確，且有一定的問題須用多元迴歸來解決時，當然也不必猶豫.

若遇小樣本而含有幾個變數的資料，分析時須十分小心. 其理由可從幾何學方面來說明. 在平面裏，任何一對(X, Y)值表示一點，故兩對(X, Y)的觀察值定一迴歸線，其餘各對觀察值則用以改進此迴歸線，並可藉此求得估計之誤差. 在三度空間裏，則須有三組 $(X,$

Y, Z)的觀察值纔能定一迴歸平面(regression plane). 依次類推,則有四個變數時,就要四個觀察所得的點纔能決定一迴歸. 同理,由六個變數構成的迴歸就有六個點子完全適合,而無估計的誤差(這時 $R=1$). 因此若僅有 10 組觀察值,而求一含有 6 個變數的迴歸,結果得 $R=0.9$,這時不要太高興了,因爲這 R 值實在是不顯著的. 這裏要告訴讀者,在小樣本中對於一切顯著性測驗務須特別謹慎,尤其在總次數 N 與變數的個數相差很近時,不要被大的 R 值所欺騙了.

11-8. 四元迴歸之計算法 在研究肺活量(vital capacity)時,我們也需要知道健康人的肺活量該有多少. 普通根據身高體重來估計肺活量作爲具有該身高體重者肺活量之平均值,如郭祖超與吳襄二氏所發表的推算國人肺活量之多元迴歸方程式(6)是. 又因肺活量與胸圍亦有密切關係,所以我們也可以根據身高、體重與胸圍三者來推算肺活量,這便是四元迴歸方程式了.

表11-6所列統計數,是從174個我國青年(15—16歲)男子的資

【表 11-6】 計算四元迴歸方程式所需之統計數

	均數	離均差平方和	相關係數
體 重(仟 克)X_1	$\bar{x}_1=47.20$	$\Sigma(X_1-\bar{x}_1)^2=1,295.34$	$r_{12}=.6451$
			$r_{13}=.8274$
身 高(厘 米)X_2	$\bar{x}_2=162.25$	$\Sigma(X_2-\bar{x}_2)^2=1,520.72$	$r_{23}=.4713$
胸 圍(厘 米)X_3	$\bar{x}_3=74.95$	$\Sigma(X_3-\bar{x}_3)^2=535.91$	$r_{Y1}=.5660$
			$r_{Y2}=.5524$
肺活量$\left(\begin{array}{c}100\text{ 立}\\ \text{方厘米}\end{array}\right)$ Y	$\bar{y}=31.06$	$\Sigma(Y-\bar{y})^2=912.86$	$r_{Y3}=.6650$

料算出來的(7),每一被試者都有體重、身高、胸圍與肺活量的實測紀

錄,紀錄的式樣與表 11–1 相似.茲因篇幅關係,從略.

四元迴歸方程式中所求的迴歸係數,可從下列標準方程式中解出來:

$$
\begin{array}{c|lll}
 & X_1 \qquad X_2 \qquad X_3 \qquad Y & & \\
\hline
X_1 & \beta_{Y1\cdot23}+r_{12}\,\beta_{Y2\cdot13}+r_{13}\,\beta_{Y3\cdot12}=r_{Y1} & & \\
X_2 & r_{12}\beta_{Y1\cdot23}+\beta_{Y2\cdot13}+r_{23}\,\beta_{Y3\cdot12}=r_{Y2} & & \\
X_3 & r_{13}\beta_{Y1\cdot23}+r_{23}\,\beta_{Y2\cdot13}+\beta_{Y3\cdot12}=r_{Y3} & &
\end{array}
\Bigg\}\text{公式(11–12)}
$$

此處 $\beta_{Y1\cdot23}$ 是'由 X_1 推算 Y 之標準迴歸係數而與 X_2 及 X_3 皆無關者',餘類推.我們可以看到三個方程式沿左上至右下方對角線的對稱情形.方程式上端與左側的變數是幫助我們寫出 r 右下角的數目的.

將表 11–6 所列相關係數代入公式(11–12),則得三個一次方程式:

$$\beta_{Y1\cdot23}+.6451\,\beta_{Y2\cdot13}+.8274\,\beta_{Y3\cdot12}=.5660, \qquad \text{(a)}$$

$$.6451\beta_{Y1\cdot23}+\beta_{Y2\cdot13}+.4713\,\beta_{Y3\cdot12}=.5524, \qquad \text{(b)}$$

$$.8274\beta_{Y1\cdot23}+.4713\beta_{Y2\cdot13}+\beta_{Y3\cdot12}=.6650. \qquad \text{(c)}$$

以下便是用普通代數方法,把三個 β 值解出來.

(a) ÷ .8274, $\quad 1.2086\beta_{Y1\cdot23}+.7797\beta_{Y2\cdot13}+\beta_{Y3\cdot12}=.6841,$ (d)

(b) ÷ .4713, $\quad 1.3688\beta_{Y1\cdot23}+2.1218\beta_{Y2\cdot13}+\beta_{Y3\cdot12}=1.1721,$ (e)

(d) − (c), $\quad .3812\beta_{Y1\cdot23}+.3084\beta_{Y2\cdot13}=.0191,$ (f)

(e) − (d), $\quad .1602\beta_{Y1\cdot23}+1.3421\beta_{Y2\cdot13}=.4880,$ (g)

(f) ÷ .3084, $\quad 1.2361\beta_{Y1\cdot23}+\beta_{Y2\cdot13}=.0619,$ (h)

(g) ÷ 1.3421, $\quad .1194\beta_{Y1\cdot23}+\beta_{Y2\cdot13}=.3636,$ (i) (h)－(i)， $1.1167\beta_{Y1\cdot23}=-.3017,$ (j)

$$\beta_{Y1\cdot23}=-.2702,$$

代入(i)， $\beta_{Y2\cdot13}=.3636-.1194(-.2702)=.3959,$

代入(d)， $\beta_{Y3\cdot12}=.6841-1.2086(-.2702)-.7797(.3959)$

$$=.7020.$$

各 β 值是以標準差爲單位的，今以相當的離均差平方和開方後乘除之，則化成原來單位，

$$b_{Y1\cdot23}=-.2702\frac{\sqrt{912.86}}{\sqrt{1,295.34}}=-0.2268,$$

$$b_{Y2\cdot13}=.3959\frac{\sqrt{912.86}}{\sqrt{1,520.72}}=0.3067,$$

$$b_{Y3\cdot12}=.7020\frac{\sqrt{912.86}}{\sqrt{535.91}}=0.9162.$$

將各迴歸係數代入下列公式，即得所需迴歸方程式：

$$Y=\bar{y}+b_{Y1\cdot23}(X_1-\bar{x}_1)+b_{Y2\cdot13}(X_2-\bar{x}_2)+b_{Y3\cdot12}(X_3-\bar{x}_3),$$

公式(11–13)

$$Y=31.06-0.2268(X_1-47.20)+0.3067(X_2-162.25)$$
$$+0.9162(X_3-74.95)$$
$$=-0.2268X_1+0.3067X_2+0.9162X_3-76.67.$$

在計算過程中，肺活量以百立方厘米爲單位，今將上式各項以100乘之，並以文字表示，則爲

肺活量＝－22.68體重＋30.67身高＋91.62胸圍－7,667. (v)

（立方厘米） （仟克） （厘米） （厘米）

這裏可以注意的，體重前面的部分迴歸係數是負號，這是指當身高胸圍固定不變時（或有一羣身高相等、胸圍相等的被試者時），體重每增加一仟克，則肺活量卽平均減少 22.68 立方厘米. 在 P. H. Stevenson 氏發表的一套迴歸方程式(8)裏；由身高、胸圍、體重推算肺活量的方程式，其體重前面的係數也是負的. 通常體重、身高、胸圍三者總是連帶的變化，幾乎沒有兩個固定而一個單獨變化的情形，因此很少人注意到身高胸圍不變時，肺活量與體重間呈反比例的關係，這現象留待生理學家來解釋吧.

若將肺活量與體重、身高、胸圍之簡單相關係數（表 11–6）和各 β 值作一比較，就可發現他們所表示的關係不同. 從 β 方面看，由體

肺活量	體重	身高	胸圍
Y	X_1	X_2	X_3
Y 與各 X 之簡單相關係數(r)	.5660	.5524	.6650
Y 在 X 上之標準迴歸係數(β)	$-.2702$	.3959	.7020

重、身高、胸圍三者推算肺活量，則胸圍所占的分量最重，身高次之，而體重與肺活量卻呈相反的關係. 至於簡單相關係數，則因有其他兩因子夾雜在內（如計算肺活量與體重之簡單相關時，並未使身高與胸圍固定不變），故不能表明眞正的關係，其數值的大小，更不能表示估計肺活量時所占分量的輕重. 這點在解釋各變數間的關係時應當注意的.

11–9. 四元迴歸之顯著性測量 四元迴歸係數的顯著性，也可以分做兩方面來測驗：一是 R 的顯著性，一是三個 β 的顯著性。R 的計算法和公式(11–4)相似：

$$R^2_{Y\cdot123}=r_{Y1}\beta_{Y1\cdot23}+r_{Y2}\beta_{Y2\cdot13}+r_{Y3}\beta_{Y3\cdot12}.\qquad \text{公式}(11\text{–}14)$$

這裏的 $R_{Y\cdot123}$ 是 Y 的觀察值與由公式(11–13)所得估計值間的相關係數，將 r 及 β 值代入，得

$$R^2_{Y\cdot123}=.5660(-.2702)+.5524(.3959)+.6650(.7020)$$
$$=.53260,$$
$$R_{Y\cdot123}=.7298.$$

測驗 R 的顯著性時，仍將總變異析為兩部分：一為 $R^2\Sigma(Y-\bar{y})^2$ 是由迴歸而來的變異，一為 $(1-R^2)\Sigma(Y-\bar{y})^2$ 是估計誤差的平方和，茲將變異數分析列後：

【表 11–7】 R 之顯著性測驗

變 異 來 源	自由度	離均差平方和	均 方	F	1%點
總變異(表 11-6)	173	912.86			
迴歸 $R^2\Sigma(Y-\bar{y})^2$	3	486.19	162.06	64.57	3.91
誤差 $(1-R^2)\Sigma(Y-\bar{y})^2$	170	429.67	2.51		

自由度的計算仍適用第 11–4 及 11–5 節的原則. 查表 9–3，當 $n_1=3$，$n_2=150$（當 170 最近）時 1% 點為 3.91，此處求得之 F 值遠過於 1% 點，故 R 為非常顯著. 另法可查表 11–5 變數為 4 個的一行，當自由度等於 150 時，R 的 1% 點為 0.270，此處求得者為 .7298，故為非常顯著.

由上節求得的三個 β 值可知估計肺活量時，胸圍所占的分量最重，身高次之，而體重則呈相反的關係. 那麼我們可否只用身高，胸圍來估計肺活量. 換句話說，在迴歸方程式裏，我們多用一個自變數(體重)是否較有進步呢？這點就要看四元迴歸比三元迴歸所減少的

誤差是否顯著. 若以表 11–6 所列相關係數, 按公式(11–1)求得由身高、胸圍推算肺活量的兩 β 值:

$$\beta_{Y2\cdot3}=\frac{.5524-(.6650)(.4713)}{1-(.4713)^2}=.30723,$$

$$\beta_{Y3\cdot2}=\frac{.6650-(.5524)(.4713)}{1-(.4713)^2}=.52020.$$

再由公式(11–4), 求得

$$R^2_{Y\cdot23}=(.5524)(.30723)+(.6650)(.5202)=.51564.$$

故由身高、胸圍以推算肺活量, 則估計誤差之平方和爲

$$(1-.51564)(912.86)=442.15.$$

用四元迴歸時, 其估計誤差之平方和(表 11–7)比三元迴歸減少

$$(442.15-426.67)=15.48.$$

這部分相當於 1 個自由度(因四元迴歸有 3 個自由度, 三元迴歸有 2 個, 相差 1 個, 參閱表 11–7 及 11–8), 是屬於四元迴歸的, 茲測驗其顯著性如下:

【表 11–8】 四元迴歸與三元迴歸之比較

變 異 來 源	自由度	離均差平方和	均 方	F	5%點	1%點
總變異 $\Sigma(Y-\bar{y})^2$	173	912.86				
三元迴歸 $R^2_{Y\cdot23}\Sigma(Y-\bar{y})^2$	2	470.71	235.36	90.873	——	4.75
誤差 $(1-R^2_{Y\cdot23})\Sigma(Y-\bar{y})^2$	171	442.15	2.59			
四元迴歸與三元迴歸之相差 $(R^2_{Y\cdot123}-R^2_{Y\cdot23})\Sigma(Y-\bar{y})^2$	1	15.48	15.48	6.167	3.91	6.81
誤差 $(1-R^2_{Y\cdot123})\Sigma(Y-\bar{y})^2$	170	426.67	2.51			

由三元迴歸所減少的變異爲非常顯著(F 值大於 1% 點), 四元迴歸

較三元迴歸更減少 15.48，這減少的部分仍為顯著(6.167 在 5%點與 1%點之間)，所以由體重、身高、胸圍來推算肺活量，比僅用身高胸圍來推算的要進步些.

現在要測驗三個 β 的顯著性了. 測驗的步驟也是先求各 β 值的標準誤，再計算 t 值. 可是在四元迴歸，各 β 的標準誤是不相等的，因此計算過程就複雜多了. 這裏先要解三組標準方程式，左邊各項的係數與上節求 β 時方程式 (a)，(b)，(c) 內的是相同的，祇是 β 換成了 k_1，k_2 與 k_3(注意這三個 k 與第七章常態性測驗中所用的無關，勿混淆)；而等號右邊的並不是相關係數，卻是三組數目：(1, 0, 0)，

【表 11-9】 k 值之解法

標準方程式之左側各項			方程式及解答之組號 1	2	3	方程式號碼	運算
k_1	$.6451k_2$	$.8274k_3$	1	0	0	(1)	
$.6451k_1$	k_2	$.4713k_3$	0	1	0	(2)	
$.8274k_1$	$.4713k_2$	k_3	0	0	1	(3)	
$1.2086k_1$	$.7797k_2$	k_3	1.2086	0	0	(4)	(1)÷.8274
$1.3688k_1$	$2.1218k_2$	k_3	0	2.1218	0	(5)	(2)÷.4713
$.3812k_1$	$.3084k_2$		1.2086	0	−1	(6)	(4)−(3)
$.1602k_1$	$1.3421k_2$		−1.2086	2.1218	0	(7)	(5)−(4)
$1.2361k_1$	k_2		3.9189	0	−3.2425	(8)	(6)÷ .3084
$.1194k_1$	k_2		−.9005	1.5810	0	(9)	(7)÷1.3421
$1.1167k_1$			4.8194	−1.5810	−3.2425	(10)	(8)−(9)
k_1			4.3158	−1.4158	−2.9036	(11)	
	k_2		−1.4158	1.7501	.3466	(12)	
		k_3	−2.9036	.3466	3.2390	(13)	

(0,1,0),(0,0,1),每組數目定一組方程式.解方程式的步驟與上節相同,爲簡單起見,這裏用表解法(此爲 Gauss 氏法,參考文獻 11),讀者如將表 11–9 與上節對照,可更爲清楚.該表最右邊一欄所註明的運算,是施於該横行內的每一分子的.如方程式(1)÷.8274時,左側各項的係數爲:

$$1/.8274=1.2086,$$
$$.6451/.8274=.7797,$$
$$.8274/.8274=1;$$

右側各項爲:

$$1/.8274=1.2086,$$
$$0/.8274=0,$$
$$0/.8274=0.$$

又如方程式(5)−(4)時,左側各項係數爲:

$$1.3688-1.2086=.1602,$$
$$2.1218-.7797=1.3421,$$
$$1-1=0;$$

右側各項爲:

$$0-1.2086=-1.2086$$
$$2.1218-0=2.1218,$$
$$0-0=0.$$

這樣繼續下去一直到方程式(10),所經過的步驟與上節由(a)至(j)完全一樣的.第(11)式的數值是將第(10)式各項除以 1.1167 的結果.其中 4.3158 是第 1 組方程式解得的 k_1, −1.4158 是第 2 組的 k_1, −2.9036是第 3 組的 k_1.於是將 k_1 代入前式,求得各組的 k_2 與 k_3,玆舉四例於後.

(甲)以第1組的 $k_1=4.3158$ 代入式(8),

$$1.2361(4.3158)+k_2=3.9189,\ \text{解之,得}\ k_2=-1.4158,$$

寫在第(12)式第1組解答行內.

(乙)將第2組的 $k_1=-1.4158$ 亦代入(8)式,

$$1.2361(-1.4158)+k_2=0,\quad k_2=1.7501,$$

此值寫在第(12)式第2組解答的一行內.

(丙)將第2組解答的 $k_1=-1.4158$, $k_2=1.7501$ 代入(3)式,

$$.8274(-1.4158)+.4713(1.7501)+k_3=0,\quad k_3=.3466,$$

此爲第2組解答中之 k_3.

(丁)同理,將第3組解答中之 k_1 與 k_2 代入(3)式,

$$.8274(-2.9036)+.4713(.3466)+k_3=1,\quad k_3=3.2390.$$

表 11–9 末所列的九個 k 值是沿着左上至右下方的對角線呈對稱的,不過爲核對起見,還是逐個計算一下比較妥當,至於末位小數微有出入是免不了的.

這樣我們求得了三組(k_1, k_2, k_3)值,每組可算出一 β 值,例如

$$\begin{aligned}\beta_{Y1\cdot 23}&=k_1r_{Y1}+k_2r_{Y2}+k_3r_{Y3}=(4.3158)(.5660)\\&\quad+(-1.4158)(.5524)+(-2.9036)(.6650)\\&=-.2703.\end{aligned}\qquad\text{公式(11–15)}$$

同理
$$\begin{aligned}\beta_{Y2\cdot 13}&=(-1.4158)(.5660)+(1.7501)(.5524)\\&\quad+(.3466)(.6650)=.3960,\\ \beta_{Y3\cdot 12}&=(-2.9036)(.5660)+(.3466)(.5524)\\&\quad+(3.2390)(.6650)=.7020.\end{aligned}$$

β 值的計算及其顯著性測驗,可用表 11–9 一併解法,但若無上節對

照，則該表演算原理，恐不容易明白了.

有了三組 k 值，就可計算 β 的標準誤. 讓我們先求一個 v 值:

$$v=\frac{1-R^2}{N-m}, \qquad \text{公式(11–16)}$$

式中 R 是本節開始時求得的多元相關係數，N 為總次數，m 為變數的個數. 將各值代入此式，得

$$v=\frac{1-.5326}{174-4}=.0027494.$$

再乘以適當的 k 值，得各 β 的變異數，

$$V_1=vk_{11}, \qquad V_2=vk_{22}, \qquad V_3=vk_{33}. \qquad \text{公式(11–17)}$$

所謂 k_{11} 是指第 1 組解答的 k_1 值，餘類推.

$$V_1=(.0027494)(4.3158)=.011866,$$
$$V_2=(.0027494)(1.7501)=.004812,$$
$$V_3=(.0027494)(3.2390)=.008905.$$

開方得標準誤

$$S_1=.1089, \qquad S_2=.0694, \qquad S_3=.0944.$$

因得

$$t_1=\beta_{Y1\cdot 23}/S_1=-.2702/.1089=-2.480,$$
$$t_2=\beta_{Y2\cdot 13}/S_2=.3959/.0694=5.707,$$
$$t_3=\beta_{Y3\cdot 12}/S_3=.7020/.0944=7.439.$$

查 t 值表(表 4–4)，當自由度為 $174-4=170$ 時，5%點約為 1.976，1%點約為 2.609. 此處 t_1 為顯著 (注意: 比較時用 t 之絕對值，不計符號)，t_2 與 t_3 均為非常顯著. 易言之，當身高、胸圍固定時，體重與肺活量間有顯著的相反關係; 而當其他兩項固定時，身高與肺活量，及

胸圍與肺活量間皆有非常顯著的正關係．這結論與表 11-7—11-8 所得的相同．

在四元迴歸中，測驗兩 β 相差之顯著性亦用 t 值，所需標準誤可利用上面求得的 v 與 k 值求得，如

$$\left.\begin{array}{l}(\beta_{Y1\cdot23}-\beta_{Y2\cdot13})\text{之標準誤爲}\sqrt{v(k_{11}+k_{22}-2k_{12})}\\(\beta_{Y2\cdot13}-\beta_{Y3\cdot12})\text{之標準誤爲}\sqrt{v(k_{22}+k_{33}-2k_{23})}\\(\beta_{Y1\cdot23}-\beta_{Y3\cdot12})\text{之標準誤爲}\sqrt{v(k_{11}+k_{33}-2k_{13})}\end{array}\right\}.\text{公式(11–18)}$$

上例

$$v(k_{11}+k_{22}-2k_{12})=.0027494[4.3158+1.7501-2(-1.4158)]$$

$$=.024463,\quad \sqrt{.024463}=.1564,$$

$$\beta_{Y2\cdot13}-\beta_{Y1\cdot23}=.3959-(-.2702)=.6661,$$

$$t=.6661/.1564=4.259.$$

又 $$v(k_{22}+k_{33}-2k_{23})=.0027494[1.7501+3.2390-2(.3466)]$$

$$=.011811,\quad \sqrt{.011811}=.10868.$$

$$\beta_{Y3\cdot12}-\beta_{Y2\cdot13}=.7020-.3959=.3061,$$

$$t=.3061/.1087=2.816.$$

查 t 值表，當自由度爲 170 時，1% 點約爲 2.609，此處求得之 t 值皆大於 1% 點，故 β 值之相差爲非常顯著．

公式(11–18)實爲公式 (11–8) 之推廣，讀者若將三元迴歸的兩個標準方程式(第 11–2 節末)仿表 11–9 解之，可得

$$k_{11}=k_{22}=1/(1-r_{12}^2),\quad k_{12}=k_{21}=-r_{12}/(1-r_{12}^2),$$

又 $v=(1-R^2)/(N-3)$，代入公式(11–18)，卽可化爲公式(11–8)．

11–10. 多元迴歸之一般解法　上面已講過三元與四元迴歸方

程式的求法. 若所含的變數更多時, 其標準方程式仍與公式(11-12)相似, 只是項數增多而已. 普通用行列式(determinant)來解這套標準方程式, 不過計算時不很便利, 所以這裏不預備介紹. 讀者要明白其原理, 可參考 H. L. Rietz 等著作(9). 下面所講的是 Doolittle 氏法(10), 當迴歸方程式中含有五個或以上的變數時, 這是最簡便的一種方法.

此法所用的表格如表 11-10, 爲了印刷上的便利計, 這裏只列到七個變數 (六個自變數, 一個因變數), 若變數更多時, 可依此類推, 因爲此表的計算方法是很有規則的, 只要把'運算'欄的說明仔細看過, 就會懂得其中的原則. 若算六元迴歸, 則將標有 F 的縱行取消, 五元迴歸則將 E 與 F 兩直行取消. 但不論有幾個變數, I 行總是保留的. 同時該表下段也隨着取消, 如六元迴歸自 25 行起刪去, 五元迴歸自 18 行起刪去. 又表末之 β 若用前面的寫法, 則 β_6 即爲 $\beta_{Y6\cdot12345}$, 玆爲簡單計, 只寫作 β_6. 若在六元迴歸中, 則只有五個自變數, 故 β 值也只有五個, 表末各式中凡含有 β_6 的各項皆取消, 餘可類推.

運算欄內所用的符號, 如 $b_2^{\cdot}$, 係指直行 b 與橫行 2 相交處方格內的數值, D_{11} 係直行 D 與橫行 11 相交處方格內的數值. 至於該表最右邊的 X 直行是爲核對用的. 凡需要乘、加或除的各橫行在 X 直行內都要算的. 但以一數乘某行總和時, 須注意起訖的地位, 凡不包括在內的幾個直行裏的數值必須扣去後再乘. 例如 14 橫行註明'以 D_6 乘 5 行各項目, 自 D 至 I', 我們先要在總和 X_5 內扣去了 b_5 與 C_5, 然後以 D_6 乘之, 即 $(X_5-b_5-C_5)D_6$, 此值應與 14 橫行內自 D_{14} 至 I_{14} 諸值之和相等. 其餘如 4, 8, 9, 13 等行均仿此. 這種核對是必要的.

【表 11-10】 七元迴歸之表解法

行號	運　　　算	A	b	C	D	E	F	I	X
1	填入 r 值	1	r_{12}	r_{13}	r_{14}	r_{15}	r_{16}	$-r_{Y1}$	自 A 至 I 之總和
2	以 -1 除 1 行								
3	填入 r 值		1	r_{23}	r_{24}	r_{25}	r_{26}	$-r_{Y2}$	自 b 至 I 之總和
4	以 b_2 乘 1 行各項目，自 b 至 I								
5	3,4 兩行相加								
6	以 $-b_5$ 除 5 行								
7	填入 r 值			1	r_{34}	r_{35}	r_{36}	$-r_{Y3}$	自 C 至 I 之總和
8	以 C_2 乘 1 行各項目，自 C 至 I								
9	以 C_6 乘 5 行各項目，自 C 至 I								
10	7,8,9 三行相加								
11	以 $-C_{10}$ 除 10 行								
12	填入 r 值				1	r_{45}	r_{46}	$-r_{Y4}$	自 D 至 I 之總和
13	以 D_2 乘 1 行各項目，自 D 至 I								
14	以 D_6 乘 5 行各項目，自 D 至 I								
15	以 D_{11} 乘 10 行各項目，自 D 至 I								
16	12,13,14,15 四行相加								
17	以 $-D_{16}$ 除 16 行								
18	填入 r 值					1	r_{56}	$-r_{Y5}$	自 E 至 I 之總和
19	以 E_2 乘 1 行各項目，自 E 至 I								
20	以 E_6 乘 5 行各項目，自 E 至 I								
21	以 E_{11} 乘 10 行各項目，自 E 至 I								
22	以 E_{17} 乘 16 行各項目，自 E 至 I								
23	18,19,20,21,22 五行相加								
24	以 $-E_{23}$ 除 23 行								
25	填入 r 值						1	$-r_{Y6}$	自 F 至 I 之總和
26	以 F_2 乘 1 行各項目，自 F 至 I								
27	以 F_6 乘 5 行各項目，自 F 至 I								
28	以 F_{11} 乘 10 行各項目，自 F 至 I								
29	以 F_{17} 乘 16 行各項目，自 F 至 I								
30	以 F_{24} 乘 23 行各項目，自 F 至 I								
31	25,26,27,28,29,30 六行相加								
32	以 $-F_{31}$ 除 31 行								

$\beta_6 = I_{32}$;

$\beta_5 = (\beta_6)F_{24} + I_{24}$;

$\beta_4 = (\beta_6)F_{17} + (\beta_5)E_{17} + I_{17}$;

$\beta_3 = (\beta_6)F_{11} + (\beta_5)E_{11} + (\beta_4)D_{11} + I_{11}$;

$\beta_2 = (\beta_6)F_6 + (\beta_5)E_6 + (\beta_4)D_6 + (\beta_3)C_6 + I_6$;

$\beta_1 = (\beta_6)F_2 + (\beta_5)E_2 + (\beta_4)D_2 + (\beta_3)C_2 + (\beta_2)b_2 + I_2$.

將全表算畢代入表末所附公式，即得各β值. 最後以β值、離均差平方和及各均數代入下式，即得所需之迴歸方程式:

$$Y=\bar{y}+\beta_1\frac{\sqrt{\Sigma(Y-\bar{y})^2}}{\sqrt{\Sigma(X_1-\bar{x}_1)^2}}(X_1-\bar{x}_1)$$

$$+\beta_2\frac{\sqrt{\Sigma(Y-\bar{y})^2}}{\sqrt{\Sigma(X_2-\bar{x}_2)^2}}(X_2-\bar{x}_2)+\cdots\cdots. \quad \text{公式(11-19)}$$

此時多元相關R的公式，可由(11-4)及(11-14)兩公式類推. 部分相關的公式亦與公式(11-9)相似，如

$$r_{Y1\cdot23456}=\sqrt{b_{Y1\cdot23456}\quad b_{1Y\cdot23456}}. \qquad \text{公式(11-20)}$$

這是指當$X_2, X_3, \cdots\cdots, X_6$五個變數固定時，$Y$與$X_1$間之部分相關，根號內的$b_{Y1\cdot23456}$是這五個變數固定時由$X_1$推算$Y$的部分迴歸係數，而$b_{1Y\cdot23456}$則爲由$Y$推算$X_1$的部分迴歸係數. 凡有$p$個變數固定者，即稱爲$p$級部分相關係數(partial correlation coefficient of the pth order). 故簡單相關係數r_{12}可稱爲零級相關係數，$r_{Y1\cdot2}$或$r_{Y2\cdot1}$稱爲一級部分相關係數，而$r_{Y1\cdot23456}$則爲五級部分相關係數. 各級部分相關係數的顯著性測驗都可用公式(11-11).

11-11. 多元共變數(multiple covariance) 在多元迴歸中，亦可應用共變數分析法. 例如以動物飼養實驗比較幾種飼料的優劣，若欲將各組動物的食量與原始體重都化爲相等，然後看各組最後體重有無顯著的差別，這時就要用多元共變數. 又如研究幾個民族的基底代謝，欲將年齡與身體表面積都化爲相等，然後比較各民族在基底狀況下每小時的發熱量，這也要用多元共變數了. 現在我們就利用表 11-1 的資料，將被試者按年齡大小分爲三組，當各被試者的

身高、體重化爲相等後，三組的心像面積有顯著的差別否？前面我們已經提過：在成人方面，年齡與心像面積之間並無顯著的相關，故三組心像面積的均數可預料到沒有眞正差別的，這裏只用來說明多元共變數的算法吧了.

我們把表 11-1 的資料，姑分爲 '19—24 歲，25—30 歲，31 歲及以上' 三組，將各組身高、體重、心像面積的總和列於表 11-11，爲求 '組間' 變異及共變異之用. 因爲前面用的是縮簡法，故此處仍沿用它(當然也可不用縮簡法). 表末 '總計' 一行裏的數值與表 11-2 所載的總和相同，這也是核對的一法. 現在可以計算各組均數間的離均差平方和及積和了(參閱公式 9-3 及 10-1)，玆舉二例於下:

【表 11-11】 各年齡組身高體重與心像面積之總和

年齡組	人數	身高 $\Sigma(X_1-160)$	體重 $\Sigma(X_2-50)$	心像面積 $\Sigma(Y-85)$
19—24 歲	31	140	21.6	20.0
25—30 歲	22	103	43.2	−4.0
31 歲及以上	24	11	15.0	−32.3
總計	77	254	79.8	−16.3

離均差平方和:

身高 $$\frac{(140)^2}{31}+\frac{(103)^2}{22}+\frac{(11)^2}{24}-\frac{(254)^2}{77}=281.66.$$

同理，求得體重的爲 26.56，心像面積的爲 53.65.

離均差積之和:

身高與體重

$$\frac{(140)(21.6)}{31}+\frac{(103)(43.2)}{22}+\frac{(11)(15.0)}{24}-\frac{(254)(79.8)}{77}=43.44.$$

同理，身高與心像面積的爲 110.56，體重與心像面積的爲 2.79.

我們把表 11-3 總的離均差平方和及積和寫在表 11-12 的第一行，把剛纔求得的組間離均差平方和及積和塡入第二行，將'總計'與'組間'兩行相減，即得'組內'各值(參閱公式 9-2 及第 10-2 節). 若分

【表 11-12】 組間與組內之離均差平方和及積和

	離均差之平方和			離均差積之和		
	身高(X_1)	體重(X_2)	心像面積(Y)	身高與體重	身高與心像面積	體重與心像面積
總 計	4,724.13	4,290.94	8,622.24	2,311.86	3,951.67	3,484.80
組 間	281.66	26.56	53.65	43.44	110.56	2.79
組 內	4,442.47	4,264.38	8,568.59	2,268.42	3,841.11	3,482.01

別求各組的組內變異或共變異亦可，但以上述方法爲便. 於是用公式(6-1)求'組內'各相關係數，如

$$r_{12}=\frac{2,268.42}{\sqrt{4,442.47}\sqrt{4,264.38}}=.5212.$$

同理，求得 $r_{Y1}=.6226,\quad r_{Y2}=.5760.$

再將此三個 r 值代入公式(11-1)，得

$$\beta_{Y1\cdot2}=.4426,\quad \beta_{Y2\cdot1}=.3453.$$

按公式(11-2)化爲原來單位，

$$b_{Y1\cdot2}=.4426\frac{\sqrt{8,568.59}}{\sqrt{4,442.47}}=.6147,$$

$$b_{Y2\cdot1}=.3453\frac{\sqrt{8,568.59}}{\sqrt{4,264.38}}=.4895.$$

將表(11-3)所列均數及此處之 b 值代入公式(11-3)，得組內三元迴

歸方程式

$$Y_e = 84.79 + .6147(X_1 - 163.30) + .4895(X_2 - 51.04).$$

我們所需要的是各組的修正均數，在多元迴歸中修正值的計算法與公式(5–7)相似，即

$$Y_a = Y - b_{Y1\cdot2}(X_1 - \bar{x}_1) - b_{Y2\cdot1}(X_2 - \bar{x}_2). \qquad \text{公式(11–21)}$$

變數更多時可以類推．將各值代入，則得

$$Y_a = Y - .6147(X_1 - 163.30) - .4895(X_2 - 51.04)$$
$$= Y - .6147X_1 - .4895X_2 + 125.38.$$

又自表 11–11，用公式(2–2)求得各組觀察的均數爲：

	身高(X_1)	體重(X_2)	心像面積(Y)
19—24 歲	164.516	50.697	85.65
25—30 歲	164.682	51.964	84.82
31 歲及以上	160.458	50.625	83.65

以各組均數代入上式，得各組修正均數如下：

19—24 歲	25—30 歲	31 歲及以上
85.08	83.53	85.61

這是把身高、體重都化爲相等時 ($\bar{x}_1 = 163.30$, $\bar{x}_2 = 51.04$) 各組心像面積的平均值．此三個修正均數有無顯著的差別，可以 F 值測驗之，其原理與第 10–1—10–3 節相同．表11–13的左半部是測驗三個觀察均數間相差的顯著性的，其離均差平方和，錄自表 11–12，組內自由度爲$(31-1)+(22-1)+(24-1)=74$，組間自由度爲 $3-1=2$．又組間均方爲 26.83，乃表示三個觀察均數間的變異，而組內均方爲 115.79，則表示各組內的個別差異或誤差，今前者小於後者，故知三

【表 11-13】 各年齡組心像面積之觀察均數及修正均數之顯著性

變異來源	自由度	心像面積 平方和	心像面積 均方	R^2	估計之誤差 自由度	估計之誤差 平方和	估計之誤差 均方
總　計	76	8,622.24		.47171	74	4,555.04	
組　內	74	8,568.59	115.79	.47450	72	4,502.79	62.54
組　間	2	53.65	26.83		2	52.25	26.13

個觀察均數並無眞正的差別.（本來顯著性測驗至此就可停止，而不必再求各組修正均數並作進一步分析了，但此處仍繼續下去，俾讀者得窺全豹.）

表 11-13 的右半部是爲測驗三個修正均數間相差的顯著性的.在第 11-3—10-4 兩節，已求得總的多元相關爲

$$R^2 = .47171, \quad (1-R^2)\Sigma(Y-\bar{y})^2 = 4,555.04.$$

同理，組內之多元相關爲

$$R^2 = .6226(.4426) + .5760(.3453) = .47450.$$

其估計誤差之平方和爲

$$(1-.4745)(8,568.59) = 4,502.79.$$

在計算三元迴歸的估計誤差時，除已用均數外，又用了兩個部分迴歸係數，故其自由度較左半部又各少 2 個.‘總計’與‘組內’兩估計誤差平方和之相差爲 52.25，附 2 個自由度是表示三個修正均數間的變異的.化成均方後，我們可以看到組間均方反較組內（誤差）均方爲小，故知三個修正均數間的差別，還不如各組組內個別差異之大，這樣便證實了本節最初的預料.

【練　習　題】

1. 玆以 X_1 為體重(仟克)，X_2 為身高(厘米)，Y 為肺活量(百立方厘米). 試用下列統計數(6)，求一由體重、身高推算肺活量之迴歸方程式.

我國十九歲以上女子 258 人體重、身高
及肺活量之均數相關係數等

$\bar{x}_1=49.02$,	$\bar{x}_2=155.16$,	$\bar{y}=24.961$.
$\Sigma(X_1-\bar{x}_1)^2=8,451.88$,	$r_{12}=.5993$,	
$\Sigma(X_2-\bar{x}_2)^2=7,261.16$,	$r_{Y1}=.4108$,	
$\Sigma(Y-\bar{y})^2=4,147.60$,	$r_{Y2}=.5645$.	

2. 測驗前題所得兩 β 之顯著性. 若有一 β 為不顯著，即表示在迴歸方程式中，此 β 所屬之變數可以除去. 試將該變數除去後，另求一推算肺活量之直線迴歸方程式.（參閱公式 6–9，式中所需之標準差可由前題之離均差平方和求得.）

3. 問在實驗資料中，下列情形為可能否？(參閱公式 11–5.)

$r_{Y1}=0.6$, $r_{Y2}=0.8$, $r_{12}=-0.5$.

4. 已知 $r_{Y1}=0.6$, $r_{Y2}=0.8$, 試求各 r_{12} 之值對於 $R_{Y\cdot12}$ 之影響. r_{12} 之值可用 $-0.8, -0.6, -0.4, -0.2, 0, 0.2, \cdots\cdots$.

5. 下表為自 1918 年美洲流行性感冒 (influenza) 大流行時 34 個城市之紀錄中(12)求得之相關係數. 其中 X_1 為人口之年齡分配，即某城市人口中少年所占的成分；X_2 為該城市的緯度；X_3 為每一萬人口中心臟病患者之死亡數；Y 為 25 星期內每一萬人口中患流行性感冒者之死亡數. 試用 Gauss 法 (表 11–9)，求 $\beta_{Y1\cdot23}$, $\beta_{Y2\cdot13}$, $\beta_{Y3\cdot12}$，並測驗其顯著性. 此表看法如 $r_{12}=0.0780$, $r_{Y1}=-0.0258$,

餘類推.

$N=34$	X_1	X_2	X_3
X_2	0.0780		
X_3	−0.6217	−0.2955	
Y	−0.0258	−0.3482	0.4705

6. 仿公式(11–10)寫出 $r_{14\cdot3}$, $r_{24\cdot3}$, 然後用下列公式求第5題之 $r_{Y1\cdot23}$, 並用公式(11–11)及表 11–5 測驗其顯著性. 問結果與上題 $\beta_{Y1\cdot23}$ 之顯著性相同否?

$$r_{12\cdot34}=\frac{r_{12\cdot3}-(r_{14\cdot3}\times r_{24\cdot3})}{\sqrt{1-r^2_{14\cdot3}}\sqrt{1-r^2_{24\cdot3}}}$$

注意第5題中之 Y 相當於此公式中之 X_1, 而第5題之 X_1 相當於此公式中之 X_2, 餘類推. 學者須養成活用公式之習慣.

7. P. H. Stevenson 氏(8)曾測量中國男子之身高(厘米)、坐高(厘米)、胸圍(厘米)、體重(仟克)及肺活量(升). 試用Doolittle氏法(表 11–10), 求一推算肺活量之五元迴歸方程式. 所需均數、標準差及相關係數如下:

	均數	標準差	相關係數			
			身高	坐高	胸圍	體重
身高	166.03	5.18				
坐高	89.97	3.02	0.720			
胸圍	80.85	4.94	0.389	0.327		
體重	56.08	6.60	0.569	0.529	0.780	
肺活量	3.47	0.56	0.348	0.331	0.336	0.306

註. $b_{Y1\cdot234}=\beta_{Y1\cdot234}\dfrac{S_Y}{S_1}$, 餘類推.

8. 下表爲飼養實驗之紀錄(13),所用動物有五組,每組予以一種

第一組			第二組			第三組			第四組			第五組		
原始體重	所食飼料	最後體重	原始體重	所食飼料	最後體重	原始體重	所食飼料	最後體重	原始體重	所食飼料	最後體重	原始體重	所食飼料	最後體重
30	674	195	26	699	194	39	708	203	41	716	226	41	831	242
21	628	177	24	626	204	34	614	190	35	769	230	36	754	225
21	661	180	20	668	200	32	733	221	32	733	218	32	722	205
33	694	200	35	668	201	35	663	173	34	742	235	35	728	228
27	713	197	25	707	195	32	607	185	32	624	197	32	646	196
24	585	170	26	651	187	35	745	225	35	710	210	36	678	196
20	575	150	20	672	191	30	637	190	30	742	217	30	763	230
29	638	180	31	660	200	29	662	201	28	648	205	28	625	170
28	632	192	29	769	208	32	609	174	34	628	200	32	710	216
26	637	184	27	666	218	25	596	180	26	601	191	26	651	175

飼料.若將原始體重及食量均化爲相等後,各組最後(實驗終)體重之均數有顯著的差別否?試以多元共變數法測驗之.

【參　考　文　獻】

(1) 鄒仲、陳又新:國人心臟及主動脈之X光測量,中華醫學雜誌,26, 1:16–37(二十九年).

(2) 郭祖超:推算國人心臟面積之公式,現代醫學,1,2:109–119(三十三年八月).

(3) Hodges, P. C. and Eyster, J. A. E., Estimation of cardiac area in man, Am. J. Roentgenol. and Rad. Therap., 12:252 (1924).

(4) Mills, F. C., Statistical Methods, pp. 544–545, Henry Holt & Co., New York (1938).

(5) Wallace, H. A. and Snedecor, G. W., Correlation and Machine Calculation, revised edition, Iowa State College Official Publication, 30, No. 4 (1931).

(6) Kuo, T. C. and Wu, C. H., Regression equations for estimating body weight and vital capacity of normal Chinese students, Proceedings of Chinese Physiological Society, Chengtu Branch, 2, 4 and 5 : 101–104 (1944).

(7) Tsai, C. and Wu, C. H., A statistical study of the vital capacity of senior middle school and college students, Chinese Journal of Physiology, 14, 1 : 95–116 (1939). 原來資料係二氏所供給.

(8) Stevenson, P. H., Prediction formulae for height, sitting height, chest girth, weight and vital capacity for adult male Chinese, Chinese Journal of Physiology, 9, 3 : 213–222 (1935).

(9) Rietz, H. L., Editor-in-chief, Handbook of Mathematical Statistics, pp. 139–142, Houghton Mifflin Co., New York (1924).

(10) Doolittle, M. H., United States Coast and Geodetic Survey Report, 1878 : 115.

(11) Carl Friedrich Gauss, Theoria Combinationis Observatorium, Pars Posterior (1821), Werke, 4 : 31; Supplement (1826), Werke, 4 : 71.

(12) Raymond Pearl, Public Health Reports, 34:1743 (1919); 36:289 (1921).

(13) Crampton, E. M. and Hopkins, J. W. J., Nutrition, 8: 329-340 (1934).

第十二章　曲線迴歸

12-1. 引言　本書前面所講的直線迴歸，在很多地方固然非常合用，但在有些資料中，變數與變數間的關係卻沒有這樣簡單．要把兩個或幾個變數間相隨着變異的情形精確地表示出來，這是曲線配合(curve fitting)中的一個問題，也就是本章所要討論的曲線迴歸(curvilinear regression)．在曲線配合中，直線的配合只是一個最簡單的特例罷了．

對於非直線的資料配合——曲線的動機有幾種：有時對於自變數(如X)的任何特殊數值，要推算其因變數(如Y)的相當的估計值，那麼就得把不規則的資料予以修勻(smoothing)，而將觀察資料中所未備的X值，用內插法求得其相當的Y值．有時配合曲線的目的，卻在發現或試驗變數所遵循的法則，如生長曲線(growth curve)等．但有時我們的目的，卻祇在除去相關係數或實驗誤差中由於非直線迴歸而引起的不正確性，至於變數間具有何種關係，卻是不去考慮的．

將資料配合曲線的問題曾從許多不同的方向去發展．除了試驗某種法則以外，大概主要的困難，是在選擇一個適當的方程式．方程式的類別，可粗分爲兩大類：

甲．多項式

$$Y=a+bX,$$

乙．對數曲線

$$Y=a+b\log X,$$

$$Y=a+bX+cX^2, \qquad \log Y=a+bX,$$

$$Y=a+bX+cX^2+dX^3, \qquad \log Y=a+b\log X,$$

$$\cdots\cdots\cdots\cdots. \qquad \cdots\cdots\cdots\cdots.$$

多項式中第一個是簡單的直線方程式(linear equation),此式卽爲第五章公式(5-3)的變相,第二個爲簡單抛物線(parabola)或二次曲線,第三個爲三次曲線.簡單抛物線祇有一個極大點(maximum point)或極小點(minimum point),而無轉向點(point of inflection).三次曲線則有極大點、極小點和轉向點各一個.曲線的次數愈高,則極大點與極小點也愈多,且轉曲亦愈多而愈快.

對數曲線可認爲多項式之變形.如在 $Y=a+bX$ 中,若以 $\log X$ 代 X,卽得對數曲線中的第一個方程式.在生物現象中,有不少資料可配合對數曲線,惟因對數方程式中 X 不能有負值,故其應用不若多項式之廣.

12-2. 對數曲線之配合法 有些比較簡單的生長現象,有一種特性,便是在任何時刻所增加的數目與當時已有的大小成正比例.這種生長的情形,與銀行中所用複利的法則相同.當培養細菌時,在生長的某一階段中,細菌的數目是依照這個法則而增加的.這一種對數曲線,特名之爲指數生長曲線(exponential growth curve).例如表 12-1 第 1,2 兩行爲普通變形桿菌(proteus vulgaris)在實驗室中生長的紀錄(1).爲測驗是項資料是否適合於對數曲線起見,先將第 2 行的數值化成普通對數,記於第 3 行.然後把第 1 行與第 3 行的數值,繪成點子,見圖 12-1 上半部.我們可以看到這些黑點,差不多在一根直線上,這表示手頭的資料可以配合 $\log C=a+bX$ 形式的

【表12–1】 指數生長曲線之配合法(普通變形桿菌之生長)

(1) 孵養時間 (分鐘, X)	(2) 細胞計數 (10^7, C)	(3) C 之對數 ($Y=\log C$)	(4) X^2	(5) Y^2	(6) XY
0	1.53	0.185	0	.03423	0
130	13.60	1.134	16,900	1.28596	147.420
175	26.40	1.422	30,625	2.02208	248.850
225	59.20	1.772	50,625	3.13998	398.700
252	77.60	1.890	63,504	3.57210	476.280
315	148.20	2.171	99,225	4.71324	683.865
1,097		8.574	260,879	14.76759	1,955.115

$$\bar{x}=182.83, \qquad \bar{y}=1.429.$$

$$\Sigma(X-\bar{x})^2=260,879-\frac{(1,097)^2}{6}=60,310.83,$$

$$\Sigma(Y-\bar{y})^2=14.76759-\frac{(8.574)^2}{6}=2.51534,$$

$$\Sigma(X-\bar{x})(Y-\bar{y})=1,955.115-\frac{(1,097)(8.574)}{6}=387.502.$$

$$Y_e-1.429=\frac{387.502}{60,310.83}(X-182.83)=.006425X-1.1747,$$

$$Y_e=.006425X+.2543. \qquad \text{(A)}$$

曲線. 在配合其他形式的對數曲線時, 可將自變數 X 化爲對數, 或將自變數與因變數一併化爲對數, 若其中有一種可使圖上各點幾乎在一直線上, 那麼就合用了. 否則便不能配合這一類的對數曲線. 這樣把資料化成對數或其他函數使各點近於一直線的手續, 叫做直線化(rectification).

經過直線化的測驗以後, 就可進行配合的手續了. 計算的過程

與第五章配合簡單直線時相同(參閱表 5–1).結果得

$$Y_e = .006425X + .2543. \quad \text{(A)}$$

在計算過程中,我們以 Y 代 $\log C$,故此式卽為

$$\log C_e = .006425X + .2543.$$

將表 12–1 第 1 行各 X 值代入前式,得相當的 Y_e 值,若代入後式,再由對數表查得眞數,則得 C_e 值.茲將 Y_e 與 C_e 各估計值列後:

X	Y_e	C_e
0	.254	1.80
130	1.090	12.29
175	1.379	23.92
225	1.700	50.11
252	1.873	74.71
315	2.278	189.75

以 X 與 Y_e 兩行的數值繪圖,則得圖 12–1 上半部的一條直線,又以 X 與 C_e 兩行的數值繪圖,則為該圖下半部的曲線.曲線附近的小圓圈表示 C 的觀察值,卽由表 12–1 第 1, 2 兩行的數值所繪成.

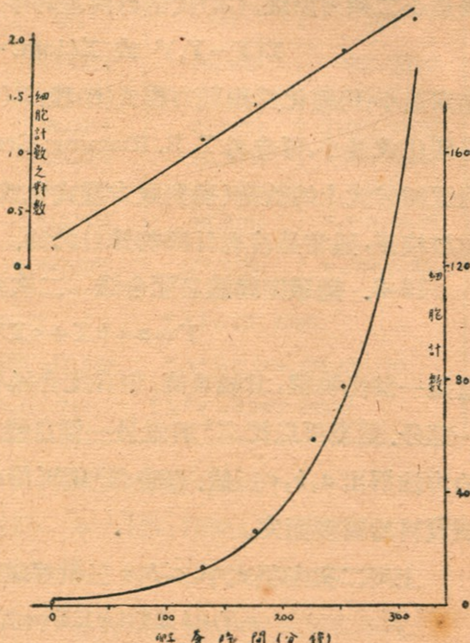


圖 12-1. 普通變形桿菌之生長及所配合之對數曲線

迴歸方程式 $\log C_e = .006425X + .2543$ 可以化成另一個形式,

因爲 $\log 1.0149 = .006425$，又 $\log 1.796 = .2543$，於是此式簡化如下:

$$\log C_e = .006425X + .2543 = X\log 1.0149 + \log 1.796$$

$$= \log(1.0149)^x + \log 1.796 = \log[1.796(1.0149)^x],$$

$$\therefore \quad C_e = 1.796(1.0149)^x. \qquad \text{(B)}$$

按最小二乘方原理，(A)式估計誤差的平方和，即

$$\Sigma(Y - Y_e)^2 \text{ 或 } \Sigma(\log C - \log C_e)^2$$

應爲最小，但經化爲指數方程式(B)後，$\Sigma(C - C_e)^2$ 卻並不是最小，若必須化成最小，可參考 T. R. Running: Empirical Formulas 等書(1). 但這種形式上的改變(將對數方程式化成指數方程式)，以及將常數稍加校正，通常是沒有什麼特殊目的的.

12-3. 簡單拋物線之配合法 二次多項式的一般形式是

$$Y = a + bX + cX^2.$$

這是一個拋物線，其軸垂直. 但在配合過程中，往往只用到拋物線的一部分. 這裏可以把 X^2 看做另一個自變數，於是即可應用多元迴歸的方法解出 a, b, c 三值. 有時 X^2 項可用 $\sqrt{X}$, $\log X$ 或 $1/X$ 等代替，視資料的需要而定.

上列二次多項式中，a, b, c 三數可以下列標準方程式解出來:

$$\left.\begin{aligned} aN + b\Sigma X + c\Sigma X^2 &= \Sigma Y \\ a\Sigma X + b\Sigma X^2 + c\Sigma X^3 &= \Sigma XY \\ a\Sigma X^2 + b\Sigma X^3 + c\Sigma X^4 &= \Sigma X^2 Y \end{aligned}\right\}. \qquad \text{公式(12-1)}$$

玆將氫游子濃度與雞蛋白凝固百分數(3)的資料，配合一簡單拋物線(表12-2第1, 2兩行)，各 X^2, X^3, X^4 等數值可由 Barlow's Tables(4)

查到，相加後卽得所需要的總和，惟 XY 及 X^2Y 等乘積則須逐一計算。各乘方及乘積的總和，見表 12–2 之末。

【表 12–2】　簡單拋物線之配合法

（氫游子濃度對於雞蛋白凝固之影響）

溶液之 pH X (1)	凝固百分數 Y (2)	XY (3)	X^2 (4)	X^2Y (5)
4.51	47.3	213.323	20.3401	962.08673
4.58	51.8	237.244	20.9764	1,086.57752
4.69	54.6	256.074	21.9961	1,200.98706
4.83	55.5	268.065	23.3289	1,294.75395
4.98	51·2	254.976	24.8004	1,269.78048
5.13	42.5	218.025	26.3169	1,118.46825
28.72	302.9	1,447.707	137.7588	6,932.65399

$N=6$,　　$\Sigma X=28.72$,　　$\Sigma X^2=137.7588$,

$\Sigma X^3=662.157,748$,　　$\Sigma X^4=3,189.434,081,16$.

$\Sigma Y=302.9$,　$\Sigma XY=1,447.707$,　$\Sigma X^2Y=6,932.65399$.

將各總和值代入公式(12–1)，得標準方程式如下：

$$6a + 28.72b + 137.7588c = 302.9,$$

$$28.72a + 137.7588b + 662.1577c = 1,447.707,$$

$$137.7588a + 662.1577b + 3,189.4341c = 6,932.654.$$

解之（參閱前章第 11–8 節），得

$$a=-2,151.625,\qquad b=921.016,\qquad c=-96.103.$$

因得　$Y_c=-2,151.625+921.016X-96.103X^2$.　　(C)

將各 X 值代入，得 Y 之估計值如下：

X	4.51	4.58	4.69	4.83	4.98	5.13
Y_e	47.41	50.73	54.05	54.91	51.64	44.05

此拋物線繪成圖形，則如圖12–2所示，曲線旁的小圓圈是根據表12–2第1,2兩行的觀察資料所繪製的.

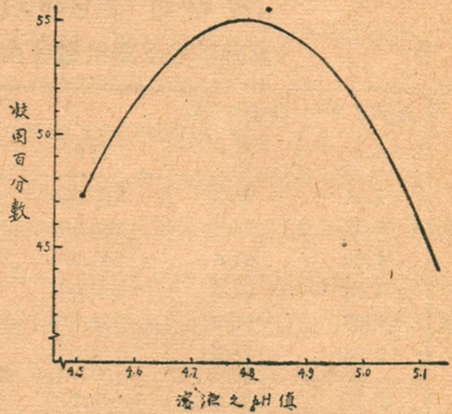


圖 12-2. 氫游子濃度對於蛋白質凝固之影響

欲求雞蛋白凝固之百分數最高時，其 pH 值為若干，則可令(C)式之微分等於0，而解得 X 值. 按常數 a 之微分為0，又 bX 之微分等於 b，又 cX^2 之微分等於 $2cX$(參閱初等微積分). 因得下式:

$$921.016+2(-96.103)X=0, \quad 921.016-192.206X=0,$$

$$X=\frac{921.016}{192.206}=4.79.$$

因知當 $pH=4.79$ 時，雞蛋白因震盪(shaking)而凝固之百分數為最高.

由圖12–1及12–2，可見由實際資料畫成之點子(可稱為觀察點，即圖上的小圓圈)都在理論曲線的附近，卻很不容易全部都在理論曲線上，因為實驗資料總受到抽樣變動的影響，所以各點與理論曲線稍有出入是不可避免的. 本書作者曾見我國雜誌上發表之研究報告中若干隨手畫成而通過所有觀察點的曲線，他們就根據這樣得

來的曲線，推論極大點或極小點在那裏．要知道隨手畫成的曲線，就是形狀也不一定對，若據以下結論，更是很危險的！

曲線迴歸對於相關的影響，我們可以在此提一下．若將表 12–2 第 1，2 兩行的原來資料求相關係數，則得 $r=-.37411$；若求觀察值 Y 與曲線迴歸的估計值 Y_e 間的相關係數，則得 $R=.99207$（此處以 X 與 X^2 當做兩個變數，故按前章第 11–3 節的定義，此 R 即爲多元相關係數）．爲測驗全體中迴歸之曲直計，可用變異數分析法解答之．先假設全體中 pH 與雞蛋白凝固百分數具有直線的關係，求得估計誤差之平方和（公式 6–12）爲

$$(1-r^2)\Sigma(Y-\bar{y})^2=[1-(-.37411)^2](118.23)=101.68.$$

同理，假設全體中呈曲線的關係時，其估計誤差之平方和（參閱前章第11–4節）爲

$$(1-R^2)\Sigma(Y-\bar{y})^2=[1-(.99207)^2](118.23)=1.87.$$

前者之自由度爲 $6-2=4$，後者爲 $6-3=3$．將平方和及自由度列於表 12–3．由此可見，引用二次曲線以後，其估計誤差之平方和減少

【表 12–3】　迴歸曲直之測驗（直線迴歸與曲線迴歸之比較）

變異來源	自由度	平方和	均方	F
直線迴歸之估計誤差	4	101.68		
曲線迴歸之估計誤差	3	1.87	.62	
迴歸之曲線性	1	99.81	99.81	160.98

了 99.81，其相當的自由度爲 1．經顯著性測驗後，其減少之部分爲非常顯著（$n_1=1$，$n_2=3$，1％點$=34.12$）．故直線迴歸之假設應予放棄，並知 pH 與雞蛋白凝固百分數間的迴歸呈非常顯著的曲線性．

讀者或許會發生一個疑問，覺得前節將資料直線化與本節配合二次曲線，似乎大相逕庭。其實這種差別只是表面的。我們若在三度空間裏繪一圖形，以 Y，X 與 X^2 用三根軸分別表示，則(C)式所代表的迴歸面(regression surface)將成爲一平面。所以前節與本節對於資料的處理其實並無二致，只在二度空間的圖形上看來好像不同吧了。

還有一點要注意的，不論直線或曲線迴歸，若在全距以外作外補插(extrapolation)，都要十分謹慎。因爲手頭的資料不足以告訴我們在其全距以外的趨勢怎樣。若將圖 12–2 的拋物線延長開去，其兩端都會伸展到橫軸以下，這時雞蛋白凝固的百分數將爲負數，這是很費解的。所以外補插往往得到錯誤的結果，必須謹慎小心纔好。

12–4. 與直線迴歸相離之測驗 前節所述測驗迴歸曲直的方法，是先要把曲線配合好了，然後進行變異數分析的。若資料分成若干組，並且只注意到是否呈直線迴歸的問題，那麼測驗步驟可以簡單得多了。B. J. Vos 與 W. T. Dawson 兩氏，曾將美國標準哇巴因(United States standard ouabain)徐徐注射於貓之靜脈內，以測定其致死量(lethal dose, Y)。注射率(X)有四種，各倍於前。此處所研究之問題爲 Y 與 X 間是否呈直線迴歸。茲測驗其顯著性如次。

先用變異數分析法測驗四個均數(18.1, 28.2, 48.3, 70.2)間有無顯著的差別(參閱第九章第 9–2 節)，結果 F 值大於 1% 點，故四均數間之差別爲非常顯著。此種差別之主要原因，係由注射率 X 而來。我們要進一步研究 $\bar{y}$ 與 X 間的關係，是否與直線迴歸有顯著的差別。

【表12-4】 哇巴因對於貓之致死量(表內為簡縮數，各減去50單位)

	注射率 X $\left(\frac{毫克/仟克/分鐘}{1,000}\right)$				總計
	1.04575	2.0915	4.183	8.366	
Y	5	3	34	51	
	9	6	34	56	
	11	22	38	62	
	13	27	40	63	
	14	27	46	70	
	16	28	58	73	
	17	28	60	76	
	20	37	60	89	
	22	40	65	92	
	28	42			
	31	50			
	31				
ΣY	217	310	435	632	1,594
k	12	11	9	9	41
$\bar{y}$	18.1	28.2	48.3	70.2	38.9
ΣY^2	4,727	10,788	22,261	45,940	83,716

【表 12-5】 變異數分析(哇巴因四種注射率之致死量)

變異來源	自由度	離均差平方和	均 方	F	1%點
總 變 異	40	21,744			
組間(注射率)	3	16,093	5,364.3	35.13	4.36
組內(誤 差)	37	5,651	152.7		

在迴歸中需用 $\Sigma(X-\bar{x})^2$ 及 $\Sigma(X-\bar{x})(Y-\bar{y})$ 等，惟因 Y 值分成四組，同組各 X 值皆相同，故計算時須加權(weighted)，如

$$\Sigma X=12(1.04575)+11(2.0915)+9(4.183)+9(8.366)=148.496,$$

同理 $\Sigma X^2=12(1.04575)^2+\cdots\cdots+9(8.366)^2=848.628,$

又 $\Sigma XY=(1.04575)(217)+(2.0915)(310)$

$$+(4.183)(435)+(8.366)(632)=7,982.21.$$

由此計算離均差平方和及積和:

$$\Sigma(X-\bar{x})^2=848.628-(148.496)^2/41=310.797,$$

$$\Sigma(X-\bar{x})(Y-\bar{y})=7,982.21-(148.496)(1,594)/41=2,208.98.$$

在 Y 方面,由於直線迴歸而來之變異爲(參閱第五章第 5-2 節)

$$(2,208.98)^2\div 310.797=15,700,$$

相當於 1 個自由度. 於是與表 12-5 注射率間之變異並列於表 12-6,

【表 12-6】 與直線迴歸相離之測驗

變異來源	自由度	平方和	均方	F	5% 點
注射率(表 12-5)	3	16,093			
直線迴歸	1	15,700			
直線迴歸之估計誤差	2	393	196.5	1.3	3.26
誤差(表 12-5)	37		152.7		

二者自由度之相差爲 2, 平方和相差 393. 此值卽爲直線迴歸之估計誤差,其均方與誤差之均方相比較,並無顯著的差別. 故各點與直線迴歸之距離實與隨機抽樣所得者無異,易言之, $\bar{y}$ 與 X 間的關係與直線迴歸,並無眞正差別.

12-5. 正交多項式 (orthogonal polynomials) 之配合法 前面所講的直線和拋物線方程式是多項式中最簡單的兩個,現在我們要講多項式的一般配合法. 配合時方法便利,且在理論上有充分根據者,當推 R. A. Fisher 氏之正交多項式(5),

$$Y=A+BX_1+CX_2+DX_3+EX_4+\cdots\cdots, \qquad \text{公式(12-2)}$$

式中 $A,B,C,\cdots\cdots$ 爲待定的數值，X_1,X_2 等各爲 X 的多項式〔不同次(degree)之多項式所代表之效果彼此獨立，換言之，不同多項式間無相關之存在. 由幾何學上解釋，若兩點代表無相關之事實，則其與球心之連線相正交〕，故配合所得方程式的最後形式仍爲

$$Y=a+bX+cX^2+dX^3+\cdots\cdots.$$

正交多項式爲一善於變化的曲線，能配合很多種的資料，它也包括直線和簡單拋物線在內. 本節所舉的例子，將限於 X 以 1 遞進，且一 X 值祇相當於一 Y 值的資料. 若 X 以異於 1 的等距離遞進，則可以此公共距離除各 X 值使化成單位距離. 倘 X 值的距離不等，或一個 X 值相當於幾個 Y 值時，則可用第 12–4 節的方法以測驗迴歸之曲直.

此法配合非常便利而有規則，循序前進，無須經嘗試錯誤(trial and error)之過程；且每配合一項，就可測驗其顯著性，故其配合適度如何隨時可以知道，這樣就不會枉費很多工夫而一無所成.

下面所用的資料是在實驗室中釀母(yeast)的生長紀錄. 實驗時將釀母種植於特殊培養基中，逐日計算其單位容量內的細胞數，記錄如下表:

日數 (X)	1	2	3	4	5	6	7	8	9	10
釀母細胞數 (Y)	52.5	87.8	147.6	203.6	289.7	365.0	414.3	447.4	469.0	481.9

爲使讀者醒目起見，我們將分三小節來討論:(一)配合的步驟，(二)估計值的計算，及(三)方程式的計算. 請分述於後.

(一)配合的步驟　若將上列原來資料繪成圖形，將見類似一個

拉長的 S 字，故此資料大概可配合三次多項式，不過在過程中需要配合到四次曲線，由顯著性測驗的結果，若四次曲線比三次曲線沒有多大進步，那樣就可以停止了.

先將 Y 值列於表 12-7 的第 1 行，其總和為 $S_1=2,963.8$. 此值

【表 12-7】 配合正交多項式所需之計算

Y	2	3	4	5
52.5	52.5	52.5	52.5	52.5
87.8	140.3	192.8	245.3	297.8
147.6	287.9	480.7	726.0	1,023.8
208.6	496.5	9 7.2	1,703.2	2,727.0
289.7	786.2	1,763.4	3,466.6	6,193.6
365.0	1,151.2	2,914.6	6,381.2	12,574.8
414.3	1,565.5	4,480.1	10,861.3	23,436.1
447.4	2,012.9	6,493.0	17,354.3	40,790.4
469.0	2,481.9	8,974.9	26,329.2	67,119.6
481.9	2,963.8	11,938.7	38,267.9	105,387.5
$S_1=2,963.8$	$S_2=11,938.7$	$S_3=38,267.9$	$S_4=105,387.5$	$S_5=259,603.1$
$a=296.3800$	$b=217.0673$	$c=173.9450$	$d=147.3951$	$e=129.6719$
$a'=296.3800$	$b'=79.3127$	$c'=-6.9319$	$d'=-3.5493$	$e'=0.6351$
$A=296.3800$	$B=52.8751$	$C=-2.88829$	$D=-.985917$	$E=0.1323125$

以項目數 $n=10$ 除之，得 $a=296.38$，寫在 S_1 的下面. 此為下列公式中求得之第一數. 其餘公式將在後面用到.

$$\left.\begin{aligned} a&=\frac{S_1}{n}=\bar{y} \\ b&=\frac{(1)(2)}{n(n+1)}(S_2) \\ c&=\frac{(1)(2)(3)}{n(n+1)(n+2)}(S_3) \end{aligned}\right\}. \quad \text{公式(12-3)}$$

$$\left.\begin{aligned}
d&=\frac{(1)(2)(3)(4)}{n(n+1)(n+2)(n+3)}(S_4)\\
e&=\frac{(1)(2)(3)(4)(5)}{n(n+1)(n+2)(n+3)(n+4)}(S_5)
\end{aligned}\right\}$$

....

再將該表的 a 值抄下來作為 a'，此為第二組公式中的第一個:

$$\left.\begin{aligned}
a'&=a\\
b'&=a-b\\
c'&=a-3b+2c\\
d'&=a-6b+10c-5d\\
e'&=a-10b+30c-35d+14e\\
f'&=a-15b+70c-140d+126e-42f\\
g'&=a-21b+140c-420d+630e-462f+132g\\
h'&=a-28b+252c-1,050d+2,310e-2,772f\\
&\qquad +1,716g-429h
\end{aligned}\right\}. \text{公式(12–4)}$$

上列公式中各字母前之係數，可將下式依次累乘求得，式中 r 為該項多項式之次數(公式 12–2 之每項代表一多項式，參閱後面第三小節).

$$\frac{r(r+1)}{1\cdot 2},\ \frac{(r-1)(r+2)}{2\cdot 3},\ \frac{(r-2)(r+3)}{3\cdot 4},\ \cdots\cdots. \text{公式(12–5)}$$

例如 h' 屬於 7 次多項式，以 $r=7$ 代入第一式，則得 28，代入第一、二兩式而乘之，得 252，餘類推.

倘有第三組公式，即為正交多項式中各項之係數:

$$\left.\begin{aligned}
A &= a' \\
B &= \frac{6}{n-1}(b') \\
C &= \frac{30}{(n-1)(n-2)}(c') \\
D &= \frac{142}{(n-1)(n-2)(n-3)}(d') \\
E &= \frac{630}{(n-1)(n-2)(n-3)(n-4)}(e')
\end{aligned}\right\}. \quad \text{公式(12-6)}$$

此後 F 之分子爲 2,772，G 之分子爲 12,012，H 之分子爲 51,480.
各分子之數值可用下式計算:

$$\frac{(2r+1)(2r)(2r-1)(2r-2)\cdots\cdots 2\cdot 1}{[r(r-1)(r-2)\cdots\cdots 2\cdot 1]^2}, \quad \text{公式(12-7)}$$

r 仍爲該項多項式之次數.

最後，第四組公式爲每配合一項所減少之離均差平方和:

$$\left.\begin{aligned}
&nA^2 = \frac{(\Sigma Y)^2}{n} \\
&\frac{n(n^2-1)}{12}(B^2) \\
&\frac{n(n^2-1)(n^2-4)}{180}(C^2) \\
&\frac{n(n^2-1)(n^2-4)(n^2-9)}{2,800}(D^2) \\
&\frac{n(n^2-1)(n^2-4)(n^2-9)(n^2-16)}{44,100}(E^2)
\end{aligned}\right\}. \quad \text{公式(12-8)}$$

此後三項之分母爲 698,544, 11,099,088, 176,679,360. 各分母之計算可用下式:

$$\frac{(2r+1)[2r(2r-1)(2r-2)\cdots\cdots2\cdot1]^2}{[r(r-1)(r-2)\cdots\cdots2\cdot1]^4}. \quad \text{公式(12-9)}$$

上例 $\Sigma Y^2=1,116,915.76$, 減去校正數 $nA^2=10(296.38)^2$, 則得 238,504.716, 卽爲 Y 之總離均差平方和 $\Sigma(Y-\bar{y})^2$. 若將原來資料配合一水平線 $Y=296.38$, 則其估計誤差之平方和卽等於 $\Sigma(Y-\bar{y})^2$.

第二步的配合自表 12-7 的第 2 行開始. 該行數值係將第 1 行各數累計得來, 如

$$52.5+87.8=140.3,$$

$$140.3+147.6=287.9,$$

$$287.9+208.6=496.5,$$

餘類推. 該行之末一數 2,963.8 應與 S_1 相等. 再將第 2 行各數相加, 得 $S_2=11,938.7$. 於是按公式(12-3), (12-4), (12-6), (12-8)計算下列各值:

$$b=\frac{1\times2}{10\times11}(11,938.7)=217.0673,$$

$$b'=296.3800-217.0673=79.3127,$$

$$B=\frac{6}{9}(79.3127)=52.8751,$$

$$\Sigma(Y-Y_e)^2=238,504.716-\frac{10(10^2-1)}{12}(52.8751)^2$$

$$=238,504.716-230,651.5365=7,853.1795.$$

最後一數爲配合直線迴歸後之估計誤差, 其自由度爲 $10-2=8$. 至

於由迴歸而減少的部分 230,651.5365，則有一個自由度. 茲以變異數分析法測驗其顯著性，見表 12–8 第一欄. 結果所得 F 值大於 1%

【表 12–8】 配合四次曲線時各階段所用自由度之顯著性

配合之次數	變異來源	平方和	自由度	均方	F	5%點	1%點
1	總變異	238,504.7160	9				
	迴歸	230,651.5365	1	230,651.5365	234.96		11.26
	誤差	7,853.1795	8	981.647			
2	迴歸	4,404.6917	1	4,404.6917	8.012	5.59	12.25
	誤差	3,848.4878	7	549.7839			
3	迴歸	3,002.4135	1	3,002.4135	21.29	5.99	13.74
	誤差	846.0743	6	141.0124			
4	迴歸	288.3967	1	288.3967	2.586	6.61	
	誤差	557.6776	5	111.5355			

點，故知由迴歸直線所減少之變異爲非常顯著.〔註：此處配合適度之測驗，與第 12–4 節與直線迴歸相離之測驗切勿混淆. 迴歸直線也許配合得並不好，但是只要它比均數 $Y=\bar{y}$ 的水平線配得好些，那麼在配合適度方面，由迴歸直線所減少的部分將爲顯著. 同時非直線性測驗（表 12–6）的結果可能表示各組 Y 的均數與迴歸直線有顯著的差別. 故遇有分組的資料（如表 12–4），當配合直線後，最好作表 12–6 那樣的測驗，若並無非直線的證據，便可不必再配合較高次的曲線了.〕

第三步爲配合含有二次多項式之項（公式 12–2 中之 X_2）. 由表 12–7 第 3 行開始，該行數值係將第 2 行各數累計而得，方法同前，結果得 $S_3=38,267.9$，

$$c=\frac{1\times2\times3}{10\times11\times12}(38,267.9)=173.945,$$

$$c'=296.38-3(217.0673)+2(173.945)=-6.9319,$$

$$C=\frac{30}{9\times8}(-6.9319)=-2.88829,$$

$$\Sigma(Y-Y_e)^2=7,853.1795-\frac{10(10^2-1)(10^2-4)}{180}(-2.88829)^2$$

$$=7,853.1795-4,404.6917=3,848.4878.$$

估計誤差之自由度爲 $8-1=7$（因較前又多配合一項），至於因迴歸而繼續減少之部分 4,404.6917，其自由度爲 1. 經顯著性測驗（表12–8 第二欄）後，此減少之部分爲顯著（此處與表 12–3 之測驗相同，請比較之）.

第四步爲配合含有三次多項式之項（公式 12–2 中之 X_3）. 依前法計算表 12–7 第 4 行之各累計值，得

$$S_4=105,387.5,$$

$$d=\frac{1\times2\times3\times4}{10\times11\times12\times13}(105,387.5)=147.3951,$$

$$d'=296.38-6(217.0673)+10(173.945)-5(147.3951)$$

$$=-3.5493,$$

$$D=\frac{140}{9\times8\times7}(-3.5493)=-.985917,$$

$$\Sigma(Y-Y_e)^2=3,848.4878-\frac{10(10^2-1)(10^2-4)(10^2-9)}{2,800}(-.985917)^2$$

$$=3,848.4878-3,002.4135=846.0743.$$

其自由度爲 $7-1=6$，乃因迴歸而減少之部分仍爲 1 個自由度．其顯著性測驗見表 12–8 第 3 欄，結果此減少之部分爲非常顯著．

第五步爲配合含有四次多項式之項（公式 12–2 中之 X_4）．由表 12–7 第 5 行，得 $S_5=259,603.1$，

$$e=\frac{1\times2\times3\times4\times5}{10\times11\times12\times13\times14}(259,603.1)=129.6719,$$

$$e'=296.38-10(217.0673)+30(173.945)-35(147.3951)$$
$$+14(129.6719)=.6351,$$

$$E=\frac{630}{9\times8\times7\times6}(.6351)=.1323125,$$

$$\Sigma(Y-Y_e)^2=846.0743-\frac{10(10^2-1)(10^2-4)(10^2-9)(10^2-16)}{44,100}(.1323125)^2$$

$$=846.0743-288.3967=557.6776.$$

其自由度爲 $6-1=5$．經顯著性測驗後，知因迴歸而減少之部分爲不顯著．故以此資料配合四次曲線，在精密度方面並不比三次曲線有何進步．配合工作至此即可停止，而以三次曲線表示其一般的趨勢．在有些問題中，得知理論曲線爲何種形式後，即可竣事．但有時需要製圖，有時需要方程式，故下文將分兩小節繼續討論．若只需作圖而不需方程式者，可閱小節（二）．需要方程式者，可逕閱小節（三）．

（二）估計值的計算　根據表 12–7 求得的 a'，b'，c' 等數值，我們可用統計學上的有限差數（finite differences）法，把估計值 Y_e 直接算出來，卻無須先求曲線的方程式．計算時先參照表 12–9，該表可用到 10 次曲線．表右半部所列多項式，按所配合曲線的次數決定需

【表 12-9】 計算末數及差數時所需之多項式及係數

多項式之次數	係數（n 表示樣本之大小）	多項式 0	1	2	3	4	5	6	7	8	9	10
0	（末　　數）	a'	$+3b'$	$+5c'$	$+7d'$	$+9e'$	$+11f'$	$+13g'$	$+15h'$	$+17i'$	$+19j'$	$+21k'$
1	$-\frac{2\cdot3}{n-1}$		b'	$+5c'$	$+14d'$	$+30e'$	$+55f'$	$+91g'$	$+140h'$	$+204i'$	$+285j'$	$+385k'$
2	$\frac{3\cdot4\cdot5}{(n-1)(n-2)}$			c'	$+7d'$	$+27e'$	$+77f'$	$+182g'$	$+378h'$	$+714i'$	$+1254j'$	$+2079k'$
3	$-\frac{4\cdot5\cdot6\cdot7}{(n-1)(n-2)(n-3)}$				d'	$+9e'$	$+44f'$	$+156g'$	$+450h'$	$+1122i'$	$+2508j'$	$+5148k'$
4	$\frac{5\cdot6\cdot7\cdot8\cdot9}{(n-1)(n-2)(n-3)(n-4)}$					e'	$+11f'$	$+65g'$	$+275h'$	$+935i'$	$+2717j'$	$+7007k'$
5	$-\frac{6\cdot7\cdot8\cdot9\cdot10\cdot11}{(n-1)(n-2)\cdots\cdots(n-5)}$						f'	$+13g'$	$+90h'$	$+442i'$	$+1729j'$	$+5733k'$
6	$\frac{7\cdot8\cdot9\cdots\cdots13}{(n-1)(n-2)\cdots\cdots(n-6)}$							g'	$+15h'$	$+119i'$	$+665j'$	$+2940k'$
7	$-\frac{8\cdot9\cdot10\cdots\cdots15}{(n-1)(n-2)\cdots\cdots(n-7)}$								h'	$+17i'$	$+155j'$	$+952k'$
8	$\frac{9\cdot10\cdot11\cdots\cdots17}{(n-1)(n-2)\cdots\cdots(n-8)}$									i'	$+19j'$	$+189k'$
9	$-\frac{10\cdot11\cdot12\cdots\cdots19}{(n-1)(n-2)\cdots\cdots(n-9)}$										j'	$+21k'$
10	$\frac{11\cdot12\cdot13\cdots\cdots21}{(n-1)(n-2)\cdots\cdots(n-10)}$											k'

用幾項. 如我們在上面配合的是 3 次曲線，所以該表各多項式祇用到 d' 項爲止(若配合的是 5 次曲線，那麼要用到 f' 項爲止，餘類推). 其中第一個多項式爲計算末了一個估計值用的，簡稱末數(terminal value, 上例卽爲相當於 $X=10$ 的 Y_e). 又該表第二欄的'係數'，是與右邊多項式相乘用的. 由'係數'與'多項式'的乘積，得到所需的各個差數. 請看下面實例，當可明瞭:

三次曲線所需之末數及各差數，

$$\text{末　　數} = a' + 3b' + 5c' + 7d'$$
$$= 296.38 + 3(79.3127) + 5(-6.9319) + 7(-3.5493)$$
$$= 474.8135,$$

$$\text{第一差數 (first difference)} = -\frac{2\cdot 3}{n-1}(b' + 5c' + 14d')$$
$$= -\frac{2\cdot 3}{9}\, 79.3127 + 5(-6.9319) + 14(-3.5493)]$$
$$= 3.3580,$$

$$\text{第二差數 (second difference)} = \frac{3\cdot 4\cdot 5}{(n-1)(n-2)}(c' + 7d')$$
$$= \frac{3\cdot 4\cdot 5}{9\cdot 8}[-6.9319 + 7(-3.5493)] = -26.4808,$$

$$\text{第三差數 (third difference)} = -\frac{4\cdot 5\cdot 6\cdot 7}{(n-1)(n-2)(n-3)}(d')$$
$$= -\frac{4\cdot 5\cdot 6\cdot 7}{9\cdot 8\cdot 7}(-3.5493) = 5.9155.$$

於是將末數及上列三個差數寫在表 12-10 的末一橫行. 將第三

【表 12-10】 用有限差數法求估計值 Y_e

X	Y_e	第一差數	第二差數	第三差數
1	48.6	−22.011	26.7587	
2	91.5	−42.854	20.8432	
3	149.3	−57.782	14.9277	
4	216.1	−66.794	9.0122	
5	285.9	−69.891	3.0967	
6	353.0	−67.072	−2.8188	
7	411.4	−58.338	−8.7343	
8	455.0	−43.688	−14.6498	
9	478.2	−23.123	−20.5653	
10	末數 474.8135	3.358	−26.4808	5.9155

差數與第二差數相加，寫在'第二差數'直行的末第二位，

$$-26.4808+5.9155=-20.5653,$$

同理，

$$-20.5653+5.9155=-14.6498.$$

這樣每次都加上一個第三差數，直至最高的一數．其次將第一差數與第二差數由下向上加，不過現在各第二差數之值是不相等了．

$$3.358+(-26.4808)=-23.123,$$

$$-23.123+(-20.5653)=-43.688,$$ 餘類推．

最後自'末數'起與第一差數，依次向上累加，卽得所求的估計值：

$$474.8135+3.358=478.172,$$

$$478.172+(-23.123)=455.049,$$

…………………………………………

將 X 與 Y_e 兩行數值作圖，則得一光滑的三次曲線(圖 12-3)．

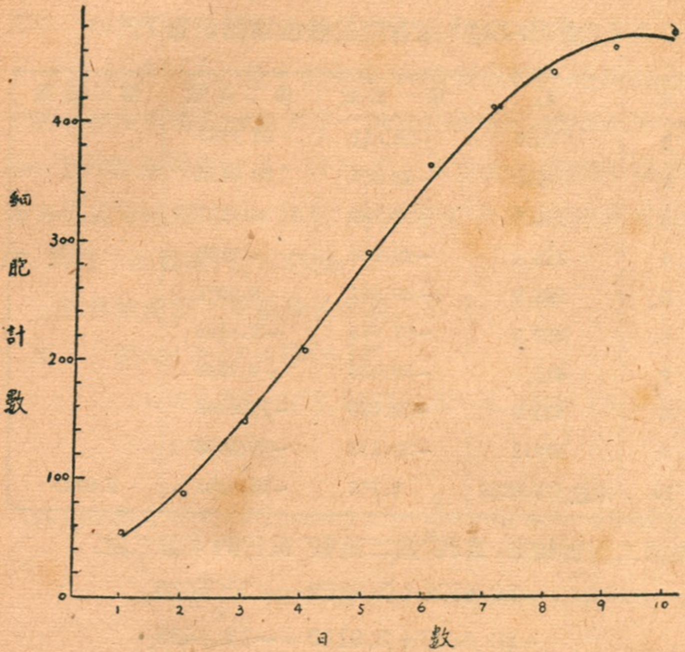


圖 12-3. 釀母之生長及其配合之三次曲線

(三)方程式的計算　前面已經說過，方程式

$$Y=A+BX_1+CX_2+DX_3+EX_4+\cdots\cdots \quad (\text{見公式 } 12\text{–}2)$$

中 $A, B, \cdots\cdots$ 爲待定的數值，就是表 12–7 末行的數值；又 $X_1, X_2, \cdots\cdots$ 等各爲 X 的多項式，其公式如下：

$$\left.\begin{aligned} X_1&=X-\bar{x} \\ X_2&=X_1^2-\frac{n^2-1}{12} \\ X_3&=X_1^3-\frac{3n^2-7}{20}(X_1) \end{aligned}\right\} \text{公式}(12\text{–}10)$$

$$\left.\begin{aligned}X_4&=X_1^4-\frac{3n^2-13}{14}(X_1^2)+\frac{3(n^2-1)(n^2-9)}{560}\\X_5&=X_1^5-\frac{5(n^2-7)}{18}(X_1^3)+\frac{15n^4-230n^2+407}{1{,}008}(X_1)\end{aligned}\right|$$

上例 $\bar{x}=5.5$, $n=10$, 代入公式, 得

$$X_1=X-5.5,$$

$$X_2=(X-5.5)^2-\frac{10^2-1}{12}=X^2-11X+22,$$

$$X_3=(X-5.5)^3-\frac{300-7}{20}(X-5.5)$$

$$=X^3-16.5X^2+76.1X+639.375.$$

我們所要的是三次方程式, 故其餘不必再計算. 最後連表 12–7 末行 A, B, C, D 等數值代入公式(12–2), 得方程式為

$$Y_e=296.38+52.8751(X-5.5)-2.88829(X^2-11X+22)$$

$$-.985917(X^3-16.5X^2+76.1X-85.8)$$

$$=26.62+9.6180X+13.37934X^2-.985917X^3.$$

將 $X=1, 2, 3, \cdots\cdots$, 10 代入, 所得各估計值應與表 12–10 的結果相同. 故若需要曲線的方程式, 則各估計值可由方程式求得, 不必再用小節(二)的方法了.

12–6. Logistic 曲線　此曲線因 Raymond Pearl 與 L. J. Reed 兩氏用以研究人口而著名, 故亦稱 Pearl-Reed 氏生長曲線. 此曲線當接近某一限度時, 其上升逐漸緩慢. Logistic 曲線的方程式可寫做下列形式:

$$y-d=\frac{K}{1+Ce^{rt}}.\qquad\text{公式(12–11)}$$

又上式可改成
$$\frac{K}{y-d}-1=Ce^{rt},$$

故該曲線方程式又可變爲
$$\log_e\frac{K-(y-d)}{y-d}=\log_e C+rt, \qquad \text{公式}(12\text{–}12)$$

此處 e 爲自然對數的基數，K 與 d 爲配合前由嘗試錯誤法擬定的數值，而 C 與 r 則爲待定的數值，t 爲時間. 該曲線上下有兩根漸近線(asymptote)，卽 $y=d+K$ 及 $y=d$. 又由公式(12–12)，可見 $\log_e[K-(y-d)]/(y-d)$ 爲 t 之直線函數. 所以在配合以前，我們要設法猜 d 與 K 的數值，儘量使 $[K-(y-d)]/(y-d)$ 的對數與 t 呈直線的關係. 又此曲線有一轉向點在原來時間 $T=-\frac{a}{r}$ 及 $y=\frac{K}{2}$ 處. 轉向點求得後，卽將曲線之原點遷至此處. 各時間與原點之距離爲 t.

茲仍以前節釀母生長之資料爲例，說明該曲線的配合法. 將原來資料在精細的方格紙上繪成圖形，藉作圖用曲線板之助，可看到該曲線之'上漸近線'大約在 $y=490$ 與 500 之間，'上漸近線'約在 $y=10$ 與 20 之間. 經幾次嘗試後，發現 $y=497$ 與 $y=12$ 兩線最爲合用，因得 $d=12$，$K=497-12=485$. 其嘗試的步驟見表 12–11. 該表的 T 爲原來時間，Y 爲細胞計數，這兩行是原來資料. 第三行 Y' 是把各 Y 值減去 12 後的簡縮數，第四行是 485 減去 Y'，第五行是前兩行的比例，卽 $(K-Y')/Y'=[K-(Y-d)]/(Y-d)$，此爲公式(12–12)等號前之分數，茲暫以符號 z 代表此比例. 若 d 與 K 值取得適當，那麼 z 的對數，與其相當的 T 值所繪成的點應該差不多在一直線上. 若不能成一直線，則另換 d 或 K 值，或兩數一同換過，總以近於

【表 12-11】 d 與 K 值之嘗試

T	Y	$Y'=Y-d$ $(d=12)$	$K-Y'$ $(K=485)$	$z=\frac{K-Y'}{Y'}$
1	52.5	40.5	444.5	10.975
2	87.8	75.8	409.2	5.398
3	147.6	135.6	349.4	2.577
4	208.6	196.6	288.4	1.467
5	289.7	277.7	207.3	.746
6	365.0	353.0	132.0	.374
7	414.3	402.3	82.7	.206
8	447.4	435.4	49.6	.114
9	469.0	457.0	28.0	.061
10	481.9	469.9	15.1	.032

直線爲止. 爲了省卻檢查對數的麻煩，並爲以後計算上的需要計，我們可用算術對數紙(arith-log paper)，這種紙的縱橫兩種尺度不同，一是對數的尺度，一是算術的尺度. 現在把表 12-11 的 T 值放在算術的尺度上，z 值放在對數的尺度上. 這裏要注意，普通對數是以十進的. 例以圖 12-4，最底下的一組數值是 .01，.02，.03，……，.10，再上去卻不是 .11，.12，……（這些數值是擠在 .10 與 .20 的短距離內的)，而是 .20，.30，.40，……，1.00，同理再向上去應爲 2.00，3.00，……. 這種規則是應用對數尺度者所必須知道的. 尺度既定，乃將各對 T 與 z 值畫成點子，如第一點爲(1，10.975)，第二點爲(2，5.398)等. 這十個點子畫好以後，可以看到它們差不多在一直線上. 於是我們設法畫一根直線，使它能夠通過大多數的點子，或各點與該直線的誤差爲最少.（可用一根細線，兩端拉緊，在所畫的點子附近嘗試，或用透明的尺、三角板等亦可. 這樣可以找到一個比較最適

宜的位置.)

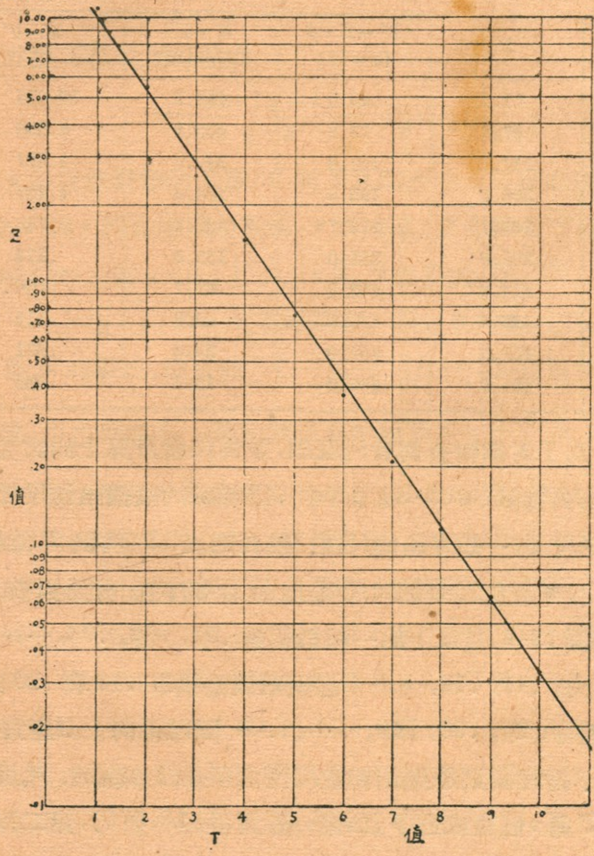


圖 12-4. 試驗值之直線性(藉知兩假定值 K 與 d 是否合用)

前面已經說過: logistic 曲線在 $T=-a/r$, $Y=K/2$ 處有一轉向點. 我們已求得 $K=485$, 則 $K/2=242.5$. 從原來資料所構成的曲線上看(或參閱圖 12-3)知道相當於 $Y=242.5$ 時的 T 值約爲 4.45.

我們就取這點(4.45, 242.5)爲原點. 各 T 值與 4.45 之距離爲 t.

現在要定該直線的坡度 m, 以及公式內 r 與 C 等的數值了. 我們在圖 12-4 的直線上任意取最方便的兩點, 一點在 $T=1$, 另一點在 $T=9$; 其相當的 t 值爲 $1-4.45=-3.45$ 及 $9-4.45=4.55$. 從圖上可以看到當 $t=-3.45$ 時, $z=10.00$, 又 $t=4.55$ 時, $z=.061$. 求兩 z 值之對數, $\log z_{-3.45}=1$, $\log z_{4.55}=\bar{2}.785,3298$.

$$m^*=\frac{\log z_{4.55}-\log z_{-3.45}}{4.55-(-3.45)}=\frac{\bar{3}.785,3298}{8}$$

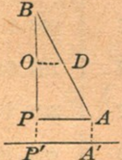
$$=\frac{-2.214,6702}{8}=-.276,833,775,$$

$$r⁑=2.30259m=-.637,4352,$$

$$^{**}\log z_0=\log z_{-3.45}-(-3.45)m$$

$$=1-.955,0765=.044,9235.$$

化成眞數 $z_0=1.108980$, 此值卽爲公式(12–12)中之 C. 將 d, K, C, r 諸值代入公式(12–12), 得所需之方程式爲



* 直線 BA 之坡度爲 $\tan A=\frac{PB}{AP}$, 試與圖 12-4 相比較, 可見 $\log z_{4.55}-\log z_{-3.45}$ 相當於 PB, 而 $4.55-(-3.45)$ 則相當於 AP.

** $\log z_0$ 相當於此圖中之 OP', 而 $OP'=BP'-BO=BP'-OD\tan\angle BDO=BP'-OD\tan A$, 與 $\log z_{-3.45}-(-3.45)m$ 相當.

⁑ 按普通對數, 以 $\log_e 10=2.30259$ 乘之, 則化爲自然對數, 故公式(12-12)中之 r 實爲圖 12-4 之縱軸化成自然對數後, 所作迴歸直線之坡度.

$$Y_e - 12 = \frac{485}{1+1.10898e^{-.637,4352t}}. \qquad \text{(L)}$$

有了方程式，就可求估計值，計算的過程詳表 12–12，其中第 3 行是 m 與 t 的乘積，如 $(-.276,833,775)(-3.45)=.955,0765$，此乘積等於 $\log e^{rt}$. 因 $r=2.30259m$（見前），故

$$mt=\frac{rt}{2.30259}=rt(0.43429)=rt\log_{10}e=\log_{10}e^{rt}.$$

又第 4 行爲相當於左行各對數之眞數. 於是將 $C=1.10898$ 乘第 4 行之值，再加 1，即爲第 5 行之 $1+Ce^{rt}$. 再將第 5 行各數除 $K=485$，得 y'，此爲第 6 行之數值. 最後在 y' 上各加 $d=12$，即得估計值 Y_e. 將 T 與 Y_e 之值繪成曲線，如圖 12–5.

【表 12–12】 由方程式(L)計算估計值 Y_e.

T (1)	t $(T-4.55)$ (2)	$\log e^{rt}=mt$ (3)	e^{rt} (4)	$1+Ce^{rt}$ (5)	$y'=\frac{K}{1+Ce^{rt}}$ (6)	$Y_e=y'+d$ (7)
1	−3.45	.955,0765	9.017,300	11.000,000	44.00	56.09
2	−2.45	.678,2427	4.766,973	6.286,476	77.15	89.15
3	−1.45	.401,4090	2.520,049	3.794,683	127.81	139.81
4	−.45	.124,5752	1.332,218	2.477,402	195.77	207.77
5	.55	$\bar{1}$.847,7414	.704,274	1.781,025	272.32	284.32
6	1.55	$\bar{1}$.570,9076	.372,312	1.412,887	343.27	355.27
7	2.55	$\bar{1}$.294,0739	.196,822	1.218,272	398.10	410.10
8	3.55	$\bar{1}$.017,2401	.104,050	1.115,389	434.83	446.83
9	4.55	$\bar{2}$.740,4063	.055,006	1.061,000	457.12	469.12
10	5.55	$\bar{2}$.463,5725	.029,079	1.0 2,247	469.85	481.85

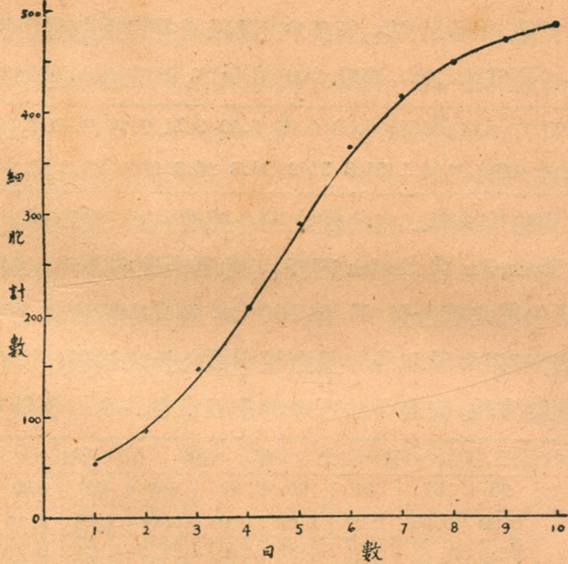


圖 12-5. 釀母之生長及其配合之曲線

【練 習 題】

1. 將 $X=1, 2, 3, \cdots\cdots, 20$ 代入下列方程式,並各繪一曲線:

(1) $Y=2.58+0.84X$, (2) $Y=2.58+8.4\log X$,

(3) $\log Y=0.258+0.058X$, (4) $\log Y=0.213+0.662\log X$.

2. 將 $X=0, 5, 10, 20, 30, 45, 60$ 代入下列方程式並繪一曲線:

$$\log Y=2.432-0.02227X.$$

3. 下表爲釀母在 30°C. 時之生長紀錄(1),試配合一對數曲線:

孵養時間(分鐘,X)	0	150	195	240	288	330
細胞計數(10^6,Y)	1.80	4.95	6.50	12.30	19.30	29.00

4．吳憲與楊恩福(6)兩氏在'以尿素引起蛋白質變性率受酸度之影響'一研究中，有一實驗之紀錄如下，試配合一拋物線：

pH (X)	6.59	6.83	6.97	7.03	7.14	7.25	7.40	7.68	7.80	7.96
蛋白質凝固百分數 (Y)	44.1	31.8	31.2	28.8	26.8	26.2	24.8	21.3	22.0	22.5

5．測驗第4題二次曲線之配合適度（參考表12-3）.

6．Swanscn 與 Smith 兩氏(7)曾測定鼠在不同年齡（X，日數）時其血漿中氮之總含量（Y，每 100 c.c. 血漿中所含氮之克數），結果載於下表．試以年齡與各年齡血漿中氮含量之均數（共有九對數值）配合一迴歸直線，配合方法可參考第五章表 5-1，但各數及其平方與

	鼠之年齡（日數）								
	25	37	50	60	80	100	130	180	360
血漿中氮之總含量（克）	0.83	0.98	1.07	1.09	0.97	1.14	1.22	1.20	1.16
	0.77	0.84	1.01	1.03	1.08	1.04	1.07	1.19	1.29
	0.88	0.99	1.06	1.06	1.16	1.00	1.09	1.33	1.25
	0.94	0.87	0.96	1.08	1.11	1.08	1.15	1.21	1.43
	0.89	0.90	0.88	0.94	1.03	0.89	1.14	1.20	1.20
	0.83	0.82	1.01	1.01	1.17	1.03	1.19	1.07	1.06
	0.84	0.95	0.93		0.98	1.08	1.19	1.13	1.29
	0.75	0.92	1.07		0.99	0.98	1.14	1.12	1.25
	0.67	0.87	1.03			0.98	1.13		1.23
	0.70		1.13			1.10	1.06		1.22
	0.77		1.05				1.11		1.17
	0.76		0.96				1.04		
	0.75		1.01				1.29		
	0.78		1.01				1.09		
	0.76		0.94				1.14		
	0.86		1.03				1.12		
	0.78		1.06				1.29		
	0.84		1.08				1.28		
	0.85		1.09				1.10		
	0.83		0.93				1.07		
	0.85		0.95				1.10		
	0.85								
	0.83								
	0.83								
	0.70								
鼠數	25	9	21	6	8	10	21	8	11

乘積，均須以該組鼠數乘之(卽加權)，卽將 $\Sigma fX, \Sigma fY, \Sigma fX^2, \Sigma fY^2$, ΣfXY 代入各公式卽可．並就已求得之離均差平方和及積和，仿公式(6-1)求得相關係數 r，再由公式(6-12)求得直線迴歸之估計誤差，九個氮含量均數之總變異與估計誤差之相差，卽爲由直線迴歸所減少之變異．最後仿表 12-8 第一欄以測驗其配合迴歸直線之適度．

7．測驗第 6 題各組均數與迴歸直線相離之變異(參閱表12-6)．並解釋第 6, 7 兩題測驗之結果．

8．下列資料表示 $pH(X)$ 與天門冬(素)酵素(asparaginase)的活潑性(activity, Y)間之關係(8)．試配合一四次曲線(正交多項式)，配合時計算至 S_5, e, e' 及 E 卽可．並測驗各階段之配合適度(表12-8)．

X	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Y	0.2	0.4	1.4	4.1	6.6	8.7	9.8	9.9	9.5	8.2	6.4	3.3	0.3	0.1

9．用有限差數法求第 8 題四次曲線之估計値 Y_e．

10．試用下列資料配合一 logistic 曲線，並繪圖．

X	1	2	3	4	5	6	7	8	9	10	11	12
Y	18	36	68	98	131	170	206	228	247	250	254	254

【參 考 文 獻】

(1) Margaret E. Greig and J. C. Hoogerheide, The correlation of bacterial growth with oxygen consumption, Journal of Bacteriology, 41 : 549-556 (1941).

(2) Running, T. R., Empirical Formulas, John Wiley and Sons, Inc., New York (1917).

(3) Hsien Wu and Schmorl M. Ling, Studies on denaturation of proteins, V. Factors controlling coagulation of proteins by shaken, Chinese Journal of Physiology, 1, 4 : 407–430 (1927).

(4) Peter Barlow, Tables of Squares, Cubes, Square Roots, Cube Roots, and Reciprocals, E. and F. N. Spon, Ltd., London (1921).

(5) Fisher, R. A., Statistical Methods for Research Workers, Oliver and Boyd, Edinburgh (1925), Section 27.

(6) Wu, H. and Yang, E. F., Studies on denaturation of proteins, XI. Effect of hydrogen ion concentration on rate of denaturation of egg albumin by urea, Chinese Journal of Physiology, 5, 4 : 301–308 (1931).

(7) Snedecor, G. W., Statistical Methods, 3rd edition, Iowa State College Press, Ames, Iowa (1940), p. 320.

(8) Geddes, W. F. and Hunter, A., Journ. Biol. Chem., 77 (1928).

第十三章　單一自由度

關於均數間或其他統計數間的相互比較，本書已討論過很多．本章對於兩均數間的比較將引申幾個新的觀念．同時對於新的計算方法及有關的實驗設計等都要提及．因爲這種比較所用的離均差平方和只有一個自由度，故本章標題爲單一自由度．

13–1．計算離均差平方和之新法　鄭集氏等在黄豆蛋白之研究中(1)，曾用白鼠三組，分别與以某種飼料，內有二組：一飼黄豆蛋白質，一飼黄豆蛋白質略加酪蛋白．下表爲該兩組白鼠在三十日內所

飼料	各鼠所增之體重							總計
黄豆蛋白質	23.5	15.9	6.6	9.6	8.4	11.4	20.4	95.8
黄豆蛋白質加酪蛋白	41.3	40.7	28.8	18.8	21.7	25.1	22.2	198.6

增之體重(克)，問兩組之均數有顯著的相差否？照以前的慣例，是用 t 值來測驗兩均數相差的顯著性，詳細過程見第四章第 4–7 節，茲將計算結果列後：

	白鼠數	自由度	均數	離均差平方和
黄豆蛋白質	7	6	13.6857	246.37
黄豆蛋白質與酪蛋白	7	6	28.3714	504.23
		12	$\bar{d}=14.6857$	750.60

$$v=750.60\div12=62.55,$$

$$S_{\bar{d}}=\sqrt{62.55}\sqrt{\frac{7+7}{7\times7}}=4.2275,$$

$$t=14.6857\div4.2275=3.4739,$$

$$\text{自由度}=12;\quad 1\%\text{點}=3.055.$$

此處所得 t 值大於 1%點，故知兩均數有非常顯著的差別.

以上所列資料，亦可以 F 值測驗之. 十四個白鼠所增體重的總和爲 294.4 克，故其總均數爲 $294.4\div14$. 又各組的均數爲 $95.8\div7$ 及 $198.6\div7$. 因得各組均數與總均數相差的平方和爲

$$\left(\frac{95.8}{7}-\frac{294.4}{14}\right)^2+\left(\frac{198.6}{7}-\frac{294.4}{14}\right)^2=107.8351,$$

再以每組鼠數 7 乘之，得

$$7(107.8351)=754.85.$$

此值表示兩均數間之變異，因其自由度爲 1，故離均差平方和卽等於組間變異數(均方). 再以組內變異數 62.55(見前)除之，卽得 F 值，

$$F=754.85\div62.55=12.068.$$

當自由度等於 1 與 12 時，F 值之 1%點爲 9.33，故結果亦爲非常顯著. 此處值得注意者，卽 $\sqrt{12.068}=3.4739$，恰等於前面求得的 t 值. 故知 F 值表(表 9–3)內第 1 直行的數值，恰爲 t 值表（表 4–4）中相當數值的平方. 當組間自由度等於 1 時，用 t 與 F 測驗的結果是一樣的.

計算兩均數間的變異數，我們曾用過一種簡便的方法(公式9–1)

$$\frac{(95.8)^2+(198.6)^2}{7}-\frac{(294.4)^2}{14}=754.85.$$

現在我們要介紹一個更簡便的算法，即將兩個總數相減的平方以每組鼠數的 2 倍除之，即得所需的離均差平方和．其公式爲

$$\frac{(a-b)^2}{2k}, \qquad \text{公式(13-1)}$$

此處 $a=95.8$, $b=198.6$, $k=7$. 代入公式，得

$$\frac{(95.8-198.6)^2}{2(7)}=754.85.$$

當祇有兩個均數的時候，用此法計算均數間的變異數，要簡便多了．此變異數的自由度爲 1．在求 F 值時，當然還需誤差的變異數，這點是和以前一樣的．讀者應該知道本章所說的顯著性測驗，在原理上，與前面各章所講的並無不同，祇是計算方法非常簡便罷了．

13-2．兩組以上之個別比較 若一個實驗含有三、四組或更多時，可作幾種個別的比較．例如第九章表 9-6 所用白兔，其中有兩組（第一與第三）注射垂體加壓劑，另一組（第二）則不曾注射．今若比較注射者與不注射者肝臟脂酸之碘指數有無顯著的差別，則可將第一、三兩組合併，與第二組相比較．我們還是先用老方法（公式9-1），再用新的簡法：

【表 13-1】 注射垂體加壓劑者與不注射者之比較

	動物數 (k)	碘指數總和	(總和)$^2/k$
注射者	20	2,155.00	232,201.25
不注射者	10	1,178.12	138,796.67
總計	30	3,333.12	370,997.92
校正數	$(3,333.12)^2/30$		370,322.96
離均差平方和			674.96
$F=674.96\div88.89=7.593$，自由度$=1$與25，			
5%點$=4.24$，1%點$=7.77$．			

求 F 時所用的誤差變異數 88.89，見表 9-8，其自由度為 25. 結果得 F 值為 7.593，此值在 5% 點與 1% 點之間，故兩組碘指數之均數有顯著的差別.

若用新法計算，則可用下列公式:

$$\frac{(a-2b+c)^2}{6k}. \qquad \text{公式(13-2)}$$

$$a=1,100, \qquad b=1,178.12, \qquad c=1,055,$$

$$\frac{[1,100-2(1,178.12)+1,055]^2}{6(10)}=\frac{(201.24)^2}{60}=674.96.$$

上面兩個公式有類似的地方，在公式(13-2)中，我們可以看到把 $a+c$ 與 b 的兩倍相比較，此與均數 $(a+c)\div 2$ 減去 b 的意義是一樣的. 至於分母部分 k 前面的係數，其計算方法有一簡便的規則，即為分子內各項係數的平方和. 例如 $a-b$ 各項係數之平方和為 $1^2+(-1)^2=2$; 又 $a-2b+c$ 各項係數之平方和則為 $1^2+(-2)^2+1^2=6$. 此兩數值恰為公式(13-1)及(13-2)中分母 k 之係數.

茲再以第十章第 3 題之資料為例，實驗時的情境，計有休息、讀故事、恐懼及看電影四種. 在此四種情境下，被試者每分鐘平均呼吸次數為: 17.83, 17.30, 20.20 及 18.30. 其中以恐懼時之呼吸速率為最快. 現在要比較恐懼與其他情境之下之呼吸速率有無差別，則可用下列公式:

$$\frac{(a+b-3c+d)^2}{12k}. \qquad \text{公式(13-3)}$$

分子部分的符號更可簡寫為

$$1 + 1 - 3 + 1,$$

各數平方之和爲 $1^2+1^2+(-3)^2+1^2=12$, 此卽分母方面 k 之係數. 又此實驗所用被試者各組都有 30 人, 故 $k=30$. 因得兩均數間之變異數爲

$$\frac{[535+519-3(606)+549]^2}{12(30)}=\frac{(-215)^2}{360}=128.40.$$

又組內變異數(誤差)爲 $2,123.57\div116=18.31$〔註: 呼吸速率之組間離均差平方和爲 143.42, 組內離均差平方和爲 2,123.57, 其相當之自由度爲 3 與 116〕. 故

$$F=128.40\div18.31=7.013,$$

查 F 値表, 當自由度等於 1 與 116 時, F 値之 1% 點約爲 6.84, 今求得之 F 値大於此數, 故知當恐懼時與在其他情境下之呼吸速率有非常顯著的差別.

前節與本節所舉的三個例題, 我們可見一相同之點, 卽在同一實驗中各組所用被試者的數目(k)必須相等, 纔能用本章的新方法來計算. 本書第一章第 1-3 節曾提到 '相互比較的各組所用被試者的數目應相等', 讀者看了本章將益感此種需要.

13-3. 均衡的比較 (orthogonal comparisons) R. A. Fisher 曾謂: 每一自由度, 可有一相當的離均差平方和, 而包括 p 種獨立比較的總離均差平方和則可析爲 p 個部分, 每個部分可用以測驗其一種比較的顯著性.

表 9-6 白兎肝臟脂酸碘指數之研究, 除上節注射垂體加壓劑與不注射者之比較外, 又可將注射垂體加壓劑者(第一組)與兼食氯化膽素者(第三組)相比較. 以符號表示之, 則爲

比較組別	第一組	第二組	第三組
一、三與二	+1	−2	+1
一與三	+1	0	−1

這裏請注意兩點：(1)每種比較(卽各橫行)的係數之和爲0.(2)兩種比較的相當係數乘積之和爲 0，如

$$(1)(1)+(-2)(0)+(1)(-1)=0.$$

前者與離均差之總和爲零(公式 3-1)相似，後者則與共變數等於零相似，共變數等於零卽表示兩種比較之間無相關之存在，因此稱爲正交的或均衡的比較(參閱公式 6-2 及前章第 12-5 節首段).

關於第一、三兩組與第二組之比較，已見前節. 茲再續作第一組與第三組之比較.

$$\frac{(1,100-1,055)^2}{(1^2+1^2)10}=101.25,$$

$$F=101.25\div 88.89=1.139,$$

$$\text{自由度}=1\text{ 與 }25\text{ 時}, \qquad 5\%\text{點}=4.24.$$

所以知該兩組碘指數之均數並無顯著的差別. 最後請讀者再注意一點，卽此處之離均差平方和 101.25，與上節的 674.96 相加，

$$101.25+674.96=776.21,$$

結果恰與表 9-7 組間變異之離均差平方和相等. 這點就說明了本節開頭的幾句話：'包括 p 種獨立比較的總離均差平方和，可析爲 p 個部分，每個部分可用以測驗其一種比較的顯著性'.

茲再用表 9-9 六組公鼠所增體重的資料，來說明多種的獨立比較. 組間的自由度爲 5，其離均差平方和爲 4,612.93(表 10-5). 茲將

此5個自由度析爲5種獨立的比較:

組別	1	2	3	4	5	6
高蛋白與低蛋白	1	+1	+1	−1	−1	−1
牛肉與豬肉,高	1	0	−1	0	0	0
肉類與穀類,高	1	−2	+1	0	0	0
牛肉與豬肉,低	0	0	0	+1	0	−1
肉類與穀類,低	0	0	0	+1	−2	+1

在此仍請注意兩點: (1)每種比較內各係數之和爲零; (2)在五種比較中,每次任取兩種,則共有10對,在此10對中之任何一對,其相當係數乘積之和爲零.

用每組所增體重之總和求得五種比較之均方(每種比較之自由度爲1,故離均差平方和與均方相等)如下:

$$\frac{(1,000+859+995-792-839-787)^2}{(1^2+1^2+1^2+1^2+1^2+1^2)10}=3,168.27$$

$$\frac{(1,000-995)^2}{(1^2+1^2)10}=1.25$$

$$\frac{[1,000-2(859)+995]^2}{(1^2+2^2+1^2)10}=1,278.82$$

$$\frac{(792-787)^2}{(1^2+1^2)10}=1.25$$

$$\frac{[792-2(839)+787]^2}{(1^2+2^2+1^2)10}=163.35$$

組間離均差平方和 $=4,612.94$

五個自由度之離均差平方和,其總計與表10–5所得者相同.若各鼠之食量略而不計,則所增體重之誤差爲 $11,586\div 54=214.6$ (參閱

表 10-5). 因得比較高低蛋白之 F 值爲 $3,168.27 \div 214.6 = 14.8$, 此值爲非常顯著; 又高蛋白之肉類與穀類相比較, 其 F 值爲 $1,278.82 \div 214.6 = 5.96$, 此值爲顯著. 其餘三種比較之均方皆比誤差之均方爲小, 故爲不顯著. 這樣用單一自由度的比較, 對於營養學家或許是很有用的. 同時我們也可以看到, 如果一個實驗設計得良好, 就可得到更多的有價值的知識而增加實驗的效能.

個別比較的數目, 在理論上是無限制的. 如包括六組的實驗, 若任取兩組, 則有15種比較之多(其中有一部分是不合於均衡原則的). 但在實用上, 則比較必須有意義, 而且在設計時, 就應該考慮到實驗結果如何用統計方法去處理, 這點是必須注意的.

13-4. 不均衡的比較(non-orthogonal comparisons) 有許多實驗的設計, 其所用獨立的比較是不均衡的. 這種設計只要能達到實驗的目的, 也不能認爲一種缺點. 不過這時幾種比較的離均差平方和, 其總計並不等於組間的離均差平方和, 這點是與上節不同的. 看了下面的例題, 就可明瞭.

第十章第 3 題的資料, 又可將休息時之呼吸速率與其他三種情形下分別比較, 如

休息與讀故事 $(535-519)^2 \div 2(30) = 4.27$,

休息與恐懼時 $(535-606)^2 \div 2(30) = 84.02$,

休息與看電影 $(535-549)^2 \div 2(30) = 3.27$.

誤差變異數仍爲 18.31(見第 13-2 節), 上列三個變異數祇有休息與恐懼時爲顯著, 其 $F = 84.02 \div 18.31 = 4.589$. 其餘兩種比較皆爲不顯著. 此三種比較, 若以符號表示之, 則爲:

實驗號數	1	2	3	16
休息與讀故事	1	-1	0	0
休息與恐懼時	1	0	-1	0
休息與看電影	1	0	0	-1

這裏每種比較的係數之和雖等於零,但兩種比較中相當係數乘積之和卻並不等於零,如

$$(1)(1)+(-1)(0)+(0)(-1)+(0)(0)=1.$$

又三種比較之離均差平方和共為 $4.27+84.02+3.27=91.56$,此值較組間離均差平方和 143.42 為小.

凡一個實驗包括若干組,內有一組為對照組,其他各組逐一與之比較者,皆可應用本節不均衡的比較法.此時可應用第四章第 4–5 節可信限的原理求出兩均數間最小而顯著的相差,則比較時可更為簡便.查 t 值表(表 4–4),當自由度等於 116(上例誤差之自由度)時,5% 點約為 1.981(用內插法).茲以 d 代兩組總數之相差,按本章第 13–1 節,當自由度 $n_1=1$ 時,F 值恰為相當 t 值的平方,因得

$$\sqrt{[(d)^2\div 2(30)]\div 18.31}=1.981,$$

$$\sqrt{[(d)^2\div (30)^2]\div 2(18.31)}=1.981\div\sqrt{30},$$

$$d\div 30=1.981\sqrt{2(18.31)\div 30}=2.189.$$

因各組人數皆為 30,故 $d\div 30$ 即等於兩均數之相差.凡他組均數與休息組均數 17.833 之相差在 2.189 以上者皆為顯著.易言之,凡均數之小於

$$17.833-2.189=15.644,$$

或大於

$$17.833+2.189=20.022$$

者皆爲顯著. 試檢視其他三組的均數僅恐懼時(20.20)大於 20.022, 故與休息時有顯著的差別, 餘二組則否.

13–5. 迴歸中之個別比較 有些資料是每隔若干時間的觀察紀錄, 或爲根據量的水準(如 pH)而分組的. 在此可以研究組均數($\bar{x}_i, \bar{y}_i$)間是否有迴歸的影響. 當一個實驗中包含三組時, 我們可用下述兩種獨立的比較以測量(1)直線迴歸, (2)第二個均數與直線迴歸的距離. 若第二個均數與直線迴歸有顯著的差別, 則可有兩種解釋: 如果第一、二兩組間與第二、三兩組間的距離已知其相等, 則表示有遞減或累進律(即曲線迴歸)之現象. 若假定迴歸爲直線的, 則表示各組間之距離不等.

有兔六頭, 於注射腎上腺素(adrenaline)後, 每隔一小時測量其血糖(每 100 立方厘米血液所含糖之毫克數)結果如下表(2):

【表 13–2】 兔被注射腎上腺素後之血糖

注射後之時間(小時)			總計
1	2	3	
380	501	532	1,413
203	348	348	899
400	460	540	1.400
342	410	345	1,097
430	508	571	1,509
396	544	597	1,537
2,151	2,771	2,933	7,855

【表 13-3】 變異數分析(兔被注射腎上腺素後之血糖)

變異來源	自由度	離均差平方和	均　　方	F
總　變　異	17	179,462.28		
時　　間	2	56,787.11	28,393.56	19.40
兎　　間	5	108,041.61	21,608.32	14.77
誤　　差	10	14,633.56	1,463.36	

我們先用第九章第9-3節的方法作變異數分析(表 13-3),求得誤差之均方爲 1,463.36. 由表 13-2,可見在注射腎上腺素後之最初三小時內,血糖數隨時間而增高. 現在我們要問:(1)血糖與時間的迴歸是否顯著,又(2)注射後第 2 小時之血糖數是否與直線迴歸有顯著的差別. 前者可用 $(-1, 0, +1)$ 的比較來解答,後者則可用$(1, -2, +1)$的比較來解答(詳細說明見次節). 茲將兩種比較均方及 F 值列後:

$$\frac{(-2,151+2,933)^2}{(1^2+1^2)6}=50,960.33,$$

$$F=\frac{50,960.33}{1,463.36}=34.82;$$

$$\frac{(2,151-2\times 2,771+2,933)^2}{(1^2+2^2+1^2)6}=5,826.77,$$

$$F=\frac{5,826.77}{1,463.36}=3.98;$$

自由度=1 與 10, 5%點=4.96, 1%點=10.04.

第一個 F 值 34.82 爲非常顯著,這表示血糖與時間間的迴歸爲非常顯著. 而第二個 F 值 3.98 則爲不顯著,這表示注射後第 2 小時之血糖均數與直線迴歸無甚差別. 這是意料中事,因爲前者之離均差平

方和 50,960.33 占去了時間方面離均差平方和(表 13-3) 的最大部分，餘下的一個自由度便不顯著了. 上列兩種比較的離均差平方和總計與表 13-3 時間方面的平方和相等，

$$50,960.33+5,826.77=56,781.10.$$

此例時間方面兩個自由度的均方爲非常顯著，而直線迴歸亦然. 但在其他例子中，兩個自由度的均方可爲不顯著，而在個別比較時，可能有一個自由度是顯著的.

關於迴歸中之個別比較，下節將作更詳細的討論. 但討論時，自變數之距離以相等者爲限.

13-6. 正交多項式與迴歸中之個別比較 前章第12-5 節，正交多項式的配合法與個別比較是有密切關係的. 我們還記得每配合一較高次的曲線，則離均差平方和更減少一些，而當時所用的許多公式便是從這樣減少的離均差平方和得來的. 下面我們要把這些公式引申出來.

令 x 爲自變數，y 爲因變數；當 x 等於某兩數值時，y 之値爲 u 與 v. 按第 12-5 節的方法，可計算如下:

【表 13-4】 正交多項式中公式之說明(一)

y 之值	第二總和
u	u
v	$u+v$
$S_1=u+v,$	$S_2=2u+v;$
$a=\dfrac{u+v}{2},$	$b=\dfrac{2u+v}{3};$
$a'=\dfrac{u+v}{2},$	$b'=\dfrac{u+v}{2}-\dfrac{2u+v}{3}=\dfrac{-u+v}{6};$
$A=\dfrac{u+v}{2},$	$B=-u+v.$

又由直線迴歸而減少的離均差平方和為

$$\frac{B^2}{2}=\frac{(-u+v)^2}{2} \quad 或 \quad \frac{(-1+1)^2}{1^2+1^2}.$$

後者為本章所用的新的符號. 於是正交多項式的幾組公式增加了新的意義, 但其用途卻毫無改變.

當 y 有 u, v 與 w 三個數值, 且其相當的三個 x 值距離相等時, 則配合正交多項式時之計算如下表:

【表 13-5】 正交多項式中公式之說明(二)

y 之值	第二總和	第三總和
u	u	u
v	$u+v$	$2u+v$
w	$u+v+w$	$3u+2v+w$
$S_1=u+v+w,$	$S_2=3u+2v+w,$	$S_3=6u+3v+w;$
$a=\frac{u+v+w}{3},$	$b=\frac{3u+2v+w}{6},$	$c=\frac{6u+3v+w}{10};$
$a'=\frac{u+v+w}{3},$	$b'=\frac{-u+w}{6},$	$c'=\frac{u-2v+w}{30};$
$A=\frac{u+v+w}{3},$	$B=\frac{-u+w}{2},$	$C=\frac{u-2v+w}{2}.$

由直線迴歸所減少之離均差平方和為

$$\frac{(-u+w)^2}{2} \quad 或 \quad \frac{(-1+0+1)^2}{1^2+0^2+1^2},$$

而由二次曲線迴歸所減少者則為

$$\frac{(u-2v+w)^2}{6} \quad 或 \quad \frac{(1-2+1)^2}{1^2+2^2+1^2}.$$

若每個 y 值為 k 次測定的總和, 則每種比較所用公式的分母都要乘以 k, 像前面幾節裏的公式那樣. 這兩種比較的離均差平方和等於 y

三個數值的總變異，因爲拋物綫能通過任何三點，而沒有估計誤差的. 所以這一對的個別比較供給我們關於三組資料所具迴歸的全部知識. 我們在前節(第 13-5 節)中的計算，至此獲得充分的意義.

茲將包含 2, 3, 4, 5 各組及配合 1, 2, 3, 4 各次多項式時迴歸中所用個別比較列後. 例如含有五組之資料，當配合二次正交多項式時，其迴歸之個別比較爲(2-1-2-1+2). 餘類推.

【表 13-6】 包含 2, 3, 4, 5 各組(距離相等)之獨立比較中所用係數

多項式次數	組數			
	2	3	4	5
1	-1+1	-1+0+1	-3-1+1+3	-2-1+0+1+2
2		1-2+1	1-1-1+1	2-1-2-1+2
3			-1+3-3+1	-1+2+0-2+1
4				1-4+6-4+1

用了這些公式，計算的手續要簡便得多. 至於比四次更高的多項式，其所需的各套係數當然也可依法推廣. 但事實上一個實驗中的處理水準很少用到五個以上的，故從略. 又在不常見的多組相比較時，須應用本書前面的普通迴歸與共變數分析法.

13-7. 迴轉實驗(reversal experiments) 若有兩種處理方法對於兩組動物(或人類)交互應用，則此類實驗稱爲迴轉實驗. 迴轉實驗的結果，可適用本章個別比較的方法(3). 金鑄、朴杜秉二氏曾作麻黃素(ephedrine)、假性麻黃素(pseudoephedrine)與麻黃甲素(ephetonin) 三者之比較研究(4)，其中關於麻黃素與麻黃甲素之部分，計有犬六頭，等分爲甲、乙兩組，甲組每犬先注射麻黃素，次注射麻黃

甲素，最後再注射麻黃素；乙組則最先與最後皆注射麻黃甲素，而中間一次則注射麻黃素. 茲將注射後所增加之血壓(指最大作用 maximal effect 時)及個別比較之計算列後:

【表 13-7】 犬注射麻黃素及麻黃甲素後增加之血壓

組別	注射次序			比較	總計
	a	b	c	$a-2b+c$	
甲	麻黃素	麻黃甲素	麻黃素		
	30	10	18	28	
	34	26	29	11	
	43	22	22	21	60
乙	麻黃甲素	麻黃素	麻黃甲素		
	24	40	28	−28	
	16	18	16	−4	
	14	14	22	8	−24

每犬既分三次注射，則按第 13–5 與 13–6 兩節的原理，由

$$a+0-c$$

與

$$a-2b+c$$

兩種比較，可獲得全部的知識. 前者$(a+0-c)$測量注射次序與所增血壓之間有無關係(直線迴歸)存在，而非比較兩種藥物之效能，故可不必討論. 後者$(a-2b+c)$則比較第二期與直線迴歸之距離，卽爲本實驗所欲研究者. 各犬$(a-2b+c)$之值見表 13–7, 甲組之總和爲 60, 係(麻黃素)−(麻黃甲素)；乙組之總和爲 −24, 則係(麻黃甲素)−(麻黃素). 故二值之相差,

$$60-(-24)=84,$$

表示麻黃素之效能大於麻黃甲素. 顯著性測驗時所用之誤差可用兩

組合併的變異數，其自由度爲 $2+2=4$(參閱第四章第4–7節)，

$$(28)^2+(11)^2+(21)^2-(60)^2/3=146,$$

$$(-28)^2+(-4)^2+(8)^2-(-24)^2/3=672,$$

$$\frac{146+672}{2+2}=204.5.$$

麻黄素與麻黄甲素兩均數間之變異數爲

$$\frac{[60-(-24)]^2}{2(3)}=1,176,$$

因得
$$F=1,176\div 204.5=5.75,$$

$$\text{自由度}=1\text{ 與 }4,\quad 5\%\text{點}=7.71.$$

結果麻黄素與麻黄甲素在最大作用時所增血壓並無顯著的差別．至於在持久性方面，其情形如何，則爲另一問題．

此種迴轉實驗的設計在營養或其他研究方面都可適用．且不限於三個時期，例如有 A, B 兩種處理分成四個時期輪流施用，則甲組各期爲 A, B, A, B; 而乙組爲 B, A, B, A. 但須注意各期所經過的時間必須相等．又比較時須用表 13–6 四組三次多項式之係數，卽仿表 13–7 計算時，其右起第二直行應爲$(-a+3b-3c+d)$，餘可類推．

13–8．析因實驗(factorial experiments)　析因實驗的設計，是企圖研究種種處理的因子，以增進實驗的效能，並擴大其範圍．例如營養實驗中，要同時作飼料間質與量的比較研究，那麼質、量兩因子就要用兩種或幾種水準的所有組合．在有些實驗中也常把幾個因子同時研究．下列資料爲一藥理的實驗(5)．將一白色來航公雞(leghorn rooster)麻醉後，以垂體後葉溶液 (posterior pituitary solution) 施

行靜脈注射，然後按 Coon 氏方法(6)記錄其血壓. 所用藥劑有兩種，一為 U.S.P. 標準垂體後葉溶液，一為待試驗之未知溶液. 用藥量分為高低兩種，高者為低者之兩倍. 於是藥劑與用量共有四種組合. 低量標準溶液之符號為 S_1，高量為 S_2；未知溶液之低量為 U_1，高量為 U_2. 每種注射四次，共 16 次. 兩次注射間相隔十分鐘. 各種注射之排列採用隨機原則，且每一橫行與每一直行無一重複之處理. 例如

$$
\text{甲}\quad \begin{array}{cccc} S_1 & S_2 & U_2 & U_1 \\ S_2 & U_2 & U_1 & S_1 \\ U_1 & S_1 & S_2 & U_2 \\ U_2 & U_1 & S_1 & S_2 \end{array}
\qquad\qquad
\text{乙}\quad \begin{array}{cccc} U_2 & U_1 & S_2 & S_1 \\ S_1 & U_2 & U_1 & S_2 \\ U_1 & S_2 & S_1 & U_2 \\ S_2 & S_1 & U_2 & U_1 \end{array}
$$

甲為原著者所用之排列次序. 其實不限於此種形式，依照上述原則，可另列形式如乙. 是項設計稱為拉丁方 (Latin square) 與第九章第 9–3 節之隨機區組最初同為農業上之田間設計，而今則已廣為應用，並不限於田間了.

按拉丁方可用任何幾種處理，每橫行與直行的項目數應與處理數相同. 此例用兩種溶液與兩個水準，二者之組合相當於四種處理，這是析因實驗中的拉丁方.

實驗結果見表 13–8.

【表13–8】 來航雞注射垂體後葉溶液後下降之血壓(毫米，水銀柱)

組 別	S_1	S_2	U_1	U_2	總 和
1	18	34	25	36	113
2	12	35	18	32	97
3	11	30	17	36	94
4	16	28	19	40	103
總 和	57	127	79	144	407

拉丁方的變異數分析，本須將總變異析爲四個部分，卽橫行、直行、處理及誤差四種. 在本例中，這四部分的自由度爲 3, 3, 3, 6(參閱表 13–10). 表 13–10 的直行相當於組別，但橫行間的變異在本問題中沒有多大意義，故將橫行與誤差兩部分合併爲'實驗誤差'，而仿隨機區組的原則(第九章第 9–3 節)作變異數分析如下:

【表 13–9】 變異數分析(來航雞降低之血壓)

變異來源	自由度	離均差平方和	均方
總變異	15	1,371.94	
處理	3	1,235.69	411.90
組別	3	52.69	17.56
實驗誤差	9	83.56	9.284

繼作析因實驗之計算，以求估計的藥力(estimated potency) 及其標準誤. 此處的獨立比較計有三種(同時參閱所附四格表):

	S_1	S_2	U_1	U_2
S 與 U 間	−1	−1	+1	+1
高與低間	−1	+1	−1	+1
交互影響	+1	−1	−1	+1

	低 1	高 2	
S	57	127	184
U	79	144	223
	136	271	

因得各種比較之均方爲

$$S\text{ 與 }U\text{ 間}\quad \frac{(-57-127+79+144)^2}{4(1^2+1^2+1^2+1^2)}=\frac{(39)^2}{16}=95.06,$$

$$\sqrt{95.06}=9.75=D;$$

$$\text{高與低間}\quad \frac{(-57+127-79+144)^2}{4(1^2+1^2+1^2+1^2)}=\frac{(135)^2}{16}=1,139.06,$$

$$\sqrt{1,139.06}=33.75=B;$$

交互影響 $\dfrac{(57-127-79+144)^2}{4(1^2+1^2+1^2+1^2)}=\dfrac{(-5)^2}{16}=0.31.$

比較時所用的誤差爲 9.284(表 13-9),於是得 F 值如下:

S 與 U 間 $F=95.06\div 9.284=10.24,$

高與低間 $F=1,139.06\div 9.284=122.69,$

自由度 $=1$ 與 9 時, 5%點 $=5.12$, 1%點 $=10.56.$

前者僅略低於 1%點,爲顯著;後者遠過於 1%點,爲非常顯著.至交互影響之均方小於誤差均方,故爲不顯著.不過這裏的目的,是在求得估計藥力及其標準誤,故需要下面的計算.

$$估計藥力之對數=M=\frac{(\log 2)D}{B}=\frac{(.3010)(9.75)}{33.75}=.0870,$$

式中 D 與 B 是前面算出來的,又藥量高低之比爲 2,故分子須用 $\log 2$ 乘之.於是得估計藥力爲

$$估計藥力=\text{antilog}\,.0870=1.22 \quad 或 \quad 122\%.$$

估計藥力已經求得,現在要計算它的標準誤了.其公式爲

$$S_M=\frac{S(\log n)\sqrt{B^2+D^2}}{B^2}, \qquad 公式(13\text{–}4)$$

此處 S 爲實驗所得的標準誤差,卽

$$S=\sqrt{9.284}=3.047,$$

又 $$B^2=1,139.06, \qquad D^2=95.06,$$

$\log n$ 爲藥量高低比例之對數,此處 $n=2$.將各數代入公式,得

$$S_M=\frac{3.047(\log 2)\sqrt{1,139.06+95.06}}{1,139.06}=0.0283.$$

由此化爲相對藥力(relative potency)之標準誤，令其符號爲 s.e.，則

$$\text{s.e.}=2.303S_M(\text{antilog}\,M). \qquad \text{公式(13-5)}$$

式中 2.303 爲由普通對數化成自然對數時所乘之常數. 於是

$$\text{s.e.}=2.303(.0283)(1.22)=.0795 \quad \text{或} \quad 8\%.$$

下面我們已求得估計藥力爲 122%，其標準誤爲 8%. 估計藥力上下兩個標準誤的距離爲 $122\pm2(8)=106\rightarrow138$. 若另用生物方法(biological method)求得實際藥力爲 117.6，此數在兩個標準誤的範圍內，故知實際藥力與估計藥力並無顯著的區別.

在拉丁方中遇有缺項時，則用下列公式填補之:

$$X=\frac{k(R+C+T)-2S}{(k-1)(k-2)}, \qquad \text{公式(13-6)}$$

此處 k＝横行數(或直行數，或處理數，因三者相等)，

R＝與缺項同一横行之各項目總和，

C＝與缺項同一直行之各項目總和，

T＝與缺項同一處理之各項目總和，

S＝所有觀察項目之總和.

例如表 13-8 的資料按原著者之排列次序，則成表 13-10 的形式. 假

【表 13-10】 拉丁方之實驗紀錄

組		別	
1	2	3	4
S_1 : 18	S_2 : 35	U_2 : 36	U_1 : 19
S_2 : 34	U_2 : 32	U_1 : 17	S_1 : 16
U_1 : 25	S_1 : 12	S_2 : 30	U_2 : 40
U_2 : 36	U_1 : 18	S_1 : 11	S_2 : 28

定第 3 組之 U_1 遺失，則

$$k=4,\quad R=82,\quad C=77,\quad T=62,\quad S=390,$$

代入公式，得 $$X=\frac{4(82+77+62)-2(390)}{(4-1)(4-2)}=17.3.$$

拉丁方中填入缺項後，在變異數分析時需要校正，若祇缺一項，則校正方法比較簡單，否則就非常複雜了. 詳見 Goulden 及 Yates 等著作(第九章參考文獻 9 及 13).

關於析因實驗及其他更複雜的實驗近年來用途日廣，進展極速. 將來在醫學研究方面占一重要地位，自不待言. 本章僅作初步之介紹，讀者欲知其詳，可參考 Fisher, Goulden, 與 Yates 等著作，書名見參考文獻(7)—(11)，不贅.

【練　習　題】

1. 設有兩數為 27 與 37, $k=1$. 試用第 13–1 節所述三種方法，求其均方.

2. F 值表(表 9–3)內第 1 直行的數值，恰為 t 值表(表 4–4)中相當數值的平方. 試在 F 值表第 1 行中任選十個數值，以說明此定理.

3. 設 $a=20$, $b=30$, $c=35$, $d=40$, 各數皆為 5 個項目之總和. 試求下列兩種比較之均方:

(1) $a+b-c-d$, (2) $3a-b-c-d$.

4. 設 $a=12$, $b=18$, $c=20$, $d=30$, $e=32$, 每一數值為 4 個項目之總和. 試求下列兩種比較之均方:

(1) $4a-b-c-d-e$,　　(2) $2a+2b+2c-3d-3e$.

5．第 3,4 兩題各種比較是否爲均衡的比較，並說明其理由．

6．試以代數方法引申下列公式（參閱第 13-2 節）

$$\frac{(2a-b-c)^2}{6k}.$$

7．假定有 A, B, C, D 四種飼料，其中 B 爲一標準飼料．實驗用動物共有 40 頭，先將原始體重之相等或甚相近者，每 4 頭集於一組，共得 10 組．然後用隨機方法，將每組中之四個動物分配於四種飼料之下，每動物各得一種．分配完畢，則每種飼料各有 10 個動物．經過一定時間後，量其所增體重，實驗結果如下表．

各動物在同一期間內所增之體重

按原始體重所分之組別	飼料 A	飼料 B	飼料 C	飼料 D	總計
1	1.40	1.31	1.40	1.96	6.07
2	1.79	1.30	1.47	1.77	6.33
3	1.72	1.21	1.37	1.62	5.92
4	1.47	1.03	1.15	1.76	5.46
5	1.26	1.45	1.22	1.88	5.81
6	1.28	0.95	1.48	1.50	5.21
7	1.34	1.26	1.31	1.60	5.51
8	1.55	1.14	1.27	1.49	5.45
9	1.57	1.25	1.22	1.77	5.81
10	1.26	1.00	1.36	1.27	4.89
總計	14.14	11.95	13.25	16.62	56.46

試先用變異數分析法求得誤差之均方，再作下列比較，並解釋統計結果之意義：

比較	飼		料	
	A	B	C	D
1	−1	+1	0	0
2	0	+1	−1	0
3	0	+1	0	−1

8. 用前章配合正交多項式的方法，演化表 13-6 中包含 4 組及 5 組之各套獨立比較.

9. 下列數值爲一動物在四個相等期間內之體重(磅):

$$3.20,\quad 3.29,\quad 3.43,\quad 3.50.$$

先將四點畫出，並猜度在三種獨立比較中，何者之離均差平方和爲最大. 然後計算三種獨立比較之均方，並用通常求 $\Sigma(X-\bar{x})^2$ 的方法，以核對三種比較之總和.

10. Cannon 氏(12)等用兩種方法使乳牛飲水，一在牛舍之內(I)，一在牛舍之外(O)，視其對於奶油產量有無影響而定. 實驗時間共分四期. 下表爲兩組牛(每組 5 頭)在各期中奶油之產量(磅).

組別	時		期		組別	時		期	
	a	b	c	d		a	b	c	d
甲	(I)	(O)	(I)	(O)	乙	(O)	(I)	(O)	(I)
	40	32	28	14		51	55	40	40
	27	19	20	14		32	32	32	28
	35	27	31	25		36	38	29	31
	30	28	27	26		32	30	25	25
	37	30	32	27		21	22	15	15

先求各牛四期之比較 $(-a+3b-3c+d)$，再求甲乙兩組之總計(參閱表 13-7). 最後用 F 值測驗兩種飲水法相差之顯著性.

11. 試根據'隨機區組'、'拉丁方'、'迴轉實驗'與'析因實驗'之原理，各擬一醫學或生物科學上之實驗設計.

【參 考 文 獻】

(1) Tai, T. K. and Cheng, L. T., Studies on soy bean proteins, IV. The supplementary effect of albumin and casein on the nutritive value of glycinin, Journal of Chinese Chemical Society, 11, 1 (1944).

(2) Bachmann, C. and Toby, G., The responses of normal and hypophysectomized rabbits to adrenaline, The Journal of Physiology, 87 : 1–10, the Cambridge University Press (1936).

(3) Brandt, A. E., Iowa Agricultural Experiment Station Research Bulletin, 234 (1938).

(4) King, T. and Pak, C., Comparative studies of ephedrine, racemic ephedrine and pseudoephedrine, III. Effects on the nasal mucous membrances, Chinese Journal of Physiology, 3, 1 : 95–108 (1929).

(5) R. Blackwell Smith, The biological assay of posterior pituitary solution, The Journal of Pharmacology and Experimental Therapeutics, 78, 1 : 72 (1943).

(6) Coon, J. M. Arch., Internat. Pharmacodyn, 62 : 79 (1939).

(7) Fisher, R. A., Design of Experiments, Oliver and Boyd,

Edinburgh (1936).

(8) Goulden, C. H., Methods of Statistical Analysis, Burgess Publishing Co., Minneapolis (1939).

(9) Goulden, C. H., Modern methods of testing a large number of varieties, Dominion of Canada Department of Agriculture, Technical Bulletin 9 (1937).

(10) Yates, F., The design and analysis of factorial experiments, Imperial Bureau of Soil Science, Harpenden (1937).

(11) Forster, H. C., Iowa Agricultural Experiment Station Research Bulletin 226 (1937).

(12) Cannon, C. Y., Hansen, E. N. and James, R. O'Neal, Iowa Agricultural Experiment Station Research Bulletin 292 (1932).

公 式 彙 錄

公式號碼	公式
(2-1)	$\bar{x}=\frac{\Sigma X}{N}$.
(2-2)	$\bar{x}=G+\frac{\Sigma(X-G)}{N}$.
(2-3)	$\bar{x}=G+\frac{\Sigma fd}{N}(i)$.
(2-4)	$md=L+\frac{\frac{N}{2}-A_1}{f_{md}}(i)$.
(2-5)	$md=U-\frac{\frac{N}{2}-A_2}{f_{md}}(i)$.
(3-1)	$\Sigma(X-\bar{x})=0$.
(3-2)	$V=\frac{(X-\bar{x})^2}{N-1}$.
(3-3)	$S=\sqrt{\frac{\Sigma(X-\bar{x})^2}{N-1}}$.
(3-4)	$S=\sqrt{\frac{\Sigma X^2-\frac{(\Sigma X)^2}{N}}{N-1}}$.

(3–5) $$\Sigma(X-\bar{x})^2=\Sigma X^2-\frac{(\Sigma X)^2}{N}.$$

(3–6) $$\Sigma(X-\bar{x})^2=\Sigma(X-G)^2-\frac{[\Sigma(X-G)]^2}{N}.$$

(3–7) $$S=\sqrt{\frac{\Sigma fd^2-\frac{(\Sigma fd)^2}{N}}{N-1}}(i).$$

(3–8) $$V=\frac{\left[\Sigma fd^2-\frac{(\Sigma fd)^2}{N}\right](i)^2}{N-1}.$$

(3–9) $$\text{C.V.}=\frac{S}{\bar{x}}\times 100.$$

(4–1) $$S_{\bar{x}}=\frac{S}{\sqrt{N}}=\sqrt{\frac{\Sigma(X-\bar{x})^2}{N(N-1)}}.$$

(4–2) $$V_{\bar{x}}=\frac{V}{N}=\frac{\Sigma(X-\bar{x})^2}{N(N-1)}.$$

(4–3) $$t=\frac{\bar{x}-m}{S_{\bar{x}}}.$$

(4–4) $$m=\bar{x}\pm S_{\bar{x}}t.$$

(4–5) $$S=\sqrt{\frac{\Sigma(X_1-\bar{x}_1)^2+\Sigma(X_2-\bar{x}_2)^2}{N_1+N_2-2}}.$$

(4–6) $$S_{\bar{x}_1-\bar{x}_2}=S\sqrt{\frac{N_1+N_2}{N_1N_2}}.$$

(4–7) $$t=\frac{\bar{x}_1-\bar{x}_2}{S_{\bar{x}_1-\bar{x}_2}},\quad \text{自由度}=N_1+N_2-2.$$

(4–8) $$\text{P.D.}=.6745S.$$

(4–9) $\text{P.E.}_{\bar{x}} = .6745 S_{\bar{x}}$.

(5–1) $\Sigma(X-\bar{x})(Y-\bar{y}) = \Sigma XY - \dfrac{(\Sigma X)(\Sigma Y)}{N}$.

(5–2) $b = \dfrac{\Sigma(X-\bar{x})(Y-\bar{y})}{\Sigma(X-\bar{x})^2}$.

(5–3) $Y_e - \bar{y} = b(X-\bar{x})$.

(5–4) $\Sigma(Y-Y_e)^2 = \Sigma(Y-\bar{y})^2 - \dfrac{[\Sigma(X-\bar{x})(Y-\bar{y})]^2}{\Sigma(X-\bar{x})^2}$.

(5–5) $V_{y\cdot x} = \dfrac{\Sigma(Y-Y_e)^2}{N-2}$.

(5–6) $S_{y\cdot x} = \sqrt{\dfrac{\Sigma(Y-Y_e)^2}{N-2}}$.

(5–7) $Y_a = Y - b(X-\bar{x})$.

(5–8) $\Sigma(X-\bar{x})(Y-\bar{y}) = \Sigma(X-G)(Y-H) - \dfrac{\Sigma(X-G)\cdot\Sigma(Y-H)}{N}$.

(5–9) $\Sigma(X-\bar{x})(Y-\bar{y}) = \left[\Sigma fxy - \dfrac{(\Sigma fx)(\Sigma fy)}{N}\right](i_x)(i_y)$.

(5–10) $S_b = \dfrac{S_{y\cdot x}}{\sqrt{\Sigma(X-\bar{x})^2}}$.

(5–11) $t_b = \dfrac{b}{S_b}$, 自由度 $= N-2$.

(5–12) $S_{1-2} = \sqrt{S_1^2 + S_2^2}$.

(5–13) $t = \dfrac{b_1 - b_2}{S_{1-2}}$, 自由度 $= (N_1-2)+(N_2-2)$.

(6–1) $r = \dfrac{\Sigma(X-\bar{x})(Y-\bar{y})}{\sqrt{\Sigma(X-\bar{x})^2}\sqrt{\Sigma(Y-\bar{y})^2}}$.

(6-2) $$r=\frac{\dfrac{\Sigma(X-\bar{x})(Y-\bar{y})}{N-1}}{\sqrt{\dfrac{\Sigma(X-\bar{x})^2}{N-1}}\sqrt{\dfrac{\Sigma(X-\bar{y})^2}{N-1}}}=\frac{CV}{\sqrt{(V_x)(V_y)}}.$$

(6-3) $$r=\frac{\Sigma fxy-\dfrac{(\Sigma fx)(\Sigma fy)}{N}}{\sqrt{\Sigma fx^2-\dfrac{(\Sigma fx)^2}{N}}\sqrt{\Sigma fy^2-\dfrac{(\Sigma fy)^2}{N}}}.$$

(6-4) $t_r=\dfrac{r\sqrt{N-2}}{\sqrt{1-r^2}}$, 自由度$=N-2$.

(6-5) $$z=\frac{1}{2}\left\{\log_e(1+r)-\log_e(1-r)\right\}.$$

(6-6) $$S_z=\frac{1}{\sqrt{N-3}}.$$

(6-7) $$S_{z_1-z_2}=\sqrt{\frac{1}{N_1-3}+\frac{1}{N_2-3}}.$$

(6-8) $t_{z_1-z_2}=\dfrac{z_1-z_2}{S_{z_1-z_2}}$, 自由度$=N_1+N_2-6$.

(6-9) $$Y_e-\bar{y}=r\frac{S_y}{S_x}(X-\bar{x}).$$

(6-10) $$X_e-\bar{x}=r\frac{S_x}{S_y}(Y-\bar{y}).$$

(6-11) $$r=\sqrt{\left(r\frac{S_y}{S_x}\right)\left(r\frac{S_x}{S_y}\right)}=\sqrt{b_{yx}\cdot b_{xy}}.$$

(6-12) $$\Sigma(Y-Y_e)^2=(1-r^2)\Sigma(Y-\bar{y})^2.$$

(7–1) $$(q+p)^k=q^k+kq^{k-1}p+\frac{k(k-1)}{2\cdot 1}q^{k-2}p^2+\frac{k(k-1)(k-2)}{3\cdot 2\cdot 1}q^{k-3}p^3+\cdots\cdots+\frac{k(k-1)\cdots\cdots(k-r+1)}{r(r-1)\cdots\cdots 2\cdot 1}q^{k-r}p^r+\cdots\cdots+p^k.$$

(7–2) $$y=\left(\frac{N}{\sigma\sqrt{2\pi}}\right)e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2}.$$

(7–3) $$z=\frac{e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2}}{\sqrt{2\pi}}.$$

(7–4) $$V_{\bar{x}_1-\bar{x}_2}=V_{\bar{x}_1}+V_{\bar{x}_2}.$$

(7–5) $$S_{\bar{x}_1-\bar{x}_2}=\sqrt{V_{\bar{x}_1}+V_{\bar{x}_2}}.$$

(7–6) $$m=kp.$$

(7–7) $$\sigma=\sqrt{kpq}.$$

(7–8) $$m=p.$$

(7–9) $$\sigma=\sqrt{\frac{pq}{k}}.$$

(7–10) $$S=\sqrt{\frac{\Sigma fd^2-(\Sigma fd)^2/N}{N-1}-\frac{1}{12}}(i).$$

(7–11) $$y=\frac{N}{\sigma}z.$$

(7–12) $$\text{偏態}=\frac{\text{均數}-\text{衆數}}{\text{標準差}}$$

(7-13) $$\begin{cases} a_1=\dfrac{\Sigma fd}{N}, & a^2=\dfrac{\Sigma fd^2}{N}, \\ a_3=\dfrac{\Sigma fd^3}{N}, & a_4=\dfrac{\Sigma fd^4}{N}. \end{cases}$$

(7-14) $$\begin{cases} v_1=a_1, \\ v_2=a_2-a_1^2, \\ v_3=a_3-3a_1a_2+2a_1^3, \\ v_4=a_4-4a_1a_3+6a_1^2a_2-3a_1^4. \end{cases}$$

(7-15) $$\begin{cases} k_1=v_1, \\ k_2=\left(\dfrac{N}{N-1}\right)v_2, \\ k_3=\left[\dfrac{N^2}{(N-1)(N-2)}\right]v_3, \\ k_4=\dfrac{N^2}{(N-1)(N-2)}\left[\dfrac{(N+1)v_4-3(N-1)v_2^2}{N-3}\right]. \end{cases}$$

(7-16) $$\begin{cases} k'_1=k_1, & k'_1=k_2-\dfrac{1}{12}, \\ k'_3=k_3, & k'_2=k_4+\dfrac{1}{120}. \end{cases}$$

(7-17) $$\begin{cases} g_1=\dfrac{k'_3}{(k'_2)^{\frac{3}{2}}}, & S_{g_1}=\sqrt{\dfrac{6N(N-1)}{(N-2)(N+1)(N+3)}}, \\ g_2=\dfrac{k'_4}{{k'_2}^2}, & S_{g_2}=\sqrt{\dfrac{24N(N-1)^2}{(N-3)(N-2)(N+3)(N+5)}}. \end{cases}$$

(8-1) $$\chi^2=\Sigma\frac{(A-T)^2}{T}.$$

(8–2) $\chi^2=\frac{(ad-bc)^2(a+b+c+d)}{(a+b)(c+d)(a+c)(b+d)}$.

(8–3) $\chi^2=\frac{(b'^2-4a'c')^2(a'+b'+c')}{(2a'+b')^2(b'+2c')^2}$.

(8–4) $\chi=\frac{(b'^2-4a'c')\sqrt{a'+b'+c'}}{(2a'+b')(b'+2c')}$.

(8–5) $R\times C$ 表中之自由度$=(R-1)(C-1)$.

(8–6) $C=\sqrt{\frac{\chi^2}{N+\chi^2}}$.

(8–7) $P_x=\frac{(a+b)!(c+d)!(a+c)!(b+d)!}{(a+b+c+d)!}\ \frac{1}{a!b!c!d!}$.

(9–1) $\Sigma k(\bar{x}_i-\bar{x})^2=\frac{(\Sigma X_1)^2+(\Sigma X_2)^2+\cdots+(\Sigma X_n)^2}{k}-\frac{(\Sigma X)^2}{nk}$.

(9–2) $\Sigma\Sigma(X_i-\bar{x}_i)^2=\Sigma(X-\bar{x})^2-\Sigma k(\bar{x}_i-\bar{x})^2$.

(9–3) $\Sigma k_i(\bar{x}_i-\bar{x})^2=\frac{(\Sigma X_1)^2}{k_1}+\frac{(\Sigma X_2)^2}{k_2}+\cdots\cdots+\frac{(X_n)^2}{k_n}-\frac{(\Sigma X)^2}{\Sigma k_i}$.

(9–4) $X=\frac{tT+bB-S}{(t-1)(b-1)}$.

(10–1) $\Sigma k_i(\bar{x}_i-\bar{x})(\bar{y}_i-\bar{y})=$

$$+\cdots\cdots\frac{(\Sigma X_1)(\Sigma Y_1)}{k_1}+\frac{(\Sigma X_2)(\Sigma Y_2)}{k_2}$$

$$+\frac{(\Sigma X_n)(\Sigma Y_n)}{k_n}-\frac{(\Sigma X)(\Sigma Y)}{\Sigma k_i}.$$

(11–1) $\beta_{Y1\cdot2}=\frac{r_{Y1}-r_{Y2}r_{12}}{1-r_{12}^2}$, $\beta_{Y2\cdot1}=\frac{r_{Y2}-r_{Y1}r_{12}}{1-r_{12}^2}$.

(11–2) $b_{Y1\cdot 2}=\beta_{Y1\cdot 2}\dfrac{\sqrt{\Sigma(Y-\bar{y})^2}}{\sqrt{\Sigma(X_1-\bar{x}_1)^2}},\quad b_{Y2\cdot 1}=\beta_{Y2\cdot 1}\dfrac{\sqrt{\Sigma(Y-\bar{y})^2}}{\sqrt{\Sigma(X_2-\bar{x}_2)^2}}.$

(11–3) $Y_e=\bar{y}+b_{Y1\cdot 2}(X_1-\bar{x}_1)+b_{Y2\cdot 1}(X_2-\bar{x}_2).$

(11–4) $R^2{}_{Y\cdot 12}=r_{Y1}\beta_{Y1\cdot 2}+r_{Y2}\beta_{Y2\cdot 1}.$

(11–5) $R^2{}_{Y\cdot 12}=\dfrac{r_{Y1}^2+r_{Y2}^2-2r_{Y1}r_{Y2}r_{12}}{1-r_{12}^2}.$

(11–6) $S_{Y\cdot 12}=\sqrt{\dfrac{\Sigma(Y-Y_e)^2}{N-3}}=\sqrt{\dfrac{(1-R^2)\Sigma(Y-\bar{y})^2}{N-3}}.$

(11–7) $S_\beta=\sqrt{\dfrac{1-R^2}{(1-r_{12}^2)(N-3)}}.$

(11–8) $S_{\beta_1-\beta_2}=\sqrt{\dfrac{2(1-R^2)}{(N-3)(1-r_{12})}}.$

(11–9) $r_{12\cdot 3}=\sqrt{b_{12\cdot 3}b_{21\cdot 3}}.$

(11–10) $r_{12\cdot 3}=\dfrac{r_{12}-r_{13}r_{23}}{\sqrt{(1-r_{13}^2)(1-r_{23}^2)}}.$

(11–11) $t=\dfrac{r}{\sqrt{1-r^2}}\sqrt{N-p-2},$ 自由度$=N-p-2.$

		X_1	X_2	X_3	Y
	X_1	$\beta_{Y1\cdot 23}+$	$r_{12}\beta_{Y2\cdot 13}+$	$r_{13}\beta_{Y3\cdot 12}$	$=r_{Y1}$
(11–12)	X_2	$r_{12}\beta_{Y1\cdot 23}+$	$\beta_{Y2\cdot 13}+$	$r_{23}\beta_{Y3\cdot 12}$	$=r_{Y2}$
	X_3	$r_{13}\beta_{Y1\cdot 23}+$	$r_{23}\beta_{Y2\cdot 13}+$	$\beta_{Y3\cdot 12}$	$=r_{Y3}$

(11–13) $Y=\bar{y}+b_{Y1\cdot 23}(X_1-\bar{x}_1)+b_{Y2\cdot 13}(X_2-\bar{x}_2)+b_{Y3\cdot 12}(X_3-\bar{x}_3).$

(11-14) $R^2_{Y\cdot 123}=r_{Y1}\beta_{Y1\cdot 23}+r_{Y2}\beta_{Y2\cdot 13}+r_{Y3}\beta_{Y3\cdot 12}.$

(11-15) $\beta_{Y1\cdot 23}=k_1 r_{Y1}+k_2 r_{Y2}+k_3 r_{Y3}.$

(11-16) $v=\dfrac{1-R^2}{N-m}.$

(11-17) $v_1=vk_{11},\quad v_2=vk_{22},\quad v_3=vk_{33}.$

(11-18) $(\beta_{Y1\cdot 23}-\beta_{Y2\cdot 13})$之標準誤為$\sqrt{v(k_{11}+k_{22}-2k_{12})}$,

$(\beta_{Y2\cdot 13}-\beta_{Y3\cdot 12})$之標準誤為$\sqrt{v(k_{22}+k_{33}-2k_{23})}$,

$(\beta_{Y1\cdot 23}-\beta_{Y3\cdot 12})$之標準誤為$\sqrt{v(k_{11}+k_{33}-2k_{13})}$.

(11-19) $$Y=\bar{y}+\beta_1\frac{\sqrt{\Sigma(Y-\bar{y})^2}}{\sqrt{\Sigma(X_1-\bar{x}_1)^2}}(X_1-\bar{x}_1)+\beta_2\frac{\sqrt{\Sigma(Y-\bar{y})^2}}{\sqrt{\Sigma(X_2-\bar{x}_2)^2}}(X_2-\bar{x}_2)+\cdots\cdots.$$

(11-20) $r_{Y1\cdot 23456}=\sqrt{b_{Y1\cdot 23456}\cdot b_{1Y\cdot 23456}}.$

(11-21) $Y_a=Y-b_{Y1\cdot 2}(X_1-\bar{x}_1)-b_{Y2\cdot 1}(X_2-\bar{x}_2).$

(12-1) $$\begin{cases} aN+b\Sigma X+c\Sigma X^2=\Sigma Y, \\ a\Sigma X+b\Sigma X^2+c\Sigma X^3=\Sigma XY, \\ a\Sigma X^2+b\Sigma X^3+c\Sigma X^4=X^2Y. \end{cases}$$

(12-2) $Y=A+BX_1+CX_2+DX_3+EX_4+\cdots\cdots.$

$$\begin{cases} a=\dfrac{S_1}{n}=\bar{y}, \\ b=\dfrac{(1)(2)}{n(n+1)}(S_2), \end{cases}$$

(12-3)
$$\begin{cases} c=\dfrac{(1)(2)(3)}{n(n+1)(n+2)}(S_3), \\ d=\dfrac{(1)(2)(3)(4)}{n(n+1)(n+2)(n+3)}(S_4), \\ e=\dfrac{(1)(2)(3)(4)(5)}{n(n+1)(n+2)(n+3)(n+4)}(S_5), \text{etc.} \end{cases}$$

(12-4)
$$\begin{cases} a'=a, \\ b'=a-b, \\ c'=a-3b+2c, \\ d'=a-6b+10c-5d, \\ e'=a-10b+30c-35d+14e, \\ f'=a-15b+70c-140d+126e-42f, \\ g'=a-21b+140c-420d+630e-462f+132g, \\ h'=a-28b+252c-1,050d+2,310e-2,772f \\ \qquad\qquad +1,716g-429h. \end{cases}$$

(12-5)
$$\frac{r(r+1)}{1\cdot 2},\quad \frac{(r-1)(r+2)}{2\cdot 3},\quad \frac{(r-2)(r+3)}{3\cdot 4},\quad \cdots\cdots.$$

(12-6)
$$\begin{cases} A=a', \\ B=\dfrac{6}{n-1}(b'), \\ C=\dfrac{30}{(n-1)(n-2)}(c'), \\ D=\dfrac{140}{(n-1)(n-2)(n-3)}(d'), \\ E=\dfrac{630}{(n-1)(n-2)(n-3)(n-4)}(e'), \end{cases}$$

$$\Big\{ F, G, H \text{ 之分子計為: } 2,772,\ 12,012,\ 51,480.$$

(12–7) $$\frac{(2r+1)(2r)(2r-1)(2r-2)\cdots\cdots2\cdot1}{[r(r-1)(r-2)\cdots\cdots2\cdot1]^2}.$$

(12–8) $$\begin{cases} nA^2=\dfrac{(\Sigma Y)^2}{n}, \\[2ex] \dfrac{n(n^2-1)}{12}(B^2), \\[2ex] \dfrac{n(n^2-1)(n^2-4)}{180}(C^2), \\[2ex] \dfrac{n(n^2-1)(n^2-4)(n^2-9)}{2,800}(D^2), \\[2ex] \dfrac{n(n^2-1)(n^2-4)(n^2-9)(n^2-16)}{44,100}(E^2). \\[2ex] \text{此後三項分母為：} 698,544,\ 11,099,088,\ 176,679,360. \end{cases}$$

(12–9) $$\frac{(2r+1)[2r(2r-1)(2r-2)\cdots\cdots2\cdot1]^2}{[r(r-1)(r-2)\cdots\cdots2\cdot1]^4}.$$

(12–10) $$\begin{cases} X_1=X-\bar{x}, \\[2ex] X_2=X_1^2-\dfrac{n^2-1}{12}, \\[2ex] X_3=X_1^3-\dfrac{3n^2-7}{20}(X_1), \\[2ex] X_4=X_1^4-\dfrac{3n^2-13}{14}(X_1^2)+\dfrac{3(n^2-1)(n^2-9)}{560}, \\[2ex] X_5=X_1^5-\dfrac{5(n^2-7)}{18}(X_1^3)+\dfrac{15n^4-230n^2+407}{1,008}(X_1). \end{cases}$$

(12–11) $$y-d=\frac{K}{1+Ce^{rt}}.$$

(12–12) $$\log_e \frac{K-(y-d)}{y-d}=\log_e C+rt.$$

(13–1) $$\frac{(a-b)^2}{2k}.$$

(13–2) $$\frac{(a-2b+c)^2}{6k}.$$

(13–3) $$\frac{(a+b-3c+d)^2}{12k}.$$

(13–4) $$S_M \doteq \frac{S(\log n)\sqrt{B^2+D^2}}{B^2}.$$

(13–5) $$\text{s.e.}=2.303\,S_M(\text{antilog}\,M).$$

(13–6) $$X=\frac{k(R+C+T)-2S}{(k-1)(k-2)}.$$

度量衡表

類別	英文名稱及符號	度量衡局譯名	教育部頒布譯名	換算		
度	Meter(m.)	公尺	米	1 米	=	100 厘米
	Centimeter(cm.)	公分	厘米	1 厘米	=	10 毫米
	Millimeter(mm.)	公釐	毫米	1,000 毫米	=	1 米
	Micrometer(μ)		微米	1,000 微米	=	1 毫米
量	Liter(l.)	公升	升			
	Milliliter(ml.)	公撮	毫升	1 升	=	1,000毫升
	Cubic centimeter(c.c.)	公撮	立方厘米	1,000 立方厘米	=	1 升
	Cubic millimeter(c.mm.)		立方毫米	1.000 立方毫米	=	1 立方毫米
	Cubic micrometer(μ^3)		立方微米	1×10^9 立方微米	=	1 立方毫米
衡	Kilogram(kg.)	公斤	仟克	1 仟克	=	1,000克
	Gram(gm.)	公分	克	1 克	=	1,000毫克
	Milligram(mgm.)	公絲	毫克	1 毫克	=	1,000微克
	Microgram(γ)		微克	1 微克	=	1,000微微克
	Micro-Microgram($\gamma\gamma$)		微微克	1×10^{-12}克	=	1微微克

國際制與我國市用制之互換

1米 = 3市尺	1升 = 1市升	1仟克 = 2市斤
1市尺 = 33.33 厘米	1市斤 = 500 克	1市兩 = 31.25 克

國際制與英國制之互換

1米 = 39.37 吋(Inches)	1吋 = 2.54 厘米
1克 = 15.40 克冷(Grains)	1磅 = 453.6 克
1仟克 = 2.20 磅(Pounds)	1加侖(Gallon, U.S.) = 3,785.4 立方厘米
1升 = 1.06 夸(Quarts)	1盎斯(Fluid ounce) = 29.57 立方厘米

附錄　中英文統計學名詞索引

(表內數目如 2-0 係指第二章爲首之引言，7-5 係指第七章第五節，餘類推.)

D

E

F

G

H

I

K

L

M

Q

R

S

T

U

V

Z

另以英文字母爲首者

中華民國三十七年九月初版

醫學與生物統計方法

全一册　定價金圓券二元二角五分
(精裝本定價另加金圓券七角五分)
(外埠酌加運費匯費)

著　者　郭祖超
發行人　蔣志澄
印刷所　正中書局
發行所　正中書局
(2327)

整自
校向

滬・本　　2/2－0.15

合

中華民國捌拾陸年柒月拾捌日