R. A. Fisher의
實驗計劃法의 方法論的 考察

崔　俊　楨

目　　次

Ⅰ. 序　　論

統計學은 最近 數十年間에 特히 第二次大戰을 契機로 하여 面目을 一新하고 劃期的 發展과 進步를 가져왔다.

일찌기 William Gosset(1876～1936)에 曙光을 찾았고 R. A. Fisher(1890～)에 이르러 近代推測統計學의 理論的基礎가 確立되고 그 體系的 大成을 보게된 小標本論은 K. Pearson(1857～1936)의 記述統計論的 立場으로서는 到底히 現象의 本質을 的確하게 認識할수 없게 되었다. 그러므로 여기에 記述統計論的 立場과 全혀 그 理論構造를 달리하는 統計推理의 方法으로서 近代 推測統計學의 새로운 局面을 開拓하게 되었고 이리하여 統計學은 顯著히 實用的인 것으로되어 計劃이라던가 管理라고 하는 行動的觀點에서, 다시 말하면 計劃의 學問으로서 管理의 方法論을 規定하는 推測統計學의 驚異的 發展을 約束하게 되었다.

如斯한 發展의 뒷받침으로서 指摘할點은 무엇보다도 最近 約一世紀間에 걸친 圃場試驗技術의 發展이 現代推測統計學의 學問的 基盤을 닦아 놓게 되었다는 事實이다. 1843年 London 의 北方 約 25哩 地點에 Rothamsted 農事試驗場이 設立되었으며 이農事試驗場에서 當面한 여러가지 問題가 드디어 R. A. Fisher 에 依하여 革命的인 科學的轉換을 가져오게 됨으로써 圃場試驗에서 解決하지 않으면 안될 問題가 Fisher 의 確率化無作爲法의 思想을 導入함으로써 完全히 解決되어 이것이 이른바 實驗計劃法(Method of Random Arrangement)을 大成하게 하였다.

特히 20世紀 初葉以來 顯著하게 經濟學 前面에 크로즈·앞된 經濟現象으로서의 量產方式의 全面的 支配는 統計的管理狀態를 招來하게 되었으며 品質管理라든가 拔取檢査等의 統計的 方法이 利用되게 되었던 것이다. 이와같은 大量生產樣式에 있어서의 統計的方法은 標準化—生產—檢査라고 하는 生產過程을 假說의 設定—實驗의 遂行—假說의 檢定이라는 科學的 認識의 進行에 適應시킴으로써 前述한바 圃場試驗法에서의 科學的 成果를 品質管理및 拔取檢査라고 하는 理論的으로는 推定論 또는 假說檢定論에 基礎한 推測統計學 適用의 科學的 論理論의 當然한 歸法을 가져오게 하였다.

이와같은 推測統計學의 發展은 今日에 이르러서는 前記한 圃場試驗 또는 量產管理에서 뿐만 아니라 온갖 科學全分野에 걸처 重要한 硏究手段으로 된다.

Ⅱ. 數理統計學의 發展沿革

Buckle 가 그著 History of Civlization in England 에서 Quetelet 의 統計學에 關하여 다음과 같이 評價하고 있다. 卽 Buckle 는 「從來의 無數한 學者들의 倫理學에 關한 硏究와 發表는 Quetelet 의 出現으로 말미암아 無와 다름 없이 되었다……Quetelet 를 비롯한 道德統計 硏究家들은 不過 十年間에 思辨的科學이 數千年의 業績보다 더욱 훌륭한 成果를 남겨 놓았다」라고 極口 讚揚하였던 Quetelet 的 統計學도 그自體가 機械論的 唯物論의 領域을 벗어나지 못하는 限 그進路에는 스스로 限界가 있었으며 마침내 Wilhelm Lexis(1837～1914)에 依하여 批判되었다.

卽 W.Lexis 에 依하면 從來 驚異的인 規則性을 지닌다고 믿어지던 Quetelet的 統計量의 大部分이 그 安全度가 標準以下 卽 過大分散이라고 하는것이 밝혀졌으며(이것을 Lexis 의 標準이라함)이로 말미암아 Quetelet 에 大成된 古代統計學은 近代統計學에로의 出發點을 마련하여 주고 스스로의 歷史的役割을 數理統計派學를 위하여 統計方法論에 있어서 보다 數學的이며 確率論的인 터전에서 새로운 統計學의 發展을 期約하였다.

이와같은 歷史的 過程에서 所謂 數理統計學派는 Quetelet 때 까지의 統計的 認識方法을 止揚하고 그들은 統計的 認識의 諸段階에서 卽 記錄 分類 標識化의 過程을 걸처 數量化에 이르는 사이에 있어서 數量化의 槪念을 이미 數學外部의 問題로서 統計的 認識의 諸段階에서 가장 重要한 問題點을 解決하였던 것이다.

뿐만 아니라 이미 Quetelet 以來 古典統計學에 있어서의 古典確率論의 發達과 더부러 統計學에 있어서 數學의 援用이 顯著하게 發展하고 있다는데서 찾아 볼수 있다. 그 몇가지 例로서 Quetelet 에 對한 Lexis 의 批判以來 數理統計學이 開拓하게 된 重要한 原因은 역시 當時에 이미 Gibbs 의 統計集團에 關한 近代的槪念 및 Boltzman 의 Ergodic theory 等 古典確率論의 範疇를 克服하고 確率의

近代化에 關한 鼓舞的 貢獻이 있었던 것이다. 이밖에도 自然科學의 發展은 卽

i) C. R. Darwin(1809~1882)에 依하여 獨創的 完成을 이룩한 進化論은 記述統計學의 誕生과 發展을 期約하게 되었으며

ii) Energy 恒存法則의 確立은 J. C. Maxwell(1831~1879) Ludwig Boltzman(1844~1906) J. W. Gibbs(1839~1903)等에 依하여 統計力學에 까지 到達하여 여기에 近代確率論이 統計學의 數理的基礎를 이룩하였으며

iii) 한便 細胞學의 發達은 遺傳學을 通해서 間接的이 나마 推測統計學의 發達을 促進하게끔 되었던 것이다.

이와같이 하여 數理統計學은 記述統計學과 推測統計學의 두分野에서 成立되는 것이며 具體的으로는 前者는 Karl Pearson 에 依하여 大成되고 後者는 R. A. Fisher 에 依하여 確立된 近代統計學으로서 Person Jr. 美國의 Neyman, Wald 等에 依하여 더욱 活潑하게 開拓되어가고 있다.

特히 最近 美國 英國等에 있어서 가장 뚜렷한 事實은 그곳 統計學者들이 한걸음 더 나아가서 統計學을 元來 數理統計學이라고 하는 學問的立場을 堅持하고 있으며 이와같은 傾向은 美國의 Wilks, S. S의 Theory of Statistical Inference(1937) Mathematical Statistics(1943) Elementary Statistical Analysis(1948)에 强調하고 있으며 亦是 記述統計論과 統計推理論이 包含되는 것으로 보아진다.

Ⅱ. Ronald A. Fisher(1890~)의 學問的歷程

Fisher 는 1890年 2月 London 東北方에 있는 East Finchely 에서 Fisher 家의 第4男으로 出生하였다. 兩親이 모두 當時 英國社會에 있어 中流以上 生活과 其家系 및 環境은 英國的良識과 實務的인 健全한 家庭生活을 營爲하였으며 如斯한 家庭環境에서의 影響은 Fisher 自身의 學問的 發展의 素地를 마련하였다.

그는 幼時부터 視力이 弱하였기 때문에 그의 數學 家庭敎師는 電燈아래에서의 工夫를 禁하고 종이와 鉛筆을 使用시키지 않고 暗記式으로 問題를 풀기하는 訓練을 거듭하였다. 이와같은 天才的 敎育方法이 後年에 이르러 그가 高次元 幾何學的 方法으로 問題를 풀수있는 素養을 길러 주었을 것이다.

Fisher 는 有名한 Harrow School 을 마치고 1909年 Cambridge 大學에서 가장 由緖깊은 Gonville and Caius College 에 入學하였으며 그의 天禀은 드디어 1912年 Mathematical Tripos 第2部를 首位로서 合格하였으나 適當한 일자리가 없어 不得已 大學에 머물고 約1年間 James Jeans 에서 統計力學과 量子論 其外 誤差論等을 硏究할 機會를 얻었으며 이 1年間 大學에서의 聽講이 그로 하여금 物理 數學에 對한 깊은 關心과 憧憬을 가지게 함으로써 그의 學問的歷程에 一大轉機를 區劃하게 하였다.

當時는 Mendelism 의 再發見後로서 W. Bateson 의 遺傳學書가 19世紀의 歷

史的科學的 創業的遺産이었던 Darwin 의 偉大한 思想 卽 進化論을 壓倒하고 있던 때였다. 젊은 Fisher 는 이와같은 學問的 雰圍氣속에서 그의 學問에 對한 불타는 情熱은 마침내 Mendelism 的 立場에서 進化論을 定量的 數理的으로 體系化할려고 決心하였던 것이다.

이와같은 Fisher 의 學究的努力이야 말로 Karl Pearson 의 記述統計學的 立場을 止揚하고 Mendelism 의 學問的 立場을 肯定함으로써 그의 統計推理論의 偉大한 學問的體系로서 推測統計學發展의 大路를 開拓하게 되었던 것이다.

그리하여 그의 學問的 成果와 統計思想이 集大成되어 1938年 Statistical Theory of Estimation, Calcutta University Readership Lecture 로서 推測統計學의 新生面을 힘차게 開拓하게 하였다.

Fisher 는 역시 統計實務家로서 Mercantile & General Investment Co. 에 勤務하고 있을때 卽 1913年 "Biometrica"誌上에 실린 H. E. Soper 의 相關係數의 標本分布에 關한 研究論文을 非常한 觀心으로 읽고 Soper 가 問題를 提起하고 있는 觀點은 研究的이나 아직 問題解決에 關하여 抽象的인點에 着眼하고 Fisher 의 斯學에 對한 熱意는 드디어 高次元 幾何學的 方法으로 共問題의 解答을 밝힘으로써 그의 推計學的 理論體系를 整理 發表한 것이 後에 W. Snedector 가 Fisher 의 推計論 卽 Snedector 自身이 誘導한 Z-分布 바로 그것을 이미 創見的努力으로 解決한 바를 깊이 간직하고 F-分布하고 命名하여 Fisher 의 學問的業績을 높이 評價하였다.

마침내 Fisher 는 前記 會社를 辭任하고 1915年부터 Public School 에서 物理學과 數學을 講義하면서 그는 이 學究的인 몇年을 理論에서 이것을 實驗化하는 努力을 거듭하여 實驗技術을 體得하였고 이것이 그의 偉大한 學問的成果인 實驗計劃法의 創始에 커다란 도움이 되었다.

그리하여 Fisher 가 1919年부터 London 郊外에 자리잡은 Rothamsted 農事試驗場에서의 實驗的努力은 그의 獨特한 研究成果와 더불어 近代推測統計學의 萠芽가 힘차게 싹트기 始作하였다.

이農事試驗場에서 Fisher는 E. S. Beaven 의 Rod-Rows Method 또는 Half-Drill strip Method 等 高貴한 試驗成果에 對한 資料를 蒐集 研究하는 過程에서 그는 Karl Pearson 式 記述統計學的方法이 往往 生物現象中에서 利用될수 있는 統計資料가 極히 制限되어 小數의 경우에 부디치게되면 現象의 本質을 的確하게 認識할수 없는 K. Pearson 的 數理統計學의 重大한 矛盾을 發見하고 그는 記述統計學을 大成한 Pearson 的 統計學的 世界를 克服하여 새로운 推測統計學的世界에서 問題의 解決點을 摸索하면서 推測統計學(inductive statistics) 또는 統計的推理論(theory of statistical inference)라고 불리우는 새로운 分野를 創始하게 되었으며 그 理論的 根據로서 母數推定論 X^2-分布의 自由度 歸無假說및 그

檢定法等 理論은 모두 前述한바 그의 努力의 偉大한 結晶이며 이와같은 理論的 바탕을 背景으로하여 우리는 Fisher 의 實驗計劃法을 硏究하고저 한다.

Ⅲ. 實驗計劃法(Design of Experiments)

統計學에 있어서 實驗計劃法이 適用되게 된것은 19世紀 中葉以來 農事試驗技術로서 利用되었으나 1923年 以降 R. A. Fisher 에 依하여 처음으로 創始되어 今日 統計學에 있어서 가장 重要한 硏究方法으로 되고 있다.

實驗計劃法은 이것을 簡約하게 規定하면 標本을 選擇하는 方法이다. 從來 우리들은 標本의 任意抽出이라고 하는 것을 흔히 論議하여 왔지만 이것만으로서는 到底히 統計學의 所期한 目的을 達成할수 없음을 이미 밝힌바 있다. 그 理由는 우리가 取扱하는 母集團全體가 等質하다고 하는 것이라면 標本의 任意抽出法 만으로서도 充分하겠지만 그러나 우리들의 實際의 經驗은 母集團이 異質的成分으로 構成되어 있는 일이 許多한 것이다. 그러므로 우리들은 母集團을 처음부터 그와같은 成分을 가진 階層으로 分類하여 各層에서 標本을 抽出하는 것보다 標本의 크기가 같을 경우에는 精度가 向上되는 것이다. 그러나 異質的인 成分으로 構成되어 있는 경우 그標本이 各階層에서 精確하게 選擇되지 않는 경우에는 大端히 重大한 結果를 惹起할수 있다는 것을 깊이 注意하지 않으면 안된다.

여기서 우리들이 當面하는 가장 緊要한 問題點은 母集團의 各層을 包含하는 抽出計劃이 必要하게 提起되는 것이며 이와같은 技術과 方法을 硏究하는 것이 Fisher 의 所謂 Design of Experiments 라고 하겠다.

그러므로 다음에 우리들은 實驗計劃法에 關한 統計學的 問題點을 解明하여 보기로 하겠다.

A) 實驗計劃法의 槪念

Rothamsted 農事試驗場에 있어서의 農藝技術은 最近 一世紀以來 急速度로 發展하여 農事試驗은 이미 技藝의 領域에서 科學의 領域에 進步하였으며 同試驗場에 있어서 從來까지 試用하던 classical fields 라고 하는 試驗方法으로 到底히 所期의 成果를 保障하기가 困難하였다.

卽 同方法에 依한 試驗은 主로 面積 20a以下의 耕地를 一區로 劃定하고 肥料의 効果 品種의 優劣을 試驗하였는데 地力이라든가 地質의 差異로 말미암아 그 差가 顯著하였으며 이것은 서로 隣接한 두試驗區에 있어서 試驗始作以來 同一한 施肥를 하였음에도 不拘하고 產額에 있어서 10%~20%의 差異를 가져오게 되었던 것이다.

그러므로 同試驗場에서는 有能한 農業技師 Mercer, Hall 및 Montgomery 等의 不斷한 試驗硏究 結果 收量의 變動은 地域的으로 的確한 at random 로 나타나는 것이 아니라 各各 相異한 試驗區에서의 收量의 變動은 그收量에 關해서 等

高線을 그리게 되면 꼭 地圖에서 나타나는 것같은 丘陵形이 곳곳마다 나타나게 되며 역시 相異한 試驗區間에는 正의 相關關係가 示顯된다는 事實을 알수 있었던 것이다.

여기서 同試驗場의 Beaven 博士를 비롯한 技師들은 地力의 變動을 的確하게 各試驗區마다 影響 받을 수 있도록 하여야 하겠다는 事實과 그러기 위하여 試驗區를 可能한 限 적게 設計하여야 하겠다는 結論을 얻고 이에 關한 研究를 거듭하여 Beaven 의 여러가지 配置法이 考案되었다.

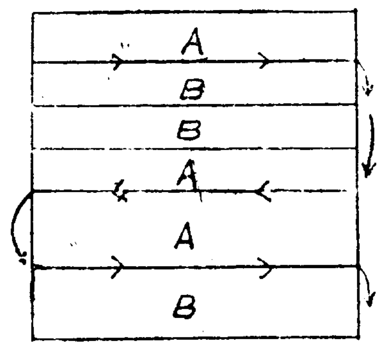
ㄱ. Rod-Rows Method

그러한 一聯의 研究의 한 方法으로서 Rod-Rows Method 가 適用되었는데 이 方法은 1 Rod 卽 5.5碼의 길이를 가지는 밭고랑을 만들어 한고랑마다 植付品種을 바꿔가는 思想인데 그러나 實地問題로서는 各고랑마다 交互作用이 있으므로 이것을 除去하기 위하여 세고랑 程度까지 同一品種을 植付하고 한고랑 떼어서 다른 品種을 植付하게 하였다.

ㄴ. Half-Drill Strip Method

이方法은 播種機의 種入箱의 內部를 左右로 區分하여 A. B. 二種의 相異한 品種을 넣어가지고 한번에 A. B를 밭에 植付하고 同時에 두고랑이 植付가 끝나면 다음은 反對쪽에서 다음 第一圖와 같이 植付하는 方法이며 이 方法으로 植付하면 第一圖에서 보는 바와 같이 A. B. B. A. A와 같이 植付한 고랑이 생기게 된다.

第一圖



ㄷ. Beaven 의 正方形試驗區의 棋盤方式(chessboard system of square yard plots)

이方案은 E.S. Beaven 博士가 案出한 方法인데 이것은 一邊 四呎의 長方形으로 되는 試驗區를 만들어 其中에 六吋幅으로 八列의 고랑을 만들고 각고랑 마다 2吋 間隔으로서 심어가는 試驗區를 將棋盤과 같이 配置하였다. 例를들

〔第二圖〕

A	F	C	H	E	B	G	D
B	G	D	A	F	C	H	E
C	H	E	B	G	D	A	F
D	A	F	C	H	E	B	G
E	B	G	D	A	F	C	H

〔第三圖〕

A	F	F	A	E	G	B	D
B	G	E	A	F	F	A	E
C	H	D	B	G	E	A	F
D	H	C	C	H	D	B	G
E	G	B	D	H	C	C	H

면 A. B. C. D. E. F. G. H의 八種의 品質을 調査하는 경우에는 第二圖와 같이 되지만 여기에서는 A가 H에 比하여 지나치게 左側으로 치우쳐 있음으로 配列方法을 改良하여 第三圖와 같이 하였다.

그러나 이와같은 方法으로서는 그配列方法이 系統的이므로 誤差가 크게 될 念慮가 있는것이다.

마침 이와 때를 같이하여 Fisher가 同試驗場에 關係하게 되었으며 그는 1923年 圃場試驗法에 關한 그의 革命的인 論文을 發表하여 이미 이論文에서 그의 實驗計劃法에 關한 學問的 理論體系의 基礎를 마련하고 그의 同試驗場에서의 高貴한 研究成果는 1925年 R. A. Fisher: Statistical Methods for Research Workers로 公表되었다.

ㄹ. Fisher의 亂塊法(randomized block method)

이方法은 前述한 그의 論文에서 그 片鱗를 찾아 볼수 있는바 이方法은 試驗場內에 均等한 地力條件을 가지고 있다고 認定되는 4個의 地帶를 選定하여 이地帶區劃內에서 例를 들면 위에서 말한 八種의 品種을 亂數表에 依한다든가 또는 주사위를 던지든가 하여 第四圖와 같이 任意로 配置하는 方法이다.

〔第四圖〕

C	D	F	C	E	H	A	G
H	A	B	H	G	C	C	H
B	E	E	D	D	A	F	B
F	G	A	G	B	F	D	E

Fisher의 이方法은 系統的이 아니므로 前示한 Beaven의 方法에 比하여 優秀하며 各地帶間에 地力의 差가 있어도 其中에서는 一定한 때에 有効하지만 假令 亂數表로서 定하여도 어떤 特定한 하나가 大多數의 地帶에서 다른것 보다 恒常 有利한 地位를 차지하지 않는다고 하는 保證은 없는 것이다.

ㅁ. Fisher의 Latin square

여기서 Fisher는 다시 Latin 文字 A B C D等을 各行 各列에 同一한 文字가 單한번만 記錄될 수 있도록 正方形으로 配置하는 方法을 研究하였다(이경우 文字의 數가 n個 있는 것을 n×n Latin 方格이라고 한다)

例를 들어 說明하면 3×3 Latin 方格은 第五圖와 같이 12個(1×3!×2!)뿐이며

〔第五圖〕

A	B	C
B	C	A
C	A	B

B	C	A
A	B	C
C	A	B

C	A	B
B	C	A
A	B	C

A	B	C
C	A	B
B	C	A

A	C	B
B	A	C
C	B	A

B	A	C
A	C	B
C	B	A

C	B	A
B	A	C
A	C	B

A	C	B
C	B	A
B	A	C

B	A	C
C	B	A
A	C	B

C	A	B
A	B	C
B	C	A

B	C	A
C	A	B
A	B	C

C	B	A
A	C	B
B	A	C

이것中의 어느 것이든지 確率的으로 選擇한다.

第五圖에서 보는 바와 같이 第一行 第一列이 모다 自然的 順序로 配列되어 있는 것을 標準型이라고 하며 餘地는 모두 이것을 行 列 文字에 關해서 交換한 것에 不過하다.

따라서 例를 들어 說明하면 假令 여기에 六種의 肥料의 實驗에 있어서는 大略 正方型의 試驗場을 六行 六列로 分割한 後에 이 36個의 試驗區에서 確率的으로 定한 이것들中에서 어느것이 든지 使用하여 割當하면 되는 것이다. 그러나 여기서 우리가 注意하여야 할 點은 實際問題에 있어서는 研究對象이 되는 要因은 許多하게 많을 것이며 例를 들어 보면 어떤 作物의 品種試驗에 있어서도 植付時期 土壤의 質, 其肥沃度 肥料의 量等 重要한 變數를 生覺하지 않으면 안될 것이다.

더욱이 生物學, 農學等 部門에 있어 한가지 實驗이 一年間이라든가 甚한 경우에는 一世代에 걸치는 일도 있을 것이므로 한가지 實驗을 可能한 限 完全하게한 다는것이 重要하다. 그러므로 이것들의 要因의 變數의 數에 副應하여 Latin 方格은 Graces-Latin 方格 hyper-Latin 方格 및 要因分析法에로 發展하게 되는 것이 Fisher 의 學問的 發展過程인 것이다. 여기에 添言하여 둘 일은 實驗計劃法의 中心的인 指導的 方法은 分散을 比較 檢討하는 일이다.

이것은 實驗計劃法 理論의 가장 中心的인 課題로서 Ficher 의 學問的理論 의 精華로서 所與의 各項에서 詳論하겠다.

이와같은 Rothamsted 試驗場에서의 가지가지의 研究上 成果는 그것이 非單 農事試驗에만 局限 된 것이 아니라 科學의 全分野에 걸쳐 其理論 研究 分析의 가장 重要한 武器로 되는 것이며 이것이 現代 推測統計學 段階에서 統計的 應用 分野에서 우리가 깊은關心을 가지지 않을 수 없는 問題點이 되는 것이다.

B) 實驗計劃法의 理論構成

R. A. Fisher 의 實驗計劃法에 關한 理論은 그에 依하여 創始된 새로운 學問 分野로서 主로 1923年 發表한 그의 勞作 Statistical Methods for Research

Workers(1925)및 The design of experiments(1935初版)에서 그의 實驗論理의 構造를 闡明하고 있다.

本稿에서는 主로 前述한 그의 獨創的이며 革命的인 二大勞作에 開陳되어 있는 實驗計劃法의 理論構成要素로서의 所謂 Fisher 의 三大原理 卽

ㄱ. 細分化의 原理(級內變動과 級間變動으로 區分하되 級內變動은 可能한 限 等質化할수 있도록 한다)

ㄴ. 反覆數의 原理(同一條件으로서 實驗을 되푸리 한다.)

ㄷ. 確率化의 原理(結果를 攪亂시키는 攪亂因子를 實驗實施中 固定시켜 놓고 이와는 何等의 關聯도 맺음이 없이 獨自的으로 可能한 限 無作爲化한다. 卽 이것은 前記한 擾亂因子의 影響이 實驗에 作用할 수 없도록 하기 위한 操作이다. 이것이야말로 實驗計劃法의 核心을 構成하는 原理이다)

以上 三大原理에 立脚하여 實驗計劃理論이 實際에 利用되는 計劃에 關하여 硏究하여 보기로 함으로써 實驗計劃理論을 考察하기로 하겠다.

a. 一元配置法(one-way layout)

이것은 Fisher 의 第一原理에 基礎하여 實際計劃에 利用함에 있어서는 實測値全體의 變動을 必然性을 나타내리라고 민어지는 級間變動과 偶然性을 나타내리라고 민어지는 級內變動으로 分離하여 歸無假說을 檢定하는 方法이다.

至今 알기 쉽게 說明하기 위하여 다음 例로써 說明하겠다.

우리들은 假令 雜種의 品質 肥料의 質및 量 植付時期等 相互 質이 다른 測定値라 하여도 處理라고 하는 名目으로서 結付시킬수 있다. 여기서 表現을 加一層 簡便하게 하기 위하여 模型的인 세가지 處理가 되어있는 5個의 個體를 測定하였다고 하자. 卽 Xij를 處理j가 주어질수 있는 i番째의 測定値로 하고 所與의 data가 第一表와 같이 配列되고 그總計와 平均이 求해졌다고 한다. 卽 $T._1$은 第一列의 모든 測定値의 合計이며 $T._2$는 第二列의 $T._3$는 第三列의 各各의 計이다.

$T..$는 모든 列別 合計의 總計이며 따라서 이것은 勿論 모든 測定値의 總計와 같다.

〔第一表〕

處理	I	II	III	計
	X_{11}	X_{12}	X_{13}	
	X_{21}	X_{22}	X_{23}	
	X_{31}	X_{32}	X_{33}	
	X_{41}	X_{42}	X_{43}	
	X_{51}	X_{52}	X_{53}	
計	$T._1$	$T._2$	$T._3$	$T..$
平均	$\bar{X}._1$	$\bar{X}._2$	$\bar{X}._3$	$\bar{X}..$

卽

$$T._1 = X_{11} + X_{21} + X_{31} + X_{41} + X_{51}$$

$$T._2 = X_{12} + X_{22} + X_{32} + X_{42} + X_{52}$$

$$T._3 = X_{13} + X_{23} + X_{33} + X_{43} + X_{53}$$

$$T.. = T._1 + T._2 + T._3$$

$$\bar{X}._1 = \frac{T._1}{5},\quad \bar{X}._2 = \frac{T._2}{5},\quad \bar{X}._3 = \frac{T._3}{5}$$

$\bar{X}.. = \frac{T..}{6}$이다.

여기서 檢定되는 假說은 「處理間의 平均에는 差가 없다」 換言하면 「세個群의 5個는 모두 同一平均을 가진 母集團에서 抽出한 것이다」라고 하는 것이며 有意水準은 實驗을 하기 前에 決定된 것으로 한다.

모든 數値가 同一한 母集團에서 抽出되었다고 하면 平均의 標本分布의 分散은 $\frac{\delta^2}{n}$이 될 것이다.(여기서 δ^2=母分散 n=各各群의 個體의 數이며 이경우 n=5)

다음에 檢討하여야 할 點은 第一表의 Ⅰ, Ⅱ, Ⅲ項의 各處理에 있어서 實測値의 數가 n_1, n_2, n_3로서 서로 다른 경우 그 結果를 擴張하는 問題에 關하여 보기로 하겠다.

爲先 먼저 두가지 方法으로서 母數 δ^2을 推定하기로 하겠다.

第一의 方法은

$$S^2{}_p=\frac{\sum_{i=1}^{n1}(Xi_1-\bar{X}._1)^2+\sum_{i=1}^{n2}(Xi_2-\bar{X}._2)^2+\sum_{i=1}^{n3}(Xi^3-\bar{X}._3)^2}{n_1+n_2+n_3-3}\cdots\cdots\cdots\cdots(1)$$

第二의 方法은 平均의 分散에 關한 推定은 $\frac{\delta^2}{n}$이므로 平均의 分散에 n를 곱하면 δ^2의 評價가 될것이니 直接 平均에 關한 分散을 計算하여 n倍한다. 따라서 여기에서는

$$S^2=\frac{\sum_{i=1}^{k}ni(\bar{X}.i-\bar{X}..)^2}{k-1}\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots(2)$$

〔여기서 k는 處理의 數이며 上例에서 k=3이다.〕

또 第一表에서와 같이 各處理의 實測値의 數가 모두 相等하다고 하면

$$S^2=\frac{n\sum_{i=1}^{k}(\bar{X}.i-\bar{X}..)^2}{k-1}\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots(2')$$

로 할수 있다.

이것들 두개의 δ_2의 評價에 有意한 差가 있는가 하는 與否는 F分布를 利用함으로써 檢定된다.

卽 萬若 各群이 相異한 平均을 가지는 母集團에서 抽出된 것이라고 하면 (2)및 (2′)에서 推定된 것은 (1)에서 推定된 것 보다 훨신 크게 될 것이다. 實際의 우리들의 當面하는 經驗的要求에서 (1)및 (2)式은 計算에 便利하도록 簡單한 式으로 고칠수 있다. 다음에 k=3의 경우에 關하여 證明하여 보이기로 하겠다(但 k의 數値가 相異한 경우에도 꼭 같이 類推 할 수 있다)

公式 (1)의 分子를 普通 級內平方計라고 하며 (2)式의 分子를 級間平方計라한다.

$$級內平方計=\sum_{i=1}^{n1}Xi_1{}^2-\frac{(\Sigma Xi_1)^2}{n_1}+\sum_{i=1}^{n1}Xi_2{}^2-\frac{(\Sigma Xi_2)^2}{n_2}+\sum_{i=1}^{n3}Xi_3{}^2-\frac{(\Sigma Xi_3)^2}{n_3}$$

$$=\sum_{i=1}^{n_1} Xi_1^2-\frac{T._1^2}{n_1}+\sum_{i=1}^{n_2} Xi_2^2-\frac{T._2^2}{n_2}+\sum_{i=1}^{n_3} Xi_3^2-\frac{T._3^2}{n_3}\ (\because \sum_{i=1}^{n_j} Xij=T.j)$$

$$=\Sigma\Sigma Xij^2-\left(\frac{T._1^2}{n_1}+\frac{T._2^2}{n_2}+\frac{T._3^2}{n_3}\right) \quad \cdots\cdots(3)$$

$$級間平方計=n_1(\bar{X}._1\bar{X}..)^2+n_2(\bar{X}._2-\bar{X}..)^2+n_3(\bar{X}._3-\bar{X}..)^2$$

$$=n_1\bar{X}._1^2-2n_1\bar{X}._1\bar{X}..+n_1\bar{X}..^2+n_3\bar{X}._2^2-2n_2\bar{X}._2\bar{X}..+n_2\bar{X}..^2$$

$$+n_3\bar{X}._3^2-2n_3\bar{X}._3\bar{X}..+n_3\bar{X}..^2$$

$$=\frac{n_1^2\bar{X}._1^2}{n_1}+\frac{n_2^2\bar{X}._2^2}{n_2}+\frac{n_3^2\bar{X}._3^2}{n_3}-2\bar{X}..(n_1\bar{X}._1+n_2\bar{X}._2+n_3\bar{X}._3)$$

$$+\bar{X}..^2(n_1+n_2+n_3)$$

$$=\frac{T._1^2}{n_1}+\frac{T._2^2}{n_2}+\frac{T._3^2}{n_3}-2\frac{T..}{N}(T._1+T._2+T._3)+\frac{T..^2}{N}\ (\because ni\bar{X}.i=T.i)$$

$$=\frac{T._1^2}{n_1}+\frac{T._2^2}{n_2}+\frac{T._3^2}{n_3}-\frac{T..^2}{N} \quad \cdots\cdots(4)$$

또 總平均에 對한 모든 數値의 偏差의 平方計를 全平方計라고 하며 따라서

$$全平方計=\Sigma\Sigma Xi^2j-\frac{(\Sigma\Sigma Xij)^2}{N}$$

$$=\Sigma\Sigma\, Xij-\frac{T..^2}{N} \quad \cdots\cdots(5)$$

여기서 ③式과 ④式을 合하면 ⑤式을 얻을 수 있음으로 級內平方計와 級間平方計를 求하면 全平方計는 容易하게 求할 수 있다. 그것은 級內平方計가 實測値의 그리고 級間平方計가 平均値間의 撤布를 나타내고 있음으로 全體의 撤布는 이 總計가 되는것은 當然하기 때문이다.

다음에 이것을 實際의 計算例로서 풀어 보기로 하겠나.

〔例〕 第二表에(說明을 위한 架室 data이다) 表示된 4個의 處理間에는 何等의 差가 없다, 卽 平均이 서로 相等하다는 것을 5%의 有意水準으로서 檢定하라

〔解〕 計算의 便利上 세개의 平方計를 別途로 計算하면

〔第二表〕

處理	I	II	III	IIII	計
	7	6	8	7	
	2	4	4	4	
	4	6	5	2	
				5	
計	13	16	17	18	64

$$級間平方計:\left(\frac{13^2}{3}+\frac{16^2}{3}+\frac{17^2}{3}+\frac{18^2}{4}\right)-\frac{64^2}{13}$$

$$=319-315.08=3.92$$

$$級內平方計:7^2+2^2+4^2+6^2+4^2+6^2+8^2+4^2+5^2+7^2+4^2+2^2+5^2-\left(\frac{13^2}{3}+\frac{16^2}{3}+\frac{17^2}{3}+\frac{18^2}{3}\right)=37.00$$

$$全平方計:7^2+2^2+4^2+6^2+4^2+6^2+8^2+4^2+5^2+7^2+4^2+2^2+5^2-\frac{64^2}{13}=356-315.08=40.92$$

이 計算의 結果를 第三表와 같이 整理 分析한 것을 分散分析表라고 하며 이에 關해서는 다시 論하겠다.

끝으로 總括的으로 一元配置法의 手順을 要約하여 理解를 도웁기로 하겠다.

〔第三表〕

要 因	平方計	自由度	不扁分散	分 散 比
級間變動	3.92	3	1.31	$F=\frac{1.31}{4.11}=0.32$
級內變動	37.00	9	4.11	$F_9^3(0.05)=3.86$
全 變 動	40.92			

첫째 假說: $M_1=M_2=M_3=\cdots Mk$ 卽 k個의 階級의 平均이 모다 相等하다.

둘째 有意水準: α를 適當히 定한다.

셋째 統計量 F로하고 級間變動과 級內變動의 不偏分散의 比 를利用한다.

넷째 $\delta_1{}^2=\delta_2{}^2=\cdots\cdots=\delta k^2$ 卽 等分散의 正規母集團부터의 標本으로서 假說이 眞이면 統計量 F는 自由度($k-1$, $\Sigma ni-k$)의 F-分布가 된다.

다섯째 棄却域은 $F>F^{k-1}_{\Sigma ni-k}(\alpha)$

여섯째 棄却域에 있는가 그렇치 않는가 하는데 따라 假說을 棄却 또는 採擇한다.

b. **二元配置法tow-way layout**

二元配置法도 역시 Fisher 의 第一原理에 基礎하여 實際計劃에 利用됨에 있어서는 二種類의 要因에 依하여 支配되고 있다고 生覺되는 實驗計劃에 對하여 分析하는 方法이다.

至今 알기쉽게 測定值가 分類된 要因의 모든 組合에 關하여 하나式 있는 경우를 想定하겠다. 例를 들면 Ⅰ, Ⅱ, Ⅲ, Ⅲ 의 四個의 品種과 a, b, c 의 세가지

〔第四表〕

	Ⅰ	Ⅱ	Ⅲ	Ⅲ	V	計	平 均
1	X_{12}	X_{12}	X_{13}	X_{14}	X_{15}	$T_{1\cdot}$	$\overline{X}_{1\cdot}$
2	X_{21}	X_{22}	X_{23}	X_{24}	X_{25}	$T_{2\cdot}$	$\overline{X}_{2\cdot}$
3	X_{31}	$X_{?2}$	X_{33}	X_{34}	X_{35}	$T_{3\cdot}$	$\overline{X}_{3\cdot}$
計	$T_{\cdot 1}$	$T_{\cdot 2}$	$T_{\cdot 3}$	$T_{\cdot 4}$	$T_{\cdot 5}$	$T_{\cdot\cdot}$	
平 均	$\overline{X}_{\cdot 1}$	$\overline{X}_{\cdot 2}$	$\overline{X}_{\cdot 3}$	$\overline{X}_{\cdot 4}$		$\overline{X}_{\cdot\cdot}$	

肥料를 組合시켰을 때 如斯한 경우에 있어서 收穫高를 나투는 것과 같다. 其測定値가 다음 第四表와 같이 주어져 있다고 하자.

第四表에 있어서 $T._1$은 第一行 測定値의 計이며 $T_1.$은 第一列의 測定値의 計이며 $T..$은 모든 測定側의 總計이다. 또 $\overline{X}i.$은 第 i 行의 平均 $\overline{X}.j$는 第 j 列의 平均 $\overline{X}..$는 總平均이다. 行의 數를 K 列의 數를 γ로 한다.

이것들의 測定値의 分散은 母集團의 分散과 恒常 存在하는 基本的 實驗誤差에 依할 뿐만 아니라 二組의 要因에 依한 差에 依하여서도 惹起 될수 있을 것이다. 그러므로 이것들의 處置를 說明하기 위하여 前節과 같은 模型的인 架空의 例를 들기로 하겠다.

여기 第五表와 같은 data 가 주어졌다

〔第五表〕

	I	II	III	IIII	計
1	7	6	8	7	28
2	2	4	4	4	14
3	4	6	5	3	18
計	13	16	17	14	60

이때 列의 平均에서 얻은 母集團分散의 推定에 使用하는 平方計는 前節의 (4)式과 같이 하여

$$級內平方計=\frac{13^2}{3}+\frac{16^2}{3}+\frac{17^2}{3}+\frac{14^2}{3}-\frac{60^2}{12}$$

$$=303.33-300=3.33$$

行의 平均에서 母集團分散을 推定하기 위하여서는 $\frac{28^2}{4}+\frac{14^2}{4}+\frac{18^2}{4}-\frac{60^2}{12}$

$$=326-300=26.00$$

또 全平方計는 $7^2+2^2+4^2+6^2+\cdots+3^2-\frac{60^2}{12}$

$$=336-300=36.00$$

本計算에 있어서 分母는 分子로서의 數値를 얻기 위하여 덧셈한 個體의 數라고 하는 點을 添言하여 둔다.

分散分析表는 前節과 多少 달라진다.

第六表에 보는 分散分析表와 같이 結果의 分析을 檢討하겠다.

殘差平方計는 行間變動의 平方計와 列間變動의 平方計를 全平方總計에서 빼기 함으로서 얻어진다.

各變動에 있어서 自由度는 行間變動이 (行의數)−1 卽 k−1이며 列間變動은 (列의 數)−1 卽 $\gamma-1$이다. 또 全變動에 있어서는 N−1 卽 kr−1이다. 殘差變動은 全變動과 다른 두個와의 差이지만 이것은 (N−1)−(k−1)−(γ−1)

〔第六表〕

要 因	平方計	自由度	不偏分散
行間變動	26.00	2	13.00
列間 〃	3.33	3	1.11
殘差 〃	6.67	6	1.11
全 變 動	36.00	11	

$=kr-k-\gamma+1=(k-1)(\gamma-1)$과 같이 된다.

換言하여 이것은 行間變動과 列間變動의 各 自由度의 積이다. 또 不偏分散은 各各의 平方計를 그 自由度로서 나눈것이다. 그리하여 行 또는 列의 平均間에 有意한 差가 있는가 없는가의 檢定은 殘差의 不偏分散을 分母로 하고 行間變動 또는 列間變動의 不偏分散을 分子로 하고 分散比 F를 求하면 이것은 自由度가 各各〔$K-1:(k-1)(\gamma=1)$〕及〔$\gamma-1:(k-1)(\gamma-1)$〕의 F—分布를 하므로 前例에서는 行의 有意한 差에 對한 分散比는 $F=\frac{13.00}{1.11}=11.7$ 그런데 棄却域은 $F>F_6^2(0.05)=5.14$故로 相異한 行의 平均間에 差가 없다고 하는 假說을 棄却한다.

또 列에 關해서는 $F=\frac{1.11}{1.11}=1.00$이며 棄却 域은 $F>F_6^3(0.05)=4.76$이므로 行에 있어서의 差와는 獨立으로 列의 平均에는 差가 없다고 하는 假說은 reject 할수 없게 되는 것이다.

二元配置法의 分散分析表의 計算은 위에서 說明한 바와 같지만 이것을 다시 吟味하여 보기로 하겠다.

即 前示한 第六表의 주어진 data에 行과 列의 平均 및 이것들의 平均의 總平均에서의 偏差를 記入하면 다음 第七表에 分析 整理 된바와 같이 된다.

〔第七表〕

	I	II	III	IIII	Ti.	$\bar{X}i.$	$\bar{X}i.-\bar{X}..$
1	7	6	8	7	24	7.0	2.0
2	2	4	4	4	14	3.5	—1.5
3	4	6	5	3	18	4.5	—0.5
T.j	13	16	17	14	60		
$\bar{X}.j$	4,333	5,333	5,667	4,667		5.00	
$\bar{X}.j-\bar{X}..$	—0.667	0.333	0.667	—0.333			

따라서 本表에서 行平均에서의 分散을 推定하면

$$\frac{\gamma\Sigma(\bar{X}i.-\bar{X}..)^2}{k-1}$$

$$=\frac{4}{2}[2.0^2+(-1.5)^2+(-0.5)^2]=13.00$$ 로서

分散分析表에서 얻은 結果에 一致하게 되며 또 列平均으로 부터의 分散의 推定은

$$\frac{k\Sigma(\bar{X}.j-\bar{X}..)^2}{\gamma-1}$$

$= \frac{3}{3}[(-0.667)^2+(0.333)^2+(0.667)^2+(-0.333)^2] \fallingdotseq 1.11$ 로서

이것도 역시 一致하게 되는 것이다. 殘差平方計는 行 及 列의 平均에 依하여 各各의 測定値에 包含되는 差가 消去된 然後의 平方計를 計算하여 求할수 있을 것이다. 여기에서는 第一行의 各各의 數値에서 2.0를 빼기하고 第二行의 各各의 數値에는 1.5를 더하고 第3行의 各各의 數値에는 0.5를 더한다. 이와같은 交換을 하여도 加減한 數値가 相等하므로 列의 平均도 總平均도 不變하게되며 더욱이 各行의 平均은 모다 5가 되어 버린다. 第八表는 이變換을 實施한 경우의 數値表이다 다음에는 같은 方法으로서 列平均이 相等하게끔 測定値를 變換한다. 卽 第一列의 各各의 數値에 0.667를 더하고 第2列에서는 0.333를 빼기하고 第3列에서는 0.667를 빼기하고 第4列에는 0.333를 더하여 주면 된다. 이結果를 表示한것이 第9表이다. 本表에서는 모든 行과 列의 平均이 相等하므로 이것들의 測定値에는 行 及 列 處理의 差는 없게된다. 이것들의 數値의 全平方計를 計算하면 이것이 實驗誤差의 平方計이며 分散分 析表에는 殘差平方計欄에 記錄된 것이며 平均의 差에 對하여 修正한 數値의 分散이다. 計算하면

$$5.667^2+4.167^2+5.167^2+3.667^2+\cdots\cdots+3.833^2-\frac{60^2}{12}=306.67-300=6.67$$

第9表의 12個의 數値는 4個의 列에 對하여 修正되었지만 列의 平均은 $\bar{X}..$가 되지 않아서는 안된다. 또 行의 平均은 역시 $\bar{X}..$이다. 따라서 其平方計의 自由度는 總平均에 對한 1自由度가 減少되는 外에 行에서 2自由度 列에서 3自由度가 減少하게 되므로 12−1−2−3=6 自由度가 되어 이平方計를 6으로 나누면 2개의 要因에 對하여 存在 할지도 알수 없는 偏差에 對하여 獨立한 δ^2의 不偏推定値가 얻어지게 되는 것이다. 그리하여 殘差平方計와 行 및 列의 2個의 平方計의 合計로서의 全平方計가 얻어진다.

卽 6.67+26.00+3.33=36.00

〔第八表〕

	I	II	III	IIII	Ti'.	$\bar{X}$i'.
1	5.0	4.0	6.0	5.0	20	5.0
2	3.5	5.5	5.5	5.5	20	5.0
3	4.5	6.5	5.5	3.5	20	5.0
T.j'	13	16	17	14	60	
X.j'	4,333	5,333	5,667	4,667		5,00

〔第九表〕

	I	II	III	IIII	Ti."	$\bar{X}$i."
1	5,667	3,667	5,333	5,333	20	5.0
2	4,167	5,167	4,833	5,833	20	5.0
3	5,167	6,167	4,833	3,833	20	5.0
T.j"	15	15	15	15	60	
$\bar{X}$.j"	5.0	5.0	5.0	5.0		5.0

至今 以上에서 說明한 二元配置法에 依한 檢定方法을 綜合整理하여 보기로 하겠다.

a. 假說: i. k行의 影響은 0

假說: ii. r列의 影響은 0

(두假說 모다 列, 行의 影響과는 別途로 獨立 檢定되는 것이다)

b. 有意水準: α을 定한다.

c. 統計量 F를 利用한다. 但 假說 i. 에 對해서는 行間變動에 對한 不偏分散과 殘差變動의 不偏分散과의 分散比, 假說 ii. 에서는 列間變動에 依한 不偏分散과 殘差變動의 不偏分散과의 分散比를 使用한다.

d. 測定値가 均齊한 分散을 가진 正規母集團에서 抽出된 것이며 더우이 行 及 列의 影響이 加法的이라고 假定하면 위에서 計算한 統計量 F는 自由度가 各各 $[k-1,\ (k-1)(\gamma-1)$ 及 $[\gamma-1,\ (\gamma-1)(k-1)]$의 F-分布가 된다.

e. 따라서 假說 i의 棄却域은 $F > F^{k-1}_{(k-1)(\gamma-1)}(\alpha)$ 또 假說 ii의 棄却域은 $F > F^{\gamma-1}_{(k-1)(\gamma-1)}(\alpha)$

f. F가 e項의 棄却域에 包含되는가 또는 안되는가에 따라서 假說은 reject or allow 된다.

c. 要因分析法(factorial experiment)

本項에서는 主로 우리가 檢定할려는 主題에 關하여 問題가 될수 있는 要因의 모든 組合을 生覺하고 그것을 同一條件下에 여러번 反覆實驗을 할려고 하는 問題를 研究하겠다. 例를 들며 어떤 農作物에 硫安 加里 堆肥의 세가지 肥料를 施肥하는 경우를 생각하여 보겠다. 硫安의 施肥量을 A와 a의 두가지 加里도 역시 B와b 堆肥도 C와c와 같이 各各 두가지라고 하면 이것들의 모든 組合은

ABC, ABc, AbC, aBC, Abc, abc, abC, abc

의 八가지가 있게 될 것이다. 그리고 同一條件에서도 若干의 變動은 있을 것이므로 이것들을 n回 反覆하여 合計 8n가지의 實驗을 한다. 이로부터 A와a와의 作用을 比較함에 있어서는

ABC,	aBC
ABc,	aBc
AbC,	abC
Abc,	abc

을 對應시켜 双方 모다 4n가지式 比較하는 것이 된다. 따라서 其收量을 같은 文字로서 表示하기로 하면 A와a의 比較는 式

$$(ABc-aBC)+(ABc-aBc)+(AbC-abc)+(Abc-abc)$$

로서 주어지게 된것이다. 또 B와 b, C와c도 역시 上式의 文字를 變換함으로써

容易하게 얻어 지게 될것이다.

그리고 硫安과 加里와의 交互作用을 檢定할 수도 있다. 卽 A와a와의 比較를 B와b와의 경우에 나눠서

B	b
ABC−aBC	AbC−abC
ABc−aBC	Abc−abc

를 比較하면 좋을 것이다. 그 收量을 같은 文字로서 表示하면 式

$$\{(ABC-aBC)+(ABc-aBc)\}-\{(AbC-abc)+(Abc-abc)\} \quad \cdots\cdots(1)$$

에 依하여 硫安과 加里의 交互作用을 檢定할수가 있다. 式(1)에서 項의 順序를 바꿔넣기를 하여

$$\{(ABC-aBC)+(ABc-Abc)\}-\{(aBc-abC)+(aBc-abc)\}$$

라고 하면 B와b의 作用이 A와a에 依하여 如何히 變動하는가 하는 것을 말하는 式이며 換言하면 加里와 硫安과의 交互作用이 된다. 卽 交互作用은 加里와 硫安이 래도 硫安과 加里라고 해도 같은 것이며 三要因의 경우에는 二要因式의 交互作 用(이것을 二因子交互作用)이 結局 세가지로 생각 될수 있다.

다음에 硫安과 加里의 影響이 堆肥의 多少에 依하여 如何히 變動하는가 하는 問題를 硏究해보면 꼭 같은 方法으로 要因을 整理함으로써 그에 依하여

$$\overbrace{\underset{ABC-aBC}{B}\qquad\underset{AbC-abC}{b}}^{C}\qquad\overbrace{\underset{ABc-aBc}{B}\qquad\underset{Abc-abc}{c}}^{c}$$

이므로 이와같은 三者의 交互作用은

$$\{(ABC-aBC)-(AbC-abC)\}-\{(ABc-aBc)-(Abc-abc)\} \quad \cdots\cdots(2)$$

로서 表示할수 있다.

(2)式도 역시 項의 順序를 바꿔 넣으면 加里와 堆肥의 硫安의 多少에 依한 作用 堆肥와 硫安의 加里의 多少에 依한 影響으로도 될수 있으며 이 세가지 要因全部에 亘한 交互作用은 單 한가지만 있을 것이다(三因子 交互作用)

이와같이 세가지 要因에 關하여 各各 두가지 경우를 取扱한 것을 $2\times2\times2$(or 2^3)要因分析法이라고 하며 이比較는 前述한 바와 같이 一要因의 主効가 세가지 二因子交互作用이 세가지 三因子交互作用이 한가지 合計 일곱가지가 된다.

다음에 2^3要因分析法에 關하여 實際의 例를 設定하여 說明하겠다.

〔例〕 水稻甲種에 關한 肥料試驗에서 各反當 硫安 1貫 5百을 施肥하였을때(A) 加里 3貫7百을 施肥하였을 때(B) 堆肥 530貫을 施肥하였을 때(C), 또 이것들을 全혀 施肥하지 않았을 때를 各各 a. b, c로 하고 640坪의 試驗水田에 이것을 4個區劃으로 나눠 各各의 區域에 20坪式 8試驗區를 取하여 要因 組合 여덟가지를 亂數表에 依하여 任意로 配置하였다. 其配列과 收量은 第十表에 記錄된 것과 같

다. 여기서는 abc의 組合은 1로 하고 其以外의 肥料가 施肥되지 않은 경우는 記錄되어 있지 않다. 例하면 ABc는 AB aBC는 BC라고 記錄되어 있다.

〔第十表〕

Ⅰ	AB 39.4	BC 53.9	C 42.3	AC 50.5	BC 55.1	C 43.9	B 36.8	AB 41.5	Ⅱ
	1 13.7	B 35.9	A 14.4	ABC 73.2	A 12.1	ABC 60.8	AC 45.8	1 14.4	
Ⅲ	C 43.8	1 11.8	AC 43.9	BC 57.3	AC 48.9	AB 36.8	A 14.0	C 43.9	Ⅲ
	AB 45.2	B 37.8	A 17.3	ABC 63.8	B 40.9	1 17.7	ABC 59.2	BC 60.3	

〔解說〕

本表에서 處理의 効果를 檢定하기 위해서는 먼저 第十一表를 만들고 二元配置法과 같은 分散分析表를 作成한다.

卽

全平方計 $=13.7^2+14.4^2+35.9^2+\cdots\cdots 59.2^2-\dfrac{1276.3^2}{32}$

$=60132.25-50904.43$

$=9227.82$

區域間平方計 $=\dfrac{1}{8}(323.3^2+310.4^2+320.9^2+321.7^2)-\dfrac{1276.3^2}{32}$

$=50917.34-50904.43=12.91$

處理間平方計 $=\dfrac{1}{4}(57.6^2+57.8^2+151.4^2+\cdots\cdots+257.02^2)-\dfrac{1276.3^2}{32}$

$=59878.39-50904.43=8973.96$

〔第十一表〕

處理＼區域	Ⅰ	Ⅱ	Ⅲ	Ⅲ	Ti.
1	13.7	14.4	11.8	17.7	57.6
A	14.4	12.1	17.3	14.0	57.8
B	35.9	36.8	37.8	40.9	151.4
AB	39.4	41.5	45.2	36.8	162.9
C	42.3	43.9	43.8	43.9	173.9
AC	50.5	45.8	43.9	48.9	189.1
BC	53.9	55.1	57.3	60.3	226.6
ABC	73.2	60.8	63.8	59.2	257.0
T.j	323.3	310.4	320.9	321.7	1276.3

따라서 第十二表와 같은 分散分析表를 얻을 수 있다. 本表에 依하면 區域間의 變動은 有意하지 않으나 處理間의 變動은 有意하다고 하는것이 明白하다.

〔十二表〕

變動因	平方計	自由度	不編分散
區域間	12.91	3	4.30
處理間	8973.96	7	1281.99
殘 差	240.95	21	11.47
全變動	9227.82	31	

여기서 이 有意性이 무엇에 起因하는가를 檢討하기로 하겠다.

前示한바 處理間平方計의 自由度 7을 主効 3, 交互作用 3, 三因子交互作用 1, 로 再構成하고자 하는 것이다.

卽 變動因A에 關한 平方計는 第十一表의 Ti.의 列이며 A의 成分이 包含되어 있는 것과 그렇지 않은 것의 差를 平方하여 32로 나눔(除)으로써 얻어진다. 記號的으로는

$$\frac{1}{32}(-1+A-B+AB-C+AC-BC+ABC)^2 \cdots\cdots\cdots\cdots(3)$$

(3)式의 括孤內를 形式的으로 因數分解하면

$$(A-1)(B+1)(C+1) \cdots\cdots\cdots\cdots(4)$$

實際로 第十一表의 Ti.의 列에서 數値를 求하면

$$\frac{1}{32}(-57.6+57.8-151.4+162.9-173.9+189.1-226.6+257.0)^2$$

$$=\frac{1}{32}57.3^2 \fallingdotseq 102.60$$

B, C에 關한 平方計도 같은 輪環的인 것이므로 形式的으로는 各各

$$(A+1)(B-1)(C-1),\ (A+1)(B+1)(C-1) \cdots\cdots\cdots\cdots(5)$$

(5)式을 展開함으로써 얻어지는 式에 따라 第十一表의 數値에 依하여 求하며 이것을 平方하여 全體의 數를 32로 나누면(除)된다. 二因子交互作用도 역시 形式的으로는 各各

$$(A-1)(B-1)(C+1),\ (A-1)(B+1)(C-1),\ (A+1)(B-1)(C-1) \cdots(6)$$

을 展開한 것이다. 이에 依하여 역시 第十一表에 所要數値를 求하고 平方하여 32로 나누면 된다. 이와같이 하여 三因子交互作用은 (2)式에서 얻어지지만 이것도 括孤內는 式

$$(A-1)(B-1)(C-1) \cdots\cdots\cdots\cdots(7)$$

을 展開한 것이다. 역시 平方하여 32로 나누면 좋다.

以上 各式을 展開한 것의 係數는 모다 ±1이므로 그符號는 第十三表와 같다.

〔第十三表〕

効果 \ 項	1	A	B	AB	C	AC	BC	ABC
A	−	+	−	+	−	+	−	+
B	−	−	+	+	−	−	+	+

AB	+	−	−	+	+	−	−	+
C	−	−	−	−	+	+	+	+
AC	+	−	+	−	−	+	−	+
BC	+	+	−	−	−	−	+	+
ABC	−	+	+	−	+	−	−	+

이와같이 하여 處理間의 平方計를 各 要因의 効果別로 分割하면 第十四表가 얻어지는 것이다. 이것을 第十二表 殘差의 分散에 比較하여 其 比가 棄却域 F^{1}_{21} (α)를 超過하면 有意하다는 것이 된다. 危險率이 5%로서는 棄却域은 4.32 1%는 8.02이므로

〔第十四表〕

變動因	平方計	自由度
A	102.60	1
B	3190.01	1
AB	21.95	1
C	5431.43	1
AC	35.91	1
BC	191.59	1
ABC	0.47	1
處理間	8973.96	7

i) 5%의 경우는

$$11.47 \times 4.32 = 49.55$$

ii) 1%의 경우는

$$11.47 \times 8.02 = 91.99$$

를 判定水準으로 하면 可할 것이다. 本例로서는 A, B, C의 主要 및 BC의 交互作用만이 有意하며 其他 交互作用은 有意하지 않음이 自明하다.

V. 簡單한 結論

要는 實驗計劃法은 R, A, Fisher에 依하여 歷史的으로 創始된 새로운 學問으로서 近代數理統計學이 K. Pearson의 記述統計學的 限界 卽 Pearson 系分布曲線 相關係數 回歸係數 百分位數等 理論的 바탕으로서는 現象의 本質을 認識할 수 있는 學問的限界가 이미 스스로 沈滯되어 있으므로 Pearson의 學的影響에서 Pearson의 學問世界를 克服하고 全혀 새로운 理論構造를 創造的으로 體系化한 統計推理의 方法으로서만이 現象에 關한 本質을 的確하게 認識할 수 있게 되었다.

이와같은 學問的推移를 우리들은 Fisher의 實驗論理의 構造를 本稿에서 槪觀함으로써 그輪廓을 考察하였던 것이다.

簡約하게 實驗計劃理論을 說明하면 前述한 Fisher의 三大原理에 立脚하여 먼저 現象의 構造를 設定하고 가장 믿음성있는 推理를 하기 위하여 實驗에 關與될 것으로 믿어지는 因子를 如何히 다룰 것인가 하는 問題를 硏究하고 그方法으로서 實驗에 先行하여 構造模型을 設定하고 다음에 確率模型을 想定하는 것이다.

卽 歸無假說을 設定하고 構造模型 確率模型에 可能한 限 가장 가깝게 接近할 수 있도록 實驗條件을 設定하여 實驗을 實施한 후 여러가지 分析法을 使用하여 檢定하는 것이다.

이렇듯 實驗計劃法이 推測統計學의 理論的 back-born 로서 이에 體系化된 推測統計學은 그歷史的 過程으로 보아 스스로 重要한 學問的 課題를 自覺하여야 하겠다.

끝으로 一言하고저 하는 問題는 推測統計學的 方法의 計量經濟學에의 導入인데 이 問題는 經濟計劃의 可能性과 必然性과의 有機的 關聯性위에서 만이 把握되어야 할 것이지만 여기에는 現代統計學의 三大原理인 ㄱ. 有限性의 原理 ㄴ. 逐次推測의 原理 ㄷ. 方略의 原理等 原理的 基礎위에서 特히 方略論의 展開가 必要하며 이때 經濟行爲의 結果를 確率變數라고 規定하여도 좋을 것이다.

이와같은 原理를 通하여 推測統計學이 計量經濟學에로의 接近 그有機的連關위에 새로운 經濟學의 一側面이 開拓되지 않을까 하는 可能性을 우리는 期待하며 操心性있게 注視하고저 한다.

〔附記〕

本稿에 參考된 書籍을 參考로 紹介하여 보겠다.

1.	統計學辭典	東洋經濟新報社
2.	確率論	河田敬義 （共立）
3.	近代確率論	國澤淸典 （岩波）
4.	數理統計學要說	成實淸松，坂井忠郎共著 （掌華房）
5.	統計學要論	有澤廣己 （明善社）
6.	統計學汎論	森田優三 （日本評論）
7.	統計學	內 藤 勝 （東大出版）
8.	新統計學	安川數大郞(共立出版)
9.	L. H. C. Tippett	Statistics (Home University Library 1943)
10.	R. A. Fisher	Statistical methods for Research Workers (1950)
11.	H. Westergard	Contributions to the History of Statistics (1932)

<SAMMARY>

Regarding The Development of The "The Design of Experiments" by R. A. Fisher

The design of experiment can be defined as a method of sampling. It has been employed and developed in the Rothamsted Experimental Station for agriculture as a technique of agricultural experiment in the later nineteenth century and was elaborated and introduced by R.A. Fisher into the field of study where experiments are done.

In earlier days, the random sampling method has exclusively been used in the Pearsonian descriptive statistical method but the achievements of Fisher made it clear that the discriptive method is very imperfect one. That is to say, in the Pearsonian system, the descriptive method holds so long as the universe were to be homogeneous. In the practical case, however, the universe we ar ealways facing with is, in most cases, not homogeneous but heterogeneous and under these circumstances we can hardly grasp the characteristics of the phenomena by means of descriptive method with any precision. Henceforth the method of inductive statistics was deviced and developed. The industive statistics has something in its theoretical construction quite different from that of descriptive statistics.

As, in many practical cases, the universe is constituted from heterogeneous elements, the precision can be expected to be improved if we classify the heterogenous univrse into a number of homogeneous strata, and select sampling from this group than to select it at random from the universe.

On this account, we are obliged to be in need of the universe. When we classify the universe into a number of strata and the frequency of observation is apportioned to these strata, the number of samples selected is hoped to be proportionate to the size of the universe.

The main theme of this paper is to examine the above-said technique of sample design in sampling method. The design of experiment of R.A. Fisher, in effect, is the core in the problems of modern statistic's and Fisher's achievements in this field is a monumental contribution to the systematication of statistics together with the analysis of variance and these theory of estimation. It constitutes the theoretical backbone of inductive statistics together with Neyman-Pearsonian Theory of testing hypothesis.

Modern statistics (industive statistics) is originated from the classical works Gosset and is developed through E.S. Pearson J. Neyman to R.A. Fisher and the studies in the theory and application of its method is enriched and refined by the works of E.S. Pearson, J. Neyman and A. Wald.

As the application of the design of experiment is to the field of the control of experiment and mass production, it is an inevitable outcome that Fisher's design of experiment came to qualify the methodology of control.