

「臨床データの信頼性と妥当性」

補遺： Excel と JMP による解析

芳賀 敏郎

目次

| | | |
|-----|---------------------------|----|
| 0 | まえがき | 1 |
| 1 | 連続データにおける信頼性：級内相関係数 (ICC) | 2 |
| 1.1 | 繰返し2回の場合 | 2 |
| (1) | データと散布図 | 2 |
| (2) | ICC の計算 | 3 |
| (3) | 相関係数の変形 (補足) | 4 |
| (4) | データのグラフによる検討 (補足) | 5 |
| (5) | x, y の回帰式 (補足) | 6 |
| 1.2 | くり返し3回以上の一般の場合 | 8 |
| (1) | データと折れ線グラフ | 8 |
| (2) | ICC の定義による計算 | 8 |
| (3) | ICC の直接計算 | 9 |
| (4) | 1元配置データの解析 | 10 |
| (5) | ICC と ρ の関係 (補足) | 12 |
| (6) | ρ の区間推定 | 13 |
| (7) | 検出力の計算と実験の大きさの決め方 | 14 |
| (8) | 外れ値の検出 (補足) | 17 |
| (9) | JMP による解析 | 18 |
| 1.3 | 2元配置データの解析 | 21 |
| (1) | データと ICC の計算 | 21 |
| (2) | 分散分析と ρ の計算 | 22 |
| (3) | JMP による解析 | 24 |
| (4) | 別の例の解析 | 24 |

ii 目次

| | | |
|----------|-------------------------|-----------|
| 2 | 分類データにおける信頼性： 係数 | 27 |
| 2.1 | くり返し2回の場合 | 27 |
| (1) | Excel 関数による解析 | 27 |
| (2) | VBA マクロによる解析 | 29 |
| (3) | JMP による解析 | 30 |
| 2.2 | 繰返し3回以上の一般の場合 | 31 |
| (1) | データ | 31 |
| (2) | VBA マクロによる解析 | 31 |
| (3) | 別の例 | 32 |
| 2.3 | 順序分類データの解析 | 33 |

0 まえがき

標題の本を手に入れ、目を通して、これはすばらしい本だと感じた。

最近の解説書は、結果だけを示しているものが多い。読者はその手法の根拠を知らないまま、市販の統計解析プログラムにデータを入力して、その出力を鵜呑みにして報告書にまとめる。そのために、不適切な適用や誤用が少なくない。この本は、手法だけでなく、その根拠についても詳しく述べており、この本を熟読して適用すれば上のような誤用は生じないであろう。

私は大学では化学を専攻して、統計の基礎教育を受けていない。従って、数式の展開による解説を理解するのは苦手である。そこで、この種の本を読んで、それを自分のものにするために、常にやる方法は「自分で計算して確かめる」である。計算には通常 Excel と JMP を用いる。

この本の読者の中には、私と同じような方が少なくないと思われる。そこで、私がこの本を理解するために試行錯誤した過程を公開することは、この本の読者にも役立つのではないかと思い、「補遺」としてまとめた。

以下、「臨床データの信頼性と妥当性」を原著と呼ぶことにする。

なお、このメモは、原著を購入した人にもみ配布することにする。原著と照らし合わせながら読んで頂きたい。したがって、原著に書かれていることは、必要最小限とし、原則として、原著のページを示して、省略する。

原著の図表は 図_{x.x}、表_{x.x} として引用し、補遺の図表は 図 と 表 を区別しないで、表示_{x.x} とする。

筆者は、データを見たら先ずグラフ化して眺めるというのがデータ解析の基本であると考え、自分でも忠実に実行している。この資料でも、原著には示されていないグラフがいくつか示され、グラフによる検討が付記されている。

添付される Excel ファイルで一つ一つ確かめながら読んで頂きたい。Excel シートの左上の「名前ボックス」をクリックすると、プルダウンメニューに Fig_{x.x} が表れる。見たい表示番号をクリックすると、テキストの該当部分が表示される。

資料の標題の「補遺」の範囲を超える部分は「補足」または「蛇足」とした。適当に読み捨てて頂きたい。そこには、原著と私の意見が合わない個所についてのコメントも述べられている。

第I部 信頼性研究

1 連続データにおける信頼性：級内相関係数(ICC)

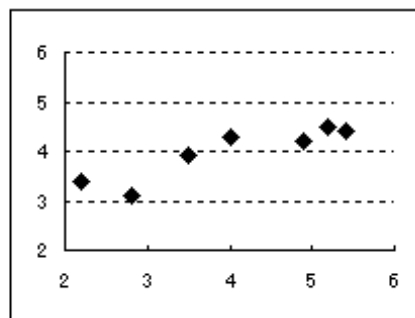
1.1 繰返し2回の場合

(1) データと散布図

原著 表1.1 (p.3) のデータを100で割った値を表示1.1の右に示す。

表示1.1: データと散布図

| | x1 | x2 |
|------|-------|-----|
| 1 | 2.8 | 3.1 |
| 2 | 5.4 | 4.4 |
| 3 | 4.0 | 4.3 |
| 4 | 4.9 | 4.2 |
| 5 | 5.2 | 4.5 |
| 6 | 2.2 | 3.4 |
| 7 | 3.5 | 3.9 |
| 相関係数 | 0.897 | |



原著では、1回目、2回目の測定値を x, y で表しているが、次節以降に繰返し回数を3回以上に拡張することを考慮して、 x_1, x_2 で表すことにする。 i 番目の患者の j 回目の測定値は x_{ij} である。

表示1.1の右は、 x_1 を横軸に、 x_2 を縦軸に取って描いた散布図(図1.2)である(縦軸と横軸を入替えた散布図は表示1.4に示す)。

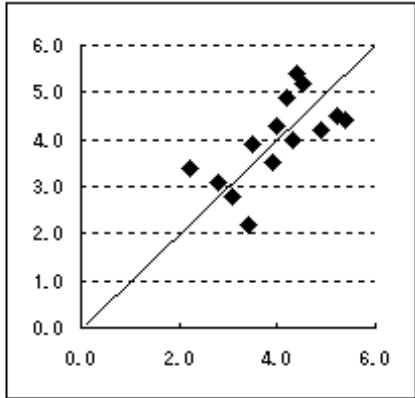
相関係数 0.897 が左下に =CORREL(x1, x2) で求められている。

(2) ICCの計算

表示1.1のデータの下に、左右を入替えた表を追加したのが表示1.2の u , v の列である。 $u-v$ の列は後の説明に用いられる。

表示1.2: u , v , $u-v$ の表, 相関係数と散布図

| i | u | v | u-v |
|------|--------|--------|-------|
| 1 | 2.8 | 3.1 | -0.3 |
| 2 | 5.4 | 4.4 | 1.0 |
| 3 | 4.0 | 4.3 | -0.3 |
| 4 | 4.9 | 4.2 | 0.7 |
| 5 | 5.2 | 4.5 | 0.7 |
| 6 | 2.2 | 3.4 | -1.2 |
| 7 | 3.5 | 3.9 | -0.4 |
| 8 | 3.1 | 2.8 | 0.3 |
| 9 | 4.4 | 5.4 | -1.0 |
| 10 | 4.3 | 4.0 | 0.3 |
| 11 | 4.2 | 4.9 | -0.7 |
| 12 | 4.5 | 5.2 | -0.7 |
| 13 | 3.4 | 2.2 | 1.2 |
| 14 | 3.9 | 3.5 | 0.4 |
| 平均 | 3.99 | 3.99 | 0.00 |
| 平方和 | 10.857 | 10.857 | 7.520 |
| 積和 | 7.097 | | |
| 相関係数 | 0.654 | | |
| b | 0.654 | | |



右に u と v の散布図が求められている。

表の下に表示1.1と同様に、`=CORREL(u, v)` で相関係数 0.654 が計算されており、原著のp.7のICCの値と一致する。

この相関係数は次の式で計算される。

$$r = \frac{\sum_i (u_i - \bar{u})(v_i - \bar{v})}{\sqrt{\sum_i (u_i - \bar{u})^2 \sum_i (v_i - \bar{v})^2}} = \frac{S_{uv}}{\sqrt{S_{uu}S_{vv}}}$$

u , v の添え字の i は患者番号ではなく、表示1.2のデータ表の行番号である。

S_{uu} , S_{vv} は平方和, S_{uv} は積和である。

平方和 S_{uu} , S_{vv} が表示1.2の平方和の行に`=DEVSQ(u)`, `=DEVSQ(v)` で求められている。当然のことながら, 2つの平方和 S_{uu} , S_{vv} は等しい。

積和を求めるExcel関数はないので, 定義に従って

`=SUMPRODUCT(u-uの平均, v-vの平均)`

`=SUMPRODUCT(B17:B30-B31, C17:C30-C31)`

として計算し, 7.097 が得られる。

4 1 連続データにおける信頼性：級内相関係数(ICC)

相関係数は

$$r = \frac{S_{uv}}{\sqrt{S_{uu}S_{vv}}} = \frac{S_{uv}}{S_{uu}} = \frac{7.097}{10.857} = 0.654$$

として得られる。

相関係数の下の傾斜 b の欄には、 $\hat{v} = a + bu$ の傾斜 b が Excel の `=SLOPE(vの値, uの値)` で求められており、これは相関係数に等しいことが分かる。これは、原著 p.9 の下に書かれていることの数値による確認である。

一般化すると、「相関係数 r は x, y の分散を等しくしたとき（平均と標準偏差で基準化したときを含む）の傾斜に等しい」と言うことができる。

(3) 相関係数の変形（補足）

u と v の差を d で表す $d_i = u_i - v_i$ 。

表示 1.2 の $u - v$ の列に d が求められている。 d の平均 0 と、平方和 7.520 が下に求められている。

d の平方和 S_{dd} は次のように変形できる。

$$\begin{aligned} S_{dd} &= \sum_i (d_i - \bar{d})^2 = \sum_i ((u_i - v_i) - (\bar{u} - \bar{v}))^2 = \sum_i ((u_i - \bar{u}) - (v_i - \bar{v}))^2 \\ &= \sum_i (u_i - \bar{u})^2 + \sum_i (v_i - \bar{v})^2 - 2 \sum_i (u_i - \bar{u})(v_i - \bar{v}) \\ &= S_{uu} + S_{vv} - 2S_{uv} \end{aligned}$$

これから、積和 S_{uv} は

$$S_{uv} = \frac{S_{uu} + S_{vv} - S_{dd}}{2}$$

として求めることができる。

$S_{uu} = S_{vv}$ であるから、

$$S_{uv} = S_{uu} - \frac{S_{dd}}{2} = 10.857 - \frac{7.520}{2} = 7.097$$

となり、相関係数は

$$r_{uv} = \frac{S_{uv}}{S_{uu}} = \frac{S_{uu} - S_{dd}/2}{S_{uu}} = 1 - \frac{S_{dd}}{2 \times S_{uu}} = 1 - \frac{7.520}{2 \times 10.857} = 1 - 0.346 = 0.654$$

として求めることができる。

ここで、 S_{uu} 、 S_{dd} の意味を考える。

S_{uu} は表示 1.2 の u の列の 14 個の値の平方和である。この 14 個の値は、 x_1, x_2 を合わせたものであるから、

$$S_{uu} = S_{xx} = \sum_{i=1}^7 \sum_{j=1}^2 (x_{ij} - \bar{x}_{.})^2$$

すなわち、測定値全体の平方和 S_{xx} である。

S_{dd} は $u - v$ の列の平方和である。 $u - v$ の平均値は 0 であるから（偏差）平方和ではなく、単純な 2 乗和である。また、上半分と下半分は u, v を入替えたものであるから、 $u - v$ の絶対値は同じで符号を逆にしたものである。したがって、上半分の 7 個を 2 乗して加えて、2 倍して求めることができる。これは、

$$S_{dd} = 2 \sum_{i=1}^7 (x_{i1} - x_{i2})^2$$

としても良い。

これから、ICC は、表示 1.3 の左に示すように、 $x_1, x_2, x_1 - x_2$ の表から

$$\text{ICC} = r_{uv} = 1 - \frac{S_{dd}}{2 \times S_{uu}} = 1 - \frac{\sum_{i=1}^7 (x_{i1} - x_{i2})^2}{S_{xx}} = 1 - \frac{3.760}{10.857} = 0.654$$

と、直接求められることが分かる。

ICC は、1 から 不一致度 を引いたものであると解釈することができる。この解釈は、§2 で ICC の意味を考えるとときに役立つであろう。

(4) データのグラフによる検討（補足）

x_1 と x_2 の関係を Excel の折れ線グラフで描くと、表示 1.3 右のグラフが得られる。

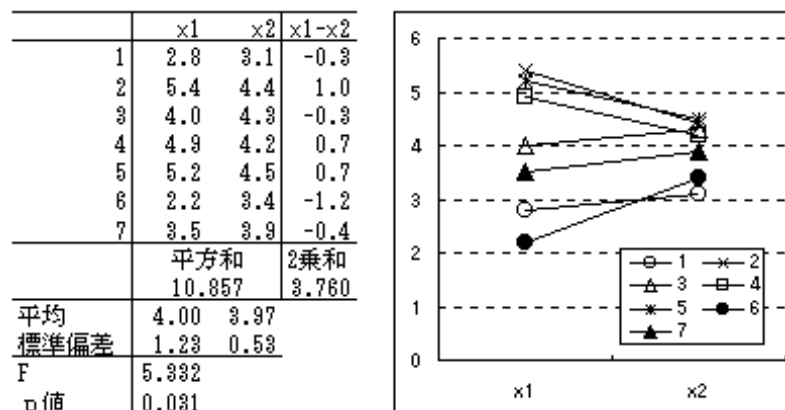
このグラフから、2 回目の測定値 x_2 は中心に集まっているように見える。

この傾向を数値で確かめるため、表示 1.3 の左に、 x_1, x_2 の平均と標準偏差が計算されている。平均値は 4.00, 3.97 とほぼ等しいにもかかわらず、標準偏差は 1.23, 0.53 と 2 回目の標準偏差は 1 回目の標準偏差の半分以下になっている。

標準偏差間に有意差があるかどうかを、分散比 F の検定で確かめる。 F 比を求めると $F = (1.23/0.53)^2 = 5.332$ で、上側 p 値は 0.031 となる。これから両側検定では有意とならないが、かなり有意に近い。

このような傾向が見られるときには、その原因について多面的に検討する必要があるであろう。

表示1.3: データと折れ線グラフ



たとえば、「患者は他の患者の検査を見ており，自分が他の人よりも値が小さいことを知ったとき，第2回には頑張る，逆に第1回目の検査値が大きい患者は，第2回には手を抜く」というようなことが考えられる．

このように，あらゆる可能性を追求して，現象の裏にあるモデルを追求することによって自然科学は進歩してきた*1．医薬開発の分野でもこのような姿勢は必ず役立つであろう．

(5) x, y の回帰式 (補足)

原著の図1.2 (p.4) の散布図には，回帰式 $y = 241.8 + 0.388x$ が求められ，本文にも引用されている．

本来回帰式は

$$y_i = \mu_i + \varepsilon_{ij} = \alpha + \beta x_i + \varepsilon_i$$

というモデルが成立する場合に意味がある．すなわち， x が原因で， y は x によって直線的に変化する μ_i に誤差 ε_i が加わったものであると考える．

上の挙げた例はこのようなモデルは成立しない．どちらを横軸に取るかも一意には決まらない．

そこで，縦軸と横軸を変えて，2つの散布図を描き，回帰直線を当てはめたのは表示1.4の左である．

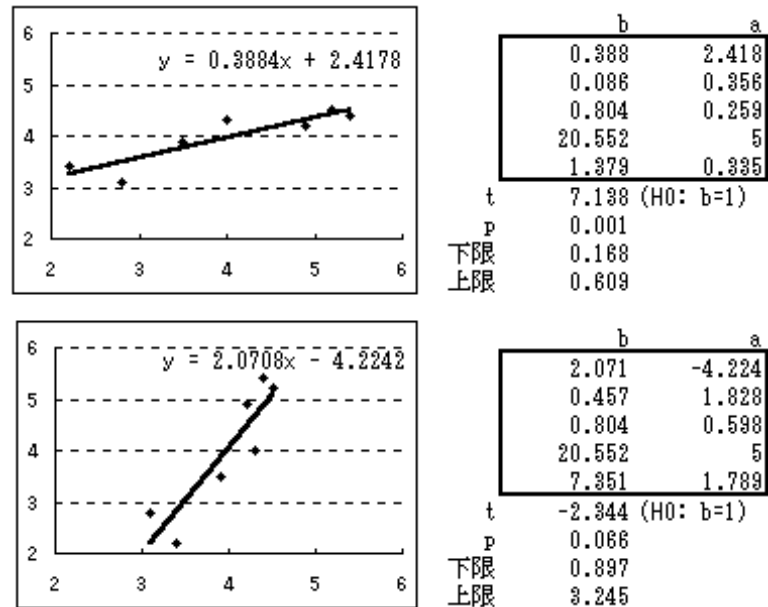
傾斜は 0.388, 2.071 と 1.0 からかなり外れている．

*1 このようなアプローチについては，次の本が参考になるであろう．

石居進「カエルの鼻」，八坂書房（1997）

第**回 医薬安全性研究会で，著者による内容紹介があり，好評を得た本である．

表示1.4: x, y の散布図, 傾斜の検定と区間推定



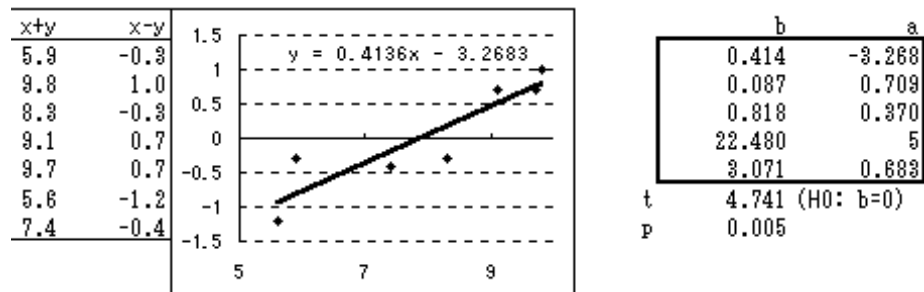
LINEST 関数で回帰式を当てはめ、傾斜 b とその標準誤差を求める。その結果から、 $H_0: \beta = 1$ を検定する t 値と p 値を計算すると、0.001, 0.066 と一方は有意、他方は有意ではない。

このような違いが出るのは、そもそも回帰式を当てはめるのが不適切であるためである。

このような場合、 $x + y$ と $x - y$ の散布図を描き、 $\beta = 0$ を検定することが考えられる。

表示1.5 はその結果である。

表示1.5: $x + y, x - y$ の散布図, 傾斜の検定



p 値は 0.005 となった。これから、第1回の測定値と第2回の測定値の間にはなんらかの関係があるのではないかと考えられる。

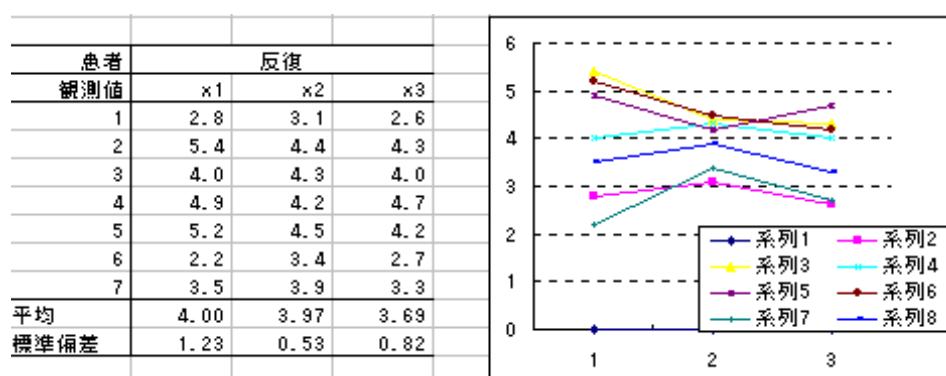
1.2 くり返し3回以上の一般の場合

この節は，前節の拡張である．

(1) データと折れ線グラフ

原著 p.12 のデータを 100 で割った値を表示 1.6 の左に示す．

表示 1.6: データと折れ線グラフ



また，右に §1.1 (4) に対応する折れ線グラフを示す．

(2) ICC の定義による計算

3回の反復測定値を x_1, x_2, x_3 で表わす．それらは，縦に7個の測定値が並んでいる．

原著の図 1.4 (p.14) を描くために，元のデータを表示 1.7 左のように， x を配列して u, v の表を作成する．

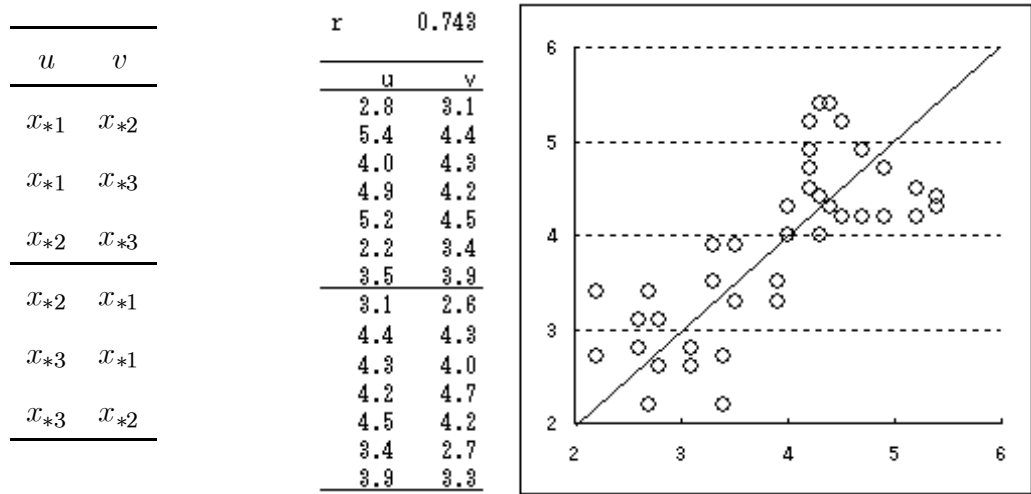
1行目の x_{*1}, x_{*2} は，表示 1.3 の x_1, x_2 の $i = 1 \sim 7$ の7個の値である．

下半分は 上半分の 左右を入れ替えたものである．縦に $6 \times 7 = 42$ 個の測定値が並んでいる．その一部を表示 1.7 の中央に示す．

このデータから散布図を描くと，表示 1.7 右の散布図 (図 1.4) が得られる．

u, v のデータ表から， u と v の相関係数を計算すると，0.743 が得られる．

表示1.7: u, v のデータ表と散布図



(3) ICC の直接計算

くり返し数が3回であれば, x_1, x_2, x_3 のデータ表を u, v のデータ表に展開するのはそれ程煩雑ではないが, 4回以上になると, 大変である.

そこで, §1.1 (3) で説明した算式を使って, 元のデータ表から ICC を計算する方法を示す.

表示1.8: ICC を直接計算するための計算表

| | x_1 | x_2 | x_3 | x_1-x_2 | x_1-x_3 | x_2-x_3 | 平方和 |
|-----|-------|--------|-------|-----------|-----------|-----------|-------|
| 1 | 2.8 | 3.1 | 2.6 | -0.3 | 0.2 | 0.5 | 0.127 |
| 2 | 5.4 | 4.4 | 4.3 | 1.0 | 1.1 | 0.1 | 0.740 |
| 3 | 4.0 | 4.3 | 4.0 | -0.3 | 0.0 | 0.3 | 0.060 |
| 4 | 4.9 | 4.2 | 4.7 | 0.7 | 0.2 | -0.5 | 0.260 |
| 5 | 5.2 | 4.5 | 4.2 | 0.7 | 1.0 | 0.3 | 0.527 |
| 6 | 2.2 | 3.4 | 2.7 | -1.2 | -0.5 | 0.7 | 0.727 |
| 7 | 3.5 | 3.9 | 3.3 | -0.4 | 0.2 | 0.6 | 0.187 |
| | 平方和 | 15.346 | | 2乗和 | 7.880 | | 2.627 |
| ICC | 0.914 | | | | | | |

表示1.8の左の3列には, 元のデータ x_1, x_2, x_3 が入力され, 下に全体の平方和 $S_{xx} = 15.346$ が求められている.

右には差 $x_1 - x_2, x_1 - x_3, x_2 - x_3$ を求める. これらの要素全体の2乗和 (平方和ではないことに注意) が2乗和の欄に $S_{dd} = 7.880$ と求められている.

表示1.7の左の表で, u の列には x_1, x_2, x_3 が2つずつある. この2は, 列数 $p = 3$ から1を引いたものである.

これから，ICC は

$$\text{ICC} = 1 - \frac{S_{dd}}{(p-1) \times S_{xx}} = 1 - \frac{7.880}{2 \times 15.346} = 1 - 0.257 = 0.743$$

として求められる．この値は， u, v の表から計算した相関係数の値に等しい．

表示1.8 の右の 平方和 は，一人の患者の3回の測定値の繰返し誤差の平方和である．これを縦に足すと 2.627 となる．この値を3倍すると 差の2乗和 $S_{dd} = 7.880$ に等しい．

この関係を用いると，表示1.9 で x の差の列が不要となり，ICC は次の式で計算される．

$$\text{ICC} = 1 - \frac{p \times \text{繰返し誤差の平方和}}{(p-1) \times S_{xx}} = 1 - \frac{3 \times 2.627}{(3-1) \times 15.346} = 0.743$$

補足 上の関係は次のようにして導かれる． x の個数を n とする．

$$\begin{aligned} \sum_{i=1}^n \sum_{i'=1}^n (x_i - x_{i'})^2 &= \sum_{i=1}^n \sum_{i'=i}^n ((x_i - \bar{x}) - (x_{i'} - \bar{x}))^2 \\ &= \sum_{i=1}^n \sum_{i'=i}^n ((x_i - \bar{x})^2 + (x_{i'} - \bar{x})^2 - 2(x_i - \bar{x})(x_{i'} - \bar{x})) \\ &= 2n \sum_{i=1}^n (x_i - \bar{x})^2 = 2nS_{xx} \end{aligned}$$

2行目の式の最後の項は

$$\sum_{i=1}^n \sum_{i'=i}^n (x_i - \bar{x})(x_{i'} - \bar{x}) = \sum_{i=1}^n (x_i - \bar{x}) \underbrace{\sum_{i'=i}^n (x_{i'} - \bar{x})}_{=0}$$

となり，消える．

最初の $\sum_{i=1}^n \sum_{i'=1}^n (x_i - x_{i'})^2$ で， $i = i'$ の組み合わせは 0 になり， $x_1 - x_2$ が2回含まれるので， $\sum_{i=1}^{n-1} \sum_{i'=i+1}^n (x_i - x_{i'})^2$ は，上の式の半分になる．

$$\sum_{i=1}^{n-1} \sum_{i'=i+1}^n (x_i - x_{i'})^2 = nS_{xx}$$

(4) 1元配置データの解析

表示1.6 のデータを1元配置データとして解析する．

計算の過程を表示1.9 に示す．

計算の手順は次の通りである．

- 右上のデータ表で各患者の平均 \bar{x}_i と総平均 $\bar{x}..$ を計算する．

表示1.9: 1元配置分散分析と の計算

| 患者 | 反復 | | | |
|----|-----|-----|-----|-----|
| x | 1 | 2 | 3 | 平均 |
| 1 | 2.8 | 3.1 | 2.6 | 2.8 |
| 2 | 5.4 | 4.4 | 4.3 | 4.7 |
| 3 | 4.0 | 4.3 | 4.0 | 4.1 |
| 4 | 4.9 | 4.2 | 4.7 | 4.6 |
| 5 | 5.2 | 4.5 | 4.2 | 4.6 |
| 6 | 2.2 | 3.4 | 2.7 | 2.8 |
| 7 | 3.5 | 3.9 | 3.3 | 3.6 |
| 平均 | | | | 3.9 |

| e | 1 | 2 | 3 | a |
|---|------------|------|------|------|
| 1 | 0.0 | 0.3 | -0.2 | -1.1 |
| 2 | 0.7 | -0.3 | -0.4 | 0.8 |
| 3 | -0.1 | 0.2 | -0.1 | 0.2 |
| 4 | 0.3 | -0.4 | 0.1 | 0.7 |
| 5 | 0.6 | -0.1 | -0.4 | 0.7 |
| 6 | -0.6 | 0.6 | -0.1 | -1.1 |
| 7 | -0.1 | 0.3 | -0.3 | -0.3 |

| | | | |
|----|---|----|---|
| 患者 | 7 | 反復 | 3 |
|----|---|----|---|

| | | |
|---------|------------|------|
| 係数 | 1.8 | |
| 標準誤差 | 0.35 | 0.78 |
| 外れ値の基準値 | 0.64 | 1.40 |

| | 平方和 | 自由度 | 平均平方 | F | 分散成分 | ρ |
|------|-------|-----|-------|-------|-------|--------|
| 患者 | 12.72 | 6 | 2.120 | 11.30 | 0.644 | 0.774 |
| 誤差 | 2.63 | 14 | 0.188 | 1.00 | 0.188 | |
| 全体 | 15.35 | 20 | 0.767 | | | |
| (検算) | 15.35 | 20 | | | | 0.832 |

| | 下限 | 上限 |
|-----------------|-------------|--------|
| F($\alpha/2$) | 3.501 | 0.189 |
| C(p) | 3.227 | 59.847 |
| ρ | 0.426 | 0.951 |
| α (両側) | 0.05 | |

| 患者 | 平方和 | 自由度 | 平均平方 | F | 分散成分 | ρ |
|----|-------|-----|-------|---|-------|--------|
| 患者 | 12.72 | 7 | 1.817 | | 0.543 | 0.743 |
| 誤差 | 2.63 | 14 | 0.188 | | 0.188 | |

- 左上のくり返し測定の誤差 $e_{ij} = x_{ij} - \bar{x}_i$ を計算する (右上に求められている) .
- 各患者の平均と総平均の差 $a_i = \bar{x}_i - \bar{x}_{..}$ を求める .
- 分散分析表を作成するために, 患者数 (水準数) $n = 7$ とくり返し数 $p = 3$ を求める .
- 患者間の平方和 S_S を次の式で求める .

$$S_S = p \times \sum_i a_i^2 = 12.72$$

- 誤差の平方和 S_e を次の式で求める .

$$S_e = \sum_i \sum_j e_{ij}^2 = 2.63$$

- 全体の平方和 S_T

$$S_T = \sum_i \sum_j (x_{ij} - \bar{x}_{..})^2 = 15.35$$

を Excel 関数 =DEVSQ(データ全体) として求める*2 .

*2 原著では, p.15 で S_T , p.16 で S_S を求める式が示されている . その中で, いずれも修正項が用いられている . この補遺では, 平方和を, その定義に従って直接求めており, 修正項は用いていない .

原著では, p.18 で S_e を $S_T - S_S$ として求めている . この補遺では, 上の方針により, 誤差の成分 e_{ij} から直接求めている .

12 1 連続データにおける信頼性：級内相関係数(ICC)

○ 患者間の平方和と誤差の平方和の合計を（検算）の行に求め、全体の平方和と一致することを確認する。

○ 3つの自由度 f_S, f_e, f_T を次の式で求める。

$$f_S = n - 1 = 6, f_e = n(p - 1) = 14, f_T = np - 1 = 20$$

○ 平方和を自由度で割って平均平方を求める。

○ 原著には書かれていないが、後に必要となるので、F比を計算する。

$$F = V_S/V_e = 2.120/0.188 = 11.30$$

○ 患者間と誤差の分散の推定値 $\hat{\sigma}_S^2, \hat{\sigma}_e^2$ を次の式で計算する。

$$\hat{\sigma}_S^2 = \frac{V_S - V_e}{p} = \frac{2.120 - 0.188}{3} = 0.644$$

$$\hat{\sigma}_e^2 = V_e = 0.188$$

○ 信頼性係数 ρ の推定値を次の式で求める。

$$\hat{\rho} = \frac{\hat{\sigma}_S^2}{\hat{\sigma}_S^2 + \hat{\sigma}_e^2} = \frac{0.644}{0.644 + 0.188} = \frac{0.644}{0.832} = 0.774$$

(5) ICC と ρ の関係（補足）

分散分析表の下の表は、患者の自由度 $n - 1$ の代わりに n として平均平方を計算し、同様の式で信頼性係数 ρ を求めたものである。 $\rho = 0.743$ が得られる。これは、 u, v の相関係数 ICC の値に等しい（原著 p.20 参照）。

この関係は、上の計算過程から比較的容易に導くことができる。

表示1.8の右の平方和の合計 2.627 は、分散分析表の誤差の平方和 S_e に等しい。

また、ICC を計算するために用いた S_{xx} は分散分析表の S_T に相当する。

これから、ICC は次のように変形できる。

$$\text{ICC} = 1 - \frac{S_{dd}}{(p-1)S_{xx}} = 1 - \frac{pS_e}{(p-1)S_T}$$

一方 分散分析表から計算した ρ を変形すると

$$\rho = \frac{\frac{V_S - V_e}{p} - V_e}{\frac{V_S - V_e}{p} + V_e} = \frac{V_S - V_e}{V_S - V_e + pV_e} = 1 - \frac{pV_e}{V_S + (p-1)V_e}$$

となる。

ここで、分散分析表で患者間の自由度を $n - 1$ でなく n とすると、

$$S_T = S_S + S_e = nV_S + n(p-1)V_e$$

となり、これは ρ の式の分母は S_T/n となることが分かる。

分子の V_e は S_e をその自由度 $n(p-1)$ で割ったものである。 $V_e = (S_e/(n(p-1)))$ を代入すると、分子は

$$pV_e = p \frac{S_e}{n(p-1)}$$

となり、 ρ は

$$\rho = 1 - \frac{pS_e/(n(p-1))}{S_T/n} = \frac{pS_e}{(p-1)S_T}$$

となり、上に求めた ICC と一致することが確かめられる。

(6) ρ の区間推定

以下の説明は、原著 §13.1 (p.163) の内容に数値例を加えたものである。

分散分析表のF比は次のように変形できる。

$$F = \frac{V_S}{V_e} = \frac{\hat{\sigma}_e^2 + n\hat{\sigma}_S^2}{\hat{\sigma}_e^2} = 1 + n \frac{\hat{\sigma}_S^2}{\hat{\sigma}_e^2}$$

これから、

$$\frac{\hat{\sigma}_S^2}{\hat{\sigma}_e^2} = \frac{F-1}{n}$$

の関係が得られる。

ρ を求める式に上の関係を導入すると、

$$\rho = \frac{\hat{\sigma}_S^2}{\hat{\sigma}_S^2 + \hat{\sigma}_e^2} = \frac{1}{1 + \frac{\hat{\sigma}_e^2}{\hat{\sigma}_S^2}} = \frac{1}{1 + \frac{n}{F-1}} = \frac{F-1}{F-1+n} \quad (1.1)$$

が導かれる。

この例について計算すると

$$\rho = \frac{11.30 - 1}{11.30 - 1 + 3} = \frac{10.30}{13.30} = 0.774$$

として、前に求めた値が得られる。

帰無仮説 $H_0: \sigma_S^2 = 0$ が正しいとき、F比は自由度 (f_S, f_e) のF分布に従う。帰無仮説が成立しないとき、 $F/(1 + \frac{n\sigma_S^2}{\sigma_e^2})$ が自由度 (f_S, f_e) のF分布に従う。

これから， F の95%信頼区間 ($\alpha = 0.05$) を求めると

$$\frac{F}{F(\alpha/2; f_S, f_e)} < 1 + \frac{n\sigma_S^2}{\sigma_e^2} < \frac{F}{F(1 - \alpha/2; f_S, f_e)}$$

$$\frac{11.30}{3.501} = 3.227 < \frac{11.30}{0.189} = 59.847$$

が得られる．

この上限・下限値から式(1.1)を使って ρ の上限・下限値を計算すると

$$\frac{3.227 - 1}{3.227 - 1 + 3} = 0.426 < \rho < \frac{59.847 - 1}{59.847 - 1 + 3} = 0.951$$

が得られる．

この計算が表示1.9 の右下に示されている．

ρ の下限 (片側信頼区間) だけを求めたいときは， (両側) のセルに 0.10 を入力する．

$$0.497 < \rho$$

が得られる．

(7) 検出力の計算と実験の大きさの決め方

前項で求めた信頼区間は予想以上に広いと思われるであろう．

ρ を希望する信頼区間の幅で推定するために必要な対象数 n ，くり返し数 p を決める考え方は，原著 §4.3 (p.84) に，その基礎については §13.2 (p.166) に説明されている．

以下，Excel を使って具体的に計算する方法を説明する．

表示1.10 は，表の左のA:E のセルに条件を入力すると，F「検出力」のセルに検出力が求められるように作られた Excel の計算表である．G:I には計算の中間値が記録されている．

2行目の各セルに記録されている関数が下に示されている．

2行目には原著の表13.2 (p.169) の左上の検出力が，3行目には右下の検出力が求められている．なお，原著は小数点以下3桁目が切捨てられているのに対し，この計算表は小数点4桁以下を四捨五入している．

のセルには，片側の を入力する． $\alpha = 0.05$ の両側検定の検出力を求めたいときには $\alpha/2 = 0.025$ を入力する．

ρ_0, ρ_1 は $\rho_0 < \rho_1$ でなければならない．

n, p から 検出力を計算する表を，逆に使って，希望する検出力を得るために必要な n, p を求めることができる．

表示1.10: 検出力の計算表

| | A | B | C | D | E | F | G | H | I | J | K | L |
|----|-----------------------|----------|----------|------------|-------------|-------|------|----------------|----------------------|---------------|---------------|-------|
| 1 | α (片側) | ρ_0 | ρ_1 | 対象 数(n) | 繰返し 数(p) | 検出力 | fS | f _e | C(ρ_0) | C(ρ_1) | F(α) | |
| 2 | 0.025 | 0.8 | 0.9 | 10 | 2 | 0.215 | 9 | 10 | 4.00 | 9.00 | 3.779 | 1.680 |
| 3 | 0.025 | 0.9 | 0.95 | 30 | 20 | 0.812 | 29 | 551 | 181.00 | 381.00 | 1.603 | 0.762 |
| 4 | | | | | | | | | | | | |
| 5 | 0.025 | 0.8 | 0.9 | 40.3 | 3 | 0.80 | 39.3 | 80.6 | 13.00 | 28.00 | 1.685 | 0.782 |
| 6 | 0.025 | 0.8 | 0.9 | 25 | 18.1 | 0.80 | 24 | 429 | 73.60 | 164.34 | 1.673 | 0.749 |
| 7 | 0.05 | 0.8 | 0.9 | 32.4 | 3 | 0.80 | 31.4 | 64.8 | 13.00 | 28.00 | 1.630 | 0.757 |
| 8 | 0.05 | 0.8 | 0.9 | 20 | 24.5 | 0.80 | 19 | 470 | 98.98 | 221.46 | 1.609 | 0.719 |
| 9 | | | | | | | | | | | | |
| 10 | G2: =D2-1 | | | | | | | | K2: =FINV(A2,G2,H2) | | | |
| 11 | H2: =(D2-1)*(E2-1) | | | | | | | | L2: =I2/J2*K2 | | | |
| 12 | I2: =1+\$E2*B2/(1-B2) | | | | | | | | | | | |
| 13 | J2: =1+\$E2*C2/(1-C2) | | | | | | | | F2: =FDIST(L2,G2,H2) | | | |

$\rho_0 = 0.8$, $\rho_1 = 0.9$ のとき検出力が 0.8 になる n , p を求めることを考える.

表示1.10 の5行目に2行目をコピーし, $\rho_0 = 0.8$, $\rho_1 = 0.9$ を入力する. $n = 10$, $p = 2$ のままの検出力は 0.189 である. 検出力を 0.80 にするためには, n , p のいずれか一方, または, 両方を増やす必要がある.

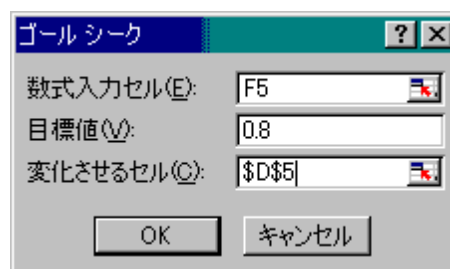
原始的には, n , p を試行錯誤で変化させて, 検出力が希望する値に近くなるような n , p の組合せを求めることができる. たとえば, $n = 30$, $p = 6$ にすると, 検出力は 0.803 になる.

この試行錯誤を自動的に実行してくれるのが, Excel の ゴールシーク である.

n , p を一緒に求めることはできないので, まず, $p = 3$ を固定して必要な n を求める. そのために, p に3を入力する.

検出力のセルをクリックして, トップメニューの「ツール」, 「ゴールシーク」を選択する. 表示1.11のパラメータ入力画面が現れ, クリックした「検出力」のセルが「数式入力セル」に表示されている。「目標値」のセルに 0.8 を, 「変化させるセル」に n のセルを入力する.

表示1.11: ゴールシークのパラメータ入力画面



「OK」をクリックすると、「解答がみつかりました」というメッセージが表示されるので、「OK」をクリックする。

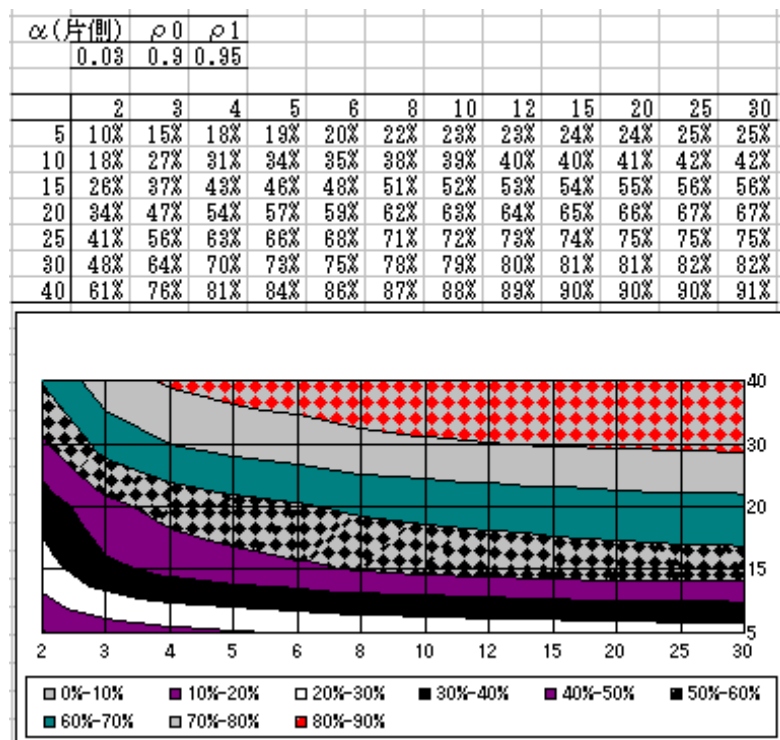
この結果が5行目に示されており、 $n = 40.3$ が得られる。切上げて41人の患者が必要となる。今度は望ましい患者数 $n = 25$ を固定して、必要な p を同様の手順で求めた結果が6行目で、 $p = 25$ となる。

これらの結果を参考にして、 n と p のバランスした実験を計画することができる。

原著の表13.2 (p.169) には、4つの ρ_0, ρ_1 組合せについて、実行可能と考えられる n, p についての検出力が表になっている。この表から、 n, p のおよその見当を付けることができる。

表示1.12の計算表は、任意の ρ_0, ρ_1 を指定することができ、表の周辺の n, p は行数・列数も含め、自由に変更できるものである。

表示1.12: 検出力の計算表と等高線



検出力は%表示で、%未満は四捨五入している（原著は切捨て）。

下には、縦軸に対象数 n 、横軸に繰返し数 p を取った検出力の等高線グラフが描かれている。等高線は10%間隔であり、上と中央の格子の下限は80%、50%である。

上の α, ρ_0, ρ_1 を修正すると、グラフも自動的に変更される。

(8) 外れ値の検出(補足)

(5), (6)項の計算は「F比の分布がF分布に従う」ことを利用している。これは、データ x_{ij} が

$$x_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

$$\tau_i \sim N(0, \sigma_S^2)$$

$$\varepsilon_{ij} \sim N(0, \sigma_e^2)$$

というモデルに従う、すなわち、患者間のばらつき、くり返し観測の誤差が共に正規分布に従うことが前提である。

個々の測定値の中に外れ値が含まれたり、異常な患者が含まれるとき、(5)項の信頼区間が不確かなものとなる。

以下、正規分布から外れている患者や測定値を見出す方法を説明する。

分散分析の計算を表示1.9のような手順で実行すると、各平方和の個々の構成要素を見ることができる。

残差平方和の構成要素 $|x_{ij} - \bar{x}_{i.}|$ の中に異常に大きい値が含まれるときは、 ρ が小さくなる。モデル式の ε_{ij} が正規分布に従うという前提が成立するかどうかを確認するためには、残差 e_{ij} の分散が必要である。

$$V[e_{ij}] = V[x_{ij} - \bar{x}_{i.}] = \left(1 - \frac{1}{p}\right) \sigma_e^2 = \frac{p-1}{p} \sigma_e^2$$

この式に σ_e^2 の推定値 $V_e = S_e / (n(p-1))$ を代入すると、

$$\hat{V}[e_{ij}] = \frac{p-1}{p} V_e = \frac{(p-1)S_e}{p * n(p-1)} = \frac{S_e}{np}$$

が得られる。これは、残差平方和 S_e を、自由度ではなく、残差の個数で割ったものである。この値は =VARP(データの集まり) で、この値の平方根(残差の標準誤差)は =STDEVP(データの集まり) で直接求めることができる。

表示1.9の右中(残差の下)に残差 e の標準誤差が 0.35 と求められている。

残差の絶対値が標準誤差の2.5倍以上であるとき、表示の字体や色を変えて表示するようしておく。外れ値が容易に発見できる。ここではこの機能を説明するために、1.8倍以上を指摘することにし、「係数」に 1.8 を入力する。「外れ値の基準値」に $1.8 \times 0.35 = 0.64$ が求められている。

残差のセル全体について、トップメニューから「書式」、「条件つき書式」を選び、「次の値の間以外」、「-基準値」、「基準値」を指定し、「書式」からスタイルで「太字斜体」を、色で「赤」を

18 1 連続データにおける信頼性：級内相関係数(ICC)

指定する．その結果が表示1.9 の右上で， $e_{21} = 0.7$ の字体が変わっている．なお，係数の 1.8 を 2.5 にすると，元に戻る．

患者間のばらつきも正規分布に従うことが前提である．

$|\hat{x}_i - \hat{x}_{..}|$ が異常に大きい患者が含まれるときは， ρ が大きくなる．

同様の考えで，主効果の値についても同様のチェックをする． a の標準誤差が 0.78 で，その 1.8 倍の 1.40 よりも大きい $|a_i|$ は認められない．

(9) JMP による解析

元のデータを患者番号と測定値の 21 行 2 列のデータ表にして，JMP データを作成する（信頼性 1-2.JMP では変数名「群」，「y」としてある）．

変数「群」の尺度タイプを 名義尺度 (N) に指定する．

トップメニューから「分析」，「モデルのあてはめ」を選択する．

表示 1.13: モデルの指定 画面 (入力後)



表示 1.13 の「モデルの指定」画面で，

- 列 y を Y に指定し，列 群 をモデル効果の構成 に 追加 する．

○ モデル構成の群を選択して、下の属性から変量効果を選択する。モデル構成の群の次に &変量効果 が追加される。この指定により、群間の分散が求められる。

変量効果のある場合の解析として、JMP では2つの方法、REML (REstricted または REsidual Maximum Likelihood) と、EMS (Expected Mean Square) が準備されている。JMP は REML 法の使用を推奨しているが、ここでの解析目的には EMS 法で十分である。

○ 右上の方法の右のレをクリックし、EMS (従来) を選択する。

「モデルの実行」をクリックすると、表示1.14 が得られる。

表示1.14: JMP の出力 (1)

| モデル全体 | | | | |
|------------------------------|----------|-----------|---------|------------|
| あてはめの要約 | | | | |
| R2乗 | | | | 0.828834 |
| 自由度調整R2乗 | | | | 0.755477 |
| 誤差の標準偏差(RMSE) | | | | 0.43315 |
| Yの平均 | | | | 3.885714 |
| オブザベーション(または重みの合計) | | | | 21 |
| 分散分析 | | | | |
| 要因 | 自由度 | 平方和 | 平均平方 | F値 |
| モデル | 6 | 12.719048 | 2.11984 | 11.2986 |
| 誤差 | 14 | 2.626667 | 0.18762 | p値(Prob>F) |
| 全体(修正済み) | 20 | 15.345714 | | 0.0001 |
| 分散成分推定値 | | | | |
| 成分 | 分散成分推定値 | 全体に対する百分率 | | |
| 群&変量効果 | 0.644074 | 77.441 | | |
| 残差 | 0.187619 | 22.559 | | |
| 合計 | 0.831693 | 100.000 | | |
| 平均平方がその期待値に等しいものとして推定したものです。 | | | | |

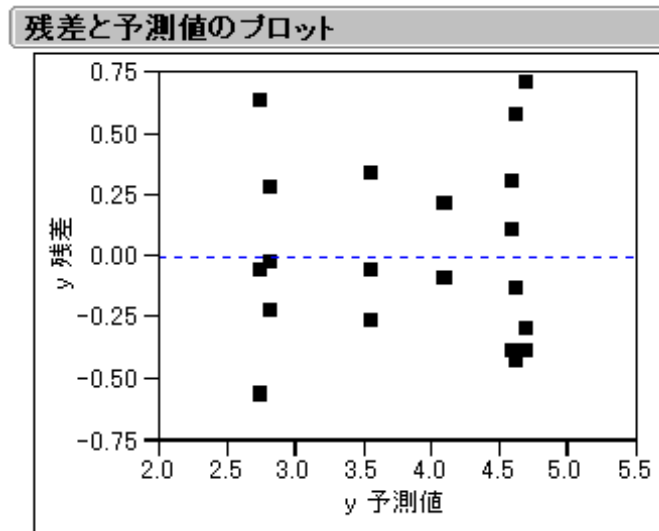
「分散成分推定値」に2つの分散成分の推定値が求められ、 ρ の値 0.774 が百分率で表示されている。

JMP は表示1.15 に示すような「残差と予測値」のグラフを表示する。

このグラフを見ると、予測値(各患者の平均値)が平均値の近傍では誤差が小さく、予測値が小さいまたは大きいとき残差が大きくなる傾向がありそうである。これは表示1.3のグラフで、第2回の測定値は第1回の測定値が縮小している傾向の見られたことと対応している。

もし、点が大きい方に広がる扇の形をしているとき、すなわち、予測値(平均値)が大きいと

表示1.15: JMP の出力 (2)



き残差のばらつきが大きくなる傾向が見られるときは、等分散性が成立しない危険性がある。

このようなときは、繰返し誤差の変動係数が一定であり、測定値を対数変換すると等分散性が満たされる場合が少なくない。

1.3 2元配置データの解析

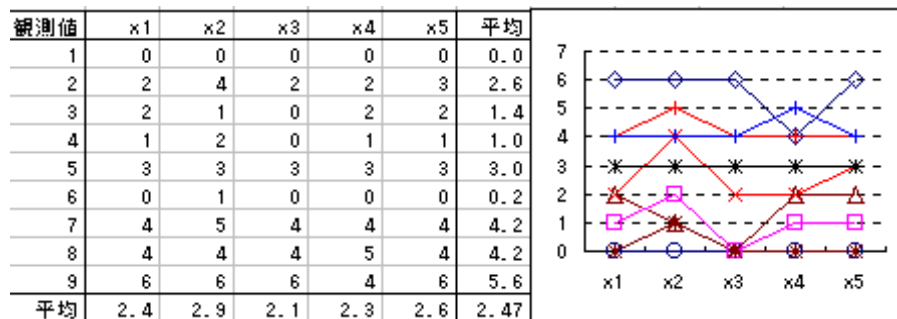
(1) データと ICC の計算

前節の実験データが、単純な反復測定ではなく、同じ患者を複数の医師が評価したり、複数の測定器で測定した場合にはどのように解析したら良いかを考えるのがこの節である。

原著 表1.8 (p.23) のデータは9人の患者を5人の医師が評価した結果である。

データを表示1.16 左に示す。

表示1.16: データ, 平均値と折れ線グラフ



前節までは、患者の平均だけが意味を持ったが、このデータでは列の平均も意味を持つので、横平均・縦平均・総平均が求められている。

右に折れ線グラフが示されている。

これから、患者間のばらつきが大きく、評価する医師間のばらつきは比較的小さいように見える。

このデータから、表示1.8の方法で計算すると^{*3}、表示1.17左に示すように $ICC = 0.895$ が得られる。

このICCに対応する ρ の分母 S_T には、医師間のばらつきが含まれる。そこで、このばらつきを除くために、 x_{ij} の代わりに $x_{ij} - \bar{x}_{.j}$ について同様の計算をすると、表示1.17右に示すように $ICC = 0.916$ が得られる。

当然のことながら、上に求めた $ICC=0.895$ よりも大きくなる。

^{*3} §1.2 (3) の最後で説明した性質を利用して、 x の差を求めずに S_{dd} を計算するように改善している。

表示1.17: ICC の計算

| 列数 | 5 | | | | | 平方和 | | x1 | x2 | x3 | x4 | x5 | 平方和 |
|-----|--------|---|---|---|---|------|-----|--------|------|------|------|------|------|
| 1 | 0 | 0 | 0 | 0 | 0 | 0.0 | 1 | -2.4 | -2.9 | -2.1 | -2.3 | -2.6 | 0.3 |
| 2 | 2 | 4 | 2 | 2 | 3 | 3.2 | 2 | -0.4 | 1.1 | -0.1 | -0.3 | 0.4 | 1.7 |
| 3 | 2 | 1 | 0 | 2 | 2 | 3.2 | 3 | -0.4 | -1.9 | -2.1 | -0.3 | -0.6 | 3.0 |
| 4 | 1 | 2 | 0 | 1 | 1 | 2.0 | 4 | -1.4 | -0.9 | -2.1 | -1.3 | -1.6 | 0.8 |
| 5 | 3 | 3 | 3 | 3 | 3 | 0.0 | 5 | 0.6 | 0.1 | 0.9 | 0.7 | 0.4 | 0.3 |
| 6 | 0 | 1 | 0 | 0 | 0 | 0.8 | 6 | -2.4 | -1.9 | -2.1 | -2.3 | -2.6 | 0.3 |
| 7 | 4 | 5 | 4 | 4 | 4 | 0.8 | 7 | 1.6 | 2.1 | 1.9 | 1.7 | 1.4 | 0.3 |
| 8 | 4 | 4 | 4 | 5 | 4 | 0.8 | 8 | 1.6 | 1.1 | 1.9 | 2.7 | 1.4 | 1.4 |
| 9 | 6 | 6 | 6 | 4 | 6 | 3.2 | 9 | 3.6 | 3.1 | 3.9 | 1.7 | 3.4 | 3.0 |
| 平方和 | 167.20 | | | | | 14.0 | 平方和 | 164.22 | | | | | 11.0 |
| ICC | 0.895 | | | | | | ICC | 0.916 | | | | | |

(2) 分散分析と ρ の計算

上に求めた ICC の意味を理解するために、分散分析表から得られる ρ と比較する。
分散分析表を作成するために、表示1.9 と同様の計算表を作成する。

表示1.18: データと3つの成分への分解

| 観測値 | 1 | 2 | 3 | 4 | 5 | 平均 | 誤差 | 1 | 2 | 3 | 4 | 5 | 主効果 |
|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|------|
| 1 | 0 | 0 | 0 | 0 | 0 | 0.0 | 1 | 0.0 | -0.4 | 0.4 | 0.1 | -0.1 | -2.5 |
| 2 | 2 | 4 | 2 | 2 | 3 | 2.6 | 2 | -0.6 | 1.0 | -0.2 | -0.5 | 0.3 | 0.1 |
| 3 | 2 | 1 | 0 | 2 | 2 | 1.4 | 3 | 0.6 | -0.8 | -1.0 | 0.7 | 0.5 | -1.1 |
| 4 | 1 | 2 | 0 | 1 | 1 | 1.0 | 4 | 0.0 | 0.6 | -0.6 | 0.1 | -0.1 | -1.5 |
| 5 | 3 | 3 | 3 | 3 | 3 | 3.0 | 5 | 0.0 | -0.4 | 0.4 | 0.1 | -0.1 | 0.5 |
| 6 | 0 | 1 | 0 | 0 | 0 | 0.2 | 6 | -0.2 | 0.4 | 0.2 | -0.1 | -0.3 | -2.3 |
| 7 | 4 | 5 | 4 | 4 | 4 | 4.2 | 7 | -0.2 | 0.4 | 0.2 | -0.1 | -0.3 | 1.7 |
| 8 | 4 | 4 | 4 | 5 | 4 | 4.2 | 8 | -0.2 | -0.6 | 0.2 | 0.9 | -0.3 | 1.7 |
| 9 | 6 | 6 | 6 | 4 | 6 | 5.6 | 9 | 0.4 | 0.0 | 0.8 | -1.5 | 0.3 | 3.1 |
| 平均 | 2.4 | 2.9 | 2.1 | 2.3 | 2.6 | 2.5 | 主効果 | 0.0 | 0.4 | -0.4 | -0.1 | 0.1 | 2.5 |

左のデータと平均の表は、表示1.17と同じである。

測定値 x_{ij} を

$$\begin{aligned}
 x_{ij} &= \bar{x}_{..} + (\bar{x}_{i.} - \bar{x}_{..}) + (\bar{x}_{.j} - \bar{x}_{..}) + (x_{ij} - (\bar{x}_{..} + (\bar{x}_{i.} - \bar{x}_{..}) + (\bar{x}_{.j} - \bar{x}_{..}))) \\
 &= \bar{x}_{..} + a_i + b_j + (x_{ij} - (\bar{x}_{..} + a_i + b_j)) \\
 &= \bar{x}_{..} + a_i + b_j + e_{ij}
 \end{aligned}$$

のように、各成分に分解した結果が右の表である。

右下に $\bar{x}_{..}$ が、右に a_i が、下に b_j が、表本体に e_{ij} に求められている。

原著 p.25 に書かれている各平方和は 表示 1.18 の x_{ij} , a_i , b_j , e_{ij} から次の式で計算される .

$$S_T = \sum_{i=1}^n \sum_{j=1}^p (x_{ij} - \bar{x}_{..})^2$$

$$S_S = p \sum_{i=1}^n a_i^2$$

$$S_R = n \sum_{j=1}^p b_j^2$$

$$S_e = \sum_{i=1}^n \sum_{j=1}^p e_{ij}^2$$

S_T は DEVSQ 関数で , 残りの 3 つの平方和は SUMSQ で計算される .

表示 1.19 の分散分析表の平方和はこの方法で求められている .

表示 1.19: 分散分析表

| | SS | DF | MS | F | 分散成分 | ICC(ρ) | |
|---------|--------|----------|--------|-------|-------|---------------|---------------|
| | | | | | | Fleiss | Ebel |
| 患者 | 153.20 | 8 | 19.150 | 55.60 | 3.761 | 0.9063 | 0.9161 |
| 評価者 | 2.98 | 4 | 0.744 | 2.16 | 0.044 | | |
| 誤差 | 11.02 | 32 | 0.344 | 1.00 | 0.344 | | |
| 全体 | 167.20 | 44 | 3.800 | | | | |
| 分散成分の合計 | | | | | | 4.150 | 4.106 |
| 患者 | 153.20 | 9 | 17.022 | | 3.327 | 0.8953 | |
| 誤差* | 14.00 | 36 | 0.389 | | 0.389 | | |
| | | | | | | 0.8953 | 0.9161 |

それぞれの自由度で割って平均平方を求められ , 原著の表 1.10 と同じ値が得られている . また , 原著 p.25 の式により分散成分が計算されている .

分散成分の合計の行には , 3 つの分散成分の和と , 患者と誤差の分散成分の和が計算されている . 合計に対する 患者 の分散成分の割合 $ICC()$ が 0.906, 0.916 と求められている . これらの値は , 原著 p.26 の上に求められている値に一致する .

原著 p.20 の下に , 「級内相関係数 ICC は 患者間の自由度を n とした ρ に一致する」と書かれており , 前節の例で確認した . ここで取り上げたデータについて確認した結果が 表示 1.19 の下の表である . ICC の行に表示 1.17 で求めた 2 つの ICC が転記されている .

分散分析表の 患者間の自由度が 9 になっており , $\rho = 0.895$ が得られる . これは , 前に求めた ICC に一致する .

Ebel の 0.9161 は , 表示 1.17 右で評価者の平均値の差を修正した ICC 0.9161 に一致する .

(3) JMP による解析

前節と同様に、行番号（患者）、列番号（評価者）、測定値の3列のデータを準備する（信頼性 1-3.JMP）。

行番号、列番号を名義尺度 (N) とし、表示 1.13 モデル指定画面で、行番号と列番号をモデルに入れ、変量効果を指定する。

表示 1.20 が得られる。

表示 1.20: 分散分析表と の計算

| モデル全体 | | | | |
|------------------------------|----------|-----------|---------|------------|
| 分散分析 | | | | |
| 要因 | 自由度 | 平方和 | 平均平方 | F値 |
| モデル | 12 | 156.17778 | 13.0148 | 37.7849 |
| 誤差 | 32 | 11.02222 | 0.3444 | p値(Prob>F) |
| 全体(修正済み) | 44 | 167.20000 | | <.0001 |
| 分散成分推定値 | | | | |
| 成分 | 分散成分推定値 | 全体に対する百分率 | | |
| 患者&変量効果 | 3.761111 | 90.629 | | |
| ID&変量効果 | 0.044444 | 1.071 | | |
| 残差 | 0.344444 | 8.300 | | |
| 合計 | 4.15 | 100.000 | | |
| 平均平方がその期待値に等しいものとして推定したものです。 | | | | |

「分散成分推定値」に 90.629% と Fleiss の ρ が求められている。Ebel の ρ は求められない。分散成分推定値から計算する。

(4) 別の例の解析

原著 §4.4 の表 4.4 (p.89) に 20 人の患者を 11 人の医師が評価したデータが示されている。これは、順序カテゴリー変数であるが、連続量として解析する。

原著でも連続量として解析して得られた分散分析表が表 4.5 (p.90) に示されている。

この分散分析表を求めるために、 $\bar{x}_{..}$, a_i , b_j , e_{ij} を求めた結果を表示 1.21 にしめす。この計算を求める元になる x_{ij} , $\bar{x}_{i.}$, $\bar{x}_{.j}$, $\bar{x}_{..}$ は添付する Excel ファイルのシート「例題」に示す。

主効果の右と下の平方和は後の説明で用いられる。

これから、分散分析表を求めた結果を表示 1.22 に示す。

このようにして、原著 p.90 の 2 元配置の 2 つの ICC が求められた。

この例のように、患者数や評価者数が大きいときは、分散分析表だけの解釈で解析を終わらせるのは、重要な情報を見逃す恐れがある。

表示1.21: データの成分への分解

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 主効果 | 平方和 |
|-----|------------|-------------|-------------|------|------|------|-------------|-------------|-------------|-------------|-------|-------|-------------|
| 1 | 0.3 | -0.7 | -0.6 | 0.5 | -0.5 | 0.5 | -0.4 | 0.6 | -0.4 | 0.7 | -0.2 | -2.94 | 2.82 |
| 2 | 0.3 | -0.7 | -0.6 | 0.5 | -0.5 | -0.5 | -0.4 | 0.6 | -0.4 | 0.7 | 0.8 | -2.94 | 3.52 |
| 3 | -0.8 | -0.8 | 0.3 | 0.3 | 0.3 | 0.3 | 0.4 | -0.6 | 0.5 | 0.6 | -0.3 | -2.76 | 2.92 |
| 4 | 0.5 | -0.5 | -0.4 | -0.3 | -0.3 | 0.7 | -0.2 | 0.8 | -0.2 | -0.1 | 0.0 | -2.12 | 2.00 |
| 5 | 0.1 | 0.1 | -0.8 | 0.2 | 0.2 | 0.2 | 0.3 | 0.3 | 0.4 | -0.5 | -0.4 | -1.67 | 1.65 |
| 6 | 1.4 | 0.4 | -0.5 | -0.4 | -0.4 | -0.4 | -0.3 | 0.7 | -0.3 | -0.2 | -0.1 | -1.03 | 3.65 |
| 7 | -0.1 | 0.9 | 1.0 | 0.0 | -1.0 | -1.0 | 0.1 | 0.1 | 0.2 | 0.3 | -0.6 | -0.49 | 4.16 |
| 8 | -0.1 | -0.1 | -1.0 | 0.0 | 0.0 | 0.0 | 0.1 | 0.1 | 0.2 | 0.3 | 0.4 | -0.49 | 1.36 |
| 9 | -0.2 | 0.8 | 0.9 | 0.9 | 0.9 | 0.9 | -2.0 | 0.0 | -0.9 | -0.8 | -0.7 | -0.40 | 9.94 |
| 10 | -0.8 | 1.3 | 0.4 | -0.6 | -0.6 | -0.6 | 0.5 | -0.5 | 0.6 | 0.6 | -0.3 | 0.15 | 4.62 |
| 11 | -0.2 | -0.2 | -0.1 | -0.1 | -0.1 | -0.1 | 0.0 | 0.0 | 0.1 | 0.2 | 0.3 | 0.60 | 0.24 |
| 12 | -0.5 | -0.5 | -0.4 | -0.3 | 0.7 | 0.7 | 0.8 | -0.2 | -0.2 | -0.1 | 0.0 | 0.88 | 2.30 |
| 13 | 0.3 | 0.3 | 0.4 | 0.5 | 0.5 | -0.5 | 0.6 | -0.4 | -1.4 | -0.3 | -0.2 | 1.06 | 3.62 |
| 14 | 0.3 | 0.3 | 0.4 | -0.5 | 0.5 | -0.5 | -0.4 | -0.4 | 1.6 | -1.3 | -0.2 | 1.06 | 5.82 |
| 15 | -0.8 | 0.2 | 0.3 | 0.3 | 0.3 | 0.3 | 0.4 | 0.4 | -0.5 | -0.4 | -0.3 | 1.24 | 2.02 |
| 16 | 0.2 | 0.2 | 0.3 | 0.3 | 0.3 | 0.3 | 0.4 | -1.6 | -0.5 | 0.6 | -0.3 | 1.24 | 3.82 |
| 17 | 0.0 | 0.0 | 0.1 | -0.9 | 0.1 | 0.1 | 0.2 | 0.2 | 0.3 | -0.6 | 0.5 | 1.42 | 1.60 |
| 18 | 0.2 | 0.2 | 0.3 | -0.7 | -0.7 | -0.7 | -0.6 | 0.4 | 0.5 | 0.6 | 0.7 | 2.24 | 3.02 |
| 19 | 0.0 | -1.0 | 0.1 | 0.1 | 0.1 | 0.1 | 0.2 | 0.2 | 0.3 | -0.6 | 0.5 | 2.42 | 1.90 |
| 20 | -0.1 | -0.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.1 | -0.9 | 0.2 | 0.3 | 0.4 | 2.51 | 1.06 |
| 主効果 | 0.20 | 0.20 | 0.10 | 0.05 | 0.05 | 0.05 | -0.05 | -0.05 | -0.10 | -0.20 | -0.30 | 4.40 | |
| 平方和 | 5.21 | 6.67 | 5.44 | 4.22 | 4.77 | 5.13 | 6.79 | 6.61 | 7.47 | 6.37 | 3.62 | | 62.06 |

表示1.22: 分散分析表と の計算

| | | | | | | | |
|------|--------|-----|-------|--------|-------|--------|--------|
| 患者数 | 20 | | | | | | |
| 評価者数 | 11 | | | | | | |
| | SS | DF | MS | F | 分散成分 | Fleiss | Ebel |
| 患者 | 655.69 | 19 | 34.51 | 105.65 | 3.108 | 0.9028 | 0.9049 |
| 評価者 | 4.85 | 10 | 0.48 | 1.48 | 0.008 | | |
| 誤差 | 62.06 | 190 | 0.33 | 1.00 | 0.327 | | |
| 全体 | 722.60 | 219 | | | | | |

誤差の平方和は表示 1.21 の表本体の e_{ij} を 2 乗して加えたものである。この中から外れ値を見つける方法は §1.2 (6) で説明した。

ここでは、測定値が離散値であるから、誤差が 1 以上の誤差を太字・イタリック (Excel 画面では赤) で示した。

誤差の絶対値が 2 である要素が一つ見られる。

誤差の平方和は、 e_{ij}^2 を横に足し縦に足しても、逆に縦に足し横に足しても求められる。この過程が 表示 1.21 の周辺に求められている。

この中に特に大きいものがあるときには、その原因を追求することにより、貴重な情報の得ら

れることがある。

この例では、評価者 9 が大きいようであるが、異常というほどではない。

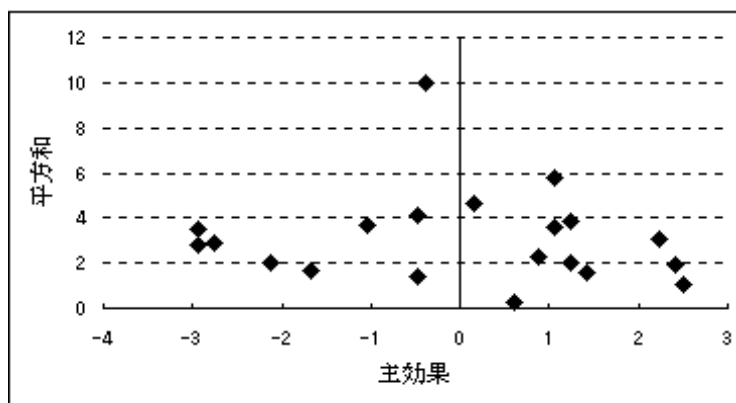
次に患者別の平方和を観察する。

患者 9 の平方和だけが 9.94 と他の患者とは大きく異なっている。医師 7 の誤差が 2.0 と大きいのが原因と思われるが、患者 9 がどのような病状なのかを追求すると何か分かるかもしれない。

患者の改善度 \bar{x}_i によって評価のばらつきが異なるかもしれない。

そこで、患者の主効果 $b_i = \bar{x}_i - \bar{x}_{..}$ を横軸に、誤差 e_{ij} を縦軸に取って散布図を描いたのが表示 1.23 である。

表示 1.23: 患者による誤差の違い



この図では、上に挙げた 9.94 以外の点には傾向が見られない。

もし、この散布図の点が山形になったり、両側に山ができたり、傾斜が見られたりしたときには、症状によって評価の分かれ方がことなるという可能性が疑われる。

医師の分散分析については、医師の所属する病院、出身大学などで層別して解析することにより、面白い結果が得られるかもしれない。

補遺の筆者は、データの背景について全く知識がない。データを取った人は沢山の背景情報を持っているであろう。これをフルに活用して、上に述べた探索的な解析をすることにより、固有技術が蓄積され、実験結果を活用して、信頼性の高い評価法を構築できるであろう。

2 分類データにおける信頼性： 係数

前章では、測定値が連続変数である場合を取り上げた。

この章では、測定値がいくつかのカテゴリーのどこに属するかである場合を取り上げる。このような変数を分類データ、カテゴリカルデータと呼ぶ。

分類データは、カテゴリーの順序に意味のある 順序尺度 (Ordinal Scale) と意味のない 名義尺度 (Nominal Scale) に分けられる。

原著では、順序尺度の例として 重症度 (ごく軽度, 軽度, 中等度, 高度) が、名義尺度の例として 診断 (脳出血, 脳梗塞, くも膜下出血) があげられている。

その何れであるかによって解析の方法や、結果の見方が大きく異なる。

JMP では、変数のタイプを C(Continuous 連続), O, N のいずれであるかを指定しなければならない。

もう一つの例として、治療の効果 (有効, 無効) が挙げられているが、2 カテゴリーの場合はどちらとしても解析結果は変わらないはずである。

§2.1, §2.2 で 名義尺度の場合を、§2.3 で 順序尺度の場合を取り上げる。

2.1 くり返し2回の場合

(1) Excel 関数による解析

原著の表2.1 から、Excel で 表2.2 を作成するためには、通常 ピボットテーブル が用いられるが、ここでは、Excel 関数を使う原始的な方法を用いる。

原著の表2.1 のデータを B3:C32 に入力する。

データの一部を表示2.1 の左に示す。

表示2.1 の右上のような表を準備する。左上のセルに

```
=COUNT(IF(($B$3:$B$32=$F3)*($C$3:$C$32=G$2),1,""))
```

```
=COUNT(IF(( x1 = 表側の値)*( x2 = 表頭の値),1,""))
```

という式を入力し、右と下にコピーする。

原著の表2.2 が得られる。

この表を、「編集」、「形式を選択してはりつけ」で、「行列をいれかえる」、[値]を指定すると、

表示2.1: データ（一部）とクロス表

| | A | B | C | D | E | F | G | H | I | J | K | L |
|----|----|----|----|---|----|----|----|----|----|---|----|--------|
| 2 | | x1 | x2 | | | 1 | 2 | 3 | 4 | 5 | 合計 | |
| 3 | 1 | 1 | 1 | | 1 | 4 | 1 | 0 | 0 | 0 | 5 | |
| 4 | 2 | 1 | 1 | | 2 | 2 | 3 | 1 | 0 | 0 | 6 | |
| 5 | 3 | 1 | 1 | | 3 | 1 | 1 | 4 | 1 | 0 | 7 | |
| 6 | 4 | 1 | 1 | | 4 | 0 | 0 | 1 | 6 | 1 | 8 | |
| 7 | 5 | 2 | 1 | | 5 | 0 | 0 | 0 | 0 | 4 | 4 | |
| 8 | 6 | 2 | 1 | | 合計 | 7 | 5 | 6 | 7 | 5 | 30 | |
| 9 | 7 | 3 | 1 | | | | | | | | | |
| 10 | 8 | 1 | 2 | | | 1 | 2 | 3 | 4 | 5 | 合計 | |
| 11 | 9 | 2 | 2 | | 1 | 4 | 2 | 1 | 0 | 0 | 7 | |
| 12 | 10 | 2 | 2 | | 2 | 1 | 3 | 1 | 0 | 0 | 5 | |
| 13 | 11 | 2 | 2 | | 3 | 0 | 1 | 4 | 1 | 0 | 6 | |
| 14 | 12 | 3 | 2 | | 4 | 0 | 0 | 1 | 6 | 0 | 7 | |
| 15 | 13 | 2 | 3 | | 5 | 0 | 0 | 0 | 1 | 4 | 5 | |
| 16 | 14 | 3 | 3 | | 合計 | 5 | 6 | 7 | 8 | 4 | 30 | |
| 17 | 15 | 3 | 3 | | | | | | | | | |
| 18 | 16 | 3 | 3 | | | 1 | 2 | 3 | 4 | 5 | 合計 | D |
| 19 | 17 | 3 | 3 | | 1 | 8 | 3 | 1 | 0 | 0 | 12 | 0.2000 |
| 20 | 18 | 4 | 3 | | 2 | 3 | 6 | 2 | 0 | 0 | 11 | 0.1833 |
| 21 | 19 | 3 | 4 | | 3 | 1 | 2 | 8 | 2 | 0 | 13 | 0.2167 |
| 22 | 20 | 4 | 4 | | 4 | 0 | 0 | 2 | 12 | 1 | 15 | 0.2500 |
| 23 | 21 | 4 | 4 | | 5 | 0 | 0 | 0 | 1 | 8 | 9 | 0.1500 |
| 24 | 22 | 4 | 4 | | 合計 | 12 | 11 | 13 | 15 | 9 | 60 | |

表示2.1の右中（原著 表2.3）が得られる．

上の2つの表の対応する要素を加えると右下の表（原著 表2.4）が得られる．

表示2.1の右下の表の横計の値と，対角要素の合計 $8 + 6 + 8 + 12 + 8 = 42$ が以下の計算に必要となる．

これらの値は，上のような手順を踏まなくても，元データから比較的簡単に求められる．

表示2.2の左上に， x_1, x_2 が1~5である個数を COUNTIF 関数で求める．横に合計すると，度数計の欄に表示2.1の右下の合計が得られる．

度数計を度数の総合計 60 で割って，割合が求める．割合を2乗して加えると， $P_e = 0.2056$ が求められる．

クロス表の対角要素の合計は，COUNT 関数と IF 関数を組合わせて， $x_1 = x_2$ の個数を数えることによって求められる．結果が Q11 に 42 と求められている．42を総度数60で割った割合が $P_0 = 0.7000$ である．

係数が表示2.2のS11に $\kappa = (P_0 - P_e)/(1 - P_e) = 0.6224$ と求められている．

以上は x_1, x_2 が単純な繰返し（1元配置データ）としての解析である． x_1, x_2 が2人の評価者

表示2.2: 統計量の直接計算

| | N | O | P | Q | R | S | T |
|----|------|----------------------------------|----|-----|--------|-----------------|----------|
| 2 | | x1 | x2 | 度数計 | 割合 | 割合 ² | 割合1*2 |
| 3 | 1 | 5 | 7 | 12 | 0.2000 | 0.0400 | 0.0389 |
| 4 | 2 | 6 | 5 | 11 | 0.1833 | 0.0336 | 0.0333 |
| 5 | 3 | 7 | 6 | 13 | 0.2167 | 0.0469 | 0.0467 |
| 6 | 4 | 8 | 7 | 15 | 0.2500 | 0.0625 | 0.0622 |
| 7 | 5 | 4 | 5 | 9 | 0.1500 | 0.0225 | 0.0222 |
| 8 | 合計 | 30 | 30 | 60 | | 0.2056 | 0.2033 |
| 9 | | | | | | | |
| 10 | | | | 度数 | 割合 | κ | κ |
| 11 | 対角要素 | | | 42 | 0.7000 | 0.6224 | 0.6234 |
| 12 | | | | | | | |
| 13 | O3: | =COUNTIF(B\$3:B\$32,\$N3) | | | | | |
| 14 | Q11: | =2*COUNT(IF(B3:B32=C3:C32,1,"")) | | | | | |
| 15 | S11: | =(R11-S8)/(1-S8) | | | | | |
| 16 | T3: | =O3*P3/O\$8^2 | | | | | |
| 17 | T11: | =(R11-T8)/(1-T8) | | | | | |

の評価値である(2元配置データ)の場合は,原著表2.3 p.36 から が計算される.

P_e は, x_1, x_2 それぞれの度数計から求めた割合の積を合計して求められる.これが,表示2.2のT列に求められ, $P_e = 0.2033$ となる.これを使って求めた は 0.6234 となる.

(2) VBA マクロによる解析

表示2.2の計算表を使えば,1元配置,2元配置の計算が比較的簡単に実行できる.しかし,繰返し数が3以上になると,かなり複雑となり,表形式での計算は困難である.

そこで,繰返し数が3以上でも使えるVBAマクロを準備した.その基本的な使い方を上の例題で説明する.

データを表示2.3の左のように準備する(データの一部は省略).左上に列数と行数を入力する.行数のセル(太字で示す)をクリックして,マクロ kappa を実行する*2.

マクロの処理結果は右に表示される.

上は表示2.1の右下のクロス表である.

中央に,1元配置としての P_0, P_e, κ が求められている.

下は,2元配置としての解析結果で,左下に P_0 が右上に κ が求められている.

*2 トップメニューの「ツール」から「マクロ」,「マクロ」を選択し,「マクロ名」で kappa を指定して実行する.

表示2.3: VBA マクロによる計算

| |
|-----------|
| 2 |
| 30 |
| 1 |
| 2 |
| 3 |
| 4 |
| 5 |
| 6 |
| 7 |
| 8 |
| 9 |
| 10 |
| 11 |
| 12 |
| 13 |
| 14 |
| 15 |
| 16 |

| | | |
|----|----|----|
| | x1 | x2 |
| 1 | 1 | 1 |
| 2 | 1 | 1 |
| 3 | 1 | 1 |
| 4 | 1 | 1 |
| 5 | 2 | 1 |
| 6 | 2 | 1 |
| 7 | 3 | 1 |
| 8 | 1 | 2 |
| 9 | 2 | 2 |
| 10 | 2 | 2 |
| 11 | 2 | 2 |
| 12 | 3 | 2 |
| 13 | 2 | 3 |
| 14 | 3 | 3 |
| 15 | 3 | 3 |
| 16 | 3 | 3 |

| | | | | | | |
|----|----|----|----|----|---|----|
| | 1 | 2 | 3 | 4 | 5 | 合計 |
| 1 | 8 | 3 | 1 | 0 | 0 | 12 |
| 2 | 3 | 6 | 2 | 0 | 0 | 11 |
| 3 | 1 | 2 | 8 | 2 | 0 | 13 |
| 4 | 0 | 0 | 2 | 12 | 1 | 15 |
| 5 | 0 | 0 | 0 | 1 | 8 | 9 |
| 合計 | 12 | 11 | 13 | 15 | 9 | 60 |

| | |
|----------|-------|
| P0 | 0.700 |
| Pe | 0.206 |
| κ | 0.622 |

| | |
|--------------|------------|
| P0, κ | |
| | x1 x2 |
| x1 | 0.623 |
| x2 | 0.700 |

(3) JMP による解析

原著の表2.2 と表2.4 のクロス表の行番号 u , 列番号 v , 度数 $f1, f2$ を並べて , 表示2.4の左に示すようなデータを準備する (信頼性2-1.JMP) .

表示2.4: JMP 用データと出力

| | | | |
|---|---|----|----|
| u | v | f1 | f2 |
| 1 | 1 | 4 | 8 |
| 1 | 2 | 1 | 3 |
| 1 | 3 | 0 | 1 |
| 2 | 1 | 2 | 3 |
| 2 | 2 | 3 | 6 |
| 2 | 3 | 1 | 2 |
| 3 | 1 | 1 | 1 |
| 3 | 2 | 1 | 2 |
| 3 | 3 | 4 | 8 |
| 3 | 4 | 1 | 2 |
| 4 | 3 | 1 | 2 |
| 4 | 4 | 6 | 12 |
| 4 | 5 | 1 | 1 |
| 5 | 4 | 0 | 1 |
| 5 | 5 | 4 | 8 |

uとvの分割表に対する分析

度数: n

検定

| | | | |
|----------|-----|-----------|--------|
| 要因 | 自由度 | (-1)対数尤度 | R2乗(U) |
| モデル | 16 | 25.416637 | 0.5301 |
| 誤差 | 10 | 22.531607 | |
| 全体(修正済み) | 26 | 47.948243 | |
| N | 30 | | |

検定 カイ2乗 p値(Prob>ChiSq)

尤度比 50.833 <0001

Pearson 55.776 <0001

警告:セルのうち 20%の期待度数が5未満です。カイ2乗にある可能性があります。

警告:平均セル度数が5未満です。尤度比カイ2乗に問題がある可能性があります。

カッパ 標準誤差

0.623431 0.104884

カッパは一致度の測度。

u, v を名義尺度 (N) に指定する。「二変量の関係」で, u を X に, v を Y に, $f1$ を度数に指定して, 実行すると, 表示2.4 の右の出力が得られる .

一番下に カッパ (一致度の測度) として 0.623431 が得られる .

f2 を 度数 に指定すると , 同様な出力で , カッパが 0.622378 となる .

2.2 繰返し3回以上の一般の場合

(1) データ

表示2.5 の左は 原著 表2.5 (p.41) のデータである .

表示2.5: VBA マクロによる計算 (くり返し3回)

| 3 | x1 | x2 | x3 |
|----|----|----|----|
| 1 | 1 | 1 | 1 |
| 2 | 1 | 1 | 2 |
| 3 | 2 | 1 | 2 |
| 4 | 1 | 2 | 3 |
| 5 | 2 | 2 | 2 |
| 6 | 3 | 3 | 2 |
| 7 | 3 | 3 | 3 |
| 8 | 4 | 3 | 4 |
| 9 | 4 | 4 | 4 |
| 10 | 4 | 4 | 4 |

| | 1 | 2 | 3 | 4 | 合計 |
|----|----|----|----|----|----|
| 1 | 8 | 5 | 1 | 0 | 14 |
| 2 | 5 | 8 | 3 | 0 | 16 |
| 3 | 1 | 3 | 8 | 2 | 14 |
| 4 | 0 | 0 | 2 | 14 | 16 |
| 合計 | 14 | 16 | 14 | 16 | 60 |

| | | | |
|----------|-------|--|-------------|
| P0 | 0.633 | | 0.633 |
| Pe | 0.251 | | 0.240 |
| κ | 0.510 | | 0.519 0.518 |

| P0, κ | | | |
|--------------|-------|-------|-------|
| | x1 | x2 | x3 |
| x1 | | 0.600 | 0.605 |
| x2 | 0.700 | | 0.351 |
| x3 | 0.700 | 0.500 | |

骨萎縮度が「正常」、「I度」、「II度」、「III度」を1, 2, 3, 4 と変換して表している . この4つのカテゴリーは , 順序尺度と思われるが , ここでは 名義尺度 として解析する . 順序尺度としての解析は次の章で説明する .

(2) VBA マクロによる解析

繰返しが3回以上になると , 計算表上で計算するのは実用的ではない . そこで , 前節で説明した VBA マクロによる解析結果を示す . 上には 原著 表2.7 p.42 のクロス表が , その下には , 原著 p.43 の上の式で計算される P_0 , P_e , κ が求められている .

下の表は , 原著 表2.7 の 評価者の組合わせごとの P_0 , κ である .

ここに得られた評価者の組合わせごとの κ を総合化する方法が原著のp.44 に書かれている。
 一つは、全部の κ の平均を取る方法で、表示2.5 の中央の右に 0.519 と求められている。
 他方は、それぞれの組合わせごとの P_0, P_e からそれらの平均値を求め、それから κ を計算する方法である。 P_0, P_e の平均が 0.633, 0.240 で、 κ は 0.518 となる。
 これは繰返しが3回以上のときにのみ出力される。

(3) 別の例

原著 §4.4 (p.88) には、患者数 $n = 20$, 評価者 $p = 11$ のデータが取り上げられている。このように、 n, p が大きいデータでも、マクロを使えば簡単に解析できる。
 表示2.6 に出力を示す。

表示2.6: VBA マクロによる計算 (くり返し11回)

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 合計 |
|----|-----|-----|-----|-----|-----|-----|-----|------|
| 1 | 72 | 88 | 0 | 0 | 0 | 0 | 0 | 160 |
| 2 | 88 | 144 | 51 | 2 | 5 | 0 | 0 | 290 |
| 3 | 0 | 51 | 130 | 50 | 29 | 0 | 0 | 260 |
| 4 | 0 | 2 | 50 | 154 | 59 | 24 | 1 | 290 |
| 5 | 0 | 5 | 29 | 59 | 248 | 134 | 5 | 480 |
| 6 | 0 | 0 | 0 | 24 | 134 | 232 | 60 | 450 |
| 7 | 0 | 0 | 0 | 1 | 5 | 60 | 204 | 270 |
| 合計 | 160 | 290 | 260 | 290 | 480 | 450 | 270 | 2200 |

| | | |
|----------|-------|-------------|
| P_0 | 0.538 | 0.538 |
| P_e | 0.159 | 0.156 |
| κ | 0.451 | 0.454 0.453 |

| | P_0, κ | | | | | | | | | | |
|----|---------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 1 | | 0.463 | 0.351 | 0.521 | 0.408 | 0.407 | 0.349 | 0.640 | 0.412 | 0.405 | 0.472 |
| 2 | 0.550 | | 0.634 | 0.405 | 0.568 | 0.337 | 0.450 | 0.353 | 0.425 | 0.304 | 0.366 |
| 3 | 0.450 | 0.700 | | 0.459 | 0.640 | 0.399 | 0.578 | 0.193 | 0.532 | 0.403 | 0.407 |
| 4 | 0.600 | 0.500 | 0.550 | | 0.645 | 0.641 | 0.642 | 0.461 | 0.412 | 0.508 | 0.405 |
| 5 | 0.500 | 0.650 | 0.700 | 0.700 | | 0.758 | 0.756 | 0.242 | 0.489 | 0.253 | 0.432 |
| 6 | 0.500 | 0.450 | 0.500 | 0.700 | 0.800 | | 0.641 | 0.471 | 0.359 | 0.290 | 0.413 |
| 7 | 0.450 | 0.550 | 0.650 | 0.700 | 0.800 | 0.700 | | 0.298 | 0.599 | 0.355 | 0.366 |
| 8 | 0.700 | 0.450 | 0.300 | 0.550 | 0.350 | 0.550 | 0.400 | | 0.292 | 0.281 | 0.472 |
| 9 | 0.500 | 0.500 | 0.600 | 0.500 | 0.550 | 0.450 | 0.650 | 0.400 | | 0.522 | 0.578 |
| 10 | 0.500 | 0.400 | 0.500 | 0.600 | 0.350 | 0.400 | 0.450 | 0.400 | 0.600 | | 0.514 |
| 11 | 0.550 | 0.450 | 0.500 | 0.500 | 0.500 | 0.500 | 0.450 | 0.550 | 0.650 | 0.600 | |

上には原著 表4.7 (p.92) の左が、中央の左には p.92 の本文中の計算式の値が、下には表4.6 が得られている。

中央の右には2つの総合した κ が求められている (原著には示されていない)。

2.3 順序分類データの解析

表示2.5のデータを順序尺度として解析する方法を検討する。

原著 p.58 には、順位に置換えて、順位を連続量として §1 の方法で解析する方法が説明されている。

ここでは、カテゴリー番号をそのまま連続量として ICC を計算した結果と、順位に変換してから連続量として ICC を計算した結果を示す。

表示2.7: 連続量としての解析

| 列数 | 3 | | | 平方和 | | | | | 平方和 |
|-----|--------|----|----|-----|-----|--------|------|------|-------|
| | x1 | x2 | x3 | | | x1 | x2 | x3 | |
| 1 | 1 | 1 | 1 | 0.0 | 1 | 4.0 | 4.0 | 4.0 | 0.0 |
| 2 | 1 | 1 | 2 | 0.7 | 2 | 4.0 | 4.0 | 11.5 | 37.5 |
| 3 | 2 | 1 | 2 | 0.7 | 3 | 11.5 | 4.0 | 11.5 | 37.5 |
| 4 | 1 | 2 | 3 | 2.0 | 4 | 4.0 | 11.5 | 19.0 | 112.5 |
| 5 | 2 | 2 | 2 | 0.0 | 5 | 11.5 | 11.5 | 11.5 | 0.0 |
| 6 | 3 | 3 | 2 | 0.7 | 6 | 19.0 | 19.0 | 11.5 | 37.5 |
| 7 | 3 | 3 | 3 | 0.0 | 7 | 19.0 | 19.0 | 19.0 | 0.0 |
| 8 | 4 | 3 | 4 | 0.7 | 8 | 26.5 | 19.0 | 26.5 | 37.5 |
| 9 | 4 | 4 | 4 | 0.0 | 9 | 26.5 | 26.5 | 26.5 | 0.0 |
| 10 | 4 | 4 | 4 | 0.0 | 10 | 26.5 | 26.5 | 26.5 | 0.0 |
| 平方和 | 37.5 | | | 4.7 | 平方和 | 2107.5 | | | 262.5 |
| ICC | 0.8132 | | | | ICC | 0.8132 | | | |

表示2.7の左には表示1.17と同じ手順で、元データから ICC が計算されている。ICC は 0.813 である。

右には、測定値を順位に変換した値が求められている。

x1 の列の最初の値 4.0 は

$$=RANK(C6, \$C\$6:\$E\$15, 1) + (COUNTIF(\$C\$6:\$E\$15, C6) - 1) * 0.5$$

という関数で計算されている。

RANK 関数の最初のパラメータは x_{11} のセル、2番目のパラメータは x_{ij} 全体で、最後の 1 は昇順に順位をつけることを表している。RANK 関数は同値があるとき、最初の順位が出力される。

1 は 7 個あるので、その平均順位は 1 ~ 7 の平均 4.0 となる。平均順位は、最初の順位 1 に (同値の個数 - 1) * 0.5 = (7 - 1) * 0.5 = 3.0 を加えて求められる。

同じ値の個数を数えるのが COUNTIF 関数である。パラメータは RANK 関数の逆である。

$x = 1, 2, 3, 4$ に対する平均順位は 4.0, 11.5, 19.0, 26.5 となる。

原著では、2組の x_{ij} について順位を求めているので、対応する平均順位は 7.5, 22.5, 37.5, 52.5

となっている．この数値は上の平均順位の2倍から 0.5 を引いて変換できる．

表示2.7 の右で，平均順位についてICC が計算され，0.813 が得られる．この値は 原著 p.59 の中央に求められている値に一致する．

カテゴリ番号と平均順位のいずれを使っても，同じ ICC が得られた．これは 平均順位の間隔がすべて 7.5 と等間隔になっているためで，いつも同じになるものではない．

いずれが良いかについては，次に私見を述べる．

蛇足 1 順位を使うことの可否について

この例では， $x_1 = 1 \sim 4$ の個数は 7, 8, 7, 8 個 とほぼ等しい．もし，この個数が 3, 12, 12, 3 のように，山のある分布になっているときには，平均順位は 2.0, 9.5, 21.5, 29.0 となり，間隔は 7.5, 12.0, 7.5 と中央の間隔が広がる．

診断で 2 を 3 と誤る損失と，1 を 2 と誤る損失を比較すると，後者の方が大きいであろう．ICC は，観測値の差の大きさから不一致度を求め，1 から引いて求めるものであった．

順序尺度を連続量の評点に変換するときは，評点の差が 評価の誤りによる損失に比例するようにするのが良いと思われる．

連続量を順位に変換して検定する ノンパラメトリック検定の中に，順位をそのまま使わず，正規分布に近くなるように変換する方法がある (Van del Waerden の検定) ．

$n = 30$ のとき，1位と2位の差は 15位と16位の差よりも大きいという考えに基づくものである． i 番目は標準正規分布の下側確率が $i/(n+1)$ である u の値に変換される．

$i = 1, 2, 15, 16$ に対しては

| i | 1 | 2 | 15 | 16 |
|-----------|--------|--------|--------|-------|
| $i/(n+1)$ | 0.048 | 0.095 | 0.476 | 0.524 |
| u | -1.667 | -1.309 | -0.060 | 0.060 |
| 差 | 0.359 | | 0.119 | |

1, 2 位の差は，15, 16 位の差の3倍以上となっていることが分かる．

個数の分布の中央付近に山があり，正規分布に近い場合は，Van del Waerden の検定 で用いられている評点は，元のカテゴリ番号と高い相関を持つ．このような場合は，順位に変換するよりも，カテゴリの順序番号をそのまま使う方が良いかと思われる．

蛇足 2 重みつき評価

原著の p.60 以降には，評価のずれの大きさを考慮して，§1, §2 で取り上げた κ を求める方法が説明されている．

そこでは、1 ランクの違いは合格とする方法、ランクの違い 1, 2, 3, 4, ... に対する重みを 1, 2, 3, 4, ... , または 1, 4, 9, 16, ... とする方法などが説明されている。

この重みのつけ方は、数理統計学の立場で決められるものではなく、評価の誤りがどれだけの損失（誤診による死亡など）を考慮して決められるべきであろう。

無難な方法として、上に述べたカテゴリーの番号をそのまま使う方法があるものと思われる。

順序に意味のない場合にも、重みつき評価を用いることが好ましい場合があるであろう。

以下は筆者（素人）の常識的な知識に基づく考えである。

脳卒中は、脳血管が破れた 脳出血 と、脳血管が詰まった 脳出血 に大別される。脳出血であれば、出血を止めるために、血液凝固薬が投与され、脳梗塞であれば、血栓を溶かすために、血液溶解薬が投与される。二つの薬剤は逆の効果があり、両者の誤診は命にかかわる。

脳出血は、出血した部位によって分けられる。くも膜下出血 であるかそれ以外の部位の出血であるかが診断される。

診断が「脳梗塞」、「くも膜下出血」、「その他の部位の出血」の3つであるとき、誤診による損失は大きく異なると思われる。

このように、カテゴリーが 名義尺度 である場合にも、重みつきの κ が役立つ場合があると思われる。

