

Title	歯科データの統計的分析について
Author(s)	菅野, 隆三
Journal	歯科学報, 101(2): 205-218
URL	http://hdl.handle.net/10130/253
Right	

歯科データの統計的分析について

菅野 隆三

東京歯科大学数学研究室

1. はじめに

近年、パソコンの性能がよくなり、統計ソフトも普及してきたので、歯科データの分析にもかなり複雑な統計の手法が使える状況にある。しかし、統計的分析が役立つためには目的に合った適切なデータを収集することが基本的に重要であり、その上で、データを分析するのにどのような統計の手法を適用すべきかが問題となる。さらに、得られた結果に対する正しい判断と適切な解釈がなされる必要がある。

本稿では、筆者がこれまでに学内で統計相談を受けてきて感じたことをもとに、上記のような観点からデータ収集と統計的分析に当たって注意すべきこと、さらに、歯科データに適用して有効と思われるいくつかの統計的手法について述べてみた。

データの収集については、データの尺度 (scale) に注意すべきであることと、特に実験データの収集に関しての注意を述べることにした。動物実験や臨床試験での処理効果の有意差を検定する問題に関する統計相談が多いが、データに偏り(バイアス bias)や系統誤差があったとすると、検定での有意差が、処理効果によるものかどうか疑わしくなる。データをとるに当たっては、偏りや大きな系統誤差が混入しないように実験の場を管理し、処理の割り付けにも細心の注意が必要となる。

統計的検定については、検定に当たっての注意事項をいくつか述べておいた。なお、多群の処理効果の検定がしばしば問題となるので、それに関して多重比較法の中で最も一般的な Tukey の方法を具体例で紹介することにした。歯科の研究報告を見ると、多重比較法を適用すべきなのに使っていない場合が多いように思われるからである。

次に、統計相談で要因分析に関するものがしばしば見受けられるので、多数の特性を同時に取り上げて分析する多変量解析法の中から、特に重回帰分析と判別分析の適用について述べることにした。これらの手法は適用範囲が広く、その分析結果の判断もしやすく、初心者でも大きな誤りを犯すことなく使えるので便利である。この他に、主観的評価の尺度化に関して、Thurstone の一対比較法を紹介することにし、その使用法を具体的な例で示しておいた。歯科の分野でも QOL が重要視されてきているので、人による好き嫌いや良し悪しの感覚的判定を数量的に尺度化する統計的手法も必要になると思われるからである。

2. データの収集について

歯科データといっても、研究目的と研究対象によって、実験データか観察データか、あるいは対象集団についての調査データかの違いがある。いずれの場合であってもデータを収集するに当たってはそれぞれの予備知識が必要になる。実験デー

データの収集には、統計学の一分野である実験計画法の知識が役立つし、調査データの場合には、標本調査法、アンケート調査法などの知識を活用すべきである。観察データでは、横断的観察(cross sectional)か縦断的観察(longitudinal)かの選択が問題になる。また、疫学で問題となる原因-結果に対する関連性を研究する場合には、前向き研究か後向き研究かでデータの収集の仕方が異なってくる。これらのことに関しては、多くの書物(例えば、奥野ほか⁵⁾、林ほか²⁵⁾、辻ほか²¹⁾、J. L. フライス²⁷⁾、宮原ほか²⁹⁾)に説明があるので参照されたい。ここでは、データの尺度(scale)の問題と、実験誤差の制御に関する注意だけを述べておくことにする。

2.1 データの尺度

取り扱う対象によってデータの尺度に制約がある。どの尺度を用いれば信頼のおける情報が得られるかを考えるべきである。データの尺度としては、次のものがある。

- (a) 分類尺度(名義尺度)
- (b) 順序尺度
- (c) 間隔尺度(距離尺度)
- (d) 比例尺度

データのもつ情報の詳しさの水準は、(a)から(d)へと順に高くなる。分類尺度では、カテゴリーに分類されたデータになるが、もし、カテゴリーに順序がある場合には順序尺度のデータとなる。データを質的データと数量データに分類する場合は、分類尺度と順序尺度は質的データであり、間隔尺度と比例尺度は数量データに属する。間隔尺度は数量データで距離をもつが、意味のある原点0をもたない。例えば、温度の摂氏Cなど。一方、比例尺度は距離と原点0に意味がある(比に意味がある場合で、例えば、長さ、重さ、金額など。一般に、実験や観測で得られる数量データは比例尺度にあり、間隔尺度にとどまるものは少ない。ただし、人の感覚的な嗜好の程度などを尺度化したものは間隔尺度になる。データが上記(a)~(d)のいずれであるかによって、適用可能な統計的分析の手法も異なってくる。一般には、情報の

多い数量データで統計的分析を行うのが望ましいが、統計的手法には、付帯条件があるのが普通であり、もし、その条件を満たさない場合には、尺度の水準を順序尺度か分類尺度へ下げて、データの順序だけを利用するか、または、分類した度数を用いる統計的手法を適用することになる。

2.2 実験データの収集に関する注意

信頼のおける良質のデータを収集するためには、実験に当たって次の2つの点を考慮する必要がある。

- (i) データの情報を増やすための工夫
- (ii) データの精度を上げるための工夫

この内の(i)については実験計画法の書物(例えば、奥野ほか⁵⁾)を見ていただくことにして、ここでは(ii)について述べておく。

実験データの誤差には、系統誤差(systematic error)と偶然誤差(random error)がある。系統誤差の発生要因としては、例えば、実験の場における温度や湿度の変化、実験操作の不慣れ、動物実験での個体差や生育環境などがある。測定においては、系統誤差が発生すると思われる要因を制御し、その発生をできるだけ防ぐことが必要である。系統誤差が完全に除去されたとしても、なおかつ測定値には微小な誤差が残るが、それは人為的には制御できない偶然的な要因から生じるもので偶然誤差とよばれるものである。それは確率的な変動をするものとして取り扱われる。

実験誤差の制御については、Fisherの3原則(反復、無作為化、局所管理)がよく知られており、それを考慮して実験計画を立てるのが望ましい。反復測定はデータのバラツキの大きさ(誤差分散)を評価するのに必要なことである。実験での処理効果を評価するために統計的推測(推定や検定)を行うときに、この誤差分散が使われる。無作為化とは、処理の割り付けを無作為化(randomization)することによって系統誤差を偶然誤差へ転化することを目的としている。例えば、臨床実験で2つの処理A、Bの効果を比較するとき、被験者を無作為に2群に分け、一方の群に処理Aを施し、他方の群に処理Bを施して

その効果を比較することが行われるが、これは無作為化によって被験者の個体差などから生じる系統誤差を両群で均一化するためである。なお、無作為に2群に分けることによって、系統誤差が両群での測定値に偶然性をもって混入するから偶然誤差と見做せるというわけである。これは系統誤差を偶然誤差へ転化したことになる。最後にの局所管理とは、系統誤差の除去を目的とするものである。上記の臨床試験の場合であれば、各被験者ごとに処理A、Bを施して、その効果を比較することである。これによって系統誤差の大部分を占める個体差が除去されることになる。ここでのA、Bの組はブロック(block)とよばれるが、これは「被験者」をブロック因子として取り上げA、Bの処理をブロック内で局所管理していることになる。なお、ブロック内での処理の割り付けは無作為化する。例えば、実験計画法における乱塊法(randomized block design)の実験はFisherの3原則を満たすものになっている。

3. 統計的検定について

統計的検定の目的は、標本データから予想される結果が標本変動を考慮しても一般的に成り立つ(母集団で成り立つ)としてよいかどうかを確率的にチェックすることにある。それには検定統計量の分布が用いられるのであるが、母集団分布の型を設定して導いた検定統計量の分布を用いる場合はパラメトリック検定といい、母集団分布の型には無関係に導いた検定統計量の分布を用いる検定はノンパラメトリック検定とよばれる。一般に、データの尺度が分類尺度か順序尺度であるときは、ノンパラメトリック検定が用いられる。例えば、分割表の度数データによる適合度検定(χ^2 検定)はノンパラメトリック検定である。検出力の点では、パラメトリック検定の方が検出力が高いので、それを用いる方が望ましいとされる。しかし、それが使えるためには付帯条件がいろいろあり、その条件からのずれを無視して使えば当然検出力は悪くなる。一方、ノンパラメトリック検定は付帯条件なしで使えるものが殆どであり、その

検出力もパラメトリック検定と比べてそれほど劣らないもの(順位を使った検定など)もあるので、もっと活用すべきであると思われる。ノンパラメトリック法についての書物としては、例えば、Siegel / Castellan¹⁴⁾, E. L. レーマン³³⁾, 岩原³⁾が例題も多く出ており参考になる。以下に検定に当たった際の注意事項を参考までに述べておく。

3.1 検定法の選択について

検定すべき内容によって検定法が異なってくる。検定する内容が分布の中心的位置(平均値や中央値)なのかバラツキの度合い(散布度:レンジ, 分散など)なのか、あるいは割合(比率)なのかによって、用いる検定統計量が異なるからである。

また、データの尺度によって検定法が制限される。分類尺度の場合は、度数データを用いた検定(一般に、 χ^2 検定)、順序尺度のときには順位を用いた検定がノンパラメトリック法として開発されている。間隔尺度や比例尺度の場合は数量データとして、一般のパラメトリック検定法の適用を考えることになる。しかし、それが使えるための付帯条件をチェックする必要がある。なお、その条件が正確には満たされず、多少のずれがあっても使える(検定の頑健性 robustness) 場合が多いことは知っているべきである。しかし、データの分布が明らかに偏っているときには、正規性を仮定した検定を使うことには無理がある。その場合は、データの変換を考えるかノンパラメトリック検定を使うことになる。

[例] 数量データによる2群の平均値の比較

独立な2標本であるか、それとも対応のある2標本(例えば、2つの処理を同一の被験者に施して得られたデータなど)であるかの区別をして検定する必要がある。

対応のある2標本データのとき、対(ついで)になったデータの差が正規分布をすると見てよければ、1標本でのStudentのt検定が使える。正規分布の仮定に無理があれば、ノンパラメトリック法の「Wilcoxonの符号付き順位和検定」を用いることになる。

独立な2標本のとき、それぞれのデータが正規分布に従うと仮定でき、しかも「等分散の検定」で有意ならなければ、等分散と見做して「2標本のStudentのt検定」が使える。もし、等分散が仮定できないときは、「2標本のWelchの検定」を用いることが多い。正規分布の条件は、検定の頑健性から厳密に成り立つ必要はないが、データの分布がかなり偏っていて正規分布の仮定が最初から無理なときは、ノンパラメトリック法の検定である「Wilcoxonの順位和検定」(Mann-Whitneyの検定と同じ)を用いることになる。

3.2 検定と標本の大きさについて

この問題については、検定における2種類の過誤のことを知っておく必要がある。有意水準 α で帰無仮説 H_0 を棄却すると判定したときは、「 H_0 が本当は正しいのに、それを正しくないと判定してしまう誤り」を犯す可能性がある。これを検定における第1種の過誤(α 過誤)といい、その誤りを犯す確率が有意水準 α である。また、帰無仮説 H_0 を棄却しなかったときには、「 H_0 が本当は正しいのに、それを正しいと判定してしまう誤り」を犯す可能性があり、これを第2種の過誤(β 過誤)という。これを犯す確率 β は対立仮説 H_1 に依存する。なお、 $1 - \beta$ は「 H_0 が正しくないとき、それを棄却する確率」(有意差を検出する確率)になり検出力(power)と呼ばれるもので、その値は大きいほど望ましい検定となる。検定としては、2種類の過誤を犯す確率をできるだけ小さくするようにしたいが、両方を同時に小さくすることはできないので、普通は α 過誤の起こる確率(有意水準)を一定にしておいて、 β 過誤の起こる確率をなるべく小さくする(検出力 $1 - \beta$ を大きくする)ような検定方式を考えている。なお、 α を一定にしておけば、標本の大きさ n を大きくすることによって β を小さくする(検出力 $1 - \beta$ を大きくする)ことができる。

以上のことから、検定に必要な標本の大きさ n を決めるには、実質的に有意な差 d を見積もっておき、有意水準 α を一定にしたとき、この差 d

を検出する検出力 $1 - \beta$ をどの程度に設定するかを考える必要がある。標本の大きさの決め方については丹後¹⁸⁾に詳しく説明が出ているので参照されたい。

ここでは、1つの例として独立な2標本での平均値の差の検定における「2標本のstudentのt検定」の場合を取り上げてみる。データは2つの正規母集団 $N(\mu_1, \sigma_1^2)$ 、 $N(\mu_2, \sigma_2^2)$ より標本抽出された、それぞれ大きさ n_1 と n_2 の標本であるとする。2標本のstudentのt検定では等分散 $\sigma_1^2 = \sigma_2^2 = \sigma^2$ が仮定されている。いま、両方のデータの大きさが等しく $n_1 = n_2 = n$ であるとして、検出したい有意差 d を次のようにとる。

$$\text{片側検定のとき, } d = (\mu_1 - \mu_2) / \sigma$$

$$\text{両側検定のとき, } d = |\mu_1 - \mu_2| / \sigma$$

このとき、有意水準 α を一定として、検出力 $1 - \beta$ と有意差 d の値を種々に設定したときの標本の大きさ n が、Cohen⁹⁾によって計算され数表に示されている。丹後¹⁸⁾にも引用されて出ている。例えば、片側検定で $\alpha = 0.05$ (両側検定で $\alpha = 0.10$)のとき、種々に設定した d と $1 - \beta$ の値に対する標本の大きさ n は次のようになる。

検出力 $1 - \beta$	有意差 d				
	0.5	0.6	0.7	0.8	1.0
0.70	38	27	20	15	10
0.80	50	35	26	20	13
0.90	69	48	36	27	18

これを見ると、有意水準5%の片側検定で差 $d = 1.0$ を有意差として80%の確率で検出するためには、 $n_1 = n_2 = n = 13$ の標本データが必要になる。

3.3 多数の処理効果の検定

(i) 分散分析法による検定

多数の処理効果の有意差を検定するとき、一般に分散分析法の検定が用いられる。その検定法は、データがどのような実験配置によって得られたものであるかによって異なってくる。分散分析法による検定は、Fisherによって開発されたも

のであるが、それは問題とする処理の分散と誤差分散との比を検定統計量として用いるもので、F分布を使った検定になる。ただし、データについて正規性と等分散性を前提としている。なお、分散分析法による検定は、多数の処理効果(処理因子の各水準での効果)を一括して有意差を検定するものであり、その検定で有意になったときは、多数の処理の内のいずれかに有意差があることを示すだけである。「歯科学報」にも分散分析法を適用した研究報告がときどき見受けられる。例えば、吉成ほか³¹⁾は2因子実験の分散分析を適用している。また、武田¹⁶⁾は乱塊法の分散分析を使った報告をしている。種々の実験配置に対する分散分析法による検定が実験計画法の書物に出ているので参照されたい。

(ii) 多重比較法による検定

分散分析法による検定で、多数の処理効果を一括して有意差を検定した結果、処理間に有意差ありと判定されたとき、それではどの処理の間に有意差があるかを調べるのが次の問題となる。それを調べる時多重比較法の検定が用いられる。その検定の特徴は、各処理の多数の組合せにおける有意差を同一の判定規準で検定できるという便利さと、多数の同時検定の結果を総合して判断するときの有意水準を一定にしてあるので合理性があることである。もし、処理の組合せ毎に有意水準を、例えば5%ずつで、単独の検定として多数繰り返す、その検定の結果を総合して判断をしたとすれば、全体の有意水準は5%より大きくなってしまい理論的に間違いとなる。

多重比較法の検定では、a個の処理をどのような組合せで比較するかによって検定法が別れる。代表的なものを上げると次のようになる。

a個の処理の全ての対(pair)を比較するときは Tukey の方法

1つの処理(対照群)と、残りの(a-1)個の処理(実験群)の各々との対を比較する時は Dunnett の方法

a個の処理の全ての対比(処理の線形結合で、その係数の和が0)の比較をするときは

Scheffé の方法

なお、Tukeyの方法とDunnettの方法では、各処理でのデータの大きさは等しいと仮定されているが、多少の違いがあっても使って支障はない。多重比較法の各手法についての説明と適用例が出ている書物としては、吉村³²⁾が大変参考になると思われる。また、丹後¹⁹⁾には、ノンパラメトリック検定での多重比較法も出ているので参照されたい。

多重比較法を適用した研究報告を近年の「歯科学報」で見ると、Tukeyの方法を使ったものとして武田¹⁶⁾があり、Schefféの方法を用いたものとして安東¹⁾がある。

以下に、多重比較法の使い方をTukeyの方法の場合で示しておく。データは上記の武田¹⁶⁾より一部借用したものである。

[例] 各種食品の咀嚼時間の比較

データは、実験計画法の乱塊法による一因子実験(処理:食品, ブロック:被験者)によって取られたものである。乱塊法の実験配置では、Fisherの3原則が全て適用されている。すなわち、被験者をブロックとすることで個体差の影響を除去しており、因子の各水準(各食品)での測定は5名の被験者で反復されており、さらに、各被験者での食品の割り付けの順序が無作為化されている。表3-1に実験でのデータが示されている。被験食品の大きさは各1cm³であり、データの時間単位は秒である。

この表3-1のデータに乱塊法の分散分析による検定を適用した結果は表3-2の分散分析表に示されている。それによると、食品間に有意水準1%で有意差が認められる。次に、どの食品間に有意差があるかを8種類の全ての対で比較してみる。ここでは、各食品におけるデータの大きさは全て同じであり、すべての対の比較をするのでTukeyの多重比較法を用いるのが妥当である。

この検定における判定規準は、第i番目と第k番目の食品の平均咀嚼時間 \bar{X}_i , \bar{X}_k の差の絶対値に関して

$$| \bar{X}_i - \bar{X}_k | > q_{\alpha}(a, ar - a) \sqrt{V_e / r}$$

で与えられる。この例では、処理の数 $a = 8$ 、ブロックの数 $r = 5$ であり、いま、有意水準 $\alpha = 0.05$ とすると $q_{0.05}(8, 32)$ の値を「student 化した範囲のパーセント点」の数表から求めればよい。ただし、数表からは $q_{0.05}(8, 30) = 4.6014$ と $q_{0.05}(8, 40) = 4.5205$ が読み取れるだけなので、ここでは、近似的に $q_{0.05}(8, 32) = 4.6014$ とする。また、 V_e の値は乱塊法の分散分析表より $V_e = 1.387$ となる。これらの値を用いて計算すると、判定規準は次のようになる。

$$|X_{i.} - X_{k.}| > 2.4235$$

表 3-3 には、各食品の平均咀嚼時間の差の絶対値 $|X_{i.} - X_{k.}|$ を求めた結果が示してある。この表 3-3 の数値で 2.4235 より大きい値である箇所は有意差ありと判定される。例えば、表 3-3 の第 1 行はトウフと他の食品との平均咀嚼時間の差の絶対値であるが、トウフと有意差が認められるのはカマボコ、ニンジン、サラミ、タクアンであることになる。なお、トウフに対してバナナ、チーズ、ダイコンの各々は有意差が認められな

い。

さて、以上の Tukey の方法で 8 種類の食品のすべての対 (28 組) を検定したとき、有意差の認められたものは 12 組あり、表 3-3 で網掛けをして示してある箇所である。この 28 組全体の同時検定における有意水準は 5% となっている。

4. 多変量解析法について

研究対象について、複数個の特性を同時に取り上げて分析するときの統計的手法として多変量解析法がある。その手法には、重回帰分析、判別分析、主成分分析、因子分析、正準相関分析、クラスター分析などがある。各手法がどの程度有効かは専門分野によって異なるし、研究目的にもよるので、その有効性は経験的に知るしかない。これらの手法を用いて有効な結果を得るためには、機械的に適用するだけでは駄目で、種々の面から検討を加えながら試行錯誤を重ねて、実質的に有効な結果を導き出す手腕が要求される。その意味で経験を要するわけであるが、具体的な適用例を上

表 3-1 各種食品の咀嚼時間(乱塊法による一因子実験)

被験者	トウフ	バナナ	チーズ	カマボコ	ダイコン	ニンジン	サラミ	タクアン
1	5.37	4.66	5.97	9.49	4.89	11.35	7.73	8.20
2	2.64	3.80	4.75	4.10	4.68	6.63	6.88	4.95
3	2.50	3.08	4.89	7.08	8.01	7.48	10.45	7.04
4	4.52	3.74	5.50	7.87	5.75	7.54	8.81	8.00
5	4.28	3.96	5.41	4.80	4.91	7.55	7.33	5.21
平均 $X_{i.}$	3.862	3.848	5.304	6.668	5.648	8.110	8.240	6.680

[注] 被験食品の大きさは各 1 cm^3 、時間単位は秒。(武田¹⁶⁾1996による)

表 3-2 各種食品の咀嚼時間に関する乱塊法の分散分析表
(処理 A: 食品 ブロック R: 被験者)

変動因	平方和	自由度	分散	分散比
処理 A	$S_A = 100.863$	$f_A = 7$	$V_A = 14.409$	$F_A = 10.389^{**}$
ブロック R	$S_R = 28.159$	$f_R = 4$	$V_R = 7.040$	$F_R = 5.076^{**}$
誤差 e	$S_e = 38.833$	$f_e = 28$	$V_e = 1.387$	
全体 T	$S_T = 167.855$	$f_T = 39$		

** : $p < 0.01$

表 3-3 各種食品の平均咀嚼時間の差の絶対値：|X_{i.} - X_{k.}|の値

	トウフ	バナナ	チーズ	カマボコ	ダイコン	ニンジン	サラミ	タクアン
トウフ		0.014	1.442	2.806	1.786	4.248	4.378	2.818
バナナ			1.456	2.820	1.800	4.262	4.392	2.832
チーズ				1.364	0.354	2.806	2.936	1.376
カマボコ					1.020	1.442	1.572	0.012
ダイコン						2.462	2.592	1.032
ニンジン							0.130	1.430
サラミ								1.560
タクアン								

[注] 多重比較法の Tukey の方法における有意性の判定規準は、この例では
 $|X_{i.} - X_{k.}| > 2.4235$
 なお、網掛けの箇所は「有意差あり」(28組の比較全体での有意水準 5%)

げて説明してある書物(例えば、奥野ほか^{6,7)}、杉山¹⁵⁾)を参考にするとよいであろう。多変量解析法の中で重回帰分析と判別分析は適用範囲が広く、その結果の判断も容易で特に大きな誤りを犯すことが少なく比較的使いやすいので、多いに活用してよい手法といえる。「歯科学報」にも重回帰分析と判別分析を適用したものが多少みられるが、まだまだ活用の余地があるように思われる。ここでは、この2つの手法について未経験な方々のために、その使用上の注意と適用例を紹介しておくことにする。

4.1 重回帰分析について

重回帰分析は、一般に、予測に用いる多変量解析法として広く知られている。現象の結果特性 y を目的変数、p 個の原因特性 x₁, x₂, ..., x_p を説明変数として、x₁, x₂, ..., x_p の値から y の値を予測する問題に重回帰分析が用いられる。その予測式が重回帰式と呼ばれるものであるが、一般に、x₁, x₂, ..., x_p の 1 次式として次のようにおき、n 組のデータ(y_i, x_{1i}, x_{2i}, ..., x_{pi}), i = 1, 2, ..., n に基づいて定数項 b₀ と係数(偏回帰係数) b₁, b₂, ..., b_p を決定する。

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_p x_p$$

ここに、 \hat{y} は y の予測値である。最小 2 乗法によれば、データ y_i と予測値 \hat{y}_i との残差 e_i = y_i

- \hat{y}_i の 2 乗和 $\sum (y_i - \hat{y}_i)^2$ を最小にするように b₀, b₁, ..., b_p の値を決めることになる。このことは、y と \hat{y} との相関を最大にすることと同じになる。この y と \hat{y} の相関係数は重相関係数 R と呼ばれるもので、常に、0 ≤ R ≤ 1 である。その 2 乗である R² は決定係数(または寄与率)と呼ばれるもので、p 個の説明変数 x₁, x₂, ..., x_p で y の変動の何%まで説明できるかを示すものである。上記のように求めた重回帰式は、決定係数 R² を最大にするように決定したのになっている。

[例] 新生児体重の予測

新生児の体重 y を、母親の体重 x₁ と母親の年齢 x₂ および懐妊期間 x₃ の 3 つの説明変数を用いて予測することを重回帰分析で行ってみよう。ここでは、佐和¹³⁾に出ている n = 15 名のデータを用いることにする。なお、佐和¹³⁾では説明変数 x₄ として母親の喫煙習慣の有無(有りならば 1, 無しならば 0)を取り上げて、重回帰式を 4 つの説明変数から導いている。ここでは、上記の 3 つの説明変数から重回帰式を求めてみると、その結果は表 4-1 のようになる。それを見ると、決定係数 R² = 0.8464 であり、この 3 つの説明変数 x₁, x₂, x₃ から y の変動の 84.64% まで説明できることがわかる。なお、表 4-1 に「自由度調整済み決定係数 R^{*2}」が示してあるが、決定係数 R² の値は説

表 4 - 1 新生児の体重の重回帰分析

説明変数 x_i	偏回帰係数 b_i	標準偏回帰係数	偏相関係数
x_1 : 母親の体重 (kg)	55.9807	1.0103	0.9187
x_2 : 母親の年齢 (歳)	-27.9464	-0.3471	-0.6335
x_3 : 懐妊期間 (日)	8.9213	0.3574	0.6615
定数 b_0	-1697.70	0	

重回帰式: $\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3$, 重相関係数 $R = 0.920$,
決定係数 $R^2 = 0.8464$, 自由度調整済み決定係数 $R^{*2} = 0.8045$

明変数の個数 p が大きいほど大きくなるので、データの大きさ n の割には p が大きいときは、その影響を除去した R^{*2} を用いるのがよい。この例では、 $R^{*2} = 0.8045$ となって寄与率は多少低下して約80%となる。

さて、ここでは、新生児体重 y の予測に重回帰式を用いたのであるが、重回帰式は要因分析を目的としても使われる。要因分析の立場からは、説明変数 x_1, x_2, x_3 の各要因がそれぞれの程度目的変数 y に寄与するのを知りたいであろう。それを見るために、表 4 - 1 に標準偏回帰係数 b'_i と偏相関係数が示されている。それらの絶対値の大小で各説明変数の y に対する影響力を見ることができる。ただし、標準偏回帰係数 b'_i について注意したいことは、説明変数が互いに無相関であれば b'_i は x_i と y との相関係数 r_{iy} に一致するので、この判断は正しいことになる。しかし、説明変数間には相関があるのが普通であるから、それは近似的な判断となる。実用的には、説明変数間の相関係数 r_{ij} の絶対値が0.3以下程度に小さいなら使ってよい (cf: 奥野ほか⁷⁾) と言われている。上の例では、表 4 - 2 のように x_1 と x_2 の相関係数 $r_{12} = 0.3614$ は0.3を少々越えているが、他は0.3より小さくなっており標準偏回帰係数で各説明変数の影響力を近似的に評価してよいであろう。一方において、 y に対する x_i の影響力を見るには x_i と y の偏相関係数を用いることが考えられる (cf: 駒沢¹⁰⁾)。 x_i と y の偏相関係数とは、 x_i と y の両方から x_i 以外の説明変数の影響を除去した上で x_i と y の相関の強さを見るも

表 4 - 2 相関行列

	x_1	x_2	x_3	y
x_1	1	0.3614	-0.2204	0.8061
x_2		1	0.0396	0.0322
x_3			1	0.1210
y				1

のである。従って、説明変数間に相関がある場合でも、 y と各説明変数との偏相関係数を求めれば、その絶対値の大きさを目的変数 y への各説明変数の影響力を評価できる。この例では、新生児体重 y に対しては母親の体重 x_1 の影響がかなり大きく、次に母親の年齢 x_2 と懐妊期間 x_3 が同程度に影響力をもつが、当然ながらその方向は逆になることがわかる。

重回帰式の妥当性は決定係数の値の大きさでわかるのであるが、その他に、重回帰式による予測値 \hat{y} とデータ y とのずれを示す残差 $e = y - \hat{y}$ を調べると、その妥当性が直観的に理解しやすくなる。残差の吟味は、重回帰分析に関する種々の問題点を検討するのに役立つので重要である。残差の検討の方法については多くの書物に出ているので、それを参照されたい。

次に、重回帰分析を適用するうえでの注意点をいくつか上げておく。これについては、奥野ほか⁶⁾の重回帰分析の章、または、適用例を中心に解説している重回帰分析の書物 (例えば、中村²⁴⁾, S.チャタジー / B.プライス²⁰⁾) を参照されたい。なお、重回帰分析に関する統計的推測 推

定と検定)については、奥野ほか⁶⁾を見れば十分であろう。

(1) 重回帰分析とデータの大きさについて

重回帰式が決定されるためには、数学的には $n - p - 1 > 0$ であればよい。しかし、重回帰式の係数の安定性と信頼性からは $n - p - 1 = 10 \sim 20$ になる程度にデータの大きさ n をとるのが望ましいと言われている。これは重回帰式を決定するに当たっての自由度が $10 \sim 20$ は欲しいということである。

(2) 重回帰分析でのデータの尺度について

重回帰分析では、一般に、説明変数 x_1, x_2, \dots, x_p と目的変数 y は共に間隔尺度とし扱われる。特別な場合として、説明変数の中に分類尺度(質的データ)の変数が入っているときは、ダミー変数を用いて重回帰分析が行われる。これについては、例えば、中村²⁴⁾、S.チャタジー/B.プライス²⁰⁾に適用例が出ているので参照されたい。なお、目的変数は数量データで、説明変数は全て質的データ(分類尺度)であるときは、数量化理論第Ⅰ類がこの場合の重回帰分析法になっている(cf: 駒沢¹⁰⁾)。

(3) 重回帰分析における説明変数の選択

上記の新生児体重の例では、予め指定した3つの説明変数から重回帰式を導いた。3つの説明変数でも自由度調整済み決定係数 R^2 が約80%になったが、さらに重回帰式の寄与率を上げることを考えるならば、他の要因を探して説明変数の個数を増やして重回帰式を作る必要がある。実用上からは、説明変数の個数 p はできるだけ少なくして目的変数 y を説明するのが望ましい。普通は $p = 5$ または 6 以内に押さえるのがよいといわれる。なお、説明変数間の相関が小さく、目的変数とは相関が高いものを説明変数として選ぶのが常識である。説明変数間に相関の高いものがあると多重共線性の存在が疑われ、偏回帰係数はデータに対して不安定となり、重回帰式の信頼性が失われることになる。ところで、重回帰分析の統計ソフトには変数選択のプログラムが入っているので、目的変数 y と関係のありそうな要因を

多数説明変数の候補にあげて、逐次選定法(stepwise method)により適切な要因を説明変数として選択することができる。逐次選定法にも種々の方法があるが、変数増減法がよく用いられるようである。それによると、寄与率が有意な増加を示す範囲で説明変数が選択されて行き、候補に上げたものから選んだ最善の最小個数の組合せになっている。しかし、そのまま採用する必要はなく、専門分野の経験的知識を駆使してよく検討を加え、実用的見地から最終的な決定を下すべきである。逐次選定法で多数の説明変数の候補を検討して重回帰式で要因分析を行った歯科関係の研究報告を近年の「歯科学報」から探してみると、宮崎²⁸⁾により矯正治療後の咬合の安定性について要因分析をしたものがあり、変数増減法により変数選択を行い説明変数の個数は高々 $5 \sim 6$ 個にしばられている。また、土屋²²⁾は咀嚼能力に影響を及ぼす要因を形態分析と機能分析の面から調べており、変数増加法で変数選択をした結果、形態分析では6個、機能分析では2個の説明変数が選ばれている。

4.2 判別分析について

個体が g 個の群のいずれか1つの群に属することはわかっているものとして、個体について得られた p 種類の測定値をもとに、この個体が g 個の群のどれに属するかを判定する方式を示す判別関数を、各群から得られた g 組の p 変量データに基づいて作ることが、多変量解析法での判別分析の問題となる。一般に、 g 個の群への判別では $(g - 1)$ 個の判別関数が必要になる。

ここでは、2群の判別に限定して述べることにする。例えば、患者についての p 種類の検査結果をもとに、その患者が2つの病名のいずれであるかを診断するような場合である。この問題に対しての判別関数は次のようにして作られる。2群 G_1, G_2 (2つの病名の患者群)のいずれに属するかが既に知られている患者を群 G_1 と群 G_2 のそれぞれから n_1 人と n_2 人を標本として抽出し、各人について求めた p 種類の検査値 x_1, x_2, \dots, x_p をデータとして、判別関数 $z = a_0 + a_1x_1 + \dots + a_px_p$

を作る。その場合、この z に関しては 2 群 G_1 と G_2 が最もよく分離され、 z の値の正負で 2 群 G_1 と G_2 のいずれに属するかの判定ができるようにする。そのために、「 z では、2 群の群内変動に対する群間変動の比が最大になる」ように判別関数の係数を定めることが行われる。このようにして判別関数が得られておれば、 G_1 であるか G_2 である病名不明の患者が来たとき、その患者の p 種類の検査値を判別関数に代入して求めた判別値 z が正なら群 G_1 へ、負なら群 G_2 へと判別できることになる。

次に、判別分析を適用する上での注意点をいくつか述べておくことにする。

(1) 判別分析でのデータの尺度について

判別分析での目的変数 y は分類尺度、説明変数 x_1, x_2, \dots, x_p は間隔尺度として扱われる。なお、説明変数が全て分類尺度(質的データ)であるときは、数量化理論第 1 類がその場合の判別分析法になっている(cf: 駒沢¹⁰⁾)。

(2) 判別分析とデータの大きさについて

2 群の判別分析では、それぞれの群からのデータの大きさを n_1, n_2 とし、説明変数の個数を p とするとき、 $n_1 + n_2 - p - 1 \geq 30$ となる程度に n_1 と n_2 の大きさをとる必要があるといわれる。なお、できるだけ $n_1 = n_2$ になるように取るのが望ましい。

(3) 判別関数の妥当性の検討について

判別関数 z は、2 群の各分布の分散・共分散行列は等しいと仮定して導かれている。しかし、多少異なっても影響は少ない(頑健性をもつ)ので、分散で 2 ~ 3 倍くらいの違いは無視でよいと言われている(cf: 奥野ほか⁷⁾)。

得られた判別関数を用いて、データを G_1 と G_2 の各群へ判別してみて、各群へ正しく判別されたものの割合(正判別率)を調べてみることである。正判別率がかなり高くないと判別関数は使えない。

(4) 判別分析における説明変数の選択について

実際の判別分析では、判別に関係すると予想さ

れる要因を多数説明変数の候補にあげておき、重回帰分析のときと同様に逐次選定法により変数の選択を行い、適切なものを最終的に 3 ~ 5 個くらいに絞るのが望ましい。なお、最終的に選択された説明変数については、説明変数間に相関の大きなものは含まれていないことを確認すること。

(5) 判別分析における各説明変数の影響力の評価

判別分析を要因分析として利用する場合には、各説明変数が判別に関する要因としてもつ影響力を比較したくなる。そのためには、標準化した説明変数に対する判別関数を作り、その係数(標準化判別係数)の絶対値の大小で各説明変数の判別への影響力を比較すればよい。ただし、前の重回帰分析のところでも述べたように、この方法は説明変数間の相関係数が小さいときに近似的に用いることができる。なお、別の方法として偏相関係数を用いることもできる(cf: 駒沢¹⁰⁾)。

判別分析を適用した歯科関係の研究報告を近年の「歯科学報」から探してみると、大野⁴⁾、稲葉²⁾、栃倉²³⁾、久永²⁶⁾などが出ている。ここでは、久永²⁶⁾における判別分析の結果を例として紹介しておく。

[例] 切歯切端部の透明感に寄与している因子の特定

透明感の強い歯群を G_1 、透明感の強くない歯群を G_2 とする。歯科医師 3 名が判定者となり、上下顎の中切歯と側切歯からなる 331 本の各切歯の透明感を視覚的に判定した結果、3 名が全員一致で G_1 か G_2 に分類できたものだけをデータとした。その結果、データの大きさは G_1 において $n_1 = 86$ 、 G_2 では $n_2 = 228$ となった。この被験歯の色および透明性に関する物理的測定値を説明変数として判別関数を作ることを試みた。なお、物理的測定に当たっては、色を 3 次元直交座標上の数値で表すことができる CIE 1976(L* a* b*) 色空間を用いた。この表色系によって明度 L* 値および色度に関する色相 a* 値と彩度 b* 値がデータとして測定される。切歯の切端部のそれらを記号で I - L*, I - a*, I - b* で示し、また、切歯の中央部のそれらを M - L*, M - a*, M - b* で示す。

さらに、切歯切端部の透明光の明度 L^* 値を $1 - T - L^*$ で示すことにする。これら7つの測定項目を説明変数の候補して変数増減法で判別関数の変数選択を行った結果、表4-3に示されている4つ変数が選ばれた。判別関数 $z = a_0 + a_1 x_1 + \dots + a_4 x_4$ の係数は表4-3の第2列目に示されている値になる。データ (x_1, x_2, x_3, x_4) を判別関数に代入して計算した判別値が $z > 0$ であれば群 G_1 へ、 $z < 0$ であれば群 G_2 へ判別される。正判別率をみるために、もとのデータをこの判別関数で判別してみた結果、正しく判別されたものの割合は群 G_1 で87.2%、群 G_2 では81.1%であった。次に、各説明変数が判別に寄与している度合いを知るために、標準化判別係数を求めると表4-3の第3列目のようになる。なお、説明変数間の相関係数は表4-4が示すように、 x_1 と x_4 、および、 x_2 と x_4 の相関係数が0.3を多少越えているが全体的に説明変数間の相関は低いので、各説明変数の寄与する程度は近似的に標準化判別係数の絶対値の大小で評価してよいであろう。その臨床的な解釈は、久永²⁶⁾の論文を参照されたい。

表4-3 切歯切端部の透明感に関する判別分析

説明変数 x_i	判別係数 a_i	標準化判別係数
x_1 : $1 - T - L^*$ 値	0.3329	0.9028
x_2 : $1 - b^*$ 値	-0.2347	-0.5615
x_3 : $1 - L^*$ 値	-0.0959	-0.2386
x_4 : $1 - a^*$ 値	0.5431	0.2333
定数 a_0	-8.05	0

判別関数 $Z = a_0 + a_1 x_1 + \dots + a_4 x_4$
 正判別率 群 G_1 : 87.2% , 群 G_2 : 81.1%

(久永²⁶⁾1997による)

表4-4 相関行列

	x_1	x_2	x_3	x_4
x_1	1	0.00	-0.13	-0.35
x_2		1	-0.23	0.39
x_3			1	-0.10
x_4				1

5. 主観的評価の尺度化について

人による感覚的な判定を尺度化する問題は、官能検査の分野でいろいろ研究されており種々の方法が知られている。ここでは、その中から Thurstone^{11),12)}の一对比較法による尺度化の方法を紹介しておく。これは比較的簡単に使えて、その尺度値の妥当性を調べる検定も Mosteller³⁰⁾によって開発されている。

5.1 一对比較実験

いま、 t 個の試料 T_1, T_2, \dots, T_t があるとき、2つずつの対 (T_i, T_j) ($i = 1, 2, \dots, t; i < j$) をつくり、各対ごとに T_i, T_j のいずれが好ましいか(または優劣)を n 人の判定者に尋ねる。その判定結果より T_i の方が T_j より好ましいと判定した人数(度数) f_{ij} の表を作製し、それをデータとして T_1, T_2, \dots, T_t 間の差を相対的な評点(間隔尺度)で格付けすることを考える。多数のものを同時に判定する場合に比べて、一对比較ではその判定の信頼性が高くデータの質がよくなる利点がある。なお、判定で「どちらとも言えない」という回答も許してよい。例として、本学口腔外科第二講座内山健志教授による一对比較実験の一部を取り上げることにする。

[例] 未手術口唇裂児の顔面模型についての審美度の一对比較実験

5個の顔面模型の全ての対(10組)を20名の判定者に順に示し、各対でいずれが審美度の上で醜いかを判定してもらった結果が表5-1である。 T_i より T_j の方が醜いと判定した人数が f_{ij} ($i < j$) であり、 $f_{ji} = n - f_{ij}$ の関係がある。また、判定で「どちらとも言えない」と回答することも許し、その場合は度数を0.5としてある。表5-2は表5-1の結果を比率 p_{ij} で表したものである。

5.2 Thurstone の一对比較法による尺度化

一对比較法による Thurstone^{11),12)}の尺度化では、試料 T_i が人に与える心理的影響の強さ X_i は人によって異なるが、人間集団として見れば正規分布に従うと仮定して尺度化を行うものである。試料 T_i と T_j を比較したときの心理的影響の差 $X_i - X_j$ の分布も正規分布になることを利用し

て、表5-2の比率 p_{ij} を正規尺度化(Z変換)する。ここでは、 p_{ij} は T_i より T_j の方が醜いと判定した者の比率であり、醜さの度合いを尺度化するので、 $X_i - X_j$ を標準化したZが標準正規分布 $N(0, 1)$ に従うことを用いて、 $Z < z_{ij}$ となる確率が p_{ij} であるような z_{ij} の値、すなわち、 $P(Z < z_{ij}) = p_{ij}$ である z_{ij} の値(図5-1を参照)を正規分布に関する数表から求める。その結果は表5-3に示されている。この表5-3の各列の値を合計し、試料の個数 t で割って平均値 z_j を求める。それが尺度化した値になるが、尺度値は間隔尺度であり原点0は任意であるから、便宜的に z_j の最小値を0になるように直したものを尺度値とする。その結果は、表5-3の最後の行に示してあるように、 $T_1 = 0, T_2 = 0.405, T_3 = 0.466, T_4 = 1.303, T_5 = 0.621$ となり、それは T_1 を規準として T_2, T_3, T_4, T_5 の各模型の醜さの尺度値

を示したものになる。その尺度値を数直線上にプロットしたものが図5-2である。

Thurstoneの一对比較法による尺度化の方法で問題になるのは、観測比率 p_{ij} が0または1になることが起こった場合である。この場合には、Z変換したものは-と+とになり、これらをZ変換の表5-3に入れたのでは尺度値を算出することは不可能になる。このようなことが起こったときの尺度値の求め方については、田中¹⁷⁾に説明が出ているので参照されたい。

表5-1 一对比較実験による度数データ f_{ij}

	T_1	T_2	T_3	T_4	T_5
T_1		14	14.5	17.5	14
T_2	6		9	17	13
T_3	5.5	11		16.5	10
T_4	2.5	3	3.5		5.5
T_5	6	7	10	14.5	

[注] データは口腔外科第二講座内山健志教授による。判定者人数は $n = 20$ である。

対角線の上方の数値は、 T_i より $T_j (i < j)$ の方が醜いと判定した者の人数 f_{ij} を示す。ただし、「どちらとも言えない」と回答した者は度数0.5としてある。対角線の下方の数値は $f_{ji} = n - f_{ij}$ (T_i より T_j の方が醜くない)を示す。

表5-2 一对比較実験：判定者比率 p_{ij}

	T_1	T_2	T_3	T_4	T_5
T_1		0.7	0.725	0.875	0.7
T_2	0.3		0.45	0.85	0.65
T_3	0.275	0.55		0.825	0.5
T_4	0.125	0.15	0.175		0.275
T_5	0.3	0.35	0.5	0.725	

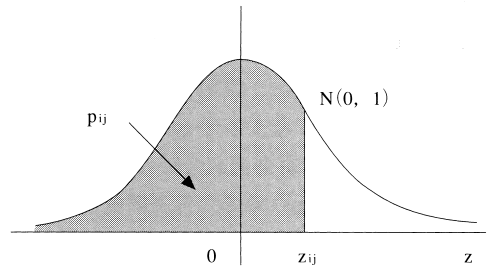


図5-1 $P(Z < z_{ij}) = p_{ij}$

表5-3 一对比較法：正規尺度化(Z変換)による z_{ij} の値

	T_1	T_2	T_3	T_4	T_5
T_1	0	0.524	0.598	1.150	0.524
T_2	-0.524	0	-0.126	1.036	0.385
T_3	-0.598	0.126	0	0.935	0
T_4	-1.150	-1.036	-0.935	0	-0.598
T_5	-0.524	-0.385	0	0.598	0
計	-2.796	-0.771	-0.463	3.719	0.311
平均値	-0.559	-0.154	-0.093	0.744	0.062
尺度値	0	0.405	0.466	1.303	0.621

[注] 尺度値は間隔尺度であり原点0は任意であるが、ここでは、 $T_1 = 0$ としてある。

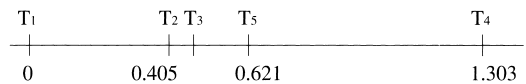


図5-2 一对比較法による尺度値の図示 ($T_1 = 0$ とした場合)

5.3 判定者相互間の一致性の検定

尺度化が可能となるためには、各判定者の判定基準に共通性があることを要する。それは判定者相互間の一致性を検定することでチェックされる。また、判定する試料の差異が小さいときには尺度化が困難となるが、その場合には判定者相互の一致性は悪くなるであろう。表5-1の度数データ f_{ij} を用いて、一対比較における判定者間の判断の一致性を検定する方法が Kendall⁸⁾ によって与えられている。それは、

$$S = n(n-1) \sum_{i < j} (t_i - t_j) + 4 \sum_{i < j} f_{ij}^2 - n \sum_{i < j} f_{ij}$$

とおくとき、次の検定統計量 χ^2 を用いて検定される。

$$\chi^2 = \frac{4}{n-2} \left[S - \frac{n(n-1) \sum_{i < j} (t_i - t_j)}{8(n-2)} \right]$$

この χ^2 は自由度

$$v = n(n-1) \sum_{i < j} [2(n-2)]$$

の χ^2 分布に従う。表5-1のデータの場合は、 $S = 1130$ 、 $\chi^2 = 51.729$ であり、自由度 $v = 12$ の χ^2 分布の上側1%点は26.22であるから有意水準1%で判定者相互間に一致性が見られることになる。

5.4 尺度値の妥当性の検定

一対比較法で得られた表5-2の観測比率 p_{ij} をもとに Thurstone の方法で尺度化を行ったのであるが、その尺度値が妥当なものであるかどうかを Mosteller³⁰⁾ の方法で検定することができる。それは、Thurstone の方法で得られた尺度値から逆に期待比率 p'_{ij} を算出し、もとの観測比率 p_{ij} との適合度を検定するものである。具体的には、観測比率 p_{ij} と期待比率 p'_{ij} の各々の平方根を逆正弦変換したものをそれぞれ θ_{ij} 、 θ'_{ij} とすると、この θ_{ij} と θ'_{ij} はそれぞれ近似的に正規分布することに着目して検定統計量を導く。

逆正弦変換： $\theta_{ij} = \sin^{-1} \sqrt{p_{ij}}$ 、 $\theta'_{ij} = \sin^{-1} \sqrt{p'_{ij}}$

検定統計量： $\chi^2 = \sum_{i < j} n(\theta_{ij} - \theta'_{ij})^2 / 821$

この検定統計量が自由度 $v = (t-1) \sum_{i < j} 2$ の χ^2 分布に従うことを用いて検定する。

この例では、表5-3に示されている尺度値から算出した期待比率 p'_{ij} は表5-4のようになる。表5-2と表5-4の比率を逆正弦変換した

表5-4 尺度値から求めた期待比率 p'_{ij}

	T ₁	T ₂	T ₃	T ₄	T ₅
T ₁		0.659	0.681	0.903	0.732
T ₂	0.341		0.524	0.813	0.583
T ₃	0.319	0.476		0.797	0.560
T ₄	0.097	0.187	0.203		0.248
T ₅	0.268	0.417	0.440	0.752	

θ_{ij} と θ'_{ij} を求めて検定統計量を計算すると $\chi^2 = 2.076$ となる。いま、自由度は $v = 6$ であり、 χ^2 分布の上側5%点は12.59である。従って、この適合度検定の結果、観測比率 p_{ij} と期待比率 p'_{ij} の適合度は大変よく、尺度値は妥当なものと判定される。

6. おわりに

歯科データの統計的分析といってもその適用範囲は広く、本稿ではごく限られた内容について述べたに過ぎない。統計相談を受けて日頃気になっていたことを中心に述べたので、その大部分は統計学の常識程度のことばかりかも知れない。また、歯科の研究に役立つと思われる統計的手法について気をつくままに述べてみたが、何か少しでも参考になることがあれば幸いである。

最後に統計的方法の有用性に関して一言すれば、統計的分析法がいつも役立つとはかぎらないし、役立つ程度も取り扱う問題によって異なるのが普通である。実際のデータ解析に当たっては、統計的分析をして得られた結果が意味のあるものかどうかを専門分野の経験をもとに判断することが大事である。統計的分析の結果を機械的に結論とすることは戒めなければならない。統計的分析法に過大な期待をもつことは禁物である。それを使うことによって研究を進める上で何かヒントになることが得られたならば、それだけで役に立ったと考えるべきではあるまいか。ところで、現実の問題に直面すると、間違った使い方をしたり経験不足のために有効な結果が出せない場合が多いのであるから、一寸使ってみて役に立たなかった

からといって、その有用性を一概に否定することは、これまた禁物であろう。パソコン用の統計ソフトが普及してきている現在、統計的分析法を上手に使うことによって歯科の研究に役立てることは1つの方法として重要であると思われる。

本稿は第270回東京歯科大学学会総会(平成12年11月5日、千葉)において特別講演したものである。

文 献

- 1) 安東基善：多核巨細胞の細胞性格に関する実験的研究 - 酵素組織化学的観察とその画像解析の応用 - , 歯科学報, 94 : 305 ~ 329, 1994 .
- 2) 稲葉敬子：顎顔面形態と咬合力に関する研究, 歯科学報, 94 : 153 ~ 176, 1994 .
- 3) 岩原信九郎：ノンパラメトリック法 第6版, 日本文化科学社, 東京, 1973 .
- 4) 大野二朗：モアレトポグラフィ法によるフィリピン人の上顎歯列弓と口蓋の形態学的研究, 歯科学報, 94 : 23 ~ 45, 1994 .
- 5) 奥野忠一, 芳賀敏郎：実験計画法, 培風館, 東京, 1969 .
- 6) 奥野忠一, 久米均, 芳賀敏郎, 吉沢 正：多変量解析法, 日科技連, 東京, 1971 .
- 7) 奥野忠一, 片山善三郎, 上郡長昭, 伊東哲二, 入倉則夫, 藤原信夫：工業における多変量データの解析, 日科技連, 東京, 1982 .
- 8) Kendall, M. G. : Rank correlation methods 4 / e, Charles Griffin, 1975 .
- 9) Cohen, J. : Statistical power analysis for the behavioral sciences, Academic Press, 1977 .
- 10) 駒沢 勉：数量化理論とデータ処理, 朝倉書店, 東京, 1982 .
- 11) Thurstone, L. L. : Psychophysical analysis, Amer. Jour. Psychol., 38 : 368 ~ 389, 1927 .
- 12) Thurstone, L. L. : A law of comparative judgement, Psychol. Rev., 34 : 273 ~ 286, 1927 .
- 13) 佐和隆光：回帰分析, 朝倉書房, 東京, 1979 .
- 14) Siegel, S. and Castellan, N. J. : Nonparametric statistics for the behavioral science 2 / e, McGraw - Hill, 1988 .
- 15) 杉山高一：多変量データ解析入門, 朝倉書店, 東京, 1985 .
- 16) 武田康雄：各種食品咀嚼時の咬筋筋活動に関する研究, 歯科学報, 96 : 695 ~ 721, 1996 .
- 17) 田中良久：心理学的測定法, 東京大学出版会, 1961 .
- 18) 丹後敏郎：医学への統計学, 朝倉書店, 東京, 1983 .
- 19) 丹後敏郎：臨床検査への統計学, 朝倉書店, 東京, 1986 .
- 20) S. チャタジー, B. プライス：回帰分析の実際, 新曜社, 東京, 1981 .
- 21) 辻 新六, 有馬昌宏：アンケート調査の方法, 朝倉書店, 東京, 1987 .
- 22) 土屋朋未：正常咬合者と反対咬合者の咀嚼能力に関する研究 - 形態的・機能的要因について - , 歯科学報, 95 : 729 ~ 755, 1995 .
- 23) 栃倉真由美：矯正治療の前後における側貌の変化に関する研究, 歯科学報, 95 : 53 ~ 93, 1995 .
- 24) 中村正一：例解回帰分析入門, 日刊工業新聞社, 東京, 1982 .
- 25) 林知己夫, 多賀保志：調査とサンプリング, 同文書院, 東京, 1985 .
- 26) 久永竜一：切歯切端部の透明感に関する色彩学的研究, 歯科学報, 97 : 243 ~ 258, 1997 .
- 27) J. L. フライス：計数データの統計学, 東京大学出版会, 東京, 1975 .
- 28) 宮崎晴代：成人抜歯症例における矯正治療後の咬合の安定性について, 歯科学報, 94 : 749 ~ 770, 1994 .
- 29) 宮原英夫, 丹後俊郎編：医学統計学ハンドブック, 朝倉書店, 東京, 1995 .
- 30) Mosteller, F. : A test of significance for paired comparisons when equal standard deviations and equal correlations are assumed, Psychometrika, 16 : 207 ~ 218, 1951 .
- 31) 吉成正雄, 大山貴司, 水田 茂：フルクラウン鑄造体による急速加熱型埋没材の特性評価, 歯科学報, 95 : 1175 ~ 1183, 1995 .
- 32) 吉村 功編著：毒性・薬効データの統計解析, サイエンス社, 東京, 1987 .
- 33) E.L. レーマン：ノンパラメトリックス - 順位にもとづく統計的方法 - , 森北出版, 東京, 1978 .