

質問と回答

高木 廣文

平成13年 8 月
超音波検査技術 Vol. 26 No. 5

統計講座 (第12回)

質問と回答

高木 廣文

Q 頸動脈超音波検査にてIMC厚を測定した。同時に年齢・性別のほか、各種検査データを調べた(表1)。

以上の結果より、IMC肥厚をもたらす動脈硬化の危険因子について、統計学的に調べるには、どのような統計処理が適当か。

表1 頸動脈超音波検査(IMC厚)と因子との関係

ID	IMC厚	年齢	性別	BMI	Hb A1c 5.8%以下	HDL-C 40~99mg/dl	T-CHO 120~239mg/dl	TG 31~219mg/dl	血圧 90~140mg/dl
1	0.8	45	M	20	4.1	50	182	145	76~132
2	1.4	56	M	24	7.2	42	180	188	88~138
3	0.6	39	M	29	5.2	48	221	450	65~123
4	1.1	70	M	21	5.2	38	125	168	106~189
5	1.1	65	M	24	4.8	60	170	120	67~118
6	0.9	55	F	21	4.2	45	130	264	98~156
7	0.6	41	F	19	5.3	80	146	96	67~124
8	1.1	53	M	17	7.6	52	196	169	78~134
9	0.8	38	M	26	5.3	63	185	300	83~137
10	0.5	30	F	17	4.5	75	210	76	77~120
11	1.3	71	M	29	5.3	32	350	551	98~178
12	1.2	58	F	22	6.9	43	201	201	96~189
13	0.6	47	F	21	5.2	60	152	480	77~127
14	1.1	47	M	19	4.3	51	240	200	65~98
15	1.1	69	F	24	5.2	56	156	208	79~134
16	1.2	36	M	28	7.8	41	304	356	88~146
17	0.7	34	F	27	4.3	63	120	185	78~134
18	1.1	69	M	19	5.4	48	282	296	76~138
19	1.3	76	F	22	7.3	56	196	176	88~139
20	0.7	48	M	24	5.2	48	160	296	68~128
21	1.1	59	F	23	6.1	60	120	165	85~123
22	0.6	46	M	20	4.5	50	156	180	63~136
23	0.7	57	F	27	5.8	53	268	489	79~166
24	1.1	68	F	16	6.9	44	204	354	89~167
25	1.1	35	M	21	5.4	54	168	198	118~196
26	1.3	70	M	20	4.5	48	178	145	98~189
27	0.7	58	M	23	5.3	68	164	176	63~120
28	1.1	74	F	27	5.5	32	285	477	85~169
29	1.2	59	F	29	5.6	31	184	322	84~140
30	0.8	53	M	25	5.9	40	112	156	88~137

Hirofumi TAKAGI: 新潟大学医学部保健学科
〒951-8518 新潟県新潟市旭町通2番町746番地
e-mail: takagi@clg.niigata-u.ac.jp
http://www.clg.niigata-u.ac.jp/~takagi/

A 基準変数であるIMC厚は数値であり、その他の変数も性別以外はすべて数値データで測定されている。そこで、最も簡単な解析手法は、IMC厚を基準変数、年齢、性別、BMIなどを説明変数とした重回帰分析を行うことである。なお、分析を行う場合、年齢はデータとして「男」、「女」として収集されているので、男=1、女=0のように数値化する必要がある。この場合、性別に関して重回帰分析で得られた偏回帰係数は、女性に対する男性のIMC厚の影響がどれだけ大きいかを示すことになる。

別の方法としては、IMC厚を「0. 正常」、「1. 異常」のように分類して、判別分析を行うか、疫学研究などでリスク評価のためによく用いられている多重ロジスティックモデルを用いるかして、解析すればよいだろう。これらの方法については、文献1か私のホームページを参

考にしてほしい。

また、年齢、BMI、HbA1cなどの説明変数については、測定データをそのまま用いて分析することが多いようだが、研究の目的によっては、適当なカテゴリに区分して（例えば、正常、境界型、異常など）解析に用いることも可能である。

ここで示した分析方法やデータの扱い方などは、統計学上の適切さというよりも、その解析の研究目的や結果の使用の仕方に依存しているといえよう。

これを読んで、ますますどの解析を行えばよいか分からなくなったかもしれないが、自分の研究にあった方法を選ぶのも研究の内といえる。

Q

超音波集団検診では、短時間の検査時間で正確な結果を出すことが求められる。

受診者A～Zの協力により、超音波検査にお

表2 検者別の検査所要時間と不一致件数

受診者	検者A		検者B		検者C		検者D		検者E	
	検査時間	不一致	検査時間	不一致	検査時間	不一致	検査時間	不一致	検査時間	不一致
A	6		10		6		5		8	
B	8		12		6		6		9	
C	6		9		6		6		6	
D	5		7		6	1	5	1	7	1
E	7		11		8		7	1	9	
F	10	2	14	1	12		8	1	15	
G	5		7		6		5		7	
H	6		8	2	6		5	2	8	
I	7		9		6		5		9	
J	7	1	9		6		5		9	
K	6		9		6		5		8	
L	8		12		6		5	1	10	
M	10		14		12		8		12	
N	6		8	1	5		6		8	
O	8		9	1	6		6	1	10	
P	7	1	9		6		6		9	
Q	7		9		6		6		10	
R	6		8		6		5	1	9	
S	7		10	2	6		7		8	
T	12		16		12		10		15	
U	6		8		6		5	1	8	1
V	8		10	1	6		5		9	
W	7		8		6		6		10	
X	5	1	8		6	1	5	1	8	
Y	6		7		6		5	1	9	
Z	7		8		6		5		10	

ける検者別の検査時間を調べた。同時に正しい検査結果との相違点を、不一致と表現した(表2)。検者の画像抽出能力や受診者に対する接遇などに差がないとしたとき、検者A～Eについて、統計学的に検査技術を評価することは可能か。

A 表をみると、「不一致」の欄に書かれた1, 2が件数を示すようだが、全体で何件の検査をしているのかよく分からないし、「正しい検査結果」なるものが、どのような基準で「正しい」のかも統計学以前の問題として、私には気になって仕方がない点である。

それはそれとして、単純に検者A～Eの5人の検査技術を比較したいのであれば、検者を要因として、一元配置分散分析で「検査時間」および「不一致」の母平均値に差がないか調べればよい。この場合、表で「不一致」に何も記載されていないデータは0として分析することになる。

または、検査時間や不一致のデータをそのまま使用するのではなく、データの順序性の情報だけを使って分析するのであれば、Kruskal-Wallis検定を用いてもよい。どちらにせよ、検定結果が有意であれば、どの検者間で差があるのかを個々に多重比較する必要がある。これらの点に関しては、文献2, 3が私のホームページに説明してあるので、参照すればよいだろう。

Q 腹部超音波検査にて、脂肪肝の有無を調べた。同時に身長、体重よりBMIを算出、体脂肪計で体脂肪を測定した(表3)。

以上の結果より、脂肪肝と肥満との関係について、統計学的に調べるにはどのような統計処理が必要か。

A 肥満と脂肪肝の関係を調べるということなので、いくつかの分析方法が考えられよう。

1) クロス集計をもとに解析

BMIの数値により、肥満と非肥満に分け、エコーによる結果(脂肪肝、軽度脂肪肝、正常)とクロス集計をし、カイ2乗検定を行う。

体脂肪についても同様に解析ができる。

表3 BMI, 体脂肪

ID	性別	BMI	体脂肪	エコー
1	M	30.1	26.4	脂肪肝
2	F	19	21.2	—
3	F	17	29.7	—
4	F	17.5	19.5	—
5	F	29.3	31.2	脂肪肝
6	M	27.5	26.9	軽度脂肪肝
7	M	22.2	22.2	—
8	M	23.9	21.2	軽度脂肪肝
9	M	21.7	15.6	—
10	M	25.4	18.7	軽度脂肪肝
11	M	19.9	19.6	—
12	M	24.6	29.7	—
13	M	24.7	20.5	脂肪肝
14	M	22.3	18.2	—
15	M	30.1	27.9	脂肪肝
16	M	18.1	10.9	—
⋮	⋮	⋮	⋮	⋮
41	M	24.7	23.8	—
42	M	22.3	20.4	軽度脂肪肝
43	M	22.4	18.2	軽度脂肪肝
44	M	22.4	21.6	軽度脂肪肝
45	M	31.2	33.4	脂肪肝
46	M	24.6	17.7	軽度脂肪肝
47	F	18.3	25.4	—
48	M	23	19.3	—
49	M	24.1	32.3	—
50	M	25.2	20.3	—
51	M	24.7	24.6	—
52	M	22.2	18.2	—
53	M	24.2	25.2	—
54	F	23.7	32.4	軽度脂肪肝
55	M	24.5	19.8	—
56	M	20	21.6	—
57	M	21.3	20.8	—
58	M	22.2	26.1	軽度脂肪肝
59	M	24.9	23.8	軽度脂肪肝
60	M	22.9	18	—
61	M	19.6	13.7	—
62	M	23.2	19.3	—
63	M	27.8	25.2	脂肪肝
64	M	24	22.8	軽度脂肪肝
肥満		26.4以上	M25.0以上 F30.0以上	

なお、この場合、BMI, 体脂肪, エコーの各結果を2分類にするか、3分類にするかなど、また分類のための基準となる数値は、いくつかの場合について、自分で実際にやってみる必要があるだろう。

2) 一元配置分散分析(またはKruskal-Wallis検定)を行う

まず、脂肪肝、軽度脂肪肝、正常の順にBMIと体脂肪の平均値が小さくなっているか確認する。そうでない場合には、脂肪肝と肥満の関係は説明が付かないだろう。

次に、回答2で説明したように、エコーの結果ごとに、BMI、体脂肪の母平均値に差があるかを、一元配置分散分析で解析する。

場合によっては、ノンパラメトリック検定(Kruskal-Wallis検定)を用いてもよい。

3) 判別分析か多重ロジスティックモデルで解析する

性別、BMI、体脂肪の3変数の脂肪肝への影響を同時に解析したい場合、最初の質問への回答で述べた方法で分析すればよい。

エコーの結果は、3区分されているが分析のためには、正常と異常のように、あらかじめ2分類にしておく。また、この分析でもBMIや体脂肪を量的データとしてではなく、正常・異常のように質的なデータとしても解析できるのは言うまでもない。

ただし、BMIと体脂肪の相関はかなり高いことが予想されるので、多重共線性の問題が起きるかもしれない。この点については実際に解析してみないと分からないが、おかしな結果が得られた場合には、この点について注意しながら、結果を解釈する必要があるだろう。

Q 我々は、各病棟・診療科からのMRSA（メチシリン耐性黄色ブドウ球菌）の分離頻度を院内感染対策として、院内の責任者へ配布しているのだが（表4）、その際

過去の分離割合と比較したいという意見があり、過去1年間の分割割合の平均をつけたところ、さらに有意な範囲を示してほしいという意見があがってきた。つまり、平均の標準偏差をつけるということであった。データ的には正規分布をとるのが少なく（0%を示すデータが多く偏っている）本当に単なる平均でよいものか迷っている。

診療科ごとのデータ値は1998年1月から2000年8月までの32か月分の各月の分離率（MRSA検出者数／検査数×100）を示す。このデータから求められる過去の実績と当該月の分離割合を比較し有意に上昇あるいは低下がみられるかを知りたいのが目的である。要は生化学などにおける精度管理のようなもので、ある診療科での過去の毎月の分離数（割合）が分かっている。そこで、どうもある科は最近多くなってきたような気がするが本当に有意に過去のデータと比較して増えているのか。増えているのなら何らかの対策を促さなければならぬ。また、減少していれば何らかの対策が有効に働いたのか、ということである。

表4 MRSAの分類頻度

診療年月	総 合	呼吸器内科	循環器内科	胸部外科	消化器内科	消化器外科
98/jan	0	0	0	6.3	12.5	4.4
98/feb	0	0	0	21.7	0	2.2
98/mar	12.5	5.9	40	7.7	0	4.9
98/apr	0	23.1	33.3	8.3	0	12.5
98/may	28.6	0	33.3	14.8	0	9.3
98/jun	18.2	10.5	22.2	8.6	0	4.5
98/jul	7.7	4.5	11.1	16.1	0	11.6
98/aug	0	13.6	16.7	11.1	16.7	10.3
98/sep	14.3	11.8	20	9.7	14.3	0
98/oct	33.3	0	0	12	0	2
98/nov	28.6	0	0	15	14.3	12.5
98/dec	10	5	0	11.4	25	7.7
⋮	⋮	⋮	⋮	⋮	⋮	⋮
99/oct	14.3	0	22.2	4.5	22.2	8.9
99/nov	26.7	12.5	22.2	15.2	8.3	16.3
99/dec	15.4	6.7	0	3.7	0	19.5
00/jan	0	7.4	0	9.1	0	8.3
00/feb	30.8	4.3	0	3.8	0	5.6
00/mar	15.4	0	27.3	4	11.1	8.6
00/apr	20	0	20	9.4	11.1	6.3
00/may	12.5	12	25	12	0	10.3
00/jun	0	15	28.6	9.1	0	11.8
00/jul	0	0	25	11.1	0	11.4
00/aug	0	3.8	0	12	0	8.8

(%)

正規分布をとっていない場合の平均（のような代表値）に代わるものは何であるか。あるいは平均でよいのか。また、平均でない値で示す場合の信頼区間の求め方などはどのようにすればよいのか。そして、表現の問題としてグラフ化したデータに信頼区間として表したい。その区間を求めるにはどのような統計手法を用いればよいのか。

A データが分離度という形に加工されているが、もともとは細菌の「あり」、「なし」という2値データだろうと思うが、そうだとすると2項分布に従うデータということになる。2項分布の平均と分散の式は決まっているので、それを使えばよいということになる。このためには検体数と「あり」の個数が分からないといけない。

検体数がNのとき、細菌がm検体で検出された場合、分離度pは、

$$p = m/N$$

となる。ここでは、100倍せずに記述する。このpの標準誤差SE(p)は、

$$SE(p) = \sqrt{[p(1-p)/N]}$$

となる。分離度95%信頼区間は正規分布近似を用いると、

$$p \pm 1.96 \times SE(p)$$

で下限と上限が求まる。%で表示したければ100倍する。基本的に信頼区間が他の部署と離れるところは、有意差があると判断できる。

また、話を簡単にして過去12か月の平均的な割合Pと、当該月での割合pを比べるものとする。12か月での総検体数をN、有所見数をMとし、当該月ではそれぞれn、mとする。

$$P = M/N, \quad p = m/n$$

はすぐに求まる。問題は、Pに比べpが有意に大きいのか小さいかということになる。これは、通常の「独立な二つの母比率の差の検定」で、正規分布近似により分析できる。もしくは、まったく同一の分析として、クロス差の形にしてカイ2乗検定すればよいことになる。

	有所見	無所見	計
過去 当該月	M m	N-M n-m	N n
計	X	T-X	T

ただし、 $T = N + n$, $X = M + m$
として、

$$\text{カイ2乗値} = \frac{T \cdot (|M(n-m) - m(N-M)| - T/2)^2}{N \cdot n \cdot X \cdot (T-X)}$$

が、自由度1のカイ2乗分布に従うことで検定できる。上記のカイ2乗値が3.84より大きければ、5%水準で有意差ありとする。

この辺の検定のやり方は、普通のクロス表の検定と同じなので、詳しくは適当な教科書を見ていただきたい。

Q 例えば、画像処理法が二つあって、どちらの方法が優れているかを評価しようとする場合、同じ対象を複数の観察者が観察したときに観察の評価が、1, 2, 3の三段階であったとして、観察者間あるいは1人の観察者の評価の再現性を評価するにはどのようにすればよいのか？

A 基本的に1人の観測者の再現性ということとは議論にならない。

何人かが同一のものを評価した結果が、再現性があるかは昔から方法がある。上記のように三段階あるとすると、

		2回目			
		1	2	3	計
1回目	1	n_{11}	n_{12}	n_{13}	N_1
	2	n_{21}	n_{22}	n_{23}	N_2
	3	n_{31}	n_{32}	n_{33}	N_3
計		M_1	M_2	M_3	T

のようなデータが得られるはずである。このような対応のあるデータからは、再現性の指標として、カッパ係数 κ が用いられている。

$$\text{一致度 } P_0 = (n_{11} + n_{22} + n_{33}) / T$$

試行が独立な場合、偶然に一致する期待値 P_c

$$P_c = (N_1M_1 + N_2M_2 + N_3M_3) / T^2$$

$$\kappa \text{ 係数} = (P_0 - P_c) / (1 - P_c)$$

によって求められる。上記の式は、三段階より段階が多い場合にも拡張できる。

なお、段階が3ではなく、もっと多くなってきた場合に単純に相関係数を求めれば、それが再現性の指標（信頼性係数）の推定値になることが知られている。

κ 係数が大きい方、または相関係数が大きい方が優れていると考えるのが普通である。

Q この度、頸動脈エコーを用いた動脈硬化症とリスクファクターの関連性についてまとめたいと思っているが、有意差などの統計処理が分からずにいる。

対象は、N群（正常）、A群（高脂血症）、B群（糖尿病）、C群（高血圧症）の単独群と、A・B群、A・C群、B・C群、A・B・C群という重複群の8カテゴリに分類し、それぞれの群の70歳以上である高齢群をN'、A'、B'、……というように区別して、加齢の関連性についても調べたいと思う。パラメータとして左右頸動脈の中内膜複合体の最大値を2で割ったものを平均ITMとして用いた。また、プラーク存在は、その性状により1点（hard）、2点（mixed）、3点（intermediate）、4点（soft）、5点（ulcer）として左右頸動脈の点数の和をプラークスコアとして表現した。各群間でのプラーク性状の有無も調べたいと思う。

A このままのデータの形式だと、グループを基準変数にして、量的な変数については「分散分析」を行う、というのが手っ取り早い方法である。質的な変数については、クロス表の検定で十分であろう。

N群とその他の群を比較するために、個別の組合せ（NとA、NとBなど）で、多重ロジスティックモデルを使う方法も考えられる。この場合、質的な変数はダミー変数にする必要がある（この辺はもう少し説明が必要であろう）。

高齢者をA' とかダッシュを付けて分類しているが、その必要はない。年齢の影響があるのならば上記の方法で出てくるはずである。もしも、年齢の影響が階段状になっていると考えているのならば、この分類でもよいのだが、そうとは考えられない。

● おわりに

今回をもって、私の統計学講義を終わらせていただきます。長くやってきた割には、重要な方法の解説が抜けていたりして、申し訳ない気がします。私のホームページには、いくつかの方法の解説もありますので、興味ある読者は、一度見に来てください。それでは、また何かの機会にお目にかかれと思います。

参考文献

- 1) 高木廣文：判別分析と多重ロジスティックモデル，医学のあゆみ1995；174(4)：285-289.
- 2) 高木廣文：3グループ以上の平均値を比べる，超音波検査技術1998；23(6)：461-466.
- 3) 高木廣文：臨床研究のための統計学-V. 多群の比較，臨床泌尿器1986；40(11)：879-885.