

統計処理について御教えください?

解答 : 小川 龍

研究計画を実行してデータ(標本)が得られたら統計解析を行います。群数が3群以上あるいは時系列で3点以上の観察点がある場合を日常よく経験します。そして任意の2群(2点)間の平均値の差を知りたいことがあります。このようなとき、すぐに2群間の平均値をt検定で検定し、有意差があれば2群(2点)間で差があると結論したくなります。これはGlantz¹⁾がしてきたように統計学の落とし穴の一つです。

またA, B, Cの3群があるときに、t検定で5%の有意水準下でA=B, B=C, A=Cの統計仮説が棄却されても、 $A \neq B \neq C$ が成立するとは限りません。

通常3群以上がある場合は、それらの平均値の一樣性は分散分析で検定します。しかし多数群のうちの2群の検定方法は、広く理解されているとは限りません。ご質問はこのような場合の統計検定方法です。以下順を追って説明します。

1. 多重検定法とはなにか

基礎的研究でも、臨床的研究でも3群以上を設定する場合があります。そして3群以上のうちから任意に2群をとりだして、平均値の差を検定する場合、一回の検定では済まず、検定を繰り返すこととなります。このよう検定を繰り返すことを多重検定(multiple comparison test)といいます。ご質問にあるFisher's LSDPなどはこの多重比較法です。原理は次のようになります。まず各群のデータが正規分布しており、分散が等しい(同じ母集団より抽出した)との前提に立ちます。任意の2群(x_i, x_j)の平均値の差($\bar{x}_i - \bar{x}_j$)がその誤差(固体差と測定誤差)より大きければ、差が

あると言えます。そして誤差をどのようにして推定するかにより様々な方法が生まれております。

本稿では多重比較のうち最も広く使われている5種類についてそれらの特徴を述べます。これらの解析法の違いを理解して、適応いただければ大変スマートな統計解析が行われます。

2. 多重検定法の実験

1) LSDP法

Fisher's least significant difference procedureの略です。通常の統計検定法の積み重ねです。

第一段階: 検定の有意水準を5%と設定します。

第二段階: m群の平均値間に差がない(一樣性)という仮説を建て、一元配置分散分析を行います。

第三段階: その結果平均値の一樣性の仮説が棄却できなければ、2群の平均値に差がないと結論します。仮説が棄却できたら次へ進みます。

第四段階: 2群を取り出し、適切なt検定を行います。その結果2群の平均値に差が見られたら、抽出した2群の平均値に差があると結論します。

実験にはこのように二段階の手順を踏むのはやや繁雑であります。また分散分析ではデータが独立で正規分布していること、標準偏差が等しいことが条件です。これを統計学者は頑健性(robustness)と呼びます。ところが臨床データはそれほどの頑健性を持ち得ない場合があります。また一部の統計学者は多数群の平均値の一樣性が否定されなくても、多重比較ができると考えております。そこで以下に述べる方法が提案されております。

2) Bonferroniの方法

LSDS法を改良した方法で、最も簡便でデータ数が少ない場合にも適用されます。また各群のデータ数が等しくなくても適用できます。質問の

ように9群があり、8つの濃度の異なる群と対照群の違いを検定したいなら、8つの組み合わせが検定の対照となります。8つの組あわせの間でt-testを行う場合、Bonferroniの考えでは有意水準を5% ($\alpha=0.05$) とするならば、群数(自由度)で除して $\alpha=0.05/8=0.00625$ 程度にまで厳しくなければならぬと考えたわけです。

実際には任意の2群を抽出し、t検定を行います。そして t_0 値が、上記の0.00625に対応するt値より大きければ、

$$t_0 > t_{0.00625}$$

2群の平均値が等しいとする仮説を棄却できます。ところで統計数値表より、 $t_{0.00625}$ を探すのは面倒です。正規分布表より簡単にこの値を計算する方法があります。

まずt分布の自由度(ν)を求めます。全部のデータ数の総和を(N)とすれば、 $\nu=N-m$ となります。9群に含まれている標本数(111)より群数9を引くと、 $\nu=102$ となります。Z値を求めるには正規分布表で、 α/m (実際には $0.05/9=0.0056$)に対応するZを求めます。ほぼ $Z=2.996$ となります。これより、

$$t=t'=Z+(Z+Z^3)/4n=3.031$$

となります。

したがって任意の2群を抽出し、独立t検定を行い、 $t_0>3.031$ ならば、仮説(2群の平均値が等しい)を棄却できます。

3) Tukeyの方法

多数群のうちの全ての組み合わせ(mC_2 組)平均値の差の検定に使われます。TukeyもBonferroniと同様にt分布を基にして、多数の群の内の任意の2群の差を検定するための、独自の臨界値を求めました。Tukey's q-testとも呼ばれます。Tukeyの臨界値は通常のテキストでは見られませ

ん。SAS, SPSSなどのコンピュータのプログラムパッケージの含まれております。通常のように取り出した2群を対応のないt-testを行い、t値がTukeyの臨界値より大きければ、2群に差があるといえます。臨界値はBonferroniのt'値とほぼ同様になります。Tukeyの方法を適用するには各群のデータ数が等しいことが条件です。

4) Dunnettの方法

多数群を特定の一部(対照群)と比較するのに便利なのはDunnettの方法です。本方法は予め指定した対照群と他の群とを比較するもので、Dunnettが計算した臨海値を使います。したがって $m-1$ 回の多重比較が行えます。対照群のデータ数は他の群より多く $\sqrt{m-1}$ 倍設定することが薦められます。実際はSASなどのプログラム・パッケージを使い、コンピュータで解析します。

5. Scheffeの方法

全ての組み合わせの検定に使えます。例えば多数群中の2群の平均値と3群の平均値の検定も可能です。しかし各群のデータ数が等しいことが条件です。

おわりに

多重比較法はこのほかにDuncan法、Student-Newman-Keul法などがあります。しかし、臨床医としてはLSDP, Bonferroni, Dunnettの方法を理解して使い分ければ十分と思われます。

文 献

- 1) Glantz SA: Biostatistics: How to detect, correct and prevent errors in the medic a literature. Circulation 61: 1~7, 1980