クロンバックの α 係数は負にもなることについて

鈴木治郎 [*1]

[*1] 信州大学 全学教育機構

[*1]szkjiro@shinshu-u.ac.jp

キーワード　多次元尺度，信頼性係数，逆転項目

1　はじめに

Yahoo!知恵袋（2012 年 12 月 2 日）に次のような質問回答例がある．

質問　**クロンバックの信頼性分析で負の値が出たとき，どういうことが考えられますか?**

回答　**項目全体の分散を項目間分散が超えることはないので，計算が間違っているとしか考えられません.**

クロンバックの信頼性係数 α は 1 に近いほうが望ましいことを質問項目の設計に生かす使い方を勧めてはいても，負にもなることを明記する書籍は見かけない．次節に α の定義をあげたが，負にもなることが明らかな定義式（次節の式 (1) はその例）を載せている書籍 [2] も，実は意外と少ない．

たとえば統計 Web[1] では（次節の定義式に合わせて記号は変えた），次のように説明している．

> n は質問の項目数，s_{ii} は各質問項目の分散，s_x^2 は各質問項目を合計した尺度得点の分散としたとき，次節の式 (2) から α を求められる．

この例に限らず，上の回答のように**項目全体の分散と項目間分散の比較**として説明してあるものを多く目にする．また，大学共通学力テストの CBT 化への研究も進む中 [3]，テスト設計の基本となる項目反応理論（Item Response Theory；IRT）の内的整合性を測るクロンバックの信頼性係数 α の正しい理解は必要であろう．

以下では数値例も参考に，上にあげた回答者の錯誤の問題を論じたい．

2　αの定義式

設問項目 i と項目 j の標本共分散を s_{ij} で表すとき，クロンバックの信頼性係数 α は次式で定義される [2].

$$\alpha = \frac{n}{n-1}\frac{\sum_{j=1}^{n}\sum_{\substack{i=1\\ i\neq j}}^{n} s_{ij}}{s_x^2} \tag{1}$$

$$= \frac{n}{n-1}\left(1-\frac{\sum_{i=1}^{n} s_{ii}}{s_x^2}\right) \tag{2}$$

右辺の式変形では標本共分散が次の関係式を満たすことを用いている．

$$s_x^2 = \sum_{j=1}^{n}\sum_{i=1}^{n} s_{ij} = \sum_{i=1}^{n} s_{ii} + \sum_{j=1}^{n}\sum_{\substack{i=1\\ i\neq j}}^{n} s_{ij}$$

とくに自身の共分散 s_{ii} とは項目 i の標本分散 s_i^2 のことであり，つまり $s_{ii} = s_i^2$ である．書籍によっては項目間分散，項目内分散などのことばを通じて説明しているものもあるが，基本的に各設問の相関関係が正である（各設問の評価に整合性がある，逆転がない）ことを測る指標が α であり，共分散を用いた式 (1) による定義の方が直観的であると思う．

定義の (1) 式からは，$\alpha < 0$ になる可能性があることは，共分散の定義から明らかである．はじめに取り上げた回答の誤解は，(2) 式の分母を項目全体の分散とみなしたためであろう．

定義式の理解の助けに，極端な例を一つ考察しておこう．すべての項目の標本分散 $s_{ii} = s_i^2$ が 1 に正規化されていると仮定するとき，共分散は相関係数に等しい．このとき，すべての相関係数が 1 であれば，式 (1), (2) におけるそれぞれの和は項数に等しいので，次のように計算できる．

$$\sum_{j=1}^{n}\sum_{i=1}^{n} s_{ij} = n^2$$

$$\sum_{i=1}^{n} s_{ii} = n$$

$$\sum_{j=1}^{n}\sum_{\substack{i=1\\i\neq j}}^{n} s_{ij} = n(n-1)$$

なので

$$\alpha = \frac{n}{n-1}\frac{n-1}{n} = 1$$

である．

3 数値例による実験

次の数値例を考察する．このデータは，実は筆者が在職する大学に在籍の大学院生が「クロンバックの信頼性係数が負になった」のだが，利用していた統計解析ソフトウェア（JMP）に誤りがないのだろうか，と相談を受けたことに端を発する．その学生も，この数値が必ず正 になると思い込んでいたから質問したのである．それを，$\alpha < 0$ が再現される限りにおいて簡略化した．

設問番号 j	1	2	3	4
得点 x_1	9	8	13	9
得点 x_2	14	13	11	9

$\alpha < 0$ であるとは，式 (2) にある分数の分子が分母より大きくなるのを確認すればよい．

以下で用いた R の関数はベクトルを引数にとる var（分散）と cor（相関係数）だけである．このデータに対する R を用いた検証コードは以下の通りである．

```
# データの準備
x1 <- c(9,8,13,9)
x2 <- c(14,13,11,9)
var(x1)+var(x2)  #（2）式の分子
[1] 9.833333
# x1+x2 と x1,x2 は同じサイズのベクトル
var(x1+x2)  #（2）式の分母
[1] 7
# x2 の得点を負にしたベクトルを準備
x3 <- -x2
var(x1)+var(x3)  #（2）式の分子
[1] 9.833333
var(x1+x3)  #（2）式の分母
[1] 12.66667
# 相関係数の計算
cor(x1,x2)
[1] -0.2881356
cor(x1,x3)
[1] 0.2881356
```

計算結果からわかるように，用意したデータの設定では $\alpha < 0$ であり，$x_3 = -x_2$ を用いた設定では $\alpha > 0$ である．

4 数値例の解釈

数値例より得られた符号の違いをどう解釈すればよいのだろうか．そのためにテストの信頼性とは何かいう基本に立ち返りたい．

もしも設問がある種の能力を測るのに妥当であるとすれば，各設問で高得点をとることは各被験者に共通となるべきである．ところが，この数値例では得点 x_1 と x_2 の相関係数は -0.29 である．これは被験者 1 に対して妥当と思える得点が，被験者 2 に対しては逆の得点を与えているのであり，それはテストそのものの妥当性が低いからである．そのことをクロンバックの信頼性係数が $\alpha < 0$ として示したのである．

だから，α の値は負の可能性があることも踏まえた上で，1 に近い値になるようテストを設計すべきなのである．

5 終わりに

さまざまな数量分析を行う上でコンピュータの利用は避けられない．計算パッケージが信頼できることはもちろんであるが，計算値の解釈は，数学的な理由も含めてできる範囲で確かめて利用したいものである．

最後に，群馬大学の青木繁浩氏 [4] に感謝したい．学生からの質問に対して自信をもって返答できたのは，私の疎い分野の解釈を確認してくれた氏のおかげである．

参考文献

[1] 統計 Web，クロンバックのアルファ，
https://bellcurve.jp/statistics/glossary/1274.html

[2] 豊田秀樹,『項目反応理論［入門編］第 2 版』，朝倉書店，2012 年
[3] 大学入試センター，『CBT について』, https://www.dnc.ac.jp/research/cbt/
[4] 青木繁浩,『R による統計解析』，オーム社，2015 年，http://aoki2.si.gunma-u.ac.jp/R/