

オプション情報： $\alpha$ 係数をめぐって

近年は内部一貫性に関して $\omega$ 係数が参照されることも多くなっていますが、依然として $\alpha$ 係数はよく利用、参照されている指標でしょう。測定対象となる概念の性質や項目数などを踏まえて多面的に検討されなければならない、かなり扱いが難しい係数だと思いますが、「.70を超えているから信頼性あり」とか、「項目数が少ないから、.68だけれどOKでしょう」といった感じの、読んでいて不安を感じる論文にも出会います。また、「先行研究で信頼性が確認されているので、(たとえ自分のデータでの $\alpha$ 係数が低くても)OK」とか、「先行研究でも $\alpha$ は低めだったので…」というようなものにも…

少々雑というか、尺度を作る側も、利用する側も、もう少し情報をもった検討、議論をする方がよいと感じることもしばしばです。何となくですが、妥当性の検討には注意が向いているのに比べ、信頼性(内部一貫性)に対しては(かなり)甘いという傾向があるように感じます。問題があれば、そこを修正してよりよいものを作ればよいだけなのではないでしょうか。こういう、尺度の「発展」に向けた動きがあまり見られないのは残念に思います。

私の感想はこれくらいにして、 $\alpha$ 係数を検討する時に使えるものはRにいくつもあります。

これまでに発表された論文はもちろん、最近の論文でも $\alpha$ 係数の95%信頼区間が明示されているものは多くありません。しかし、それを使って母集団についての検定をするならば、 $\alpha$ 係数についても95%信頼区間くらいは見ておいて損はないと思います。

psychパッケージのalpha関数を使えば、簡単に95%信頼区間は手に入ります。 $\alpha=.70$ を気にするなら(この数値の根拠がどこにあるのか私も知りませんが)、サンプルにおける $\alpha$ をこの基準と照合するより、信頼区間の下限を使って照合した方が妥当な感じがするのですが…。また尺度を構成する項目間相関係数の平均値も合わせて示してくれるので、項目数に影響を受ける $\alpha$ 係数の情報と合わせてみるとよいのではないかと。

尺度を作成する場合は、こういった結果が自分のデータから得られるので、それをじっくり見ればよいでしょう。しかし、尺度を利用しようと先行論文の記載情報からその尺度について検討する場合は、記載された $\alpha$ 係数自体とか項目数、著者の解釈などを参考にするしかありません。私もそう思っていたのですが、こういう時に便利に使える関数があることを知りましたので紹介しておきます。

#### ●先行研究で記載されている $\alpha$ 係数などから、 $\alpha$ 係数の95%信頼区間を求める

以前の論文はもちろん、最近の論文でも $\alpha$ 係数の95%信頼区間が明示されているものは多くはないでしょう。そこで、psychパッケージのalpha.ci関数を使うと、サンプルの人

数と得られている  $\alpha$  係数などから、95%信頼区間を Feldt の手順 (Feldt et al., 1987) に沿って計算した結果を出してくれます。

以下は、先行研究においてサンプルが 126 人、尺度に含まれる項目数が 5、得られた  $\alpha$  係数が.73 だったときのスクリプトです。

```
print(alpha.ci(.73, 126, n.var=5))
```

これを実行すると、この場合の 95%信頼区間は、0.65 から 0.80 であることがわかります。

つまり、先行研究での母集団における真の  $\alpha$  は 0.65 から 0.80 の間にある確率が高いわけです (真の値が.70 以下である可能性が結構ある)。こうなると (こういう情報が手に入ると)、「 $\alpha$  係数は.73」であったという論文記載情報だけよりも、この尺度の信頼性は大丈夫かな…、と慎重になれるのではないのでしょうか。また、そう判断すれば、もう少し内部一貫性が高まるような修正を試みようかという建設的な動きも生まれるような気がします。

Feldt, L. S., Woodruff, D. J., & Salih, F. A. (1987). Statistical inference for coefficient alpha. *Applied Psychological Measurement*, 11, 93-103.

●先行研究で記載されている  $\alpha$  係数から、尺度を構成している項目間相関の平均値を求める

以下のスクリプトは、psych パッケージの `alpha2r` 関数を用い、上と同じ、尺度に含まれる項目数が 5、得られた  $\alpha$  係数が.73 だったとき、その 5 項目間の相関の平均値を算出するものです。なおこの場合にサンプルの人数は関係しません。

```
alpha2r(alpha = .73, n.var=5)
```

この結果は、0.3509615 となります。つまり、この尺度を構成する項目間の相関係数は、平均すると.35 程度となります。ちなみに 95%信頼区間であった、0.65 と 0.80 の場合を計算してみると 0.2708333 と 0.4444444 となります。

ここからは、分析者の主観的判断になるのですが、測定したいひとつの構成概念を測定するための 5 つの項目間の相関係数はどれくらいを想定するのか。もし、「.35 程度以上はほしいな」と思っていたら、この尺度をそのまま使うことはかなりリスクを伴うといえるでしょう。こういった判断をしやすくなる情報を与えてくれると思います。

ちなみに、もしこれが 20 項目から構成されている尺度であれば…。計算してみてください。もし項目数 20、 $\alpha$ =.73 という尺度があったとしたら、これをどう評価すべきでしょう。 $\alpha$  係数は、項目数によって左右される値であることがよくわかると思います。

●複数の  $\alpha$  係数を検定する

複数の  $\alpha$  係数の検定は、cocron パッケージでできます。たとえば、先行研究で得られている  $\alpha$  係数と、自分が得たデータから計算した  $\alpha$  を検定する（つまり同じ母集団を仮定できるか…）ことや、再検査信頼性を相関係数からだけでなく、 $\alpha$  係数の点（検定）でも確認できるということです（たぶん…）。

先行研究で得られている  $\alpha$  係数との比較の場合（2つの  $\alpha$  係数の検定）は、以下のように  $\alpha$  係数と人数があれば計算できます（41人の.78と、151人の.71）。

```
cocron.two.coefficients(alpha=c(.78,.71), n=c(41,151), dep=FALSE)
```

同一サンプルの場合は相関係数も必要なようですが、これを使うのは他者の研究データではなく、自分のデータでしょうから、計算は問題ないでしょう。

なお、cocron パッケージはRを使っていなくても、同一製作者が以下のwebを開設していて、そこで必要なものを入力すると計算してくれます。

<http://comparingcronbachalphas.org>