

19日目：偏相関係数

さて、今日は偏相関係数の算出について紹介します。個人的には、相関と偏相関の差異を眺めてみるのは好きなのですが、論文などではあまり見かけないです…

Rでは、`psych`の中に偏相関係数を算出するコマンドはあるのですが、使い勝手があまりよくないのです…。また自作関数ですかね…

まずは`psych`にある関数を紹介します。

`partial.r(x, c(2, 3), c(4))`

このように入力すると、ファイル`x`の中の、2、3列目の変数間の偏相関係数を、4列目の変数をコントロールして計算してくれます。カッコの中は、ファイル名、偏相関を求める変数、コントロールする変数という順で並べます。

結果はシンプルに右のようになります。

```
> library(psych)
> partial.r(x, c(2, 3), c(4))
partial correlations
              親の様子  家にある本
親の様子      1.00      0.65
家にある本    0.65      1.00
> |
```

2、3列目の変数間の偏相関係数を、4、5、6列目の変数をコントロールして計算したいならば…

`partial.r(x, c(2, 3), c(4, 5, 6))`

4、5、6列目の変数間の偏相関係数を、2、3列目の変数をコントロールして計算したいならば…

`partial.r(x, c(4, 5, 6), c(2, 3))`

これだけなのでとても簡単に使えますが、表示は小数点以下2桁までのようです。もう1桁欲しい場合は、`print`でコントロールすればできるようですが、どうも無相関検定はやってくれないようです…

`print(partial.r(x, c(2, 3), c(4)), 3)`

Webを探してみると、偏相関に関する自作関数はいくつも見つけることができます。たとえば青木先生のサイトに以下のような自作関数が紹介されていました。

<http://aoki2.si.gunma-u.ac.jp/R/partial-cor.html>

これは、たとえば5変数を投入した場合、偏相関を求める2変数以外の変数をコントロールした場合の偏相関係数を計算してくれます（もし、a1からa5を投入したなら、a1とa2の偏相関の場合は、残るa3, a4, a5をコントロールし、a1とa3の偏相関の場合は、残るa2, a4, a5をコントロールして計算してくれるということです）。

偏相関係数行列

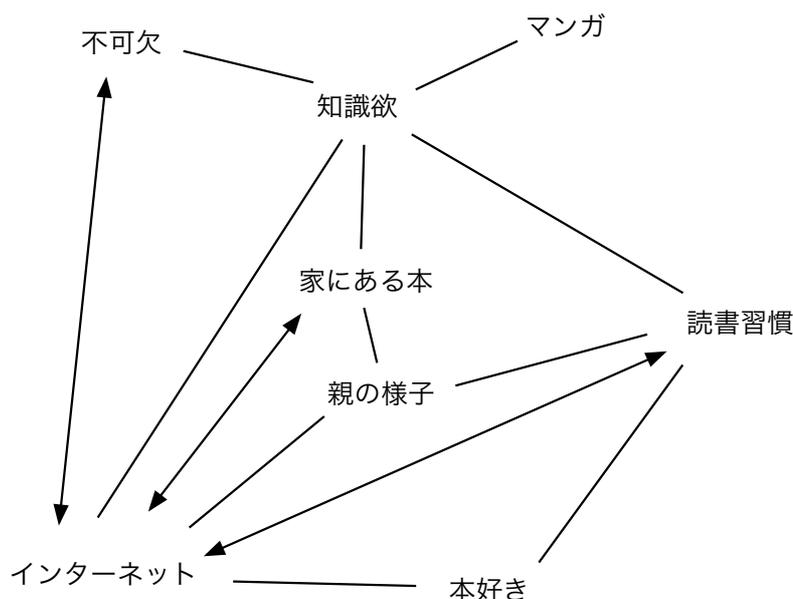
```
partial.cor <- function(x) # データ行列
{
  x <- subset(x, complete.cases(x)) # 欠損値を持つケースを除く
  i <- solve(cor(x)) # 相関係数行列の逆行列
  d <- diag(i) # 対角成分
  i <- -i/sqrt(outer(d, d)) # 偏相関係数行列
  diag(i) <- NA # 対角成分は未定義
  rownames(i) <- colnames(i) <- paste("Var", 1:ncol(x))
  return(i)
}
```

行列を使った偏相関係数の計算の仕方に忠実な形なので、計算式と照らし合わせると中身を理解しやすいのではないかと思います。なお、下から3行目、`rownames(i) <- colnames(i) <- paste("Var", 1:ncol(x))`を削除しておくと、データファイルにある変数名が表示されます。

得られた結果を相関係数とともにまとめたものが以下の表です。一昨日の相関係数の表の左下側に偏相関係数をいれてあります。相関係数と偏相関係数で大きく異なっているところもあります。

	親の様子	家にある本	知識欲	不可欠さ	インターネット	マンガ	本好き	読書習慣
親の様子	---	.687 ***	.288 ***	.325 ***	.150 ns	.044 ns	.475 ***	.493 ***
家にある本	.609	---	.340 ***	.347 ***	-.058 ns	-.046 ns	.321 ***	.422 ***
知識欲	-.157	.214	---	.549 ***	.275 ***	.312 ***	.223 **	.395 ***
不可欠さ	.132	.052	.496	---	.034 ns	.172 *	.246 **	.234 **
インターネット	.237	-.229	.352	-.215	---	.301 ***	.237 **	-.030 ns
マンガ	.053	-.123	.218	.058	.193	---	.025 ns	.006 ns
本好き	.158	.002	-.169	.173	.307	-.051	---	.547 ***
読書習慣	.235	.008	.382	-.178	-.308	-.039	.465	---

この偏相関係数を使って、変数感の関連を図示してみました。偏相関係数が.200以上のものを取り上げ、正の場合は実線で、負の場合は両端に矢印のある実線でつないでみました。変数間の構造を眺める一つの方法だと思います。…どうも、インターネットの項目がいろいろな（本質的ではなさそうな）影響を与えているような感じがしますが、でっちあげたダメデータだからですかね…



さて、やはり無相関検定の結果もほしい…。ということなら、自作する必要がありそうです。web上にも、あまり公開されていないようですが、計算自体はそれほど複雑ではないので、丁寧にtとdfを求める計算式を作っていけば、それほど難しくないような…気がします。

本日はここまでにします。明日は回帰分析を使ってみたいと思います。