

16 日目 : 2 要因分散分析 (被験者間)

さて、本日は 2 要因分散分析を行ってみます。一昨日、昨日という流れ (分析としての) からするとおかしいことになってしまいますが、因子分析の結果から構成された下位尺度得点 (`total_f1`, `total_f2`, `total_f3`) について、性と専攻を要因とする 2 要因分散分析を行うことをやってみます。

しつこいですが、まずはそれぞれの群別の平均値、標準偏差などを確認しておきます。検定をやってから、平均値を確認するのは順序が逆です!!!

方法は、`describeBy` が便利でしょう。ですが今回は、群を分ける変数が 2 つになります。`describeBy` は、複数の群を分ける変数に対応してくれるのか?

結論から言うと、やってくれます。`describeBy(x[*:*], x[**:**])` というような命令をやってくれます。しかし、「性別」と「専攻」は 2 列目と 4 列目であり、3 列目に余分なものが入っているので、`x[2:4]` とやるとおかしなことになってしまいます。

そこで、14 日目にもやった方法で、変数をまとめて指定します。つまり…

```
label_f <- c("total_f1", "total_f2", "total_f3")
label_g <- c("性別", "専攻")
describeBy(x[label_f], x[label_g])
```

これで、性別、専攻別に計算してくれます。8 日目に紹介した `mat=TRUE` を加えることもできますので、自分が作りたい形式に便利な方を選びましょう。

`total_f1` の結果を簡単にまとめると次のような表になります。

		専攻		
		group1	group2	group3
1	平均値	17.79	17.92	17.49
	標準偏差	3.93	4.63	3.93
	標準誤差	0.73	0.60	0.65
	n	29	59	37
性別	平均値	18.60	17.95	19.52
	標準偏差	3.99	3.93	3.88
	標準誤差	0.42	0.86	0.49
	n	90	21	62

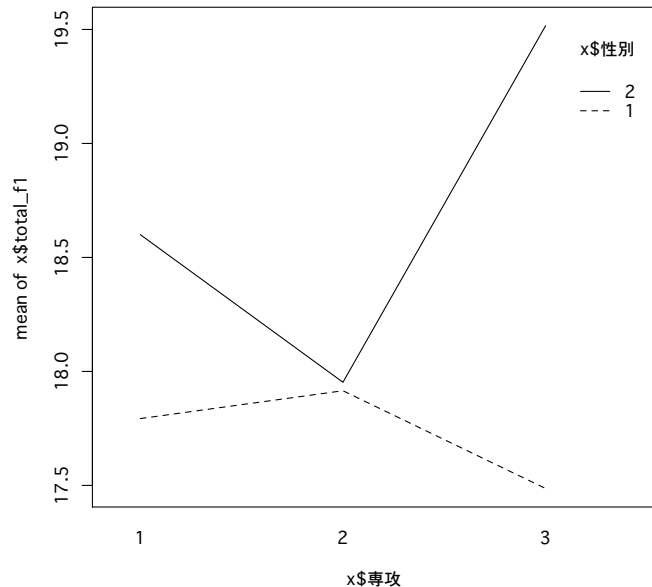
次に、これをグラフにしてみます。2 要因分散分析だと、交互作用を説明する時のグラフが見やすいかもしれません。あのようなものも作れます。

```
par(family="Osaka")  
interaction.plot(x$専攻, x$性別, x$total_f1)
```

これで右のようなグラフを作成して
くれます。

日本語が入るので、フォントの設定
をしています。interaction.plot
のカッコの中は、グラフの横軸（独立
変数の1つ）、線の区別（独立変数の
もう1つ）、グラフの縦軸（従属変数）
の順になります。

グラフからだと、交互作用がありそ
うにも思えますが、検定結果はどの
うなのでしょう…



さて、ここから分析に入りますが、1 要因分散分析同様、独立変数を factor 型にしておく
必要があります。

今回の独立変数は「性別」と「専攻」の2つですから、これに対して…

```
x$性別 <- as.factor(x$性別)  
x$専攻 <- as.factor(x$専攻)
```

これで2つを factor 型に変更します。確認の仕方は昨日と同じです。

R には2 要因分散分析を行う命令が2 種類ほどあるようですが、今回も aov を使ってみ
ます。使い方は以下のようです。カッコの中は、「従属変数」～「独立変数1」*「独立
変数2」, 「データ」と並びます。

```
aov2.f1 <- aov(total_f1 ~ 性別 * 専攻, x)  
summary(aov2.f1)
```

結果は、次の図のように表示されます。典型的な分散分析表ですので説明ははぶきます。

```
> aov2.f1 <- aov(total_f1 ~ 性別 * 専攻, x)
> summary(aov2.f1)
      Df Sum Sq Mean Sq F value Pr(>F)
性別    1     86   86.17   5.162 0.0238 *
専攻    2     15    7.45   0.446 0.6404
性別:専攻  2     39   19.69   1.179 0.3089
Residuals 292  4875   16.69
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

性別の主効果が、 $F(1,292)=5.162$ で5%水準の有意。専攻の主効果、交互作用は有意ではないということが読み取れます。先に作成したグラフでは相互作用がありそうな感じもありましたが、検定の結果、性差のみが有意ということがわかりました。

私がやるなら、このような結果で十分なのですが、必要な場合もあるでしょうから、とりあえず多重比較のやり方を書いておきます。

チューキー法なら、昨日の命令を少し変えるだけでできます。`TukeyHSD()`のカッコの中を、本日のもの書き換えればそれでOKです。

`TukeyHSD(aov2.f1)`

もしくは

`TukeyHSD(aov(total_f1 ~ 性別 * 専攻, x))`

もちろん、単純主効果を計算することもできますが、ここでは省略します。

私自身が分散分析はかなり苦手ということもあり、十分な説明はできていません。詳しいwebもありますし、ANOVA4やANOVA君といった便利なものも多く公開されているので、必要ならばそちらをあたってください。

本日はここまでです。