

14日目：t検定（対応のない場合）

さて、本日は因子分析の結果から構成された下位尺度得点について、t検定を行ってみたいと思います。

SPSSでt検定を行ったことのある人は、等分散性の検定結果を見て、上下どちらの結果を読み取るかを定める…といった手順を覚えていると思います。等分散を仮定できない場合の検定はwelchのt検定などとも呼ばれますが、Rのデフォルトはこのwelchの方法になります。

どうせRでやるなら、すべてwelchの方法でやればいいのではないかと思ったりもしますが…

やろうとしていることを確認しますが、因子分析の結果から構成された下位尺度得点（**xx**に入っている、**to.f1**, **to.f2**, **to.f3**）に対して、性についてのt検定を行おうということです。そこでまずは、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、7日目にやった**describe.by**が便利でしょう。

3つの下位尺度得点をそれぞれ…

describe.by(xx\$to.f1, xx\$性別)

とやってもよいですし、下位尺度得点の列番号を確認しておいて

describe.by(xx[26:28], xx\$性別)

とまとめてやることもできます。

出力結果を整理すると、
右のような状況になります。
このような差について検定
をするわけです。

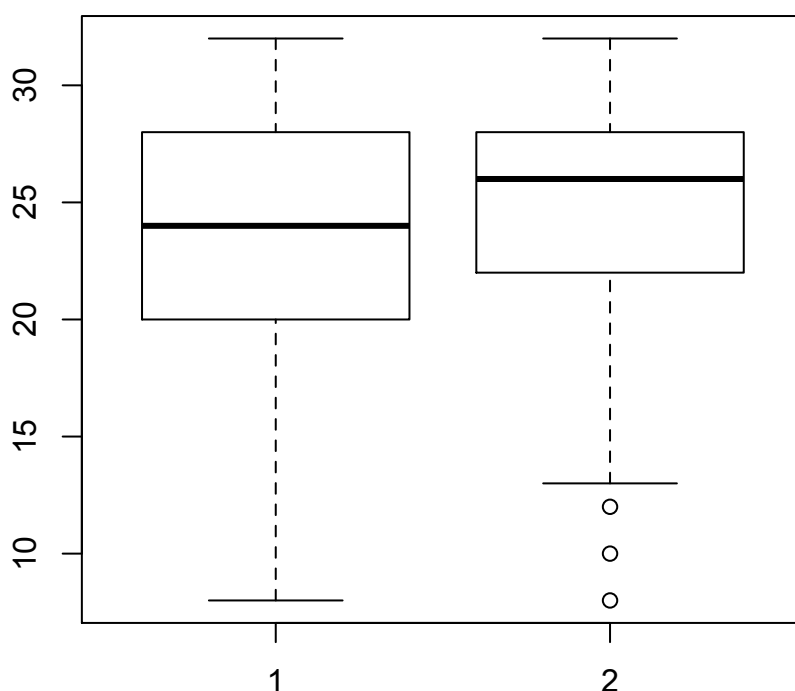
		n	mean	sd
to.f1	group:1	125	23.65	5.27
	group:2	173	24.84	4.76
to.f2	group:1	125	12.62	3.73
	group:2	172	13.82	3.89
to.f3	group:1	125	8.49	3.47
	group:2	173	7.79	2.97

また箱ひげ図を書かせてみるのもよいでしょう。**to.f1**についての箱ひげ図が次です。以下のような簡単な指示でできます。

boxplot(xx\$to.f1 ~xx\$性別)

見方は、箱の上辺が第3四分位数、下辺が第1四分位数の位置を表します。そして箱の中のちょっと太い横線が中央値（平均値ではない！）です。そして箱から伸びた「ひげ」の先は、ちょっとややこしい説明になります。多くの場合（?）、これを最大値と最小値と考えればよいのですが、今回のグループ2のように、その外側に○がプロットされている場合は注意が必要です。

この「ひげ」の先は、たとえば上側なら、第3四分位数から四分位範囲の1.5倍以内にあるデータのうちの最大値を示すようです。そこよりもさらに外れるようなデータは○でプロットされています。つまり、今回のグループ2には、点数が低い方へ大きく外れたデータが少なくとも3つあるということがわかります。



前置きが長くなりましたが、このようなチェックをした後に t 検定をやってみましょう。等分散を仮定しない `welch` の t 検定を実行する命令は以下のようです。

```
t.test(xx$to.f1 ~ xx$性別)
```

カッコ内は、前に平均値が検定される変数を、「~」で区切って、後ろに群を分ける変数を入れます。

結果は以下のように表示されます。

```
> t.test(xx$to.f1 ~ xx$性別)

Welch Two Sample t-test

data:  xx$to.f1 by xx$性別
t = -2.0035, df = 250.565, p-value = 0.0462
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.36009613 -0.02020445
sample estimates:
mean in group 1 mean in group 2
 23.64800      24.83815
```

3行目に、t 値、自由度 (df)、p 値が並んでいます。これだと、5%水準で有意な差があるという結果ですね。その下は対立仮説、95%信頼区間と並び、最後にそれぞれの群の平均値が示されています。

まとめて一気にやりたいなら、以下のように for を使ってやることもできます。でも、やはり変数名が a と表示されてしまうので、間違えてしまいそう…

```
for(a in 26:28) {
  print(t.test(xx[,a] ~ xx$性別))
}
```

結果を一覧にしたら、以下のようになります。もちろん、t 値は絶対値表示にして、df は小数点以下1桁で示しています。論文なら、「t 検定を行った」と書くところを、「welch の t 検定を行った」としておけば十分なのではないでしょうか。

		n	mean	sd	t値 自由度
to.f1	group:1	125	23.65	5.27	2.00 *
	group:2	173	24.84	4.76	250.6
to.f2	group:1	125	12.62	3.73	2.68 **
	group:2	172	13.82	3.89	273.6
to.f3	group:1	125	8.49	3.47	1.83
	group:2	173	7.79	2.97	241.0

* p<.05; * p<.01

これで十分だと思うのですが、等分散の検定をして、等分散ならその場合の t 検定をした
いという場合は、以下のようです。

- ・等分散の検定

```
var.test(xx$to.f1 ~ xx$性別)
```

- ・等分散を仮定する場合の t 検定

```
t.test(xx$to.f1 ~ xx$性別, var.equal=TRUE)
```

その他、対応のある場合の t 検定とか、平均は0でないことを検定する場合の t 検定とか
いろいろありますが、必要な場合には適宜情報をさぐってください。

本日はここまでにします。