1 introduction

In this example I’ll use fictitious data, taken from http://www.ruf.rice.edu/~mickey/psyc339/notes/RMANOVA.html. The dependent variable IQ was measured for 5 participants at 3 ages. Is there a difference in the different IQ’s measured at different ages — is there an age effect on IQ? ¹

2 “fixed” model

If all factors were fixed, then we could use the ANOVA command. Using the Syntax window of SPSS your input should be:

* fixed two-way ANOVA.
ANOVA
   iq  BY Age(1,3) Subj(1,5).

¹Because MANOVA requires contiguous factor levels, without gaps, the non-contiguous values of Age (5,25,45) were changed into contiguous codes (1,2,3).
The first factor, Age, takes three different values ("levels"), and the second factor, Subject, has five levels. This command always assumes a factorial design, and effects are always tested against the within-cell error variance. In other words, the denominator of the $F$ ratio (the so-called error term) is always the variance within cells. In this case, however, that variance is zero, because there is only one observation for each combination of Age and Subj. Hence, no $F$ ratios can be computed. This analysis is clearly not appropriate here.

<table>
<thead>
<tr>
<th></th>
<th>Sum of Sq</th>
<th>df</th>
<th>Mean Sq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main Effects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGE</td>
<td>12.133</td>
<td>2</td>
<td>6.067</td>
</tr>
<tr>
<td>SUBJ</td>
<td>23.333</td>
<td>4</td>
<td>5.833</td>
</tr>
<tr>
<td>2-Way Interactions</td>
<td>1.867</td>
<td>8</td>
<td>0.233</td>
</tr>
<tr>
<td>Model</td>
<td>37.333</td>
<td>14</td>
<td>2.667</td>
</tr>
<tr>
<td>Residual</td>
<td>0.000</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>37.333</td>
<td>14</td>
<td>2.667</td>
</tr>
</tbody>
</table>

3 “mixed” model, univariate

In other designs, some factors are not fixed, or the design is not factorial. In these cases, you cannot use the within-cell variance as error term. You should specify which effects should be tested against which error terms. You cannot do that with the ANOVA command; you must use MANOVA instead, with an explicit design specification.

* mixed ANOVA, univariate.
MANOVA IQ BY Subj(1,5) Age(1,3)
  /METHOD ESTIMATION (NOCONSTANT)
  /PRINT=HOMOGENEITY (COCHRAN)
  /DESIGN
      Subj BY Age = 1,
      Age VS 1.

In the last subcommand /DESIGN, the interaction is mentioned first, because this effect is later used as error term as well. The interaction term is abbreviated as 1 for later use. The interaction does not need to be tested, so no error term is specified. The main effect of Age is tested against the interaction effect, just abbreviated as 1. The main effect of random factor Subj is not tested, because there is no suitable error term for it.

The command above produces the following output, slightly edited here for presentation purposes:
Tests of Significance for IQ using UNIQUE sums of squares

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>Sig of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error 1</td>
<td>1.87</td>
<td>8</td>
<td>.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGE</td>
<td>12.13</td>
<td>2</td>
<td>6.07</td>
<td>26.00</td>
<td>.000</td>
</tr>
</tbody>
</table>

The last line shows that the fixed effect of Age is tested against error term 1, specified above (and not tested), and it also shows that this main effect is significant, $F(2, 8) = 26, p < .001$. The SS’s in this analysis are equal to those from the fixed model above (please check), but the $F$ ratio is computed differently here.

This univariate analysis requires the assumption of sphericity. Informally, this means that all age differences must have equal variances. Using the 3 repeated measures, we can compute 2 new difference measures. Do the two difference variables have the same variance?

```plaintext
COMPUTE D1 = IQ2-IQ1 .
COMPUTE D2 = IQ3-IQ2 .
DESCRIPTIVES
  VARIABLES=d1 d2
  /STATISTICS=MEAN STDDEV VARIANCE SEMEAN .
```

The output below shows that the variances of these differences are not equal. Because of the very small size of the sample, however, $H0: s_{D1} = s_{D2}$ cannot be rejected.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>Std.Error</th>
<th>Std.Dev.</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>5</td>
<td>1.0000</td>
<td>.3162</td>
<td>.7071</td>
<td>.500</td>
</tr>
<tr>
<td>D2</td>
<td>5</td>
<td>1.2000</td>
<td>.2000</td>
<td>.4472</td>
<td>.200</td>
</tr>
</tbody>
</table>

Nevertheless it seems appropriate to assume no sphericity for these data. Therefore we need to resort to a multivariate ANOVA.

4 repeated measures, multivariate

A better test, with higher power, uses Repeated Measures ANOVA. This does not require the assumption that the differences have equal variances. The effect of fixed factor Age is not compared as a main effect, outside participants, but only within participants. In most cases this is your best option.
Because of this within-subject comparison of Age, there is no interaction possible between Age and Subj, or more generally, between the main effect and the observation units. This interaction cannot be separated from the residual variance or measurement error. This was also noted for the fixed model above.

Unfortunately, a Repeated Measures ANOVA is not easy in SPSS. First, you need a so-called multivariate data layout. All data from one participant (observation unit) need to be on one row in the data matrix. Within each participants, data need to be sorted according to the within-subject factor. This yields the following arrangement:

```
2 4 5
5 6 7
3 3 4
1 2 4
2 3 4
```

The accompanying command for Repeated Measures ANOVA is as follows, in the Syntax window:

* mixed ANOVA, multivariate.
MANOVA iq1 iq2 iq3
  /WSFACTOR age(3)
  /MEASURE = iq
  /ANALYSIS (REPEATED).

This indicates that the 3 observations for each participant must be taken as 3 conditions of the fixed within-subject factor Age. The dependent variable gets the label IQ, and this needs to be analyzed as a Repeated Measures ANOVA.

The long output will be discussed below (with minor editing) in portions.

Tests of Between-Subjects Effects.
Tests of Significance for T1 using UNIQUE sums of squares
Source of Variation  SS    DF    MS    F Sig of F
WITHIN CELLS       23.33  4    5.83
CONSTANT           201.67  1  201.67  34.57  .004

Are there any differences between participants? This study has no between-subject factors; otherwise they would be tested here. In this analysis I’ve refrained from testing the differences between subjects, given the low number of subjects. The only term remaining is the constant, viz. the grand
mean. This is tested against the differences between participants, yielding $F(1, 4) = 34.6, p = .004$. So the grand mean is significantly different from zero. Indeed. Let’s move on quickly to more interesting results.

Tests involving 'AGE' Within-Subject Effect.

Mauchly sphericity test, $W = .61224$
Chi-square approx. = 1.47187 with 2 D. F.
Significance = .479

Greenhouse-Geisser Epsilon = .72059
Huynh-Feldt Epsilon = 1.00000
Lower-bound Epsilon = .50000

Here the sphericity assumption is investigated with a dedicated test, Mauchly’s $W$. We’ve tried to do the same test informally above. This formal test here is not significant, $p = .479$, due to the small sample size. Hence no correction is necessary for the resulting $F$ ratios.

If Mauchly’s $W$ had been significant, then we should have multiplied the degrees of freedom for the resulting $F$ ratios with the Huynh-Feldt $\epsilon$, a number between $1/(k-1)$ (here 1/2) and 1. That would make the statistical testing more conservative. But that’s not required now.

First we see the results of the multivariate ANOVA:

EFFECT .. AGE
Multivariate Tests of Significance (S = 1, M = 0, N = 1/2)

<table>
<thead>
<tr>
<th>Test Name</th>
<th>Value</th>
<th>Exact F</th>
<th>Hypoth. DF</th>
<th>Error DF</th>
<th>Sig. of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pillais</td>
<td>.92000</td>
<td>17.25000</td>
<td>2.00</td>
<td>3.00</td>
<td>.023</td>
</tr>
<tr>
<td>Hotellings</td>
<td>11.5000</td>
<td>17.25000</td>
<td>2.00</td>
<td>3.00</td>
<td>.023</td>
</tr>
<tr>
<td>Wilks</td>
<td>.08000</td>
<td>17.25000</td>
<td>2.00</td>
<td>3.00</td>
<td>.023</td>
</tr>
<tr>
<td>Roys</td>
<td>.92000</td>
<td></td>
<td>2.00</td>
<td>3.00</td>
<td></td>
</tr>
</tbody>
</table>

Note.. F statistics are exact.

The $F$ ratio for the within-subject effect of Age can be computed in many ways, all shown here, and all yielding the same result. Wilks’ $F$ ratio is often the most useable. The results are clear: $F(2, 3) = 17.25, p = .023$, hence $H_0$ can be rejected. The IQ scores are not identical at all ages.

As a bonus this command also gives you the output of a univariate analysis. Results should be the same as in the previous section, and indeed: for Age, $F(2, 8) = 26, p < .001$. 

5
AVERAGED Tests of Significance that follow multivariate tests are equivalent to univariate or split-plot or mixed-model approach to repeated measures. Epsilons may be used to adjust d.f. for the AVERAGED results.

Tests involving 'AGE' Within-Subject Effect.

AVERAGED Tests of Significance for IQ using UNIQUE sums of squares

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F</th>
<th>Sig of F</th>
</tr>
</thead>
<tbody>
<tr>
<td>WITHIN CELLS</td>
<td>1.87</td>
<td>8</td>
<td>.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGE</td>
<td>12.13</td>
<td>2</td>
<td>6.07</td>
<td>26.00</td>
<td>.000</td>
</tr>
</tbody>
</table>