Nested ANOVA

A nested ANOVA is one that might “appear” to be a two way ANOVA but isn't.

The reason for this is that not all levels of two factors occur with each other.

When one level of one factor occurs with a level of another factor, those two levels are said to be crossed.

In a factorial or regular two way ANOVA, all factors are crossed with each other.

This is not true in a nested design.

Let's give an example before we get confused.

But, we'll use our cats again. But this time we'll set things up differently.

Example: We want to compare the effects of three different anesthetics.

The problem is, we don't have enough anesthetic from supplier 1, so we need to get more from a 2nd and even 3rd supplier.

So we decide to parcel out the anesthetic from the suppliers in such a way so that we might be able to evaluate the effect of supplier. But we can't do blocking (not enough drug).

Supplier 1 gives us 5 vials of A and 5 vials of C
Supplier 2 gives us 5 vials of A, 5 vials of B, and 5 vials of C
Supplier 3 gives us 5 vials B

How do we set up our experiment?

Remember, we are mostly interested in the effect of the anesthetic. Supplier is (hopefully) not that important, but we need to make sure.

We divide our cats into 3 groups of 10 each:

Drug A    Drug B    Drug C

*each group gets 10 cats*

Now we subdivide each group to account for our supplier:

Drug A    Drug B    Drug C
supp 1    supp 2    supp 2    supp 3    supp 1    supp 2
5 cats    5 cats    5 cats    5 cats    5 cats    5 cats
Notice the setup. We subdivide our cats based on both supplier and drug, but this is not a fully crossed design:

supplier 3 only occurs with drug B.
supplier 1 only occurs with drug A and C.

For a fully two way ANOVA we'd need every supplier to be crossed with every drug.

This is a “nested” design, since not every “sub” treatment occurs at every full treatment; the subtreatments are “nested” within the treatments.

See also your text on page 308, but notice that in the example there, every “source” is present only once (but still doesn't occur with every drug).

Either design will work.

Analyzing a nested ANOVA:

We need to calculate 5 different sums of squares.

We'll skip the mathematical details (they're on pages 310 - 311, but please ignore the machine formulas) and summarize them in words:

SS\(_{\text{total}}\) The total sum of squares. We calculate this as usual by subtracting each measurement in turn from the grand mean, squaring, and adding up the total.

SS\(_{\text{groups}}\) The group sum of squares. This is for the treatment we're really interested in. We calculate it the same way as we would the SS\(_{\text{groups}}\) for a one way ANOVA (subtract the grand mean from the group mean, square, and multiply by \(n_i\), and add all of the result for each group). Above we'd have 3 groups.

SS\(_{\text{among subgroups}}\) The among subgroup sum of squares. We pretend that each subgroup is now a group, and calculate the sum of squares (subtract the grand mean from each subgroup mean, square, multiply by \(n_{ij}\), etc.). Above we'd have 6 subgroups.

SS\(_{\text{subgroups}}\) The sum of squares for the subgroups within a particular group. Think of it as trying to get at the variation between subgroups but within the same group. Think of comparing \(\bar{x}_{ij}\) with \(\bar{x}_j\). The easiest way to calculate it is to take advantage of the following:

\[
\text{SS}_{\text{subgroups}} = \text{SS}_{\text{among subgroups}} - \text{SS}_{\text{groups}}
\]

SS\(_{\text{error}}\) Since there are duplicate measurements within each group, this could be considered SS\(_{\text{within}}\). We calculate it the easy way by doing:

\[
\text{SS}_{\text{error}} = \text{SS}_{\text{total}} - \text{SS}_{\text{among subgroups}}
\]

Once we have all the various sums of squares, we arrange them into an ANOVA table as follows:
among subgroups
SS among subgroups ab-1
SS groups a-1 SS groups/(a-1) MS groups/MS subgroups
SS subgroups a(b-1) SS subgroups/a(b-1) MS subgroups/MS error
error SS error N-ab SS error/(N-ab)
total SS total N-1

The main thing we're interested is differences in our groups. If this is more variable than differences in subgroups (which is hopefully irrelevant), then we have a significant effect. That's why we divide MS groups by MS subgroups.

Let's walk through the example on page 308/309, skipping some of the math details:

Finally, let's see how to do this in R:

Start the usual way:

```
level <- scan()
102 103 108 109 104 105 104 104 110 108 106 107
drug <- scan()
1 1 2 2 3 3 1 1 2 2 3 3
srce <- scan(what = character())
a q d b l s a q d b l s
drug <- factor(drug)
srce <- factor(srce)
```

Now do the following:

```
chol <- aov(level ~ drug/srce)
```

This version of the aov command tells R to keep src inside drug:

```
summary(chol)
```

```
Df Sum Sq Mean Sq F value  Pr(>F)
drug         2  61.17   30.58  20.389 0.00211 **
drug:srce    3   1.50    0.50   0.333 0.80220
Residuals    6   9.00    1.50
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

And from this we see that the effect of subgroups (our “source” - the second line) is N.S. (source didn't make a difference). However, the first line for our groups (drug) is not quite right.
Unfortunately, R, by default, divides the $\text{MS}_{\text{drug}}/\text{MS}_{\text{residuals}}$, which is wrong. So we need to do it again and tell R to use the right denominator:

\[
\text{choll} \leftarrow \text{aov}(\text{level} \sim \text{drug} + \text{Error(drug:srce)})
\]

R will complain about the model being singular - just ignore it.

\[
\text{summary(choll)}
\]

\[
\begin{array}{lrrrr}
\text{Error: drug:srce} \\
\text{Df} & \text{Sum Sq} & \text{Mean Sq} & \text{F value} & \text{Pr} (> F) \\
\text{drug} & 2 & 61.17 & 30.58 & 61.17 & 0.0037 ** \\
\text{Residuals} & 3 & 1.50 & 0.50 & \\
\end{array}
\]

*** Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

\[
\begin{array}{lrrr}
\text{Error: Within} \\
\text{Df} & \text{Sum Sq} & \text{Mean Sq} & \text{Pr} (> F) \\
\text{Residuals} & 6 & 9 & 1.5
\end{array}
\]

Our second line now disappeared, but this time R correctly divides $\text{MS}_{\text{groups}}$ (notice that this command shoved the $\text{SS}_{\text{subgroups}}$ into $\text{SS}_{\text{residuals}}$, but it all works out).

(Comment: I haven't tried it, but I suspect you can use the “Error” above to get the correct $F$ value in a mixed model as well).

One way ANOVA when the data are not normally distributed.

(This is the same material as was presented in 312, with some extra comments for using R)

Suppose you have a one way design, and want to do an ANOVA, but discover that your data are seriously not normal?

Just like with the MWU test as “replacement” for the t-test, there is the Kruskal-Wallis test for a one way ANOVA.

In fact, if you have just two groups, the KW test will give you the identical results to a MWU test (sort of like a ANOVA gives the same result as a equal variance t-test).

Before we look at the KW test, note that there are other things that we can do:

Remember transformations. They can make non-normal data normal.

That's probably the only way to deal with more complicated designs. Things like a two way ANOVA, nested designs, etc., don't have non-parametric equivalents (well, there is one for two way designs... more on that later).

Check out your text in chapter 13 for some ideas.

Let's introduce the KW test:

Just like the MWU test, if make no assumptions other than randomness, we can use the KW test
pretty much anytime. It has no other assumptions we need to worry about.

But what are we testing now?

- \( H_0 \): All \( k \) population distributions are identical
- \( H_1 \): At least one of these is different (tends to yield larger/smaller observations than the others).

Just like last time, though, if we assume distributions are similar except in location, we can use means (or medians) instead.

Then proceed as usual:

- select \( \alpha \)
- calculate test statistic (let’s call it \( W^* \))
- compare it to a value from KW tables
- if \( W^* \geq W \text{ table} \) (but see below), then reject.

So, what is \( W^* \)?

It’s a bit messy, and is given by (note I’m doing things just a bit differently from the text):

\[
W^* = \frac{1}{S^2} \left( \sum_{i=1}^{k} \frac{R_i^2}{n_i} - \frac{N(N+1)^2}{4} \right)
\]

(Your text (on page 217) uses a slightly different formula).

Yech! But let’s not give up yet. It’s really not that difficult. Let’s figure out what the various parts of this are:

- \( N \) = total sample size (\( = n^* \), using our ANOVA notation)
- \( R_i \) = sum of the ranks of the \( i^{th} \) sample:

\[
R_i = \sum_{j=1}^{n} R(X_{ij})
\]

- this says, take the rank of each of the x’s and sum these values for each sample (this is very similar to \( K_1 \) and \( K_2 \) in the Mann-Whitney test, except now we’re using ranks instead of “number smaller in other sample”)

- notice that the very first thing we’ll have to do is to rank our observations from smallest to largest (write in the rank next to our each of our observations).
- \( S^2 \) = an analogue of the variance (notice this is capitalized). It’s given as follows:

\[
S^2 = \frac{1}{n} \left( \sum_{i=1}^{k} \frac{R_i^2}{n_i} - \frac{N(N+1)^2}{4} \right)
\]
\[ S^2 = \frac{1}{N-1} \left( \sum_{\text{all ranks}} R(X_i)^2 - N \left( \frac{N+1}{4} \right) \right) \]

- okay, now you’re really worried. But let’s stick with it for just a little longer.

- this says, Sum up all the square of each of the ranks, then subtract the other quantity (the one involving N). That’s not too bad. It’s actually very similar to the “calculator” formula for variance that we didn’t have a chance to talk about (but are in my notes).

- Now we know how to calculate W* (it’ll be much more obvious after an example). But what about W?

- This is a bit of a problem. Remember how you needed to have n and n’ in the Mann-Whitney U-test?

- now we have 2, 3, 4 ... or more samples, so we really can’t list the probabilities for everything.

- Let’s take another quick look at W*. It turns out if there are no ties, we can re-write it as:

\[ W^* = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(N+1) \]

- this is identical to the formula in your text for “no ties”.

- (if there are no ties, then use this formula - it’s a lot easier)

- the stuff in the sum symbol is kind of analogous to what we “observed” -we’re using the actual ranks and getting a kind of sum (squared and adjusted by \( n_i \), but still a sum).

- the 3(N+1) term is kind of what we would expect for the sums if the ranks were basically equal in each sample.

- the term in front of the sum symbol is another type of “expected” value. \{Note: the sum of a bunch of numbers going from 1...n = n(n+1)/2, so you might recognize that both 3(N+1) and the quantity out front are a little similar\}.

- so we have an observed quantity (-) an expected quantity, where the observed is also divided by an expected quantity. Sound just a little familiar?

- As it turns out W* has (approximately) the Chi-square distribution.

- So what do you look up for W? A Chi-square value with k-1 degrees of freedom.

- Your text takes a different approach, and for small sample sizes and less than four groups gives you the exact values in table B.13. We'll just stick with the Chi-square approach since it's much easier.
- As mentioned, this is an approximation, and exact tables do exist, but they take up a good portion of a book, and your text only does this for small $n$ and $k$ up to 4.

- So how does it all work?? An example (let’s use an example similar to the pig diet example, but this time we’re looking at sheep

  - let’s do

    H0: the diets are all the same.
    H1: at least one of the diets is different.
    $\alpha = .05$

<table>
<thead>
<tr>
<th>Diet 1</th>
<th>Diet 2</th>
<th>Diet 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>value</td>
<td>rank</td>
<td>rank$^2$</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>8.5</td>
<td>72.25</td>
</tr>
<tr>
<td>9</td>
<td>3.5</td>
<td>12.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sum ($R_i$) | 14 | 88.5 | 41 | 385 | 23 | 175

- Notice that $N = n^* = 12$; also notice that the highest rank is 12, so everything’s fine so far (if one or more values are tied, you need to use the average rank).

- Now we calculate $S^2$:

$$S^2 = \frac{1}{11}\left(649 - 12 \left(\frac{13^2}{4}\right)\right) = 12.909$$

- Now we “merely” plug all this into $W^*$ to get:

$$W^* = \frac{1}{12.909}\left(\frac{14^2}{3} + \frac{41^2}{5} + \frac{23^2}{4} - 12 \left(\frac{13^2}{4}\right)\right) = 2.075$$

- Let’s get $W_{table}$, using our Chi-square table and 2 d.f.:

$W_{table} = 5.99$

- And since $W^* < W_{table}$, we “fail to reject” and conclude we have no evidence to show the diets are different.

- Okay, let’s summarize:

  - When do you use KW? When you don’t meet the assumptions of ANOVA:

    - As usual, with a larger sample size, ANOVA will start to do better, and you don’t need to worry about the normal assumption as much.

    - What about equal variances? If you're worried about unequal variances, don't test for equal means. Use the general form of the KW test.
- How about power? Pretty good; even when the data are normal it doesn’t do too badly.

- Why not use it all the time?
  - As you probably know by now, ANOVA is much more flexible. The number of designs available for ANOVA is almost endless.
  - and even though the power isn’t bad, it isn’t the best test to use if data are normal.

- Multiple comparisons are available for the KW test, but we just don’t have time to dig into that as well.

- We didn’t do too much with the theory here. The basic idea is very similar to that of the Mann-Whitney U test, the only odd thing is that we can get away with using Chi-square tables instead of exact tables.

Doing Kruskal-Wallis in R:

The basic setup is almost identical to a one way ANOVA:

```r
gain <- scan()
8 16 9 9 16 21 11 18 15 10 17 6
diet <- scan()
1 1 1 2 2 2 2 2 3 3 3 3
diet <- factor(diet)
kruskal.test(gain ~ diet)
```

```
Kruskal-Wallis rank sum test
data:  gain by diet
Kruskal-Wallis chi-squared = 2.0748, df = 2, p-value = 0.3544
```

(R calls the test statistics “Kruskal-Wallis chi-squared”)

- Finally, what about two way or more complicated designs?
  - The Kruskal-Wallis test is designed for one-way type analyses.
  - There are other non-parametric tests for more complicated designs; if you need to use something like this, you can look up the “Quade” test or the “Friedman” test (the Friedman test is related to the sign test).

The Friedman test works for two way designs, but isn't very powerful.

The Quade test works for blocked designs, and is worth looking at if you have serious problems with normality and are using a blocked design.