NESTED FACTORS (Chapter 18)

A factor B is *nested* in factor A if each level of B occurs in conjunction with only one level of A. (In other words, if there is a completely different set of levels of B for every level of A.)

Recall: Two factors A and B are *crossed* if every level of A occurs with every level of B.

Examples: In each of the following examples:

a. Identify which pairs of factors are crossed and which are nested.
b. Identify which factors are fixed and which are random.
c. Decide which interactions between factors make sense.
d. Identify the experimental units.

1. The amount of vitamin A in a jar of baby food carrots might vary from brand to brand and might also vary between jars of the same brand. To study the effect of these two factors on vitamin A content, a researcher randomly selected five jars of baby carrots from each of the three brands of baby food sold in a particular city.

2. Gum arabic is used to lengthen the shelf length of emulsions. It comes from acacia trees and is processed for use in emulsions. Four raw gum arabic samples are obtained from each of two different varieties of acacia tree (for a total of eight samples.) Each sample is split into two subsamples. For each sample, one of the subsamples is randomly chosen and given the experimental treatment. The other serves as a control The sixteen subsamples are dried, and an emulsion made from each. The response is the time until the emulsion begins to separate.

Note: In Example 2, the word “sample” has a different meaning than the usual one in statistics.

Models for Nested Factors

Notation: Use subscript j(i) to denote that the factor indexed by j is nested in the factor indexed by i.

Examples: 1. In the baby food example above, let A denote brand and B denote jar. Use index i = 1, 2, 3 for A and index j = 1, .. , 5 for B. Suppose there are r = 2 measurements of vitamin A content of each jar. The model is

$$Y_{ijt} \text{ (or } Y_{t(ij)}) = \mu + \alpha_i + B_{j(i)} + \varepsilon_{ijt},$$

where

each $B_{j(i)} \sim N(0, \sigma_{B(A)}^2)$,

each $\varepsilon_{ijt} \sim N(0, \sigma^2)$,

and the $B_{j(i)}$'s and ε_{ijt}'s are all mutually independent random variables.

Note how the nesting subscript notation is also used in the subscript on $\sigma_{B(A)}^2$.

Also, ε_{ijt} is sometimes expressed as $\varepsilon_{t(ij)}$.

2. In the gum Arabic example, let

V = variety of tree (so v =__),

S = sample (so s = __),

A = treatment (so a = __),

and r = the number of measurements made on each subsample.

Then the model equation is

Y_{ijkt} =

For the unrestricted model, the additional assumptions are:

Fitting and Model Checks for Nested Designs:

- Fits and residuals are obtained by Least Squares.
- Residuals can be used for the usual model checks as appropriate for the model and as possible.
- Assumptions of normality of variance components are usually not possible to check.

Analysis for Nested Designs:

- The exact details will depend on the particular design, but the same general ideas as for previous designs are used.
- See Section 18.3 for more details for fixed effects models.
- See Section 18.4 for the cases when random effects are included in the model.
- Software will usually do most calculations, including giving expected mean squares

Using software for nested designs: In Minitab:

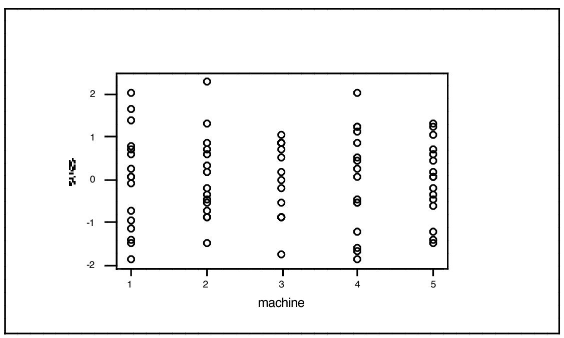
- Use B(A) to denote that B is nested in A.

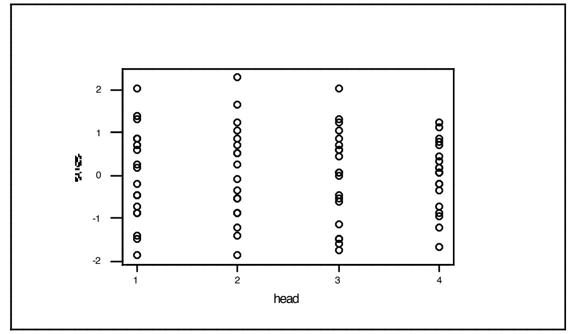
  Example: A|B(A)|C designates the model with terms
  - A
  - B(A) (i.e., B nested in A)
  - C
  - A*C
  - B(A)*C
- Note: B(A)*C is sometimes denoted as B*C(A) or (B*C)(A).
- If B is nested in A and A is random, then B is automatically considered random.

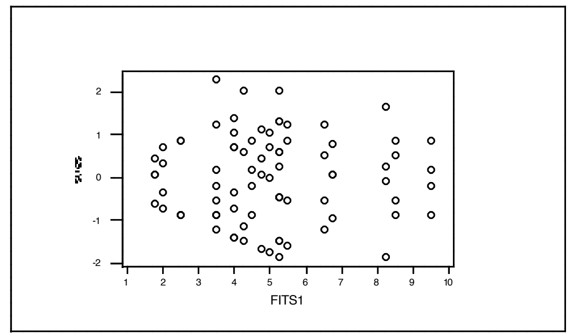
Example: A certain type of machine in a manufacturing process used heads that wear out and need to be replaced periodically. The variability of the effect of head on strain reading was studied. The only machines of interest were the five used in the study. Each head was tested on only one machine. Twenty heads were used, four on each machine.

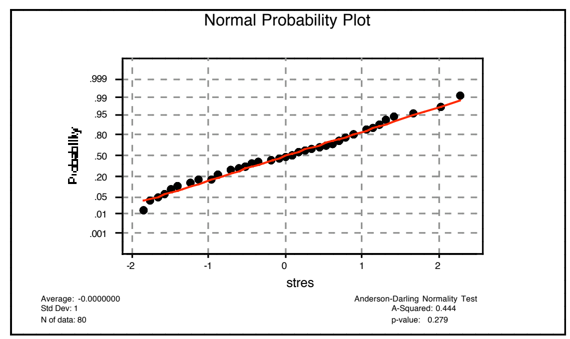
Model:

Model checking plots: What can we and can't we check?









Output:

```
Analysis of Variance (Balanced Designs)


Factor          Type Levels Values
machine        fixed      5     1     2     3     4     5
head(machine) random      4     1     2     3     4


Analysis of Variance for y

Source          DF        SS        MS       F      P
machine          4     45.08     11.27    0.60  0.670
head(machine)   15    282.88     18.86    1.76  0.063
Error           60    642.00     10.70
Total           79    969.95


Source          Variance Error Expected Mean Square
                component  term (using unrestricted model)
 1 machine                 2   (3) + 4(2) + Q[1]
 2 head(machine)    2.040  3   (3) + 4(2)
 3 Error           10.700      (3)
```

Discussion:

Note: The model in this example is a special case of what are sometimes "hierarchical models": Models that have *only* nested factors (i.e., no crossed factors).

Example: In the semiconductor industry, "gauge studies" might have 4 batches (B), 5 wafers per batch (W(B)), 3 placements per wafer (P(BW), and 2 observations per placement.

> [*Caution*: "Hierarchical model" has (at least) two other meanings in statistics:
> 1. Models which, when an interaction is included, include all smaller interactions and all factors in the interaction. For example, if ABC interaction is included in the model, then to make the model hierarchical in this second meaning, AB, AC, BC, A, B, and C effects would also be included in the model.
> 2. Any model including nested factors.]

For example, if there are three factors A, B(A), C(AB), all random, then the model is

$Y_{ij(i)k(ij)t} =$

where

The expected mean squares are

A: $\sigma^2 + r\,\sigma_{C(AB)}^2 + cr\sigma_{B(A)}^2 + bcr\sigma_A^2$

B(A): $\sigma^2 + r\,\sigma_{C(AB)}^2 + cr\sigma_{B(A)}^2$

C(AB): $\sigma^2 + r\,\sigma_{C(AB)}^2$

So the test statistics for variance components are:

For H_0^A: $\sigma_A^2 = 0$ ______________

For $H_0^{B(A)}$: $\sigma_{B(A)}^2 = 0$: ________________

For $H_0^{C(AB)}$: $\sigma_{C(AB))}^2 = 0$: ______________