Extracting correct F-tests from statistics packages

by Matthew Williams

In analysing ANOVA designs using general-purpose statistics packages such as SAS and Systat, it is usual for the package to report incorrect F values and associated probabilities, except in the case of the completely randomized design (CRD). This is because statistics packages are not smart enough to recognise the design of an experiment. The packages routinely use the "error" term in the denominator of the F ratio, whether this is appropriate for particular effects or not. Since the CRD is rarely used, it is usually necessary to persuade the package to work out the correct F tests. These must be artificially constructed by explicitly requesting the package to estimate the correct denominator terms in the model. How this is done varies, but in SAS and Systat at least, it is fairly straightforward. The method for each type of design is specified below:

Randomized complete block (1 observation per cell (treatment/block) combination):

SYSTAT
correct F-test reported

SAS
correct F-test reported

Randomized complete block over multiple sites (with one observation per cell):

SYSTAT

```
MGLH
CATEGORY SITE BLOCK TREAT
MODEL X = CONSTANT + SITE + BLOCK{SITE} + TREAT + SITE*TREAT
RUN
HYPOTHESIS
EFFECT = SITE
ERROR  = BLOCK{SITE}
TEST
```

SAS

```
PROC GLM;
CLASS SITE BLOCK TREAT;
MODEL X = SITE BLOCK(SITE) TREAT SITE*TREAT;
TEST H=SITE  E=BLOCK(SITE);
RUN;
```

Randomized complete block with split-plot arrangement (i.e. multiple observations per cell):

SYSTAT

```
MGLH
CATEGORY BLOCK TREAT
MODEL X = CONSTANT + BLOCK + TREAT + BLOCK*TREAT
RUN
HYPOTHESIS
EFFECT = TREAT
ERROR  = BLOCK*TREAT
TEST
```

SAS

```
PROC GLM;
CLASS BLOCK TREAT;
MODEL X = BLOCK TREAT BLOCK*TREAT;
TEST H=TREAT  E=BLOCK*TREAT;
RUN;
```