## 10.6 THE PSEUDO-*F* TEST

Occasionally the EMS column for a given experiment indicates that there is no exact $F$ test for one or more factors in the design model. Consider the following example.

*Example 10.3* Two days in a given month were randomly selected in which to run an experiment. Three operators were also selected at random from a large pool of available operators. The experiment consisted of measuring the dry-film thickness of varnish in mils for three different gate settings: 2, 4, and 6 mils. Two determinations were made by each operator each day and at each of the three gate settings. Results are shown in Table 10.3.

**Table 10.3** Dry-Film Thickness Experiment

| | Day | | | | | |
|---|---|---|---|---|---|---|
| | 1 | | | 2 | | |
| Gate | Operator | | | Operator | | |
| Setting | *A* | *B* | *C* | *A* | *B* | *C* |
| 2 | 0.38 | 0.39 | 0.45 | 0.40 | 0.39 | 0.41 |
| | 0.40 | 0.41 | 0.40 | 0.40 | 0.43 | 0.40 |
| 4 | 0.63 | 0.72 | 0.78 | 0.68 | 0.77 | 0.85 |
| | 0.59 | 0.70 | 0.79 | 0.66 | 0.76 | 0.84 |
| 6 | 0.76 | 0.95 | 1.03 | 0.86 | 0.86 | 1.01 |
| | 0.78 | 0.96 | 1.06 | 0.82 | 0.85 | 0.98 |

Assuming that days and operators are random effects, gate settings are fixed, and the design is completely randomized, the analysis yields Table 10.4.

**Table 10.4** Analysis of Dry-Film Thickness Experiment

| Source | df | SS | MS | EMS |
|---|---|---|---|---|
| Days $D$ | 1 | 0.0010 | 0.0010 | $\sigma_\varepsilon^2 + 6\sigma_{DO}^2 + 18\sigma_D^2$ |
| Operators $O$ | 2 | 0.1121 | 0.0560 | $\sigma_\varepsilon^2 + 6\sigma_{DO}^2 + 12\sigma_O^2$ |
| $D \times O$ interaction | 2 | 0.0060 | 0.0030** | $\sigma_\varepsilon^2 + 6\sigma_{DO}^2$ |
| Gate setting $G$ | 2 | 1.5732 | 0.7866** | $\sigma_\varepsilon^2 + 2\sigma_{DOG}^2 + 4\sigma_{OG}^2$ $+ 6\sigma_{DG}^2 + 12\phi_G$ |
| $D \times G$ interaction | 2 | 0.0113 | 0.0056 | $\sigma_\varepsilon^2 + 2\sigma_{DOG}^2 + 6\sigma_{DG}^2$ |
| $O \times G$ interaction | 4 | 0.0428 | 0.0107 | $\sigma_\varepsilon^2 + 2\sigma_{DOG}^2 + 4\sigma_{OG}^2$ |
| $D \times O \times G$ interaction | 4 | 0.0099 | 0.0025** | $\sigma_\varepsilon^2 + 2\sigma_{DOG}^2$ |
| Error | 18 | 0.0059 | 0.0003 | $\sigma_\varepsilon^2$ |
| Totals | 35 | 1.7622 | | |

** Two asterisks indicate significance at the 1-percent level.

All $F$ tests are clear from the EMS column except for the test on gate setting, which is probably the most important factor in the experiment. Only the three-way interaction shows significance. It is obvious from the results that gate setting is the most important factor, but how can it be tested? If the $D \times G$ interaction is assumed to be zero, then the gate setting can be tested against the $O \times G$ interaction term. On the other hand, if the $O \times G$ interaction is assumed to be zero, gate setting can be tested against the $D \times G$ interaction term. Although neither of these interactions is significant at the 5 percent level, both are numerically larger than the $D \times O \times G$ interaction against which they are tested. In this case any test on $G$ is contingent upon these tests on interaction. One method for testing hypotheses in such situations was developed by Satterthwaite and is given in Bennett and Franklin [4, pp. 367–368].

The scheme consists of constructing a mean square as a linear combination of the mean squares in the experiment, where the EMS for this mean squares includes the same terms as in the EMS of the term being tested, except for the variance of that term. For example, to test the gate-setting effect $G$ in Table 10.4 a mean square is to be constructed whose expected value is

$$\sigma_\varepsilon^2 + 2\sigma_{DOG}^2 + 4\sigma_{OG}^2 + 6\sigma_{DG}^2$$

This can be found by the linear combination

$$\text{MS} = \text{MS}_{DG} + \text{MS}_{OG} - \text{MS}_{DOG}$$

as its expected value is

$$\begin{aligned} E(\text{MS}) &= \sigma_\varepsilon^2 + 2\sigma_{DOG}^2 + 4\sigma_{OG}^2 + \sigma_\varepsilon^2 + 2\sigma_{DOG}^2 + 6\sigma_{DG}^2 - \sigma_\varepsilon^2 - 2\sigma_{DOG}^2 \\ &= \sigma_\varepsilon^2 + 2\sigma_{DOG}^2 + 4\sigma_{OG}^2 + 6\sigma_{DG}^2 \end{aligned}$$

An $F$ test can now be constructed using the mean square for gate setting as the numerator and this mean square as the denominator. Such a test is called a pseudo-$F$ or $F'$ test. The real problem here is to determine the degrees of freedom for the denominator mean square. According to Bennett and Franklin, if

$$\text{MS} = a_1(\text{MS})_1 + a_2(\text{MS})_2 + \cdots$$

and $(\text{MS})_1$ is based on $v_1$ df, $(\text{MS})_2$ is based on $v_2$ df, and so on, then the degrees of freedom for MS are

$$v = \frac{(\text{MS})^2}{a_1^2[(\text{MS})_1^2/v_1] + a_2^2[(\text{MS})_2^2/v_2] + \cdots}$$

In the case of testing for the gate-setting effect above, $a_1 = 1$, $a_2 = 1$, and $a_3 = -1$ and the degrees of freedom are $v_1 = 4$, $v_2 = 2$, and $v_3 = 4$. Here

MS = 0.0107 + 0.0056 − 0.0025 = 0.0138 and its df is

$$v = \frac{(0.0138)^2}{(1)^2[(0.0107)^2/4] + (1)^2[(0.0056)^2/2] + (-1)^2[(0.0025)^2/4]}$$

$$= \frac{1.9044 \times 10^{-4}}{0.4586 \times 10^{-4}} = 4.2$$

Hence the $F'$ test is

$$F' = \frac{\text{MS}_G}{\text{MS}} = \frac{0.7866}{0.0138} = 57.0$$

with 2 and 4.2 df, which is significant at the 1 percent level of significance based on $F$ with 2 and 4 or 2 and 5 df.

## 10.7 REMARKS

The examples in this chapter should be sufficient to show the importance of the EMS column in deciding just what mean squares should be compared in an $F$ test of a given hypothesis. This EMS column is also useful (usually in random models) to solve for components of variance as illustrated in Chapter 3, Section 3.6.

One special case is of interest. In a two-factor factorial when there is but one observation per cell ($n = 1$), the EMS columns of Section 10.3 reduce to those in Table 10.5.

**Table 10.5** EMS for One Observation Per Cell

| Source | EMS (Fixed) | EMS (Random) | EMS (Mixed) |
|---|---|---|---|
| $A_i$ | $\sigma_\varepsilon^2 + b\phi_A$ | $\sigma_\varepsilon^2 + \sigma_{AB}^2 + b\sigma_A^2$ | $\sigma_\varepsilon^2 + \sigma_{AB}^2 + b\phi_A$ |
| $B_j$ | $\sigma_\varepsilon^2 + a\phi_B$ | $\sigma_\varepsilon^2 + \sigma_{AB}^2 + a\sigma_B^2$ | $\sigma_\varepsilon^2 + a\sigma_B^2$ |
| $AB_{ij}$ or $\varepsilon_{ij}$ | $\sigma_\varepsilon^2 + \phi_{AB}$ | $\sigma_\varepsilon^2 + \sigma_{AB}^2$ | $\sigma_\varepsilon^2 + \sigma_{AB}^2$ |

A glance at these EMS values will show that there is no test for the main effects $A$ and $B$ in a fixed model, as interaction and error are hopelessly confounded. The only test possible is to assume that there is no interaction; then $\phi_{AB} = 0$, and the main effects are tested against the error. If a no-interaction assumption is not reasonable from information outside the experiment, the investigator should not run one observation per cell but should replicate the data in a fixed model.

For a random model both main effects can be tested whether interaction is present or not. For a mixed model there is a test for the fixed effect $A$ but