# Experimental Design 4TH EDITION

## Procedures for the Behavioral Sciences

ROGER E. KIRK

BAYLOR UNIVERSITY

SAGE

Los Angeles | London | New Delhi
Singapore | Washington DC

Step-down procedures are widely used in the behavioral sciences, health sciences, and education. Step-up procedures are used less often. Both kinds of procedures tend to be more powerful than single-step procedures. However, step-down and step-up procedures suffer from several shortcomings: (1) In general, they cannot be used to construct confidence intervals; (2) with a few exceptions, they cannot be used to test directional hypotheses; and (3) they tend to require more computation than single-step procedures.

## Five Common Hypothesis-Testing Situations

From a review of the literature in the behavioral sciences, health sciences, and education, I have identified five hypothesis-testing situations that occur with some degree of regularity: testing hypotheses about

1. $p - 1$ a priori orthogonal contrasts
2. $p - 1$ a priori nonorthogonal contrasts involving a control group mean
3. $C$ a priori nonorthogonal contrasts
4. All pairwise contrasts among $p$ means
5. All contrasts including nonpairwise contrasts that appear interesting from an inspection of the data

Contrasts in the first category are a priori and orthogonal; those in the other four categories are nonorthogonal. As discussed earlier, for contrasts in the first hypothesis-testing situation, the usual practice is to adopt the individual contrast as the conceptual unit for a Type I error. For the other four hypothesis-testing situations, it is customary to adopt the family of contrasts as the conceptual unit for a Type I error.

Statisticians have developed a variety of test statistics that can be used to control the Type I error rate in these five situations. Table 5.1-1 summarizes the test statistics that I recommend for each situation. The procedures in the upper part of the table assume normality of the population distributions, random sampling or random assignment, and homogeneity of population variances. Tukey's test, the REGW $FQ$ test, and the REGW $Q$ test also require equal-sized samples. If the assumption of homogeneity of population variances is not tenable or the requirement of equal-sized samples is not met, the multiple comparison procedures in the lower part of Table 5.1-1 can be used. As you will see, the power of the recommended procedures differs markedly. In general, test statistics that were designed for testing a select, limited number of a priori contrasts are more powerful than those designed to test all pairwise comparisons or all possible contrasts. Hence, when possible, it is to a researcher's advantage to specify in advance either orthogonal contrasts or a limited number of contrasts. The problem facing a researcher is to choose the test statistic that provides both the desired kind of Type I error protection and maximum power. The following sections describe the recommended test statistics for each of the five research situations.

**Table 5.1-1** ▪ Multiple Comparison Procedures That Are Recommended for Five Common Research Situations

Recommended Procedures When Assumptions Are Tenable

| | Orthogonal Contrasts | Nonorthogonal Contrasts |
|---|---|---|
| **A priori contrasts** | **1. Testing $p - 1$ contrasts**<br>Student's $t$ test (5.2)* | **2. Testing $p - 1$ contrasts with a control group mean**<br>Dunnett's test (5.3) |
| | | **3. Testing $C$ contrasts****<br>Dunn-Šidák test (5.4)<br>Holm's test (5.4) |
| **A posteriori contrasts** | | **4. Testing all pairwise contrasts****<br>Tukey's test (5.5)<br>Fisher-Hayter test (5.5)<br>REGW $F$, $FQ$, and $Q$ tests (5.5) |
| | | **5. Testing all contrasts**<br>Scheffé's test (5.6) |

Recommended Procedures When Assumptions Are Not Tenable

| | Orthogonal Contrasts | Nonorthogonal Contrasts |
|---|---|---|
| **A priori contrasts** | **1. Testing $p - 1$ contrasts:**<br>**Heterogeneous variances**<br>Student's $t'$ test with Welch degrees of freedom (5.2) | **2. Testing $p - 1$ contrasts with a control group mean:**<br>**Unequal sample $n$s or heterogeneous variances**<br>Dunnett's test with modifications (5.3) |
| | | **3. Testing $C$ contrasts:**<br>**Heterogeneous variances****<br>Dunn-Šidák test with Welch degrees of freedom (5.4)<br>Holm's test with Welch degrees of freedom (5.4) |

*(Continued)*

**Table 5.1-1** ▪ Multiple Comparison Procedures That Are Recommended for Five Common Research Situations (Continued)

| | | |
|---|---|---|
| **A posteriori contrasts** | | **4. Testing all pairwise contrasts:**<br>**Unequal sample sizes**<br>Tukey-Kramer test (5.5)<br>Fisher-Hayter test (5.5)<br>**Heterogeneous variances****<br>Dunnett's *T*3 test (5.5)<br>Dunnett's *C* test (5.5)<br>Games-Howell test (5.5) |
| | | **5. Testing all contrasts:**<br>**Heterogeneous variances**<br>Brown-Forsythe test (5.6) |

*Note:* The recommended procedures control the per-contrast, familywise, or per-family error rate and also have one or a combination of the following virtues: conceptual simplicity, ease of computation, excellent power, availability of confidence intervals, and robustness.

*The numbers in parentheses denote the section in which a procedure is described.

**When more than one procedure is recommended, the procedures are listed in order of increasing power.

## 5.2 Procedures for Testing $p - 1$ a Priori Orthogonal Contrasts

### Student's Multiple *t* Test

Student's $t$ statistic is a single-step procedure that can be used to test null hypotheses of the form

$$H_0\colon \psi_1 = 0$$

$$H_0\colon \psi_2 = 0$$

$$\vdots$$

$$H_0\colon \psi_{p-1} = 0$$

where the $p - 1$ contrasts are a priori and mutually orthogonal. It is not necessary to test the omnibus null hypothesis with an ANOVA $F$ statistic prior to testing the individual contrasts. An omnibus test answers the general question, "Are there any differences among the population means?" If a specific set of orthogonal contrasts has been advanced, a researcher is not interested in this general question. Rather, the researcher is interested in