TABLE* 10.3.2

PROBABILITIES OF TYPE-I ERROR WITH THE F-TEST FOR TESTING THE EQUALITY OF MEANS OF FIVE GROUPS OF FIVE EACH AT THE NOMINAL 5 PER CENT SIGNIFICANCE LEVEL, APPROXIMATED BY A CORRECTION ON THE NUMBERS OF D.F.; γ_1 AND γ_2 ARE THE SKEWNESS AND KURTOSIS OF THE POPULATION OF ERRORS

γ_1^2	Population	γ_2: −1	−0.5	0	0.5	1
0	Pearson	0.053	0.051	0.050	0.048	‡
	Edgeworth	0.053	0.051	0.050	0.049	0.048
	Edgeworth†	0.052	0.051	0.050	0.049	0.048
0.5	Pearson	0.052	0.051	0.050	0.049	‡
	Edgeworth	0.053	0.051	0.050	0.049	0.048
	Edgeworth†	0.053	0.052	0.050	0.049	0.048
1	Pearson	0.052	0.050	0.049	0.048	0.048
	Edgeworth	0.053	0.052	0.050	0.050	0.049
	Edgeworth†	0.053	0.052	0.051	0.050	0.049

* From p. 14 of "Permutation theory in the derivation of robust criteria and the study of departures from assumption" by G. E. P. Box and S. L. Andersen, *J. of the Royal Stat. Soc.*, series B, Vol. 17 (1955). Reproduced with the kind permission of the authors and the editor.

TABLE* 10.4.1

EFFECT OF INEQUALITY OF POPULATION VARIANCES ON PROBABILITY OF TYPE-I ERROR WITH TWO-TAILED t-TEST FOR EQUALITY OF MEANS AT NOMINAL 5 PER CENT SIGNIFICANCE LEVEL. J_1 AND J_2 ARE SAMPLE SIZES, $\theta = \sigma_1^2/\sigma_2^2$, AND σ_1^2 AND σ_2^2 ARE POPULATION VARIANCES

(J_1, J_2)	θ: 0	0.1	0.2	0.5	1	2	5	10	∞
(15, 5)	0.32	0.23	0.18	0.098	0.050	0.025	0.008	0.005	0.002
(5, 3)	0.22	0.14	0.10	0.072	0.050	0.038	0.031	0.030	0.031
(7, 7)	0.072	0.070	0.063	0.058	0.050	0.051	0.058	0.063	0.072

* From p. 12 of "Contributions to the theory of Student's t-test as applied to the problem of two samples" by P. L. Hsu, *Stat. Research Memoirs*, Vol. 2 (1938*a*). Values are here reproduced with the kind permission of the editors.

The rest of this section will deal with tests for means in fixed-effects models when the errors are independently normal. Some exact results are then available for the probability of type-I errors. For the two-tailed t-test for the equality of two means at the nominal 5 per cent level, Table 10.4.1 gives the probabilities[33] of type-I errors when the population

[33] Calculated by P. L. Hsu (1938*a*).

9.4. Prove that if $M = (\mu_{ij})$ is an $I \times \mathcal{J}$ matrix such that for some $J < \mathcal{J}$ every $I \times J$ submatrix of M has equal row sums (i.e., the row main effects are zero), then the rows of M are equal (i.e., every submatrix of M has zero interactions). [*Hint:* To prove that $\mu_{ij} = \mu_{i'j}$ for all i, i', j it suffices to prove that $\mu_{i1} = \mu_{i'1}$ for all i, i', since the columns of M may be permuted without affecting the above equality of row sums. Now consider the $J+1$ submatrices formed from the first $J+1$ columns of M by omitting the sth column ($s = 1, \cdots, J+1$). Express the above equality as $c_i - \mu_{is} = c_{i'} - \mu_{i's}$, where $c_i = \sum_{j=1}^{J+1} \mu_{ij}$, and sum from $s = 1$ to J.]

Scheffé, H. 1959. The analysis of Variance. John Wiley & Sons, NY. 477 pp.

BM

Brian N. Tissot

CHAPTER 10

The Effects of Departures from the Underlying Assumptions

10.1. INTRODUCTION

In this chapter we shall study the effects of violations of the following assumptions made elsewhere in various parts of this book: (i) normality of the errors, and also normality of the random effects in the models where these appear (Chs. 7 and 8), (ii) equality of variance of the errors, and (iii) statistical independence of the errors. A study of this kind cannot be exhaustive, for one reason, because assumptions like this can be violated in many more ways than they can be satisfied. Usually we shall treat the three kinds of violations one at a time. We shall not be able to treat all the basic designs under all the models we have considered. Some of our conclusions will have to be inductions from rather small numerical tables. However, we must tackle the important questions raised in this chapter even though the evidence is incomplete and we realize that standards of rigor possible in deducing a mathematical theory from certain assumptions generally cannot be maintained in deriving the consequences of departures from those assumptions.

In considering the effects of nonnormality we will find it convenient to use the measures[1] γ_1 of skewness and γ_2 of kurtosis of the distribution of a random variable x. If the mean and variance of the distribution are denoted by μ and σ^2, respectively, its skewness γ_1 is defined as

$$\gamma_1 = \sigma^{-3} E[(x-\mu)^3],$$

and its kurtosis γ_2 as

$$\gamma_2 = \sigma^{-4} E[(x-\mu)^4] - 3.$$

[1] Other commonly used measures are $\beta_1 = \gamma_1^2$ for the magnitude of the skewness, and $\beta_2 = \gamma_2 + 3$ for the kurtosis. To preserve the sign of the skewness in the β-system, γ_1 is usually denoted by $\sqrt{\beta_1}$, meaning $\pm\sqrt{\beta_1}$ with the sign of γ_1. We assume for all distributions whose skewness and kurtosis is considered in this chapter that the variance is positive and that the fourth moment (and hence the variance also) is finite.