A NOTE ON SENSITIVITY OF THE ANOVA TESTS WHEN OBSERVATIONS ARE SUBJECT TO MEASUREMENT ERROR

By

S. Kakoty

Dibrugarh University, Assam

(Received : May, 1980)

Summary

The effect of measurement error on the power function of the ANOVA test is investigated. Certain modifications in the usual test procedures have also been suggested so as to increase the power of the test.

Introduction

The well-known parametric model for testing group differences in one-way classification with one observation per cell is given by

$$x_{ij}=\mu+\alpha_i+\epsilon_{ij} \tag{1.1}$$

$$i=1, 2,\ldots\ldots\ldots, k; \qquad j=1, 2,\ldots\ldots\ldots, n,$$

where μ is the general mean, α_i is the deviation of the mean of the *i-th* group from the general mean ϵ_{ij}'s are the random residuals distributed independently as $N(0, \sigma)$ and x_{ij}'s are the true measurements. The hypothesis to be tested is $H_o : \alpha_i=0$ for all i. In practice, we get observations y_{ij}'s, which are subject to measurement error. Let

$$y_{ij}=x_{ij}+d_{ij} \tag{1.2}$$

where d_{ij} denotes measurement error. It is assumed, for simplicity, that d_{ij}'s are independent and normally distributed with a variance σ_d^2. It is clear, under these assumptions, that the measurement error increases the error variance σ^2 to $\sigma^2+\sigma_d^2$ and thus decreases the power of the test. The power function for testing H_o against $H_1 : \alpha_i\neq 0$ (for some i) is given by

$$1-\beta=\int_{F_\alpha}^{\infty} f_{\nu_1, \nu_2}(F' \mid \lambda')\, dF', \tag{1.3}$$

where F' denotes non-central F with $\nu_1=k-1$, $\nu_2=k(n-1)$ degrees of freedom,

$$\lambda'=n \sum_i \alpha_i^2/\sigma^2(1+\sigma_d^2/\sigma^2)^{-1}=\lambda\ (1+\sigma_d^2/\sigma^2)^{-1}, \qquad (1.4)$$

F_α being the upper 100 α% point of the central F – distribution. The extent to which the observed power given by (1.3) deviates from the true power is seen from Table 1 constructed for $\alpha=0.05$, $\nu_1=4$, $\nu_2=30$, $\sigma_d^2/\sigma^2=0.25$, 0.50, 0.75 and a few selected values of

$$\phi=\sqrt{\lambda}/(\nu_1+1).$$

TABLE 1

True and observed powers for ANOVA test—Parametric model, one observation per experimental unit,
$\alpha=0.05$, $\nu_1=4$, $\nu_2=30$.

ϕ	*True Power*	*Observed Power* σ_d^2/σ^2		
		0.25	0.50	0.75
0.0	0.05	0.05	0.05	0.05
0.5	0.1101	0.0956	0.0862	0.0796
1.0	0.3421	0.2790	0.2373	0.2077
1.2	0.4822	0.3931	0.3279	0.2862
1.4	0.6282	0.5197	0.4417	0.3823
1.6	0.7598	0.6496	0.5617	0.4891
1.8	0.8619	0.7649	0.6757	0.6006
2.0	0.9300	0.8562	0.7787	0.7054
2.2	0.9681	0.9210	0.8600	0.7952
2.6	0.9960	0.9838	0.9570	0.9204
3.0	0.9997	0.9988	0.9931	0.9774

2. Some Modifications in the Test Procedure

The test will be more powerful if we can separate out the measurement error variance σ_d^2 from σ^2. The models of type (1.2) donot allow for the separation of these two components. The simplest procedure to do the separation would be to consider repeated measurements on each of the experimental units. This leads to constructing a model

$$y_{ijl}=\mu+\alpha_i+E_{ij}+d_{ijl} \qquad \ldots(2.1)$$

$$i=1, 2,\ldots, k; j=1, 2,\ldots, n; l=1, 2,\ldots, r.$$

The ANOVA under this model is shown in Table 2.

TABLE 2

Source	*d.f.*	*S.S.*	*M.S.*	*E(M.S.)*
Between Groups	$k-1$	$nr\sum_i(\bar{y}_{i\ldots}-\bar{y}_{\ldots})^2$	s^2_{by}	$nr\sum_i \alpha^2/(k-1)+r\sigma^2+\sigma^2_d$
Between units within Groups	$k(n-1)$	$n\sum_i\sum_j(\bar{y}_{ij}-\bar{y}_{\ldots})^2$	s^2	$r\sigma^2+\sigma^2_d$
Between Measurements	$kn(r-1)$	$\sum_i\sum_j\sum_l(y_{ijl}-\bar{y}_{ij.})^2$	s^2	σ^2_d
Total	$knr-1$	$\sum_i\sum_j\sum_l(y_{ijl}-\bar{y}_{\ldots})^2$		

For testing the group differences, we employ the test criterion reject H_0, if

$$F=\frac{s^2_{by}}{s^2_{wy}}\geqslant F_{\alpha;\,\nu_1,\,\nu_2}. \qquad \ldots(2.2)$$

To calculate power of this test, we see that under the alternative hypothesis H_1, the ratio s^2_{by}/s^2_w follows non-central F-distribution with d.f. $\nu_1=k-1$, $\nu_2=k(n-1)$ and non-centrality parameter

$$\lambda'=n\sum_i\alpha_i^2/\sigma^2[1+\sigma^2_d/(r\sigma^2)]^{-1} \qquad \ldots(2.3)$$

If there were no measurement error in the observations, the ratio would follow, under H_1, non-central F with non-centrality parameter

$$\lambda=n\sum_i\alpha_i^2/\sigma^2$$

Hence $\gamma'\simeq\lambda$ for large r and thus the power of the test with fallible observations will be almost equal to that of the true test.

The best recourse would be to separate out σ_d^2 from the observed variances and then apply the usual test criterion. We find from Table 2 that the estimates of σ^2 and σ_d^2 are given by

$$\hat{\sigma}^2 = (s_w^2 - s_d^2)/r, \quad \hat{\sigma}_d^2 = s_d^2 \qquad \text{...(2.4)}$$

So, it will be worthwhile to subtract s_d^2 from both numerator and denominator of F in (2.2) that is, use the test statistic

$$\bar{F} = \frac{s_{by}^2 - s_d^2}{s_w^2 - s_d^2}.$$

In this case both numerator and denominator of $\bar{F}$ would provide unbiased estimate of σ^2 under the null hypothesis and also the statistic $\bar{F}$ becomes free from measurement error. However, this $\bar{F}$ will not have the F-distribution.

It can be shown that

$$\bar{F} = \frac{C_1 F_1 - 1}{C_2 F_2 - 1} \qquad \text{...(2.5)}$$

where

$$C_1 = \frac{kn(r-1)}{k-1}(1 + r\sigma_d^2/\sigma^2)$$

$$C_2 = \frac{n(r-1)}{k-1}(1 + r\sigma_d^2/\sigma^2)$$

$$F_1 = F_{\{k-1,\, kn(r-1)\}}$$

$$F_2 = F_{\{(kn-1),\, kcn(r-1)\}}$$

Following Morrison (1971), we can find

$$\text{Prob}\left\{\frac{C_1 F_1 - 1}{C_2 F_2 - 1} \geqslant k\right\} = \text{Prob}\{C_1 F_1 - C_2 k F_2 \geqslant 1 - k\} \qquad \text{...(2.6)}$$

ACKNOWLEDGEMENT

The author wishes to thank the referees for their valuable comments.

REFERENCE

Morrison, D.F. (1971) : The distribution of linear functions of independent F-variates. *Jour. Amer. Stat. Assoc.*, **66** 383-385.