

## ASYMPTOTIC QUADRATIC ESTIMATORS IN THE RANDOM, ONE-WAY ANOVA

المقدرات المحاذية من الدرجة الثانية للنموذج العشوائي لتحليل التباين بطريق واحد

د. عبد الجبار شهاب احمد البرهاوي

استاذ مساعد

قسم الاحصاء والمعلوماتية

كلية علوم الحاسبات والرياضيات

### المخلص

يتعرض هذا البحث إلى الاستقرار العشوائي لبعض المعدلات التي تستعمل لحساب مجموع المربعات في تحليل التباين. وقد استعملت هذه المعدلات لإيجاد مقدرات غير سالبة لكون التباين استعمل مقدر معدل مربعات الخطأ كمقدر لمكون تباين الخطأ. أعطي أيضا مقدرًا لمعامل التوريث. كما اشتق الاختبار المحاذي لمكون التباين وأخيرًا أخذت هذه المقالة بنظر الاعتبار كون عدد الملاحظات في كل خلية متغيرًا عشوائيًا.

### SUMMARY

The paper considers stochastic convergence of certain means used to obtain the between sum of squares in analysis of variance. These limiting random variables are used to obtain a nonnegative estimator of the between component of variance. The usual ANOVA estimator of the within component of variance is considered. A nonnegative estimator of heritability is given. Asymptotic tests are derived also. Finally, the paper extends the linear model to allow the number of observations in each cell to be random.

### Keywords:

Asymptotic estimators. Asymptotic tests. Random model with random numbers of observation in cells.

## ASYMPTOTIC QUADRATIC ESTIMATORS IN THE RANDOM, ONE-WAY ANOVA

### Introduction

Asymptotic methods are frequently used to approximate level of significance of tests, bias and mean square error of estimators. Components of variance test and estimation are complicated for the unbalanced case. Asymptotic methods are appropriate tools to derive large sample results. We consider the usual one – way analysis of variance random effect model

$$y_{ij} = \mu + a_i + e_{ij}$$

$$i = 1, 2, \dots, k,$$

$$j = 1, 2, \dots, n_i,$$

$$E(a_i) = 0, \quad v(a_i) = \sigma^2 \quad \text{for all } i.$$

$$E(e_{ij}) = 0, \quad v(e_{ij}) = \sigma^2 \quad \text{for all } i \text{ and } j$$

The random variables  $a_i$  and  $e_{ij}$  ( $i = 1, \dots, k, j = 1, \dots, n_i$ ) are independent.

This paper gives an estimator of  $\sigma^2$  that is asymptotically unbiased and derives a test for  $H_0: \sigma^2 = \sigma_0^2$ .

This paper also consider estimation of the components of variance when  $n_i$  ( $i = 1, 2, \dots, k$ ) are random variables.

The asymptotic properties of  $\bar{y}_i = \sum_j y_{ij} / n_i \quad i = 1, \dots, k$  play central role in deriving the estimator and the test of  $\sigma^2$ . It is easy to see that

$$E(y_i) = \mu \quad i = 1, 2, \dots, k,$$

$$V(y_i) = \sigma^2 1 + \sigma^2 / n_i = \sigma^2 1 (1 + \gamma / n_i) \quad i = 1, 2, \dots, k$$

### Large Sample Estimation and Test

Let

$$a_i \sim N(0, \sigma^2 1)$$

$$e_{ij} \sim N(0, \sigma^2)$$

$$i = 1, 2, \dots, k \quad ; \quad j = 1, 2, \dots, n_i$$

Then

$$y_i \sim N(\mu, \sigma^2 1 (1 + \gamma / n_i))$$

$$i = 1, \dots, k$$

It is obvious that

$$y_i \xrightarrow{d} N(\mu, \sigma^2 1) \quad i = 1, 2, \dots, k$$

The error committed when considering

$$y_i \sim N(\mu, \sigma^2 1) \quad i = 1, 2, \dots, k$$

Is

$$Er = \int_a^b \phi(z) dz$$

Where

$$a = |y - \mu| / \sqrt{\sigma^2 1 (1 + \gamma / n_i)}$$

$$b = |y - \mu| / \sqrt{\sigma^2 1}$$

$$\phi(z) \text{ is the density } N(0, 1)$$

This error is a decreasing function of  $n_i$  and goes to zero as  $n_i$  goes to infinity.

Since

$$u = \sum_i (y_i - y_{..})^2 / \sigma^2 1$$

Is continuous function of

$$y_1, y_2, \dots, y_k,$$

It follows that (van der vaart (1998) )

$$u \rightarrow \chi_{k-1}^2,$$

And

$$\bar{\sigma}^2_1 = \sum_i (y_{i.} - y_{..})^2 / (k - 1)$$

Is asymptotically unbiased estimator of  $\sigma^2_1$ . This estimator is always non-negative.

$$The\ statistic\ V = \sum_i (y_{i.} - y_{..})^2 / \sigma^2_0$$

Is asymptotically  $\chi_{k-1}^2$  under  $H_0 : \sigma^2_1 = \sigma^2_0$  and can be used to test  $H_0 : \sigma^2_1 = \sigma^2_0$ . This result justifies the use of unweighted means in analysis of variance.

This procedure is applicable for balanced data when the number of observations is large.

A non-negative estimator of heritability is

$$H = \frac{\sigma^2_1}{\sigma^2_1 + EMS}$$

We can see that

$$E [\mu + ai + \bar{e}_i - \mu - ai]^2 \rightarrow 0$$

As  $n_i \rightarrow \infty$  with the sole assumption that  $V(\bar{e}_i) \rightarrow 0$  as  $n_i \rightarrow \infty$  ( $i = 1, 2, \dots, k$ ) then  $y_{i.}$  converges in quadratic mean to  $\mu + ai$  ( $i = 1, 2, \dots, k$ ). This is in close agreement with the previous result and our estimator is distribution free for large  $n_i$  ( $i = 1, 2, \dots, k$ ) in the following sense. If  $ai$  has any distribution that has second moment, then our estimator estimates the variance of this distribution unbiasedly for large  $n_i$  ( $i = 1, 2, \dots, k$ ).

Since convergence in quadratic mean implies convergence in distribution we see that

$$F_{n_i}(x) \xrightarrow{w} F(x)$$

Where  $F_{n_i}(x)$  is the distribution of  $y_{i.}$  and  $F(x)$  is the distribution of  $\mu + ai$ .

The estimator of  $\sigma^2$  is the error mean squares of the analysis of variance

$$EMS = \sum_j \sum_i (y_{ij} - y_{i.})^2 / \sum_i (n_i - 1)$$

This is known to the consistent estimator of  $\mu^2$  and its distribution is independent of the distribution of  $a_i$  ( $i = 1, 2, \dots, k$ ).

### The Case of Random Number of Observation

In this section we assume that the number of observations  $n_i$  ( $i = 1, 2, \dots, k$ ) are *iid* as geometric distribution

$$F(n_i) = q^{n_i-1} p \quad n_i = 1, 2, 3, \dots$$

We take the parameter  $p = 0.5$  to avoid the problem of estimating it. Let the normalized eigenvectors of the matrix  $I - \frac{1}{k}J$  be  $W_0, W_1, W_2, \dots, W_{k-1}$  ; where  $W_0' = (\frac{1}{\sqrt{k}}, \frac{1}{\sqrt{k}}, \dots, \frac{1}{\sqrt{k}})$ . And let  $T' = (T_1, T_2, \dots, T_k)$  ; where  $T_i = \sum_j y_{ij}$   $i = 1, 2, \dots, k$  .

Ower estimator in this case is based on

$$\sum (T_i - \bar{T})^2 = \sum (W_i' T)^2$$

This estimator obviously is independent of the choice of

$$W_1, W_2, \dots, W_{k-1}.$$

$$E \sum_i (T_i - \bar{T})^2 = 2k(\sigma^2 + 3\sigma^2 \mathbf{1})$$

The unbiased estimator of  $\sigma^2 \mathbf{1}$  is

$$V = \frac{E(T_i - \bar{T})^2 - EMS}{3}$$

### References

- van der Vaart, A. W. (1998). Asymptotic statistics. Cambridge University Press, Cambridge.