ANOVA models for Brownian motion

Gordon Hazen†‡
Daniel Apley†
Neehar Parikh‡

†Department of Industrial Engineering and Management Sciences
‡Northwestern University Transplant Outcomes Research Collaborative

Northwestern University
December 2013

Corresponding Author:
Gordon Hazen
IEMS Department, McCormick School of Engineering
Northwestern University
Evanston IL 60208-3119

Running Head: ANOVA for Brownian Motion
Abstract

We investigate longitudinal models having Brownian-motion covariance structure. We show that any such model can be viewed as arising from a related “timeless” classical linear model where sample sizes correspond to longitudinal observation times. This relationship is of practical impact when there are closed-form ANOVA tables for the related classical model. Such tables can be directly transformed into the analogous tables for the original longitudinal model. We in particular provide complete results for one-way fixed and random effects ANOVA on the drift parameter in Brownian motion, and illustrate its use in estimating heterogeneity in tumor growth rates.

KEY WORDS: LONGITUDINAL MODELS; BROWNIAN MOTION; ANALYSIS OF VARIANCE; ANOVA; RANDOM EFFECTS; STOCHASTIC HETEROGENEOUS TUMOR GROWTH.

1. Introduction

The statistical analysis of longitudinal data has a long history (e.g., (Diggle, 1988; Diggle, Heagerty, Liang, & Zeger, 2002; Singer & Willett, 2003; Verbeke & Molenberghs, 2009)). Consider the situation in which repeated measurements $x_{ij}$ on individual $i$ are taken at times $0 = t_0 < t_1 < \ldots < t_M$. We consider here a particular linear model with parameters $\alpha_1, \ldots, \alpha_M, \sigma^2$:

$$x_y = x_{i0} + (u_i \alpha_1 + \cdots + u_M \alpha_M) t_y + \epsilon_y \quad \epsilon_y \sim \text{normal}(0, \sigma^2 t_y).$$

where $u_i, \ldots, u_M$ are specified constants or observed covariates for individual $i$. We allow the index $i$ to be lexicographically nested, so that (1) includes, among many possible examples, the additively separable model with $i = (g,h)$:

$$x_{gh} = x_{gh0} + (\nu + \kappa_g + \lambda_h) t_{gh} + \epsilon_{gh}$$

in which $(\alpha_1, \ldots, \alpha_M) = (\nu, \kappa_1, \ldots, \kappa_g, \lambda_1, \ldots, \lambda_H)$. In (1) we restrict ourselves to the specific correlation structure

$$\text{Corr}(\epsilon_j, \epsilon_{j'}) = \sqrt{t_j/t_{j'}} \quad \text{if } j \leq j'$$

that arises when observations come from a Brownian motion process (e.g. (Karlin & Taylor, 1975; Oksendal, 1998)) with drift $\mu_i = u_i \alpha_1 + \cdots + u_M \alpha_M$ and volatility $\sigma^2$. 
Random effects versions of model (1) are important when measurements are available from multiple individuals (e.g., cancer growth data from multiple patients – see section 5). In this case, one might wonder whether the population of individuals is homogeneous – all sharing the same growth rate coefficients $\alpha_1, \ldots, \alpha_M$ – or heterogeneous. In the latter case, one might suspect that individual growth rate coefficients are sampled from a hypothetical population (e.g., slow- versus fast-growing tumors). Such models are attractive because they allow an analyst to forecast adaptively based on an individual’s history. For instance, in a contribution not involving repeated measurements, Ayer et al (Ayer, Alagoz, & Stout, 2012) examine personalizing protocols for breast cancer screening, in which healthy screening outcomes for an individual yield an adaptive forecast of reduced cancer incidence. This allows avoidance of unnecessary screening, which would reduce costs and false-positive morbidities without impacting survival.

In this paper, we show that any model of the form (1)-(2) can be viewed as arising from a related “timeless” classical linear model in which sample sizes correspond to observation time increments $\Delta t_y = t_{i+1} - t_i$. Moreover, inferences from the related classical model apply directly to model (1)-(2), after modifying the degrees of freedom in the error sum of squares. When the classical linear model possesses a closed-form sum-of-squares table, this table converts directly into a sum-of-squares table for the desired model (1)-(2). In particular, we show that the special case of one way Brownian ANOVA, in which the parameters are the patient drifts themselves, arises from a classical one-way ANOVA that is simple enough to solve by hand or on a spreadsheet. Recasting the analysis in a familiar ANOVA format can provide greater insight than the “black box” results from statistical software. Any inference procedure for Brownian motion would also apply to processes such as geometric Brownian motion obtained by transformation (see section 5), since inference could occur on transformed data. We believe the practical import of this paper will be for analysts examining longitudinal data, but with limited access to or experience with statistical software or general linear models, who wish to quickly obtain ANOVA results in a familiar format without acquiring software or studying general linear models.

We structure the paper as follows. In Section 2 we transform to a general linear model with constant error variances, and derive closed-form estimates and tests. In Section 3 we specialize these general results to one-way ANOVA on a Brownian drift parameter, and also consider estimation and testing for a random-effects versions of this model. In Section 4, we examine the connection described above between model (1)-(2) and related “timeless” classical linear models. Section 5 provides a one-way Brownian ANOVA on tumor growth data. Section 6 concludes.
2. ANOVA for the general longitudinal model

In model (1)-(2), let
\[ \Delta x_{ij} = x_{ij} - x_{i,j-1} \quad \Delta t_{ij} = t_{ij} - t_{i,j-1} \quad j = 1, \ldots, J. \]

Then we have
\[ \Delta x_{ij} = \mu_t \Delta t_{ij} + \Delta \varepsilon_{ij} \quad \Delta \varepsilon_{ij} \sim \text{independent normal}(0, \sigma^2 \Delta t_{ij}) \] (3)

where \( \mu_t = u_t \alpha_t + \cdots + u_{a_t} \alpha_{a_t} \) and there are a total of \( J = \sum_i J_i \) observations. We rewrite the equation for the drift vector \( \mu = (\mu_1, \ldots, \mu_J) \) in terms of the parameter vector \( \alpha = (\alpha_1, \ldots, \alpha_{a_t}) \):
\[ \mu = U \alpha \] (4)

Notation

We consider both fixed effects and random effects models. We use the notation
\[ \Delta x_i = \sum_j \Delta x_{ij} = x_{ij} - x_{i1} \quad \Delta t_i = \sum_j \Delta t_{ij} \]
\[ \Delta x_\gamma = \sum_j \Delta x_{ij} = x_{ij} - x_{i1} \quad \Delta t_\gamma = \sum_j \Delta t_{ij} \]

The material below uses the following matrix-vector notation. The prime symbol \( ' \) denotes matrix transpose. \( I \) will denote an identity matrix of dimension \( J \). If \( A_1, \ldots, A_I \) are matrices, then \( \text{diag}(A_i) \) is the block diagonal matrix with blocks \( A_i \). If the \( A_i \) have the same number of columns, then \( \text{stack}(A_i) \) is the matrix \( \begin{bmatrix} A_1' & \cdots & A_I' \end{bmatrix} \) obtained by stacking \( A_1 \) onto \( A_2 \) onto \cdots onto \( A_I \). The following identity will be useful:
\[ \text{diag}(A_i) \cdot \text{stack}(B_i) = \text{stack}(A_i B_i) \]
if \( A_i, B_i \) are compatible under matrix multiplication.

The basic model

We treat both fixed and random effects, but initially do not distinguish which case we are considering. With the transformation
\[ y_{ij} = \frac{\Delta x_{ij}}{\Delta t_{ij}^{1/2}} \]
the basic model (3) becomes
\[ y_{ij} = \mu_t \Delta t_{ij}^{1/2} + e_{ij} \quad e_{ij} \sim \text{independent normal}(0, \sigma^2) \]

In vector notation, this is
\[ y = \text{diag}(\Delta t_{ij}^{1/2}) \mu + e \quad e \sim \text{normal}(0, \sigma^2 I_J) \] (5)

where
The resulting model is:

$$y = \text{diag}(\Delta t_{ij}) U \alpha + e.$$  \hspace{1cm} (6)

This is a special case of the linear model $y = X \alpha + e$ in which $X = \text{diag}(\Delta t_{ij}) U$. \hspace{1cm} (7)

In the following, we explore what happens to ANOVA for this linear model.

**Error sum of squares**

The least-squares estimates $\hat{\alpha}$ are the unrestricted values of $\alpha$ that minimize the sum of squares $S(\alpha)$ subject to $\mu = U \alpha$, where

$$S(\alpha) = \| y - X \alpha \|^2 = \| y - \text{diag}(t_{ij}) U \alpha \|^2 = \| y - \text{diag}(t_{ij}) \mu \|^2 = \| \text{stack}(y - \Delta t_{ij} \mu) \|^2$$

$$= \sum_i \sum_j (y_{ij} - \Delta t_{ij} \mu_j)^2 = \sum_i \sum_j \Delta t_{ij} \left( \frac{\Delta y_{ij}}{\Delta t_{ij}} - \mu_j \right)^2.$$  \hspace{1cm} (8)

Let $\hat{\mu} = U \hat{\alpha}$ The resulting error sum of squares is given by

$$SS_e = \| y - X \hat{\alpha} \|^2 = \sum_i \sum_j \Delta t_{ij} \left( \frac{\Delta y_{ij}}{\Delta t_{ij}} - \hat{\mu}_j \right)^2.$$  \hspace{1cm} (9)

$SS_e$ has $J - r$ degrees of freedom, where $r$ is the rank of $X = \text{diag}(\Delta t_{ij}) U$. The latter is equal to the rank of $U$ because if $u_i$ is the $i^{th}$ row of $U$, then $X = \text{diag}(\Delta t_{ij}) U = \text{stack}(\Delta t_{ij} u_i)$. Assuming each $u_i$ is nonzero, then each element $\Delta t_{ij} u_i$ of this stack $X$ is an array of rank 1 with rows proportional to $u_i$. If $r'$ is the rank of $U$, then there are $r'$ linearly independent rows $u_i$, and $X$ must have rank $r = r'$.

**Numerator sum of squares**

Suppose the null hypothesis to test is $H: \alpha \in \omega$ (against an alternative hypotheses with no restrictions on $\alpha$) and consists of $q$ independent linear restrictions on $\alpha$. Let $\hat{\alpha}_o$ be the least squares estimate under $H$, and let $\hat{\mu}_o = U \hat{\alpha}_o$.

The numerator sum of squares for an $F$-test of $H$ is

$$SS_o = \| X \hat{\alpha} - X \hat{\alpha}_o \|^2 = \| \text{diag}(\Delta t_{ij}) U (\hat{\alpha} - \hat{\alpha}_o) \|^2 = \| \text{diag}(\Delta t_{ij}) (\hat{\mu} - \hat{\mu}_o) \|^2.$$  \hspace{1cm} (10)

$SS_o$ has $q$ degrees of freedom.
For the fixed-effects scenario, the expectation of the mean square $MS_\mu = SS_\mu / q$ can be obtained from Scheffe’s Rule 2 ((Scheffe, 1959), p.39), and is given by

$$E[MS_\mu] = \sigma^2 + q^{-1} \sum_i \Delta t_i (\mu_i - \bar{\mu}_i)^2$$

(11)

where $\bar{\mu}_i$ is obtained from the formula for $\hat{\mu}_i$ by replacing each term $y_{ij}$ by its mean $1/2 \hat{t}_{ij}$ and $\Delta t_i$ by $\Delta t_i - \Delta \mu_i$. This is equivalent to replacing each term $\Delta x_{ij}$ by $\Delta t_{ij} - \Delta \mu_i$.

F-Test

$SS_E$ and $SS_\mu$ are independent, with $SS_E/\sigma^2$ having a chi-square distribution with the appropriate degrees of freedom, and $SS_\mu/\sigma^2$ being noncentral chi-square in general, and chi-square under $\omega$. So the usual F-test of the hypothesis $H$ applies. The noncentrality parameter $\delta$ for $SS_\mu/\sigma^2$ may be obtained from Scheffe’s Rule 1 ((Scheffe, 1959), p.39): If in $SS_\mu$ each observation $y_{ij}$ is replaced by its expected value $1/2 \hat{t}_{ij}$ (or $\Delta x_{ij}$ by $\Delta t_{ij} - \Delta \mu_i$), the result is $\sigma^2 \delta^2$.

Random effects

In a random effects model, we would have $\alpha \sim \text{normal}(\alpha_0, \Sigma_{\alpha})$. (If only some terms of $\alpha$ are random, this would be reflected by zeroes in $\Sigma_{\alpha}$.) The distributional and mean claims above remain conditionally true given $\alpha$. In particular, given $\alpha$, the quantity $SS_\mu/\sigma^2$ is conditionally $\chi^2(J-r)$ – a distribution that does not depend on $\alpha$.

Therefore, $SS_\mu/\sigma^2$ is unconditionally $\chi^2(J-r)$.

The random-effects counterpart to the fixed-effects null hypothesis $H: \alpha \in \omega$ is a null hypotheses $H_\alpha$ that places certain restrictions on $\Sigma_{\alpha}$. Namely, $H_\alpha$ restricts the variances of the $q$ independent linear contrasts of $\alpha$ corresponding to $\alpha \in \omega$ to be zero. Then under $H_\alpha$, $SS_\mu/\sigma^2$ and $SS_\mu/\sigma^2$ are conditionally independent given $\alpha$ with $\chi^2(J-r)$ and $\chi^2(q)$ distributions not depending on $\alpha$, so are therefore independent $\chi^2(J-r)$ and $\chi^2(q)$ variables under $H_\alpha$. The $F$ statistic therefore still has an $F$-distribution under $H_\alpha$, and the same $F$-test applies.

3. Special Case of One-way Brownian ANOVA

As we have noted, our framework encompasses one-way or multi-way ANOVA models, nested or crossed, with or without covariates, etc. In this section, we specialize to the simplest version of Brownian ANOVA, a one-way model, which we define as the general model (1)-(2) with the $I$ drifts $\mu_1, \ldots, \mu_I$ the parameters of interest. In the framework of the general model, for the one-way ANOVA special case we have $\alpha = \mu$ and $U = 1_I$ in (4), yielding $X$
= \text{diag}(\Delta t^{ij})$. The general results from the prior section lead directly to many of the assertions summarized in Table 1 below. The fixed-effects assertions in particular fall out in a straightforward manner when the least-squares estimates

$$
\hat{\mu}_i = \frac{\Delta x_i}{\Delta t_i}, \quad \hat{\mu}_{\omega} = \frac{\Delta x_\omega}{\Delta t_\omega}
$$

are substituted. Hence, in the next subsection, we focus attention on deriving the random-effects assertions in Table 1.

### Table 1. One-way fixed and random effects Brownian ANOVA

<table>
<thead>
<tr>
<th></th>
<th>Fixed effects</th>
<th>Random effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model parameters</td>
<td>$\mu_1, \ldots, \mu_I$</td>
<td>$\mu_1, \ldots, \mu_I$ independent normal($\nu, \sigma^2_\mu$)</td>
</tr>
<tr>
<td>Hypothesis</td>
<td>$H: \mu_1 = \ldots = \mu_I$</td>
<td>$H: \sigma^2_\mu = 0$</td>
</tr>
<tr>
<td>Error sum of squares</td>
<td>$SS_e = \sum_i \sum_j \Delta t_i \left( \frac{\Delta x_i}{\Delta t_i} - \frac{\Delta x_\omega}{\Delta t_\omega} \right)^2$ with $J-I$ degrees of freedom</td>
<td>$MS_e = \frac{SS_e}{J-I}$</td>
</tr>
<tr>
<td>Hypothesis sum of squares</td>
<td>$SS_\mu = \sum_i \Delta t_i \left( \frac{\Delta x_i}{\Delta t_i} - \frac{\Delta x_\omega}{\Delta t_\omega} \right)^2$ with $I-1$ degrees of freedom</td>
<td>$MS_\mu = \frac{SS_\mu}{I-1}$</td>
</tr>
<tr>
<td>Expected mean squares</td>
<td>$E[MS_e] = \sigma^2$</td>
<td>$E[MS_\mu] = \sigma^2 + \Delta t \cdot \sigma^2_\mu$.</td>
</tr>
<tr>
<td></td>
<td>$E[MS_\mu] = \sigma^2 + (I-1)^{-1} \sum_j \Delta t_i (\mu_i - \bar{\mu})^2$</td>
<td>$\bar{\Delta t} = (I-1)^{-1} \left( \Delta t_\mu - \Delta t^{-1} \sum_i \Delta t_i \right)$.</td>
</tr>
<tr>
<td>$F$-test</td>
<td>Reject $H$ at significance level $\alpha$ if $\frac{MS_\mu}{MS_e} &gt; F_{\alpha, I-1, J-1}$</td>
<td></td>
</tr>
<tr>
<td>Parameter estimates</td>
<td>$\hat{\sigma}^2 = MS_e$</td>
<td>$\hat{\sigma}^2_\mu = \Delta t^{-1} (MS_\mu - MS_e)$</td>
</tr>
<tr>
<td></td>
<td>$\hat{\mu}_i = \frac{\Delta x_i}{\Delta t_i}$</td>
<td>$\hat{\sigma}^2_\mu = \Delta t^{-1} (MS_\mu - MS_e)$</td>
</tr>
<tr>
<td></td>
<td>$\hat{\mu}<em>i \sim \text{normal}(\mu_i, \sigma^2</em>\mu \Delta t_i^{-1})$</td>
<td>$\hat{\sigma}^2_\mu = \Delta t^{-1} (MS_\mu - MS_e)$</td>
</tr>
<tr>
<td></td>
<td>$\text{se}(\hat{\mu}_i) = \frac{\hat{\sigma}^2}{\Delta t_i^{1/2}}$</td>
<td>$\text{se}(\hat{\sigma}^2_\mu) = \frac{\Delta t}{\hat{\sigma}^2_\mu \Delta t_i^{1/2}}$</td>
</tr>
</tbody>
</table>

†In case this quantity is negative, it is common practice to replace the estimate with zero.
A key take-away is that Table 1 is nearly identical to the analogous table for a “timeless” classical one-way ANOVA if one substitutes observation time increments $\Delta t_{ij}$ in the Brownian model for within-cell sample sizes in the classical model. This interesting connection continues to hold in the general model (1)-(2). We elaborate in section 4.

**Derivation of Random effects Results in Table 1**

In the random effects version, we assume the effects are distributed as $\{\mu_i: i = 1, 2, \ldots, I\} \sim$ independent normal($\nu$, $\sigma_{\mu}^2$). Regarding the fixed-effects expectation of the mean square $MS_{\mu}$, we have in equation (11) from the prior section

$$\bar{\mu}_{\mu} = \Delta t^{-1} E[\Delta \mu] = \Delta t^{-1} \sum_i \mu_i \Delta t = \Delta t^{-1} \sum_i \mu_i \Delta t, \quad$$

which we denote by the quantity $\bar{\mu}$ in Table 1, as it does not depend on $i$. In the random effects model, we have from (11) that

$$E[MS_{\mu}] = \sigma^2 + \Delta t \cdot \sigma_{\mu}^2,$$

Using Lemma A2, this becomes

$$E[MS_{\mu}] = \sigma^2 + \bar{\Delta t} \cdot \sigma_{\mu}^2,$$

where

$$\bar{\Delta t} = (I-1)^{-1} \left( \Delta t - \Delta t^{-1} \sum_i \Delta t_i \right),$$

as shown in the table. Here, in parallel to classical random effects ANOVA, when the total observation times $\Delta t_{ij}$ are identical across patients $i$, the quantity $\bar{\Delta t}$ will be their common value.

Combining the last equation with the fact that $MS_e$ is an unbiased estimate of $\sigma^2$, we obtain in the conventional ANOVA moment estimator for $\sigma_{\mu}^2$:

$$\hat{\sigma}_{\mu}^2 = \bar{\Delta t}^{-1} (MS_{\mu} - MS_e).$$

As in classical random effects ANOVA, it is possible that this estimate can be negative, in which case it is common practice to report a zero estimate.

A least-squares estimate of the random effects population mean $\mu$ can be obtained as follows when we treat $\sigma^2$ and $\sigma_{\mu}^2$ as known. Substituting $\mu \sim$ normal($\nu I$, $\sigma^2 I$) into the random effects model $y = \text{diag}(\Delta t_i^{1/2}) \mu + e$ gives

$$y = X \mu + g \quad \text{g} \sim \text{normal}(0, \sigma^2 I + \sigma_{\mu}^2 XX^{'})$$
Defining
\[
W = (\sigma^2 I + \sigma^2_t XX')^{-1/2}X_1,
\]
the weighted least-squares estimate of \( \nu \), treating \( \sigma^2 \) and \( \sigma^2_t \) as known, is
\[
\hat{\nu} = \frac{W'z}{WW} = \frac{I_j'X'(\sigma^2 I_j + \sigma^2_t XX')^{-1/2}(\sigma^2 I_j + \sigma^2_t XX')^{-1/2}y}{I_j'X(\sigma^2 I_j + \sigma^2_t XX')^{-1/2}(\sigma^2 I_j + \sigma^2_t XX')^{-1/2}X_1} = \frac{I_j'X(\sigma^2 I_j + \sigma^2_t XX')^{-1/2}y}{I_j'X(\sigma^2 I_j + \sigma^2_t XX')^{-1/2}X_1}.
\]

The matrix to be inverted here simplifies to
\[
(\sigma^2 I_j + \sigma^2_t XX')^{-1} = \text{diag} \left( (\sigma^2 I_j + \sigma^2_t \Delta^1/2 \Delta^1/2)^{-1} \right)
\]

Use Lemma A3 to write
\[
(\sigma^2 I_j + \sigma^2_t \Delta^1/2 \Delta^1/2)^{-1} = \sigma^{-2} I_j - \frac{\sigma^2}{1 + \sigma^2_t \Delta^1/2} (\sigma^2 I_j)^{-1} \Delta^1/2 \Delta^1/2 (\sigma^2 I_j)^{-1}
\]
\[
= \sigma^{-2} I_j - \frac{\sigma^2}{1 + \sigma^2_t \Delta^1/2} \Delta^1/2
\]
\[
= \sigma^{-2} \left( I_j - \frac{\sigma^2}{\sigma^2 + \sigma^2_t \Delta^1/2} \Delta^1/2 \right).
\]

Substitute this into the denominator of the last equation for \( \hat{\nu} \) to get
\[
\sum_i \Delta_i^1/2 (\sigma^2 I_j + \sigma^2_t \Delta^1/2 \Delta^1/2)^{-1} = \sum_i \Delta_i^1/2 \sigma^{-2} \left( I_j - \frac{\sigma^2}{\sigma^2 + \sigma^2_t \Delta^1/2} \Delta^1/2 \right) \Delta^1/2
\]
\[
= \sigma^{-2} \sum_i \Delta_i - \sum_i \frac{\sigma^2}{\sigma^2 + \sigma^2_t \Delta^1/2} \Delta^1/2 \Delta^1/2
\]
\[
= \sigma^{-2} \sum_i \Delta_i \left( 1 - \frac{\sigma^2 \Delta_i}{\sigma^2 + \sigma^2_t \Delta_i} \right)
\]
\[
= \sigma^{-2} \sum_i \Delta_i \frac{\sigma^2}{\sigma^2 + \sigma^2_t \Delta_i}
\]
\[
= \sum_i \Delta_i \frac{\Delta_i}{\sigma^2 + \sigma^2_t \Delta_i}
\]

and into the numerator to get
\[
\sum_i \Delta t_i^{i/2} (\sigma^2 \mathbf{1}_i + \sigma^2 \mathbf{1}_i \Delta \mathbf{t}^{i/2})^{-1} \mathbf{y}_i = \sum_i \Delta t_i^{i/2} \sigma^{-2} \left( \mathbf{I}_i - \frac{\sigma^2}{\sigma^2 + \sigma^2 \Delta \mathbf{t}_i} \Delta \mathbf{t}_i^{i/2} \right) \mathbf{y}_i.
\]

\[
= \sigma^{-2} \left( \sum_i \Delta \mathbf{y}_i - \sum_i \frac{\sigma^2}{\sigma^2 + \sigma^2 \Delta \mathbf{t}_i} \Delta \mathbf{t}_i \Delta \mathbf{y}_i \right)
\]

\[
= \sigma^{-2} \sum_i \Delta \mathbf{y}_i \left( 1 - \frac{\sigma^2 \Delta \mathbf{t}_i}{\sigma^2 + \sigma^2 \Delta \mathbf{t}_i} \right)
\]

\[
= \sigma^{-2} \sum_i \Delta \mathbf{y}_i \frac{\sigma^2}{\sigma^2 + \sigma^2 \Delta \mathbf{t}_i}
\]

\[
= \sum_i \frac{\Delta \mathbf{y}_i}{\sigma^2 + \sigma^2 \Delta \mathbf{t}_i}.
\]

The resulting quotient is the estimate \(\hat{\mathbf{v}}\) listed in Table 1. To obtain an estimated standard error for \(\hat{\mathbf{v}}\), calculate the numerator variance using independence of the \(\Delta \mathbf{y}_i\):

\[
\text{Var} \left[ \sum_i \frac{\Delta \mathbf{y}_i}{\sigma^2 + \sigma^2 \Delta \mathbf{t}_i} \right] = \sum_i \left( \frac{1}{\sigma^2 + \sigma^2 \Delta \mathbf{t}_i} \right)^2 \text{Var} [\Delta \mathbf{y}_i]
\]

\[
= \sum_i \left( \frac{1}{\sigma^2 + \sigma^2 \Delta \mathbf{t}_i} \right)^2 \left( \text{Var} [E[\Delta \mathbf{y}_i | \mu_i]] + \text{E}[\text{Var}[\Delta \mathbf{y}_i | \mu_i]] \right)
\]

\[
= \sum_i \left( \frac{1}{\sigma^2 + \sigma^2 \Delta \mathbf{t}_i} \right)^2 \left( \text{Var} [\mu_i \Delta \mathbf{t}_i] + \text{E}[\sigma^2 \Delta \mathbf{t}_i] \right)
\]

\[
= \sum_i \left( \frac{1}{\sigma^2 + \sigma^2 \Delta \mathbf{t}_i} \right)^2 \left( \sigma^2 \Delta \mathbf{t}_i^2 + \sigma^2 \Delta \mathbf{t}_i \right)
\]

\[
= \sum_i \frac{\Delta \mathbf{t}_i}{\sigma^2 + \sigma^2 \Delta \mathbf{t}_i}.
\]

Then substitute to obtain \(\hat{\text{se}}(\hat{\mathbf{v}})\) listed in the table.

4. Equivalence to classical ANOVA models

We noted in the prior section the near equivalence between one-way classical and Brownian ANOVAs when observation time is treated as sample size. This equivalence in fact extends to any longitudinal model (1)-(2), that is, to any parameterization \(\mathbf{u} = \mathbf{U} \mathbf{a}\). Such parameterizations include regression and ANCOVA models. In this section, we explain how this connection arises. We restrict ourselves here to fixed-effects models.

For comparison, we rewrite our longitudinal model (3) here:

\[
\Delta \mathbf{y}_i = \mu_k \Delta t_i + \Delta \mathbf{e}_i \quad \Delta \mathbf{e}_i \sim NID(0, \sigma^2 \Delta t_i) \quad i = 1, \ldots, I; \ j = 1, \ldots, J_i, \quad (#)
\]
Consider the related classical ANOVA model
\[ y_{ijk} = \mu_i + \epsilon_{ijk} \quad i = 1, \ldots, I; j = 1, \ldots, Ji; k = 1, \ldots, Kij \] (*)

having the same parameters \( \mu = (\mu_1, \ldots, \mu_I) \) and subject to the same structure \( \mu = U\alpha \). Let \( \hat{\alpha}^\#, \hat{\alpha}^* \) be the least-squares estimates\(^1\) of \( \alpha \) under models (\#) and (*), respectively; let \( \hat{\alpha}_{\omega}^\#, \hat{\alpha}_{\omega}^* \) be the least squares estimates for hypothesis \( H: \alpha \in \omega \) under (\#) and (*), respectively; and let \( SS_{\mu}^\#, SS_{\mu}^* \), \( SS_E^\#, SS_E^* \) be the corresponding sums of squares. Let \( K = \sum_i \sum_j K_{ij} \) be the total number of observations.

In the following it is assumed we can set \( K_{ij} = \Delta t_{ij} \) for all \( i, j \). Of course this requires the \( \Delta t_{ij} \) to be integers, a condition that can be achieved by suitably rescaling time in model (\#)\(^2\).

**Proposition 1:** The least-squares estimates \( \hat{\alpha}^* \) and \( \hat{\alpha}_{\omega}^* \) under model (*) depend on the variables \( y_{ijk} \) only through the sample sums \( y_{ij} \). Whenever \( y_{ij} = \Delta x_j \) and \( K_{ij} = \Delta t_{ij} \) for all \( i, j \), the following relationships hold:

\[
\hat{\alpha}^* = \hat{\alpha}^\#
\]
\[
\hat{\alpha}_{\omega}^* = \hat{\alpha}_{\omega}^\#
\]
\[
SS_{\mu}^* = SS_{\mu}^\#
\]
\[
SS_E^* = \sum_i \sum_j \sum_k (y_{ijk} - \bar{y}_{ij})^2 + SS_E^\#
\]

where \( \bar{y}_{ij} = K_{ij}^{-1} y_{ij} \).

**Proof:** We begin by introducing vector notation for (*):

\[
y_i = \text{stack}_d(\text{stack}_d(y = \text{stack}_k(y))) \\
= \text{stack}_d(\text{stack}_d(\text{stack}_k(\mu)) + \epsilon) \\
= \text{stack}_d(\text{stack}_k(\mu \ 1_{K_i}) + \epsilon)
\]

---

\(^1\) We assume side conditions are imposed to insure uniqueness of the least-squares estimates, the same side conditions in (*) as in (\#). For model (\#), the estimates are the ones discussed in Section 2, that is, the weighted least-squares estimates obtained after transforming the data to have homogeneous variances.

\(^2\) This technically may be not possible if some of the \( \Delta t_{ij} \) are irrational, but in practice data would be expressed to only a finite number of decimal places, hence are effectively rational.
With the parameterization $\mu = U\alpha$, we have

$$y = \text{diag}(1_{K_i})U\alpha + \varepsilon = X^*\alpha + \varepsilon$$

where $X^* = \text{diag}(1_{K_i})U$. Under model (*), the least-square estimates $\hat{\alpha}^*$ minimize the sum of squares $S^*(\alpha)$ subject to $\mu = U\alpha$, where:

$$S^*(\alpha) = \sum_i \sum_j \sum_k (y_{ijk} - \bar{y}_{ij} - \mu_{i})^2$$

So in effect, what must be minimized is $S_2^*(\alpha) = \sum_i \sum_j K_{ij} (\bar{y}_{ij} - \mu_{i})^2$ subject to $\mu = U\alpha$. Of course the resulting $\hat{\alpha}^*$ depends only on the sample sums via the sample means $\bar{y}_{ij}$. Moreover, the quantity $S_2^*(\alpha)$ is identical to the sum of squares $S^#(\alpha)$ of (8) (we have added the superscript # to indicate it is for the model (#)) under the assumed conditions $y_{ij} = \Delta x_{ij}$ and $K_{ij} = \Delta t_{ij}$ for all $i,j$. Therefore the resulting $\hat{\alpha}^*$ must be identical to $\hat{\alpha}^#$ under these conditions. The same argument implies that $\hat{\alpha}^*_{io} = \hat{\alpha}^#_{io}$.

That $SS_E^* = SS_E^#$ follows by noting that $SS_E^* = S^* (\hat{\alpha}^*) - S^* (\hat{\alpha}^c) = S_2^* (\hat{\alpha}^*) - S_2^* (\hat{\alpha}^c)$, which (by the arguments in the preceding paragraph) is the same as $S^# (\hat{\alpha}^*) - S^# (\hat{\alpha}^c) = SS_E^#$ under the assumed conditions $y_{ij} = \Delta x_{ij}$ and $K_{ij} = \Delta t_{ij}$ for all $i,j$. Finally, substituting $\alpha = \hat{\alpha}^*$, we obtain

$$SS_E^* = S^* (\hat{\alpha}^*) = \sum_i \sum_j \sum_k (y_{ijk} - \bar{y}_{ij} - \hat{\mu}_i)^2 = SS_{E,1}^* + SS_{E,2}^*$$

(12) $SS_{E,2}^*$ is $SS_{E}^*$ under the assumed conditions.

Notice in Proposition 1 that if in addition $y_{ijk}$ does not depend on $k$ for all $i,j$, then we have not only $(\hat{\alpha}^*, \hat{\alpha}^*_{io}, SS_E^*) = (\hat{\alpha}^*, \hat{\alpha}^*_{io}, SS_{E}^*)$, but also $SS_{E,2}^* = SS_{E}^*$. Under the assumption $K_{ij} = \Delta t_{ij}$ of the proposition, we have
{ y_{ij} = \Delta x_{ij} \text{, } y_{ijk} \text{ does not depend on } k } \iff y_{ijk} = \Delta x_{ij} / \Delta t_y \quad \text{for all } i,j

Proposition 1 therefore implies that by performing a classical ANOVA on model (*) while making the substitutions

\[ K_{ij} = \Delta t_{ij} \quad y_{ijk} = \Delta x_{ij} / \Delta t_y \]

we can obtain all relevant statistics for the longitudinal model (#). But does the (*)-ANOVA give us the correct statistical inferences for model (#)? It is apparent it cannot, because \( SS_E^*/\sigma^2 \) has a \( \chi^2(K-r) \) distribution, whereas \( SS_E^*/\sigma^2 \) has a \( \chi^2(J-r) \) distribution. To see what has happened, refer to (12) above, in which \( SS_E^* \) is decomposed into the sum of two independent \( \sigma^2 \chi^2 \) quantities \( SS_{E,1}^* \) and \( SS_{E,2}^* \) with \( K-J \) and \( J-r \) respective degrees of freedom. Assuming \( y_{ijk} \) does not depend on \( k \), as in the substitution above, results in \( SS_{E,1}^* = 0 \) and \( SS_{E,2}^* = SS_{E,2}^* \sim \sigma^2 \chi^2(J-r) \).

In general, statistical inferences for the (*)-ANOVA depend on the joint distribution of the quantities \( (\bar{\alpha}, \bar{\alpha}, SS_{\mu}^*, SS_E^*) \), joint because conclusions depend on probabilistic independencies that may exist, such as independence of \( SS_{\mu}^* \) and \( SS_E^* \) needed for hypothesis testing, and independence of \( \bar{\alpha} \) and \( MS_E^* \) needed for confidence interval construction.

One might worry that the restriction that \( y_{ijk} \) does not depend on \( k \) could do more than merely change the distribution of \( SS_E^*/\sigma^2 \) from \( \chi^2(K-r) \) to \( \chi^2(J-r) \) – perhaps it could further alter the joint distribution of \( (\bar{\alpha}, \bar{\alpha}, SS_{\mu}^*, SS_E^*) \) in some way. That this is not so is a consequence of the following result.

**Proposition 2**: The joint distribution of \( (\bar{\alpha}, \bar{\alpha}, SS_{\mu}^*, SS_{E,2}^*) \) is independent of the event \( A = \{ y_{ijk} \text{ does not depend on } k \} \) for all \( i,j \). Moreover, if \( K_y = \Delta t_{ij} \) for all \( i,j \), then this distribution is identical to the joint distribution of \( (\bar{\alpha}^#, \bar{\alpha}^#, SS_{\mu}^#, SS_E^#) \).

**Proof**: To derive the first claim, observe the following: (i) \( (\bar{\alpha}, \bar{\alpha}, SS_{\mu}^*, SS_{E,2}^*) \) depends on \( \{ y_{ijk} \} \) only via the within-sample sums \( \{ y_{ij.} \} \) or, equivalently, via the within-sample averages \( \{ \bar{y}_{ij.} = K_{ij}^{-1} y_{ij.} \} \); (ii) \( SS_{E,1}^* \) is independent of \( \{ \bar{y}_{ij.} \} \) and, therefore, of \( (\bar{\alpha}, \bar{\alpha}, SS_{\mu}^*, SS_{E,2}^*) \), because \( SS_{E,1}^* \) is a function of the within-sample variances \( K_{ij}^{-1} \Sigma_k (y_{ijk} - \bar{y}_{ij.})^2 \), which are independent of the within-sample averages for the normal samples in question; and (iii) the stated
event A is equivalent to $SS^*_{E,1} = 0$. Combining these three results proves that the joint distribution of 

$$(\alpha^*, \overline{\alpha^*}, SS^*_{\mu}, SS^*_{E,2})$$

is independent of the event A, as asserted.

To show the second claim note first that by Proposition 1, when $K_{ij} = \Delta t_{ij}$ for all $i,j$, we have $(\alpha^*, \overline{\alpha^*}, SS^*_{\mu}, SS^*_{E,2}) = g(\Delta x_{ij})$ and $(\alpha^*, \overline{\alpha^*}, SS^*_{\mu}, SS^*_{E,2}) = g(\{y_{ij}\})$ for some function $g(\cdot)$ – the same function in both cases. The second claim then follows from the easily-verified fact that the joint distribution of $\{y_{ij}\}$ under model (*) is the same as the joint distribution of $\{\Delta x_{ij}\}$ under model (#). ♦

What then is the relevance to the longitudinal model (#) of statistical inferences from classical (*)-ANOVA when we make the substitutions $K_{ij} = \Delta t_{ij}$, $y_{ijk} = \Delta x_{ij}/\Delta y$? Step back for a moment and compare two inferential exercises:

- Statistical inference for parameters $\alpha$ under model (*) and the condition $A = \{y_{ijk} \text{ does not depend on } k \text{ for all } i,j\}$. The data on which such inferences are based may be taken, by Proposition 1, to be the sample sums $\{y_{ij}\}$ and sample sizes $\{K_{ij}\}$. Any such inferences depend by Proposition 2 on the same joint distribution $F$ of $(\alpha^*, \overline{\alpha^*}, SS^*_{\mu}, SS^*_{E,2})$ that obtains in the absence of condition A, that is, the joint distribution that obtains in the usual classical ANOVA. The only difference is that under condition A, we have $SS^*_{E,1} = 0$, forcing the error sum of squares $SS^*_{E}$ to equal $SS^*_{E,2}$, and reducing its degrees of freedom from $K-r$ to $J-r$.

- Statistical inference for parameters $\alpha$ under model (#). This is based on data $\{\Delta x_{ij}\}$ and observation times $\{\Delta t_{ij}\}$, and depends on the same (by Proposition 2) joint distribution $F$ of $(\alpha^*, \overline{\alpha^*}, SS^*_{\mu}, SS^*_{E,2})$.

Because of these parallels, inference in these two situations will be identical on identical data $\Delta t_{ij} = K_{ij}$, $\Delta x_{ij} = y_{ij}$.

But as we have mentioned, these equalities are equivalent to $\Delta t_{ij} = K_{ij}$, $y_{ijk} = \Delta x_{ij}/\Delta y$. This leads us to the following.

**Conclusion:** Statistical inferences on $\alpha$ obtained by substituting $y_{ijk} = \Delta x_{ij}/\Delta y$ and $K_{ij} = \Delta t_{ij}$ for all $i,j$ into the classical (*)-ANOVA and changing the degrees of freedom for $SS^*_{E}$ from $K-r$ to $J-r$ will be valid for $\alpha$ in the longitudinal model (#).
Although this conclusion is far-reaching, its practical implications are more limited. If extreme scaling is needed to get all-integer $\Delta t_{ij}$, the resulting large sample sizes $K_{ij}$ and quantities of identical data $y_{ijk}$ for the (*)-ANOVA will be inconvenient for software packages, and the $SS_E$ degrees of freedom in software output would need to be corrected from $K-r$ to $J-r$. So this is not in general an approach we recommend for software packages. However, none of these difficulties arise if there are closed-form tables for ANOVA on model (*). Closed-form tables for unbalanced models like (*) do exist in some cases, for example, for one-way ANOVA as we have already discussed, as well as for nested hierarchical models when effects are weighted by sample size (see (Scheffe, 1959), Ch. 5.3). We discuss the latter in the example presented below. In general, to convert a closed-form classical ANOVA table for model (*) to a corresponding ANOVA table for model (#):

1. Replace the sample sizes $K_{ij}$ by observation times $\Delta t_{ij}$ everywhere.
2. In $SS_E^*$ and $SS_\mu^*$, replace $y_{ijk}$ and $\bar{y}_{ij}$ by $\Delta x_{ij}/\Delta t_{ij}$.
3. Change the degrees of freedom $K-r$ for $SS_E^*$ to $J-r$.

It should be noted that the replacements in Step 2 also entail the following replacements:

$$\bar{y}_{i.} = K_i^{-1} \sum_j K_{ij} \bar{y}_{ij}, \quad \leftrightarrow \quad \Delta t_i^{-1} \sum_j \Delta t_{ij} \frac{\Delta x_{ij}}{\Delta t_{ij}} = \Delta t_i^{-1} \Delta x_{i.} = \Delta \bar{x}_{i.},$$

$$\bar{y}_{..} = K^{-1} \sum_i \sum_j K_{ij} \bar{y}_{ij}, \quad \leftrightarrow \quad \Delta t_{..}^{-1} \sum_i \sum_j \Delta t_{ij} \frac{\Delta x_{ij}}{\Delta t_{ij}} = \Delta t_{..}^{-1} \Delta x_{..} = \Delta \bar{x}_{..}.$$  

Note in step 1 that one can omit any scaling of the $\Delta t_{ij}$ to integer values. This is because any such scaling would cancel from numerator and denominator of all $F$-statistics. Moreover, any scaled $MS_E^*$ would be an unbiased estimate of a scaled $\sigma^2$, hence the unscaled $MS_E^*$ would be unbiased for the unscaled $\sigma^2$.

**Example: A nested model**

The following illustrates how a specific unbalanced nested classical ANOVA translates to a nested longitudinal ANOVA using the methods just discussed. Scheffe ((Scheffe, 1959), Ch. 5.3) introduces a procedure for constructing classical ANOVA tables for unbalanced nested designs. His procedure applies, for example, to the following nested design.
\[ y_{hijk} = \mu_{hi} + \varepsilon_{hijk} \quad \varepsilon_{hijk} \sim \text{NID}(0, \sigma^2) \] (13)

\[ \mu_{hi} = \nu + \kappa_h + \lambda_{hi} \]

\[ h = 1, \ldots, H; \quad i = 1, \ldots, I_h; \quad j = 1, \ldots, J_{hi}; \quad k = 1, \ldots, K_{hijk}. \]

\[ \sum_h I_h = I; \quad \sum_h I_i J_{hi} = J; \quad \sum_h \sum_i \sum_j K_{hijk} = K. \]

To be concrete, this model reflects, for example, the effect \( \kappa_h \) due to hospitals \( h \) and within hospital \( h \) the effect \( \lambda_{hi} \) due to patients \( i \). The indices pair \((h,i)\) in (13) corresponds to the single index \( i \) in (*).

To eliminate parameter redundancy, Scheffe imposes the following restrictions³:

\[ \sum_h K_{hi} \kappa_h = 0 \quad \text{and} \quad \sum_i K_{hi} \lambda_{hi} = 0 \quad \text{for each } h. \]

The least-square estimates and associated \( 1 - \alpha \) confidence intervals are:

\[ \hat{\kappa}_h = \bar{y}_{h..} - \bar{y}_{..} \quad \kappa_h \in \hat{\kappa}_h \pm t_{K-I, \alpha/2} \cdot \hat{\sigma} \sqrt{K_{-\kappa}^{-1} - K_{-\kappa}^{-1}} \] (14)

\[ \hat{\lambda}_{hi} = \bar{y}_{hi..} - \bar{y}_{h..} \quad \lambda_{hi} \in \hat{\lambda}_{hi} \pm t_{K-I, \alpha/2} \cdot \hat{\sigma} \sqrt{K_{-\lambda}^{-1} - K_{-\lambda}^{-1}} \] (15)

where \( \hat{\sigma}^2 = MS_E = SSE/(K-1) \). For testing the two different null hypotheses

\[ H_{\kappa} : \kappa_h = 0 \quad \text{for all } h \]

\[ H_{\lambda} : \lambda_{hi} = 0 \quad \text{for all } h,i \]

the ANOVA table is:

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>d.f.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \kappa )</td>
<td>( SS_{\kappa} = \sum_h K_{hi} (\bar{y}<em>{h..} - \bar{y}</em>{..})^2 )</td>
<td>( H-1 )</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>( SS_{\lambda} = \sum_h \sum_i K_{hi} (\bar{y}<em>{hi..} - \bar{y}</em>{h..})^2 )</td>
<td>( I-H )</td>
</tr>
<tr>
<td>Error</td>
<td>( SSE = \sum_h \sum_i \sum_j \sum_k (y_{hijk} - \bar{y}_{h..})^2 )</td>
<td>( K-I )</td>
</tr>
<tr>
<td>Total</td>
<td>( SSTot = \sum_h \sum_i \sum_j \sum_k (y_{hijk} - \bar{y}_{..})^2 )</td>
<td>( K-1 )</td>
</tr>
</tbody>
</table>

³ It is to be noted that these specific restrictions are needed in order that the sums of squares in the ANOVA table below decompose properly.
The corresponding longitudinal ANOVA results would be obtained by following the instructions 1, 2, 3 stated earlier in this section. The index \( i \) there corresponds to the pair \((h, i)\) here, and the longitudinal ANOVA model is the same as model (13) but with \( y_{hijk} = \mu_{hi} + \varepsilon_{hijk} \) and \( \varepsilon_{hijk} \sim NID(0, \sigma^2) \) replaced by \( \Delta y_{hij} = \mu_{hi} \Delta t_{hij} + \varepsilon_{hij} \) and \( \varepsilon_{hij} \sim NID(0, \sigma^2 \Delta t_{hij}) \), respectively. The least-square estimates and confidence intervals (14)-(15) translate to

\[
\hat{k}_h = \Delta \bar{x}_{h,,} - \Delta \bar{x}_{,\cdot} \\
\kappa_h \in \hat{k}_h \pm \frac{t_{j-1,a/2}}{\hat{\sigma}} \sqrt{K_{h,,}^{-1} - K_{,\cdot}^{-1}}
\]

\[
\hat{\lambda}_{hi} = \Delta \bar{x}_{hi,} - \Delta \bar{x}_{h,} \\
\lambda_{hi} \in \hat{\lambda}_{hi} \pm \frac{t_{j-1,a/2}}{\hat{\sigma}} \sqrt{K_{hi,}^{-1} - K_{h,}^{-1}}
\]

where \( \hat{\sigma}^2 = MS_E = SSE / (J - I) \) and we have used \( J - I \) instead of \( K - I \) as the degrees of freedom for \( SSE \). The classical nested ANOVA table above translates to:

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>d.f.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \kappa )</td>
<td>( SS_{\kappa} = \sum_h \Delta t_{h,i}(\Delta \bar{x}<em>{h,} - \Delta \bar{x}</em>{,\cdot})^2 )</td>
<td>( H - 1 )</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>( SS_{\lambda} = \sum_h \sum_j \Delta t_{hij}(\Delta \bar{x}<em>{hi,} - \Delta \bar{x}</em>{h,})^2 )</td>
<td>( I - H )</td>
</tr>
<tr>
<td>Error</td>
<td>( SSE = \sum_h \sum_{ i,j } \Delta t_{hij} \left( \frac{\Delta \bar{x}<em>{hij}}{\Delta t</em>{hij}} - \Delta \bar{x}_{hi,} \right)^2 )</td>
<td>( J - I )</td>
</tr>
<tr>
<td>Total</td>
<td>( SS_{\text{Total}} = \sum_h \sum_{ i,j } \Delta t_{hij} \left( \frac{\Delta \bar{x}<em>{hij}}{\Delta t</em>{hij}} - \Delta \bar{x}_{hi,} \right)^2 )</td>
<td>( J - 1 )</td>
</tr>
</tbody>
</table>

These results arise by invoking the guidelines above to obtain the substitutions

\[
K - I \leftarrow J - I \\
K_{h,i} \leftarrow \Delta t_{hij} \\
K_{hi,} \leftarrow \Delta t_{hi,} \\
K_{,\cdot} \leftarrow \Delta t_{,\cdot}
\]

\[
y_{hijk} \leftarrow \Delta x_{hij} / \Delta t_{hij} \\
\bar{y}_{hi,} \leftarrow \Delta \bar{x}_{hi,} \\
\bar{y}_{h,} \leftarrow \Delta \bar{x}_{h,} \\
\bar{y}_{,\cdot} \leftarrow \Delta \bar{x}_{,\cdot}
\]

under which the inner sum over \( k \) in \( SSE \) simplifies since its argument no longer depends on \( k \).

5. Example Analysis

Ebara et al. (Ebara et al., 1986) published data (see Figure 1) on the growth of untreated hepatocellular carcinoma tumors. To illustrate the methods presented here, we perform a one-way random effects Brownian ANOVA (Table 1) to analyze growth-rate heterogeneity in this data.
Sophisticated differential equation models of tumor growth are available (see (Araujo & Mcelwain, 2004)). Here for simplicity we focus on an exponential growth model, in which, if $z_t$ is size at time $t$, and $\alpha_t$ is growth rate at time $t$, then

$$dz_t = \alpha_t dt \cdot z_t,$$

which yields exponential growth $z_t = z_0 e^{\alpha t}$ when $\alpha_t$ is a constant $\alpha$. A stochastic version of this model allows growth rate $\alpha_t$ to vary in random independent increments over time:

$$\alpha_t dt = \mu dt + \sigma dB_t,$$

where $\mu$ is mean growth rate, $\sigma$ is growth rate volatility, and $dB_t$ is an increment of standard Brownian motion. This assumption forces $z_t$ to be geometric Brownian motion (e.g., (Oksendal, 1998)), that is, $x_t = \ln(z_t)$ is Brownian motion with drift $\mu$ and volatility $\sigma^2$:

$$\frac{dz_t}{z_t} = d(\ln(z_t)) = \mu dt + \sigma dB_t.$$

Chia et al. (Chia, P. Salzman, Plevritis, & Glynn, 2004) employ this model to explore simulation-based parameter estimation for breast cancer tumor growth. If we allow mean growth rate $\mu$ to depend on patient $i$, then our model (3) is an immediate result.

A random-effects model allows growth rates $\mu_i$ for patients $i$ to be randomly sampled from a normal population with mean $\nu$ and standard deviation $\sigma_\mu$. Our one-way ANOVA procedures (Table 1) allow us to estimate $\nu$ and $\sigma_\mu$ and

Figure 1. Hepatocellular carcinoma tumor growth in 22 patients from Ebara et al. (1986).
test the hypothesis $H: \sigma^2 = 0$. We performed such a one-way random-effects ANOVA based on Table 1. The procedure is simple enough that we carried it out on a spreadsheet.

A word about units: Because $x_t = \ln(z_t)$, we have $dx_t = dz_t/z_t$, so that $dx_t$ (and therefore $x_t$ itself) is a unitless quantity. It follows that the quantities $\mu, \sigma^2, \nu, \sigma_\nu$ have units of time$^{-1}$, as do all sums of squares and mean squares. To make such quantities more meaningful in our results below, we convert them to percents per 100 days.

The results from applying the ANOVA of Table 1 are:

\[ \sigma^2 = 1.42\% /100 \text{ da.} \]
\[ \hat{\sigma}_\nu = 9.64\% /100 \text{ da.} \]
\[ \hat{\nu} = 13.2\% /100 \text{ da.} \]
\[ \hat{se}(\hat{\nu}) = 2.53\% /100 \text{ da.} \]

Therefore, an estimated 95% confidence interval for the mean $\nu$ of the population of mean growth rates is $13.2\% \pm 2.53\% = 13.2 \pm 4.96$ per 100 days, and the standard deviation $\sigma_\nu$ of the population of mean growth rates is estimated to be 9.64% per 100 days. The results of applying the Table 1 ANOVA test of the null hypothesis $H: \sigma^2 = 0$ are

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>118.7%</td>
<td>20</td>
<td>5.94%</td>
<td>4.174</td>
<td>1.45E-06</td>
</tr>
<tr>
<td>Error</td>
<td>126.6%</td>
<td>89</td>
<td>1.42%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

from which we conclude that the above estimate $\sigma_\nu = 9.64\% /100$ days is significantly different from 0 ($p$-value $1.45 \times 10^{-6}$). We conclude that growth-rate heterogeneity is present in the population from which these observations arose.

6. Conclusion

The primary theoretical result from this paper is that any longitudinal model (1)-(2) can be viewed as arising from a corresponding “timeless” classical ANOVA model in which sample sizes are observation times; and that statistical inferences for the former can be directly extracted from the latter once error degrees of freedom are suitably revised. The models (1)-(2) include not only ANOVA models of all types, but also regression and ANCOVA models, for example. We concede, however, that the practical impact of this result is limited because: (i) as noted in Section 4,
the required software inputs for the corresponding classical model could be cumbersome, and the software outputs would need to be revised in light of the altered degrees of freedom; and (ii) it is possible to transform any longitudinal model (1)-(2) into a general linear model, and use available statistical software to numerically derive estimates and hypothesis test results without any necessity of exploiting the connection to classical ANOVA. However, our procedure does have practical usefulness when the desired longitudinal model arises from a classical ANOVA possessing a closed-form sum-of-squares table. In that case, the classical table converts directly into a sum-of-squares table for the desired longitudinal model. The easiest version of this is likely the one-way random effects ANOVA (Table 1) in which it is desired to estimate the population of drifts from which the data have arisen, and to test for heterogeneity in this population. These estimates and tests are simple enough to be derived on a spreadsheet. We believe the practical import of this paper will be for analysts with limited access to or experience with statistical software or general linear models, who wish to quickly obtain familiar ANOVA results without acquiring software or studying general linear models.

References


Appendix

Suppose we have observations $y_1, \ldots, y_J$ and weights $w_1, \ldots, w_J$. Define the sum of squares

$$SS = \sum_j w_j (y_j - \bar{y})^2 \quad \text{where} \quad \bar{y} = \sum_j w_j y_j$$

where $w = \sum_j w_j$. The following result is standard and a consequence of elementary algebra.

**Lemma A1:** $SS = \sum_j w_j y_j^2 - w\bar{y}^2$.

**Lemma A2:** Suppose in addition that the variables $y_i$ are independent with mean $\mu$ and variance $\sigma^2$. Then

$$E[SS] = \left( w - w^{-1} \sum_j w_j^2 \right) \sigma^2.$$ 

**Proof:** We have

$$E\left[ y_j^2 \right] = \mu^2 + \sigma^2$$

$\bar{y}$ has mean $\mu$ and variance $w^2 \sigma^2 \sum_j w_j^2$, so

$$E\left[ \bar{y}^2 \right] = \mu^2 + w^2 \sigma^2 \sum_j w_j^2.$$ 

Then
$$E[SS] = E\left[\sum_j w_j y_j^2 - w\overline{y}^2\right] = \sum_j w_j (\mu^2 + \sigma^2) - w (\mu^2 + \sigma^2 \sum_i w_i^2)$$

$$= w\sigma^2 - \sigma^2 \sum_i w_i^2 = (w - \sigma^2 \sum_i w_i^2)\sigma^2$$

\* \*

The following is also a standard result (e.g., see Theorems 8.3.3 and 8.9.3 in (Graybill, 1983)).

**Lemma A3:** Let the $k \times k$ matrix $C$ be given by

$$C = D + \alpha ab'$$

where $D$ is a nonsingular symmetric matrix, $a$ and $b$ are each $k \times 1$ vectors, and $\alpha$ is a scalar such that $1 + \alpha a'D^{-1}b \neq 0$. Then

$$C^{-1} = D^{-1} - \frac{\alpha}{1 + \alpha a'D^{-1}b} D^{-1}ab'D^{-1}.$$