Measuring the Attenuation in a Subject-specific Random Effect with Paired Data

G. Jones, A. D. L. Noble, B. Schauer and N. Cogger
Massey University

Abstract: This paper is motivated by an investigation into the growth of pigs, which studied among other things the effect of short–term feed withdrawal on live weight. This treatment was thought to reduce the variability in the weights of the pigs. We represent this reduction as an attenuation in an animal–specific random effect. Given data on each pig before and after treatment, we consider the problems of testing for a treatment effect and measuring the strength of the effect, if significant. These problems are related to those of testing the homogeneity of correlated variances, and regression with errors in variables. We compare three different estimates of the attenuation factor using data on the live weights of pigs, and by simulation.

Key words: Homogeneity of variances, hierarchical model, measurement error, orthogonal regression.

1. Introduction

This paper is motivated by an investigation into factors influencing variability in live pig weights in the North Island, New Zealand. Part of this study included an investigation into the effect of short-term feed withdrawal on live weight, for a period of approximately twelve hours from the evening until the morning of the next day. It was conjectured that part of the variability in live weight between pigs was due to differences in gut fill, and that a short period of feed withdrawal would eliminate this factor, leading to a slightly less variable, and more useful, measure of pig live weight. (Gut fill is defined as the weight of contents of the stomach, small and large intestine, which represents 3% to 10% of live body weight. See Kyriazakis *et al.*, 2006).

Testing for a reduction in variance with paired data can be regarded as testing the equality of the diagonal elements in a bivariate covariance matrix. This problem was solved for a bivariate normal distribution by Pitman (1939) and Morgan (1939). The Pitman-Morgan method involves testing the correlation between the sum and difference of the two responses, with zero correlation corresponding to

equality of the two variances. Later authors (Bhoj, 1984; Harris, 1985; Shapiro and Cohen, 1990) have extended this work to include missing values, multiple repeated measures, or robustness to non-normality. A thorough review, and further references, are given by Piepho (1996).

To measure the size of the treatment effect, if any, we could consider the proportional reduction in variance. Pitman (1939) describes the calculation of fiducial limits for the ratio of the two variances, before and after treatment. However these variances will typically include a within-subject component, which if present will lead to a reduction in the apparent treatment effect, biasing any comparison of treatment effect between different groups with different error variances. Our proposed solution is to model the treatment effect directly as a proportional reduction in an individual-specific random effect. Section 2 presents the general modeling and estimation framework, with the application to pig live weight illustrated in Section 3. In Section 4 we investigate the performance of alternative estimators of the treatment effect using simulation. Finally, the results and methodology are discussed in Section 5.

2. Hierarchical Model

Denote the observed measurements on individual i before and after treatment as Y_{1i} and Y_{2i} respectively. We wish to model the reduction in variance due to treatment, allowing for correlation between Y_{1i} and Y_{2i}. To do this we write

$$Y_{1i} = \mu_1 + \theta_i + e_{1i} \tag{2.1}$$
$$Y_{2i} = \mu_2 + k\theta_i + e_{2i}, \tag{2.2}$$

where θ_i is assumed to be normally distributed with zero mean and variance σ_θ^2 and e_{1i}, e_{2i} are independent and identically normally distributed with mean zero and variance σ_e^2. This decomposes the observed variation into a between-individuals component σ_θ^2 and a within-individual component σ_e^2. The treatment effect k is thus represented as a proportional reduction in an individual-specific random effect, reducing the first component to $k^2\sigma_\theta^2$ but not affecting the second component. This can be regarded as a definition of the treatment effect, in which the covariance matrix of $(Y_{1i}, Y_{2i})'$ is parameterized as

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} \sigma_\theta^2 + \sigma_e^2 & k\sigma_\theta^2 \\ k\sigma_\theta^2 & k^2\sigma_\theta^2 + \sigma_e^2 \end{bmatrix} \tag{2.3}$$

so that the ratio of variances of the observed data is $(k^2\sigma_\theta^2 + \sigma_e^2)/(\sigma_\theta^2 + \sigma_e^2)$. The use of this variance ratio as a measure of the treatment effect ignores the influence of the within-individual variation. If this variation is large it will tend to understate the actual effectiveness of the treatment in reducing between-individual variance.

We suggest using an estimate of k as the measure of treatment effect. We first note that a test for the significance of the treatment effect involves testing the null hypothesis $k = 1$, which is equivalent to testing $\sigma_{11} = \sigma_{22}$. For this the Pitman-Morgan test (Pitman,1939; Morgan, 1939) may be used. First we define the pairwise sum and difference

$$
\begin{aligned}
S_i &= Y_{1i} + Y_{2i} = \mu_1 + \mu_2 + (1 + k)\theta_i + e_{1i} + e_{2i} \qquad (2.4)\\
D_i &= Y_{1i} - Y_{2i} = \mu_1 - \mu_2 + (1 - k)\theta_i + e_{1i} - e_{2i}, \qquad (2.5)
\end{aligned}
$$

noting that $e_{1i} + e_{2i}$ and $e_{1i} - e_{2i}$ are uncorrelated and therefore, by assumption of normality, independent. The covariance of $(S_i, D_i)'$ is

$$
\Sigma^* = \left[\begin{array}{cc} \sigma_{11}^* & \sigma_{12}^* \\ \sigma_{12}^* & \sigma_{22}^* \end{array} \right] = \left[\begin{array}{cc} (1 + k)^2 \sigma_\theta^2 + 2\sigma_e^2 & (1 - k^2)\sigma_\theta^2 \\ (1 - k^2)\sigma_\theta^2 & (1 - k)^2 \sigma_\theta^2 + 2\sigma_e^2 \end{array} \right] \qquad (2.6)
$$

so a test of $k = 1$, equivalently $\sigma_{11} = \sigma_{22}$, can be performed by testing the significance of the correlation between S and D, $\rho_{SD} = 0$. The Pitman-Morgan test calculates the sample correlation coefficient r_{SD} for n individuals and refers $r_{SD}\sqrt{n-2}/\sqrt{1 - r_{SD}^2}$ to a t-distribution with $n - 2$ degrees of freedom.

An equivalent, and perhaps more convenient, way of carrying out this test is to regress D_i on S_i and to assess the significance of the slope. This can be regarded as an "errors in variables" regression, with $e_{1i} + e_{2i}$ as the measurement error in the x variable $\mu_1 + \mu_2 + (1 + k)\theta_i$. In the absence of measurement error, the true slope of the regression would be $(1 - k)/(1 + k)$, or zero if $k = 1$. It is well-known (see Fuller, 1987; Carroll, Rupert and Stefanski,1995) that the effect of measurement error in x is to bias the estimated slope towards zero. However, as Carroll *et al.* (1995) point out, the naive test that the slope is zero, using standard regression output, is valid provided there are no other covariates and the measurement error in x is uncorrelated with the error in y.

If the treatment effect is found to be significant, we can use an estimate of k to describe the strength of the effect. We now consider some alternative estimators, using the framework of measurement error models, and noting that in this case the measurement errors $e_{1i} + e_{2i}$ and $e_{1i} - e_{2i}$ in Equation 2.4 are uncorrelated with equal variances.

2.1 Naive regression estimator

A naive approach to the estimation of k would be to simply regress Y_{2i} on Y_{1i}. In the absence of measurement error e_{1i} the regression coefficient would be unbiased for k, but from Equation 2.3 we can see that the expected slope is $\sigma_{12}/\sigma_{11} = k\sigma_\theta^2/(\sigma_\theta^2 + \sigma_e^2)$. Since $0 < \sigma_\theta^2/(\sigma_\theta^2 + \sigma_e^2) < 1$ the regression slope $\hat{k}_{nr}$ is biased towards zero. We could attempt to correct for this bias by multiplying by

an estimate of $(\sigma_\theta^2 + \sigma_e^2)/\sigma_\theta^2$, which is then equivalent to the method of moments estimator of the next section. Carroll *et al.* (1995, p33) point out that such bias-correction will also increase the variance and perhaps lead to a larger mean squared error. Since the variance of $\hat{k}_{nr}$ is not mathematically tractable, this issue is investigated using simulation in Section 4.

2.2 Method of moments

By equating the components of the covariance matrix in Equation 2.3 to their sample values s_{11}, s_{12}, s_{22} we can solve for k and the variance parameters σ_e^2, σ_θ^2. Writing $\gamma = (s_{11} - s_{22})/s_{12}$ we get

$$\hat{k}_{mm} = (\sqrt{\gamma^2 + 4} - \gamma)/2$$

and, assuming $s_{12} > 0$, the variance components are then

$$\hat{\sigma}_\theta^2 = s_{12}/\hat{k}_{mm} \text{ and } \hat{\sigma}_e^2 = s_{11} - \hat{\sigma}_\theta^2$$

If we use these estimates of the variance components to bias-correct the naive estimate $\hat{k}_{nr}$ as suggested above, we again get the method of moments estimator $\hat{k}_{mm}$. Although the sample moments are unbiased for the true values, the estimated correction factor will not be, so we expect this correction to reduce but not eliminate bias. We note too that $\hat{k}_{mm}$ is the slope of the first principal component of $(Y_{1i}, Y_{2i})'$, also the orthogonal regression estimator (Carroll *et al.*,1995, p28).

2.3 Pitman-Morgan regression estimator

A third alternative is to use the regression of the difference D_i on the sum S_i, as given in Equation 2.4. Again this is an errors-in-variables regression. In the absence of measurement error in D_i $(e_{1i} - e_{2i} \equiv 0)$ the expected regression slope would be

$$\beta = (1 - k)/(1 + k)$$

but the covariance matrix of Equation 2.6 shows that

$$\mathrm{E}(\hat{\beta}) = \frac{\sigma_{12}^*}{\sigma_{11}^*} = \frac{(1 - k^2)\sigma_\theta^2}{(1 + k)^2\sigma_\theta^2 + 2\sigma_e^2} = \left(\frac{1 - k}{1 + k}\right)\frac{\sigma_\theta^2}{\sigma_\theta^2 + 2\sigma_e^2/(1 + k)^2}$$

so again the measurement error causes a bias towards zero. If we back-transform we get the estimator

$$\hat{k}_{pm} = \frac{1 - \hat{\beta}}{1 + \hat{\beta}}$$

A first-order approximation gives

$$\mathrm{E}(\hat{k}_{pm}) = k + \frac{(1-k)\sigma_e^2}{(1+k)\sigma_\theta^2 + \sigma_e^2}$$

so assuming $k < 1$ there is a positive bias in $\hat{k}_{pm}$ which will be quite small if $\sigma_e^2 \ll \sigma_\theta^2$ or k is close to one. The next term in a higher-order approximation to $\mathrm{E}(\hat{k}_{pm})$ involves $\sigma_{\hat{\beta}}^2$ and will be much smaller than the previous term if the data set is reasonably large. We can expect then that the bias of this estimator will be small so that it may out-perform $\hat{k}_{mm}$. This too is investigated by simulation in Section 4.

2.4 Bayes estimator

If we specify prior distributions for the parameters k, μ_1, μ_2, σ_θ^2 and σ_e^2 then the posterior distribution of k is easily obtained by Markov chain Monte Carlo estimation (Gilks, Richardson and Spiegelhalter, 1996) using, for example, the BUGS program (Spiegelhalter, Thomas, Best and Gilks, 1994). This has the advantage of providing a measure of the uncertainty of the estimate, for example the standard error of the estimated posterior distribution. For the other non-Bayesian estimators, analytical expressions for the standard errors are mathematically intractable, although approximate standard errors can be calculated using the bootstrap as illustrated in the next section.
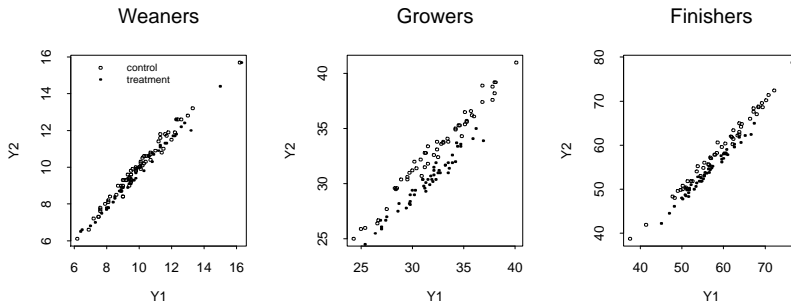
3. Analysis of Pig Data

Figure 1: Live weights of pigs before ($Y1$) and Afer ($Y2$) feed withdrawl.

To investigate the effect of feed withdrawal on live weight, pigs in three different age groups ("weaners", "growers" and "finishers") were split randomly into control and treatment groups. They were weighed in the evening and again the following morning after a time lapse of 11 hours for the weaners and 17 hours for

the other groups. Those in the control group were fed normally, but food was withheld from the treatment group. It is thought that "gut fill" is a component of the variability in live weight, and that feed withdrawal would reduce gut fill to give a less variable and more useful measure of live weight. The experiment was later repeated, with the same treatment and control groups, after a period of 20 days, giving four measures of live weight per pig. The full dataset is available at `http://www-ist.massey.ac.nz/GJones`.

Here we analyze the results of the first experiment using, in the notation of Section 2,

$$Y_{1i} = \text{Live weight in the evening}$$
$$Y_{2i} = \text{Live weight the following morning}$$

The data are plotted in Figure 1. There is a clear difference between control and treatment groups in the Growers, less so in the Finishers, and very little apparent difference in the Weaners. Table 1 gives the Pitman-Morgan p-value for each group for testing the null hypothesis of no difference in the variances of Y_1 and Y_2. These are two-sided p-values allowing for the possibility that the between-pig variance has increased, which is possible since they are continuing to grow and, in the control group, to eat. For all three age groups there is a significant decrease in variance with treatment. The results for the variance decomposition (σ_θ, σ_e) show that the within-pig variation is much smaller than the between-pig component. Both show a gradual increase with age. It is noticeable that the within-pig component is similar between control and treatment groups, whereas the between-pig component is much larger in the control group for both Growers and Finishers. This does not seem to be a failing in the model, since the discrepancy is present in the initial standard deviations σ_{Y1}. This perhaps calls into question the random allocation into treatment groups.

Table 1: Results of analysis of three groups of pigs, giving the p-value for the Pitman-Morgan test and the estimated standard deviations. * denotes the results for the Bayesian analysis using MCMC.

Group	Treatment	n	pval	$\hat{\sigma}_{Y1}$	$\hat{\sigma}_{Y2}$	$\hat{\sigma}_\theta$	$\hat{\sigma}_e$	pval*	$\hat{\sigma}_\theta^*$	$\hat{\sigma}_e^*$
Weaners	Control	66	0.7246	1.7408	1.7296	1.7315	0.1793	0.7300	1.7530	0.1827
	Withdrawl	52	0.0008	2.0704	1.9599	2.0647	0.1530	0.0004	2.0920	0.1583
Growers	Control	52	0.9485	3.7364	3.7417	3.7141	0.4075	0.9352	3.7650	0.4198
	Withdrawl	51	0.0014	2.6355	2.3880	2.6103	0.3633	0.0010	2.6470	0.3730
Finishers	Control	50	0.7216	8.1935	8.2371	8.1720	0.5933	0.7204	8.2810	0.6122
	Withdrawl	52	0.0484	5.0151	4.8302	4.9942	0.4572	0.0434	5.0640	0.4701

The Bayesian analysis was performed by running the model (2.1), with vague conjugate prior distributions, in WinBugs for 20 000 iterations after a burn-in of 10 000. Table 1 gives the estimated posterior means for σ_θ and σ_e, which are seen

to be broadly similar to those derived by the method of moments. The Bayesian "p-value" shown is really twice the posterior probability that $k \leq 0$ (or $k \geq 0$ as appropriate).

The estimated attenuation factors (k) are displayed in Table 2. Standard errors for $\hat{k}_{mm}$, $\hat{k}_{nr}$ and $\hat{k}_{pm}$ were obtained by resampling from each group of pigs, whilst $\hat{k}_b$ represents the standard deviation of the posterior distribution of k derived from the MCMC output. The downward bias in the naive regression estimator $\hat{k}_{nr}$ is apparent. The Pitman-Morgan regression estimator $\hat{k}_{pm}$ consistently gives results very similar to the method-of-moments $\hat{k}_{mm}$, but is much easier to compute as it involves only a simple regression.

The results suggest that the effect of feed withdrawal is most pronounced in the middle age group, when the growth rate of the pigs is at its highest.

Table 2: Results of analysis of three groups of pigs, giving the estimated treatment effect (standard error) by method of moments (mm), naive regression (nr), Pitman-Morgan regression (pm) and Bayesian posterior mean (b)

| Group	Treatment	n	$\hat{k}_{mm}$	$\hat{k}_{nr}$	$\hat{k}_{pm}$	$\hat{k}_b$
Weaners	Control	66	0.9935 (0.0211)	0.9829 (0.0199)	0.9935 (0.0210)	0.9940 (0.0185)
	Withdrawl	52	0.9463 (0.0163)	0.9411 (0.0163)	0.9465 (0.0163)	0.9465 (0.0149)
Growers	Control	52	1.0014 (0.0215)	0.9895 (0.0214)	1.0014 (0.0214)	1.0020 (0.0226)
	Withdrawl	51	0.9042 (0.0275)	0.8870 (0.0278)	0.9052 (0.0271)	0.9053 (0.0276)
Finishers	Control	50	1.0053 (0.0170)	1.0000 (0.0167)	1.0053 (0.0170)	1.0050 (0.0153)
	Withdrawl	52	0.9628 (0.0204)	0.9548 (0.0207)	0.9630 (0.0203)	0.9627 (0.0184)

4. Simulation Study
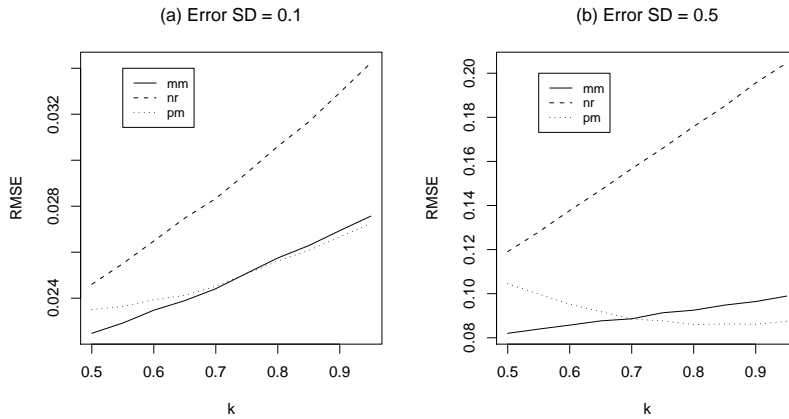
Figure 2: Root-mean-squared error for estimating k by method of moments (mm), naive regression (nr) and Pitman-Morgan regression (pm), for $n = 60$ and $\sigma_e/\sigma_\theta = 0.1$ or 0.5, based on 10 000 simulations.

The analysis of the pig data suggests that for estimating the treatment effect k the Pitman-Morgan regression estimator may be competitive with the method-of-moments estimator. The analysis of Section 2 showed that the former is biased upwards, but this bias may be compensated for by a smaller variance. In this Section we investigate this issue by simulating from the model 2.1 for various values of k from 0.5 to 0.95, and for various values of the ratio σ_e/σ_θ from 0.05 to 0.5. In each simulation 10 000 simulated values of $\hat{\beta}$ were produced to estimate $\mathrm{E}(\hat{\beta})$ and $\mathrm{Var}(\hat{\beta})$, from which the root-mean-squared-error RMSE $= \sqrt{[\mathrm{E}(\hat{\beta} - \beta)]^2 + \mathrm{Var}(\hat{\beta})}$ was calculated. Sample sizes of $n = 30$, 60 and 100 were considered.

Typical results, for $n = 60$, are shown in Figure 2. The bias term dominates the RMSE of the naive regression estimator $\hat{k}_{nr}$, which in all cases has the worst performance. There is negligible bias in the method-of-moments estimator $\hat{k}_{mm}$ but as k increases its variance increases faster than that of the Pitman-Morgan regression estimator $\hat{k}_{pm}$, which has a small positive bias. The result is that for large values of k the use of $\hat{k}_{pm}$ gives the best result in terms of RMSE. When $n = 60$ this occurs at approximately $k > 0.7$, a little earlier ($k > 0.65$) when $n = 30$ and a little later ($k > 0.75$) when $n = 100$. The difference in RMSE between the two estimators is quite small when the measurement error (within-individual) variance σ_e^2 is small, as in the pigs data, but increases as the measurement error increases.

5. Discussion

We have presented a model for paired data in which the change in variance between the first and second measurements is represented as a re-scaling of a subject-specific random effect. The purpose behind this is to measure the effect of a treatment on the variability of measurements in a way which allows for measurement error or other within-subject variation. The question arises as to whether the model is appropriate for a given data set. Since the model can be regarded as a re-parametrization of the bivariate normal distribution, testing its goodness of fit is equivalent to testing for bivariate normality. The parameter k measuring the attenuation in the subject-specific random effect can be understood as a definition of what we regard as the effect of the treatment. The within-individual component σ_e^2 could be thought of as measurement error, and in some situations it might be possible to estimate it from repeated measures on each subject. We could then test the appropriateness of the assumption that e_{1i}, e_{2i} have a common variance. However it has to be remembered that the overall variability in the measuring process is not always easy to reproduce as it may include more factors than just the repeatability of the measuring instrument.

We have investigated some alternative estimators of the size of the treatment effect. A Bayesian analysis is reasonably straightforward and would be particularly useful if prior information were available. If conjugate priors are used, the posterior distribution of the variance of $(Y_{1i}, Y_{2i})'$ is inverse gamma, and the model parameters are functions of its components which, though analytically intractable, could be derived by simulation. MCMC gives a simple alternative which converges quickly.

The naive regression estimator has been found to be significantly biased, but regression of the difference on the sum, which we have termed Pitman-Morgan regression, and then transforming the estimated slope, performs well if the treatment effect is not too large (k close to 1). This is the simplest alternative since it requires only standard regression software. Estimation based on moments is a little more complicated but also has good properties, particularly when the attenuation is large ($k \ll 1$).

Acknowledgements

References

Bhoj, D. S. (1984). On testing equality of variances of correlated variates with incomplete data. *Biometrika* **71**, 639-641.

Carroll, R. J., Ruppert, D., and Stefanski, L. A. (1995). *Measurement Eror in Nonlinear Models.* Chapman & Hall.

Fuller, W. A. (1987). *Measurement Error Models.* Wiley.

Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (Eds.) (1996). *Markov Chain Monte Carlo in Practice.* Chapman and Hall.

Harris, P. (1985). Testing for variance homogeneity of correlated variables. *Biometrika* **72**, 103-107.

Kyriazakis, I. and C. T. Whittemore (2006). *Whittemore's Science and Practice of Pig Production*, 3rd edition. Blackwell Publishing.

Morgan, W. A. (1939). A test for the significance of the difference between two variances in a sample from a normal bivariate population. *Biometrika* **31**, 13-19.

Piepho, H.-P. (1997). Tests for equality of dispersion in bivariate samples - review and comparison. *Journal of Statistical Computation and Simulation* **56**, 353-372.

Pitman, E. J. G. (1939). A Note on normal correlation. *Biometrika* **31**, 1-12.

Shapiro, A., and Cohen, A. (1990). Testing and estimation of equal variances for correlated variables. *Statistics and Probability Letters* **10**, 231-234.

Spiegelhalter, D. J., Thomas, A., Best, N. G., and Gilks, W. R. (1994). BUGS: Bayesian inference Using Gibbs Sampling, Version 0.30. Cambridge: Medical Research Council Biostatistics Unit.

G. Jones
Institute of Information Sciences & Technology
Massey University
Private Bag 11222
Palmerston North, NZ
g.jones@massey.ac.nz

A. D. L. Noble
Institute of Information Sciences & Technology,
Massey University, Private Bag 11222, Palmerston North, NZ
a.d.noble@massey.ac.nz

B. Schauer
Institute of Veterinary, Animal and Biolomedical Sciences
Massey University
Private Bag 11222, Palmerston North, NZ
b.schauer@massey.ac.nz

N. Cogger
Institute of Veterinary, Animal and Biolomedical Sciences
Massey University
Private Bag 11222
Palmerston North, NZ
n.cogger@massey.ac.nz