CHAPTER 14

Components of dispersion

14.1 Introduction

The models so far considered all have the property that the observations are assumed independent and the variability associated with each observation is determined by at most a single dispersion factor, σ^2 or ϕ, to be estimated. In many areas of application, however, the experimental or survey design is such that variation is present at several levels or strata. In the absence of external effects, units in the same cluster tend to be more alike, or positively correlated, than units in different clusters. In an educational-testing context pupils within classrooms within schools form a natural hierarchy of three strata, and it is natural to associate a random component to the units within each of the three strata. Usually these variance components, or dispersion components as we shall call them, are given contrasting descriptive labels such as 'between-schools variance' and 'between-classrooms-within-schools variance', or 'between-animals variance' and 'within-animals variance' depending on the context. Usually the between-blocks or between-animals variance is larger than the within-blocks or within-animals variance. For example in the tuberculin-assay problem discussed in section 6.3.1, where four measurements were taken on each cow, it was anticipated that the variability of reaction between different cows would be considerably larger than the variability between sites on the same animal. Consequently, provided that there is no interaction between tuberculins applied at different sites, greater precision can be achieved by comparing tuberculins on the same animal.

432

14.2 Linear models

Before proceeding to consider more complicated non-linear models it is helpful to consider first the essential ingredients of linear models with so-called 'random effects'. We take as our example the tuberculin-assay data in Table 6.1. To keep the discussion as simple as possible it is assumed initially that all effects are additive with constant variance on the log scale. This assumption is almost, but not quite, in accord with the analysis in section 6.3.1.

The standard mathematical description is to express the response as an additive function of both fixed (non-stochastic) and random effects, all random effects being assumed independent. Thus we write

$$\log Y_{ij(k)} = \alpha_i + \gamma_j + \tau_k + \epsilon_{ijk}, \qquad (14.1)$$

where k denotes the tuberculin type and volume applied at site i to cows in cow class j. Primary interest centres on the tuberculin effects τ_k, and the analysis in section 6.3 focused exclusively on the joint effect of tuberculin type and volume. Of secondary interest, but nonetheless perhaps physiologically important, is the sensitivity to tuberculin of the four sites on the cow's neck. Thus τ_k and α_i are considered here as fixed effects. Since there can be no lasting scientific interest in the sensitivity of individual unidentified cows, it is appropriate to take γ_j, the cow-class effect corresponding to differences between cows, as independent random variables. Finally, the residual errors ϵ_{ijk} are taken to be mutually independent and independent of the cow-class effects γ_j.

Since $Y_{ij(k)}$ is the sum of observed values for 30 cows, it is appropriate to take

$$\mathrm{var}(\epsilon_{ijk}) = \sigma^2/30, \qquad \mathrm{var}(\gamma_j) = \sigma_b^2/30.$$

Thus σ^2 is the 'within-cow' variance and σ_b^2 is the 'between-cow' variance. Note that $\mathrm{var}(\log Y) \simeq \mathrm{cv}^2(Y)$ (Exercises 14.10–11). Further, if Y_i are independent,

$$\mathrm{cv}^2\{\sum Y\} \simeq \mathrm{cv}^2(Y)/n,$$

where $\mathrm{cv}(Y)$ is the coefficient of variation of Y. The variances σ^2 and σ_b^2 then refer to individual cows and not to averages or totals over 30 cows.

The usual analysis-of-variance decomposition is shown in Table 14.1. From the 'error' line in this table we obtain the estimate $\tilde{\sigma}^2 = 0.0216$ or $\tilde{\sigma} = 0.147$. In other words, the within-cow coefficient of variation is approximately 15%. From the 'cow-class' line together with the 'error' line we find $\tilde{\sigma}_b^2 = 0.8821$ or $\tilde{\sigma}_b = 0.939$, showing that the between-cow coefficient of variation is approximately 94%. In other words there is very substantial variation between animals, and comparatively little variation within animals.

Table 14.1 *Analysis of variance for the tuberculin-assay data*

Source	S.S.	d.f.	M.S.	E(M.S.)
Cow class	0.47232	3	0.11833	$\sigma^2/30 + 4\sigma_b^2/30$
Sites	0.08324	3	0.02775	$\sigma^2/30 + 4\sum(\alpha_i - \bar{\alpha})^2/3$
Treatments	0.17596	3	0.05865	$\sigma^2/30 + 4\sum(\tau_k - \bar{\tau})^2/3$
Error	0.00433	6	0.00072	$\sigma^2/30$
Total	0.73584	15		

Because of the orthogonality built into the Latin square design the site effects and the tuberculin effects are uncorrelated, and independent under Normality. Further, the variances of the estimated tuberculin effects and site effects are functions only of σ^2 and not of σ_b^2. For this reason, if interest focuses on the site and tuberculin effects, the same results would be obtained if we were to condition on the cow-class effects and take them as fixed. The estimated site effects are

$$\boldsymbol{\alpha} = (0.000, 0.093, 0.128, -0.053),$$

and the standard errors of simple contrasts are $\tilde{\sigma}/\sqrt{60} = 0.019$. Thus sites 2 and 3 are significantly more sensitive than sites 1 and 4.

14.3 Non-linear models

The discussion in this section is meant to be quite general, but in order to keep the terminology concrete we use the tuberculin-assay example. The analysis that follows is performed on the original scale and avoids explicitly transforming the data. This feature is a characteristic of generalized linear models.

Suppose that the conditional mean and variance of $Y_{ij(k)}$ given the cow-class effects are as follows:

$$E(Y_{ij(k)} \mid \text{assignment of cows to classes}) = M_{ij(k)},$$
$$\text{var}(Y_{ij(k)} \mid \text{assignment of cows to classes}) = \sigma^2 V(M_{ij(k)}),$$

where $V(\cdot)$ is the known conditional variance function. All observations, whether in the same cow class or in different classes, are assumed to be conditionally independent. Suppose in addition that the conditional mean satisfies the log-linear model

$$\log M_{ij(k)} = \alpha_i + \gamma_j + \tau_k \qquad (14.2)$$

in which the treatment effects τ_k, and possibly the site effects α_i, are the parameters of interest. Evidently this conditional formulation specifies a generalized linear model, which happens in this case to be log-linear. In this example all parameters, including γ_j, can be estimated in the usual way without further assumptions, and no new issues arise. The analysis in section 6.3 corresponds to the choice $V(M) = M$, whereas the analysis in the previous section via response transformation corresponds roughly to $V(M) = M^2$.

Suppose now that we insert into the model the further, quite reasonable, assumption that the cow-class effects are independent and identically distributed random variables from a particular family of distributions, say

$$\gamma_j \sim N(0, \sigma_b^2). \qquad (14.3)$$

This additional assumption does not invalidate the conditional analysis recommended above, but it does open up the possibility that a more efficient analysis could be devised by making use of the random-effects assumption (14.3).

The most illuminating examples of this type occur when the parameters of interest are not estimable in the conditional model (14.2). Estimation is then impossible in the absence of a further assumption such as (14.3). Suppose, by way of example, that the tuberculin-assay design is such that all cows in a given class receive the same treatment at each of the sites. Suppose further that two replicates of the experiment are available, so that there are eight cow classes in all. In other words the experimental arrangement

Table 14.2 *Alternative design for tuberculin assay problem*

Cow class (j)		I	II	III	IV	V	VI	VII	VIII
Treatment (k)		A	B	C	D	A	B	C	D
Site (i)	1								
	2				$Y_{ij(k)}$				
	3								
	4								

is as shown in Table 14.2. Evidently the treatment contrasts are now aliased with a subset of the cow-class contrasts, and are not estimable in the fixed-effects model (14.2).

Taking the log-linear model (14.2) together with the random-effects assumption (14.3), however, we have

$$M_{\cdot j(k)} = \sum_i \exp(\alpha_i) \times \exp(\gamma_j) \times \exp(\tau_k)$$

so that the unconditional means for the cow-class totals are

$$\mu_{\cdot j(k)} = E(M_{\cdot j(k)}) = \sum \exp(\alpha_i) \times E(e^{\gamma_j}) \times \exp(\tau_k)$$
$$= \sum \exp(\alpha_i) \times \exp(\sigma_b^2/2) \times \exp(\tau_k).$$

On the log scale we have

$$\log(\mu_{\cdot j(k)}) = \tau_k + \text{const}, \qquad (14.4)$$

which depends only on the treatment applied. Clearly the eight cow-class totals are independent random variables. Their unconditional variances are given by

$$\text{var}(Y_{\cdot j(k)}) = \sigma^2 \sum_i E\{V(M_{ij(k)})\} + \text{var}(M_{\cdot j(k)}). \qquad (14.5)$$

If $V(M) = M$ this gives

$$\text{var}(Y_{\cdot j(k)}) = \sigma^2 \mu_{\cdot j(k)} + \mu_{\cdot j(k)}^2 \,\text{cv}^2(e^\gamma)$$
$$= \sigma^2 \mu_{\cdot j(k)} + \mu_{\cdot j(k)}^2 \big(\exp(\sigma_b^2) - 1\big).$$

On the other hand if $V(M) = M^2$ we have

$$\text{var}(Y_{\cdot j(k)}) = \sigma^2 \sum_i E(M_{ij(k)}^2) + \text{var}(M_{\cdot j(k)})$$
$$= \sigma^2 \sum_i \mu_{ij(k)}^2 \{1 + \text{cv}^2(e^\gamma)\} + \mu_{\cdot j(k)}^2 \,\text{cv}^2(e^\gamma)$$
$$= \mu_{\cdot j(k)}^2 \big\{\text{cv}^2(e^\gamma) + \frac{\sigma^2}{4}\big(1 + \text{cv}^2(e^\gamma)\big)\big(1 + \text{cv}^2(e^\alpha)\big)\big\},$$

where $\mathrm{cv}^2(\alpha') = \sum(\alpha'_i - \bar{\alpha}')^2/(n\bar{\alpha}'^2)$. In the first case the unconditional variance function is approximately quadratic provided that $\sigma_b^2\bar{\mu} \gg \sigma^2$. In the second case the unconditional variance function is exactly quadratic. Thus the parameters in (14.4) can be estimated from the eight cow-class totals by using a quadratic variance function and a log link.

From two replicates the coefficient of variation of the cow-class totals can be estimated on four degrees of freedom. It is this estimate that must be used for setting confidence limits for treatment effects.

For the estimation of site effects it is unnecessary to use the random effects assumption (14.3). The site effects are estimable in the fixed-effects model (14.2) and their variances do not depend on σ_b^2. From two replicates the residual variance σ^2 can be estimated on 21 degrees of freedom using the residual deviance from model (14.2).

14.4 Parameter estimation

Parameter estimates are obtained using the quasi-likelihood estimating equation (9.5). We use this method in preference to explicit maximum likelihood chiefly because it is often much simpler and is based only on properties of the unconditional mean and covariance matrix of the observations. Suppose that the unconditional mean and covariance matrix of $\mathbf{Y}$ are

$$E(\mathbf{Y}) = \boldsymbol{\mu}(\boldsymbol{\beta}), \qquad \mathrm{cov}(\mathbf{Y}) = \mathbf{V}(\boldsymbol{\mu}, \boldsymbol{\sigma}^2),$$

where the components of $\boldsymbol{\sigma}^2$ are the dispersion components. Then, following the notation established in Chapter 9, the estimating equations for $\boldsymbol{\beta}$ are $\mathbf{U}(\hat{\boldsymbol{\beta}}, \boldsymbol{\sigma}^2) = \mathbf{0}$, where

$$\mathbf{U}(\boldsymbol{\beta}, \boldsymbol{\sigma}^2) = \mathbf{D}^T\mathbf{V}^{-1}\big(\mathbf{y} - \boldsymbol{\mu}(\boldsymbol{\beta})\big). \qquad (14.6)$$

Ordinarily, but with some important exceptions, the solution to this equation depends on the dispersion components, usually on ratios of dispersion components. In such cases it is necessary to use a supplementary set of estimating equations for the dispersion parameters. Such a system of estimating equations is now derived.

In all of the examples that we have in mind the dispersion components are physical characteristics of identifiable populations and the covariance matrix can then be written in the form

$$\mathbf{V}(\boldsymbol{\mu}, \boldsymbol{\sigma}^2) = \sigma_1^2 \mathbf{V}_1(\boldsymbol{\mu}) + \ldots + \sigma_k^2 \mathbf{V}_k(\boldsymbol{\mu}). \qquad (14.7)$$

Usually the rank of $\mathbf{V}_j(\boldsymbol{\mu})$ is equal to the number of elements in the sample that are drawn from the population indexed by j.

By way of illustration we consider the example discussed in the previous section, in which there are two dispersion components. The response variances are given by

$$\mathrm{var}(Y_{ij(k)}) = \sigma^2 E\big(V(M_{ij(k)})\big) + \mathrm{var}(M_{ij(k)}).$$

If the conditional variance function is $V(M) = M$, this expression reduces to

$$\mathrm{var}(Y_{ij(k)}) = \sigma^2 \mu_{ij(k)} + \{\exp(\sigma_b^2) - 1\}\mu_{ij(k)}^2.$$

Similar calculations show that the only non-zero covariances have the form

$$\mathrm{cov}(Y_{ij(k)}, Y_{i'j(k')}) = \mathrm{cov}(M_{ij(k)}, M_{i'j(k')})$$
$$= \mu_{ij(k)}\mu_{i'j(k')}\{\exp(\sigma_b^2) - 1\}$$

for $i' \neq i$. On comparing these expressions with the general form (14.7) we find $\sigma_1^2 = \sigma^2$, $\mathbf{V}_1(\boldsymbol{\mu}) = \mathrm{diag}(\boldsymbol{\mu})$, $\sigma_2^2 = \exp(\sigma_b^2) - 1$, and $\mathbf{V}_2(\boldsymbol{\mu}) = \boldsymbol{\mu}\mathbf{J}\boldsymbol{\mu}^T$, where $\mathbf{J}$ is a block-diagonal unit matrix taking the value unity if two observations refer to the same cow class, and zero otherwise. Note that the rank of $\mathbf{V}_2$ is equal to the number of cow classes.

The most natural way to estimate the dispersion components is to choose k suitable quadratic forms and to equate the observed values of these to their expectations as functions of the parameters. Usually the root of (14.6) is fairly insensitive to the choice of $\boldsymbol{\sigma}^2$, so for that purpose the choice of quadratic forms is not critical. However, the asymptotic covariance matrix of $\hat{\boldsymbol{\beta}}$,

$$\mathrm{cov}(\hat{\boldsymbol{\beta}}) \simeq (\mathbf{D}^T \mathbf{V}^{-1} \mathbf{D})^{-1}$$

does depend heavily on $\boldsymbol{\sigma}^2$.

In many applications there is a natural set of quadratic forms

$$Q_r = (\mathbf{Y} - \boldsymbol{\mu})^T \mathbf{P}_r (\mathbf{Y} - \boldsymbol{\mu}), \qquad r = 1, \ldots, k, \qquad (14.8)$$

in which $\mathbf{P}_r$ is a fixed projection matrix associated with the rth random effect. Usually $\mathbf{P}_r\mathbf{Y}$ is equivalent to a set of marginal totals, in which case Q_r is the sum of squares for that set of totals. Provided that $\mathbf{V}(\boldsymbol{\mu})$ satisfies the additive decomposition (14.7), these quadratic forms have expectations

$$E(Q_r) = \sum_{j=1}^{k} \text{tr}\{\mathbf{P}_r\mathbf{V}_j\}\sigma_j^2, \qquad (14.9)$$

which are linear functions of the dispersion components. In this way we obtain k simultaneous equations for the k dispersion components. As usual with variance components, the estimates obtained need not be positive.

More generally the method of restricted maximum likelihood, as described in section 7.2, can be used here, even though the observations are usually not Normally distributed. However this method gives rise to more complicated estimating equations and requires the inversion of $\mathbf{V}(\boldsymbol{\mu}, \boldsymbol{\sigma}^2)$ for various values of the dispersion parameters. For that reason we use the simpler but less efficient method described above.

In practice, if $p = \dim(\boldsymbol{\beta})$ is positive, it is necessary to use $\hat{\boldsymbol{\mu}}$ in place of $\boldsymbol{\mu}$ in (14.8) and to iterate between (14.6) and (14.8). This method of estimation is illustrated in the following section.

14.5 Example: A salamander mating experiment

14.5.1 *Introduction*

This section describes a particular problem, involving an experiment with binary responses, for which neither the methods of Chapter 4 nor those of Chapter 9 are directly applicable. In fact straightforward application of linear logistic models could lead to misleading conclusions, so it is essential to recognize the characteristics of a design in which explicit recognition of dispersion components is necessary. The problem described here involves several dispersion components associated with different subgroups or

Table 14.3. *Design used in a salamander mating experiment*[†]

	Males					
Females	June 4	June 8	June 12	June 16	June 20	June 24
1	1	4	5	1	4	5
2	5	5	3	3	1	2
RBF 3	RB 2	WS 1	RB 1	WS 4	RB 3	WS 3
4	4	2	2	5	5	4
5	3	3	4	2	2	1
6	9	9	10	7	6	8
7	8	8	9	9	7	6
RBF 8	WS 6	RB 6	WS 7	RB 10	WS 10	RB 9
9	10	7	8	6	9	10
10	7	10	6	8	8	7
1	9	9	7	10	10	8
2	7	6	9	7	6	10
WSF 3	RB 8	WS 7	RB 6	WS 9	RB 7	WS 6
4	10	10	8	8	9	9
5	6	8	10	6	8	7
6	5	2	3	4	2	1
7	4	1	5	2	1	5
WSF 8	WS 1	RB 4	WS 2	RB 5	WS 5	RB 3
9	3	3	1	1	4	4
10	2	5	4	3	3	2

[†]RBF rough-butt female; WSF whiteside female; RB rough-butt male; WS whiteside male.
Each female is paired with three males of each type: each male is paired with three females of each type.

populations, and is not to be confused with simple over-dispersion in which the observations are independent and there is only one dispersion component.

The analysis presented here is admittedly rather rough and ready. It involves at various stages Taylor approximations whose adequacy, even for applied work, is more dubious than usual. However, we believe that problems of this nature are sufficiently common, and suitable statistical techniques either primitive or excessively complicated, that a protracted discussion of a particular example is worthwhile.

14.5.2 *Experimental procedure*

The purpose of this experiment, conducted by S. Arnold and P. Verrell of the Department of Ecology and Evolution at the University of Chicago, was to study the extent to which mountain dusky salamanders from different populations will interbreed. These populations, all belonging to the same species, are geographically isolated from one another, and are found at high elevations in the southern Appalachian mountains of the eastern United States. Thus the salamanders paired together in the laboratory would never encounter each other in their natural environment—at least not at this moment in geological time. The manner in which mechanisms that prevent interbreeding evolve is of great interest to biologists studying speciation. The question of interest here is whether these barriers to interbreeding can evolve in populations that are isolated from one another.

The data given here refer to two populations called Rough Butt (RB) and Whiteside (WS). Forty animals were used in each of three experiments, one conducted in the summer of 1986 and two in the Fall of the same year. Normal practice is to use fresh animals for each experiment but in this instance the animals used in the first Fall experiment were identical to those used in the Summer experiment. Repeat experimentation using the same animals is potentially important but, because it is the exception to normal practice, this aspect will be ignored in the analyses that follow. The three experiments will be treated *as if* different animals were used in each.

The forty salamanders available in each of the three experiments comprise

     10 Rough Butt males numbered 1–5 & 6–10 (RBM)
     10 Rough Butt females numbered 1–5 & 6–10 (RBF)
     10 Whiteside males numbered 1–5 & 6–10 (WSM)
     10 Whiteside females numbered 1–5 & 6–10 (WSF)

According to the design in Table 14.3, which was used in each of the three experiments, Rough Butt females numbered 1–5 (RBF 1–5) were sequestered as heterosexual pairs with Rough Butt males numbered 1–5 on 3 occasions June 4, 12 and 20, and with Whiteside males numbered 1–5 also on 3 occasions, June 8, 16 and 24. These Rough Butt females were never permitted to see RBM 6–10 nor WSM 6–10. Thus, in effect, the design contains two replicates as

Table 14.4. *Observed matings: Summer '86*

Females	June 4	June 8	June 12	June 16	June 20	June 24
			Date			
1	1	1	1	0	1	1
2	1	1	1	1	1	1
RBF 3	RB 1	WS 0	RB 1	WS 1	RB 1	WS 1
4	1	1	1	0	1	1
5	1	1	1	1	1	1
6	1	1	1	0	1	1
7	0	0	0	1	0	0
RBF 8	WS 0	RB 1	WS 0	RB 0	WS 1	RB 1
9	0	0	1	1	1	1
10	0	0	1	0	1	0
1	0	1	1	1	0	1
2	0	0	0	1	0	0
WSF 3	RB 0	WS 0	RB 0	WS 0	RB 0	WS 1
4	0	1	1	1	0	1
5	0	1	0	0	0	0
6	0	0	1	0	0	0
7	1	1	1	0	1	1
WSF 8	WS 1	RB 0	WS 1	RB 0	WS 1	RB 0
9	1	1	1	1	1	0
10	1	0	0	1	1	0

shown in Table 14.7. The non-zero values in Table 14.7 are the actual numbers of matings observed for the various crosses in the Summer '86 experiment. It should be pointed out that, although there are 25 possible crosses between the females RBF 1–5 and the males WSM 1–5, only 15 of these crosses are permitted by the design. For example, RBF 1 was not permitted to see RBM 2 or RBM 3. Conversely RBM 1 did not see RBF 4 or RBF 5.

In order for the design to be complete in this sense it would have been necessary to extend the experiment over 10 nights instead of 6. This was considered impractical because four days of recuperation are required between successive pairings, enabling females to transport sperm, and males to synthesize spermatophores. A complete design would have extended the duration of the experiment from 21 days to 37 days. The design used here comprises eight replicates of an incomplete Latin square, although little use is made of that design in the present analysis.

The design of the experiment permits a comparison of the

Table 14.5. *Observed matings: Fall '86 (re-runs)*

Females		Sept. 4	Sept. 8	Sept. 12	Sept. 16	Sept. 20	Sept. 24
				Date			
	1	1	0	1	0	1	0
	2	1	1	1	0	0	1
RBF	3	RB 0	WS 1	RB 1	WS 1	RB 1	WS 1
	4	1	0	0	0	0	0
	5	1	0	1	0	0	0
	6	1	1	1	0	0	0
	7	1	1	0	1	0	1
RBF	8	WS 1	RB 0	WS 0	RB 0	WS 1	RB 0
	9	1	1	1	1	0	1
	10	0	0	1	1	1	0
	1	0	1	0	0	0	1
	2	0	1	0	0	0	1
WSF	3	RB 0	WS 1	RB 0	WS 1	RB 0	WS 1
	4	0	0	1	1	0	1
	5	0	0	0	0	0	0
	6	1	0	1	1	1	1
	7	1	0	1	0	0	0
WSF	8	WS 0	RB 0	WS 0	RB 0	WS 0	RB 0
	9	1	1	1	1	1	1
	10	1	0	1	1	1	0

mating probabilities for the four possible crosses. In the analysis that follows we concentrate on comparing the mixed crosses. In other words we focus on the question of whether the RBF/WSM crosses result in mating more frequently or less frequently than the WSF/RBM crosses. The observed number of matings in the Summer '86 experiment for each of the crosses is shown in Table 14.7. For the comparison in question we observe that the RBF/WSM cross resulted in 20 matings out of 30 encounters: the WSF/RBM cross resulted in only 7 matings out of 30 encounters. While the difference is apparently quite substantial, it is difficult to set confidence limits on the difference or on the odds ratio because the Bernoulli observations in Table 14.4 are not independent. Even if there is no time trend it would be a gross over-simplification to assume that there is no variation among individual males or among females or to assume that the six observations on any one animal are independent. A binomial comparison is inappropriate

Table 14.6. *Observed matings: Fall '86*

Females		Date					
		Oct. 4	Oct. 8	Oct. 12	Oct. 16	Oct. 20	Oct. 24
	1	1	1	1	0	1	1
	2	0	0	0	1	0	0
RBF	3	RB 1	WS 1	RB 1	WS 0	RB 1	WS 1
	4	1	0	1	1	0	0
	5	0	1	1	0	1	0
	6	0	1	1	1	1	1
	7	0	1	0	0	0	0
RBF	8	WS 1	RB 1	WS 1	RB 1	WS 1	RB 0
	9	0	0	1	0	1	1
	10	0	1	1	1	1	1
	1	0	0	0	1	0	1
	2	0	1	0	1	0	1
WSF	3	RB 1	WS 0	RB 1	WS 0	RB 0	WS 1
	4	1	1	1	0	0	0
	5	0	1	1	1	0	1
	6	0	0	1	0	0	0
	7	1	0	1	0	0	0
WSF	8	WS 0	RB 0	WS 1	RB 0	WS 1	RB 0
	9	1	0	0	0	1	0
	10	0	0	1	0	1	0

and potentially misleading here because the variability in the totals is undoubtedly in excess of that predicted by the binomial model.

14.5.3 *A linear logistic model with random effects*

It is convenient to begin by considering a model for the observed data conditionally on the actual animals used in the experiment. From this conditional perspective it is reasonable to suppose that the 120 observations in Table 14.4 are independent but not identically distributed. For the sake of simplicity we assume here that there is no time trend and that all effects are additive on the logistic scale. The conditional fixed-effects model formula is therefore

individual female effect + individual male effect + cross effect.
(14.10)

The male and female effects are factors whose 20 levels identify the animals, and the cross effect has the four levels R/R, R/W, W/R and W/W in the order female/male. The cross effect is partly

aliased with the sum of the male and female effects, and hence not all cross-effect contrasts are estimable in (14.10).

Table 14.7. *Observed number of matings in an incomplete crossed design (Summer '86)*

	Males			
Females	RBM 1–5	WSM 1–5	RBM 6–10	WSM 6–10
RBF 1–5	15	12	—	—
WSF 6–10	5	12	—	—
RBF 6–10	—	—	7	8
WSF 1–5	—	—	2	9

We do not propose to fit the fixed-effects model for a number of reasons. First, most of the contrasts of interest are not estimable. Second, some of the individual effects may be estimated as $\pm\infty$. Rough Butt females 2 and 5 had success rates of 6/6, so their effects are estimated as $+\infty$ on the logistic scale. Third, the individual animal effects are of scientific importance only to the extent that they mimic the populations from which they are drawn. The main advantage of considering (14.10) explicitly is that the random-effects model of interest is a sub-model of (14.10). Consequently the statistic that is sufficient for the fixed-effects model must also be sufficient for the random-effects model.

In order to make further progress or to draw any interesting scientific conclusions it is necessary to make further assumptions regarding the relationship between the experimental animals and the populations that they are supposed to represent. We shall assume here that the experimental animals are, either in fact or in effect, a random sample from their respective populations. The individual effects in (14.10) are therefore regarded as random variables with variances σ_M^2 and σ_F^2. The male and female variances are assumed constant across populations.

The sufficient statistic, **S**, for the fixed-effects model is the set of totals for each animal and cross type. The values for Summer '86 are shown in Table 14.8.

In what follows we apply the method of quasi-likelihood either to the vector **S** or to the original observations Y. We use unconditional expectations and variances, and denote the four probabilities of interest by π_{RR}, π_{RW}, π_{WR} and π_{WW}. If we denote by Y_{ij} the observation corresponding to female i and male j, then $E(Y_{ij})$ is

Table 14.8 *Components of the sufficient statistic* **S**: *Summer '86*

Animal	Random effects				Systematic effects		
	RBF	WSF	RBM	WSM	Observed		Fitted
1	5	4	5	4	R/R	22	22
2	6	1	3	5	R/W	20	20
3	5	1	5	6	W/R	7	7
4	5	4	3	5	W/W	21	21
5	6	1	4	4			
6	5	1	2	3			
7	1	5	1	1			
8	3	3	2	5			
9	4	5	3	4			
10	2	3	1	4			
Fitted	4.2	2.8	2.9	4.1			
s^2	2.84	2.84	2.10	1.88			

equal to one of these four probabilities according to the populations from which the male and female are drawn. (At no point in the experiment was the same pair observed twice, so this notation introduces no ambiguity.) Similarly, the components of $E(\mathbf{S})$ are linear functions of the four probabilities.

The unconditional covariance matrix of Y_{ij} is assumed to have the pattern implied by (14.10), namely that observations on non-overlapping pairs of animals are independent. Selected elements of the covariance matrix are assumed to have the following form:

$$V_{ij,kl} = \begin{cases} \pi_{\mathrm{RR}}(1 - \pi_{\mathrm{RR}}) & \text{if } i = k, \, j = l, \text{ both of type R;} \\ \sigma_{\mathrm{F}}^2 g^2(\pi_{\mathrm{RR}}) & \text{if } i = k\,[= \mathrm{R}]; \, j \neq l, \text{ both R;} \\ \sigma_{\mathrm{M}}^2 g(\pi_{\mathrm{RW}})g(\pi_{\mathrm{WW}}) & \text{if } i \neq k\,[= \mathrm{R}, \mathrm{W}]; \, j = l\,[= W]; \\ 0 & \text{if } i \neq k, \, j \neq l. \end{cases}$$

$$(14.11)$$

and so on for some known function $g(\cdot)$. Other non-zero components have the form

$$\pi_{\mathrm{RW}}(1 - \pi_{\mathrm{RW}}), \quad \sigma_{\mathrm{F}}^2 g^2(\pi_{\mathrm{RW}}), \quad \sigma_{\mathrm{M}}^2 g^2(\pi_{\mathrm{RW}}), \quad \sigma_{\mathrm{F}}^2 g(\pi_{\mathrm{RR}})g(\pi_{\mathrm{RW}}), \ldots,$$

depending on whether the two observations are identical, have one Rough Butt female in common, one Whiteside male in common, one Rough Butt female in common, and so on. This covariance

matrix with $g(\pi) = \pi(1 - \pi)$ can be obtained via a Taylor approximation based on (14.10) provided that σ_M^2 and σ_F^2 are sufficiently small. For details see Exercise 14.2.

With this covariance matrix it can be shown that the quasi-likelihood estimates of the four probabilities are

$$\hat{\pi}_{RR} = \sum_{R/R} Y_{ij}/30, \qquad \hat{\pi}_{RW} = \sum_{R/W} Y_{ij}/30,$$

$$\hat{\pi}_{WR} = \sum_{W/R} Y_{ij}/30 \quad \text{and} \quad \hat{\pi}_{WW} = \sum_{W/W} Y_{ij}/30, \tag{14.12}$$

as indicated in Table 14.8. For a sketch of the derivation see Exercises 14.8–9. The remarkable aspect of this is that the quasi-likelihood estimates do not depend on the values of σ_F^2 or σ_M^2 though their distributions do. Provided that the covariance matrix has the form (14.11) the quasi-likelihood estimates are the same as if the two dispersion components were zero.

Since the parameter estimates are linear functions of the data, the covariance matrix of $\hat{\pi}$ may be obtained in terms of the dispersion components using (14.11). We find after some considerable algebraic reduction that

$$30\,\mathrm{cov}(\hat{\boldsymbol{\pi}}) = \Pi(\mathbf{I} - \Pi)$$

$$+ \mathbf{G} \begin{pmatrix} 2(\sigma_F^2 + \sigma_M^2) & 3\sigma_F^2 & 3\sigma_M^2 & 0 \\ 3\sigma_F^2 & 2(\sigma_F^2 + \sigma_M^2) & 0 & 3\sigma_M^2 \\ 3\sigma_M^2 & 0 & 2(\sigma_F^2 + \sigma_M^2) & 3\sigma_F^2 \\ 0 & 3\sigma_M^2 & 3\sigma_F^2 & 2(\sigma_F^2 + \sigma_M^2) \end{pmatrix} \mathbf{G}$$

$$\tag{14.13}$$

where $\mathbf{G} = \mathrm{diag}\{g(\pi_{RR}), g(\pi_{RW}), g(\pi_{WR}), g(\pi_{WW})\}$ and $\Pi = \mathrm{diag}\{\boldsymbol{\pi}\}$. A similar but slightly more complicated expression can be obtained if the exact covariance matrix is used in place of (14.11). For example the exact variance of $30\sqrt{\hat{\pi}_{RR}}$ is

$$\pi_{RR}(1 - \pi_{RR}) + \mathrm{cov}\big(F(\alpha_{RR} + \epsilon_1 + \delta_1), F(\alpha_{RR} + \epsilon_1 + \delta_2)\big)$$
$$+ \mathrm{cov}\big(F(\alpha_{RR} + \epsilon_1 + \delta_1), F(\alpha_{RR} + \epsilon_2 + \delta_1)\big),$$

while the covariance of $\sqrt{30}\hat{\pi}_{RR}$ and $\sqrt{30}\hat{\pi}_{RW}$ is

$$3\,\mathrm{cov}\big(F(\alpha_{RR} + \epsilon_1 + \delta_1), F(\alpha_{RW} + \epsilon_1 + \delta_2)\big), \tag{14.14}$$

where $F(\cdot)$ is the cumulative logistic function, and ϵ_i, with variance σ_{F}^2, and δ_i, with variance σ_{M}^2, are the female and male random effects. Only the first term of the Taylor expansions is included in (14.13).

The quasi-likelihood estimates are exactly the maximum likelihood estimates obtained from the linear logistic model containing only the four-level factor 'cross type'. The covariance matrix obtained from this analysis is the first term in (14.13). The second term arises entirely as a result of covariances among the observations, and could conceivably dominate the first term.

14.5.4 *Estimation of the dispersion parameters*

From the approximate covariance matrix (14.11) we see that the covariance matrix of $\{Y_{ij}\}$ is expressible in the form

$$\mathbf{V}_0(\boldsymbol{\pi}) + \sigma_{\mathrm{F}}^2 \mathbf{V}_1(\boldsymbol{\pi}) + \sigma_{\mathrm{M}}^2 \mathbf{V}_2(\boldsymbol{\pi}).$$

Consequently the expected value of any quadratic form in the residuals is linear in the dispersion components $\sigma_{\mathrm{F}}^2, \sigma_{\mathrm{M}}^2$. For purposes of estimation the easiest way to proceed is to choose two suitable quadratic forms and to equate the observed values of these to their expectations as functions of $\sigma_{\mathrm{M}}^2, \sigma_{\mathrm{F}}^2$. This gives a pair of simultaneous linear equations for the dispersion components. For reasons given in the previous section such quadratic forms should be functions of the sufficient statistic $\mathbf{S}$.

In Table 14.8 the expected value of each entry in column 1 is $3\pi_{\mathrm{RR}} + 3\pi_{\mathrm{RW}}$, estimated as 4.2, which is the column mean. Similar calculations apply to the next three columns. The variance of each entry in column 1 is

$$\begin{aligned}
\mathrm{var}(S_1) = {} & 3\pi_{\mathrm{RR}}(1 - \pi_{\mathrm{RR}}) + 3\pi_{\mathrm{RW}}(1 - \pi_{\mathrm{RW}}) \\
& + 6\sigma_{\mathrm{F}}^2\{\pi_{\mathrm{RR}}^2(1 - \pi_{\mathrm{RR}})^2 + \pi_{\mathrm{RW}}^2(1 - \pi_{\mathrm{RW}})^2 + 3\pi_{\mathrm{RR}}(1 - \pi_{\mathrm{RR}})\pi_{\mathrm{RW}}(1 - \pi_{\mathrm{RW}})\}
\end{aligned}$$

which we denote by $\kappa_2(S)$. Within column 1 some of the covariances are zero: the non-zero covariances are not all equal because some pairs of females have one male in common and others two. The sum of the covariances between S_i and S_j for $i \neq j$ in column 1 is

$$90\bar{\kappa}_{11}(S) = 60\sigma_{\mathrm{M}}^2\{\pi_{\mathrm{RR}}^2(1 - \pi_{\mathrm{RR}})^2 + \pi_{\mathrm{RW}}^2(1 - \pi_{\mathrm{RW}})^2\}.$$

It follows from the result established in Exercise 14.6 that the expected mean square for column 1 is

$$E(\text{mean square}) = \kappa_2(S) - \bar{\kappa}_{11}(S),$$

which is estimated as

$$1.2533 + 1.3080\sigma_{\text{F}}^2 - 0.0584\sigma_{\text{M}}^2.$$

Similar calculations for the second column give an expected mean square of

$$1.1667 + 1.1328\sigma_{\text{F}}^2 - 0.0507\sigma_{\text{M}}^2.$$

The estimates used here are based on the pooled mean square for females, and the pooled mean square for males. Details of the calculations are shown in Table 14.9.

Table 14.9 *Estimation of the dispersion components from* S

Source	Mean square	E(Mean square)	$\tilde{\sigma}^2$
RBF	2.8444	$1.2533 + 1.3080\sigma_{\text{F}}^2 - 0.0584\sigma_{\text{M}}^2$	
WSF	2.8444	$1.1667 + 1.1328\sigma_{\text{F}}^2 - 0.0507\sigma_{\text{M}}^2$	
Total F	5.6888	$2.4200 + 2.4408\sigma_{\text{F}}^2 - 0.1091\sigma_{\text{M}}^2$	$1.3704 = \tilde{\sigma}_{\text{F}}^2$
RBM	2.1000	$1.1233 + 1.0511\sigma_{\text{M}}^2 - 0.0468\sigma_{\text{F}}^2$	
WSM	1.8778	$1.2967 + 1.4009\sigma_{\text{M}}^2 - 0.0623\sigma_{\text{F}}^2$	
Total M	3.9778	$2.4200 + 2.4520\sigma_{\text{M}}^2 - 0.1091\sigma_{\text{F}}^2$	$0.6963 = \tilde{\sigma}_{\text{M}}^2$

The results of similar calculations for the Fall '86 experiments are summarized in Table 14.10. Evidently the mating probabilities are quite consistent across the three experiments, $\hat{\pi}_{\text{WR}}$ being consistently lower than the other three probabilities. The estimated dispersion components are similar for the first two experiments, but the pattern is reversed for the third experiment. It should be borne in mind, however, that the variability of these estimates is appreciable, and negative estimates are not impossible. Because of correlations among the components of S it is difficult to assign degrees of freedom to the sums of squares in Table 14.9.

Using the pooled estimates of all parameters together with (14.13), we find that $90\,\text{cov}(\hat{\pi})$ is estimated as

$$\begin{pmatrix} 0.2222 & & & \\ & 0.2469 & & \\ & & 0.1665 & \\ & & & 0.2222 \end{pmatrix} + \begin{pmatrix} 0.1772 & 0.1506 & 0.0977 & \\ 0.1506 & 0.2188 & & 0.1448 \\ 0.0977 & & 0.0995 & 0.1015 \\ & 0.1448 & 0.1015 & 0.1772 \end{pmatrix}.$$

Table 14.10 *Summary of parameter estimates from three experiments.*

Experiment	Parameter estimate				Dispersion component	
	$\hat{\pi}_{RR}$	$\hat{\pi}_{RW}$	$\hat{\pi}_{WR}$	$\hat{\pi}_{WW}$	$\tilde{\sigma}_F^2$	$\tilde{\sigma}_M^2$
Summer '86	0.7333	0.6667	0.2333	0.7000	1.3704	0.6963
Fall '86 (re-run)	0.6000	0.4667	0.2333	0.6667	0.9787	0.5997
Fall '86	0.6667	0.5333	0.1667	0.6333	0.3954	1.3440
Pooled estimate	0.6667	0.5556	0.2111	0.6667	0.9148	0.8800

The random effects account for about half of the total variability
in the parameter estimates, although not all contrasts among the
$\hat{\pi}$s are affected equally (Exercise 14.12). The pooled estimate of
the mixed contrast is $\hat{\pi}_{RW} - \hat{\pi}_{WR} = 0.3445$ with estimated standard
error 0.0904. Thus the evidence for a non-zero mixed contrast is
evidently very strong. From this analysis there is no evidence of
differences among the probabilities π_{RR}, π_{RW} and π_{WW}.

Despite the fact that the magnitudes of the random effects as
estimated in Table 14.10 are approximately equal for the two sexes
it appears from closer inspection that the nature of the effect for
males is quite different than for females. A comparison of the
two sets of animal totals for the Summer '86 and the first Fall '86
experiment shows a strong correlation for males but no evidence
of any correlation for females. Thus it appears that the male
random effects persist over several months at least whereas the
female effects are short-lived. This conclusion applies to both male
populations and both female populations.

14.6   Bibliographic notes

Linear models in which there is more than one variance component
have been used for many years going back at least to the work of
Fisher and Yates in agricultural experiments. For detailed accounts
of specific designs see Kempthorne (1952, Chapter 9) or Cox (1958c,
Chapters 7,8). For the estimation of variance components more
generally in cases where the design is unbalanced, the method
of restricted maximum likelihood is usually preferred to ordinary
maximum likelihood. For an account of this see section 7.2,
Patterson and Thompson (1971) or Harville (1974, 1977).

There is a parallel but extensive literature on educational test-

ing, which uses ordinary maximum-likelihood estimation for unbalanced Normal-theory linear models. See, for example, Goldstein (1986) or Bock (1989).

Although the need has long been recognized by practising statisticians, the development of analogous models and techniques of estimation for non-Normal data or for non-linear effects has proceeded very slowly. Some notable exceptions are Stiratelli, Laird and Ware (1984), who use a Bayesian argument, with a diffuse prior on the regression coefficients, for the estimation of the dispersion parameters. Gilmour, Anderson and Rae (1985) discuss a random-effects probit model. For Normal-theory linear models the previous two methods reduce to restricted maximum likelihood. Anderson and Aitkin (1985) develop an unmodified maximum-likelihood estimation procedure for nested variance components in linear logistic and probit models. Their likelihood is obtained by direct numerical integration assuming Normal random effects, and the resulting estimates are not the same as those suggested here.

In the previous three papers the effects of interest are defined as linear contrasts of the *conditional* logits. In section 14.5.3, however, the effects of interest are defined in terms of the *unconditional* probabilities or logits: the random-effects model (14.10) is used only to justify the choice of covariance matrix. A qualitatively similar argument is given by Prentice (1988). Zeger, Liang and Self (1985) and Liang and Zeger (1986) also argue in favour of specifying models for the unconditional rather than the conditional probabilities, but the choice must ultimately be governed by what the interesting parameters are in any given context. For further discussion regarding the distinction between subject-specific models such as (14.10) and population-averaged models corresponding to the estimates (14.12), see Zeger, Liang and Albert (1988).

Morton (1987) considers models for counted data in which the random effects are nested and effects are multiplicative. His method of estimation, using quasi-likelihood estimating functions, is very similar to that used here. The salamander example is rather unusual in that the random effects are crossed rather than nested.

Despite the apparent paucity of references, this subject is extremely important with a broad range of applications. At the time of writing it is evident that the topic is a rich source of good research problems.

14.7 Further results and exercises 14

14.1 Consider the vector **S** whose 44 components are displayed in Table 14.8. Show that $T_1 = S_1 - S_{RR}/10 - S_{RW}/10$ is uncorrelated with S_{RR}, S_{RW}, S_{WR} and S_{WW}, and that $E(T_1) = 0$. Hence deduce that the quasi-likelihood estimates of the parameters are given by (14.12).

14.2 Suppose that ϵ is a Normal random variable with mean zero and variance σ^2. Define the following random variable and its expectation:

$$P = \frac{\exp(\alpha + \epsilon)}{1 + \exp(\alpha + \epsilon)}, \quad \text{and} \quad \pi = E(P).$$

Justify empirically the approximation $\pi \simeq F(\alpha^*)$, where $F(x) = e^x/(1 + e^x)$, and

$$\alpha^* = \alpha - \tfrac{1}{2}\sigma^2 \tanh\{\alpha(1 + 2\exp(-\sigma^2/2))/6\}.$$

For $\sigma^2 < 2$, the maximum error of this approximation is about 0.003 on the probability scale. Show that the approximation has the correct limiting behaviour in the limit as $\alpha \to \pm\infty$ for fixed σ^2.

By differentiating the above approximation with respect to α, show that the variance of P is approximately

$$\text{var}(P) \simeq \sigma^2 \pi(1 - \pi)\pi^\dagger(1 - \pi^\dagger) \times \tfrac{1}{3}(1 + 2\exp(-\sigma^2/2)),$$

where $\pi^\dagger = F\big(\alpha(1 + 2\exp(-\sigma^2/2))/3\big)$. This approximation is considerably more accurate than the Taylor approximation in (14.11).

14.3 Suppose that ϵ_1, ϵ_2 are Normal random variables with variances σ_1^2, σ_2^2, and correlation ρ. Show that the correlation ρ' of $F(\epsilon_1)$ and $G(\epsilon_2)$ satisfies

$$\frac{|\rho|}{|\rho'|} \simeq 1 + \tfrac{1}{4}\left\{\sigma_1^2\left(\frac{F''}{F'}\right)^2 + \sigma_2^2\left(\frac{G''}{G'}\right)^2 - 2\rho\sigma_1\sigma_2\frac{F''G''}{F'G'}\right\} \geq 1$$

when terms of order σ^4 and smaller are ignored. Under what conditions is ρ' equal to ρ?

14.4 Using the notation of the previous Exercise, define the random variables P_1, P_2, together with their expectations as follows:

$$P_1 = \frac{\exp(\alpha_1 + \epsilon_1)}{1 + \exp(\alpha_1 + \epsilon_1)}, \qquad P_2 = \frac{\exp(\alpha_2 + \epsilon_2)}{1 + \exp(\alpha_2 + \epsilon_2)},$$
$$\pi_1 = E(P_1), \qquad\qquad \pi_2 = E(P_2).$$

Using the results given in the previous Exercises, find an approximation for the covariance of P_1 and P_2.

Compute the exact covariance numerically for variances and covariances in the range 0–2. Comment on the adequacy of the approximate formula.

14.5 Compute the components of the sufficient statistic **S** for the first fall '86 experiment (data in Table 14.5). Compare the individual animal totals with those in Table 14.8. Compute the correlation coefficients or regression coefficients of the Fall totals on the Summer totals for each of the four groups. Comment on your findings.

14.6 Show that if $Y_1, \ldots, Y_n$ have common mean μ and covariance matrix $\kappa_{i,j}$, then the expected value of

$$s^2 = \frac{1}{n-1} \sum (Y_i - \bar{Y})^2$$

is given by

$$E(s^2) = \frac{1}{n} \sum_i \kappa_{i,i} - \frac{1}{n(n-1)} \sum_{i \neq j} \kappa_{i,j}.$$

14.7 The pooled estimate of the dispersion components in Table 14.10 is the average estimate for the three experiments. By pooling together the sums of squares for F and M from the three experiments and equating these totals to their expectations, show that the pooled estimates are $\tilde{\sigma}_F^2 = 0.9035$ and $\tilde{\sigma}_M^2 = 0.8759$.

14.8 Suppose that the random vector **Y** satisfies the linear model

$$E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}, \qquad \text{cov}(\mathbf{Y}) = \mathbf{V}.$$

Let $\tilde{\beta}$ be the ordinary least-squares estimate $\tilde{\beta} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$ and let $\mathbf{R}$ be the residual vector

$$\mathbf{R} = (\mathbf{I} - \mathbf{P})\mathbf{Y} = \mathbf{Y} - \mathbf{X}\tilde{\beta}.$$

Let $\mathcal{X}$ denote the column space of $\mathbf{X}$. Show that if $\mathbf{V}x$ lies in $\mathcal{X}$ for each $x \in \mathcal{X}$ then $\mathbf{R}$ and $\tilde{\beta}$ are uncorrelated and

$$\tilde{\beta} = \hat{\beta} = (\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}\mathbf{Y},$$

where $\mathbf{W} = \mathbf{V}^{-1}$. Deduce also that $\mathbf{V}^{-1}x \in \mathcal{X}$. [Kruskal, 1968].

14.9 Using the result established in the previous exercise show that the maximum quasi-likelihood estimates of the parameters in the salamander example are given by (14.12) provided that the experiment is suitably balanced. Obtain the required balance condition.

14.10 Show that if Y has the gamma distribution $G(\mu, \nu)$, then

$$\mathrm{var}\{\log(Y)\} = \psi'(\nu),$$

where $\psi(x) = \Gamma'(x)/\Gamma(x)$. Under what conditions is the approximation $\mathrm{var}(\log Y) \simeq \mathrm{cv}^2(Y)$ adequate for the gamma family?

14.11 Repeat the calculations of the previous exercise under the assumption that Y has the log-Normal distribution, $\log(Y) \sim N(\mu, \sigma^2)$.

14.12 Consider the model formula (14.10) written in matrix notation in the form

$$\mathbf{Z}\mathbf{b} + \mathbf{X}\boldsymbol{\pi},$$

in which $\mathbf{Z}$ is a (120×40) incidence matrix identifying the animals paired together in each trial, and $\mathbf{X}$ is a (120×4) incidence matrix for the cross type, R/R, R/W, W/R and W/W. Show that the matrix $(\mathbf{I} - \mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^-\mathbf{Z}^T)\mathbf{X}$ has rank 1, with span corresponding to the contrast R/R − R/W − W/R + W/W. Hence justify the claim made in section 14.5.3 that most of the effects of interest are not estimable in (14.10).

Using (14.12) and (14.13) find an estimate of the above contrast and obtain a numerical estimate of the standard error, (i) under the assumption that the between-animal variances are zero, and (ii) using the estimates given in Table 14.10. Explain briefly why the standard error in (ii) is smaller than in (i).