COMPOSITIONAL HYPOTHESES OF SUBCOMPOSITIONAL STABILITY AND SPECIFIC PERTURBATION CHANGE AND THEIR TESTING

J. Aitchison
Department of Statistics
University of Glasgow
Glasgow G12 8QQ
Scotland
Email: john.aitchison@btinternet.com

K. W. Ng
Department of Statistics and Actuarial Science
University of Hong Kong
Hong Kong
People's Republic of China
Email: kaing@hku.hk

Abstract

In standard multivariate statistical analysis common hypotheses of interest concern changes in mean vectors and subvectors. In compositional data analysis it is now well established that compositional change is most readily described in terms of the simplicial operation of perturbation and that subcompositions replace the marginal concept of subvectors. To motivate the statistical developments of this paper we present two challenging compositional problems from food production processes. Against this background the relevance of perturbations and subcompositions can be clearly seen. Moreover we can identify a number of hypotheses of interest involving the specification of particular perturbations or differences between perturbations and also hypotheses of subcompositional stability. We identify the two problems as being the counterpart of the analysis of paired comparison or split plot experiments and of separate sample comparative experiments in the jargon of standard multivariate analysis. We then develop appropriate estimation and testing procedures for a complete lattice of relevant compositional hypotheses.

1. The problems

A common challenge in compositional data analysis is to attempt to characterise change in compositions and to be able to test hypotheses concerning the nature of the change. We use two easily described practical problems in food production to illustrate this special form of compositional problem, describe a simple but effective characterisation of change, and illustrate how we may test a variety of compositional hypotheses within the structure of a perturbation hypothesis lattice.

Problem 1. Cow milk production
In an attempt to improve the quality of cow's milk, milk from each of thirty cows was assessed by dietary composition before and after a strictly controlled dietary and hormonal regime over a period of eight weeks. Although seasonal variations in milk quality could probably be regarded as negligible over this period a control group of thirty cows was kept under the same conditions but on the standard regime. The sixty cows were of course allocated to control and treatment groups at random. Table 1 provides a set of typical before and after results for ten cows, five in the control group and five in the treatment group, showing the protein, milk fat, carbohydrate, calcium, sodium and potassium proportions by weight of total dietary content. The full data set is obtainable in Appendix Table 1. The purpose of the experiment is to determine whether the new regime has produced any significant change in the milk composition so it is essential to have a clear idea of how change in compositional data is

characterised by a group operation termed perturbation. We shall see later how a whole series of perturbation hypotheses emerge and how we may investigate the full lattice of such hypotheses. Meanwhile we note that because of the before and after nature of the data within each experimental unit we have for compositional data the analogue of a paired comparison situation for real measurements, where traditionally the differences in pairs of measurements are considered. We have thus to find the counterpart of difference for such paired compositions.

TABLE 1
Typical before and after dietary compositions of the milk of ten cows, C1-C5 in the control group, T1-T5 in the treatment group

	Before						After					
Cow no	pr	mf	ch	Ca 10^{-2} x	Na 10^{-2} x	K 10^{-2} x	pr	mf	ch	Ca 10^{-2} x	Na 10^{-2} x	K 10^{-2} x
C1	0.310	0.224	0.441	1.03	0.25	1.27	0.258	0.306	0.411	1.05	0.21	1.28
C2	0.268	0.369	0.338	0.84	0.30	1.44	0.238	0.395	0.336	1.12	0.30	1.68
C3	0.258	0.339	0.375	1.74	0.47	1.57	0.241	0.329	0.398	0.93	0.47	1.79
C4	0.245	0.261	0.462	0.90	0.90	1.40	0.288	0.246	0.434	1.08	0.63	1.49
C5	0.371	0.148	0.451	0.98	0.32	1.63	0.440	0.125	0.405	1.09	0.28	1.69
T1	0.327	0.196	0.450	0.68	0.83	1.23	0.357	0.178	0.436	1.07	0.85	1.13
T2	0.326	0.172	0.427	0.71	0.57	1.25	0.506	0.104	0.361	1.07	0.63	1.29
T3	0.247	0.330	0.392	0.86	0.59	1.56	0.363	0.245	0.362	0.97	0.60	1.37
T4	0.262	0.272	0..434	0.90	0.54	1.69	0.351	0.204	0.418	1.16	0.27	1.25
T5	0.281	0.270	0.423	0.42	1.08	1.12	0.225	0.303	0.442	0.71	1.16	1.21

Notation: pr = protein, mf = milk fat, ch =carbohydrate, Ca = calcium, a = sodium, K = potassium

Problem 2. Chicken carcass compositions

A study of chicken carcasses has been conducted to investigate the nature of the changes which might be encountered when chickens are subjected to a new diet regime. By the nature of the experiment it is not possible to determine the carcass composition of a chicken both before and after the treatment. In the study 60 chickens were randomly divided into three groups of 20 each. Group 1 was sacrificed at the beginning of the study and the twenty carcass compositions determined. Group 2 continued on the existing diet while group 3 was placed on the special diet, otherwise being kept under identical conditions. Both groups 2 and 3 were then sacrificed at the end of the study and their carcass compositions determined. Table 2 shows three typical proportions- by- weight compositions (breast muscle, other muscle, fat, skin, bone) for each group and the complete data set is given in Appendix Table 2. In contrast to Problem 1 we have here a situation which corresponds to separate sample

comparisons.

Before we proceed to the modelling and analysis of these problems and data sets it is essential to understand the nature of the relevant simplex sample space and the characterisation of change within the space.

TABLE 2
Typical chicken carcass compositions of the three groups of chickens

Chicken no	Composition (proportions by weight) breast muscle	other muscle	fat	skin	bone
A1	0.281	0.212	0.041	0.106	0.360
A2	0.375	0.159	0.036	0.102	0.328
A3	0.447	0.134	0.024	0.086	0.309
B1	0.396	0.163	0.024	0.085	0.332
B2	0.405	0.155	0.026	0.089	0.325
B3	0.408	0.143	0.024	0.116	0.309
C1	0.517	0.123	0.020	0.102	0.238
C2	0.365	0.197	0.020	0.127	0.291
C3	0.527	0.118	0.017	0.106	0.232

2. The role of perturbation in compositional data analysis

Statisticians are so familiar with the operation of displacement or translation, essentially vector addition, in D-dimensional real sample space R^D that they hardly require to think about its properties or indeed be aware that it is the basic internal group operation on which most of multivariate statistical analysis is based. A second operation, namely scalar multiplication also plays an important role. For example for a translation t and for a scalar multiple a on a random vector x, the mean and covariance properties,

$$E(x+t) = E(x)+t, \quad E(ax) = aE(x), \quad V(x+t) = V(x), \quad V(ax) = |a|^2 V(x),$$

are fundamental to unconstrained multivariate analysis. A similar situation exists in directional data analysis with the unit sphere as sample space, where the fundamental group operation of rotation plays a central role. For example, in the study of the movement of tectonic plates, it was recognition that the group of rotations on the sphere plays a central role and the use of a satisfactory representation of that group that led Chang (1988) to the production of the essential statistical tool for spherical regression.

Since the sample space associated with *D*-part compositions is the unit simplex:

$$S^D = \{[x_1, \ldots, x_D]: \quad x_i > 0 \ \ (i = 1, \ldots, D), \quad x_1 + \ldots + x_D = 1\},$$

we have to ask what is the basic group operation in this space which plays the counterpart of translation in real space and rotation on the sphere, and so can be used to characterise compositional change. The answer is in the group operation of perturbation (Aitchison, 1982, Section 3; 1986, Section 2.8) defined as follows. Given two *D*-part compositions $x, y \in S^D$ the perturbation $x \oplus y$ is defined by

$$x \oplus y = [x_1 y_1, \ldots, x_D y_D] / (x_1 y_1 + \ldots + x_D y_D) = C[x_1 y_1, \ldots, x_D y_D],$$

where the 'closure operator' C standardises the contained vector by dividing by the sum of its components so that the components sum to unity. It is trivial to show that the operation $\oplus$ defines an abelian group on the simplex with identity $e = (1/D)[1, \ldots, 1]$ and inverse $x^{-1} = C[1/x_1, \ldots, 1/x_D]$. These properties are important because for the solution of our practical problems we clearly have to be able to characterise the operation which changes a *D*-part composition *x* into a *D*-part composition *X*, in other words what is the perturbation *p* such that $X = p \oplus x$? The answer is clearly in the inverse operation

$$p = X \ominus x = C[X_1 / x_1, \ldots, X_D / x_D].$$

Note that in our choice of notation here we have used the symbol $\oplus$ to emphasise the analogue with vector addition in real space.

It is worth noting here, although we shall not make any substantial use of these concepts, that a further operation $\otimes$ of powering, analogous to scalar multiplication in real space and a simplicial metric can be introduced providing the statistician with an algebraic-structure on the simplex sample space – a metric vector space – analogous to that available for standard multivariate analysis. The formal definitions are as follows. Given a *D*-part composition $x \in S^D$ and a real number $a \in R^1$ the power transformed composition is

$$a \otimes x = C[x_1^a, \ldots, x_D^a].$$

The simplicial metric $\Delta_S : S^D \times S^D \to R_{\geq 0}$ defined by Aitchison (1983; 1986, p.193) is given by

$$\Delta_S(x, y) = \left[\sum_{i=1}^{D} \left\{ \log \frac{x_i}{g(x)} - \log \frac{y_i}{g(y)} \right\}^2 \right]^{1/2} \quad (x, y \in S^D),$$

where $g(.)$ denotes the geometric mean of the components of the enclosed vector.

We need a further compositional concept for investigation of our food compositional problems, that of subcomposition . As in standard multivariate analysis marginal concepts are important. For compositions and the simplex the marginal concept is a subcomposition, such as the chemical (Ca, Na, K)-subcomposition of a full milk composition in Problem 1. For example the (1, . . . *C*)-subcomposition of a *D*-part composition $[x_1, \ldots, x_D]$ is defined as

$$[s_1, \ldots, s_C] = \mathrm{C}[x_1, \ldots, x_C] = [x_1, \ldots, x_C] / (x_1 + .. + x_C).$$

For statistical modelling of compositional variation we have to consider distributions on the simplex and their characteristics. The well-established 'measure of central tendency' $x \in S^D$ is the distributional 'centre'

$$x = cen(x) = C(\exp(E(\log x))).$$

Conforming with this mean value there is a variety of equivalent forms of dispersion and covariance characteristics; see Aitchison (1986, Chapters 4 and 5). Which of these equivalent forms is used will depend on the particular nature of the application. In our perturbation analysis here it will be convenient to use the variation matrix $\mathrm{T}(x)$ whose (i, j)th component is $\mathrm{var}[\log(x_i / x_j)]$. We note here that these characteristics have properties analogous to the mean and variance matrix properties cited for real space above:

$$cen(x \oplus p) = cen(x) \oplus p, \quad cen(a \otimes x) = a \otimes cen(x), \quad \mathrm{T}(x \oplus p) = \mathrm{T}(x), \quad \mathrm{T}(a \otimes x) = |a|^2 \mathrm{T}(x)$$

We shall require to use parametric classes of distributions on the simplex such as the logistic-normal class in Aitchison and Shen (1980) and Aitchison (1986, Chapter 6) in our testing processes. These can be parameterised in terms of the centre x and variation matrix T as defined above. We shall use the notation $L^D(x, \mathrm{T})$ to identify a logistic normal distribution with centre x and variation matrix T.

These results have consequences for estimation problems for compositional data, summarised as follows. In what follows we shall be concerned with compositional data sets, typically an $N \times D$ matrix X with nth row composition x_n. First we note that the estimate $\hat{x}$ of x is given by

$$\hat{x} = C[g_1, \ldots, g_D],$$

where the g's are the geometric means of the individual components. Moreover the variance elements of $\mathrm{T} = [t_{ij}]$ are simply estimated in a standard way by

$$\hat{t_{ij}} = (N-1)^{-1} \left[\sum_{n=1}^{N} \{\log(x_{ni} / x_{nj})\}^2 - N^{-1} \left\{ \sum_{n=1}^{N} \log(x_i / x_j) \right\}^2 \right].$$

See Aitchison and others (2003) for further details of the structure of the simplex.

3. The paired comparison lattice

3.1 Testing perturbation hypotheses. The data of Appendix Table 1 are of a before- and after-nature. Each cow has had milk composition determined at the beginning and at the end of the trial and so we have essentially, in standard statistical analysis terms, paired comparisons. The major difference is that we require to use a measure of difference appropriate to compositional change and we have seen this to be perturbation. Thus for cow C1 of Table 1 we have the before and after milk compositions

$$x_B = [0.310 \;\; 0.224 \;\; 0.441 \;\; 0.0103 \;\; 0.0025 \;\; 0.0127],$$
$$x_A = [0.258 \;\; 0.306 \;\; 0.411 \;\; 0.0105 \;\; 0.0021 \;\; 0.0128],$$

so that the required perturbation is

$$p = x_B \,\Theta\, x_A = [0.139 \;\; 0.228 \;\; 0.155 \;\; 0.170 \;\; 0.140 \;\; 0.168].$$

We can similarly calculate the perturbations associated with each of the sixty cows in the trial.

As a first approach to analysis we can address the problems that we now face in three stages by posing three questions.

Question 1 Is there any evidence of seasonal change in milk composition, in other words is there any evidence of differences in the milk compositions of the control group between the beginning and end of the trial? Phrased as a compositional hypothesis this is simply a question of whether the centre of the control group perturbations is the identity perturbation. A standard way of testing such a hypothesis is through the logratio analysis of Aitchison (1986). Transformed into logratio terms this is simply asking whether the mean of the additive logratio vectors

$$q = [q_1,\ldots,q_5] = alr(p) = [\log(p_1/p_6),\ldots,\log(p_5/p_6)]$$

is a zero vector, a hypothesis easily tested under standard multivariate analysis. A standard exact test statistic is available for this purpose, namely Hotelling's T-squared: see for example, Anderson (1958, Section 5.3.1). The computed value of this is 56.6 to be compared against percentage points of the $(5(N-1)/(N-5))F(5,25)$ distribution. This comparison shows that the *alr* mean is significantly different from zero at the 0.1 percent significance level and therefore that the centre of the perturbations is significantly different from the identity perturbation. We thus conclude that there is substantial evidence of a seasonal change which justifies the insistence on having a control group. The centre of the control group perturbations is

[0.1595 0.1835 0.1599 0.1818 0.1458 0.1695].

Question 2 Is there similar evidence of a change in the treatment group?. Here the Hotelling T-squared statistic value is even larger, 331.1, again to be compared against

the same percentile value, and so we have real evidence of change, with the centre of the treatment group perturbations being

[0.1928 0.1416 0.1589 0.2309 0.1338 0.1420].

Question 3 The remaining question is to ask whether there are differences between the control and treatment group perturbations and this question can be answered by using for the two samples of perturbations a separate sample lattice identical to that for the hongite-kongite comparison in Aitchison (1986, Section 7.5).

The three Q-statistics in the same order as for the previous example are 153.7, 45.6, 212.0 to be compared against 95 percentiles of the chi-squared distribution at 20, 15, 5 degrees of freedom, all giving significant differences. Thus there is strong evidence of differences between control and treatment changes.

A good indication of what the nature of this change is can be obtained by computing the perturbation difference between the control and treatment perturbation centres, namely

[0.2015 0.1286 0.1656 0.2117 0.1529 0.1397].

Thus we can see that relatively there is enhancement of protein, carbohydrate and calcium, presumably a successful nutritional result.

Though the above analysis is probably adequate for answering the immediate questions posed it is worth pointing out that there is a complete lattice of possible hypotheses concerning the maximum model here which is that the control and treatment perturbations follow $L^6(x_C, \mathrm{T}_C)$ and $L^6(x_T, \mathrm{T}_T)$ distributions. We set out in Figure 1 a fairly complete lattice of hypotheses together with the associated parameter sizes and the corresponding maximised loglikelihoods. Note that there are essentially two routes from the simplest hypothesis of no effects and no differences, namely $x_C = x_T = e,\ \mathrm{T}_C = \mathrm{T}_T$, to the maximum model, by an equal and by an unequal covariance route. Working on the principle that we prefer a simple explanation to a complicated one we can proceed by lattice testing along the lines of Aitchison (1986, Section 7.4). With the maximised loglikelihoods determined we can use generalised likelihood ratio tests, and we would find that we would have to reject all the hypotheses of the lattice in favour of accepting as a working model the maximum model. This is in accordance with our previous finding that there is a seasonal effect but that there are differences between the control and treatment perturbations.

3.2 *Testing hypotheses of subcompositional stability.* While the above analysis was sufficient for the aim of the experiment we can use this example to illustrate another important form of compositional hypothesis, namely subcompositional stability. For example in geology in the study of the major-oxide chemistry of a series of rocks the question may arise as to whether certain oxides stay roughly constant relative to each other, in other words whether the subcomposition of these major oxides is stable. Let us place such a hypothesis within the framework of our milk composition problem. Suppose that it had been suggested that seasonal change would not affect the relative proportions of the minor elements (Ca, K, Na). This suggestion is clearly expressible

as a perturbation hypothesis, namely that the perturbation is of the form $x_C = [x_1\, x_2\, x_3\, x_4\, x_4\, x_4]$ with the last three components corresponding to the minor elements equal. Since $alr(x_C)$ takes the form $[h_1\, h_2\, h_3\, 0\, 0]$ the hypothesis of subcompositional stability is simply a linear hypothesis within standard multivariate analysis and an exact test exists. We can, however apply an easier test. Since the hypothesis refers to the (Ca, K, Na) subcomposition we can confine consideration to the before-after perturbations associated with these subcompositions and ask whether these have an *alr* mean of [0 0]. For the control group this gives a T^2 value of 3.52 to be compared against a 5 per cent critical value of 6.99; for the treatment group the T^2 value is 10.45. Thus a reasonable conclusion is that the (Ca, K, Na) subcomposition is stable against seasonal change but that there is significant instability in the presence of treatment.

4. The separate sample lattice

In the experiment of Problem 2 we cannot ascertain directly the perturbations involved in change of chicken carcass composition. We can, however, simply incorporate the change mechanism of perturbation into our statistical modelling of the experiment. Denote by x a generic composition describing chicken composition at the start of the study, and suppose that this is distributed as $L^D(x, \mathrm{T})$. Here, of course, D = 5. If we then suppose that the composition of a chicken on normal diet will be perturbed by a random perturbation p_C distributed as $L^D(a_C, \Omega_C)$ then, by simple compositional distributional results, the final composition $X_C = p_C \oplus x$ will be distributed as $L^D(a_C \oplus x, \Omega_C + \mathrm{T})$. Similarly we can imagine a random perturbation p_T with distribution $L^D(a_T, \Omega_T)$ acting on the generic composition to produce final compositions $X_T = p_T \oplus x$ distributed as $L^D(a_T \oplus x, \Omega_T + \mathrm{T})$. We note here that constant perturbation effects correspond to hypotheses such as $\Omega_C = 0, \Omega_T = 0$.

We set out in Figure 2 a complete lattice of hypotheses together with the associated parameter sizes and the corresponding maximised loglikelihoods. Note that for this lattice there are three routes from the simplest hypothesis of no effects and no differences, namely $a_C = a_T = e,\ \Omega_C = \Omega_T = 0$, to the maximum model, by an unequal.an equal and a zero perturbation covariance route. Again with the maximised loglikelihoods determined we can use generalised likelihood ratio tests, and in this case we find that we cannot proceed any higher in the lattice than the hypothesis $a_C = e,\ \Omega_C = \Omega_T = 0$. so we may conclude that a reasonable working model is that there is no seasonal effect, that the treatment appears effective and that change in treatment is associated with a constant perturbation. Corroboration of this is given by the fact that we have to reject the two hypotheses $a_T = e,\ \Omega_C = \Omega_T = 0$ and $a_C = a_T,\ \Omega_C = \Omega_T = 0$. Note that with the 'acceptance' of a low level hypothesis there is no need to perform all the computations of the lattice.

We can obtain an idea of the nature of the treatment effect here by comparing the centres of the before group and the treated group, which are

0.382 0.159 0.027 0.107 0.325

0.458 0.149 0.022 0.111 0.260

with the mean change being represented by the perturbation

0.251 0.196 0.169 0.216 0.168.

Note that we see in this perturbation a tendency for relative increase in breast muscle, skin and, less so, non-breast muscle at the expense of fat and bone.

5. Discussion

The main thrust of this paper has been to demonstrate the importance of the perturbation operator as a main and appropriate tool for measuring change in compositions and to illustrate how hypotheses about the nature of perturbations may be simply tested within existing statistical methodology. As part of such investigations we recommend that some forethought be given, with the cooperation of the investigator, to the construction of sensible lattices before any analysis is undertaken. Even better consideration of possible lattices prior to experimentation may prevent the conduct of an inadequate experiment which fails to answer the experimenter's questions. In the construction of the lattice attention can also be given to any subcompositional hypotheses such as stability so that we can ensure that the experiment provides sufficient information to investigate such hypotheses. Finally the construction of a lattice of perturbation and subcompositional hypotheses allows a clear view of the interrelations between these hypotheses and in dealing with what are multiple-hypotheses situations encourages a preference for a simple rather than a complex explanation of the compositional variability.

References

Aitchison, J., 1982, The statistical analysis of compositional data (with Discussion): J. R. Statist. Soc B, v.44, p.139-177.

Aitchison, J., 1983, Principal component analysis of compositional data: Biometrika, v. 70, p. 57-65.

Aitchison, J., 1986, The Statisitcal Analysis of Compositional Data: Chapman and Hall, London, 416 p.

Aitchison, J., 2001, Simplicial inference, in *Algebraic Methods in Statistics and Probability* (eds. M. A. G. Viana and D. St. P. Richards), pp. 1-22. Providence, Rhode Island: Contemporary Mathematics Series 287, American Mathematical Society..

Aitchison, J., Barceló-Vidal, C., Ecozcue, J. J. and Pawlowsky-Glahn, V., 2002, :A concise guide to the algebraic-geometric structure of the simplex, the sample space for compositional data analysis, in *Proceedings of the Eighth Annual Conference of the International Association for Mathematical Geology*.

Aitchison, J. and Shen, S.M., 1980, Logistic-normal distributions: some properties and uses: Biometrika, v. 67, p. 261-272.

Anderson, T.W., 1958, An Introduction to Multivariate Statistical Analysis: Wiley, New York, 374 p.

Chang, T. C., 1988, Spherical regression: Ann. Statist., v. 14, p. 907-924.

Appendix Table 1

Milk compositions

Notation:

Pr=protein
Mf=milk fat
Ch=carbohydrate
Ca=calcium
Na=sodium
K =potassium

Control group before

Pr	Mf	Ch	Ca	Na	K
0.3098	0.2237	0.4410	0.0103	0.0025	0.0127
0.2679	0.3687	0.3377	0.0084	0.0030	0.0144
0.2583	0.3392	0.3747	0.0074	0.0047	0.0157
0.2450	0.2614	0.4617	0.0090	0.0090	0.0140
0.3715	0.1477	0.4514	0.0098	0.0032	0.0163
0.2451	0.2987	0.4263	0.0104	0.0032	0.0163
0.3797	0.2268	0.3660	0.0064	0.0080	0.0131
0.2286	0.2723	0.4709	0.0097	0.0026	0.0159
0.2381	0.2182	0.5199	0.0100	0.0016	0.0122
0.3731	0.1937	0.4051	0.0109	0.0020	0.0153
0.1988	0.4113	0.3632	0.0056	0.0080	0.0131
0.3178	0.1908	0.4678	0.0058	0.0067	0.0111
0.2446	0.2976	0.4272	0.0114	0.0018	0.0175
0.2680	0.2357	0.4731	0.0041	0.0085	0.0106
0.3448	0.2428	0.3840	0.0098	0.0040	0.0148
0.2107	0.4630	0.2955	0.0154	0.0016	0.0138
0.2767	0.1796	0.5177	0.0040	0.0089	0.0130
0.3286	0.2883	0.3584	0.0065	0.0038	0.0143
0.2168	0.3149	0.4421	0.0083	0.0043	0.0136
0.2325	0.2858	0.4544	0.0049	0.0066	0.0157
0.3140	0.1600	0.4967	0.0092	0.0053	0.0149
0.3007	0.2313	0.4451	0.0084	0.0016	0.0131
0.1966	0.3840	0.3933	0.0101	0.0031	0.0128
0.1207	0.5170	0.3328	0.0075	0.0042	0.0179
0.1728	0.4103	0.3892	0.0112	0.0015	0.0150
0.1655	0.5171	0.2841	0.0094	0.0066	0.0173
0.3257	0.1735	0.4761	0.0059	0.0044	0.0142
0.2177	0.3711	0.3788	0.0147	0.0021	0.0155
0.2628	0.3019	0.4022	0.0131	0.0035	0.0164
0.3754	0.1718	0.4256	0.0112	0.0009	0.0150

Control group after

Pr	Mf	Ch	Ca	Na	K
0.2582	0.3057	0.4107	0.0105	0.0021	0.0128
0.2381	0.3954	0.3356	0.0112	0.0030	0.0168
0.2405	0.3291	0.3985	0.0093	0.0047	0.0179
0.2877	0.2461	0.4342	0.0108	0.0063	0.0149
0.4395	0.1251	0.4049	0.0109	0.0028	0.0169
0.2040	0.3285	0.4400	0.0103	0.0022	0.0149
0.3427	0.2165	0.4115	0.0070	0.0077	0.0146
0.1469	0.4245	0.4000	0.0115	0.0015	0.0156
0.1941	0.2976	0.4779	0.0135	0.0018	0.0150
0.4360	0.1699	0.3690	0.0107	0.0012	0.0132
0.2302	0.4212	0.3186	0.0069	0.0085	0.0145
0.3338	0.2230	0.4174	0.0070	0.0063	0.0123
0.2351	0.3279	0.4102	0.0101	0.0013	0.0154
0.2475	0.2789	0.4435	0.0059	0.0102	0.0140
0.2942	0.3392	0.3415	0.0086	0.0034	0.0132
0.2112	0.4724	0.2886	0.0152	0.0009	0.0117
0.2809	0.1890	0.4981	0.0055	0.0098	0.0166
0.3244	0.3192	0.3291	0.0070	0.0051	0.0153
0.2164	0.2855	0.4692	0.0097	0.0047	0.0147
0.2310	0.3091	0.4341	0.0048	0.0064	0.0145
0.2411	0.1875	0.5468	0.0082	0.0039	0.0125
0.3304	0.2364	0.4056	0.0103	0.0017	0.0156
0.2461	0.3472	0.3786	0.0115	0.0030	0.0137
0.1321	0.5356	0.3041	0.0077	0.0035	0.0170
0.1276	0.4896	0.3516	0.0136	0.0012	0.0164
0.1447	0.6130	0.2158	0.0080	0.0047	0.0139
0.3044	0.1814	0.4878	0.0068	0.0041	0.0155
0.2352	0.4027	0.3373	0.0114	0.0015	0.0119
0.2248	0.3225	0.4217	0.0117	0.0037	0.0157
0.3039	0.2252	0.4477	0.0106	0.0006	0.0119

Treatment group before

Pr	Mf	Ch	Ca	Na	K
0.3270	0.1956	0.4500	0.0068	0.0083	0.0123
0.3758	0.1720	0.4267	0.0071	0.0057	0.0125
0.2473	0.3304	0.3924	0.0086	0.0059	0.0156
0.2624	0.2719	0.4344	0.0090	0.0054	0.0169
0.2811	0.2700	0.4226	0.0042	0.0108	0.0112
0.3456	0.2318	0.4003	0.0039	0.0069	0.0115
0.4216	0.1417	0.4138	0.0080	0.0024	0.0125
0.2465	0.3286	0.3980	0.0087	0.0046	0.0135
0.2468	0.3266	0.3945	0.0092	0.0052	0.0178
0.3486	0.1670	0.4575	0.0118	0.0015	0.0135
0.3217	0.2407	0.4055	0.0069	0.0126	0.0128

0.2165	0.3268	0.4260	0.0111	0.0035	0.0161
0.3296	0.2173	0.4197	0.0092	0.0110	0.0133
0.2324	0.3370	0.4026	0.0086	0.0022	0.0172
0.2252	0.3160	0.4245	0.0099	0.0072	0.0171
0.1756	0.4177	0.3797	0.0091	0.0037	0.0143
0.3169	0.2167	0.4373	0.0051	0.0116	0.0125
0.2226	0.3809	0.3668	0.0064	0.0088	0.0145
0.2820	0.2373	0.4514	0.0085	0.0040	0.0168
0.2180	0.3414	0.4138	0.0066	0.0042	0.0161
0.3460	0.2307	0.3926	0.0106	0.0046	0.0155
0.3065	0.2337	0.4336	0.0125	0.0014	0.0122
0.2522	0.2965	0.4227	0.0141	0.0016	0.0130
0.3312	0.1541	0.4896	0.0073	0.0048	0.0130
0.2800	0.2365	0.4562	0.0115	0.0015	0.0144
0.2704	0.2809	0.4256	0.0119	0.0009	0.0104
0.5041	0.0875	0.3808	0.0104	0.0027	0.0146
0.3187	0.2490	0.4041	0.0111	0.0037	0.0134
0.2396	0.3502	0.3793	0.0106	0.0033	0.0170
0.2424	0.2725	0.4592	0.0117	0.0015	0.0127

Treatment group after

Pr	Mf	Ch	Ca	Na	K
0.3575	0.1780	0.4357	0.0090	0.0085	0.0113
0.5056	0.1038	0.3607	0.0107	0.0063	0.0129
0.3635	0.2455	0.3616	0.0097	0.0060	0.0137
0.3510	0.2040	0.4182	0.0116	0.0027	0.0125
0.2246	0.3028	0.4419	0.0071	0.0116	0.0121
0.3966	0.1662	0.4115	0.0066	0.0085	0.0107
0.5544	0.1024	0.3145	0.0146	0.0023	0.0117
0.3587	0.2107	0.3980	0.0147	0.0048	0.0130
0.2509	0.2850	0.4385	0.0108	0.0027	0.0122
0.4076	0.1332	0.4351	0.0137	0.0012	0.0094
0.2939	0.2268	0.4510	0.0099	0.0083	0.0101
0.1521	0.3636	0.4580	0.0127	0.0025	0.0111
0.4641	0.1584	0.3491	0.0085	0.0101	0.0098
0.2870	0.2738	0.4091	0.0126	0.0019	0.0157
0.2693	0.2995	0.4037	0.0135	0.0035	0.0104
0.1894	0.4421	0.3416	0.0110	0.0041	0.0117
0.2816	0.2176	0.4722	0.0071	0.0098	0.0117
0.2154	0.4184	0.3414	0.0092	0.0050	0.0105
0.2896	0.2187	0.4638	0.0097	0.0028	0.0154
0.3070	0.2707	0.3921	0.0112	0.0030	0.0160
0.3749	0.2146	0.3794	0.0145	0.0039	0.0128
0.3195	0.2214	0.4297	0.0186	0.0011	0.0097
0.2654	0.2255	0.4766	0.0206	0.0011	0.0108
0.3843	0.1460	0.4478	0.0088	0.0034	0.0096

Appendix Table 2

Chicken compositions

Notation

Bm=breast muscle
Om=other (non-breast) muscle
Fa=fat
Sk=skin
Bo=bone

Group sacrificed at beginning of study

Bm	Om	Fa	Sk	Bo
0.2815	0.2119	0.0410	0.1058	0.3598
0.3745	0.1587	0.0360	0.1024	0.3284
0.4464	0.1336	0.0244	0.0863	0.3092
0.3694	0.1606	0.0282	0.1191	0.3227
0.4870	0.1101	0.0252	0.1033	0.2743
0.4319	0.1358	0.0259	0.1036	0.3028
0.3518	0.1733	0.0330	0.1064	0.3355
0.4167	0.1546	0.0225	0.0798	0.3264
0.4117	0.1442	0.0179	0.1162	0.3100
0.3391	0.1895	0.0266	0.0927	0.3520
0.2562	0.2161	0.0373	0.1398	0.3506
0.4155	0.1362	0.0228	0.1294	0.2960
0.5204	0.1035	0.0203	0.0865	0.2693
0.3146	0.2024	0.0269	0.0946	0.3617
0.3247	0.1753	0.0326	0.1389	0.3284
0.3108	0.1957	0.0295	0.1129	0.3511
0.3900	0.1642	0.0256	0.0791	0.3411
0.4014	0.1313	0.0291	0.1458	0.2923
0.4322	0.1284	0.0203	0.1291	0.2900
0.3885	0.1673	0.0210	0.0859	0.3372

Group on standard diet sacrificed at end of study

Bm	Om	Fa	Sk	Bo
0.3965	0.1632	0.0237	0.0846	0.3321
0.4055	0.1545	0.0256	0.0893	0.3251
0.4077	0.1430	0.0245	0.1156	0.3091
0.4712	0.1224	0.0223	0.0915	0.2925
0.3845	0.1584	0.0185	0.1197	0.3188
0.4217	0.1447	0.0224	0.0951	0.3161
0.3346	0.1865	0.0268	0.1054	0.3467

0.3412	0.1833	0.0259	0.1054	0.3442
0.4132	0.1526	0.0322	0.0696	0.3325
0.5136	0.1057	0.0217	0.0869	0.2721
0.3680	0.1654	0.0328	0.0991	0.3346
0.2726	0.2186	0.0311	0.1157	0.3620
0.4612	0.1304	0.0173	0.0896	0.3015
0.3455	0.1644	0.0285	0.1469	0.3147
0.4686	0.1267	0.0198	0.0875	0.2975
0.3800	0.1529	0.0230	0.1298	0.3143
0.3409	0.1923	0.0245	0.0877	0.3545
0.4086	0.1341	0.0302	0.1313	0.2958
0.3676	0.1648	0.0218	0.1206	0.3251
0.3665	0.1763	0.0151	0.1027	0.3395

Group on new diet sacrificed at end of study

Bm	Om	Fa	Sk	Bo
0.5169	0.1232	0.0203	0.1021	0.2375
0.3650	0.1969	0.0201	0.1269	0.2910
0.5264	0.1182	0.0173	0.1061	0.2321
0.5236	0.1138	0.0122	0.1267	0.2237
0.4012	0.1821	0.0273	0.1001	0.2893
0.4677	0.1509	0.0267	0.0906	0.2641
0.5056	0.1302	0.0180	0.1011	0.2452
0.4560	0.1476	0.0216	0.1179	0.2569
0.3473	0.1850	0.0210	0.1750	0.2716
0.4979	0.1380	0.0213	0.0857	0.2571
0.4894	0.1383	0.0227	0.0934	0.2563
0.4873	0.1467	0.0229	0.0792	0.2638
0.2870	0.2342	0.0257	0.1496	0.3035
0.4172	0.1800	0.0298	0.0833	0.2898
0.4241	0.1479	0.0221	0.1520	0.2539
0.4600	0.1471	0.0262	0.1041	0.2627
0.5239	0.1252	0.0174	0.0917	0.2419
0.4573	0.1387	0.0216	0.1350	0.2474
0.5097	0.1271	0.0189	0.1039	0.2404
0.4981	0.1227	0.0253	0.1163	0.2376

40 Maximum model $(x_C, \mathrm{T}_C), (x_T, \mathrm{T}_T)$ 799.56

35 $x_C = x_T$ 694.53
35 $x_T = e$ 761.77
35 $x_C = e$ 785.32

30 $x_C = x_T = e$ 745.53

30 $\mathrm{T}_C = \mathrm{T}_T$ 776.74

25 $x_C = x_T, .\ \mathrm{T}_C = \mathrm{T}_T$ 723.55
25 $x_T - e, \quad \mathrm{T}_C = \mathrm{T}_T$ 704.14
25 $x_C = e, \quad \mathrm{T}_C = \mathrm{T}_T$ 758.69

Unequal covariances route

Equal covariances route

20 $x_C = x_T = e, \quad \mathrm{T}_C = \mathrm{T}_T$ 686.70

Figure 1. The milk composition lattice. Various hypotheses are expressed in terms of the maximum model parameters. The numbers to the left of the model and hypotheses are the numbers of parameters involved. The numbers on the right are the values of the maximised loglikelihoods For further details, see text.

42 Maximum model $(x, \mathrm{T}), (a_C \oplus x. \Omega_C + \mathrm{T}), (a_T \oplus x, \Omega_T + \mathrm{T})$ 837,00

38 $a_C = a_T$
38 $a_T = e$
38 $a_C = e$

34 $a_C = a_T = e$

32 $\Omega_C = \Omega_T$ 831.18

28 $a_C = a_T,\ \Omega_C = \Omega_T$ 712.11
28 $a_T = e,\ \Omega_C = \Omega_T$
28 $a_C = e,\ \Omega_C = \Omega_T.$

24 $a_C = a_T = e, \Omega_C = \Omega_T$

22 $\Omega_C = \Omega_T = 0$ 827.44

18 $a_C = a_T,\ \Omega_C = \Omega_T = 0$ 663.16
18 $a_T = e,\ \Omega_C = \Omega_T = 0$ 662.79
18 $a_C = e,\ \Omega_C = \Omega_T = 0$ 825.17

Unequal covariance perturbation trail | Equal covariance perturbation trail | Constant perturbation trail

14 $a_C = a_T = e, \Omega_C = \Omega_T = 0$ 652.38

Figure 2. The chicken carcass composition lattice. Various hypotheses are expressed in terms of the maximum model parameters. The numbers to the left of the model and hypotheses are the numbers of parameters involved. The numbers to the right are the maximised loglikelihoods. For further details, see text.