AN ANALYSIS OF DATA FROM LABORATORY POPULATIONS OF *DROSOPHILA MELANOGASTER*

E. NOVITSKI AND E. R. DEMPSTER

Biology Division, Oak Ridge National Laboratory,¹ and Department of Genetics, University of California, Berkeley

Received August 7, 1957

THE detailed analysis of data from biological experiments may sometimes be thwarted by the magnitude of the computational effort involved in evaluating the variables in the statistical formulation. In those instances where a precise mathematical model can be set up, this difficulty can be overcome by the use of high-speed computers. One such application is described in this paper; it will be shown that not only can they provide a more satisfactory interpretation of numerical results, but also that a greater insight into the issues involved, and possible deficiencies in the experimental procedures may be revealed by the detail and precision of the numerical output of such gadgets.

The computer used was the "Oracle," operated by the Mathematics Panel of the Oak Ridge National Laboratory. The conversion of the maximum likelihood equations developed in the course of this work into the language of the machine (programming) was the responsibility of the first author; the various routines described here, although adequate to the task at hand, are admittedly amateurish and should be regarded only as evidence that productive effort is possible even if a programmer has limited experience and mathematical talent.

THE PROBLEM

The data analyzed are those of HEXTER (1955), which were collected to test the possible existence of single locus and/or multilocus heterosis. In brief, HEXTER found that in laboratory populations of *Drosophila melanogaster* in which certain mutant genes were present in panmixia with their wild type alleles, the frequencies of the wild and homozygous recessive phenotypes, and of the homozygous wild type and heterozygous genotypes departed from the expectations based on the HARDY-WEINBERG Law and, furthermore, that the departure from expectation was generally most pronounced in the case of the heterozygote, which appeared more often than expected. From this it was concluded that such counts might serve as evidence for heterosis.

Several questions arise at this point: (1) If the population counts deviate from the HARDY-WEINBERG Law in its simple form, do they conform with the expectations when the HARDY-WEINBERG expression is modified to include viability factors for the three genotypes? (2) Is it possible that the deviations are brought about, not by an increased viability of the heterozygote, but by a decreased via-

¹ Operated by Union Carbide Nuclear Company for the U. S. Atomic Energy Commission.

bility of the homozygous recessive class, which would be not unexpected, or by some combination of the two? (3) Can the values of these viability factors be estimated from the data available? The modification of the HARDY-WEINBERG expression required to answer these questions leads to equations that are utterly intractable by ordinary methods, hence the use of the automatic computer.

The Nature of the Data

The populations analyzed involved six different mutants: eyeless, clot, vestigial, thread, hairy, and black. Each of the six experiments was divided into two lines, one in which the initial frequency of the mutant allele was high (0.80) and another in which it was low (0.20). Counts of the two phenotypes, and genetic tests of a sample of the wild type flies to determine the relative frequencies of homozygotes and heterozygotes, were made monthly for about a year. Combining the two lines in each experiment gives a total of about 24 (22–28) counts for each experiment. Because the first few counts might deviate from the HARDY-WEINBERG distribution for extraneous reasons, which HEXTER points out, the first two in both high and low lines in each experiment were excluded from the data analyzed here.

Table 1 is a reproduction of the data in the *thread* experiment. It shows the largest deviations of all six experiments. If the first two counts of both high and low lines are disregarded, there are in the remaining 22 counts, six cases where

TABLE 1

Summary of thread experiment: Frequencies of thread and wild type flies and results of progeny tests. Frequency of recessive allele (q) and test of goodness of fit to Hardy-Weinberg distribution (χ^2)

Count	High line						Low line					
	Population count		Progeny test		q	χ^2	Population count		Progeny test		q	χ^2
	+	<i>th</i>	+/+	+/th			+	<i>th</i>	+/+	+/th		
1	1547	125	22	72	0.294	39.6*	1008	74	23	65	0.289	33.2*
2	1840	168	17	60	0.303	32.1*	1368	123	20	57	0.303	25.0*
3	1745	136	30	52	0.279	13.9*	2292	234	32	50	0.299	7.7†
4	2095	202	31	57	0.304	12.2*	1841	166	17	47	0.297	20.5*
5	2269	227	26	56	0.309	15.2*	2106	203	25	36	0.300	4.2‡
6	2471	198	27	58	0.281	21.3*	2020	246	37	47	0.332	1.3
7	1807	112	32	41	0.249	8.5†	1923	147	29	42	0.273	7.9†
8	2437	154	40	45	0.249	6.3‡	2794	226	36	45	0.277	5.0‡
9	2225	102	30	32	0.216	7.4†	2580	181	37	42	0.260	4.8‡
10	2601	103	47	37	0.200	4.6‡	2512	149	44	45	0.242	5.3‡
11	2329	84	47	29	0.190	1.5	2262	112	49	37	0.221	1.9
12	1952	53	57	28	0.166	0.9	2589	96	42	43	0.198	12.6*
13	2163	51	49	30	0.159	4.9‡	2773	104	48	41	0.196	7.5†

* Significant at 0.001 level.

† Significant at 0.01 level.

‡ Significant at 0.05 level.

the test of goodness of fit to the HARDY-WEINBERG distribution shows a deviation significant at the 0.001 level, four at the 0.01 level, and seven at the 0.05 level. The data from the *vestigial* experiment (not reproduced here) have out of 22 counts, six deviating at the 0.001 level, four at the 0.01 level, and three at the 0.05 level. The data from the other experiments are less extreme although each, with the possible exception of data involving the gene *clot*, presents an unexpectedly high number of counts departing from the expected ratios.

Formulation of the Maximum Likelihood Expressions

Instead of testing the distribution of the four types (two phenotypes and two wild-type genotypes) against a simple HARDY-WEINBERG distribution, the assumption will be made that relative to the frequency of the homozygous wild-type, which is arbitrarily taken as unity, the frequency of the heterozygote is modified by a factor k_1 and that of the homozygous recessive by a factor k_2 . For each count, then, the gene frequency, q , of the recessive allele will be evaluated, and for all counts simultaneously, k_1 and k_2 will be evaluated for a "best fit." The six experiments are to be tested individually in this fashion in an attempt to answer the questions posed earlier.

If the relative frequencies of the three genotypes at each count are defined as $(1-q)^2$, $2q(1-q)k_1$ and q^2k_2 , then the relative proportions of the four classes, the phenotypically wild type and the homozygous recessive among the total counted, and the homozygous wild type and heterozygotes among the wild types tested, become:

$$\begin{aligned}
 &+ \frac{(1-q)^2 + 2q(1-q)k_1}{(1-q)^2 + 2q(1-q)k_1 + q^2k_2} \\
 a/a &\frac{q^2k_2}{(1-q)^2 + 2q(1-q)k_1 + q^2k_2} \\
 +/+ &\frac{(1-q)^2}{(1-q)^2 + 2q(1-q)k_1} \\
 +/a &\frac{2q(1-q)k_1}{(1-q)^2 + 2q(1-q)k_1}
 \end{aligned}$$

If A , B , C , and D are the actual numbers of the above four classes, in that order, in any count, then the solution for q is obtained by equating to zero the derivative with respect to q of the logarithm of the likelihood, the latter, apart from a constant multiplier, being simply the product of the last four expressions raised to the A , B , C , and D powers, respectively. This equation is

$$\frac{\delta \log L}{\delta q} = -\frac{A+C}{1-q} + \frac{(A-C-D)(2k_1-1)}{1-q+2qk_1} + \frac{2B+D}{q} - \frac{2(A+B)(qk_2+q-2qk_1+k_1-1)}{(1-q)^2+2q(1-q)k_1+q^2k_2} = 0$$

There is no simple algebraic solution to this equation. The computer was programmed to solve for q in a rather simple fashion by trial and error. By use of arbitrary numerical values of k_1 and k_2 at the outset and subsequently machine altered values as the program proceeded toward the final solution, the value $\frac{1}{2}$ was substituted for q , and the expression evaluated. If positive, indicating that q was too small, $(\frac{1}{2})^2$ was subtracted from the previous guess of $\frac{1}{2}$, and if too large, it was added. Successive powers of $\frac{1}{2}$ were added or subtracted until q was estimated accurately to $(\frac{1}{2})^{15}$.

Although q is estimated individually for each count, k_1 and k_2 are to be estimated for the best fit to all the counts in a given experiment. Consequently, the complete maximum likelihood expression involves the product of the individual expressions for each count. This product becomes the summation after the logarithm is taken. Where N is the total number of counts in the experiment being analyzed, the first partial derivatives of the log likelihood with respect to k_1 and k_2 are respectively,

$$\frac{\delta \log L}{\delta k_1} = \sum_1^N \left[\frac{(A_i - C_i - D_i)2q_i}{1-q_i+2k_1q_i} + \frac{D_i}{k_1} - \frac{2(A_i+B_i)q_i(1-q_i)}{(1-q_i)^2+2q_i(1-q_i)k_1+q_i^2k_2} \right]$$

$$\frac{\delta \log L}{\delta k_2} = \sum_1^N \left[-\frac{(A_i+B_i)q_i^2}{(1-q_i)^2+2q_i(1-q_i)k_1+q_i^2k_2} + \frac{B_i}{k_2} \right].$$

In principle, the steps involved in the computer program for the solution for any one set of data are the following:

1. k_1 and k_2 are arbitrarily set at some starting point within the "biologically acceptable range" (neither can be negative, nor can they be distastefully high). The gene frequency q is evaluated for each count individually from 1 to N .
2. The numerical values of the first partial derivatives with respect to the viability coefficients are obtained by summing the individual contributions from each count. A χ^2 total is computed for the entire set of data.
3. On the basis of the first partial derivatives, either k_1 or k_2 , or both, is adjusted to give a better fit, and the program returns the computations to step 1.
4. When values of k_1 and k_2 have been estimated such that the most recent readjustment of both k_1 and k_2 gives a change in both of less than 2^{-20} , the system is considered to be solved and the computer calculates the χ^2 total at this endpoint. (This does not imply that the over-all accuracy of all the computations was 2^{-20} .)

Although such precision could readily be achieved by the machine, we felt that the raw data did not warrant the additional time required to attain it. Consequently, the over-all accuracy of the computations was limited to give a maximum error of about 2^{-12} , or about three decimal places. Some idea of this accuracy can be obtained by noting the relative uniformity of the quotient k_1^2/k_2 within each series in Table 2).

TABLE 2

Machine output of the maximum likelihood solutions of the viability coefficients for one dummy set of data and six tables of Hexter

Experiment	Initial		Final		Initial	Final	k_1^2
	k_1	k_2	k_1	k_2	χ^2	χ^2	k_2
Dummy (D.F. = 23)	1.000	1.000	1.69791	1.23004	164.75	0.00	2.34375
	1.500	1.500	1.99277	1.69433	46.29	0.00	2.34375
	0.500	0.500	Divergent				
	1.500	0.500	0.80621	0.27737	112.13	0.00	2.34375
	0.500	1.500	Divergent				
<i>eyeless</i> (D.F. = 23)	1.000	1.000	1.26970	1.06655	58.64	27.49	1.51154
	1.500	1.500	1.50682	1.50210	27.52	27.49	1.51154
	0.500	0.500	0.92724	0.56880	242.57	27.49	1.51154
	1.500	0.500	0.35278	0.08233	258.67	27.49	1.51162
	0.500	1.500	Divergent				
<i>clot</i> (D.F. = 23)	1.000	1.000	1.14084	1.06946	35.46	26.47	1.21698
	1.500	1.500	1.29877	1.38601	38.30	26.47	1.21699
	0.500	0.500	0.88675	0.64612	210.22	26.47	1.21698
	1.500	0.500	Divergent				
	0.500	1.500	Divergent				
<i>vestigial</i> (D.F. = 21)	1.000	1.000	1.79321	0.70441	114.05	16.74	4.56513
	1.500	1.500	1.05899	0.24563	107.85	16.74	4.56531
	0.500	0.500	1.12628	0.27786	26.76	16.74	4.56535
	1.500	0.500	1.50952	0.49911	16.66	16.74	4.56531
	0.500	1.500	1.56164	0.53418	233.36	16.74	4.56537
<i>thread</i> (D.F. = 21)	1.000	1.000	1.62832	1.20081	200.82	65.51	2.20800
	1.500	1.500	1.91379	1.65878	95.86	65.51	2.20797
	0.500	0.500	1.20679	0.65958	323.26	65.51	2.20797
	1.500	0.500	0.66019	0.27920	123.30	65.51	2.20804
	0.500	1.500	Divergent				
<i>hairy</i> (D.F. = 19)	1.000	1.000	1.33236	1.20276	41.20	13.68	1.47589
	1.500	1.500	1.48172	1.48752	13.84	13.68	1.47591
	0.500	0.500	1.07431	0.78198	208.41	13.68	1.47590
	1.500	0.500	Divergent				
	0.500	1.500	Divergent				
<i>black</i> (D.F. = 17)	1.000	1.000	1.42767	1.18188	76.91	26.36	1.72460
	1.500	1.500	1.64703	1.57291	29.54	26.36	1.72461
	0.500	0.500	1.08755	0.68579	283.04	26.36	1.72462
	1.500	0.500	Divergent				
	0.500	1.500	Divergent				

Programs Used

One serious problem recognized prior to the attempt to solve these equations was that of the possible existence of multiple solutions, since it was not obvious from the form of the equations that there could exist one and only one set of values for k_1 and k_2 that would lead to a solution. This was handled empirically in several different ways. First, a "graphing" program was written that evaluated the first partial derivatives as k_1 was held constant and k_2 varied over its range, then k_1 was stepped up by a small fraction (2^{-4}) and k_2 was varied again. The process was repeated until k_1 reached its upper limit. When fed into the computer, it showed that the derivatives changed sign only once as one viability coefficient was held constant and the other varied, indicating that for any acceptable k_1 there would be one and only one acceptable k_2 , and vice versa. Second, two different programs were written for maximizing the likelihood with respect to k_1 and k_2 . The first was essentially a trial and error routine and the second, a modification of NEWTON's method. It was thought that if both routines arrived at the same solutions consistently, the existence of multiple solutions could be considered unlikely. Finally, both these methods were programmed to allow for a variety of starting values, of k_1 and k_2 , to create the possibility of arriving at different sets of final values, if such should exist. It was discovered that there are, in fact, multiple solutions, and that these bear a simple but nevertheless unexpected relation to each other. This relation was discovered during the process of "debugging."

It is not unusual for a programmer to err in his attempt at converting the operations implied in a mathematical expression to the language understood and slavishly obeyed by the computer. Such errors are, in fact, a source of constant embarrassment and consternation to the beginner. The systematic depletion of errors from a program is known in the trade as "debugging." One effective method of debugging consists in feeding in a mock problem to which the answer has previously been determined by hand.

The data used for debugging purposes were as follows: Let $A=2700$; B , 200; C , 40; and D , 50. Simple substitution shows that there is an exact solution when k_1 is $5/4$, k_2 is $2/3$ and q is $1/3$, the numbers for A , B , C , and D having been derived originally by substituting these fractions in the maximum likelihood expression to get an exact fit. The χ^2 should consequently be reduced to zero (or very close to it) when the best values of k_1 and k_2 are computed. However, if both k_1 and k_2 equal 1, the χ^2 value is about 6.8, and if repeated in tabular form 24 times to simulate a set of actual data, the χ^2 comes out to about 163. The first line of Table 2 shows a rather surprising result. With the initial values of k_1 and k_2 of 1.000, the χ^2 comes out correct, but the final values are not those anticipated, but are instead 1.698 and 1.230. Note, however, that the χ^2 has been reduced to zero, indicating that an exact fit has been achieved. Furthermore, different initial values (lines 2 and 4 of the table) lead to different final values but in each case the χ^2 is reduced to zero; these values also fit the maximum likelihood equation.

This occurrence of multiple solutions could have been anticipated because, as

is evident from the mechanics of computing χ^2 , the minimum χ^2 for 24 identical sets of data is the same as 24 times the minimum χ^2 for a single set. There are two degrees of freedom for a single set that permit calculation of two parameters for a perfect fit, but not three. In the present example either k_1 , k_2 , or q can be given an arbitrary value and an exact fit obtained by adjusting the other two, as is shown by the following algebra where R_1 is the ratio of expected frequencies of wild type to homozygous recessives, and R_2 the ratio of the expected frequencies of homozygous wild type to heterozygotes:

$$R_1 = \frac{(1-q)^2 + 2q(1-q)k_1}{k_2q^2}, \quad R_2 = \frac{(1-q)}{2qk_1}$$

then,

$$1-q = R_2 2qk_1, \quad q = \frac{1}{1 + 2R_2k_1}, \quad \text{and} \quad R_1 k_2 q^2 = (1-q)(1-q + 2qk_1). \quad (1)$$

Substituting,

$$R_1 k_2 \left(\frac{1}{1 + 2R_2k_1} \right)^2 = \left(\frac{2R_2k_1}{1 + 2R_2k_1} \right) \left(\frac{2R_2k_1 + 2k_1}{1 + 2R_2k_1} \right)$$

$$\text{and finally,} \quad \frac{R_1}{4R_2(R_2 + 1)} = \frac{k_1^2}{k_2}. \quad (2)$$

If experimental ratios are substituted for R_1 and R_2 , and if either q or k_1 is given any arbitrary value, the value of the other of these two parameters required for a perfect fit can be obtained from expression (1), and k_2 then computed from expression (2); or if k_2 is given an arbitrary value k_1 can be computed from (2) and q from (1). For any such fit, however, the quotient k_1^2/k_2 (expression 2) depends only on the two experimental ratios corresponding to R_1 and R_2 .

For any one set of data, then, consisting of four numbers A , B , C , and D , all combinations of k_1 and k_2 for which the ratio k_1^2/k_2 has the value indicated by expression (2) can provide an exact fit. The program was extended slightly by the addition of the machine computation of this ratio since it might have some relevance in the evaluation of solutions provided by the computer.

ANALYSIS OF THE DATA

Table 2 shows the output of the computer for all six tables of HEXTER. The specific routine used here was the modification of NEWTON's method for estimating two parameters attributed to FISHER by KOSHAL (1933) and GARWOOD (1941). This method was suggested to the authors by DR. ALLYN KIMBALL of the Mathematics Panel, Oak Ridge National Laboratory, to whom thanks are tend-

ered particularly because it proved to be the most efficient method of the three mentioned previously, usually leading to a final set of viability coefficients from a given starting point in about 40 seconds, whereas the other took several minutes. For each of the six experiments, five different initial values were used. In some cases, final values were not obtained because successive differences in the viability constants became larger rather than smaller (i.e., appeared to diverge). The indication of divergence does not necessarily imply that a solution is impossible from these starting values but only that arriving at the solution would probably be more time consuming than in those cases where successive values of k_1 and k_2 appeared to converge. The program included a test for such divergence such that when it occurred, a notation to that effect was recorded in the output and the computer proceeded to the next set of starting values.

When both the initial values are 1.000, the HARDY-WEINBERG expression is unadjusted, and the χ^2 computed at that initial point corresponds to the one that would be obtained by comparing the data directly with the simplest form of the HARDY-WEINBERG Law. It is seen that in every such instance where the initial values are both 1.000, the total χ^2 is reduced considerably by estimating k_1 and k_2 , indicating that the modified expression, with those viability constants, is more compatible with the data than the unmodified one. Furthermore, the χ^2 values remaining after those estimates are made are not inconsistent with expectation based on the number of degrees of freedom remaining at the conclusion of the computations, except for one case, that of thread. Since the χ^2 remaining in the latter case is still rather high (65.5 for 21 D. F.), there may be other factors responsible here for the original high χ^2 . In this case there may be evidence for multilocal heterosis, as HEXTER points out in his original account, with a breakdown of that system in later generations leading to an over-all heterogeneity of the data.

The final values for k_1 and k_2 are different within any set, although the χ^2 totals at the end points are much the same, indicating that these different final values fit the data equally well. Thus, despite the ample number of degrees of freedom (18–24 in the different experiments to estimate 11–14 parameters, respectively), multiple solutions were obtained here as with a single set of data. In the clot experiment, for example, equally acceptable fits are obtained if the heterozygote is assigned a viability of 130 percent and the homozygous recessive a viability of 137 percent (compared to the homozygous dominant), or if the first is assigned a viability of 87 percent and the second a viability of 64 percent. However, the values of the ratio k_1^2/k_2 (given in the last column of table 2) remain constant in any given experiment, being 1.2170 in the clot experiment. The regular agreement of those ratios within any set of iterations for a given experiment strongly suggests that the maximum likelihood solution occurs when that ratio is reached, and that it will occur at different points, depending on the values assigned to the viability factors at the start of each iteration.

The machine has thus forced to attention a circumstance completely unexpected by the authors and others with whom the problem had been discussed; namely,

the occurrence of multiple solutions. On reconsideration, however, the reason was not difficult to discover.

χ^2 for any experiment depends on the theoretical ratios R_1 and R_2 calculated for each set of data. These ratios in turn are determined by the chosen values of k_1 and k_2 and the individual values of q determined for each set. As shown by equations (1) and (2), any theoretical ratios (R_1 and R_2) whatsoever, dependent on a given k_1 , k_2 , and q , will also result from any other k_1 and k_2 having the same quotient k_1^2/k_2 provided that q is given a value corresponding to expression (1). Suppose, then, a minimum χ^2 results from given values of k_1 and k_2 , with q being optimum for each set of data of an experiment. If we now change k_1 and k_2 so that k_1^2/k_2 remains constant, and adjust each q in accordance with expression (1), χ^2 remains the same, and is therefore still a minimum. The parameters are thus not independent, and there is in fact a fatal interdependence among them that prevents evaluation of k_1 and k_2 but permits an estimation of the ratio mentioned.

DISCUSSION

The solutions presented in Table 2 tell a number of things about the analysis of this sort of data. In the first place, it is clear that a test of goodness of fit of the actual counts against an unmodified HARDY-WEINBERG Law can tell little about the source of any heterogeneity, if such exists. In the second place, the analysis of such data without the help of an automatic computer could lead to erroneous results, since it would not be obvious that a multiplicity of solutions existed. The backbreaking labor involved in arriving at even one such solution by hand might very well preclude the possibility of the regular repetition of the analysis by different initial values. The first values derived that satisfied the maximum likelihood equations might then be taken as the proper ones. Finally, the advantage of analysis of the data with a computer is shown also by its revelation of a simple algebraic relation inherent in the maximum likelihood formulation that was not obvious (to the authors) previously.

In a sense, such automatic devices might be used in the planning of an experiment as well as in the analysis of the data from it. If, before the beginning of an experiment, it is known what sort of analysis will be used finally and if that analysis is programmed immediately, then the introduction of sample data may reveal ambiguities not obvious at the start, as it did in this particular case. Under these conditions, it might be well to modify the experiment so as to remove these ambiguities.

Of the three questions posed in the introduction, the third remains unanswered. If we had had sufficient information about the generation time, the manner of generation overlap, and had been willing to make certain additional assumptions such as, for example, that the change in gene frequency is caused entirely by these viability differences, we might have obtained from HEXTER's data the best estimates of k_1 and k_2 instead of their ratio. It seems more practical, however, to obtain

this information on relative viabilities from a simple experiment designed to test just this point.

These comments are meant as general observations and not as contradictions to the analysis of HEXTER, whose detailed consideration of the data leads to a number of interesting conclusions in other connections. We hope that this account may call the attention of other geneticists to the greater possibilities for precise, detailed, and accurate analyses of such problems by the use of computing devices, which are becoming more available.

SUMMARY

In some cases heterosis is inferred because a population appears to contain more heterozygotes than can be accounted for by ordinary genetic theory. A mathematical model based on a maximum likelihood formulation has been set up that allows an estimate to be made of the degree to which irregularities in genetic data require the assumption of heterosis. The analysis based on such a model is not practicable by ordinary methods, and for this reason an automatic computer (the Oracle at the Oak Ridge National Laboratory) was used. The output of the computer showed that the assumption that heterosis is responsible for the increased frequency of the heterozygotes is unwarranted, and that a reduced viability of homozygous classes will equally well explain such data. This result is of some interest because it shows that one of the most commonly applied tests for heterosis, a statistical test for an excess of heterozygotes in a population, is actually ambiguous even if the data should consist of a large number of replicated trials or observations.

ACKNOWLEDGMENTS

The authors wish to thank DR. ALLYN KIMBALL of the Mathematics Panel for his help during the course of this work, MRS. NANCY M. DISMUKE, MRS. SUSIE E. ATTA, MRS. BARBARA OSBORNE for their patient tutoring in the art of programming for the computer, and DR. W. M. HEXTER for his helpful suggestions about the presentation of the problem in the manuscript.

LITERATURE CITED

- GARWOOD, F., 1941 The application of maximum likelihood for dosage-mortality curves. *Biometrika* **32**: 46-58.
- HEXTER, W. M., 1955 A population analysis of heterozygote frequencies in *Drosophila*. *Genetics* **40**: 444-459.
- KOSHAL, R. S., 1933 Application of maximum likelihood to the improvement of curves fitted by the method of moments. *J. Roy. Statist. Soc.* **96**: 303-313.