Title
PROGRAMS FOR THE ANGULAR DISTRIBUTION ANALYSIS BY A MAXIMUM-LIKELIHOOD METHOD

Author
Grard, Fernand.

Publication Date
1962-02-26

UCRL-10100

UNIVERSITY OF CALIFORNIA

Ernest O. Lawrence Radiation Laboratory

BERKELEY, CALIFORNIA

DISCLAIMER

UCRL-10100

UNIVERSITY OF CALIFORNIA

Lawrence Radiation Laboratory
Berkeley, California

Contract No. W-7405-eng-48

PROGRAMS FOR THE ANGULAR DISTRIBUTION ANALYSIS BY A MAXIMUM-LIKELIHOOD METHOD

Fernand Grard

February 26, 1962

PROGRAMS FOR THE ANGULAR DISTRIBUTION ANALYSIS BY A MAXIMUM-LIKELIHOOD METHOD

Fernand Grard[†]

Lawrence Radiation Laboratory
University of California
Berkeley, California

February 26, 1962

The two programs described in this paper have been written for the analysis of an angular distribution by using the maximum-likelihood method.

The programs are particularly suited for the cases in which the least-squares method is inapplicable, i.e., when some intervals of the distribution are so poorly populated that the distribution of assigned errors is not Gaussian. The programs, called "AD-Cos" and "AD-Leg", have been written for the 709 and 7090 IBM computer of the Lawrence Radiation Laboratory, Berkeley. They have to be run with the FORTRAN MONITOR system. A copy of each of these programs is available from Robert Harvey, Alvarez group, Lawrence Radiation Laboratory.

Description of the Method

The function fitted to the experimental data is

$$f = 1 + a_1 u_1 + a_2 u_2 + \cdots + a_n u_n, \qquad (1)$$

where $u_1, u_2, \cdots u_n$ are given functions of the angle θ defining the angular distribution.

In Program "AD-Cos" the function f is a polynomial of order n in $\cos\theta$. Thus, $u_\ell = \cos^\ell \theta$.

[†] On leave from the Institut Interuniversitaire des Sciences Nucléaires, Belgium.

In Program "AD-Leg" the functions u_i are the unnormalized Legendre polynomials[1] that satisfy the relations

$$P_\ell(\cos\theta) = \frac{2\ell - 1}{\ell}\cos\theta\, P_{\ell-1}(\cos\theta) - \frac{\ell - 1}{\ell} P_{\ell-2}(\cos\theta), \quad (2)$$

$$(2\ell + 1) P_\ell(\cos\theta) = \frac{dP_{\ell+1}(\cos\theta)}{d\cos\theta} - \frac{dP_{\ell-1}(\cos\theta)}{d\cos\theta}. \quad (3)$$

The maximum value for n is 10 for both cases.

The experimental distribution is assumed to be known within two intervals of $\cos\theta$ extending from $\cos\theta = a$ to $\cos\theta = b$ for the first interval and from $\cos\theta = c$ to $\cos\theta = d$ for the second (Fig. 1). This allows one to delete regions of the distribution where it could have been difficult or impossible experimentally to measure the events, as for example in the forward and backward directions.

The probability that a given event has a value θ_i in $d\cos\theta$ is

$$p(\theta_i)\, d\cos\theta = \frac{f(\theta_i)}{\int_a^b + \int_c^d [f(\theta)\, d\cos\theta]}\, d\cos\theta. \quad (4)$$

Then the likelihood function for all the events (total number = N) is given by

$$L = \frac{\prod_{i=1}^{N} f(\theta_i)}{\left[\int_a^b + \int_c^d [f(\theta)\, d\cos\theta]\right]^N}. \quad (5)$$

The maximum-likelihood estimates of the coefficients $a_1, a_2, \cdots a_n$ in (1) are those for which the likelihood function L, Eq. (5), is maximum. They are found by solving the following set of equations:

$$\frac{\partial \ln L}{\partial a_\ell} = 0 \text{ for } \ell = 1, 2, \cdots, n. \quad (6)$$

---

(1) Erdelyi, Magnus, Oberhettinger, and Tricomi, Higher Transcendental Functions, Vol. 1 (McGraw-Hill Book Co., Inc., New York 1953).

or

$$\sum_{i=1}^{N} \frac{u_\ell(\cos\theta_i)}{f(\cos\theta_i)} - N \frac{\int_a^b + \int_c^d [u_\ell(\cos\theta)\, d\cos\theta]}{\int_a^b + \int_c^d [f(\cos\theta)\, d\cos\theta]} = 0. \qquad (7)$$

These equations are nonlinear in the coefficients a_ℓ. They have been solved by an iterative procedure, starting from approximate values a_ℓ^* which are assumed to be known.

The exact solution is then supposed to be

$$a_\ell = a_\ell^* + \Delta a_\ell. \qquad (8)$$

A first attempt to determine the solution is made by solving the following equations, obtained by expending Eq. (7) in a Taylor series and keeping only first-order terms in Δa_ℓ:

$$\sum_{i=1}^{N} \frac{u_\ell(\cos\theta_i)}{f^*(\cos\theta_i)} - N \frac{\int_a^b + \int_c^d (u_\ell(\cos\theta_i)\, d\cos\theta)}{\int_a^b + \int_c^d (f^*(\cos\theta)\, d\cos\theta)} \qquad (9)$$

$$= \sum_{k=1}^{N} \Delta a_k \left\{ \sum_{i=1}^{N} \frac{u_\ell(\cos\theta_i)\, u_k(\cos\theta_i)}{f^*(\cos\theta_i)^2} - N \frac{\left[\int_a^b + \int_c^d (u_\ell(\cos\theta)\, d\cos\theta\right]\left[\int_a^b + \int_c^d (u_k(\cos\theta) d\cos\theta\right]}{\left[\int_a^b + \int_c^d f^*(\cos\theta)\, d\cos\theta\right]^2} \right\}.$$

Here $f^*(\cos\theta)$ represents the function f in which the coefficients $a_1 \cdots a_n$ have been replaced by the approximate values $a_1^* \cdots a_n^*$. In the "AD-Leg" case, the integrals in (9) have been reduced by using relation (2) and (3). The increments Δa_k found by solving the system (9) are added to the approximate coefficient a_k^*. The resulting values of a_k are then used as a new approximate solution and the procedure is repeated, starting from (8).

Convergence and Variances

At each iteration, the quantities

$$\delta a_k = \left[- \frac{\partial^2 \ln L}{\partial a_k^2} \right]^{-\frac{1}{2}} \tag{10}$$

are computed. These quantities, which would represent the standard errors if the coefficients a_k were not correlated, are used to build the convergence check. In the present version of the programs, the iteration procedure is stopped when the last increments Δa_k found by solving the system (9) are all smaller than a certain fraction of their respective "standard errors," δa_k. This fraction can be introduced in the programs as an initial information (XERR).

Remarks

The function f, which should represent the probability distribution of the events, can take negative values for an arbitrary set of the coefficients. When this happens in the course of the iteration procedure, new approximate coefficients a_k^* are determined by adding to the approximate coefficients of the previous step 1/10 the increments Δa_k. If necessary, the increments are reduced again by the same factor and the iterative procedure is continued.

Input Format

The experimental angular distribution can be fed into the program in two different ways:

(a) as a list of individual events, each with its corresponding value of $\cos\theta$;

(b) as a histogram. For each interval, the number of the events and the value of $\cos\theta$ corresponding to the middle of the interval are given. All the events of the interval are treated by the program as though they had the same value of $\cos\theta$. This makes easier the introduction of the data into the computer (and reduces the number of input cards). The intervals have to be chosen small enough in order to have a distribution practically equivalent to that of Case a.

The total number of events must be less than 2000.

The input data have to be given on punched cards, in the following order:

1. One card with the maximum number of iterations allowed and the factor XERR. Format (20X, I10, F10.5)

2. One card with the limits of the experimental angular distribution (see above): a, b, c, d. Format (20X, 4F10.5)

3. One card with the number of intervals in Case (b) above, with the total number of events in Case (a). Format (20X, I10)

4. For each interval or each event, one card with the corresponding value of cos θ and the corresponding number of events (one in Case (a)). Format (20X, F10.5, I10)

5. One card for the order n. Format (20X, I10)

6. The initial values of the coefficients, each on one card. Format (20X, E12.4)

At this point, if a function f of different order n has to be fitted to the same experimental data, steps 5 and 6 are repeated.

7. A blank card if another set of data has to be analyzed. Then cards 2 and through 6, inclusive, are added.

8. A blank card which is interpreted as the end of the calculations.

Output Format

Here is given an example of output for an angular distribution analyzed by the "AD-Cos" program. (The example is shown on the following page.) The output format is identical for the "AD-Leg" Program. The time required for the calculations is about 2 minutes on the IBM 709 Computer. Figure 2 shows the experimental data along with the normalized fitted distribution.

Remarks

1. At the end of each step the value of the likelihood function is computed and printed.

2. When a satisfactory set of coefficients has been found, the error matrix is computed by inversion of the matrix formed with the second derivatives of the logarithm of the likelihood function L.

3. The function f is integrated over the two intervals (a, b) and (c, d) and the results are printed.

4. The fitted distribution normalized to the number of events in the intervals (a, b) and (c, d) is tabulated for values of $\cos\theta$ from -1 to +1 for intervals of width 0.1.

5. If a satisfactory solution has not been found after the allowed number of iterations, the last increments Δa_k are printed in addition to the other information.

6. If the function f takes negative values for the initial coefficients, the calculation is stopped. The program has to be sent back to the computer with another set of initial values.

```
. DATA
                    -0.9000        13
                    -0.7000         8
                    -0.5000         7
                    -0.3000        13
                    -0.1000        14
                     0.1000        23
                     0.3000        27
                     0.5000        34
                     0.7000        33
                     0.9000        21
 RESULTS FOR ORDER N =  3

       INITIAL PARAMETERS
     0.730E 00
     0.168E-00
     0.

  L. FUNCT. AFTER ITER NO.  1 =                  -118.4885
 THE PROB. FUNCT. WAS NEG. AT ITER. NO.= 2
  L. FUNCT. AFTER ITER NO.  2 =                  -112.8384
  L. FUNCT. AFTER ITER NO.  3 =                  -111.6775
  L. FUNCT. AFTER ITER NO.  4 =                  -111.5691

 NUMBER OF ITERATIONS =  4 FOR N =  3 LIK.FNCT. =          -111.5675

  SATISFACTORY SOLUTION FOR THE PARAMETERS
     0.163E 01
    -0.165E-00
    -0.171E 01

  SQ. ROOT OF INV. OF SEC. DER.
   0.1009E-00
   0.1813E-00
   0.1674E-00
 INTEGRATED DISTRIBUTION =   0.1890E 01


  ERROR MATRIX
 0.7170E-01  0.1679E-01 -0.1018E-00
 0.1679E-01  0.4727E-01 -0.4147E-02
-0.1018E-00 -0.4147E-02  0.1815E-00


      FITTED DISTRIBUTION
      -0.950          0.790E 01
      -0.850          0.562E 01
      -0.750          0.419E 01
      -0.650          0.351E 01
      -0.550          0.348E 01


      -0.450          0.399E 01
      -0.350          0.494E 01
      -0.250          0.623E 01
      -0.150          0.774E 01
      -0.050          0.938E 01
       0.050          0.110E 02
       0.150          0.126E 02
       0.250          0.140E 02
       0.350          0.151E 02
       0.450          0.157E 02
       0.550          0.159E 02
       0.650          0.155E 02
       0.750          0.143E 02
       0.850          0.124E 02
       0.950          0.948E 01
```

Acknowledgments

These programs were developed to a great extent during a previous stay in CERN, Geneva. It is a pleasure to thank here Dr. Y. Goldschmidt-Clermont for his stimulating discussions and Dr. F. Bruyant for his collaboration in the early stage of the problem. The last phases of the work have been done under the auspices of the U. S. Atomic Energy Commission.

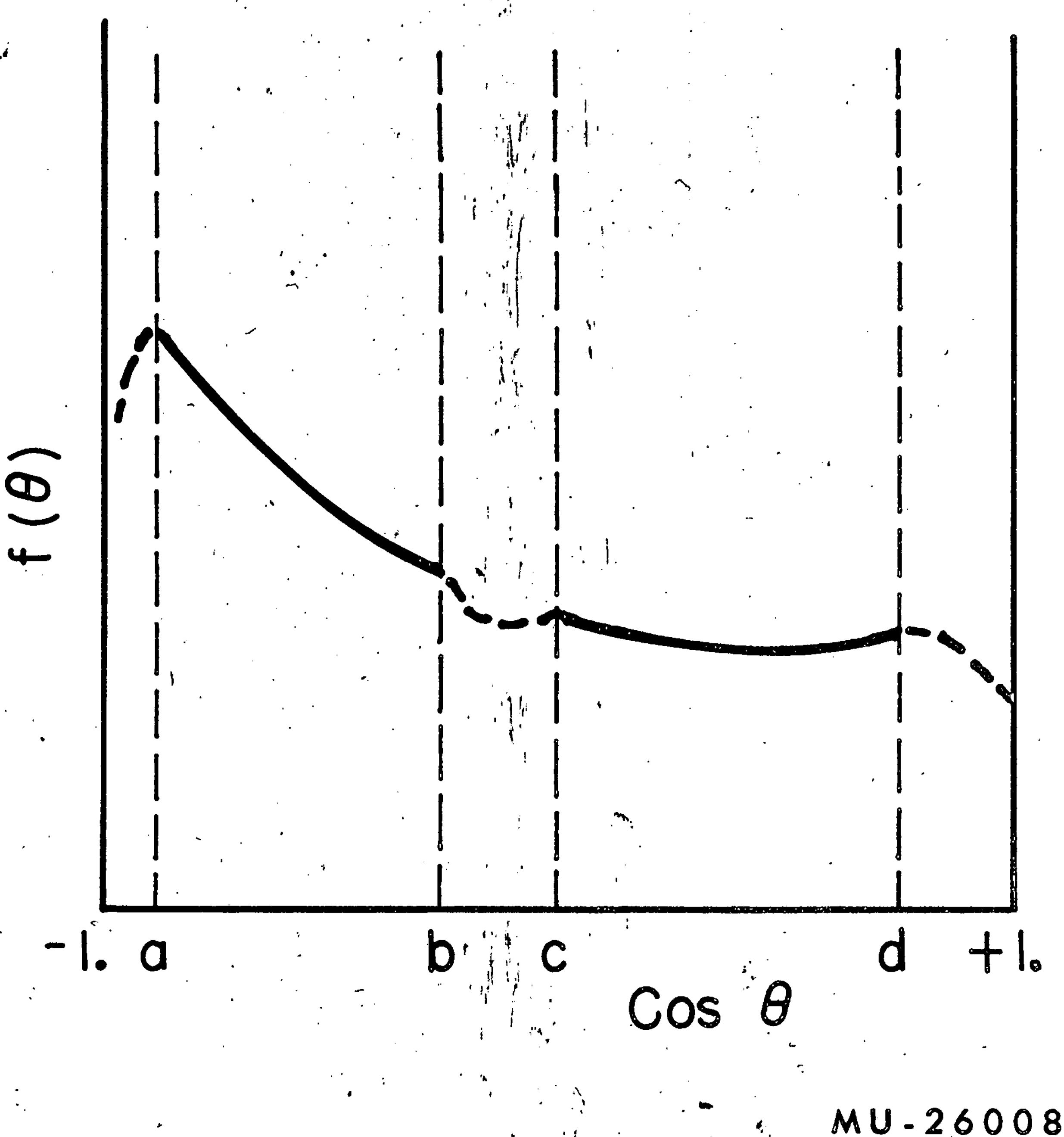

Fig. 1. Experimental distributions within two given intervals.

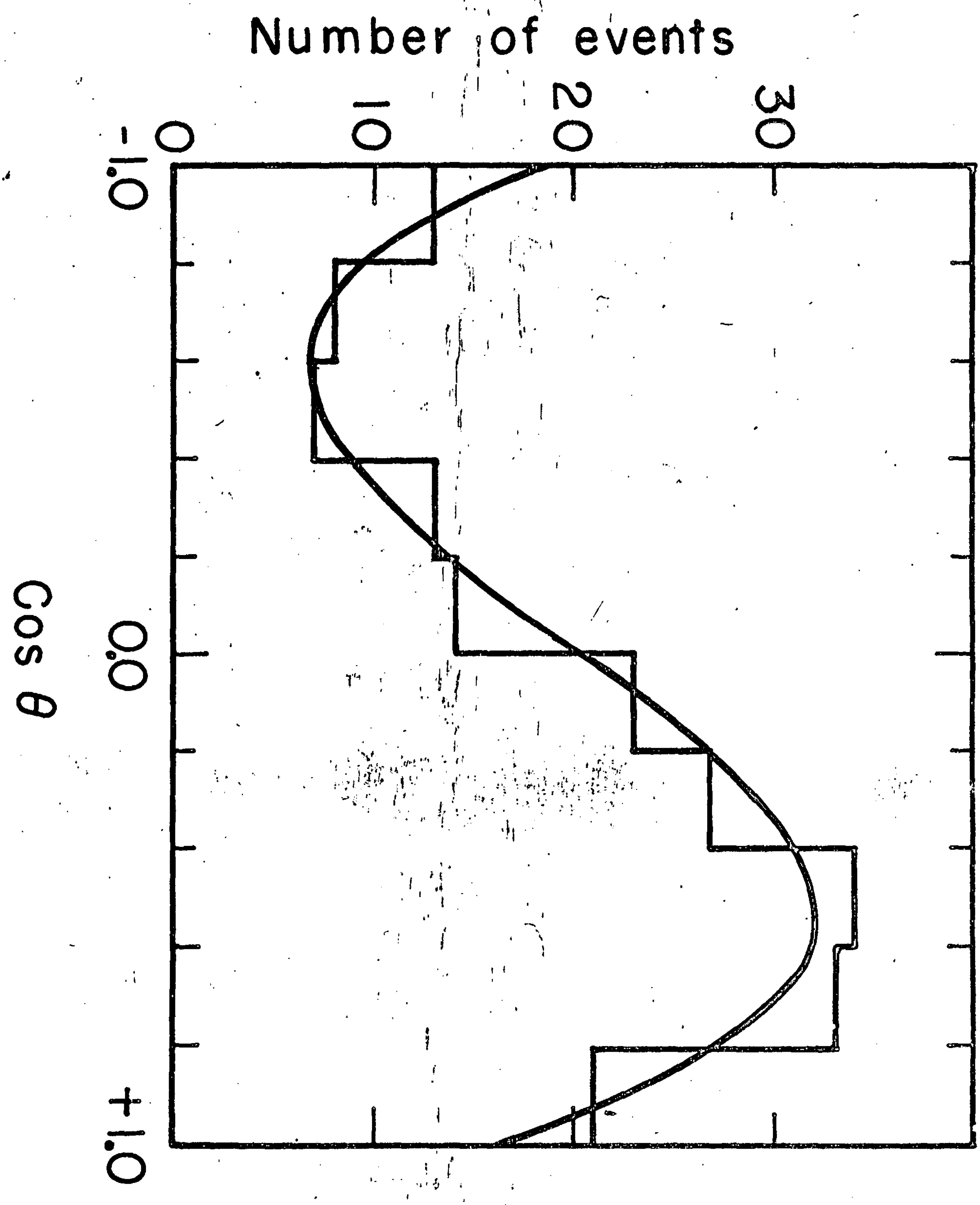

Fig. 2. Experimental and normalized fitted distribution.