Regression analysis of exponential palaeodose growth curves

V. Poljakov and G. Hütt
Institute of Geology, Estonian Academy of Sciences, Estonia pst.7, Tallinn 200101, ESSR, USSR.

When using palaeodosimetric dating methods, laboratory reconstruction of the palaeodose poses a key problem. As a rule, it is solved by the *additive dose* method, with extrapolation of the resulting experimental curve towards intersection with the x-axis. Given an experimental set of data (1)

$$\{ D_A^{(i)}, I_i \equiv I(D_A^{(i)}); i = 1, ...n \} \quad (1)$$

where D_A is the laboratory additive does, and $I(D_A)$ the intensity of the TL peak or ESR spectrum signal, reconstruction of the accumulated dose DN is realised through a statistical regression analysis. This assumes optimum choice of the regressional functional dependence.

Supposing that fading involves a first order thermally activated process, the solving of the kinetic equations leads to the following functional shape for the intensity-dose relation (Hutt and Smirnov, 1982):

$$I(D_A) = I_o [1 - e^{-\beta(D_N + D_A)}] \quad (2)$$

where I_o and β are the parameters characterising the palaeo-dosimeter. Thus, (1) and (2) represent a non-linear three-parameter regression model I_o, β, D_N are the parameters to be estimated.

It would be more convenient to perform the regression analysis rewriting the equation (2) in the following way:

$$y(x) = a + be^{cx}$$

where $x \equiv D_A$ and $y(x) \equiv I(D_A)$.

From the condition $y(-D_N) = 0$ we obtain the value of the accumulated dose:

$$D_N = \frac{1}{c} \ln (-\frac{b}{a})$$

The final model may be expressed as follows

$$y_i = f(\vec{\theta}; x_i) + \varepsilon_i , \quad i = 1, ...n$$

where y_i is the dependent variable, $\vec{\theta} = \{a,b,c\}$ is the estimated parameter vector, x_i is the independent variable, and $\vec{\varepsilon}$ is the random deviation vector. It is also assumed that x_i is error-free and that the random deviations are of normal distribution and uncorrelated,

i.e. $\quad cov(\vec{\varepsilon}) \sim N(\vec{0}, \sigma^2 \vec{1})$.

The parameters a, b, c are estimated using the least squares method, when optimal choice of these parameters $\hat{\vec{\theta}} = \{\hat{a}, \hat{b}, \hat{c}\}$ is determined by minimization of

$$S(a,b,c; x_i,y_i) = \sum_{i=1}^{n} (y_i - a - be^{cxi})^2$$

The linearised Newton-Gauss method of minimization of $S(\vec{\theta}; x_i,y_i)$ is known from the literature (Berger et al, 1987). However, the linearisation procedure has some disadvantages. In some cases it leads to a slow convergence of the iterative processes and even to divergence.

The present paper proposes a straight minimisation method of $S(\vec{\theta}; x_i,y_i)$ by solving of the corresponding system of normal equations:

$$\left.\frac{\partial S}{\partial a}\right|_{\vec{\theta} = \hat{\vec{\theta}}} = -\sum_{i=1}^{n} 2 (y_i - \hat{a} - \hat{b}e^{cxi}) = 0$$

$$\left.\frac{\partial S}{\partial b}\right|_{\vec{\theta} = \hat{\vec{\theta}}} = -\sum_{i=1}^{n} 2 (y_i - \hat{a} - \hat{b}e^{\hat{c}xi}) e^{\hat{c}xi} = 0$$

$$\left.\frac{\partial S}{\partial b}\right|_{\vec{\theta} = \hat{\vec{\theta}}} = -\sum_{i=1}^{n} 2 (y_i - \hat{a} - \hat{b}e^{\hat{c}xi}) \hat{b}x_i e^{\hat{c}xi} = 0$$

.... (3)

From the first two equations of the system (3) we have:

$$\hat{b} = \frac{\sum_{i=1}^{n} y_i e^{\hat{c}x_i} - \frac{1}{n} \sum_{i=1}^{n} e^{\hat{c}x_i} \sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} e^{2\hat{c}x_i} - \frac{1}{n} (\sum_{i=1}^{n} e^{\hat{c}x_i})^2}$$

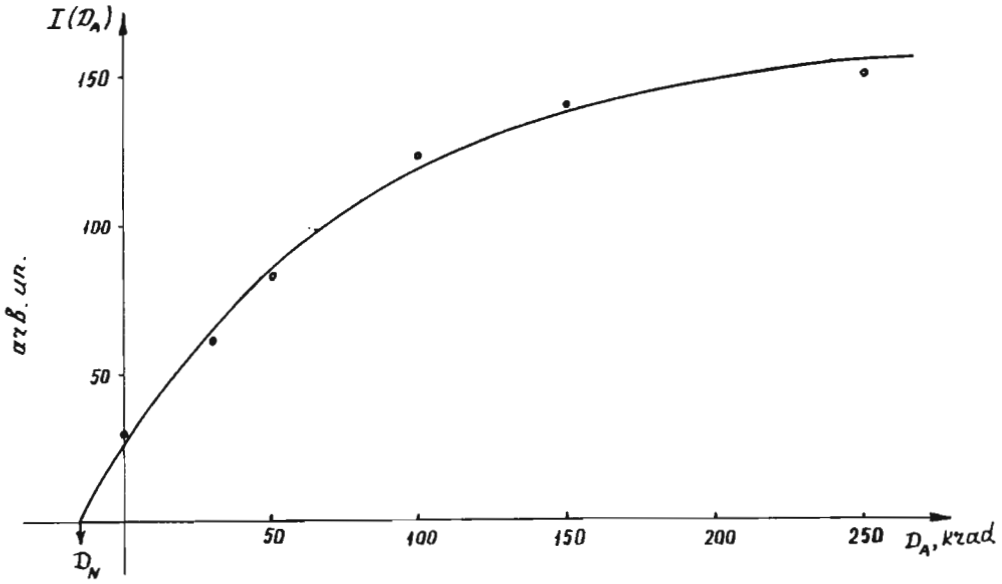
$$\hat{a} = \frac{1}{n} \sum_{i=1}^{n} y_i - \frac{\hat{b}}{n} \sum_{i=1}^{n} e^{\hat{c}x_i}$$

..... (4)

Inserting (4) into the third equation of the system (3) we obtain the equation for $\hat{c}$:

$$F(\hat{a}, \hat{b}, \hat{c} ; x_i, y_i) = 0 \quad (5)$$

Non-linear equation (5) can be solved numerically by means of the Newton-Raphson iteration method:



$$\hat{c}_K = \hat{c}_{K-1} - \frac{F(\hat{c}_{K-1})}{F'(\hat{c}_{K-1})}, \quad K = 1, \ldots ; \qquad (6)$$

where $\hat{c}_o$ is the initial approximation and

$$F'(c) \equiv \frac{\partial F}{\partial c} + \frac{\partial F}{\partial a}\frac{\partial a}{\partial c} + \frac{\partial F}{\partial b}\frac{\partial b}{\partial c}$$

Calculating the derivatives included in $F'(\hat{c}_{K-1})$ by means of the system of equations (4) we entirely determine the calculation scheme of the iterative procedure. The termination of the iterative process (6) is based on the natural proximity condition of consequent estimates for $\hat{c}$:

$$| (\hat{c}_{K^*}) - (\hat{c}_{K^*-1}) | \leq \varepsilon$$

The proposed calculation scheme is realised in the form of a BASIC programme. The working results of the programme are presented in the figure, where the regression dependence is shown by the solid line, and the experimental data by dots.

A linearised procedure for the approximate estimation of error in the accumulated dose can be carried out in the following way. Denoting with $V(y_i)$ the variance of y_i:

$$V(y_i) = \frac{1}{m_i - 1}\sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)^2 \qquad i = 1, \ldots. n;$$

$$\text{where } \bar{y}_i = \frac{1}{m_i}\sum_{j=1}^{m_i} (y_{ij} - y_i)^2 \qquad i = 1, \ldots n;$$

and m_i is the number of repeated measurements of value y_i. Writing the accumulated dose in a Taylor series in powers of $\delta y_i \equiv y_i - \bar{y}_i$, and confining ourselves to linear terms:

$$\delta_{D_N} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial a}\frac{\partial a}{\partial y_i} + \frac{\partial f}{\partial b}\frac{\partial b}{\partial y_i} + \frac{\partial f}{\partial c}\frac{\partial c}{\partial y_i}\right)\partial y_i, \qquad (7)$$

where

$$f = D_N = \frac{1}{c}\ln\left(\frac{b}{a}\right),$$

$$a = a(x_i, y_i), \; b = b(x_i, y_i), \; c = c(x_i, y_i);$$

$$\delta_{D_N} = f(a,b,c) - f(\hat{a}, \hat{b}, \hat{c})$$

and the derivatives of function f are determined at the point of

$$\hat{\vec{\theta}} = \{ \hat{a}(x_i, \bar{y}_i), \hat{b}(x_i, \bar{y}_i), \hat{c}(x_i, \bar{y}_i) \}.$$

A linearized procedure for the approximation estimate of error in the accumulated dose can be acrried out by using the law of propagation of errors (Mandel, 1964):

$$V(D_N) = \sum_{i=1}^{n} \left(\frac{\partial D_N}{\partial a}\frac{\partial a}{\partial y_i} + \frac{\partial D_N}{\partial b}\frac{\partial b}{\partial y_i} + \frac{\partial D_N}{\partial c}\frac{\partial c}{\partial y_i}\right)^2 V(y_i)$$

where $V(D_N)$ is the variance of D_N and $V(y_i)$ is the variance in the value of y_i evaluated using replicate measurements at the same dose x_i.

References

Berger, G.W., Lockhart, R.A., and Kuo, I., (1987), Regression and error analysis applied to the dose-response curves in thermoluminescence dating. *Nucl. Tracks and Radn. Measts.*, **13**, 177-184.

Hutt, G., and Smirnov, A., (1982) Detailed thermoluminescence dating studies of samples from geological reference profiles in Central Russia. *PACT J.*, **6**, 505-513.

Mandel, J. (1964) *The statistical analysis of experimental data.* Interscience, p73.

PR. Reviewed by A.D. Franklin