

# MASSACHUSETTS INSTITUTE OF TECHNOLOGY

## PHYSICS DEPARTMENT

8.13/8.14

Junior Laboratory

### STATISTICS AND ERROR ESTIMATION

The purpose of this note is to explain the application of statistics to the estimation of error in the context of an imaginary experiment that is similar to situations encountered in the Junior Lab. In physical measurements the word "error" should be interpreted as meaning uncertainty. More complete discussions of the theory and practise of error estimation can be found in the books by Bevington & Robinson (1992) and by Taylor (1982).

Imagine an experiment designed to measure the gamma-ray activity of a radioactive source with a detector connected to a counting system as illustrated in Figure 1. Assume that the emission of gamma rays by the source is isotropic. Call  $Q$  the activity of the source  $S$  in units of disintegrations per second,  $A$  the effective area of the detector (the actual area multiplied by the detector efficiency),  $x$  the distance from the source to the detector, and  $n$  the number of gamma-ray photons detected in time  $T$  as measured with a digital timer. The objective of the experiment is to derive an estimate of  $Q$  from measurements of these quantities and to estimate the error in the derived value.

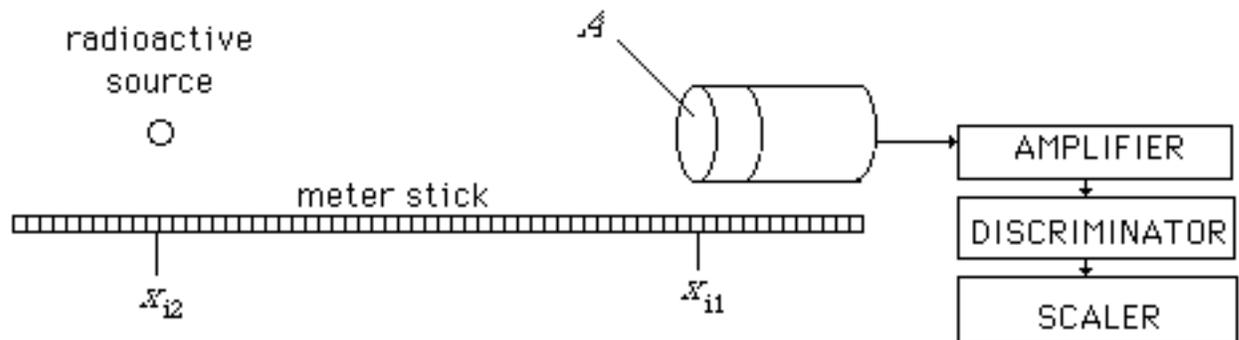


Figure 1. Experimental setup for measuring the activity of a radioactive source.

The expression for the source strength in terms of the measured quantities is

$$Q = \frac{n}{T} \frac{4}{A} x^2. \quad (1)$$

The data are listed in Table 1. The error associated with  $A$  is an estimate of the "one-sigma" uncertainty in the effective area due to a combination of random and systematic errors of measurement. The error of  $T$  is purely systematic since it is the reading of a digital timer controlled by a vibrating quartz crystal with a limited guarantee of accuracy. The data include the results of a measurement of  $x$  which was carried out with a meter stick in a manner designed to exploit the advantages of repetition. We assume that the observer, cautious to avoid psychological bias, shifted the meter stick between

measurements of the positions of the source and detector and did not subtract the two readings  $X_{i1}$  and  $X_{i2}$  before writing them down.

The following questions arise:

1. What are systematic and random errors?
2. What does "one-sigma" mean?
3. What is the best estimate and one-sigma error of the measurement of the distance  $x$ ?
4. What is the one-sigma error of the number of recorded counts  $n$ ?
5. Given the measurements and one-sigma errors of  $n$ ,  $T$ ,  $X$ , and  $A$ , how are they to be combined to obtain the estimate of  $Q$  and its error?

---

Table 1. Hypothetical data from a measurement of the intrinsic activity of a radioactive source.

$$n = 10,000$$

$$T = 100 \pm 0.001 \text{ s}$$

$$A = 40.0 \pm 0.2 \text{ cm}^2$$

$i$	$X_{i1}$ (cm)	$X_{i2}$ (cm)	$x_i = X_{i2} - X_{i1}$ (cm)
1	12.2	22.8	10.6
2	11.8	22.2	10.4
3	15.7	26.0	10.3
4	12.7	23.2	10.5
5	13.3	23.9	10.6
6	14.0	24.5	10.5

### Summary

In brief, the answers to the questions are:

1. Systematic errors are caused by calibration errors of instruments, by improper procedures, and by psychological bias that can result from prior conviction as to what the answer "should" be. Systematic errors tend to be one-sided, i.e., to result in repeated measurements that lie on one or the other side of the true value known only to Mother Nature. Random error is the uncertainty in a measurement due to everything other than systematic error such as, for examples, counting statistics, electronic noise, and uncertainties in visual interpolation between scale marks. Random error causes the results of a repeated measurement to be scattered with a distribution which is often (but not necessarily) symmetric about a mean value and describable by the Gaussian distribution.

2. A "one-sigma error" is a measure of the width of a Gaussian distribution which would be expected to fit the histogram of values of a measured quantity if the measurement were repeated many times. In the example only six statistically independent measurements of  $x$  were made. So the problem is to make the best estimate of  $x$  and its error from the available data.

3. The best estimate of  $x$  and the associated random error are

$$x = m_x \pm s_x,$$

where

$$m_x = \frac{1}{6} \sum_{i=1}^6 x_i = 10.48 \text{ cm}$$

is the mean of the six measurements

$$\text{and } s_m = \sqrt{\frac{1}{6 \cdot 5} \sum_{i=1}^6 (x_i - m_x)^2} = 0.05 \text{ cm}$$

is the standard error of the mean.

4. The error associated with the number of counts caused by the decay of a radioactive substance is governed by Poisson statistics. If the number of counts recorded is  $n$ , then the one-sigma estimate of error is  $\sqrt{n}$ . So the one-sigma error in the number of counts recorded in the experiment is 100.

5. The one-sigma error of the estimate of  $Q$  is expressed by the equation

$$s_Q = Q \sqrt{\frac{s_n^2}{n} + \frac{2s_x^2}{x} + \frac{s_A^2}{A^2} + \frac{s_T^2}{T^2}}$$

The numerical portion of the estimate and its error are

$$Q = \frac{n}{T} \frac{x^2}{A} \left( 1 \pm \frac{s_Q}{Q} \right) = 3452.6 \pm 49.8$$

The result would be reported as

$$Q = (3.45 \pm 0.05) \times 10^3 \text{ disintegrations s}^{-1}.$$

### 1. The Gaussian Distribution of Error

In many situations, repeated, independent, unbiased measurements of a physical quantity will yield values that are randomly distributed in a manner that can be described by the Gaussian distribution

$$G(\mu, \sigma; x) = \frac{1}{\sigma\sqrt{2\pi}} \exp - \frac{(x-\mu)^2}{2\sigma^2}, \quad (2)$$

where  $G(\mu, \sigma; x)dx$  is the probability that a given measurement will yield a value between  $x$  and  $x+dx$ ,  $\mu$  is the "mean" of the distribution,  $\sigma$  is the "standard deviation" of the distribution. The semicolon separates the parameters of the distribution from the distributed quantity. It can be shown that the distribution is normalized, i.e.,

$$\int_{-\infty}^{\infty} G(\mu, \sigma; x)dx = 1. \quad (3)$$

Denoting mean values of the distribution by enclosure in brackets [ ], it can also be shown (with the help of gamma functions) that

$$[x] = \int_{-\infty}^{\infty} G(\mu, \sigma; x) x dx = \mu, \quad (4)$$

$$[(x-\mu)^2] = \int_{-\infty}^{\infty} G(\mu, \sigma; x) (x-\mu)^2 dx = \sigma^2, \quad (5)$$

$$\int_{\mu^-}^{\mu^+} G(\mu, \sigma; x) dx = 0.6826 \quad (6)$$

The last of these equations shows that the probability of a single measurement yielding a value between  $\mu^-$  and  $\mu^+$  is 0.6826. Two-sigma and three-sigma ranges indicate 0.9554 and 0.9974 probabilities, respectively.

## 2. Maximum Likelihood: Sample Mean and Sample Standard Deviation

If the distribution of measured values of a quantity  $x$  can be described by a Gaussian, then we can express the probability  $P(x_i)$  of having obtained  $N$  specific values of  $x$ , each within some small fixed range  $\Delta x \ll \sigma$ , as the product of the probabilities of obtaining each one which is

$$P(x_1, \dots, x_N) = \prod_{i=1}^N G(\mu, \sigma; x_i) = \prod_{i=1}^N \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x_i-\mu)^2}{2\sigma^2}\right]. \quad (7)$$

The "true values" of the parameters  $\mu$  and  $\sigma$  are, of course, known only to Mother Nature. The best we can do is to derive estimates of  $\mu$  and  $\sigma$  from the available data. We call  $m$  the "maximum likelihood" estimate of  $\mu$  and define it to be that value which maximizes  $P$ . To maximize  $P$ , the sum in the exponential must be minimized. Setting

$$\frac{\partial}{\partial m} \sum_{i=1}^N \frac{(x_i-m)^2}{2} = 0 \quad (8)$$

we require

$$-\frac{2}{m} \sum_{i=1}^N (x_i-m) = 0, \quad (9)$$

which implies

$$m = \frac{1}{N} \sum_{i=1}^N x_i \quad (10)$$

Thus, the maximum likelihood estimate of  $\mu$  is simply the mean of the measured values, which is called the "sample mean."

Since  $N$  is always finite (6 in the present case), we cannot assume that the average value of  $(x_i - m)^2$  is a good estimate of  $\sigma^2$ . Indeed, since  $m$  is calculated to minimize the sum of the squared deviations, it follows that  $[(x_i - m)^2]$  is likely to be an underestimate of  $\sigma^2$ . To compensate for this effect and obtain a better estimate of  $\sigma^2$ , the "sample standard deviation",  $s$ , is defined by the equation

$$s^2 = \frac{1}{(N-1)} \sum_{i=1}^N (x_i - m)^2 \quad (11)$$

in which the division is by  $(N-1)$  instead of  $N$ .

The quantity  $s$  measures the spread of the measured values of  $x$ . If the series of measurements were repeated many times, the distribution of the means would obviously be substantially narrower than the distributions of the individual measurements. Thus, to get a measure of the error in the average of a finite number of measurements we need to estimate the error of the sample mean.

We first address the general problem of estimating the standard deviation of a function  $f(q_1, q_2, \dots)$  of several measured quantities  $q_i$  with known means  $\mu_i$  and standard deviations  $\sigma_i$ . We call  $\mu_f$  and  $\sigma_f$  the mean and standard deviation of  $f$ , respectively, and assume the mean value of  $f$  is  $\mu_f = f(\mu_1, \mu_2, \dots)$ . Then

$$\sigma_f^2 = [(f - \mu_f)^2] = [(\sum_i \frac{\partial f}{\partial q_i} (q_i - \mu_i))^2]. \quad (12)$$

If the fluctuations of the measured quantities are uncorrelated, then the sum of the cross terms in the expansion of the squared sum will average to zero. In this case, the standard deviation of the mean is defined by the equation

$$\sigma_f^2 = \sum_i \left( \frac{\partial f}{\partial q_i} \right)^2 \sigma_i^2 \quad (13)$$

Considering the sample mean as a function of the variables  $x_i$ , we find

$$\frac{\partial m}{\partial x_i} = \frac{1}{N}. \quad (14)$$

The estimate of  $\sigma_i$  is the sample standard deviation  $s$ . Thus

$$s_m^2 = \frac{1}{N^2} N s^2, \quad (15)$$

so the estimate of the error of the mean is

$$s_m = \frac{s}{\sqrt{N}} = \sqrt{\frac{1}{N(N-1)} \sum_{i=1}^N (x_i - m)^2} \tag{16}$$

Table 2 lists the results of numerical experiments employing 10,000 trial data sets of various sample sizes, each consisting of  $N$  random Gaussian variates with  $\mu=100$  and  $\sigma = 10$ . For each set, the mean,  $m$ , and various measures of deviation were computed and averaged. The results show that even with the  $(N-1)$  compensation, the average value of  $s$  (sample standard deviation) is less than  $\sigma$ . The rms (root of the mean square) deviation from the specified  $\mu$  is closer to  $\sigma$ , as one would expect. The tests, with  $N=3, 10$  and  $100$ , confirm the expectation that in the limit  $N \rightarrow \infty$ ,  $m \rightarrow \mu$  and  $s \rightarrow \sigma$ .

Table 2. Summary of Monte Carlo tests of Gaussian statistics. In this table brackets  $\langle \rangle$  signify the average over  $n$  trials. The asterisks mark the quantities derived from data.

	number of trials: n =	10000	10000	10000
	sample size: N =	3	10	100
	specified mean: $\mu =$	100.000	100.000	100.000
	specified standard deviation: $\sigma =$	10.000	10.000	10.000
	stdev of the mean of N variates : $\sigma/\sqrt{N} =$	5.774	3.162	1.000
*1	$\langle m \rangle = \langle \text{sample mean} \rangle$	99.969	99.917	100.004
2	$\langle \text{rms deviation of x from } \mu \rangle$	9.178	9.756	9.970
*3	$\langle s \rangle = \langle \text{sample stdev} \rangle$	8.863	9.732	9.970
4	$\langle \text{rms deviation of x from } m \rangle$	7.236	9.233	9.920
5	$\langle \text{rms deviation of mean from } \mu \rangle$	5.718	3.156	1.003
*6	$\langle sm \rangle = \langle \text{stdev of sample mean} \rangle$	5.117	3.078	0.997
*1	$\langle m \rangle = \langle (1/N) * \sum \{x_i\} \rangle$			
2	$\langle \{ [1/N] * \sum \{ (x_i - \mu)^2 \}^{1/2} \rangle$			
*3	$\langle s \rangle = \langle \{ [1/(N-1)] * \sum \{ (x_i - m)^2 \}^{1/2} \rangle$			
4	$\langle \{ [1/N] * \sum \{ (x_i - m)^2 \}^{1/2} \rangle$			
5	$\langle \{ [1/N^2] * \sum \{ (x_i - \mu)^2 \}^{1/2} \rangle$			
*6	$\langle sm \rangle = \langle \{ [1/N(N-1)] * \sum \{ (x_i - m)^2 \}^{1/2} \rangle$			

Applying the formulas for the sample mean, standard deviation and, error of the mean with the measured values of the distance between source and detector, one finds

$$\begin{aligned} m &= 10.4833 \\ s &= 0.1169 \\ s_m &= 0.0477. \end{aligned} \tag{17}$$

These results would be reported as

$$x = 10.48 \pm 0.05 \text{ cm} \quad (18)$$

where only the significant digits are quoted after rounding off.

### 3. Poisson Statistics

To estimate the error in the quantity  $n$  obtained for the number of counts recorded in time  $T$ , one must predict the distribution of the numbers of counts that might have been obtained in many trials of the same measurement. One can never know the true rate  $r$  since any measurement is subject to uncontrollable fluctuations. So we have to proceed with the only estimate we have, which is  $n$  counts in time  $T$ .

The Poisson distribution,  $P(\mu; x)$ , expresses the probability that in a given trial exactly  $x$  counts are recorded in time  $T$  when the true rate is  $r = \mu/T$ . Imagine the time interval  $T$  divided into  $M$  sub-intervals of equal duration  $T = T/M$ , where  $M \gg \mu$ . In the limit  $M \rightarrow \infty$ , the probability that exactly one count occurs in  $x$  specific sub-intervals is

$\frac{\mu}{M}^x$ , and the probability that none occurs in the rest is  $1 - \frac{\mu}{M}^{M-x}$ . Thus,

$$P(\mu; x) = \lim_{M \rightarrow \infty} \frac{\mu}{M}^x \left(1 - \frac{\mu}{M}\right)^{M-x} W(M, x), \quad (19)$$

where  $W(M, x)$  is the number of distinct ways in which  $x$  counts can occur in  $M$  sub-intervals.

To find  $W$ , imagine the record of counts laid out as a series of  $M$  boxes with  $x$  balls labeled 1 to  $x$  representing the counts placed in numerical sequence in the boxes corresponding to the sub-intervals in which the count occurred. That series would represent just one of many ways in which a record of  $x$  counts could occur. Another way would be represented by  $x$  balls numbered 1 to  $x$  scattered at random without regard to numerical order. Ball #1 could be dropped in any of  $M$  boxes. Ball #2 could be dropped in any of the remaining  $M-1$  unoccupied boxes, etc., so that the number of distinct series without regard to numerical order is readily computed as  $M(M-1)(M-2)\cdots(M-x+1)$ . But this would be a gross over-estimate of the number of distinct ways in which  $x$  counts could occur because, if the labels were erased after the dropping, there would be no distinction between all the possible series that had the same boxes occupied, but by balls with permuted labels. The number of possible permutations of  $x$  different label numbers is  $x!$ . So the number of distinct series with unlabeled balls is less than the number of distinguishable series with labeled balls by the factor  $1/x!$ . Thus

$$W(M, x) = \frac{M(M-1)(M-2)\cdots(M-x+1)}{x!} \quad (20)$$

In the limit  $M \rightarrow \infty$ ,  $M(M-1)(M-2)\cdots(M-x+1) \rightarrow M^x$ , and  $(1 - \mu/M)^{M-x} \rightarrow \exp(-\mu/M)$ . Combining equations (19) and (20) and taking the limit, we obtain for the Poisson distribution the expression

$$P(\mu; x) = \frac{\mu^x \exp(-\mu)}{x!}. \quad (21)$$

The distribution is properly normalized, i.e.,

$$\sum_{i=0}^{\infty} P(\mu; x) = \exp(-\mu) \left( \frac{1}{0!} + \frac{\mu}{1!} + \frac{\mu^2}{2!} + \dots \right) = 1. \quad (22)$$

One easily finds for the mean and standard deviations

$$\begin{aligned} [x] &= \sum_{i=0}^{\infty} x P(\mu; x) = \exp(-\mu) \left( \frac{0}{0!} + \frac{\mu}{1!} + \frac{2\mu^2}{2!} + \frac{3\mu^3}{3!} + \dots \right) \\ &= \mu \exp(-\mu) \left( \frac{1}{0!} + \frac{\mu}{1!} + \frac{\mu^2}{2!} + \dots \right) = \mu, \end{aligned} \quad (23)$$

and,

$$\begin{aligned} &= [(x - \mu)^2] = [x^2] - \mu^2 = [x(x - 1)] + [x] - \mu^2 \\ &= \mu^2 + \mu - \mu^2 = \mu. \end{aligned} \quad (24)$$

It is interesting to compare the error estimate based on the total count  $n$  with an error estimate derived from the counts recorded in each of  $N$  sub-intervals of the total time  $T$ . Call  $n_i$  the number of counts in the  $i$ th interval of duration  $T/N$ . The sample mean, sample standard deviation, and error of the mean would be, respectively,

$$m_n = \frac{1}{N} \sum_{i=1}^N n_i = \frac{n}{N} \quad (25)$$

$$s_n = \sqrt{\frac{1}{(N-1)} \sum_{i=1}^N (n_i - m_n)^2} \quad (26)$$

$$s_{m_n} = \frac{s_n}{\sqrt{N}}. \quad (27)$$

Since  $n_i$  is a Poisson variate with an estimated mean  $m_n$ , the estimate of  $\sum_{i=1}^N (n_i - m_n)^2$  is

$Nm_n = n = s^2$ . Thus the error of the mean of  $n_i$  is related to the sample standard deviation of  $n$  by

$$s_{m_n} = \frac{s}{\sqrt{N(N-1)}}. \quad (28)$$

The estimated rate and error estimate is

$$r = \frac{m_n \pm s_{m_n}}{T}$$

$$= \frac{n}{T} \pm \sqrt{\frac{N}{(N-1)}} \frac{s}{T}. \quad (29)$$

Thus the error incurred by analyzing the data in  $N$  time segments instead of all together is larger by the factor  $\sqrt{\frac{N}{(N-1)}}$ .

It can be shown that in the limit  $n$

$$P(\mu; n) \rightarrow G(\mu, \sqrt{\mu}; n) \quad (30)$$

Returning to the data analysis, the only estimate of the mean number of counts in 100 seconds available from the data is the actual number of counts. So we take

$$m=n=10,000.$$

Then

$$s_n = \sqrt{m}=100.$$

It can be shown (with the help of Stirling's approximation) that in the limit  $\mu \rightarrow$

$$P(\mu; x) \rightarrow G(\mu, \sqrt{\mu}; x),$$

where  $\mu = \sqrt{\mu}$ . Calculation of Poisson probabilities becomes unwieldy for large values of  $x$  because of the need to evaluate  $x!$ . For values of  $\mu > 20$  the Gaussian approximation is often good enough. As an example, suppose you want to know the probability of obtaining a count in the range from 121 to 130 when  $\mu=100$ . A good estimate would be

$$10G(\mu, \sqrt{\mu}; x) \text{ with } \mu=100, \sqrt{\mu}=10, \text{ and } x=125.$$

#### 4. Error Estimate of the Source Strength

The source strength is expressed in terms of four uncorrelated measured quantities by equation (1). Assuming that the errors of those quantities are all Gaussian, we can apply equation (13) to estimate the error in  $Q$ . To find the partial derivatives it is convenient first to take the natural log of both sides of equation (1). Then, taking the partial derivatives, one obtains

$$\frac{\partial Q^2}{\partial n} = \frac{Q^2}{n}, \quad \frac{\partial Q^2}{\partial x} = \frac{2Q^2}{x}, \quad \frac{\partial Q^2}{\partial A} = \frac{Q^2}{A}, \quad \text{and} \quad \frac{\partial Q^2}{\partial T} = \frac{Q^2}{T}. \quad (31)$$

Thus

$$s_Q = Q \sqrt{\frac{s_n^2}{n^2} + \frac{2s_x^2}{x^2} + \frac{s_A^2}{A^2} + \frac{s_T^2}{T^2}}. \quad (32)$$

Evaluating  $Q$  and the fractional errors, we find

$$Q = \frac{n}{T} \frac{4}{A} \frac{x^2}{1 \pm \frac{s_Q}{Q}} \quad (33)$$

$$Q = 3452.6 \left( 1 \pm \sqrt{0.00010 + 0.000083 + 0.000025 + 1.0e-10} \right) \\ = 3452.6 \pm 49.8$$

The last term under the radical sign can be ignored. The result would be reported as

$$Q = (3.45 \pm 0.05) \times 10^3 \text{ disintegrations s}^{-1}. \quad (34)$$

### 5. Curve Fitting by the Maximum Likelihood Method

The maximum likelihood method used in Section 2 can be extended to solve the problem of fitting a "model" function to a set of data that consists of  $N$  measurements of one quantity,  $y_i$ , with measurement errors,  $\sigma_i$ , as a function of another quantity,  $x_i$ , which is assumed to be exact. The data set would consist of  $N$  triplets of numbers  $(x_i, y_i, \sigma_i)$ ,  $i=1, 2, \dots, N$ .

We assume that the probability distributions of the  $y_i$  are Gaussians with standard deviations  $\sigma_i$ . Suppose there is reason to believe that  $y$  is related to  $x$  by the equation  $y=f(a, b, \dots; x)$ , where the values of the parameters  $a, b, \dots$  are not known. The function  $f$  is called the model, and the aim of the analysis is to determine the values of the model parameters that would have yielded  $N$  values of  $y$  in the ranges  $y_i \pm \sigma_i$  with maximum probability. Adapting equation (7), we express the probability as

$$G(a, b, \dots; \sigma_i; y_i) = \prod_i \frac{1}{\sigma_i \sqrt{2\pi}} \exp \left( - \frac{[y_i - f(a, b, \dots; x_i)]^2}{2 \sigma_i^2} \right). \quad (35)$$

The condition of maximum likelihood requires, as before, that the summation be minimized. Setting

$$\chi^2 = \sum_i \frac{[y_i - f(a, b, \dots; x_i)]^2}{\sigma_i^2}, \quad (36)$$

we require

$$\frac{\partial \chi^2}{\partial a} = 0, \quad \frac{\partial \chi^2}{\partial b} = 0, \quad \dots \quad (37)$$

( $\chi^2$  is called "chi-square").

To illustrate the method we consider the case of a data set that can be modeled by a linear relation

$$y = a + bx. \quad (38)$$

According to equations (37) we require

$$\frac{2}{a} = -2 \sum_i \frac{[y_i - a - bx_i]}{i^2} = 0, \quad (39a)$$

$$\frac{2}{b} = -2 \sum_i \frac{x_i[y_i - a - bx_i]}{i^2} = 0. \quad (39b)$$

This yields two simultaneous equations in  $a$  and  $b$ ,

$$\left( \sum_i 1/i^2 \right) a + \left( \sum_i x_i/i^2 \right) b = \sum_i y_i/i^2 \quad (40a)$$

$$\left( \sum_i x_i/i^2 \right) a + \left( \sum_i x_i^2/i^2 \right) b = \sum_i y_i x_i/i^2 \quad (40b)$$

which are easily solved by determinants.

As an example, suppose the exact quantities are the durations in seconds of measurements of the activity of a long-lived radioactive source. The quantities subject to random error (i.e., statistical fluctuations) are the corresponding numbers of counts recorded by the detector. Since the number of counts is a Poisson variate, the one-sigma errors are equal to the square roots of the numbers of counts. To test the method, the Gaussian approximation of equation (30) was used to generate a set of 100 variates  $y_i$  with mean values

$$\mu_i = 100x_i, \quad x_i = 1, 2, 3, \dots, 100,$$

and corresponding one-sigma errors (based on the "measured" values rather than the  $\mu_i$ )

$$\sigma_i = \sqrt{y_i}.$$

The maximum likelihood values of  $a$  and  $b$  in the model function of equation (38) were computed by the program LINFIT listed in Bevington & Robinson (1992). Figure 2 is a plot of the data and the fitted model. The results demonstrate an essential fact about the relation between a data set with correctly estimated error bars and a smooth curve fitted to the data, namely, approximately one-third of the error bars should not be intersected by the curve.

The maximum likelihood method of curve fitting is readily extended to fitting a multi-parameter non-linear function. Assume that one has fairly good initial guesses  $a_0$ ,  $c_0$ ,  $b_0$  for the approximate values of the parameters. Expand the function in a Taylor series about those values, keeping only the linear terms in  $a$ ,  $b$ ,  $c$ ,  $\dots$ , and substitute the expanded form of  $f$  in equation (35). Numerical derivatives can be computed if analytical ones are too complicated or impractical. The numerator will then be

$$y_i = f(a_0, c_0, b_0; x_i) = \frac{f}{a} a - \frac{f}{b} b - \dots^2 \quad (41)$$

Setting the partial derivatives of  $\chi^2$  with respect to  $a$ ,  $b$ ,  $c$ , ... to zero, one obtains a set of simultaneous equations that can be solved for the increments to the parameters that improve the fit. The process is then repeated with the incrementally improved parameter values till  $\chi^2$  no longer decreases. This is the method employed in several of the Junior Lab curve-fitting programs. Details of this procedure with implementing codes can be found in Bevington and in Press, et al.

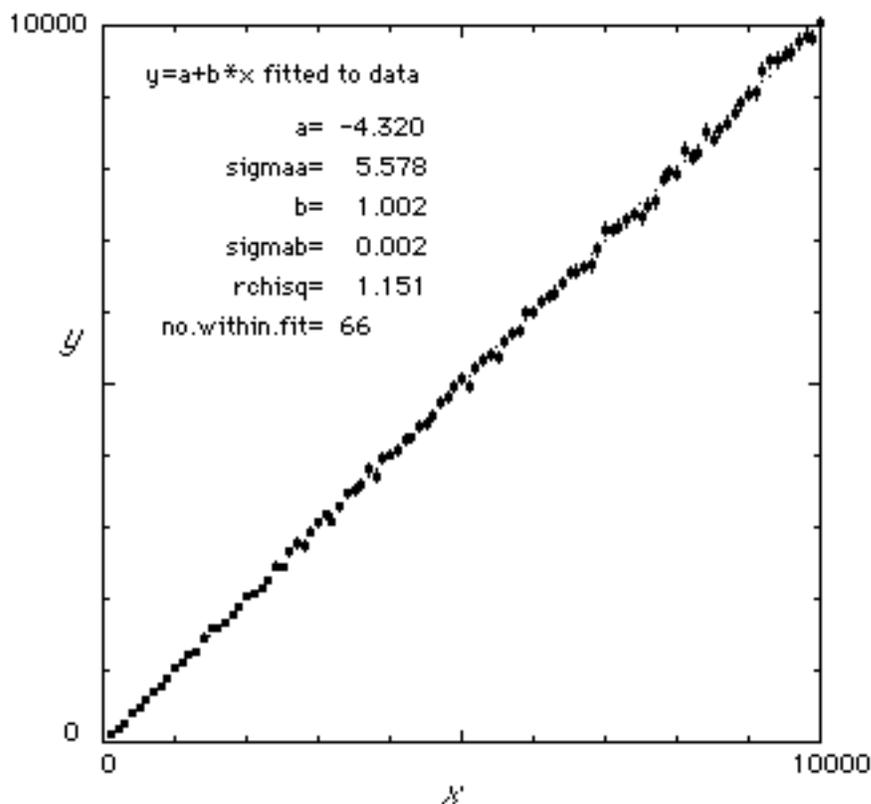


Fig. 2. Illustration of a maximum-likelihood fit of a linear relation to data produced by the Gaussian variate generator GASDEV of Press et al. (1989). The mean was  $x_i = 100i$ ,  $i=1$  to 100, and the standard deviation was  $\sqrt{x_i}$ . Note that the number of data error bars ( $y_i \pm \sqrt{y_i}$ ) intersected by the fitted line is 66, which is, as expected, about 2/3 of 100.

### References

- P. R. Bevington & D. K. Robinson, "Data Reduction and Error Analysis", (McGraw-Hill 1969)
- W. H. Press, B. P. Flannery, S. A. Teukolsky, & W. T. Vetterling, "Numerical Recipes", (Cambridge 1989)
- J. R. Taylor, "Introduction to Error Analysis", (University Science Books 1982)

