Chi-square Fitting

Zeng Qi & Qin Songmei

Chi2

If each data point (x_i, y_i) has its own, known standard deviation σ_i, then equation (15.1.3) is modified only by putting a subscript i on the symbol σ. That subscript also propagates docilely into (15.1.4), so that the maximum likelihood estimate of the model parameters is obtained by minimizing the quantity

$$\chi^2 \equiv \sum_{i=1}^{N} \left(\frac{y_i - y(x_i; a_1 \dots a_M)}{\sigma_i} \right)^2 \tag{15.1.5}$$

called the "chi-square."

Chi2 fitting & Chi2 distribution

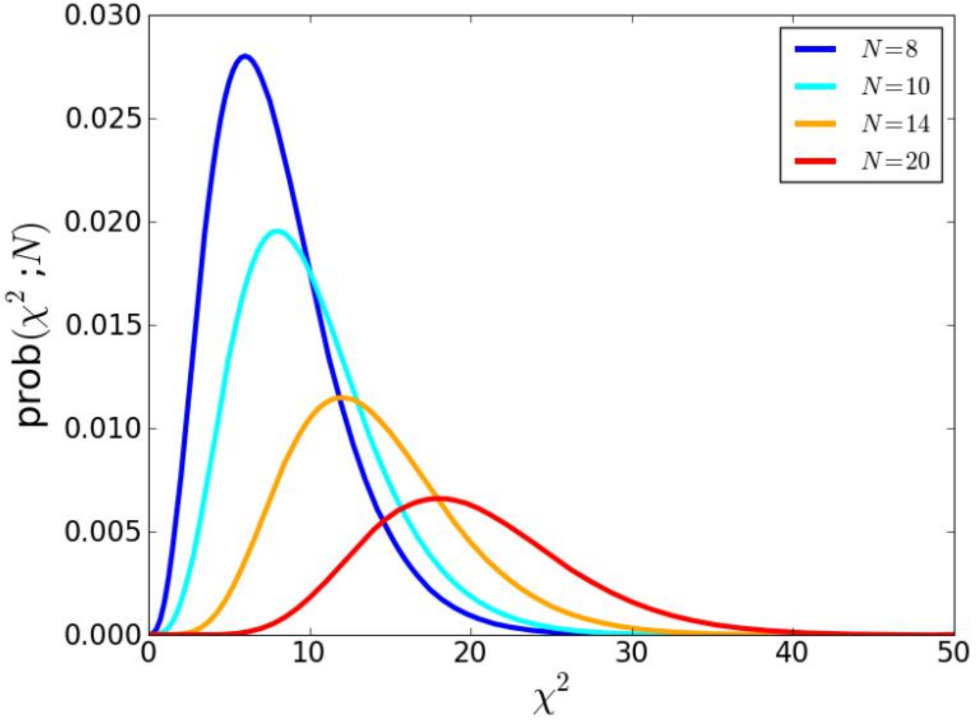
Definition: The **Chi-Square Distribution**, denoted as χ^2 is related to the standard normal distribution such as, if the independent normal variable, let's say Z assumes the standard normal distribution, then the square of this normal variable Z^2 has the chi-square distribution with 'K' degrees of freedom.

$$Q = \sum_{i=1}^{K} Z_i^2$$

Freedom

K=N-M

Definition: The **Degrees of Freedom** refers to the number of values involved in the calculations that have the freedom to vary. In other words, the degrees of freedom, in general, can be defined as the total number of observations minus the number of independent constraints imposed on the observations.



χ_2 –distributions for different values of $K = N$ degrees of freedom

Reference: arXiv:1012.3754

Freedom

A rule of thumb is that a “typical” value of χ^2 for a “moderately” good fit is $\chi^2 \approx k$

$\chi^2 << k$

Data errors are overestimated.
Errors are too large.

$\chi^2 >> k$

Data errors are underestimated.
Intrinsic dispersion

Freedom

Dos and don'ts of reduced chi-squared

René Andrae[1], Tim Schulze-Hartung[1] & Peter Melchior[2]
[1] Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
[2] Institut für Theoretische Astrophysik, ZAH, Albert-Ueberle-Str. 2, 69120 Heidelberg, Germany
e-mail: andrae@mpia-hd.mpg.de

$$\chi^2_{\mathrm{red}} = \frac{\chi^2}{K} \, . \tag{2}$$

χ^2_{red} is a quantity widely used in astronomy. It is essentially used for the following purposes:

1. Single-model assessment: If a model is fitted to data and the resulting χ^2_{red} is larger than one, it is considered a "bad" fit, whereas if $\chi^2_{\mathrm{red}} < 1$, it is considered an overfit.

2. Model comparison: Given data and a set of different models, we ask the question which model fits the data best. Typically, each model is fit to the data and their values of χ^2_{red} are compared. The winning model is that one whose value of χ^2_{red} is closest to one.

3. Convergence diagnostic: A fit is typically an iterative process which has to be stopped when converged. Convergence is sometimes diagnosed by monitoring how the value of χ^2_{red} evolves during the iteration and the fit is stopped as soon as χ^2_{red} reaches a value sufficiently close to one. Sometimes it is claimed then, that "the fit has reached noise level".

4. Error estimation: One fits a certain model to given data by minimising χ^2 and then rescales the data's errors such that the value of χ^2_{red} is exactly equal to one. From this one then computes the errors of the model parameters. (It has already been discussed by Andrae (2010) that this method is incorrect, so we will not consider it any further here.)

Chi2 & Maximum likelihood estimation

Maximum likelihood estimation distribution of data for a given model.

$$P(\mu,\sigma)= \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

$$L \propto \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

Chi2 & Maximum likelihood estimation

If experimental data are subject to measurement error not only in the y_i's, but also in the x_i's, then the task of fitting a straight-line model

$$y(x) = a + bx \tag{15.3.1}$$

We assume that the uncertainty σ_i associated with each measurement y_i is known, and that the x_i's (values of the dependent variable) are known exactly.

$$\chi^2(a,b) = \sum_{i=1}^{N} \left(\frac{y_i - a - bx_i}{\sigma_i} \right)^2$$

If the measurement errors are normally distributed, then this merit function will give maximum likelihood parameter estimations of a and b.

Errors in both x and y

If experimental data are subject to measurement error not only in the y_i's, but also in the x_i's, then the task of fitting a straight-line model

$$y(x) = a + bx \tag{15.3.1}$$

is considerably harder. It is straightforward to write down the χ^2 merit function for this case,

$$\chi^2(a,b) = \sum_{i=1}^{N} \frac{(y_i - a - bx_i)^2}{\sigma_{y\,i}^2 + b^2\sigma_{x\,i}^2} \tag{15.3.2}$$

where $\sigma_{x\,i}$ and $\sigma_{y\,i}$ are, respectively, the x and y standard deviations for the ith point.

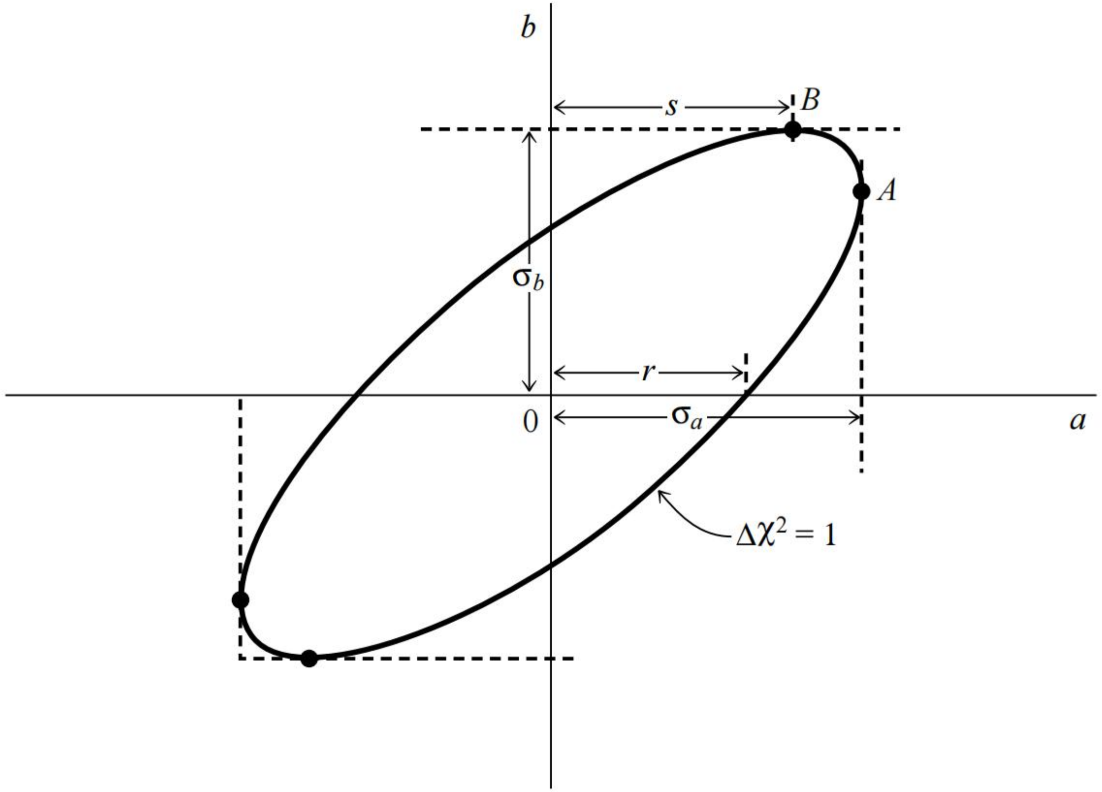
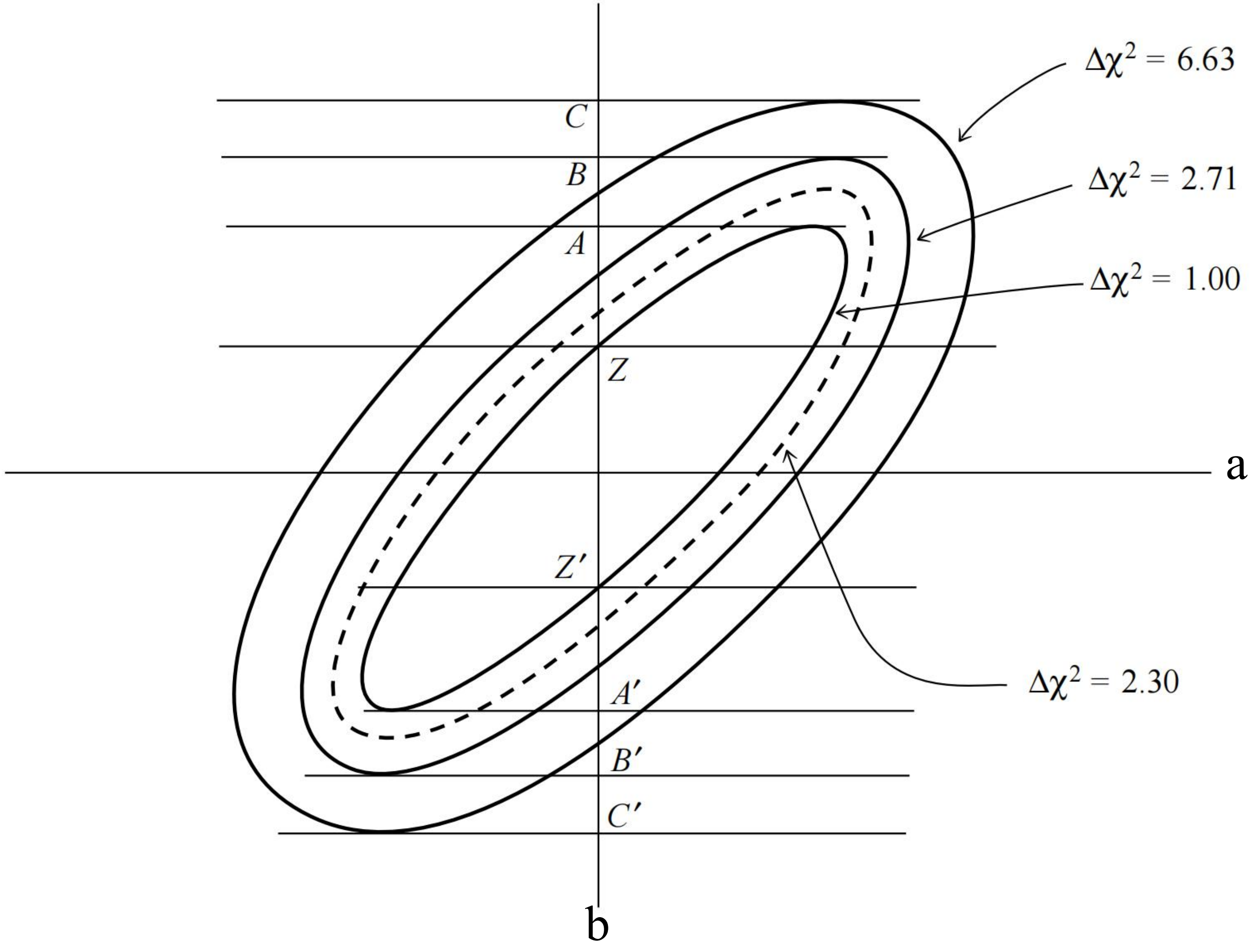


Figure 15.3.1. Standard errors for the parameters a and b. The point B can be found by varying the slope b while simultaneously minimizing the intercept a. This gives the standard error σ_b, and also the value s. The standard error σ_a can then be found by the geometric relation $\sigma_a^2 = s^2 + r^2$.

$$\mathrm{Var}(y_i - a - bx_i) = \mathrm{Var}(y_i) + b^2\mathrm{Var}(x_i) = \sigma_{y\,i}^2 + b^2\sigma_{x\,i}^2$$

Confidence region and confidence level



$$\Delta\chi^2_\nu \equiv \chi^2_\nu - \chi^2_{\min}$$

Confidence region and confidence level

$\Delta\chi^2$ as a Function of Confidence Level and Degrees of Freedom						
	ν					
p	1	2	3	4	5	6
68.3%	1.00	2.30	3.53	4.72	5.89	7.04
90%	2.71	4.61	6.25	7.78	9.24	10.6
95.4%	4.00	6.17	8.02	9.70	11.3	12.8
99%	6.63	9.21	11.3	13.3	15.1	16.8
99.73%	9.00	11.8	14.2	16.3	18.2	20.1
99.99%	15.1	18.4	21.1	23.5	25.7	27.8

Nonnormal errors

- to fit for parameters by minimizing χ^2
- to use a contour of constant $\Delta\chi^2$ as the boundary of your confidence region
- to use Monte Carlo simulation or detailed analytic calculation in determining *which* contour $\Delta\chi^2$ is the correct one for your desired confidence level
- to give the covariance matrix C_{ij} as the "formal covariance matrix of the fit."

Useful webpages

https://github.com/statsmodels/statsmodels/blob/main/examples/python/chi2_fitting.py

https://probfit.readthedocs.io/en/latest/api.html#probfit.costfunc.Chi2R