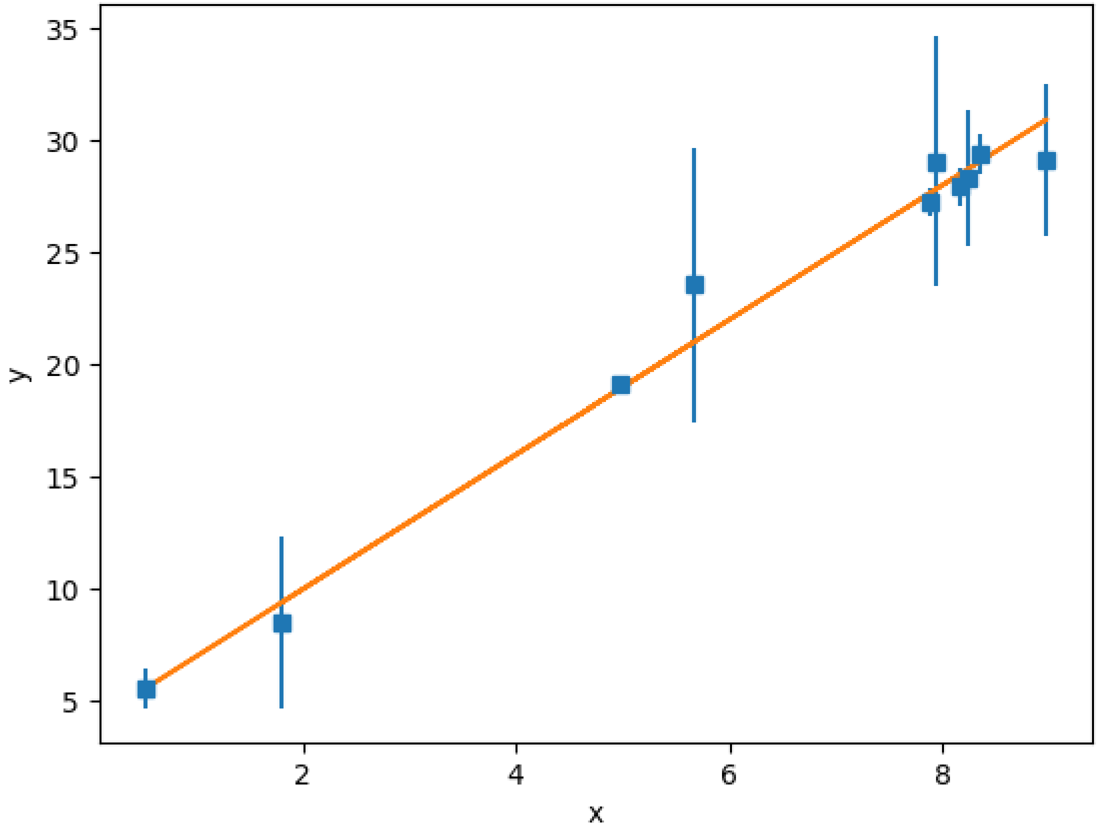
Modeling of data

from least square fitting to Bayesian approach

Data and model

- Data: (x_i , y_i) i=1,N
  - Uncertainty of y_i : $yerr_i$
- Model: $y=f(x)=ax+b$



- x=[4.97 5.67 1.8 7.89 7.94 8.98 8.35 8.24 0.51 8.17]
- y=[19.12 23.55 8.5 27.23 29.08 29.14 29.43 28.34 5.6 27.94]
- yerr=[0.17 6.12 3.86 0.62 5.54 3.36 0.9 3.03 0.89 0.84]

Least square fitting

Suppose that we are fitting N data points (x_i, y_i) $i = 1, \ldots, N$, to a model that has M adjustable parameters a_j, $j = 1, \ldots, M$. The model predicts a functional relationship between the measured independent and dependent variables,

$$y(x) = y(x; a_1 \ldots a_M) \tag{15.1.1}$$

where the dependence on the parameters is indicated explicitly on the right-hand side.

What, exactly, do we want to minimize to get fitted values for the a_j's? The first thing that comes to mind is the familiar least-squares fit,

$$\text{minimize over } a_1 \ldots a_M : \quad \sum_{i=1}^{N} [y_i - y(x_i; a_1 \ldots a_M)]^2 \tag{15.1.2}$$

Chi-square fitting

- Data points can not have the same error

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

Understand as a weight of each data in least square fit

- Chi-square distribution with $\nu = N - M$ degrees of freedom
  - Model is perfect
  - Measurement error is right and Gaussian
  - Sample is not biased
- What does it mean if we get $\chi^2=30$ for freedom 10?
  - $P(\chi^2 > 30|10)=0.001$: the probability we reject one of above assumptions is wrong (At least one of the above assumption is wrong at 99.9% level.)

χ^2 distribution

χ^2 distribution with freedom of K is the a sum of the squares of K independent standard normal random variables.

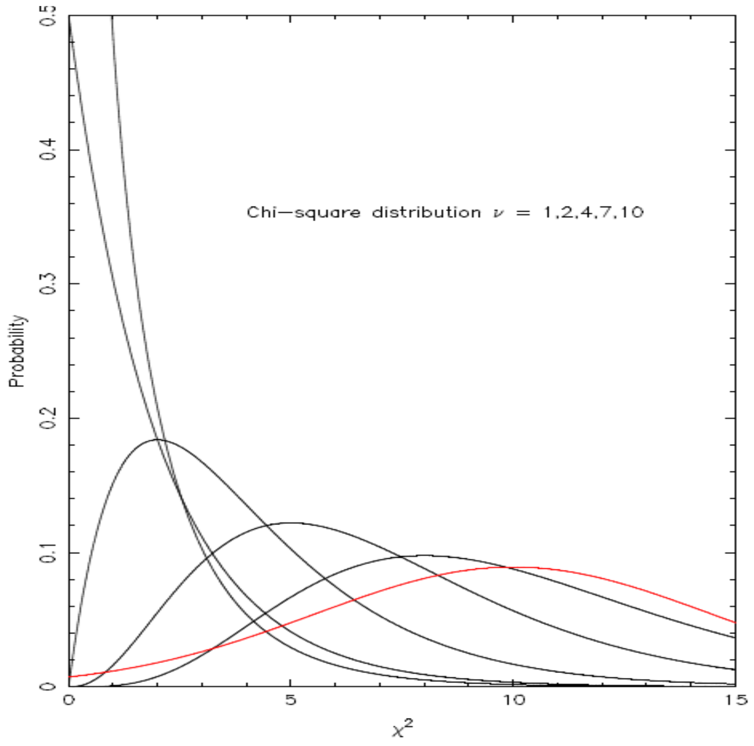


Figure 1: Examples of χ^2 distribution - N(10,20) Gaussian in red

$$P(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2}$$

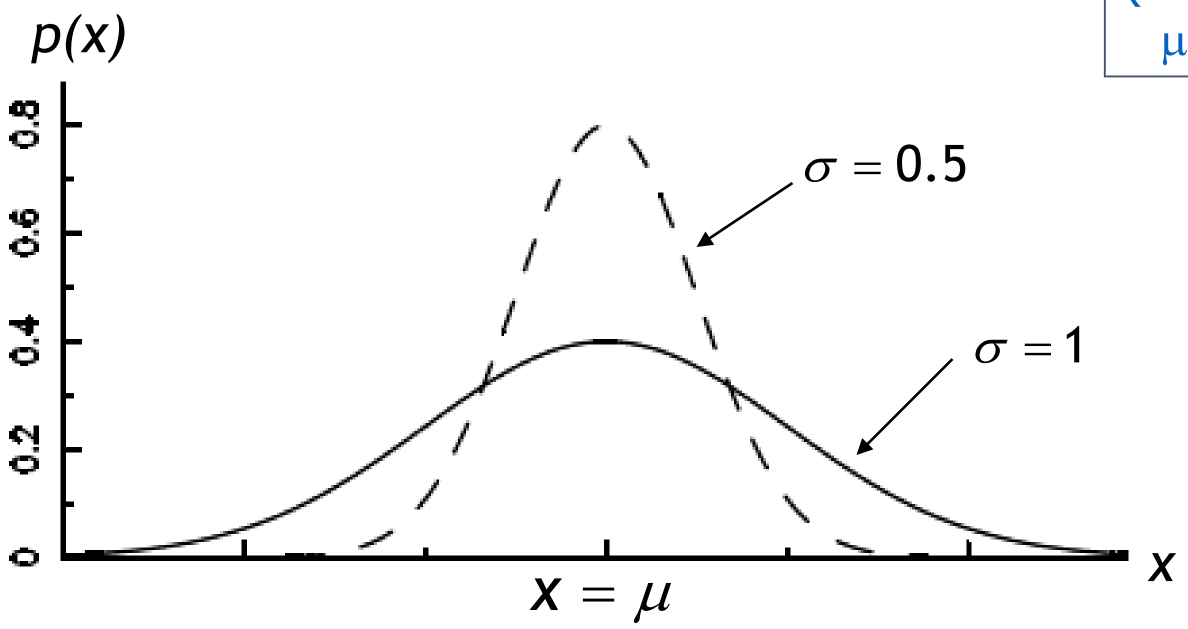
Approximate Gaussian again when K is large

Mean: *K* Variance: *2K*

Gaussian (normal) distribution

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

X follow Gauss (μ,σ)
(X-u)/σ follow Standard normal
μ=0 σ=1



Mean: μ Variance: σ^2

Notes on χ^2

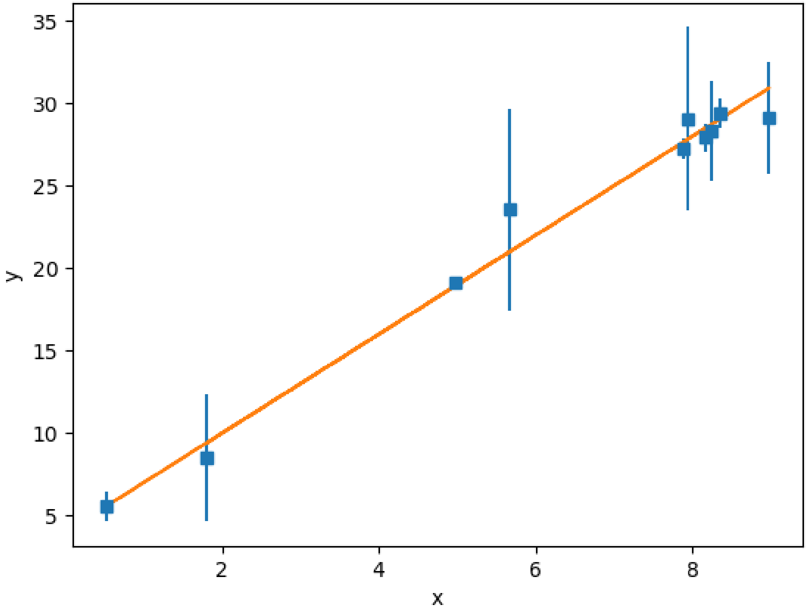
- If $\chi^2 \sim \nu$, model is reasonable
- If $\chi^2 << \nu$
  - Error is overestimated
  - Error is too large to distinguish models
- If $\chi^2 >> \nu$ (most of cases)
  - Ideal model never exists, e.g. Scaling relations
  - Data errors are underestimated
  - Data are always biased

freedom is a question! see arXiv:1012.3754

Maximum likelihood estimation

- $P(y_i | x_i) = \text{Gauss}(\mu, \sigma)$
  - $\mu = a*x_i + b$
  - $\sigma = yerr_i$
- $L \propto \prod_{i=1,N} P(y_i|x_i)$
  - $\ln L = -\frac{1}{2}\sum_{i=1,N}\left(\frac{y_i-(ax_i+b)}{yerr_i}\right)^2 + const$

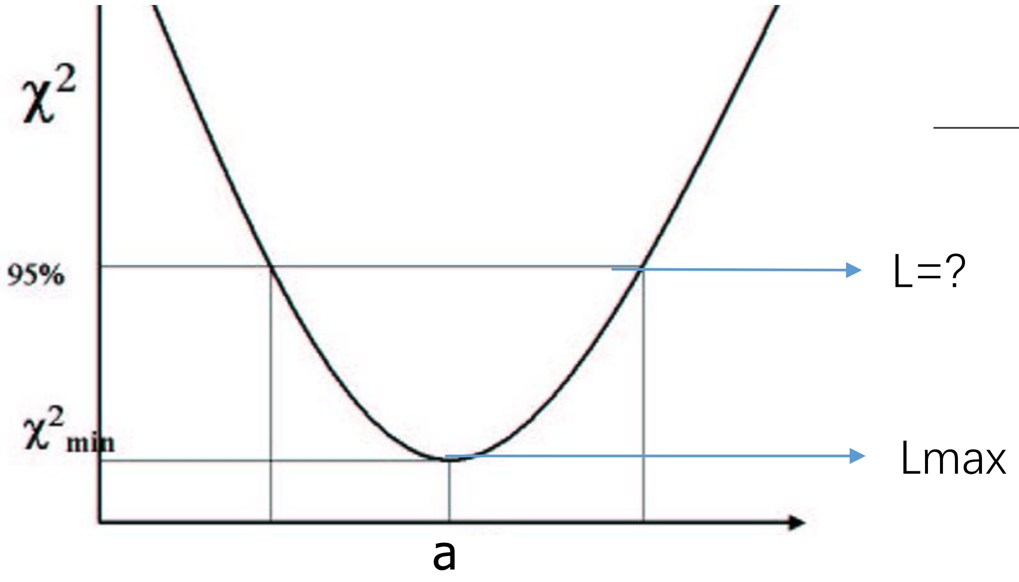
    $= -\frac{1}{2}\chi^2$ +const !!!
- L(max) $\leftrightarrow \chi^2$(min)

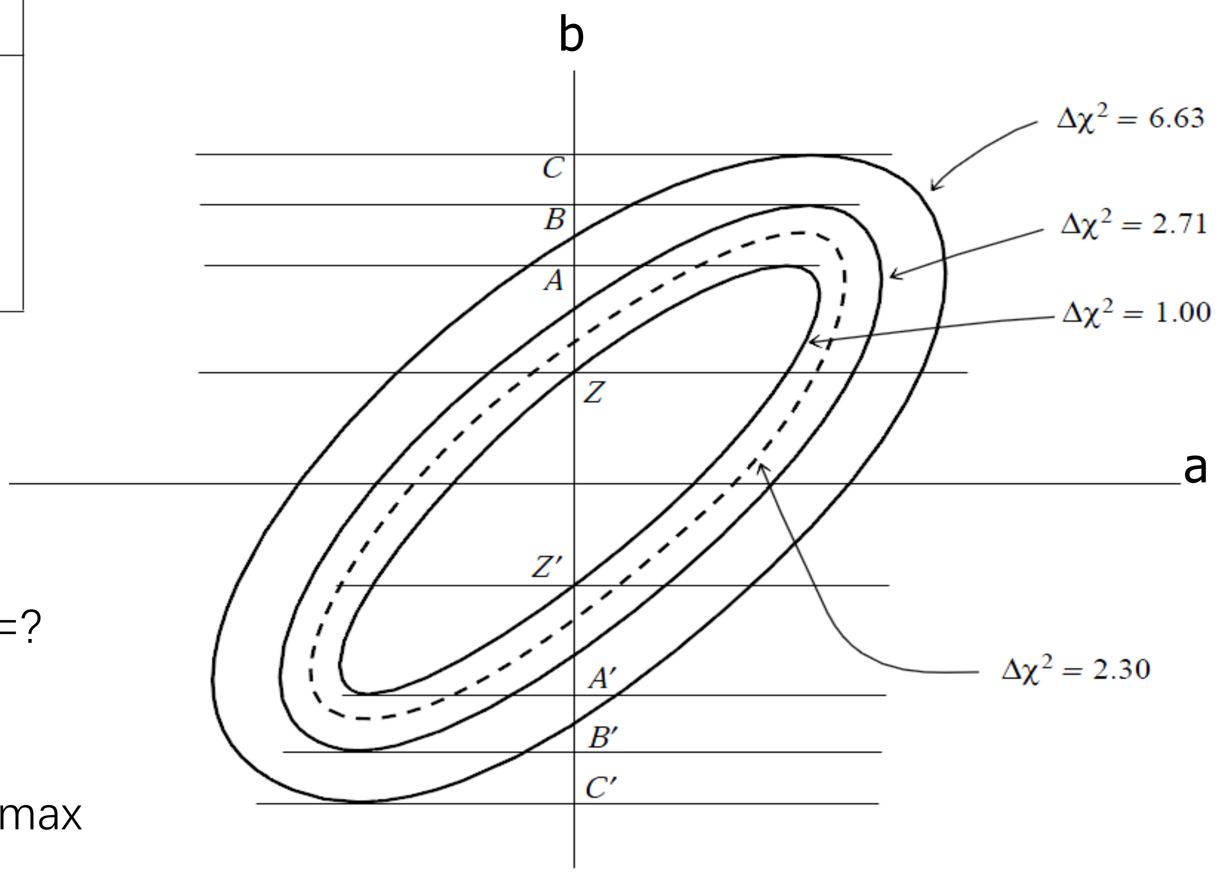


$$p(x) = \frac{1}{\sqrt{2\pi\sigma}}\exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

Confidence level of parameter estimation

$\Delta\chi^2$ 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





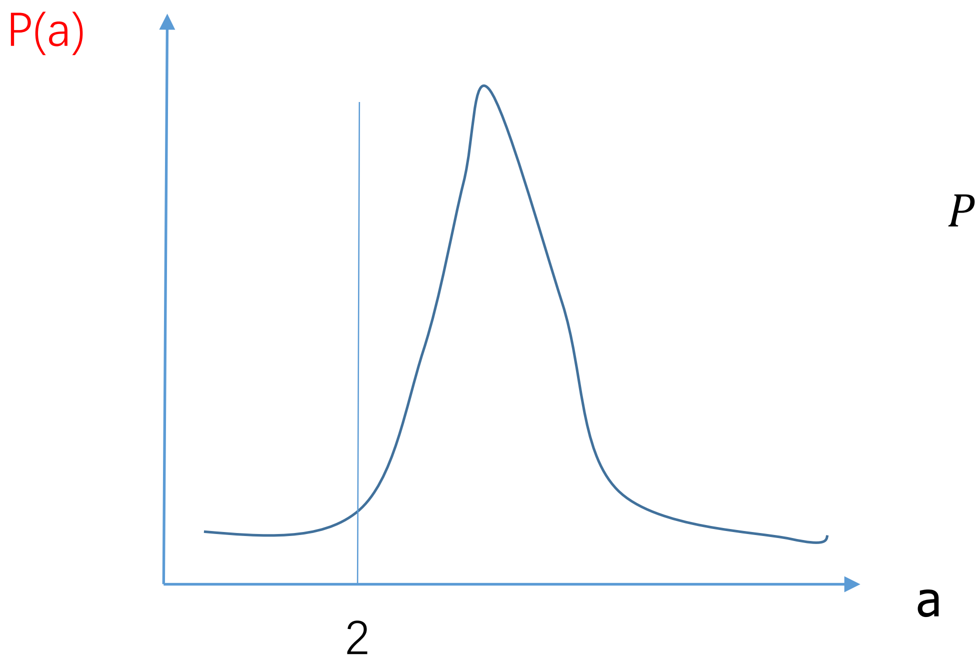
Bayesian probability

- Frequentist
  - Data are random, probability is frequency of data
  - Cannot refer to the probability of a hypothesis (either true or false)
- Bayesian
  - Data are not random (in astronomy!)
  - Evaluate the probability of a hypothesis in light of data (and prior information)
    - P(M|D)
      - $\int P(M|D)dM = 1$

Bayesian approach

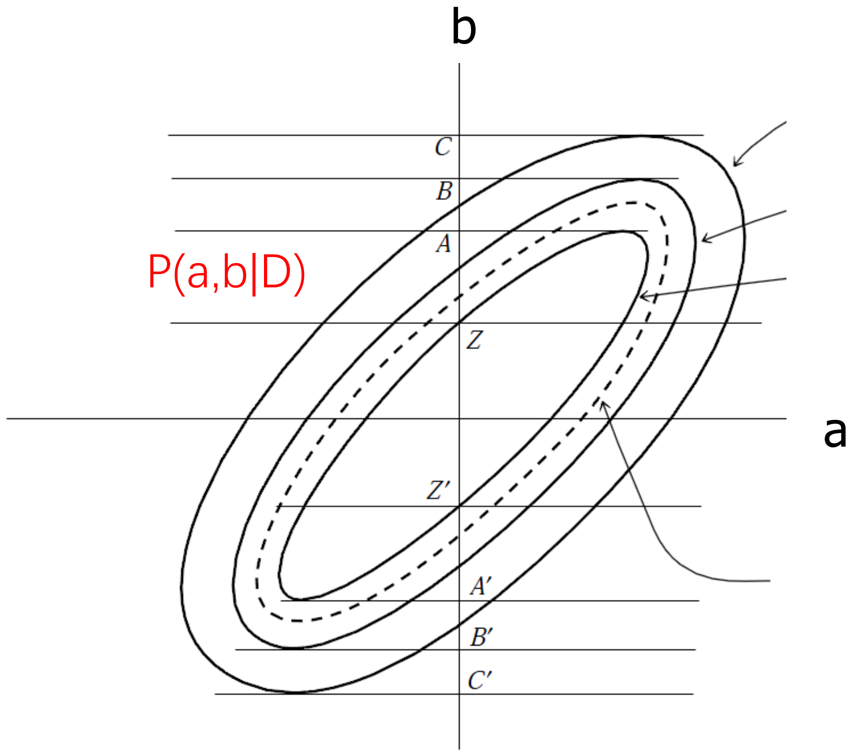
- Bayesian theorem: $P(M|D) = \frac{P(M)P(D|M)}{P(D)}$
  - P(M|D): P(a,b)|(x_i , y_i ,$yerr_i$)
  - P(D|M)
    - P(y|x,yerr,a,b): i.e. maximum likelihood !
  - P(D): = 1
  - P(M): prior on model
    - If unbiased to any models P(M|D)=P(D|M)
    - if I know a>2, i.e. P(a<2)=0 ?

Uncertainty of Bayesian estimation



$$P(a) = \int P(a,b)db$$

$$\int P(a)da = 1$$



Summary

- Bayesian estimation are consistent with ML estimation
  - when priors are uniform

- ML estimation are consistent with χ^2 fitting
  - When data and model are perfect

- χ^2 fitting are consistent with least square fitting
  - When errors are considered as weight

Differences: uncertainty estimation!