ASTR21200

Observational Techniques in Astrophysics
Lecture-7

Bradford Benson

Next ~couple weeks

- Today: Brief lecture to setup concepts for Homework-3 and Lab-2
- Homework-3 posted on Canvas:
  - Due Tuesday Apr-29 at 5pm
- Lab-2 posted on Canvas:
  - Due Tuesday May-6 at 5pm
  - Will have to do several observations, so important to plan to take observations mostly in the next ~week.

Code Issues Pull requests Actions Projects Wiki Security Insights Settings

Schedule Spring 2025

bradfordbenson edited this page now · 25 revisions

Edit New page

Week	Date	Topic	Lecture	Homework / Lab	Tutorial
1	Mar-25	Intro to Astro Observing	Lect-1	HW-1, Due Apr-1	Python-1: Visibility
	Mar-27	Practical Observing	Lect-2		
2	Apr-1	CCDs and Astronomical Images	Lect-3	HW-2, Due Apr-8	Python-2: CCD Images
	Apr-3	Intro to Stone Edge	Lect-4		Python-3: Astropy Fits
3	Apr-8	Intro to Labs and Lab1	Lect-5	Lab-1, Due Apr-22	SEO Cheat Sheet
	Apr-10	(Analysis and Help/Hack Session)			Python-4: RGB Images
4	Apr-15	(Analysis and Help/Hack Session)	Lect-6	HW-3, Due Apr-29	Python-5: Gaussian Fits
	Apr-17	(Analysis and Help/Hack Session)			
5	Apr-22	Intro to Lab2, More Statistics	[Lect-7]	Lab-2, Due May-6	
	Apr-24	(Analysis and Help/Hack Session)			
6	Apr-29	(Analysis and Help/Hack Session)			
	May-1	(Analysis and Help/Hack Session)			
7	May-6	Intro to Lab 3, Project Ideas	[Lect-8]	Lab-3, Due May-22	
	May-8	(Analysis and Help/Hack Session)			
8	May-13	(Analysis and Help/Hack Session)			
	May-15	(Analysis and Help/Hack Session)			

Pages 17

General Information

- Syllabus
- Schedule
- Stone Edge Observatory (SEO)
- SEO Observing Calendar
- SEO Data Archives

Labs and Observing

- Lab Report Guidelines
- Lab1
- Lab2
- Lab3
- Astro Data Archives

Computing Resources

- Astronomy Software
- Python
- GitHub
- ds9
- Source Extractor
- Jupyter
- LaTeX
- Topcat
- Astrometry.net
- Coadd or Stack Images
- Image Add, Sub, Arithmetic
- Awk and Sed

Clone this wiki locally

https://github.com/bradfordbenso

Types of Errors

Statistical errors — Uncertainty due to known statistical effects, such as:
1) photon counting statistics (e.g., Poisson/shot),
2) detector or measurement noise (e.g., thermal noise), or
3) other sources of variance.

Systematic errors — Uncertainty not due to statistical uncertainty. Typically related to some unquantified measurement bias, or instrumental effect. These typically are multiplicative or additive corrections to the signal, and do not follow Gaussian statistics.

Suggested References:
Barlow *SLAC Lecture 5– systematic errors*
Barlow: *systematic errors: facts and fictions*

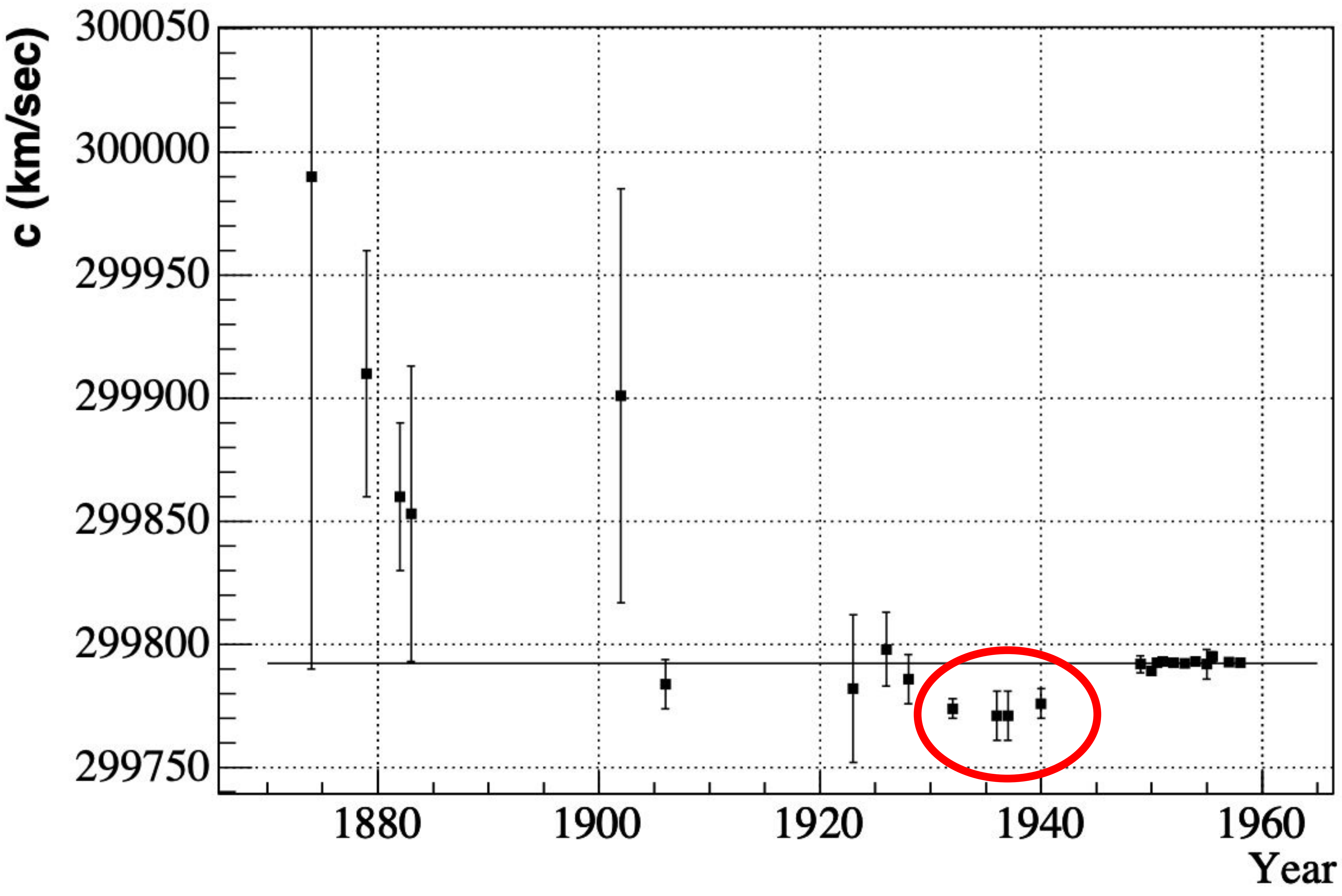
Systematic and statistical errors are quoted separately

You will often see results like: $\sigma = 45 \pm 4 \pm 1mb$

- The **first** number is the **most-likely value**
- The **second** number is the **statistical uncertainty** (typically the 68% confidence range)
- The **third** number is the **systematic uncertainty** (also typically the 68% confidence range, but for just systematic uncertainty).

This facilitates combining uncertainty (or errors), but also shows the relative importance of the two factors. For example, systematic uncertainty often is not reduced by taking more data.

Systematic errors: The Speed of Light



Summary of measurements of the speed of light from 1870-1960:

- One striking feature is the of ~17 km/sec, between 1930-1940 and later determinations.
- Later analyses concluded this was due to ***"observer bias"***,
    - i.e., experimenters didn't want to get a result "inconsistent" with the accepted value
    - Beware "observer bias" in your labs!

Calculating the "Chi-squared": χ^2

Definition:

$$\chi^2 = \sum_{i=1}^{N} \frac{(x_i - \mu)^2}{\sigma_i^2}$$

x_i, are the measurements (N total)
σ_i is the measurement uncertainty
μ is the mean (or model prediction)

We expect $\chi^2 = N_{dof} = N - N_{model}$ (so each data point contributes χ^2~1)

Where N_{dof} is the number of degrees of freedom.

Calculating a χ^2 is useful to:

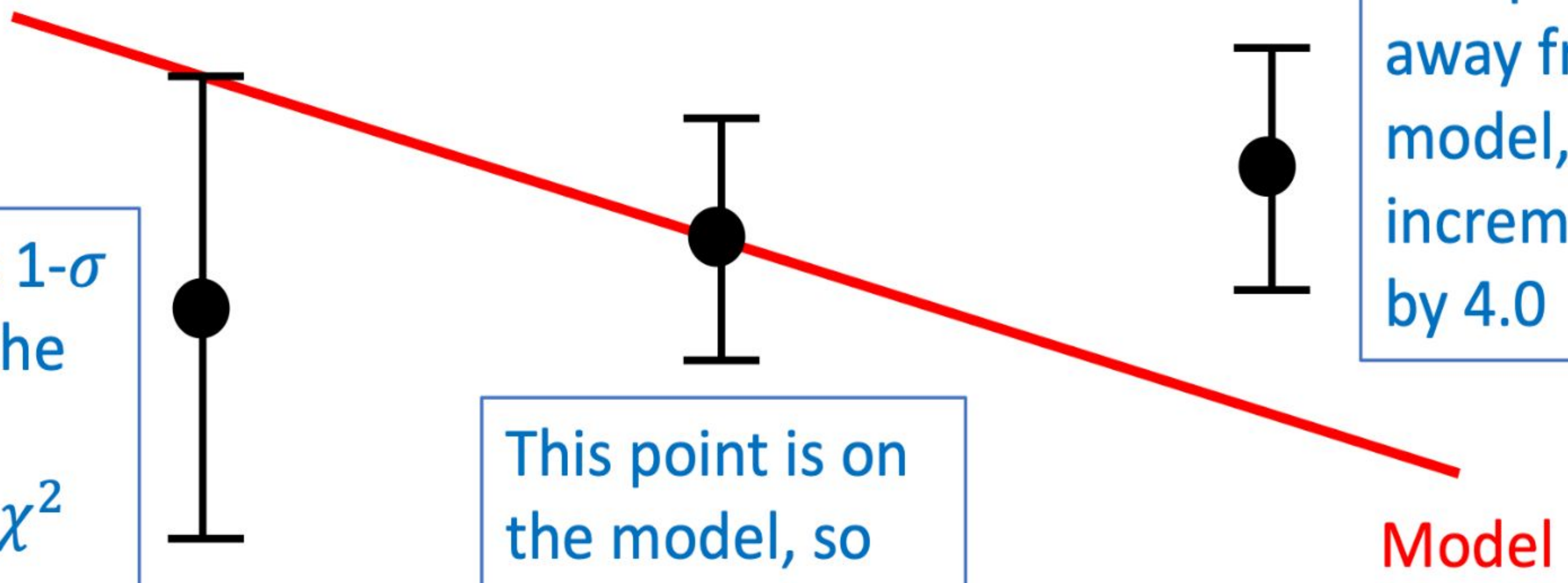
(1) **Test how good the model is.** There is a statistical expectation of what "chi-squared" should be (i.e., the chi-squared distribution), from which we can calculate a ***probability to exceed (PTE)***, which depends on χ^2 and N_{dof}

(2) **Estimate the "true" uncertainty/error**. We can compute the chi-squared and set it equal to N to estimate the uncertainty.

Calculating the “Chi-squared”: χ^2

$$\chi^2 = \sum_{i=1}^{N} \frac{(x_i - \mu)^2}{\sigma_i^2}$$

χ^2 is a sum over the data points



This point is 1-σ away from the model, so increments χ^2 by 1.0

This point is on the model, so increments χ^2 by 0.0

This point is 2-σ away from the model, so increments χ^2 by 4.0

Probability To Exceed (PTE) (the Pearson Test)

A chi-squared result should follow a chi-squared probability distribution ***f(χ²)*** with N degrees of freedom.

On-average, the measured ***χ²*** is expected to be roughly equal to the N degrees of freedom, i.e., $\chi^2 = N_{dof} = N - N_{model}$

The "PTE" tells you how likely you would have been to get your result,. e.g.,

The PTE on average should be 0.5, and there should be only a 2% chance that you find a PTE < 0.01 or > 0.99

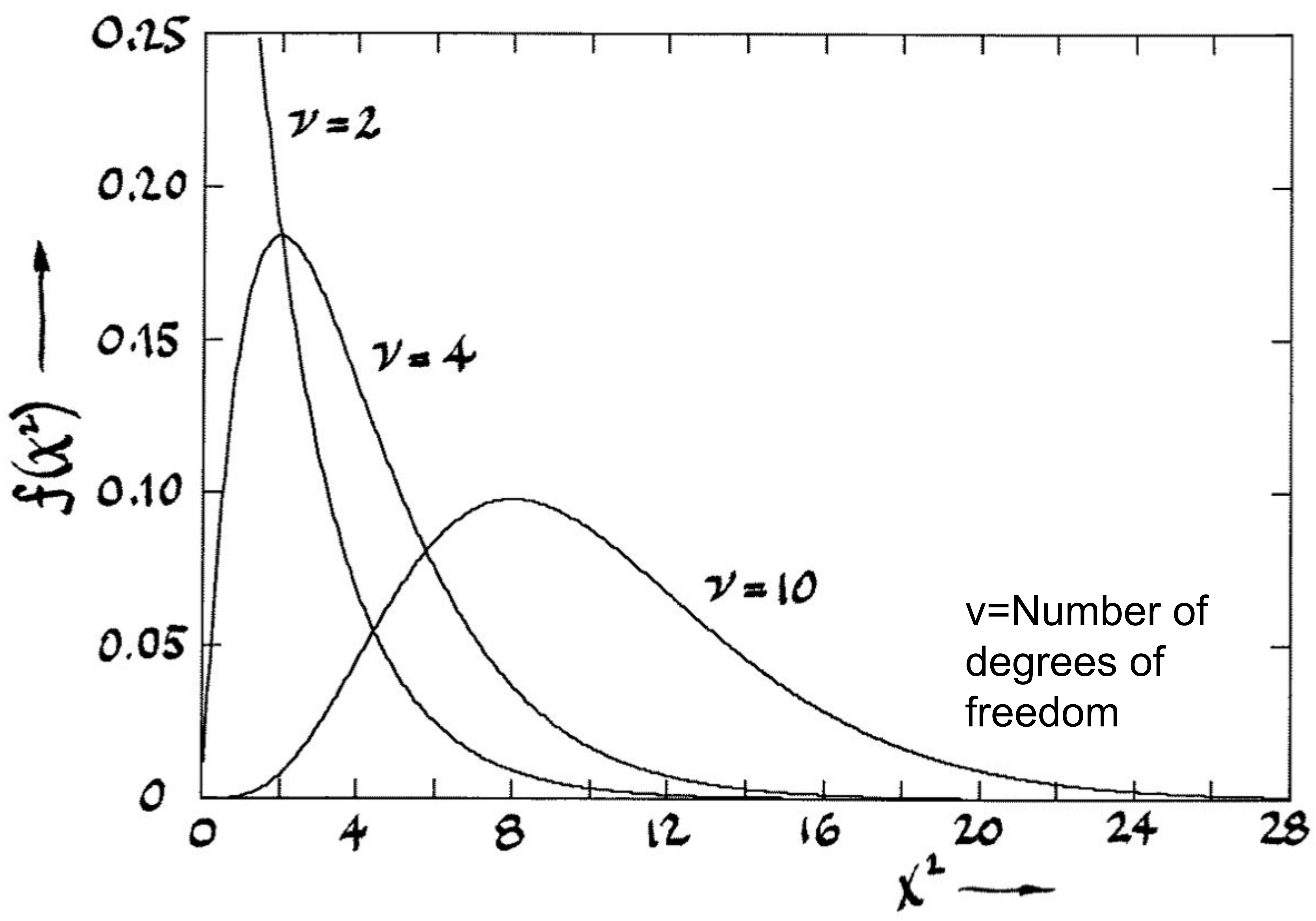


Figure 1 — The chi-square distribution for $\nu = 2$, 4, and 10.

e.g., Bevington (Data Reduction and Error Analysis for the Physical Sciences)

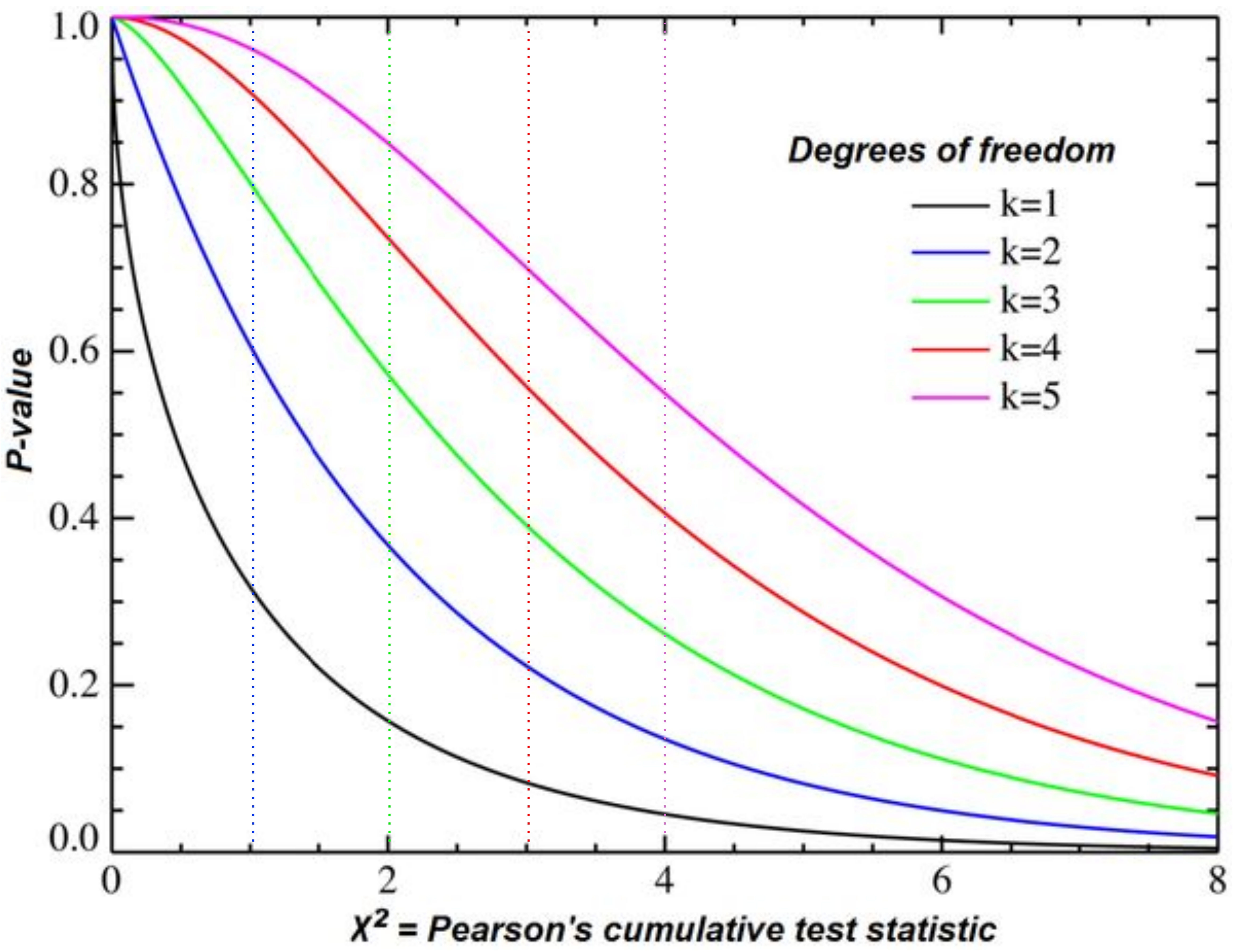
Probability To Exceed (PTE) (the Pearson Test)

A chi-squared result should follow a chi-squared probability distribution ***f(χ²)*** with N degrees of freedom.

On-average, the measured ***χ²*** is expected to be roughly equal to the N degrees of freedom, i.e., $\chi^2 = N_{dof} = N - N_{model}$

The "PTE" tells you how likely you would have been to get your result,. e.g.,

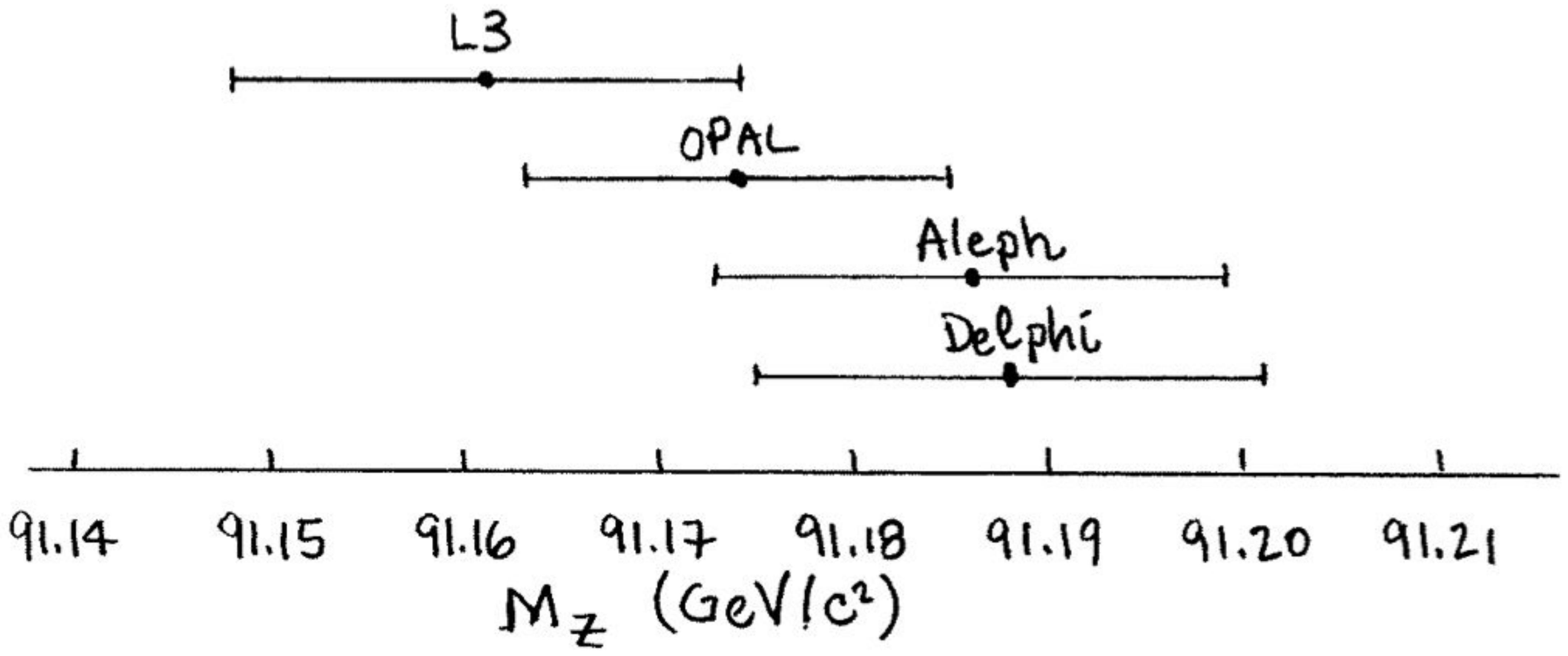
> The PTE on average should be 0.5, and there should be only a 2% chance that you find a PTE < 0.01 or > 0.99



e.g., Bevington (Data Reduction and Error Analysis for the Physical Sciences)

PTE Example: Four measurements of Z-Boson

Detector	Mass in GeV/c^2
L3	91.161 ± 0.013
OPAL	91.174 ± 0.011
Aleph	91.186 ± 0.013
Delphi	91.188 ± 0.013



Measurements of the Z^0 boson.

Calculate a ***weighted average (M)*** and ***uncertainty (σ)*** from the 4x data points:

$$\overline{M}_Z = \frac{\sum M_i/\sigma_i^2}{\sum 1/\sigma_i^2} \qquad \sigma^2_{\overline{M}_Z} = \frac{1}{\sum 1/\sigma_i^2}$$

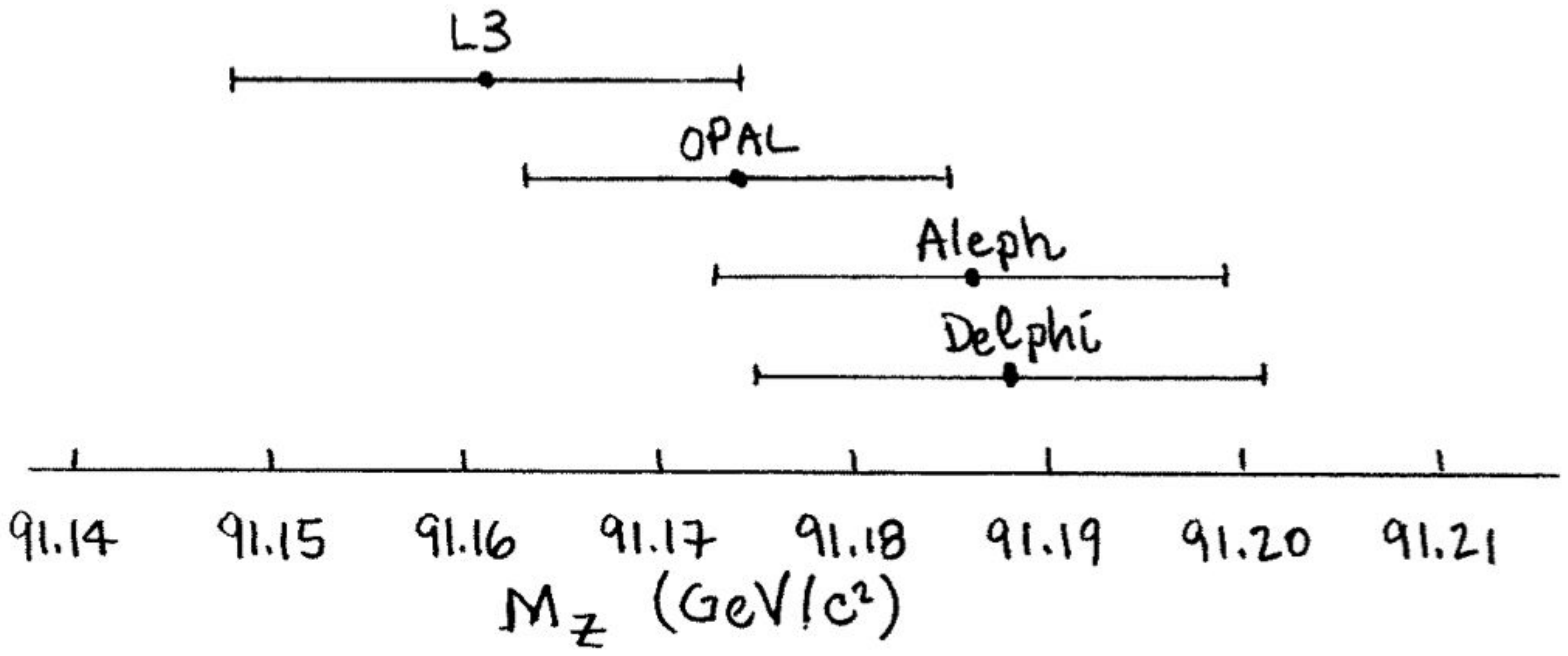
Each data point will contribute a "**weight**" based on the inverse of its uncertainty squared ($1/\sigma_i^2$), e.g., 10x more uncertainty means 100x less weight.

$$\overline{M}_Z \pm \sigma_{\overline{M}_Z} = 91.177 \pm 0.006$$

We have 4x measurements with similar uncertainty, so expect the weighted average's uncertainty to be ~ sqrt(4) = 2x less

PTE Example: Are the data consistent?

Detector	Mass in GeV/c^2
L3	91.161 ± 0.013
OPAL	91.174 ± 0.011
Aleph	91.186 ± 0.013
Delphi	91.188 ± 0.013



Measurements of the Z^0 boson.

Calculate via the chi-squared/chi2:

$$\chi^2 = \sum_{i=1}^{4} \frac{(M_i - \overline{M}_Z)^2}{\sigma_i^2} \approx 2.78$$

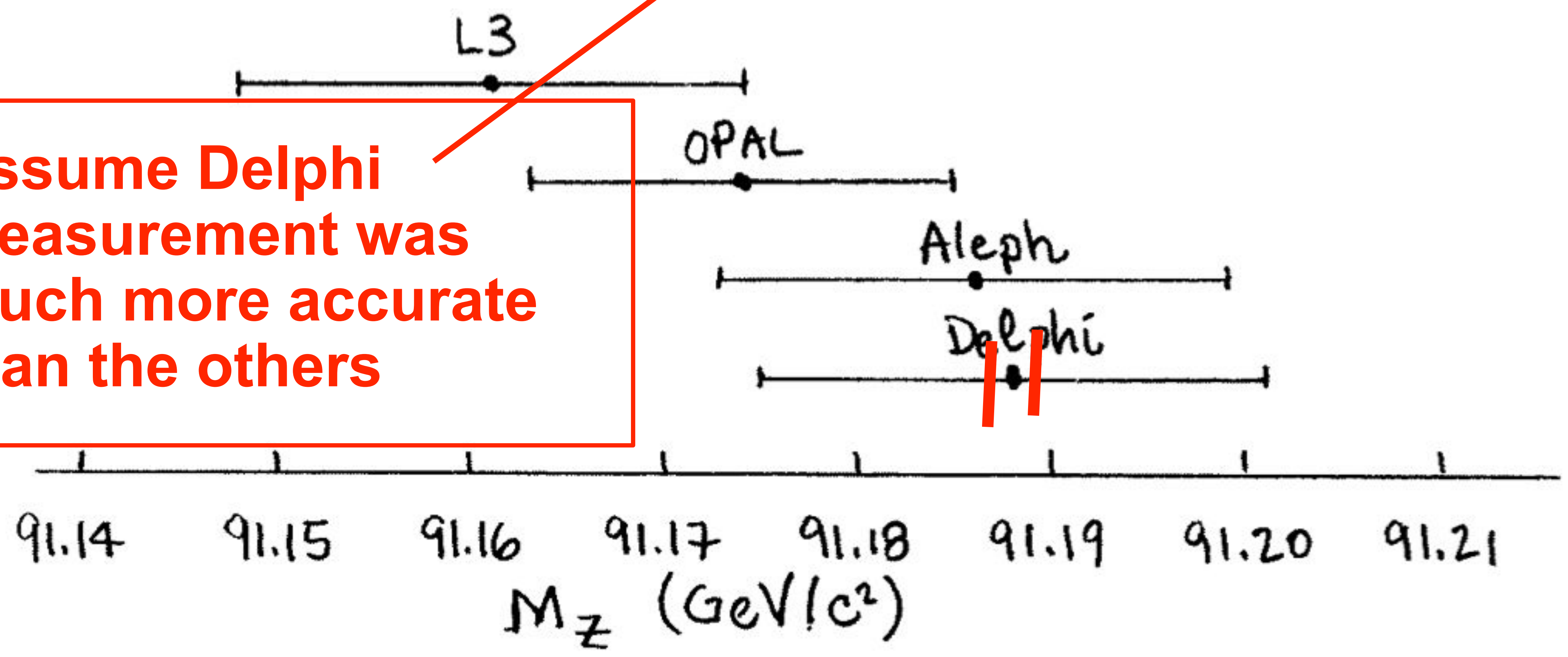
We expect this chi2 to be drawn from a chi-squared distribution with 3 degrees of freedom (DOF) (i.e., 4x data points – 1x model parameter, the mean value)

The integrated probability ***f*(χ^2)** to get a chi2 of 2.78 or higher with 3 DOF is 0.42 (42%)

PTE Example: Four measurements of Z-Boson

Detector	Mass in GeV/c^2
L3	91.161 ± 0.013
OPAL	91.174 ± 0.011
Aleph	91.186 ± 0.013
Delphi	$\cancel{91.188 \pm 0.013}$

91.188+/-0.002



Measurements of the Z^0 boson.

Assume Delphi measurement was much more accurate than the others

Calculate a ***weighted average (M)*** and ***uncertainty (σ)*** from the 4x data points:

$$\overline{M}_Z = \frac{\sum M_i/\sigma_i^2}{\sum 1/\sigma_i^2} \qquad \sigma^2_{\overline{M}_Z} = \frac{1}{\sum 1/\sigma_i^2}$$

$$\overline{M}_Z \pm \sigma_{\overline{M}_Z} = \cancel{91.177 \pm 0.006}$$

91.187+/-0.0019

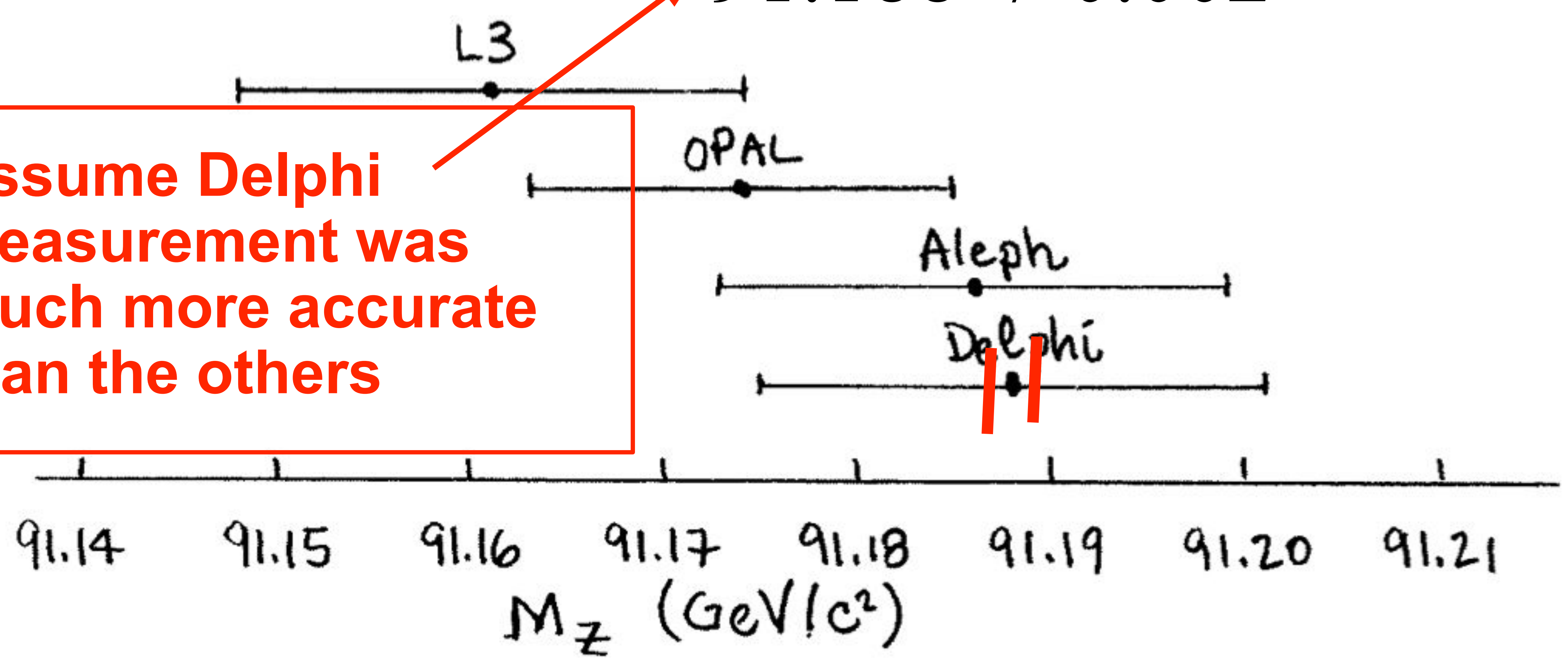
Now the Delphi measurement is weighted higher by a factor of ~30-40 than the other measurements and dominates final results.

Is this still a good chi2?

PTE Example: Four measurements of Z-Boson

Detector	Mass in GeV/c^2
L3	91.161 ± 0.013
OPAL	91.174 ± 0.011
Aleph	91.186 ± 0.013
Delphi	~~91.188 ± 0.013~~
	91.188+/-0.002

Assume Delphi measurement was much more accurate than the others



Measurements of the Z^0 boson.

$$\overline{M}_Z \pm \sigma_{\overline{M}_Z} = \cancel{91.177 \pm 0.006}$$

91.187+/-0.0019

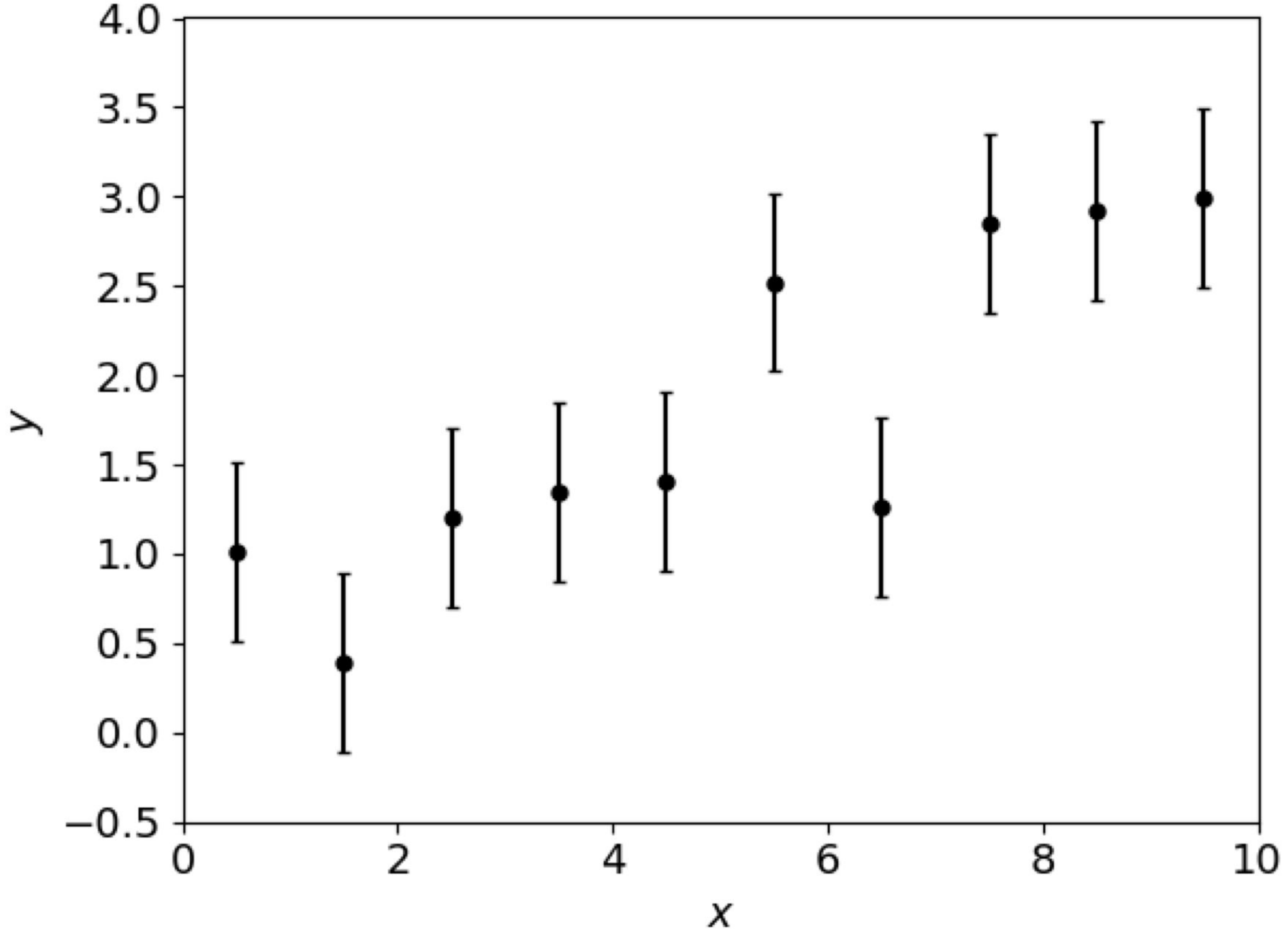
Calculate via the chi-squared/chi2:

$$\chi^2 = \sum_{i=1}^{4} \frac{(M_i - \overline{M}_Z)^2}{\sigma_i^2} \approx 5.65$$

The integrated probability ***f(χ²)*** to get a chi2 of 5.65 or higher with 3 DOF is 0.13 (13%), which is still a reasonably likely result.

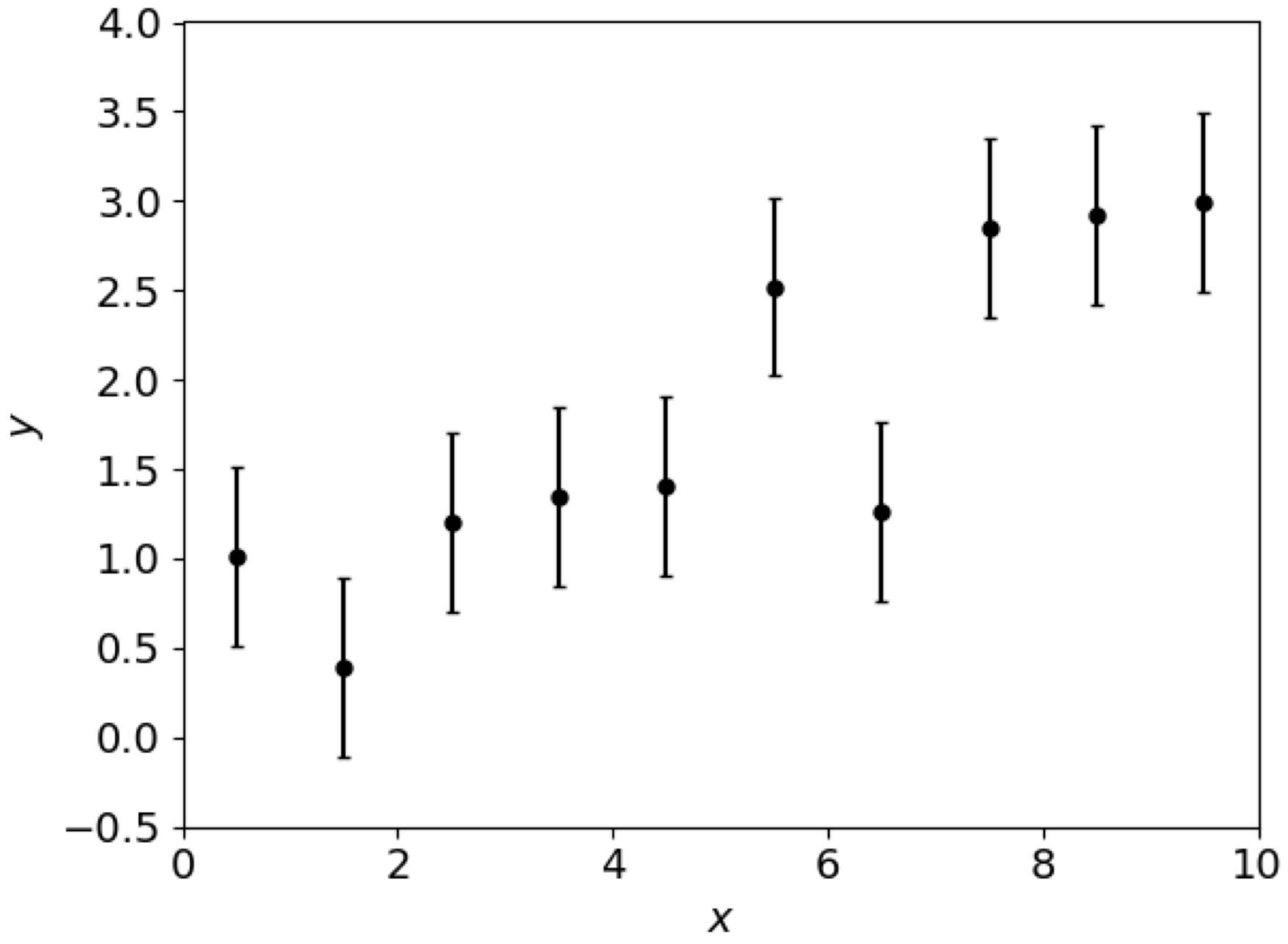
Parameter Estimation using the Likelihood & χ^2

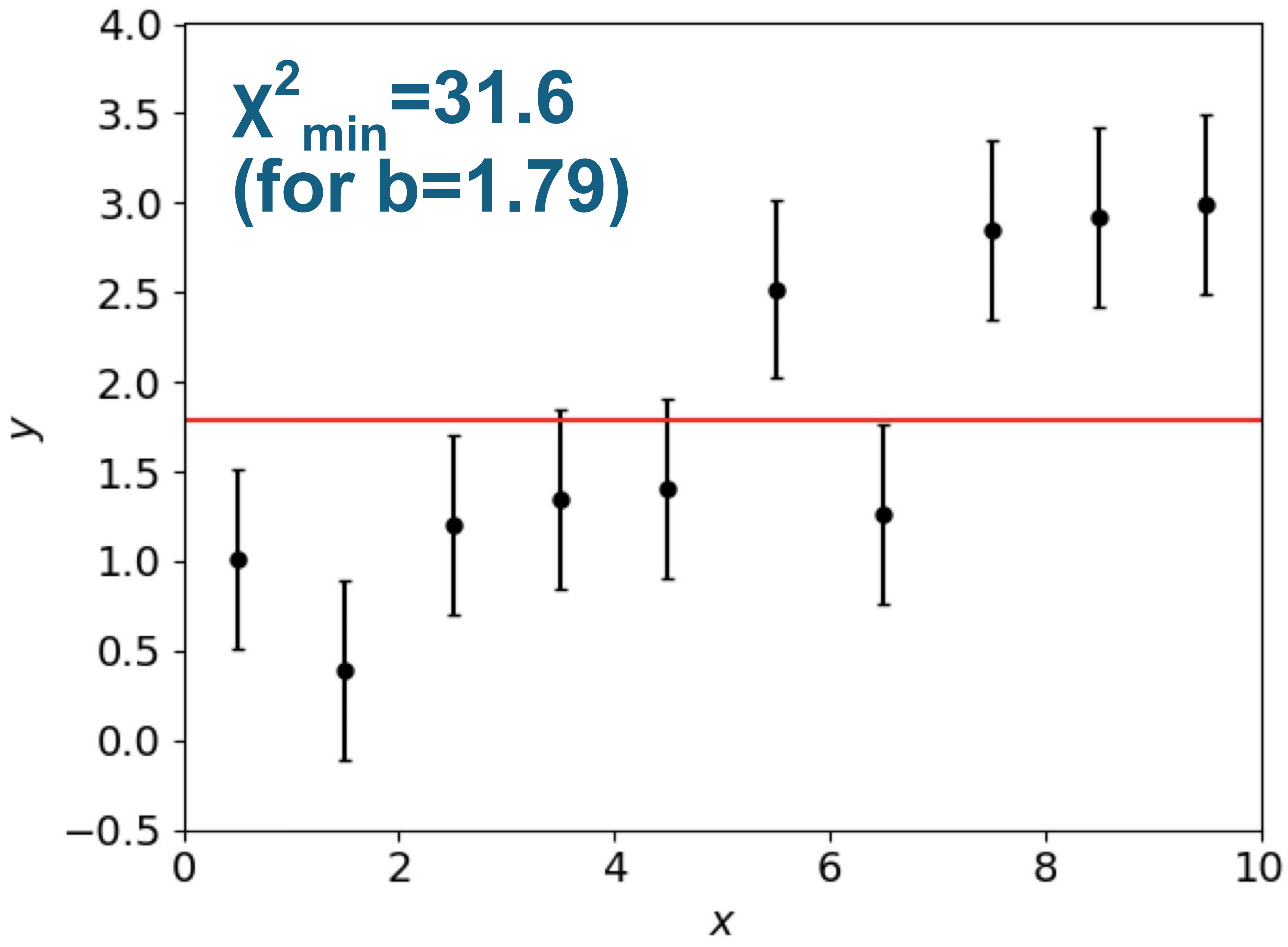
We have a data set with 10 points. Are they well fit by a constant (y=b)?



Parameter Estimation using the Likelihood & χ^2

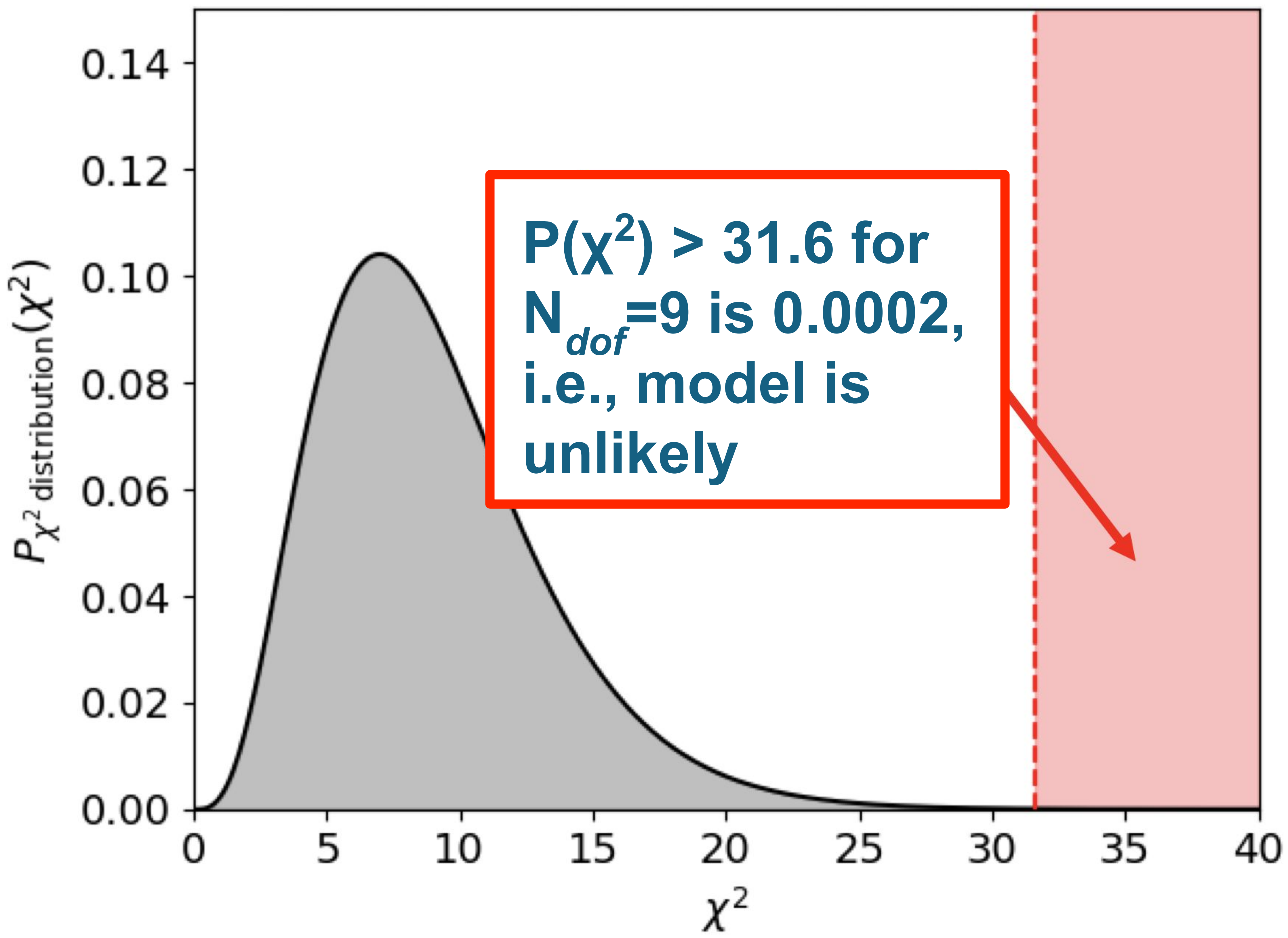
We have a data set with 10 points. Are they well fit by a constant (y=b)?

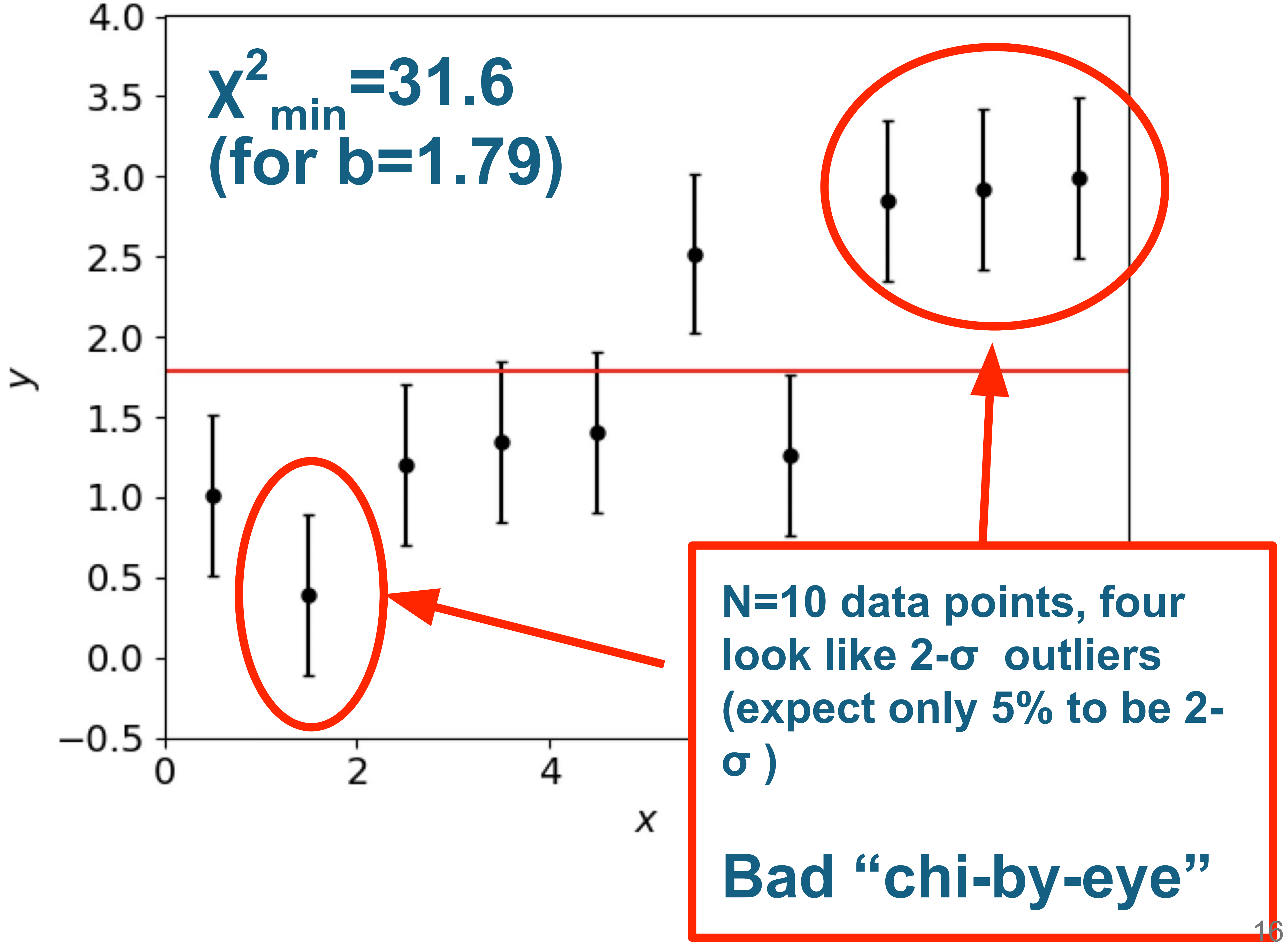




Parameter Estimation using the Likelihood & χ^2

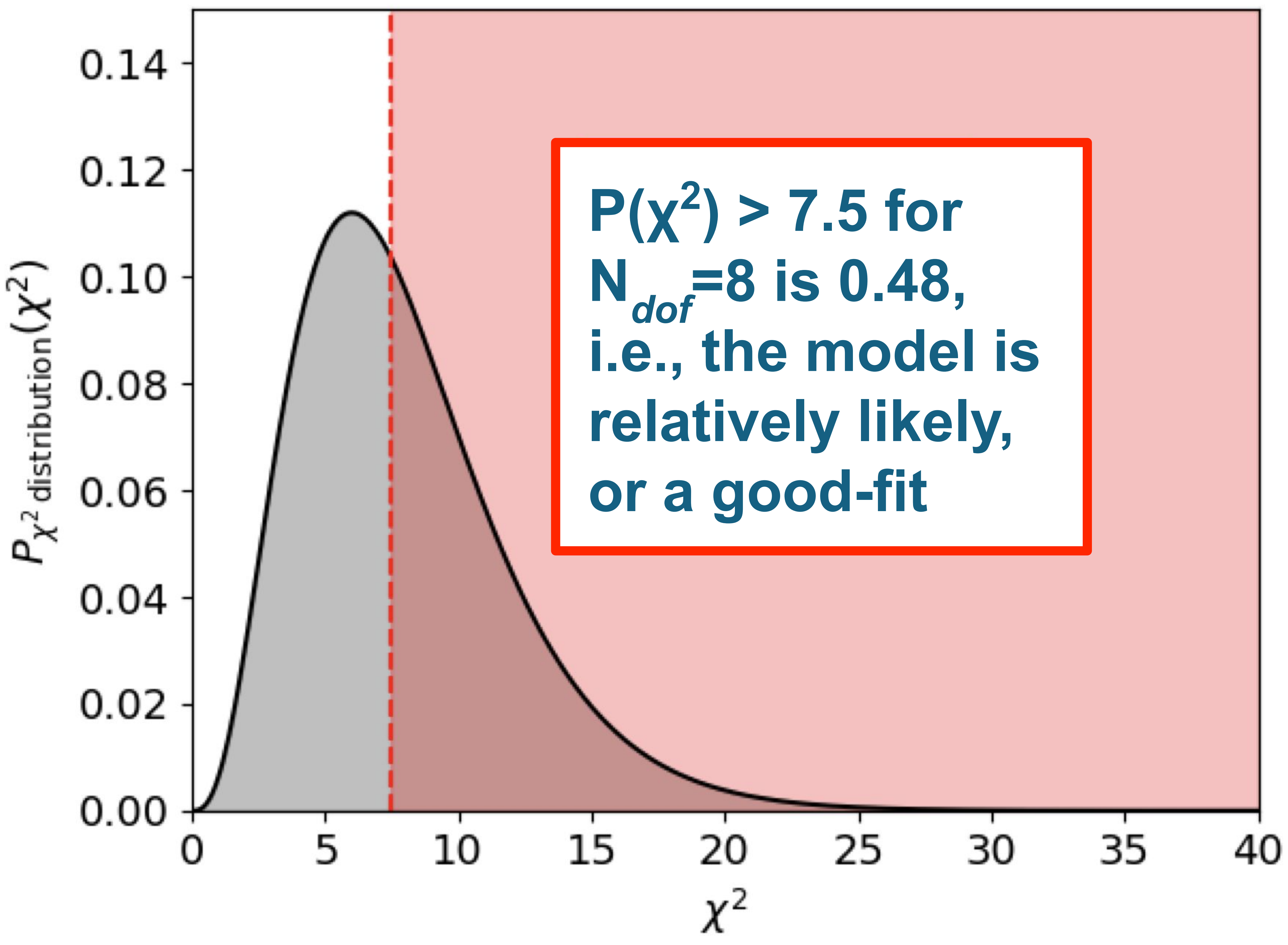
We have a data set with 10 points. Are they well fit by a constant (y=b)?

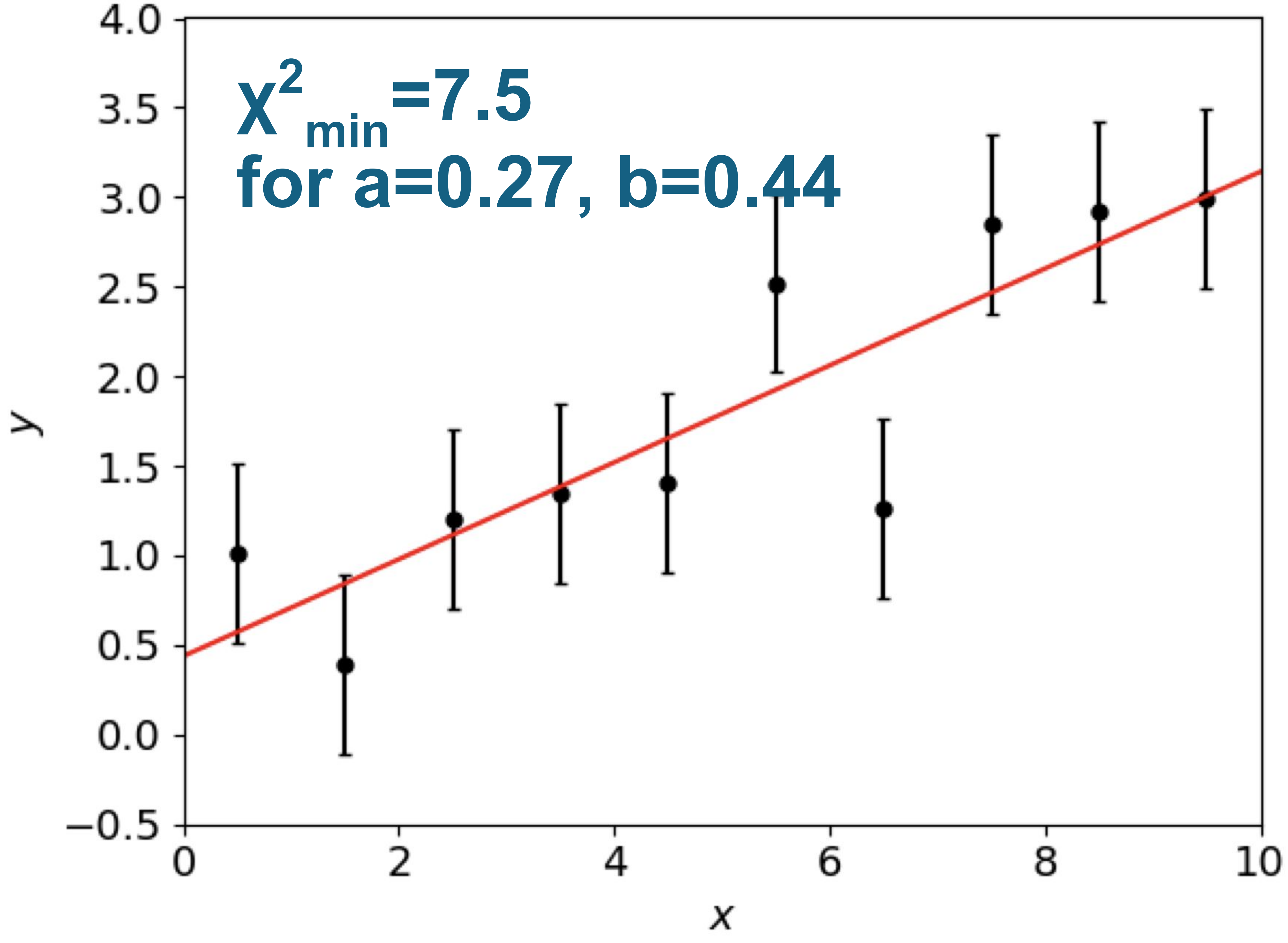




Parameter Estimation using the Likelihood & χ^2

We have a data set with 10 points. Are they well fit by a line: y=ax +b?





Estimating the Likelihood given the χ^2

- Assuming Gaussian errors, the χ^2 gives you Likelihood (L) of the model parameters

$$L = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N e^{-\frac{1}{2}\boxed{\sum_{i=1}^N \frac{\left(x_i - \mu\right)^2}{\sigma_i^2}}}$$

$= \chi^2$

$$P(\chi^2) \propto e^{-\chi^2/2}$$

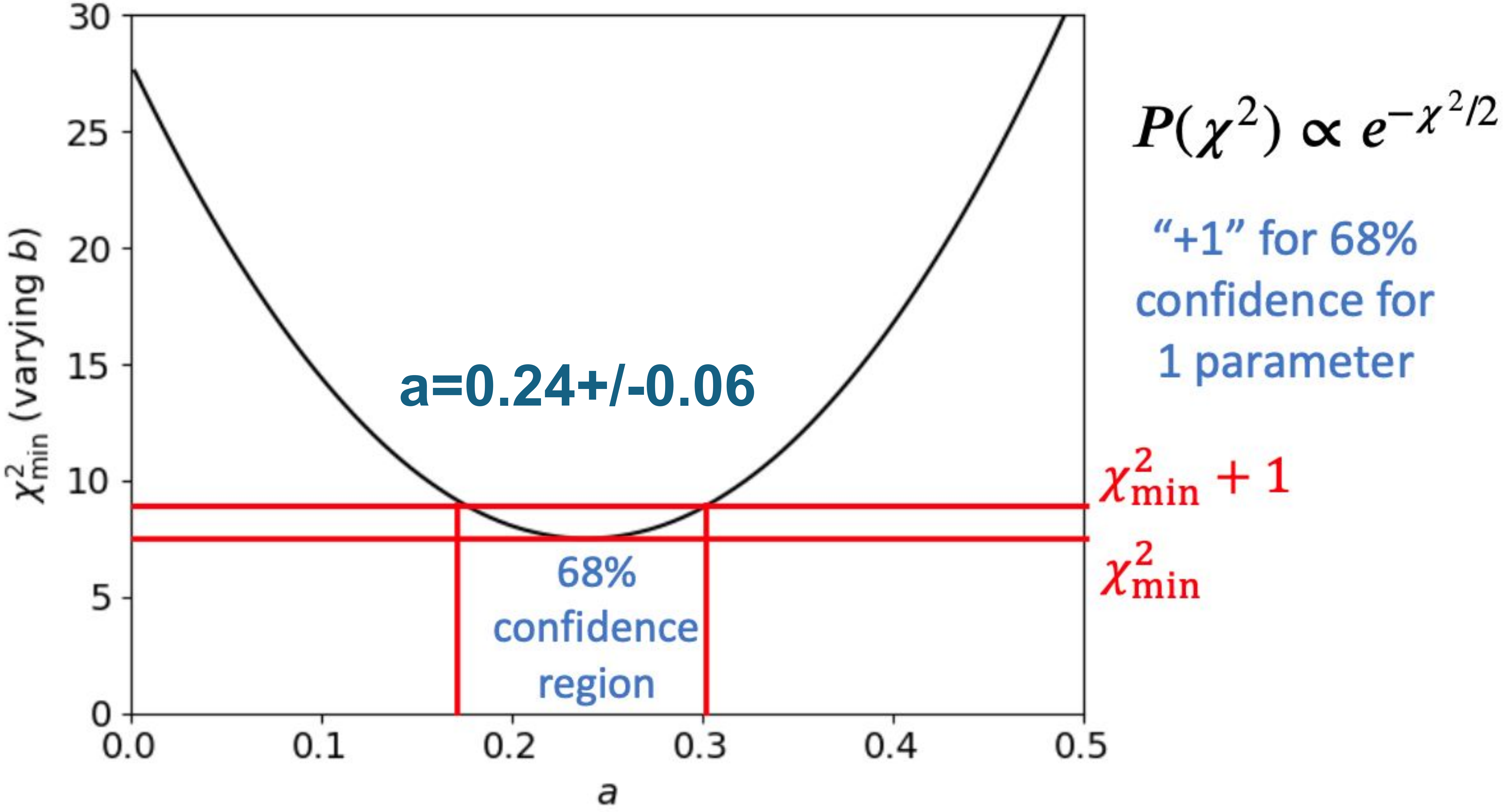
The **boxed** quantity is the chi-squared (χ^2)

The “best-fit” (or most-likely) model is defined as when the χ^2 is minimized.

Change in chi2 ($\Delta\chi^2$) tells you relatively likelihood or probability [$P(\chi^2)$]

Using χ^2 for parameter fitting

Recalculating χ^2 as you vary your model parameter(s), will give you the best-fit value and confidence range for your model parameter(s):



Lab-2: Measuring the Hubble Constant using "Standard Rulers"

1 Introduction

In 1929, Edwin Hubble measured that distant galaxies were systematically redshifted relative to galaxies that were closer (see Figure 1). From this data, Hubble inferred that the universe was expanding, an idea initially worked out by Georges Lemaitre using EinsteinÕs theory of gravity. In this lab, you will conduct a measurement similar to Hubble's and will produce your own version of his famous Hubble diagram shown below.

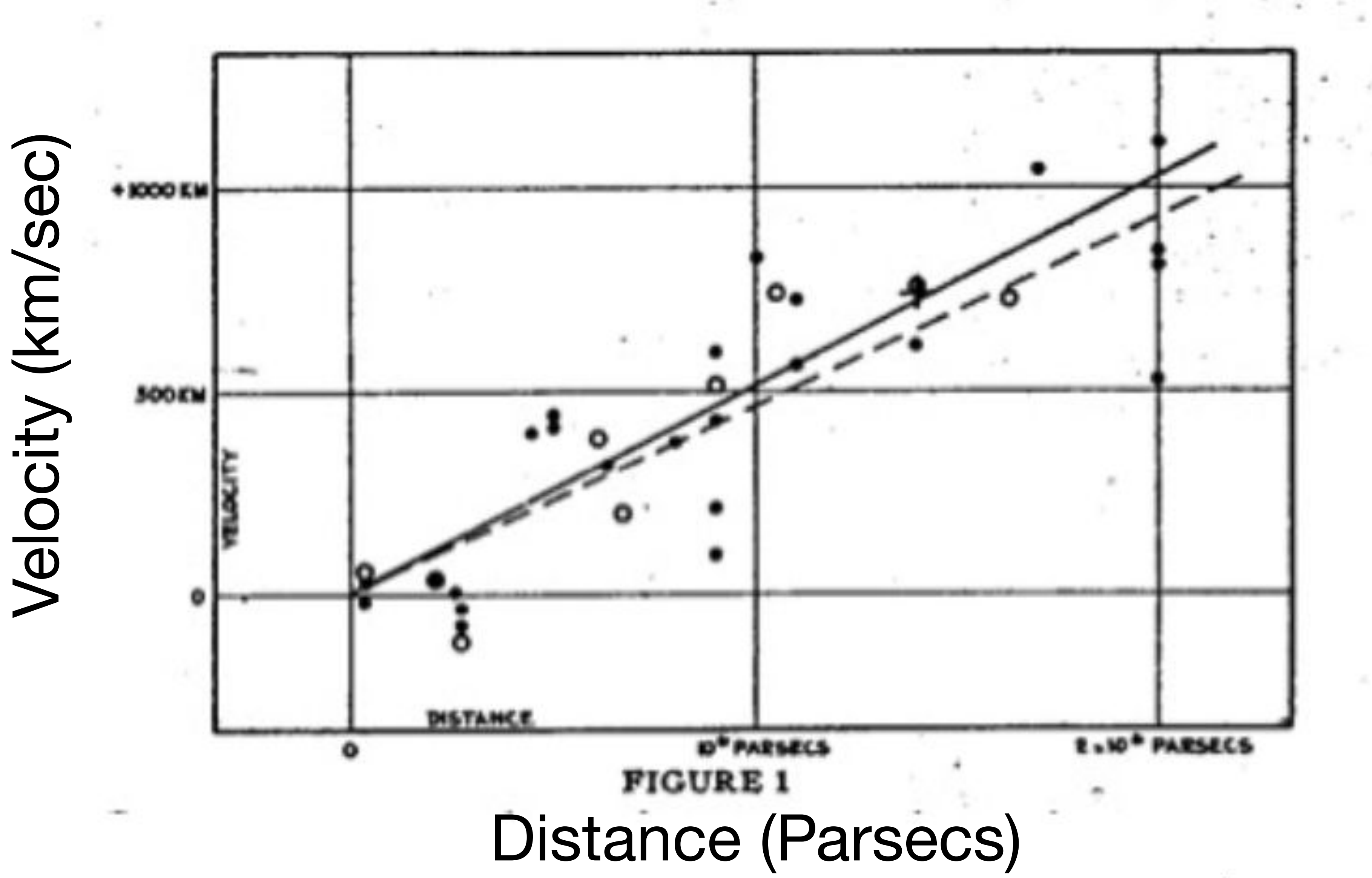


Figure 1: Velocity-Distance Relation among Extra-Galactic Nebulae (Hubble et al. 1929).

- The Universe is expanding; galaxies that are further away, also appear to be moving away from us at larger velocity.
- One way we measure this expansion is through "**Standard Rulers**" or "**Standard Candles**", i.e.,
  - Observe objects of known size (or luminosity), by measuring its apparent size (or brightness), we can infer its **distance**.
  - By also measuring its **redshift** (or recessional **velocity**), we can then measure the expansion of the Universe (or Hubble constant).
  - Typically Hubble constant measured in units of [km / s / Mpc]

Lab-2: Method

2.1 Measuring Distances

We will use geometry to measure the distance of our galaxies, by assuming that they can be used as *standard rulers*, objects with the same physical size. Galaxies that are closer will look bigger and subtend a larger angle. Whereas galaxies that are further away will look smaller and subtend a smaller angle. This relationship between the angular size of the galaxy and its distance is illustrated in Figure 3.

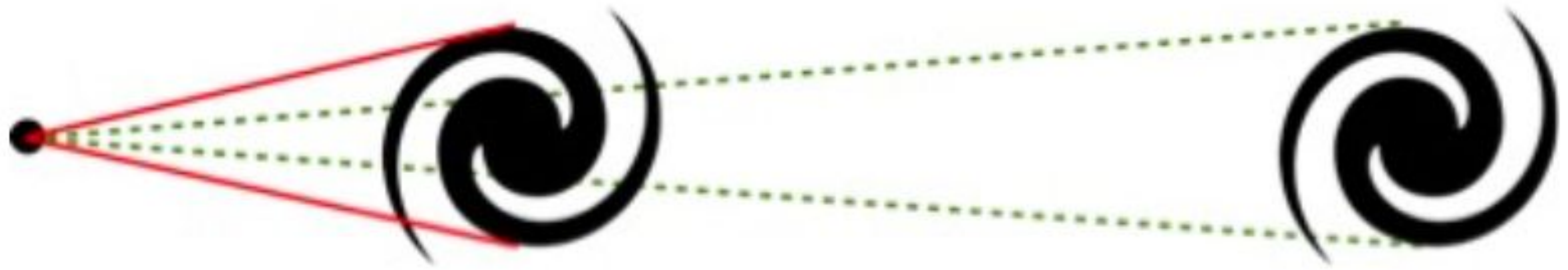


Figure 3:

Typical galaxies are about 22 kpc across. Using the geometry shown in the Figure 3, we can arrive at the following relationship:

$$Angular\,Size = \frac{22\,\mathrm{kpc}}{Distance} \tag{1}$$

So, by measuring the angular size of our galaxy images, we can use the above equation to determine the distance to the galaxy.

- Assume “galaxies” are standard rulers of fixed size.
- By assuming galaxies are a **fixed physical size** (22 kpc) and measuring their **apparent angular size**, we can infer the galaxy’s distance.
- By also looking up the galaxy’s “**redshift**” **(z)**, we can infer an expansion rate [velocity/distance] from each galaxy.

Lab-2: Observations

Observe at least 6 galaxies, between:

$0.002 <$ redshift, $z < 0.15$, and

$15 <$ magnitude, $M < 6$

3.3 Instructions for Scheduling Observations ASTR 21200: Hubble Diagram

Figure 5: Two galaxies that you might observe for this lab. (Left) NGC 1357 is an isolated spiral galaxy situated in the constellation of Eridanus. Optical image taken with the 32-inch Schulman Telescope, image source: http://www.caelumobservatory.com/gallery/n1357.shtml. (Right) NGC 4486, also known as M87, is a nearby elliptical galaxy in the constellation Virgo. It is known for having a large population (~10,000) of globular clusters, about 100 times more than the Milky Way galaxy, and made even more famous through observations of its supermassive black hole by the Event Horizon Telescope (EHT). For scale, the above image is 97 arcminutes across. Image source: http://www.eso.org/public/images/eso1525a/

Galaxy Name	z	R.A. (J2000)	Decl. (J2000)	Notes
NGC 2683	0.00137	08:52:41.3	33:25:19	
NGC 2775				
NGC 2903				
NGC 3034				
NGC 3147				
NGC 3184	0.001975	10:18:17.0	41:25:28	
NGC 3227				
NGC 3368				
NGC 3516				
NGC 3627				
NGC 3941				
NGC 4486				
NGC 4565	0.004103	12:36:20.8	25:59:16	
NGC 4631				
NGC 4775				
NGC 5248				
NGC 5548				
NGC 5907	0.002225	15:15:53.8	56:19:44	
NGC 6181				
NGC 6217				
NGC 6643				
NGC 6764				
NGC 7331	0.00272	22:37:04.1	34:24:56	

Table 1: A list of potential galaxies that should be reasonable to observe for this lab, with space given for the redshift, Right Ascension (R.A.), and Declination (Decl.) for each object. You should choose a mix of spiral and elliptical galaxies.

Lab-2: Analysis & Conclusions

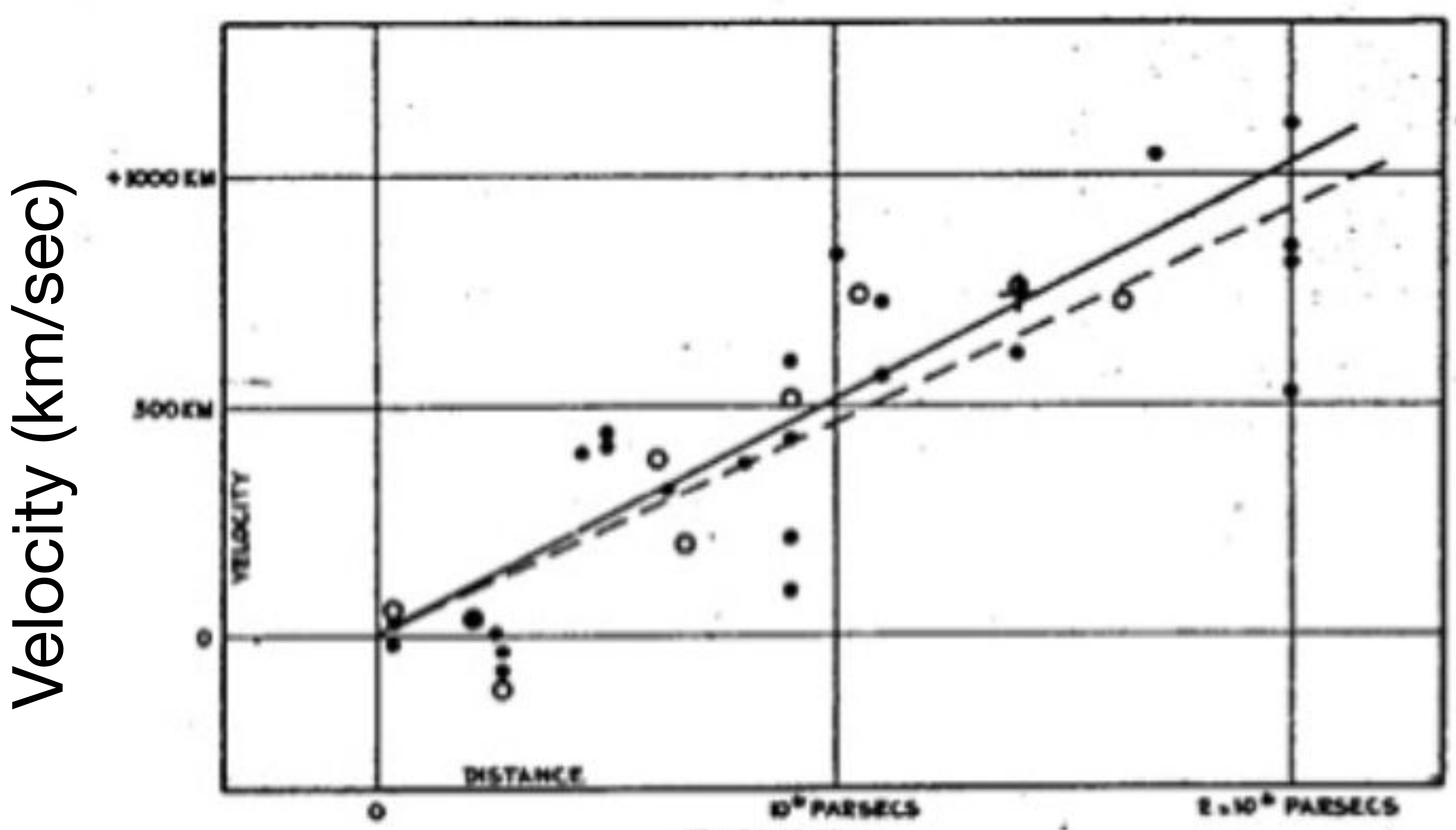


FIGURE 1

Distance (Parsecs)

Figure 1: Velocity-Distance Relation among Extra-Galactic Nebulae (Hubble et al. 1929).

- Fit a Hubble Constant to your measurements of velocity/redshift vs distance.
- See sections 4.2, 4.3 of the Lab Instructions;
  - make sure to answer the questions in your submitted Lab Report.
- Write up your results in Jupyter Notebook:
  - Include sections for: 1) Intro, 2) Data & Observations, 3) Data Analysis, and 4) Conclusions
- **Note: You will want to start observing galaxies this week**, to give plenty of time for the data analysis part of this report

Extras

Note: Generalizable to multidimensional models $\mu(x|a,b,c)$ and correlated noise (C_{ij}) between data points (x_i, x_j)

- The covariance matrix, Cij, tells you the level of correlated noise between data points x_i and x_j
- In the presence of data correlations:

$$\chi^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{(x_i - \mu)^2}{\sigma_{ij}} = (x-\mu)^T \left[C_{ij}^{-1}\right](x-\mu)$$

- And the gaussian likelihood is (still) given by:

$$L = Ne^{-\chi^2/2}$$

Where N is a normalization constant, such that the integral of the Likelihood = 1

Correlated noise example:
https://astrobites.org/2014/07/01/beyond-chi-squared-an-introduction-to-correlated-noise/