Introduction to Junior Lab

Junior lab web page:

web.mit.edu/8.14

Who am I
(and what do I do at MIT?)

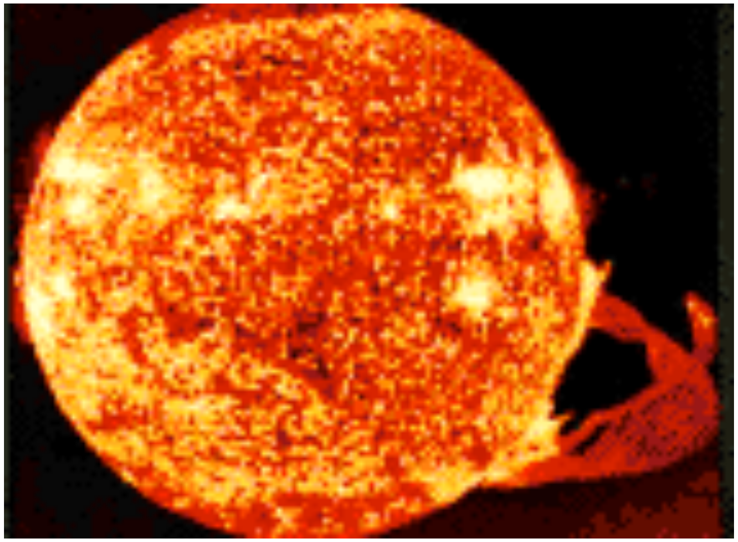
- Prof. Gunther Roland
- Office: 24-504
- Phone: x3-9735
- AIM: PhobosRolandG
- E-mail: rolandg@mit.edu
- When I'm not teaching:
  - Basic research in High-Energy Nuclear and Particle Physics

- Find these slides on JLab page after class

When I’m not in my office,
I’m usually here

What do we do at LHC ?

- Make the ***Hottest Matter in the Universe***
  - 100.000.000 times hotter than the surface of the sun



$\times 10^8$

The Junior lab team

- Prof. Ulrich Becker Rm 44-123B x3-5822 becker@lns.mit.edu
  - JLab lead instructor
- Dr. Scott Sewell Rm 4-432 x3-5082 sewell@mit.edu
  - JLab expert
- Dr. Sean Robinson Rm 6c-203 x3-9702 spatrick@mit.edu
  - JLab expert
- Chester Chu cpchu@mit.edu
  - Teaching assistant

- Prof. Gunther Roland Rm 24-504 x3-9735 rolandg@mit.edu
  - Section leader

For more info, see web.mit/edu/8.14

Getting started

- Find a good partner
  - 18 units: 6 in lab, 12 outside lab
    - this is not an overestimate
  - make sure you + your partner can coordinate schedule and work together efficiently
  - need to form partnership by next Monday
    - 1st experiment starts next Wednesday

Section organization

- Max 16 students/section
- This section is not too crowded
  - amount of help you can get is ~ 1/enrollment
  - If you consider switching:
    - Contact Dr. Sewell if you want to switch
    - Sections need to be fixed before 1st experiment

Introduction

- See first pages of 8.14 reader
  - pick up from copytech

Experiments

- Major advances in science (e.g. Nobel prizes)
- Refine your skills in the art + science of experimental physics
  - How to obtain good data
  - How to document your work
  - <u>How to estimate errors</u>
  - How to present your results
- As close to real life as possible

Experiments

- 26 sessions total
  - attendance is required
- You will do
  - 4 out of 8 exp's (5 sessions each)
    - prep questions
    - oral presentation + paper for each
    - 2x30min per partnership for oral
  - Last 4 weeks: Possible "challenge" work on selected experiment

Experiments

- If needed (emergencies etc)
  - extra time on Fri
  - signup sheet (Scott Sewell)
  - never work alone

Ethics in Science

- Fabrication/Falsification of data
  - document everything as you go (Notebook)
  - complete record of everything you have done, including mistakes
- Plagiarism
  - never use other work without acknowledgement
  - mark quotes as quotes
  - do not import text (from web resources)
  - Comparison to known values is ok, but not substitution/modification of your data, unless clearly marked
- No tolerance in JLab

Reading Material

- 8.14 reader (Copytech)
- Bevington: Data reduction and error analysis
- Other books (e.g. Melissinos): Reading room
- Original papers: See e-library on 8.14 webpage
  - Greater reliance on original literature than for 8.13

Safety

- Electrical safety
  - be careful
  - never work alone
- Laser safety
  - Wear goggles
  - Use common sense
- Cryo Safety
- Radiation Safety

Grading Scheme

- 10% attendance/lab performance
  - change 'lead' from exp to exp
- 8% Notebooks
  - graded by Scott Sewell (2x in semester)
- 10% prep problems
  - come prepared, you will need the time
- 40% orals
  - 20' are short
  - split topic between partners (but not along theory/experiment)
- 32% papers
  - < 4 pages, due morning after oral
  - both partners have to write their own paper

Homework

- Find partner
- Look at web.mit.edu/8.14
- Decide on experimental line
- Matlab homework: Hypothesis testing using fits

Introduction to Data Analysis

c.f. Bevington

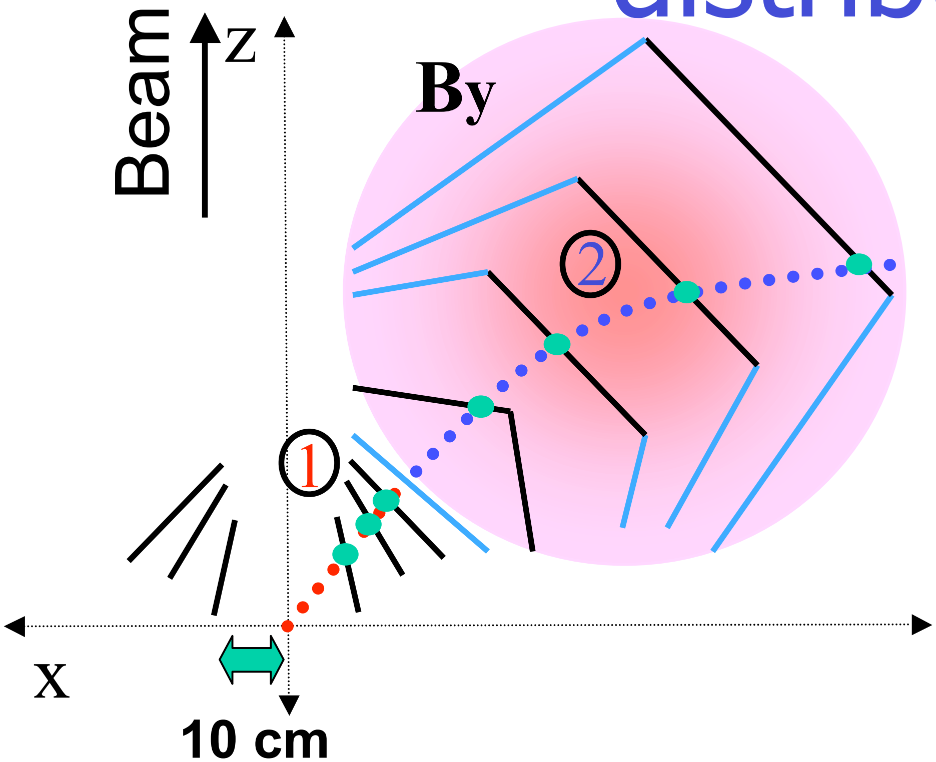
Chapters 1-3

Feb 6 2008

Data Reduction

- Translate measured data into one or more physical variables of interest
- Obtain
  - best estimate for physical variable
  - estimate for precision and accuracy of measurement (systematic and statistical uncertainty)
- Example (from my experiment):

Example: Determination of particle momentum distribution



$R = p_{zx}/(q*B_y)$

Histograms

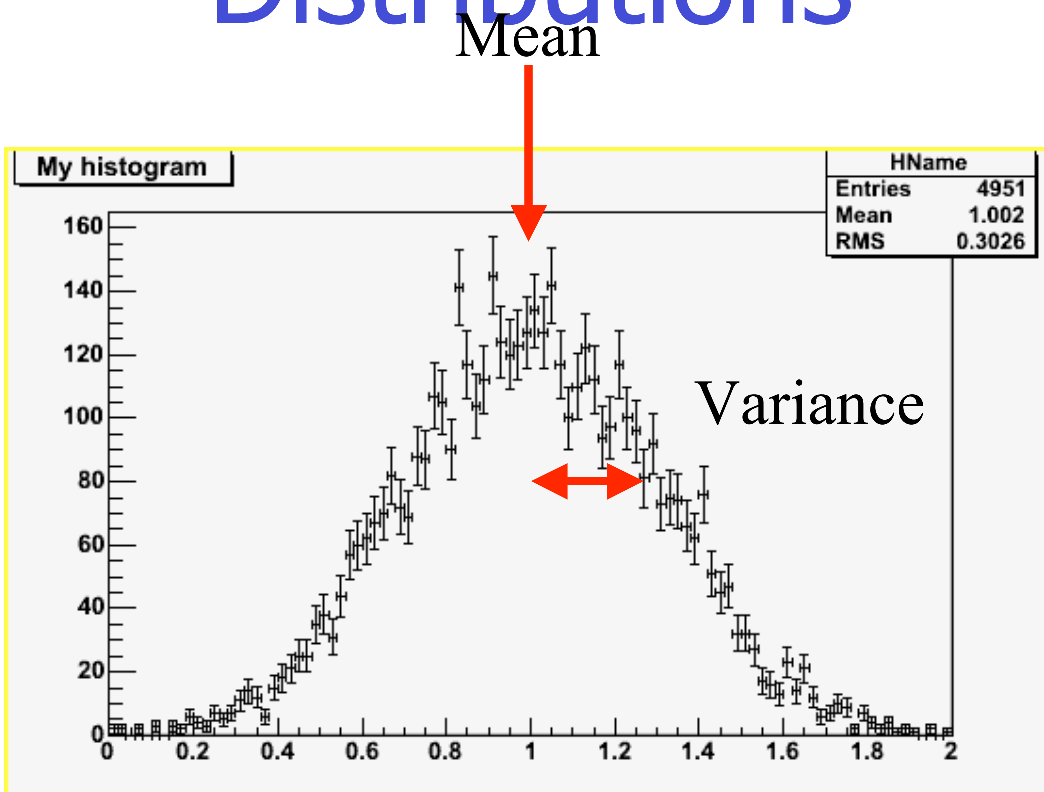
Binned representation of data in 1, 2 or 3 dimensional variable space

Statistical and Systematic Error

- Systematic error
  - inherent to measurement, apparatus, methods
  - estimate magnitude by comparing different approaches
  - limits accuracy
- Statistical error
  - Measurements jitter around truth
  - Average many measurements to improve estimate (if they are independent)
  - limits precision, but lim(N-> inf) = truth

Distributions

Sample Distribution



$$\text{Mean} \quad \langle x \rangle = \frac{1}{N}\sum_{i=1}^{N} x_i$$

$$\text{Variance } s_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \langle x \rangle)^2$$

for N -> inf :
Parent distribution

Binominal Distribution

$$P(x:n,p) = \frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}$$

Mean $\mu = n \cdot p$

Std. Dev $\sigma = \sqrt{n \cdot p(1-p)}$

P(x:n,p): Probability to get ‘yes’ *x* times out of *n* tries, if probability of ‘yes’ for single try is *p*

Poisson Distribution

$$P(x, \mu) = \frac{\mu^x}{x!} e^{-\mu}$$

Mean $\mu = \mu$

Std. Dev $\sigma = \sqrt{\mu}$

Derives from binomial distribution for p << 1 with n*p = m
Important for counting experiments with low count rate

Gaussian Distribution

$$P(x, \mu, \sigma) = \frac{1}{2\pi\sigma} e^{-(x-\mu)^2/(2\sigma^2)}$$

Mean $\mu \quad = \mu$

Std. Dev $\sigma \quad = \sigma = \sqrt{\mu}$

Derives from Poisson distribution for n*p >> 1

Seen everywhere b/c of Central Limit Theorem

Statistical Error on Mean

$$\langle x \rangle \pm \sigma_x^2/\sqrt{N} = \frac{1}{N}\sum_{i=1}^{N} x_i \pm \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \langle x \rangle)^2/\sqrt{N}$$

- Repeated measurements increase precision
  - but only as sqrt(N)
  - ultimate limit may come from systematic uncertainty

Error propagation

Interested in error on ‘x’, but measure ‘u’, with x = f(u)

$$\sigma_x^2 = \sigma_u^2 (\frac{\partial x}{\partial u})^2$$

What if x = f(u,v):

$$\sigma_x^2 = \sigma_u^2 (\frac{\partial x}{\partial u})^2 + \sigma_v^2 (\frac{\partial x}{\partial v})^2$$

Errors add in quadrature, provided error in u and v is independent

Fitting of Data

c.f. Bevington

Chapters 4-6

Feb 6 2008

Fitting

- Fitting: Find a functional form that describes data within errors
- Why fit data?
  - extract physical parameters from data
  - test validity of data or model
  - interpolate/extrapolate data

How to fit data?

- Vary parameters of function until you find a global maximum of goodness-of-fit criterion
- Goodness-of-fit
  - chi-square ← most common
  - likelihood
  - (Kolmogorov-Smirnov test)

Chi-square fit

Data points

Model expectation

Chi-square

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

Estimated uncertainty

Best Fit: Global minimum of Chi-square

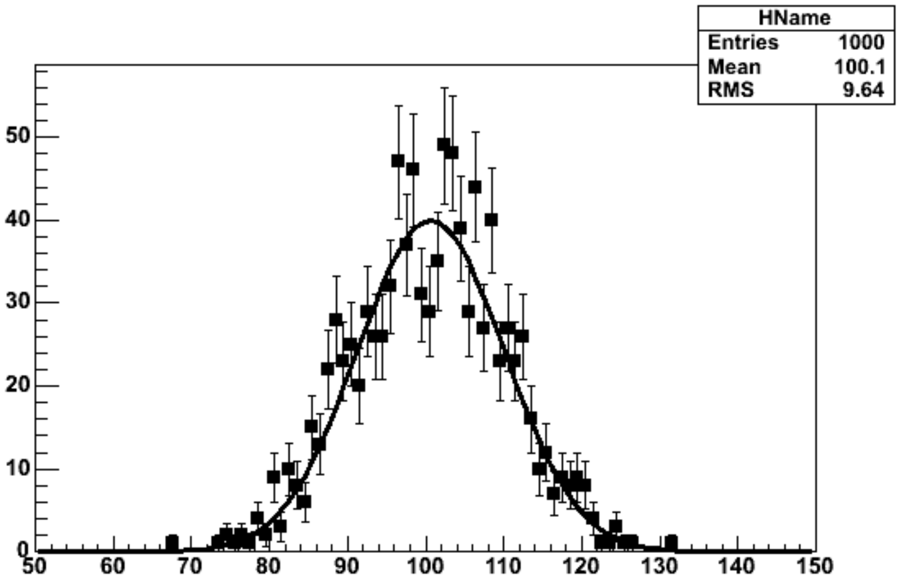
Goodness of fit

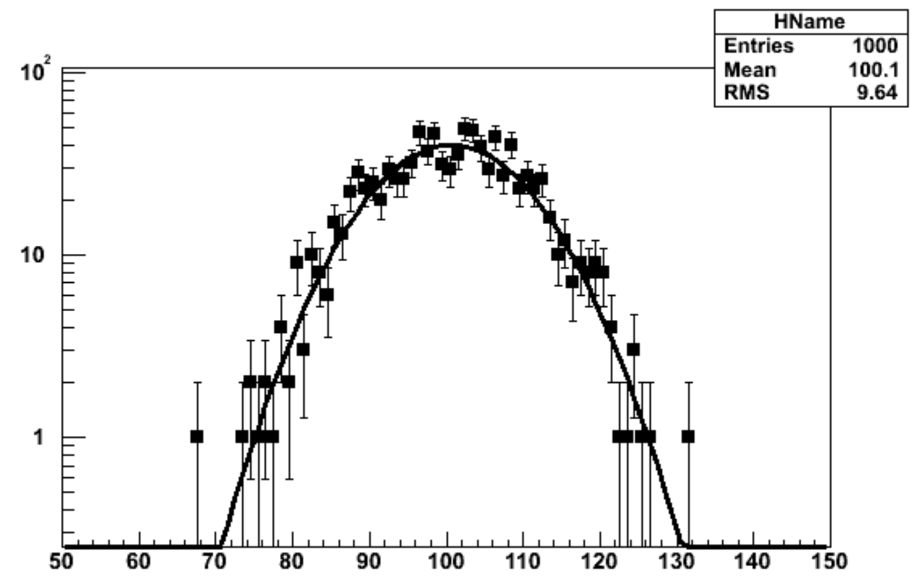
- You EXPECT data to fluctuate by ~ error
- For Gaussian errors, only 68% of your data points should agree with model/fitted function by better than 1s
- Chi-square per Degree of Freedom (DoF) should be ~ 1
  - DoF: Number of data points - number of fitted parameters
- Chi-square allows you to test if model/fitted function is compatible with data

Example

- Suppose you have 10^4 atoms
- Probability of decay is 10^{-2}/sec
- Count #of decays/second
- Repeat experiment many times (always starting with 10^4 atoms)
- What distribution do you expect for the number of observed counts?

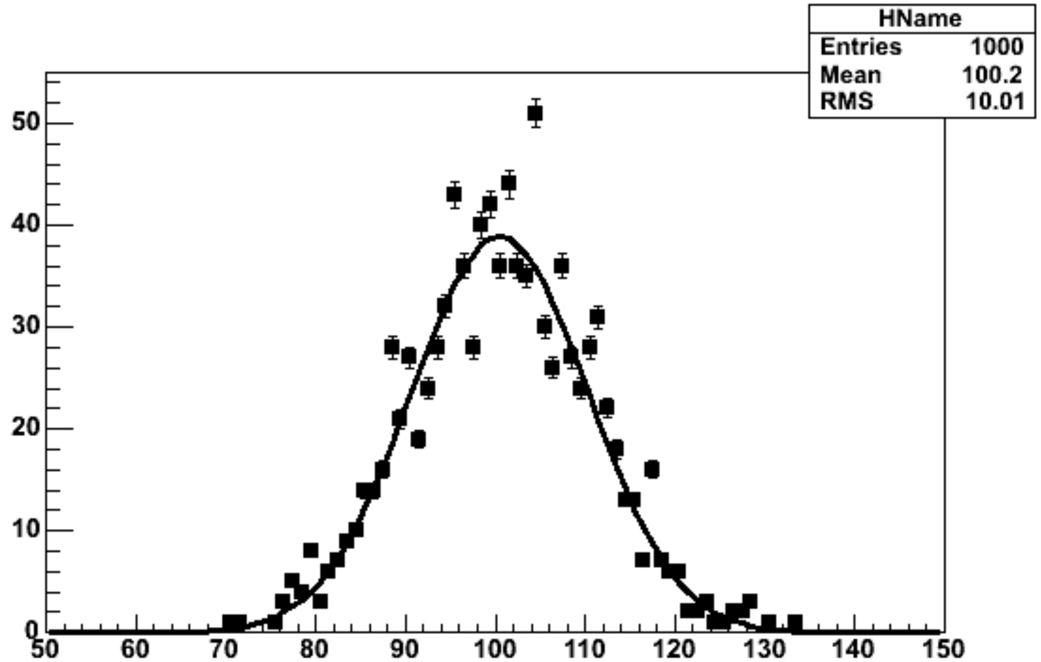
Example





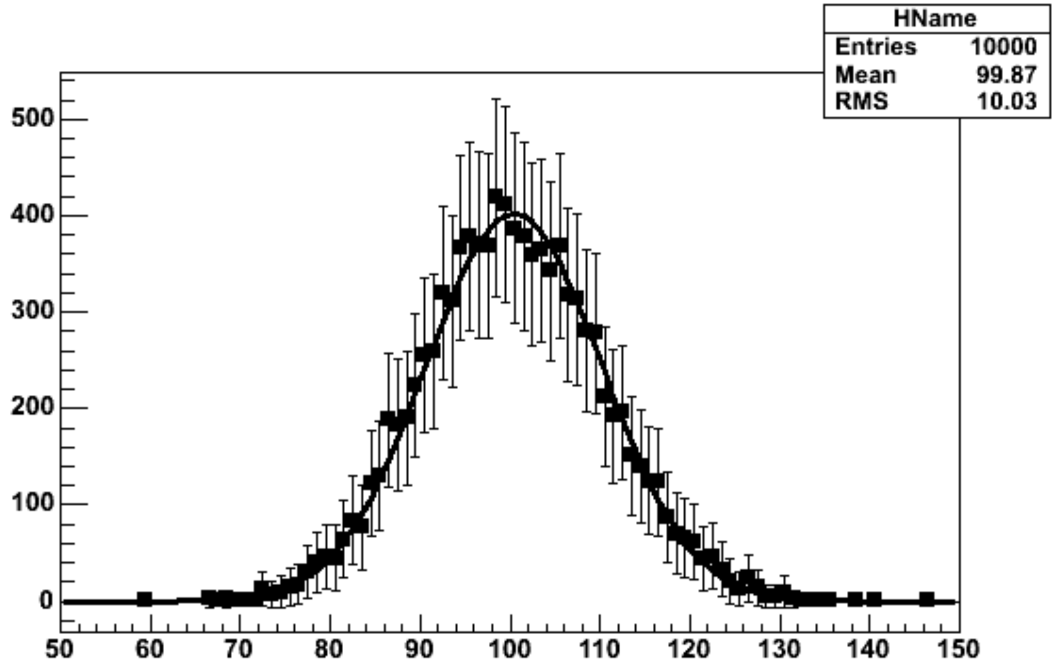
Chi-square/DoF = 60/53

Example



Chi-square/DoF = 1200/55

Example



Chi-square/DoF = 4.5/71

Chi-square distribution

- Even if your error estimate is correct and the model is correct
  - Chi-square will fluctuate from chi2/DoF < 1 to chi2/Dof >1
  - The shape of the chi2-distribution depends only on the number of degrees of freedom
  - allows calculation of probability that data and model agree
- Chi-square probability
  - Percentage of all measurements that you expect to have a worse chi-square than what you see

Fitting

- Linear relationship: Analytical
- In general: Numerical minimization of chi-square
  - Many methods
    - simple grid-search
    - Non-linear Newton
    - Levenberg-Marquardt
    - Simulated annealing
  - Criteria
    - Dimension of parameter space
    - Data statistics
  - Challenge
    - Speed vs finding true global minimum