Physics 371
Intermediate Lab I: Electrons
…(And Error)

Prof. Justin Frantz

frantz@ohio.edu

9/6/11

First: a) SIGN UP SHEET

b) Go over syllabus…

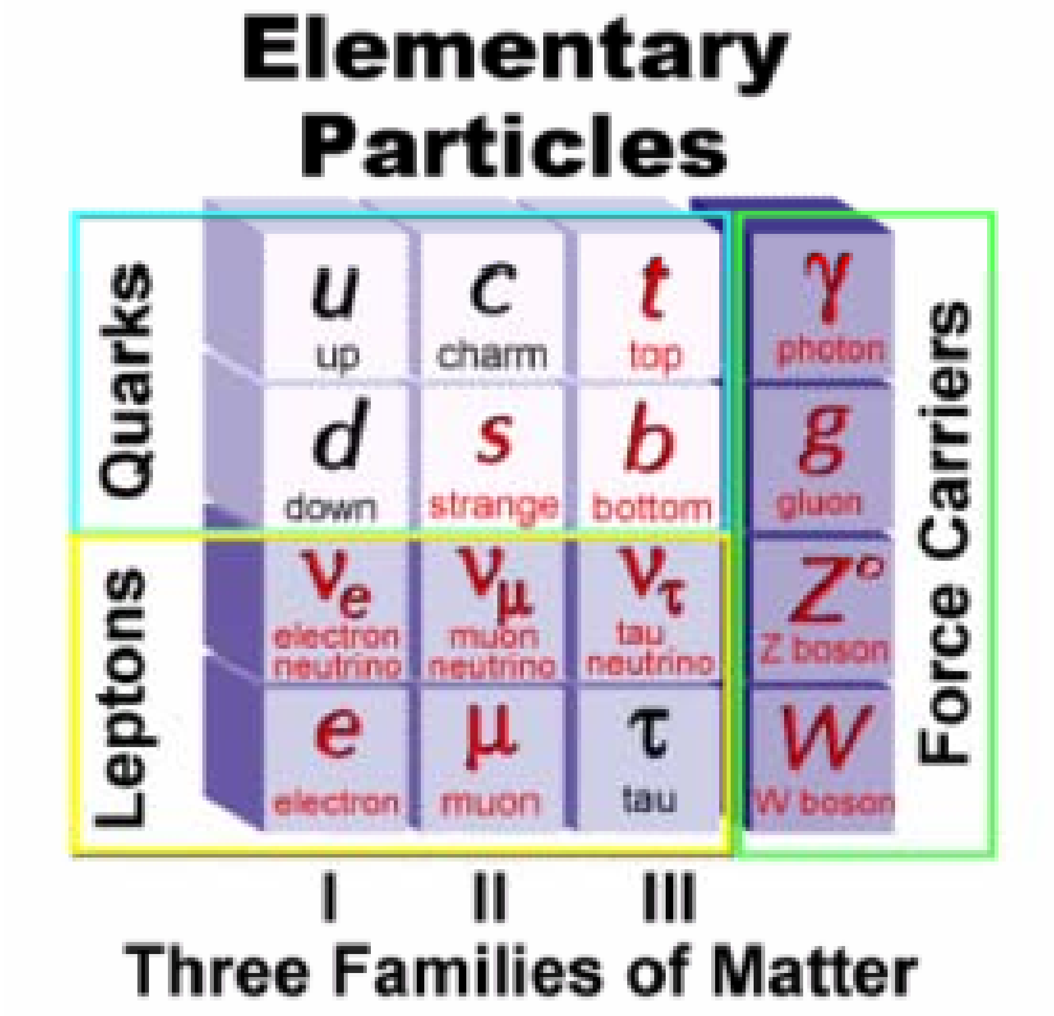
c) Bevington & Robinson Text

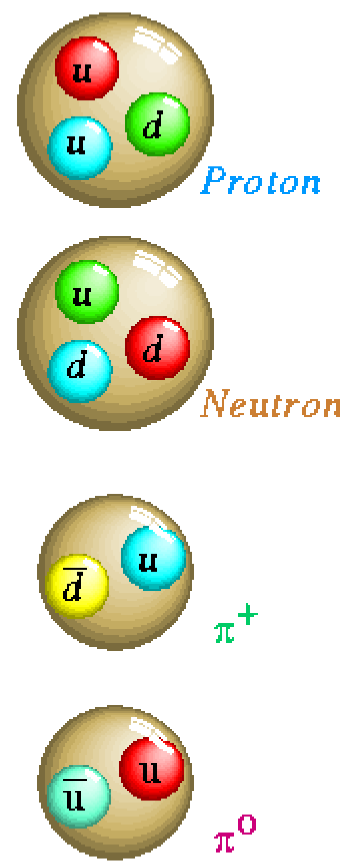
d) Go upstairs

Physics 37X Series

- 371: Electrons, 372: Photons, 373: Nucleons
- I am a high energy particle physicist so I may give the physics a slightly biased slant
- This (371) is first intro:
  - focuses a lot on error and uncertainty analysis
  - and report writing
- Electron: Subatomic particle
  - Elementary
  - Charge: Quantized
  - Mixture of classic experiments
- Photon: In almost every situation we will use photons to make our observations
  - 372: Modern condensed matter (e.g. Scanning Electron Microscopes)
- What is a Nucleon? N = n, p (neutron, protons)
  - 373: Nuclear and Particle physics
  - Counting experiments

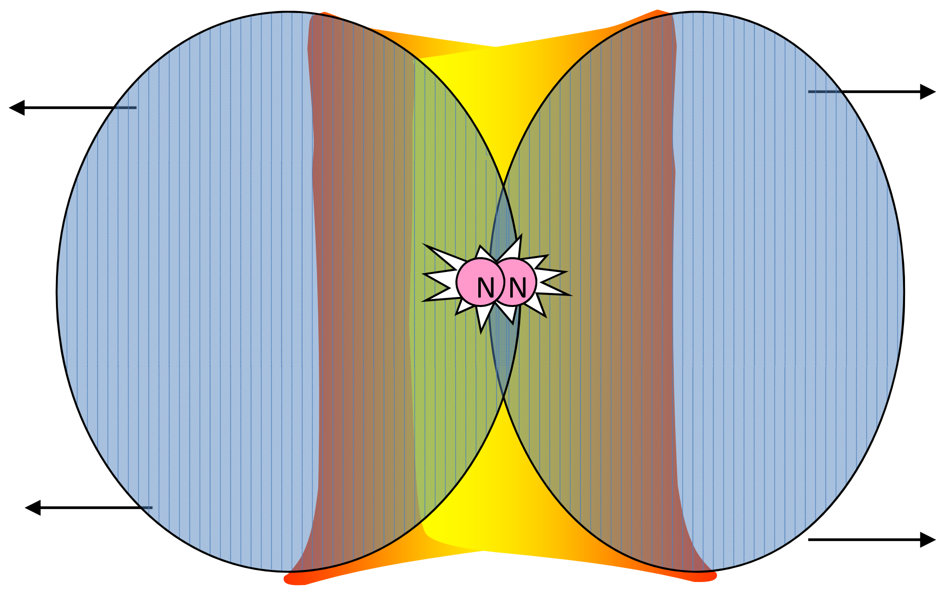
Elementary Particles



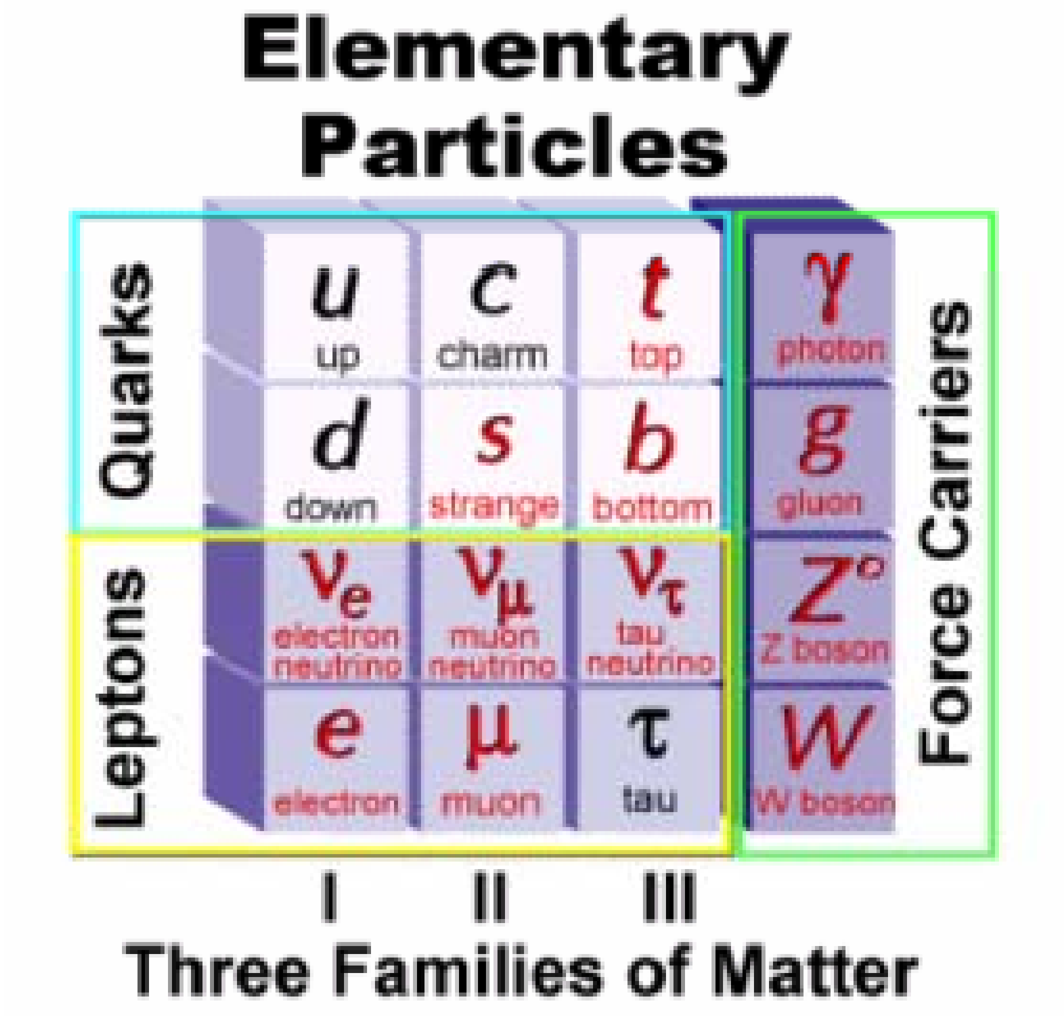


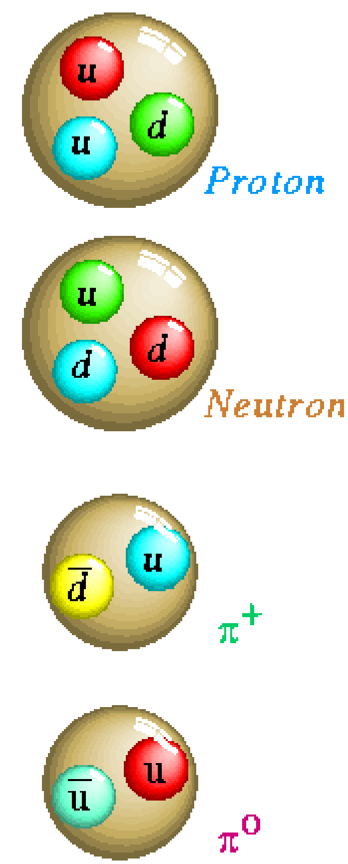
What I do: Relativistic Heavy Ion Physics

- Combination of Particle and Nuclear
- Nucleus + Nucleus Collisions at 200 GeV: RHIC



Elementary Particles



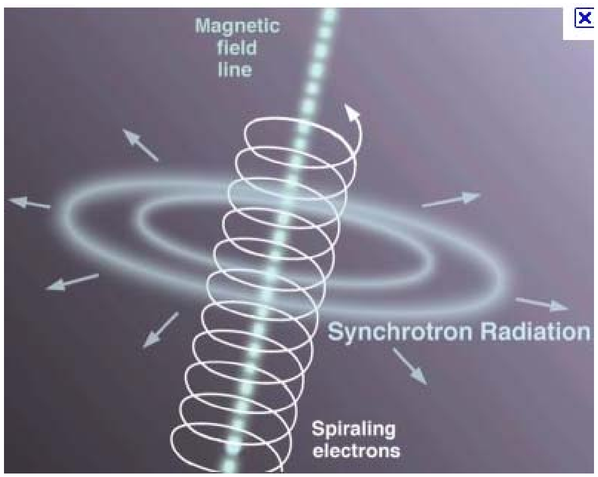


Electrons

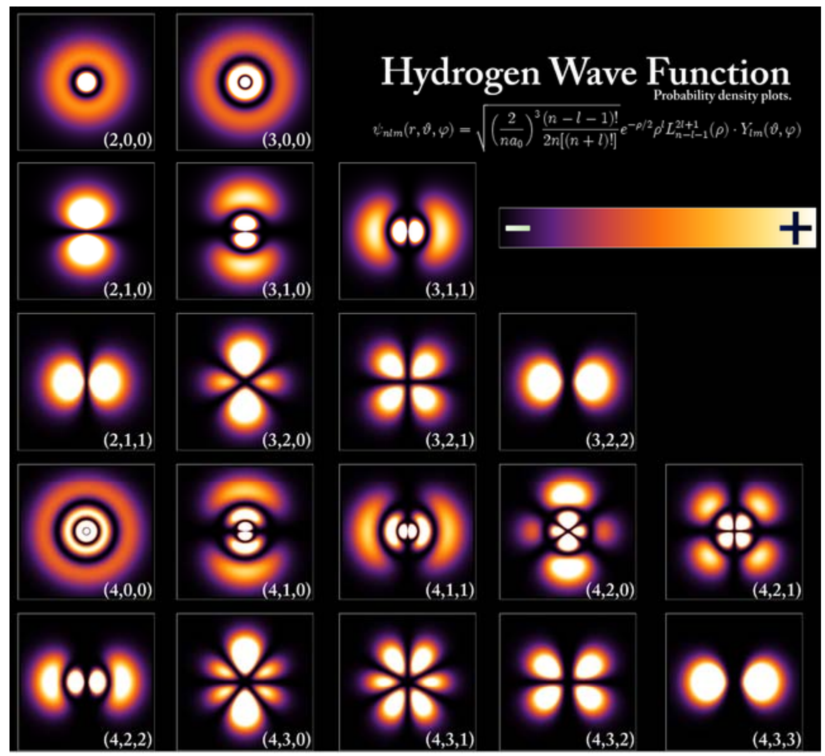
- Mass: 511 KeV **$9.10938291 \times 10^{-31}$ kg**
- 1 eV = 10^{-19} J $E = mc^2$
- Atomic interactions (e_{cloud}-nucleus/lattice)
  - 1 eV to 1 keV (γ = "Photon" or "X-ray")
  - Conduction Band (10's of eV)
- Nuclear Structure interactions (e.g. n,p shell transitions)
  - 1 keV ~ 10 MeV (γ = "Gamma Ray")
- Nucl**eon** Structure Interactions
  - $m_{nucleon}$ = 938 MeV (**$1.67262158 \times 10^{-27}$** kg)
  - m_{pion} = 135 MeV
- "Space (Vacuum) Structure" Effects
  - Higg's Boson Mass: ~100's GeV (top quark mass 173 GeV)

Electrons

- Particle Nature: e/m, movement in fields
- Quantum Effects: electron diffraction, Wave Fn



Picture from AstronomyOnline.org



- This Lab will explore both sides

Other interesting facts

- Discovered by J. J. Thomson (1897)
- Has anti-particle: called positron e^+
- called Weakly Interacting: does not feel "Strong Nuclear Force)
- Governed by Quantum Field Theory (it is a field, just like photon)
- Common component of most prolific *plasmas* (Earths/Sun mag field--)

Aurorae, Coronae

Report Writing

- Why Important?
  - Convince Other Scientists
    - You did the measurement
    - And you did it right
  - **Reproducibility**
    - In order to be considered a real result must be reproducible

Excel Tutorial

- Auto-complete
- Copy
- Definite
- Functions: Histogramming

Why is it important?

- A Measurement is meaningless without uncertainties!!!!!
  - http://deathbyvaccination.com/
  - Theoretical AOL Article about war waste $
- Typically:
  - ~10% of time/work spent making measurement
  - ~90% of time/work spent evaluating/estimating uncertainties

Physics 371
Intermediate Lab I: Electrons
…(And Error)

Prof. Justin Frantz
frantz@ohio.edu

9/8/11

First Lecture I Error

b) Tutorials

c) Go over to new location Clipp

d) Lecture on Distributions

Reading Assignement Ch 2.

Intersection of Statistics and Error

- Random (Stat) Error → Bev. Ch 1 Model:
  - Single quantity measured N times

- Measurement = Truth + Error

  "Random Variable"

- Part I of course : Consider at Distributions of measurements
  - Histograms are way to view Distributions
  - Parameters (e.g. moments) describing these distributions

Table …

Parent Distribution/Sample

- Sample Distribution: Set of (N) measurements of "one" quantity
  - (often: same measurement repeated)
- Parent: "functional form" in limit of infinite number of measurements $N \rightarrow \infty$
  - (Distribution)

- Histogram → Visualization of Sample Distribution
  - Drawing to demonstrate

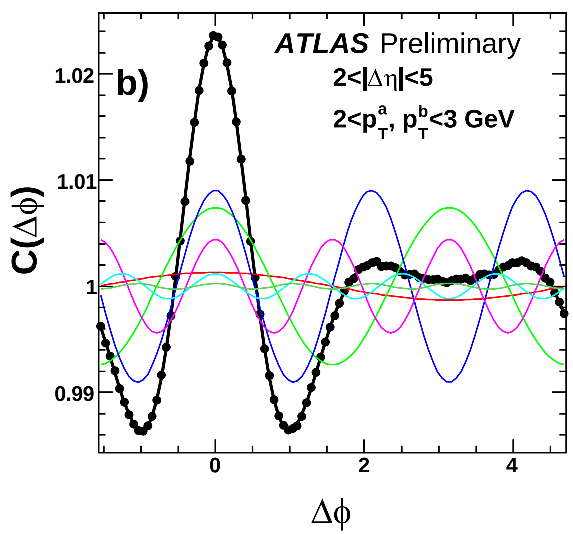
Parameters of Distribution

- You: fill in
- Mean
- Median
- Mode: Most Prob. Value
- Some Weird Examples …. Flat, prob neq mean
- Deviations
  - Standard (RMS) σ
- Variance σ^2
- See next slides: s^2

What to use these parameters for?

- To estimate your measurement: (which parameter?)
- To characterize how accurate your measurement should be (come back to weird)
- If you reproduce experiment, how accurate might you expect your experiment to be.
  - This is why we make a difference btw s^2 and σ^2

- However note that quite often these same parameters are used for other purposes

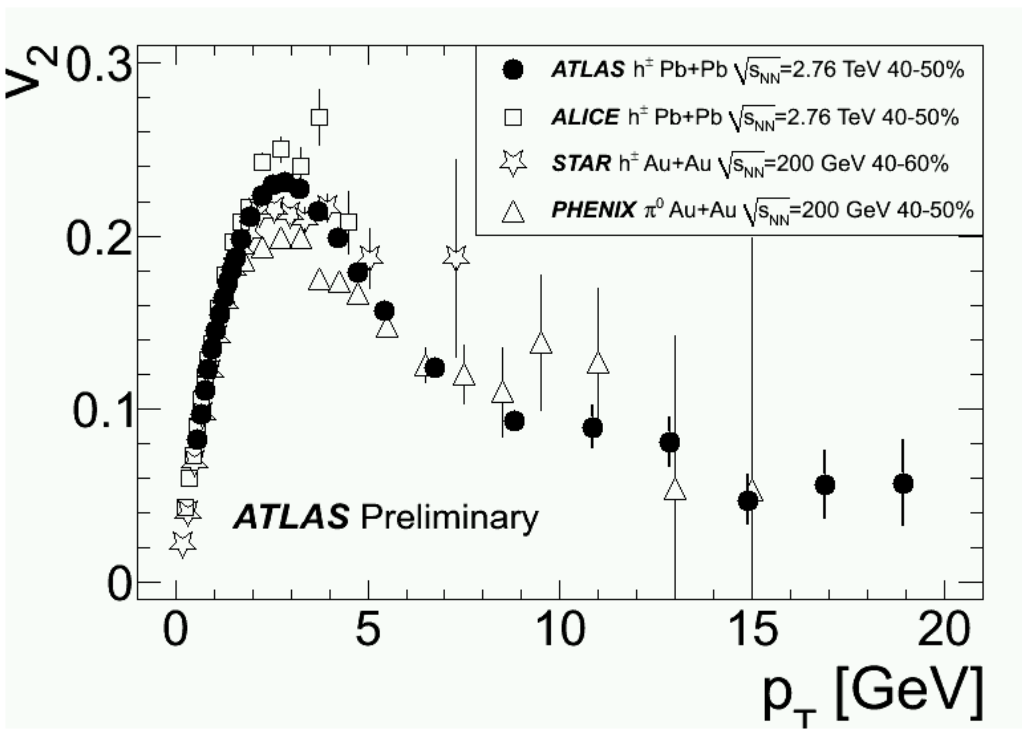


Error Tutorial

- Columbia Error Tutorials Go to website
- Perform excercises in Excel: make new file, email me at the end. Put group names in first page.
- Make Histograms in Excel of : Histogram example, Exercise 3. see link on Justin's page
- Prof Frantz: example using ROOT

Bigger picture

- Usually we make "histograms" (in particle physics/counting experiment) /graphs of measurements as a function of some changing input.
- How do these distribs/params relate to these? (Draw)
- Accuracy vs Precision



Physics 371
Intermediate Lab I: Electrons
…(And Error)

Prof. Justin Frantz
frantz@ohio.edu

9/13/11

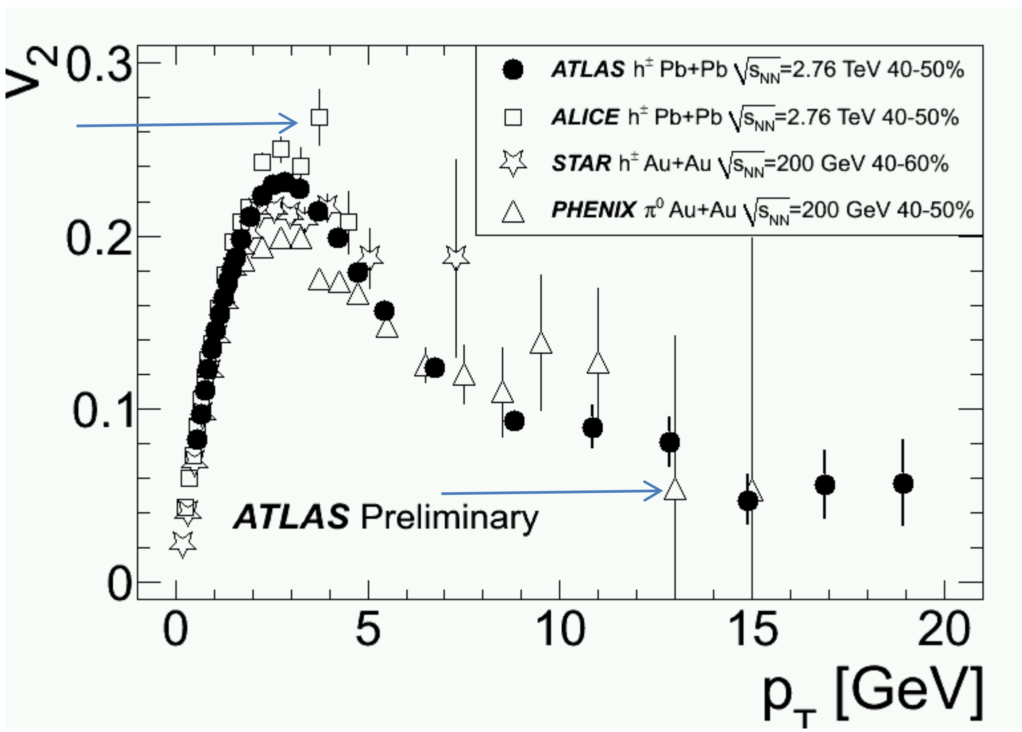
First Lecture I Error
Lecture on Distributions

Reading Assigement Ch 3.

Bigger picture

- Usually we make “histograms” (in particle physics/counting experiment) /graphs of measurements as a function of some changing input.
- How do these distribs/params relate to these? (Draw)

Accuracy vs Precision : make sure you know the difference



Parent Distribution→ Probability

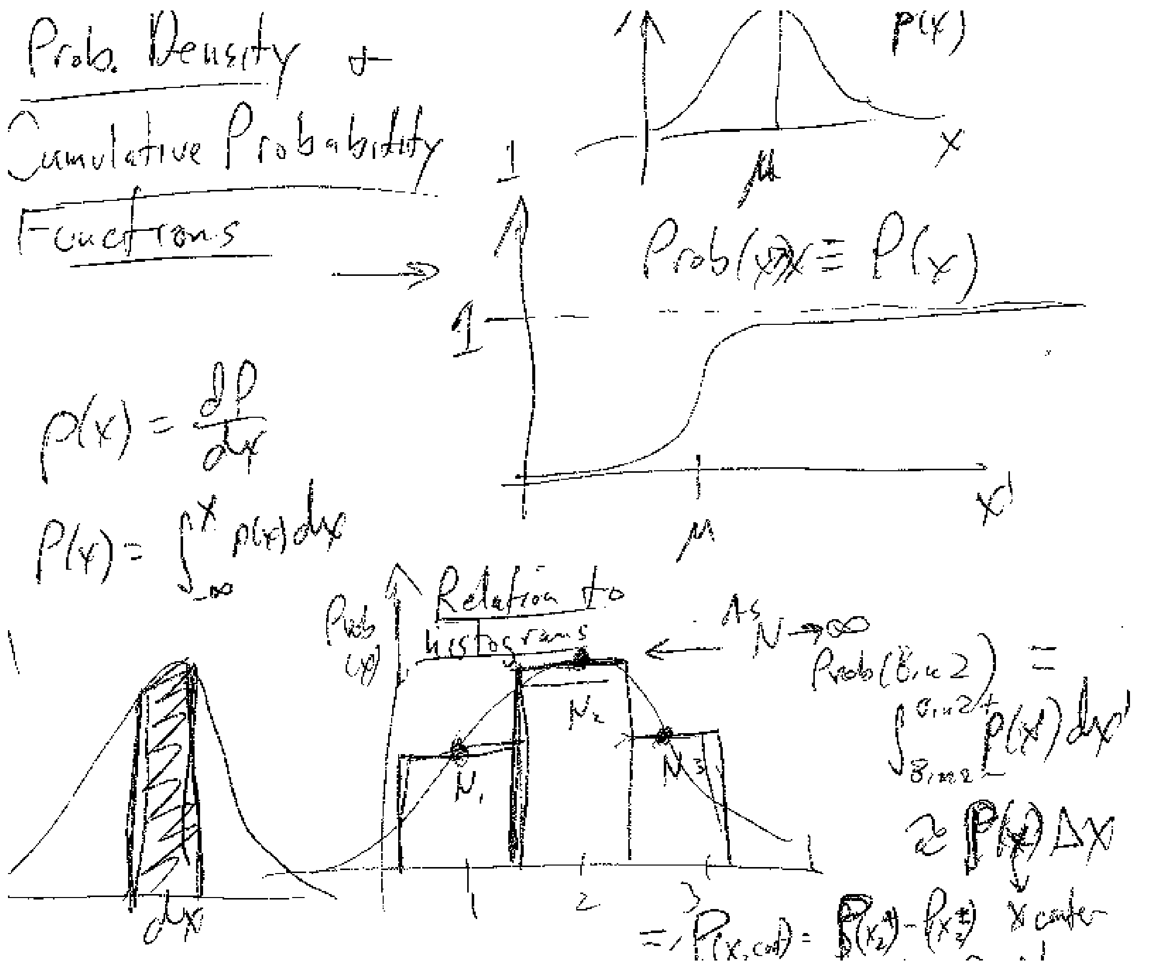
- p(x) : Probability Density Function
  - "Continuous" (limit N→ inf, so do # bins in hist)
- Cumulative Probability Function P(x)
  - Integral
- In usual cases Parent Dist == p(x)
- P(x) = 0 -→ P(x) dx
- Histogram P(x) Δx

- Relation between btw <x> formula

- Expectation value <f(x)>
  - Can estimate from any histogram

- We will discuss several common examples

Notes on board



$= N_2/N$

More notes on board

Note: clarification made in class

Drawing above is impossible in reality only drawn for demonstration purposes

Prob(x) (curve) ≤ 1 always but

$N_i \to \infty$ so drawing assumed some scale factors (unplotted) e.g. $\frac{1}{\Delta x}$

I referred to this one in class. The other is a factor of N e.g. the histogram would need divided by N to make drawing realistic.

This should clarify the possible confusion regarding the unnormalized "$p(x_2)$" $= N_2$ & the real $p(x)$

$$\langle x \rangle = \frac{1}{N}\sum_{\text{MEASUREMENTS}} x_i$$

$$= \frac{1}{N}\left(x_1 + x_2 + x_3 + x_4 + x_5 \ldots\right)$$

Regroup sum:

$$= \frac{1}{N}\Big[\underbrace{(x_1+x_2+x_3)}_{\text{"}N_1\text{" group}} + \underbrace{(x_4+x_5\ldots)}_{\text{"}N_2\text{" group}} + \ldots\Big]$$

$$= \frac{1}{N}\Big(N_1 \underbrace{\bar{x}_1^{\text{mean}}}_{\equiv \text{"}x_1\text{"}} + N_2(\bar{x}_2^{\text{mean}}) + \ldots\Big]$$

$$= P_1(x_1)x_1 + P_2(x_2)x_2 + \ldots$$

$$= \sum_{\text{BINS}} P(x_i)\, x_i$$

Equivalently Expectation Value

$$\langle f(x) \rangle = \sum P(x_i) f(x_i)$$

or for continuous distribution

$$= \int P(x) f(x)\, dx$$

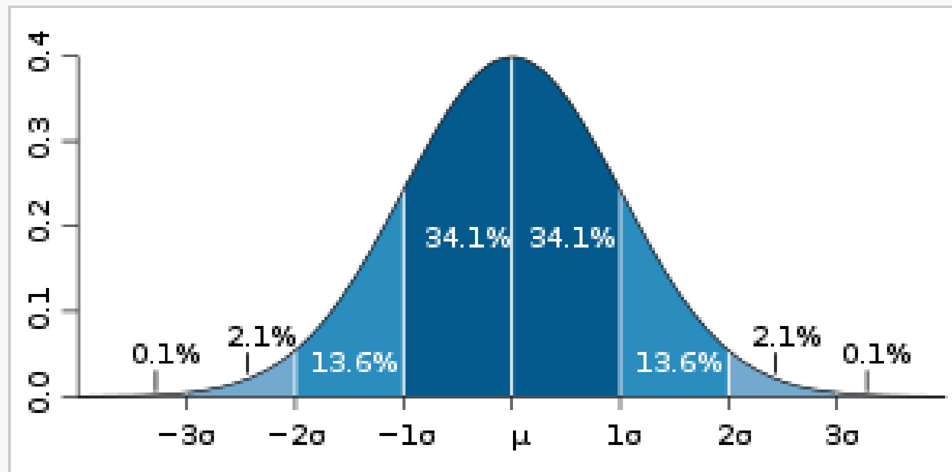
Normal or Gaussian

- #1 Normal or Gaussian Distribution
- Origin of 2/3 rule
- Origin of

"Standard Statistical Interpretation"

Many errors are assumed to be Gaus like

Example: flat



Dark blue is less than one standard deviation from the mean. For the normal distribution, this accounts for about 68% of the set, while two standard deviations from the mean (medium and dark blue) account for about 95%, and three standard deviations (light, medium, and dark blue) account for about 99.7%.

Gauss

- Functional Form

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

- Memorize
- Continuous Distributions
- Mean is mu
- Sigma is RMS
- Often used for convenience

Errors vs. Uncertainty

Slides from Tues

- Often used interchangably (by everyone)
- Error: measurement are always inherently off by some amount
  - Example 1: measuring lengths (locations of lines)
  - Example 2: determining number of cosmic rays/second
- Random Errors   Fluctuations
  - Systematic : error somehow always the “same” (usually has some random element though initially)
  - Random errors: always different each measurement.
  - Statistical Error

Counting Experiments

- Binomial and Poission
  - Two important non-Normal Error Distributions
- Integer cases: (discrete not continuous)
  - Now Histogram-like Sum becomes exact
- In limits of large N → Gaussian
- Any time we histogram something we are counting things
- Often underlying physics processes are discrete
  - Example: Energy measurements in particle detectors

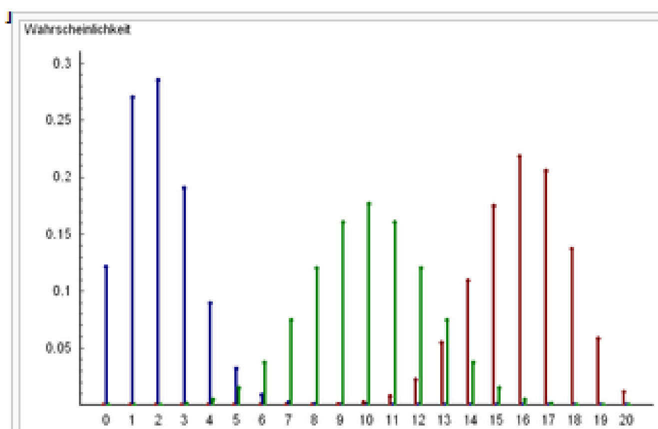
Binomial

- 2-outcome experiments: prob "success" is p
- Count # of Success
- Combinations Symbol $C(n,x) = \frac{Pm(n,x)}{x!} = \frac{n!}{x!(n-x)!} = \binom{n}{x}$

$$P_B(x;n,p) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

- q = 1-p
- Mean np
- Variance npq

$$\mu = \sum_{x=0}^{n} \left[x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \right] = np$$



Binomial distribution for $n = 20$
$p = 0.1$ (blue), $p = 0.5$ (green) and $p = 0.8$ (red)

Physics 371
Intermediate Lab I: Electrons
…(And Error)

Prof. Justin Frantz
frantz@ohio.edu
9/15/11

-Note about assignments in lab manual
Lecture : - Distributions continue
-Error propagation
Tutorials

-Reading Assignment Ch 4.
-Start looking through Lab Choices—decide if you like one better than others.

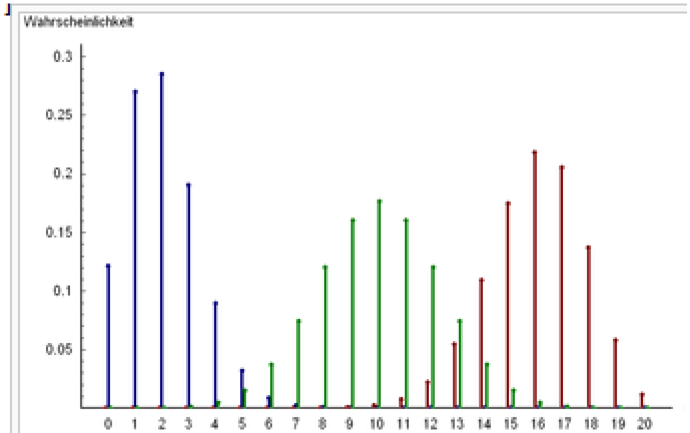
Binomial

- 2-outcome experiments: prob "success" is p
- Count # of Success
- Combinations Symbol $C(n,x) = \frac{Pm(n,x)}{x!} = \frac{n!}{x!(n-x)!} = \binom{n}{x}$

$$P_B(x; n, p) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

- q = 1-p
- Mean np
- Variance npq

$$\mu = \sum_{x=0}^{n} \left[x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \right] = np$$



Binomial distribution for $n = 20$
$p = 0.1$ (blue), $p = 0.5$ (green) and $p = 0.8$ (red)

Poisson Distribution

- Poisson distribution→ limit of Binomial Distribution when *p (q)* is small *p* → R Δt

In counting random events, the number of events occurring within a time interval is described by a Poisson distribution, dependent on the rate

(number expected: $\lambda = R\ \Delta t$).

The variance of a Poisson distribution is equal to the mean

For a given measurement of N events in a given time interval, the standard deviation is $\sqrt{N}$

1 measurement of N : “mean”
N sigma = sqrt(N) !

Count things !

$$P(n) = \frac{\lambda^n e^{-\lambda}}{n!} \qquad P_P(x;\mu) \equiv \frac{\mu^x}{x!} e^{-\mu}$$

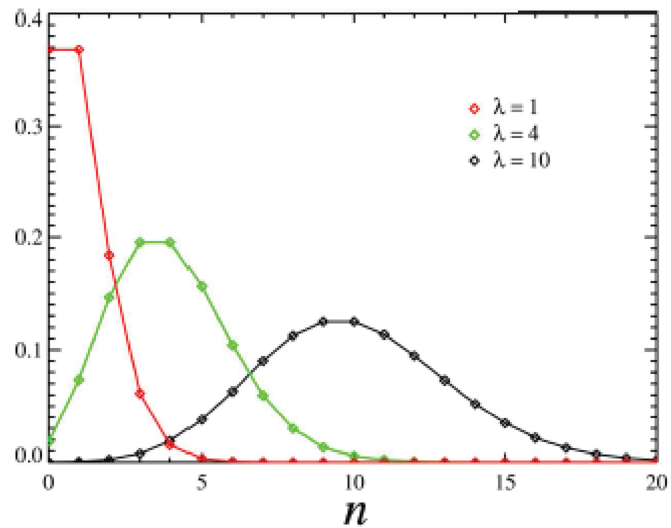


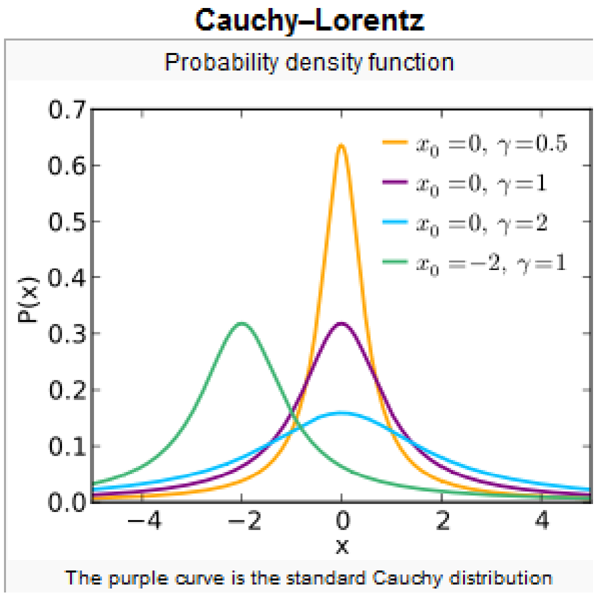
Figure from “Poisson distribution” Wikipedia entry

(Measurement of N = N independent 1 measurements)

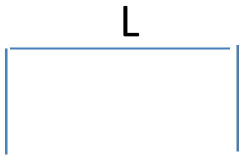
Other Important Non-normals

- Lorentizian: "Gaussian with tails"
  - unfortunately mean/variance of Lorentzian is infinite! (tails don't fall off fast enough somewhat like f(x) = 1/x)

$$\frac{1}{\pi}\left[\frac{\gamma}{(x-x_0)^2+\gamma^2}\right]$$

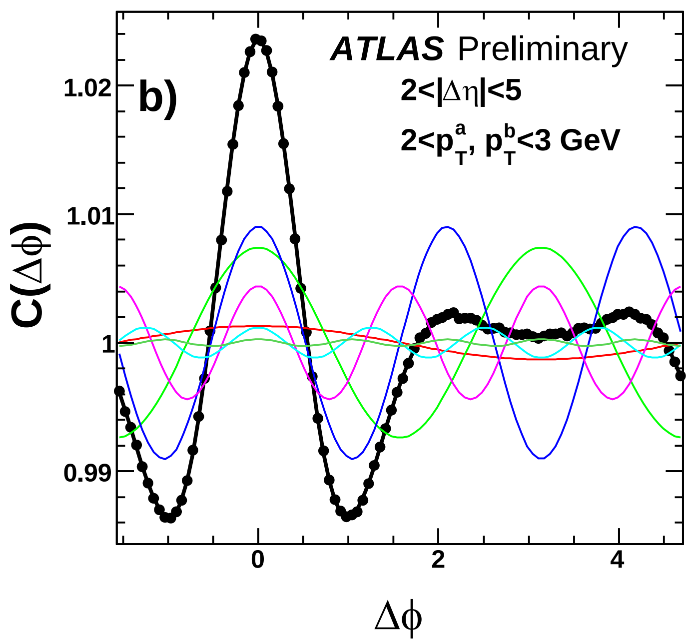


- Flat Distribution: sigma = L/√12



More details …

- Gaussian: Integral Normalizations trick
- First guess for peak shape: Central Limit Theorem
- Lorentzian: used to describe resonance peaks



Continue Tutorial

- See pages

Data Analysis with ROOT

- C++: Object Oriented Language
  - Don’t call functions as in fortran
  - rather: create object variables, tell them to run functions
  - variable type called “class”, formally defined
- ROOT is a Free C++ Analysis Tool / Framework
  - A set of “C++ class libraries” Do virtually anything!
  - A command line analyzer (C++ intepreter)
- Most physics (all types) analysis is done with a similar tools: IDL (astro) MatLab (condensed matter other science too) Root (Older version PAW/Fortran)
- Links on webpage

Things about C++/root to remember

- int, float char, arrays (int a[3]);   char *   pointers
- constructors;   text strings char *
- tab for "autocomplete" (class/type name)
- TH1F h1   TF1   cout << "hi"
  - .Fill()   .Draw()   "guass", "expo", "pol"
- TRandom r .Rndm()   ->Eval,   "[0] + [1]*x"
- Reference root.cern.ch ("Reference")
- .root_hist
- macros
-

Error Propagation

- Consider y = f(x)
  - Calc $y_i = f(x_i)$ from measurments x_i
- What is the distribution of y ?
  - The answer to this is complicated, (standard Ans: Gaus)
  - First focus instead on μ, σ
- Mean: $<y> = <f(x)>$
- Sigma/RMS: First how much should any deviation Δx modify f?
  - "Taylor" Series $f(x + \Delta x)$

Error Propagation: Multi-D

- Function of several variables f(u,v,x,…)
  - Ignore higher order terms
- "Multi-Dimensional"
  - Common Assumption: u, v, x, … orthogonal
  - More important: errors on each: independent (orthogonal)
- Main Error Formula
  - Variance and Covariance
  - Covariance: need to worry if variables have "CORRELATION"

Physics 371
Intermediate Lab I: Electrons
…(And Error)

Prof. Justin Frantz
frantz@ohio.edu
9/20/11

-

Lecture : -Finish Error Prop   -
-Ch 4 & 6 Points
-Big Tutorial

-Reading Assignment Ch 6.
-1 Homework Due Th
-Lab Manual

Propagation of Error/Uncertainty (σ)

- From Wikipedia ("Propagation of Uncertainty")

Function	Variance
$f = aA$	$\sigma_f^2 = a^2\sigma_A^2$
$f = aA \pm bB$	$\sigma_f^2 = a^2\sigma_A^2 + b^2\sigma_B^2 \pm 2ab\,\mathrm{cov}_{AB}$
$f = AB$	$\left(\frac{\sigma_f}{f}\right)^2 = \left(\frac{\sigma_A}{A}\right)^2 + \left(\frac{\sigma_B}{B}\right)^2 + 2\frac{\sigma_A\sigma_B}{AB}\rho_{AB}$
$f = \frac{A}{B}$	$\left(\frac{\sigma_f}{f}\right)^2 = \left(\frac{\sigma_A}{A}\right)^2 + \left(\frac{\sigma_B}{B}\right)^2 - 2\frac{\sigma_A\sigma_B}{AB}\rho_{AB}$
$f = aA^{\pm b}$	$\frac{\sigma_f}{f} = b\frac{\sigma_A}{A}$
$f = a\ln(\pm bA)$	$\sigma_f = ab\frac{\sigma_A}{A}$
$f = ae^{\pm bA}$	$\frac{\sigma_f}{f} = b\sigma_A$
$f = a^{\pm bA}$	$\frac{\sigma_f}{f} = b\ln(a)\sigma_A$

Partial derivatives

[edit]

Given $X = f(A, B, C, \ldots)$

Absolute Error	Variance
$\Delta X^2 = \left\|\frac{\partial f}{\partial A}\right\|^2 \cdot \Delta A^2 + \left\|\frac{\partial f}{\partial B}\right\|^2 \cdot \Delta B^2 + \left\|\frac{\partial f}{\partial C}\right\|^2 \cdot \Delta C^2 + \cdots$	$\sigma_X^2 = \left(\frac{\partial f}{\partial A}\sigma_A\right)^2 + \left(\frac{\partial f}{\partial B}\sigma_B\right)^2 + \left(\frac{\partial f}{\partial C}\sigma_C\right)^2 + \cdots$ [3]

Distribution: Gaussian Assumption?

- Some say these error prop formula's assume Gaussian errors
- This is true only in the sense that we usually still expect the result f(x) to be normally distributed
  - Mathematically it can be shown this is not true in many cases
- Central limit theorem: in many cases (e.g. sums) even if you don't start Gauss → get Gauss
  - also "contrary" to the above…

SEE DISTRIBUTIONS TUTORIAL

Error Prop: Usually however…

- Using the derivatives (like ones on previous page) is often a pain…
- AND it can be more exact to just use
  - $f(x)$, $f(x+\sigma_x)$, $f(x-\sigma_x)$
  - For multi-variables (computer) one can calculate all ± permutations of $f(u\pm\sigma_u, v\pm\sigma_v, x\pm\sigma_x \ldots)$
- Good Practical idea anyway:
  - When the differentiation gets complicated, checks for mistakes
  - If the two methods don't agree, then maybe other systematics are not completely understood
- In real practice this usually equates to doing the measurement several times with different input conditions
  - "f()" is like an operation: Example perform entire analysis w/ slightly different "cuts" , Raw * Corr = Measurement
  - Mostly this is for studying systematics (sys errors)

Important Applications of Error Prop

- …

Standard Distribution of the Mean

- So far we only discussed one sample at a time
- If one considers multiple samples *n* one may ask what is the distribution of the means in those samples $\mu_i = \bar{x}_i$
  - Again let's focus on mean and sigma
- Mean: 1/n Sum μ_i? No
  - Weighted Sum where weight higher for smaller uncertainty
  - $\bar{x}_\mu = \Sigma w_i \mu_i / \Sigma w_i$
  - Weight : $1/\sigma^2$ (For "usual" cases e.g. counting, will be propto N_i num of measurements in each sample i)
- Using error prop for sum
  - For identical sigma: sigma/sqrt(*n*) can be derived from Error prop of above we

- Demonstration: Poisson sqrt(N) → N independent 1+/- 1 measurements:

$$\sum_i^N 1 \pm 1 \rightarrow N \pm \sqrt{N} \rightarrow \bar{x} = 1 \pm \sigma_{\bar{x}} = \frac{1}{\sqrt{N}}$$

Standard Error : σ_μ

- Usually the error we quote should be su
- So far: use just σ or s ?
  - **Most important USE:** Sigma/RMS : gives idea how close single reproduction of measurement should be expected to be → Probability Tests
  - When small number of measurements (especially changing parameters for sys error studies) we usually use sigma for error
- When to quote σ_μ:
  - Remember σ_μ = 1/sqrt(N) assumes knowledge of parent distribution
  - Circular process: 1) several measurements 2) confirm distribution (measure sigma) 3) Then take 1/sqrt(N)
  - ….why Student's *t* Distribution for Prob Tests is needed later

Combining Measurements

- Generally Different Measurement/Samples will have different uncertainties
- Mean:
- Sigma:
- This is method used by Particle Data Group e.g. Pion Lifetime

Outliers/Probability Tests

- Chauvenet's criterion for throwing out outliers
  - If Prob < 0.5 "Events"

- Confidence Levels: Statement about how probable "Truth" is to sample μ based ***on*** $\boldsymbol{\sigma_\mu}$
  - ***Different than what to expect in reproduction experiment***
  - 1 sigma : 68% confidence level:
  - 2 sigma: ~95% (Whatever normal dist tells)
  - Very often these are used when no actual measurement is possible (e.g. a search for an effect that yields a null result, limits can be placed how large it could be)
- Student's *t* Distribution : Another Distribution of like gaussian but in terms of *t*
  - t = # sigma deviations (sigma = σ_μ, but requires N ie uses σ info)
  - More accurate for generating confidence levels samples with lower numbers of measurements or when outliers present

Physics 371
Intermediate Lab I: Electrons
…(And Error)

Prof. Justin Frantz
frantz@ohio.edu
9/22/11

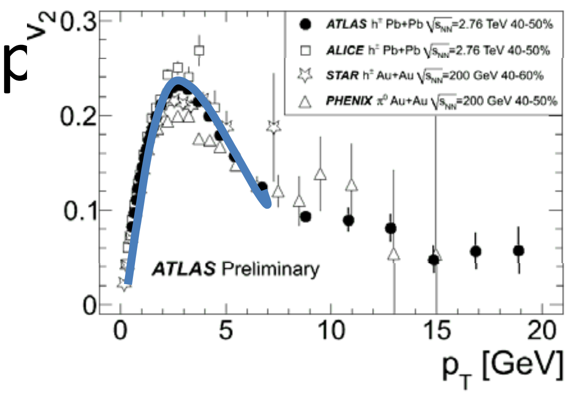
-

Lecture : -Last Error lecture: Fitting -
-Ch 6 Points
-Big Tutorial
-Intro into Labs.

-Due Tuesday: Preliminary Questions
-Lab Plan: Due Thursday

Intro to Curve Fitting

- This is for cases where you're changing parameter x_i and measuring y_i, x_i ($y_i(x_i)$)
  - (not y = f(x) calc → That's Error Prop
  - what prob density distribution describes measured values y_i,x_i
    - parent dist
    - A,B,C called Fit Parameters



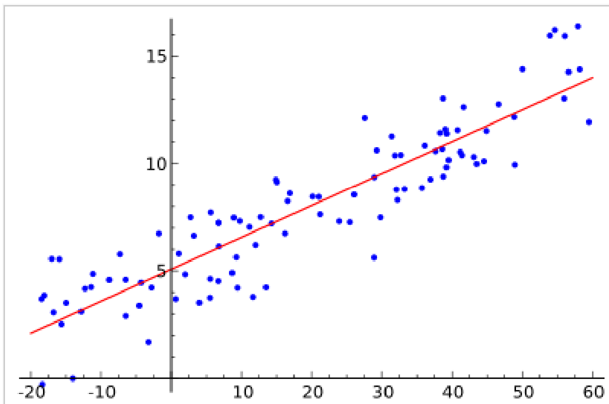
y = f(x, A,B,C…)

- Maximum Likelihood
- Minimize $\chi 2 = \sum_{i=1}^{k} \left(\frac{X_i - \mu_i}{\sigma_i} \right)^2$

Reduced χ^2 : χ_ν^2
= χ^2/num degrees of freedom
ndof = k – # fit parameters

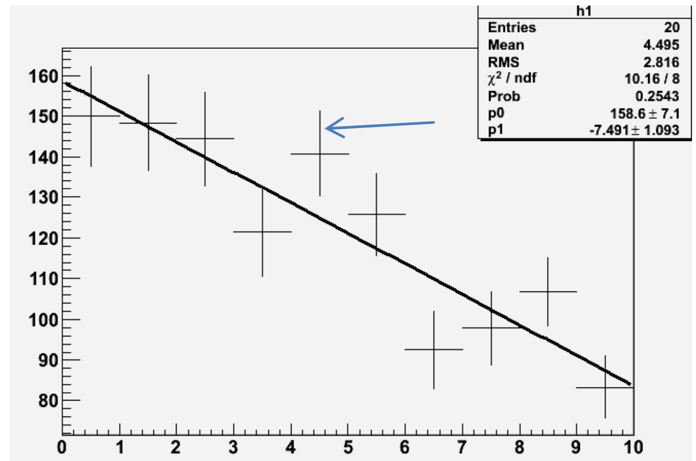
(Linear) Regression vs Physics Fitting

- “Regression” vs. Physics Fittting



Example of simple linear regression, which has one independent variable.

Regression



Physics ChiSq Min Fit

- Depends on what you’re minimizing:
  - Least squares fitting (ambiguous)
  - Regression: min $(y_i - y(x_i)^2)$ or $(\mathbf{ri} - \mathbf{r}(x))$
  - In physics we assign an error to every measurement so we think it only makes sense to min $\chi 2$

How to do practically

- With ROOT or any Analysis Tool, automatic, very easy… any functional form
  - Provide automatic calc's of x2 and most importantly σA,σB, … Error estimates of fitted parameters A, B, C
- Usually other functions polynomials can be accomodated with linear fit
  - y = eBx x3→ y = A + Bx^3
- Linear Fit functionality in Excel:
  - Trendline can be used for when uncertainties are all the same (absolute size)
  - Use solver: see web link

Data Analysis with ROOT

- C++: Object Oriented Language
  - Don't call functions as in fortran
  - rather: create object variables, tell them to run functions
  - variable type called "class", formally defined
- ROOT is a Free C++ Analysis Tool / Framework
  - A set of "C++ class libraries"   Do virtually anything!
  - A command line analyzer (C++ intepreter)
- Most physics (all types) analysis is done with a similar tools: IDL (astro) MatLab (condensed matter other science too) Root (Older version PAW/Fortran)
- Links on webpage

Things about C++/root to remember

- int, float char, arrays (int a[3]); char * pointers
- constructors; text strings char *
- tab for "autocomplete" (class/type name)
- TH1F h1 TF1 cout << "hi"
  - .Fill() .Draw() "guass", "expo", "pol"
- TRandom r .Rndm() ->Eval, "[0] + [1]*x"
- Reference root.cern.ch ("Reference")
- .root_hist
- macros
-