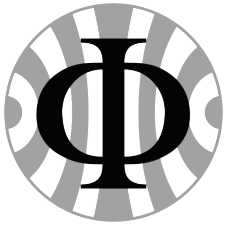
Statistical Methods in Particle Physics / WS 13

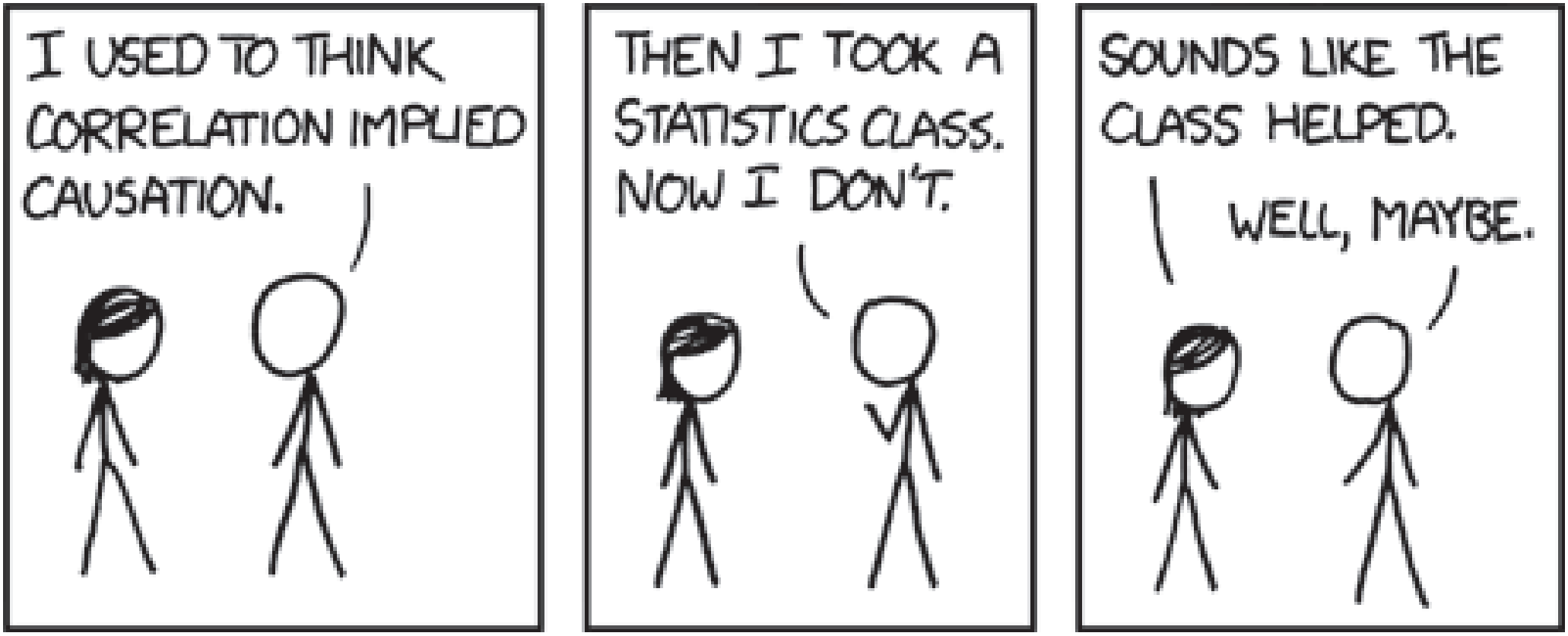
Lecture I

Probability Theory

Niklaus Berger

Physics Institute, University of Heidelberg





http://xkcd.com/522/

Aims of the course

- Understand Statistical Concepts

  Ability to read analysis papers

Read an analysis paper...

Acceptances and efficiencies are obtained mostly from full simulations of the ATLAS detector [81] using GEANT4 [82]. These simulations include a realistic modelling of the pile-up conditions observed in the data.

The uncertainty on the electron energy scale results in an uncertainty of ±0.7% (±0.5%/±0.2%) on the mass scale of the $m_{4\ell}$ distribution for the $4e$ ($2e2\mu/2\mu2e$) channel.

For the 7 TeV data, this information is combined in a neural network, tuned to achieve a similar jet rejection as the cut-based selection described in Ref. [95], but with higher photon efficiency.

The parameter of interest is the global signal strength factor μ, which acts as a scale factor on the total number of events predicted by the Standard Model for the Higgs boson signal. This factor is defined such that $\mu = 0$ corresponds to the background-only hypothesis and $\mu = 1$ corresponds to the SM Higgs boson signal in addition to the background. Hypothesised values of μ are tested with a statistic $\lambda(\mu)$ based on the profile likelihood ratio [122]. This test statistic extracts the information on the signal strength from a full likelihood fit to the data. The likelihood function includes all the parameters that describe the systematic uncertainties and their correlations.

Exclusion limits are based on the CL_s prescription [123]; a value of μ is regarded as excluded at 95% CL when CL_s is less than 5%. A SM Higgs boson with mass m_H is considered excluded at 95% confidence level (CL) when $\mu = 1$ is excluded at that mass.

Aims of the course

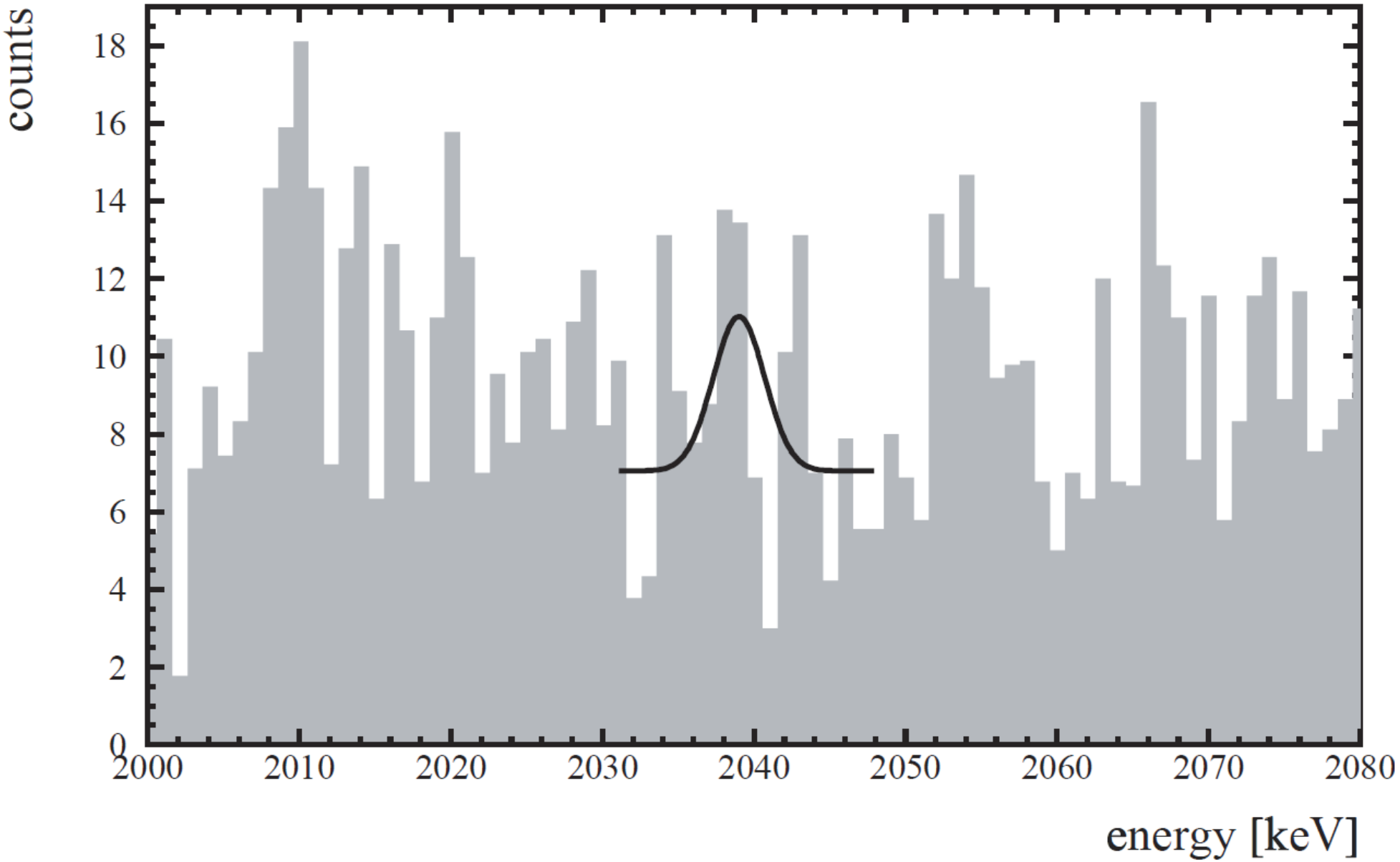
- Understand Statistical Concepts

  Ability to read analysis papers

- Develop Intuition

  Do I believe a statistical claim?

Do you believe a statistical claim



Aims of the course

- Understand Statistical Concepts
  Ability to read analysis papers
- Develop Intuition
  Do I believe a statistical claim?
- Know Methods
  Your statistical toolbox

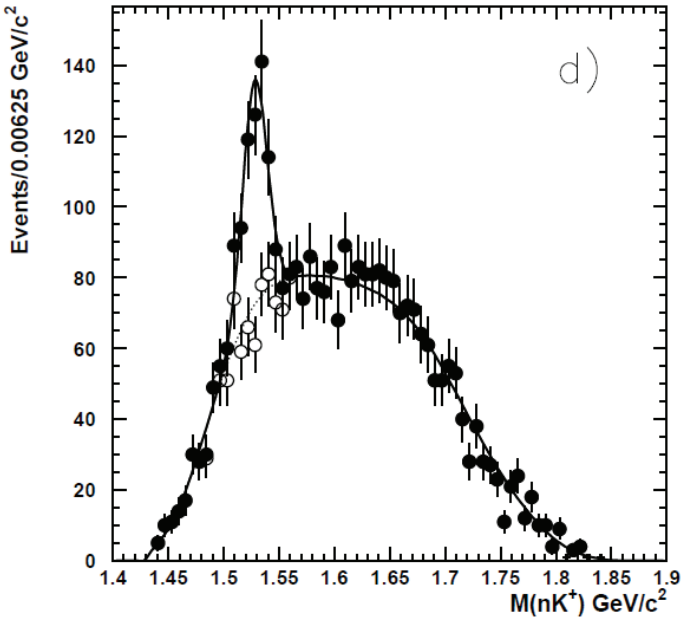
Your statistical toolbox

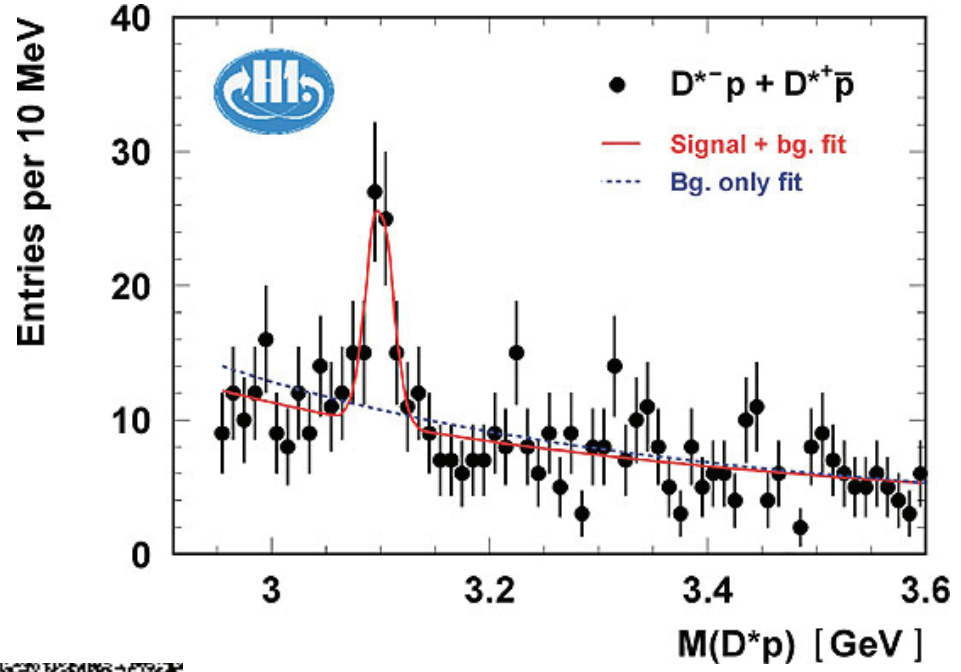
Photo: David I Poole

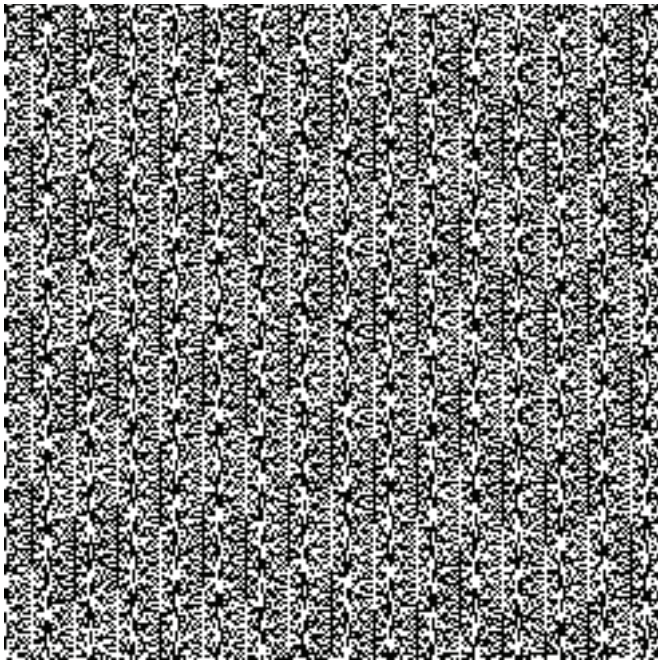
Aims of the course

- Understand Statistical Concepts
  Ability to read analysis papers
- Develop Intuition
  Do I believe a statistical claim?
- Know Methods
  Your statistical toolbox
- Know Pitfalls
  Avoid mistakes already made by others

Avoid Mistakes Made by Others



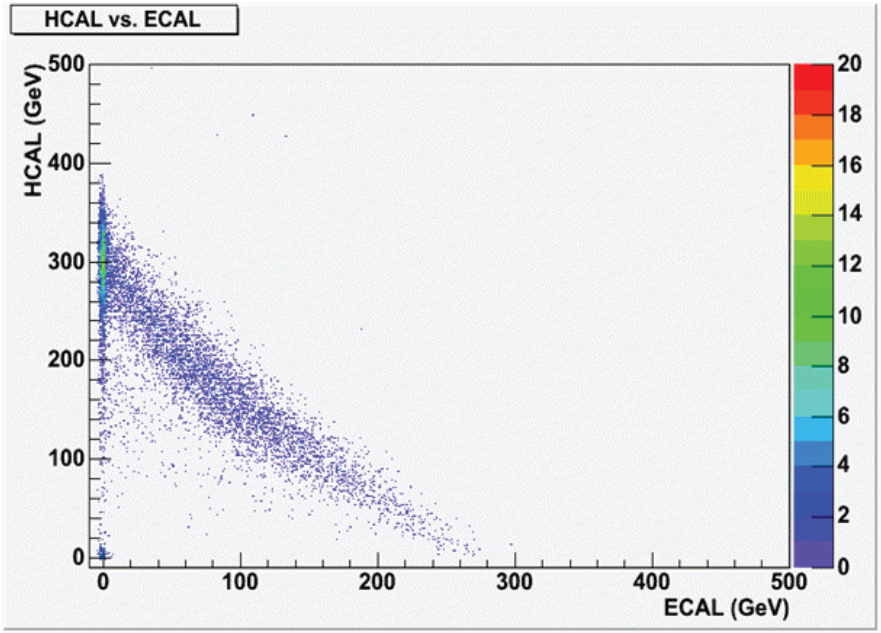


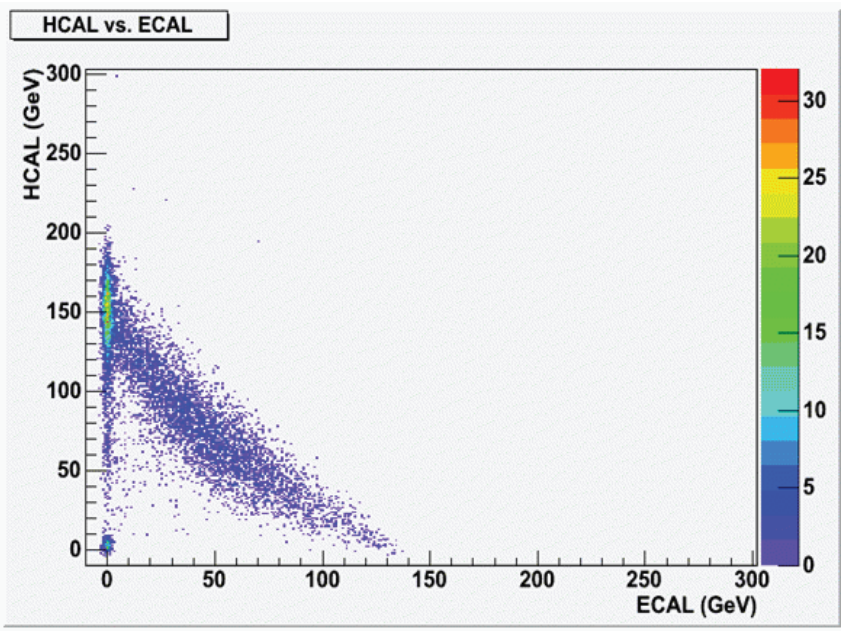


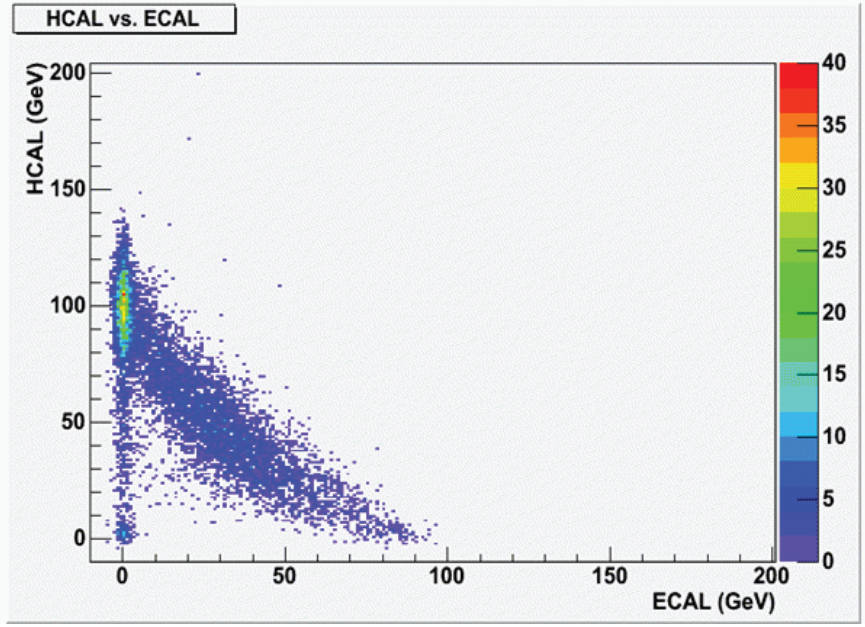
Aims of the course

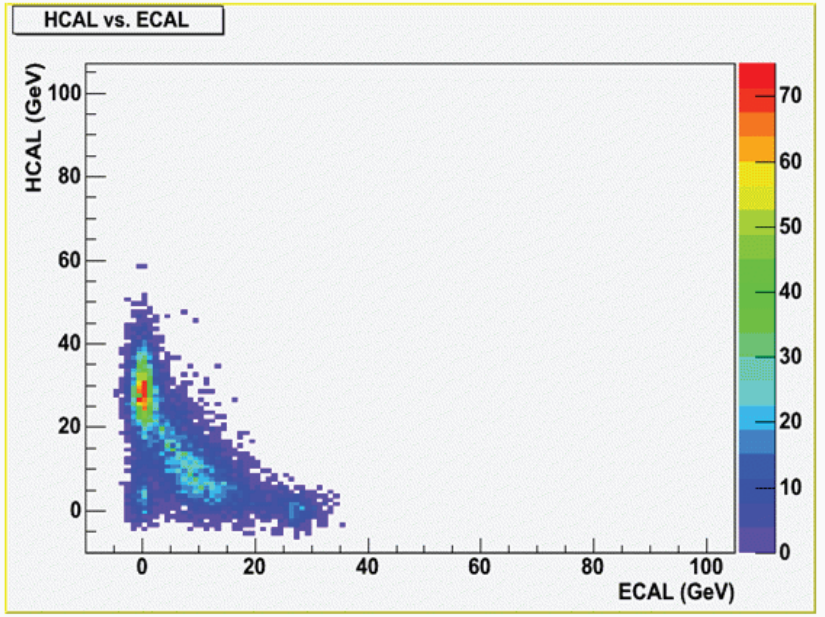
- Understand Statistical Concepts
  Ability to read analysis papers
- Develop Intuition
  Do I believe a statistical claim?
- Know Methods
  Your statistical toolbox
- Know Pitfalls
  Avoid mistakes already made by others
- Use Tools
  Learn to use root

Learn to use root





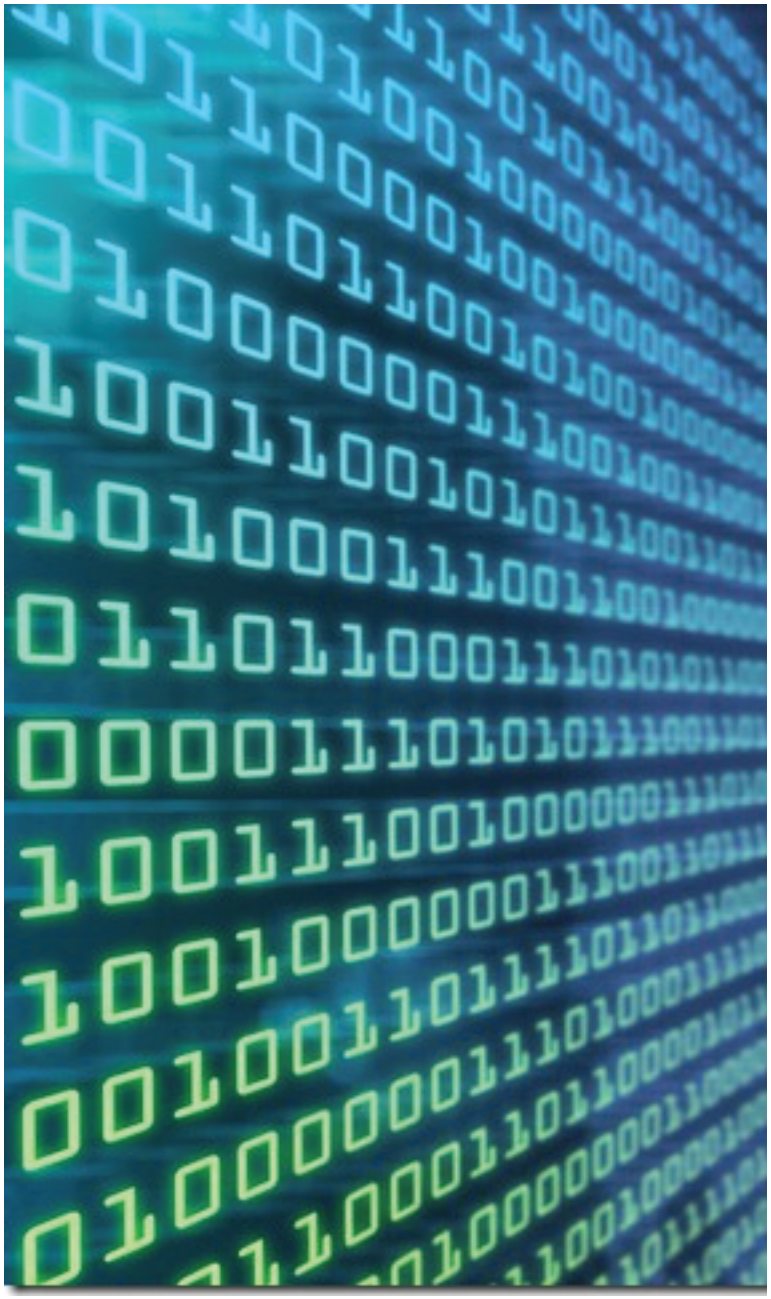




Aims of the course

- Understand Statistical Concepts
  Ability to read analysis papers
- Develop Intuition
  Do I believe a statistical claim?
- Know Methods
  Your statistical toolbox
- Know Pitfalls
  Avoid mistakes already made by others
- Use Tools
  Learn to use root
- Practice
  Get ready for your own data analysis

This is an applied course!



- We will use lots of examples from "real life" particle physics
- We will sometimes talk about implementation on a computer
- You should ask questions, discuss
- You will write lots of code

Practicalities

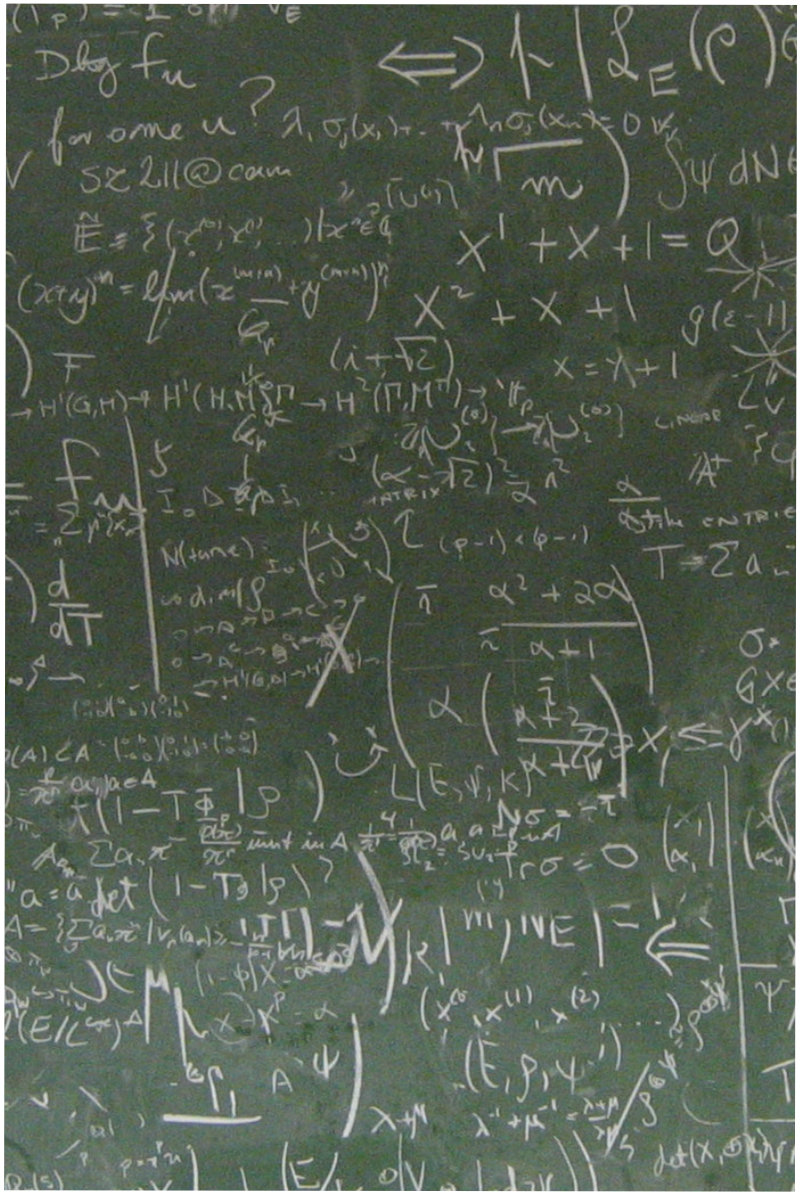
- No exam, no marks
- Tutorials are an integral part
- You get your credit points for coding
- Details from Oleg

Material

- There is a course website:

`http://www.physi.uni-heidelberg.de/~nberger/teaching/ws13/statistics/statistics.php`

- You will get the slides
- When I use the blackboard, you should take notes...
- There are a few good books out there



Literature

- Lyons:
  Statistics for Nuclear and Particle Physicists
  (Cambridge University Press)
- Cowan:
  Statistical Data Analysis
  (Oxford Science Publications)
- Barlow:
  Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences
  (Manchester Physics Series)
- Blobel, Lohrmann:
  Statistische Methoden der Datenanlyse
  (Teubner, in German)
  Ebook: `http://www.desy.de/~blobel/eBuch.pdf`

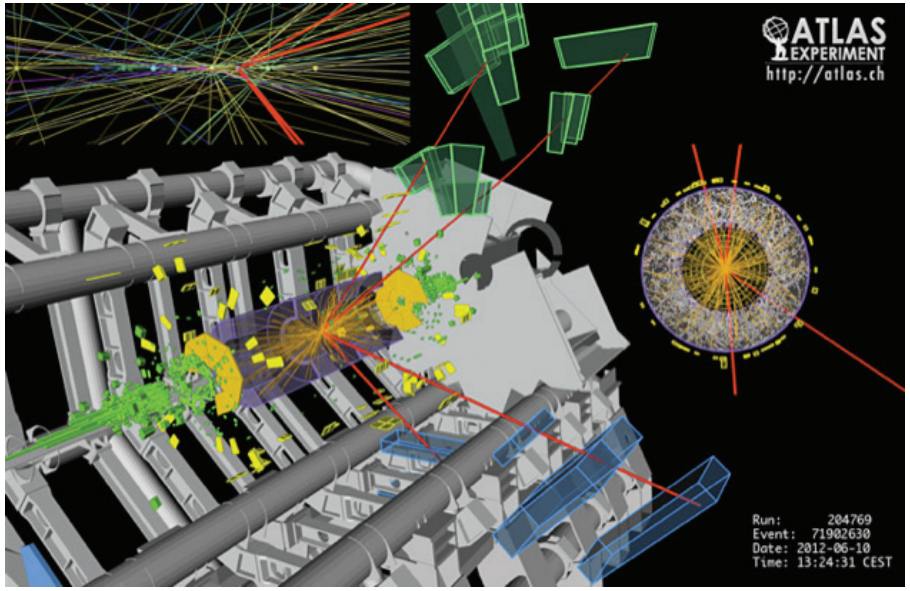
Curriculum

- Probability theory
- Probability density functions and their properties
- Error propagation, correlations
- Monte Carlo techniques
- Estimators
- Fitting:
  - Least Squares
  - Maximum Likelihood
- Confidence intervals and limits
- Multivariate methods
- Unfolding

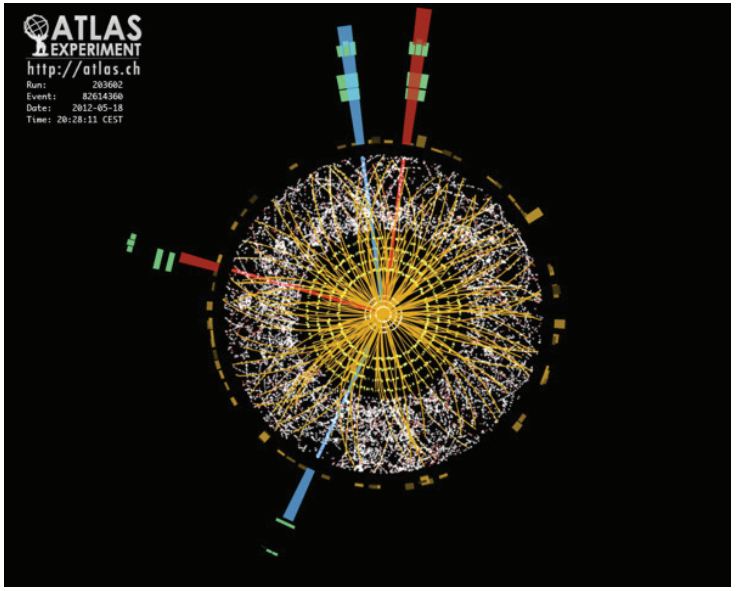
Recurring topics:

- Bayesian vs. Frequentist interpretation
- Examples from our work
- Examples from your work
- Discussion of numerical implementations and programming techniques
- The role of the physicists judgement

Statistics in Particle Physics

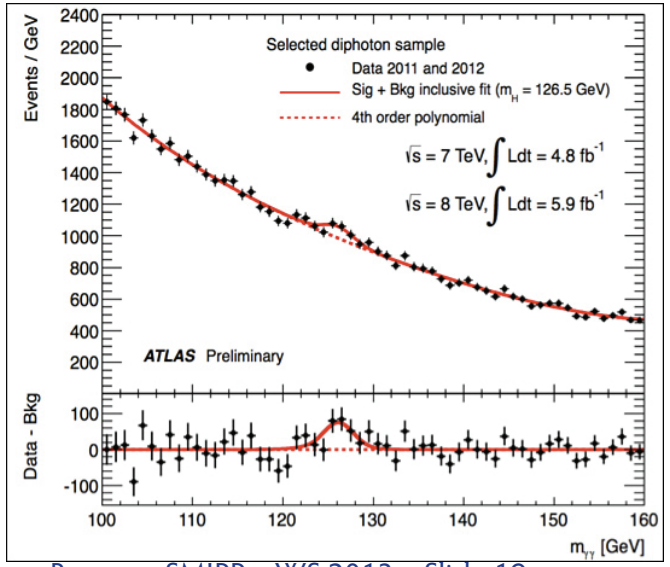


Detector modelling (Monte Carlo methods)

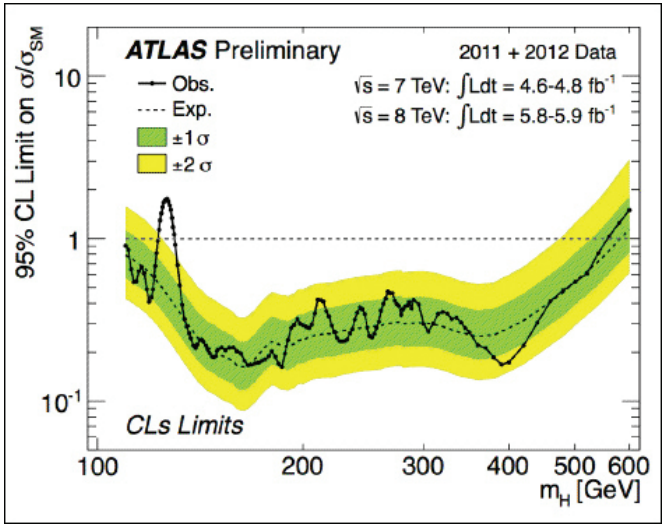


Event reconstruction (e.g. track fitting)

Data analysis (e.g. event classification)



Interpretation of results (e.g. limits)

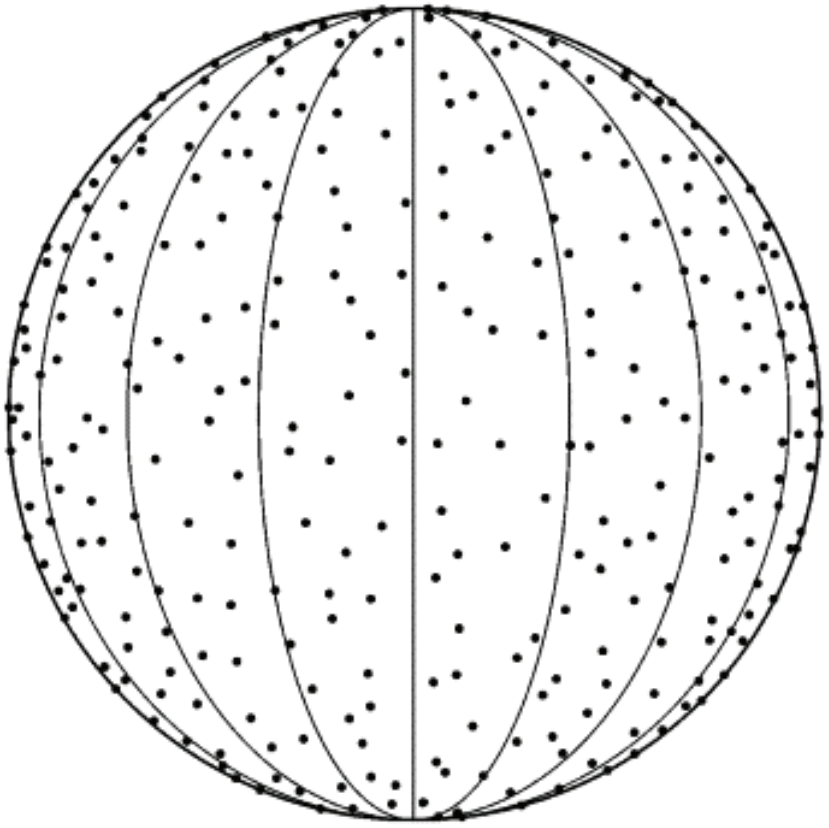


Part I:

Probability Theory

The Borel-Kolmogorov Paradox

- Distribute points equally over a sphere
- Points on the equator will also be equidistributed
- Points on any great circle, e.g. a meridian will be equidistributed



The Borel-Kolmogorov Paradox

- Distribute points equally over a sphere (Earth)
- Points on the equator will also be equidistributed
- Points on any great circle, e.g. a meridian will be equidistributed

- A quarter of every meridian lies north of 45° N
- Integrate over all meridians: A quarter of all points lie north of 45° N, points are equi-distributed, thus a quarter of the earth's surface lies north of 45° N

Contradiction?

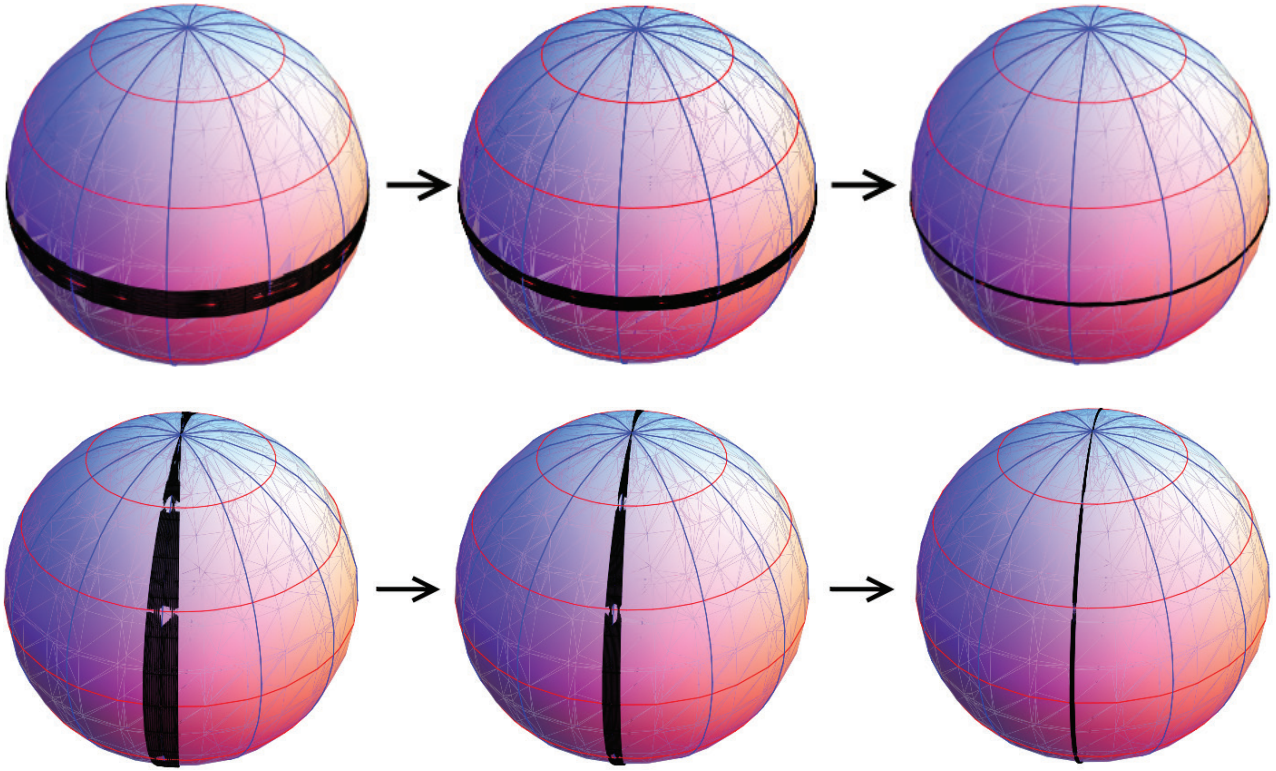
The Borel-Kolmogorov Paradox

Where did this go wrong?

The Borel-Kolmogorov Paradox

Where did this go wrong?

- It is not allowed to condition conditional probabilities on a set of probability zero
- Instead take finite surface and the proper limit



Images Greg Gandenberger