PHYSTAT 2003

Statistical Problems in Particle Physics, Astrophysics and Cosmology

Louis Lyons

Oxford & KEK

ACAT03, Dec 2003

PHYSTAT 2003

SLAC
Stanford, California
8th - 11th September 2003

Conference on Statistical Problems in Particle Physics, Astrophysics and Cosmology

An interdisciplinary gathering of Physicists in these fields and Statisticians

Topics of relevance include:

- Signal significance
- Setting limits
- Systematics
- Non-parametric estimation
- Cluster analysis
- Spatial data
- Time series analysis
- Unfolding/deconvolution
- Blind analyses
- Multivariate classification
- Machine learning
- Goodness of fit
- Hypothesis testing

http://www-conf.slac.stanford.edu/phystat2003/ for further details and for registration information.

PHYSTAT 2003

- History of Conferences
- Overview of PHYSTAT 2003
- Specific Items
  - Bayes and Frequentism
  - Goodness of Fit
  - Systematics
  - Signal Significance
  - At the pit-face
- Where are we now ?

HISTORY

Where	CERN	Fermilab	Durham	SLAC
When	Jan 2000	March 2000	March 2002	Sept 2003
Issues	Limits	Limits	Wider range of topics (incl "Partons")	Wider range of topics
Physicists	Particles	Particles +3 astrophysicists	Particles +3 astrophysicists	Particles + Astro + Cosmo
Statisticians	3	3	2	Many

Statistical Problems in Particle Physics, Astrophysics, and Cosmology

PHYSTAT 2003 *an interdisciplinary gathering of statisticians and physicists in the*

Home
Committees
Program
Recommended Reading
Registrants
Author Instructions
Reviewed Paper Submission
Reviewer Acceptance Form
Paper Status

SLAC, Stanford, California :: 8th - 11th September

The aim of the Conference is to discuss advanced statistical techniques as used in experimental analyses of data in Particle Physics, Astrophysics and Cosmology.

This Conference concentrates on the statistical issues related to both measurements and searches. It follows on from the Workshops on Limits at CERN and at Fermilab in 2000, and the Conference at Durham in 2002. However it differs from the earlier meetings in that we plan to have much more significant participation of Astro-Particle Physicists and Statisticians at this Conference.

Topics of relevance include:

- Signal significance
- Setting limits
- Systematics
- Non-parametric estimation
- Spatial data
- Unfolding convolution
- Blind analyses
- Multivariate classification
- Machine Learning
- Variability of sources
- Hypothesis testing
- Goodness of fit
- Cluster analysis

There will be invited talks and submitted contributions. We do not intend to include the reporting of current experimental results or standard statistical analyses.

The following is a list of invited speakers and titles:

Roger Barlow	"Introduction to statistical issues in Particle Physics"
Persi Diaconis	"A place for Philosophy? Bayesian Statistics in 2003"
Seth Digel	"Statistical Issues in High-Energy Gamma-Ray Astronomy for GLAST"
Jonathan Dorfan	Welcome
Brad Efron	"Bayesians, Frequentists, and Physicists (Keynote address)"
Eric Feigelson	"Statistical challenges in Modern Astronomy"
Jerry Friedman	"Modern Developments in Machine Learning"
Christopher Genovese	"Doing Cosmology with balls and envelopes"
Fred James	"Unified Approach to Understanding Statistics"

Bob Nichol	"Data mining cosmological data sets: the need for fast and efficient algorithms"
Frank Porter	"Statistical issues in Particle Physics: a view from BaBar"
Harrison Prosper	"Multivariate Analysis: Past, Present and Future"
Nancy Reid	"Likelihood based inference with nuisance parameters"
John Rice	"Conference Summary"
Pekka K. Sinervo	"The mysteries of systematic uncertainties: theory and practice"
Philip Stark	"Using what we know: inference with physical constraints"
Daniel Stump	"Statistical and theoretical uncertainties of parton distribution functions"
David van Dyk	"Hierarchical models, data augmentation and Markov Chain Monte Carlo"
Licia Verde	"Statistical challenges of WMAP cosmic microwave background analysis"

Scientific Committee:

- Roger Barlow
- Bob Cousins
- Glen Cowan
- Seth Digel
- Brad Efron
- Jerry Friedman
- Fred James
- Dean Karlen
- Jim Linnemann
- Tom Loredo
- Louis Lyons
- Bill Murray
- Vahe Petrosian
- Frank Porter
- John Rice
- Jeffrey Scargle
- Peter Shawhan
- Pekka Sinervo
- Diego Torres
- Guenther Walther
- Steve Yellin
- Harrison Prosper

Local Organizing Committee

- Richard Mount - Chairman
- Arla LeCount
- Joseph Perl
- David Lee

The meeting will take place in the Auditorium at SLAC. About 100 rooms have been reserved on a 'first come first served' basis in the new SLAC Guest House.

The fee for the Conference is $175 if paid before June 30th, or $225 afterwards. This does not include accommodation or meals. Payment can be made online via this web-site. We are also planning to have a welcoming reception and barbeque.

For questions not dealt with on this web-site, contact phystat2003-l@slac.stanford.edu
design/webmaster: D.Lee

Workshop on 'Confidence Limits'

17-18 January, 2000
CERN Council Chamber

Workshop Homepage

Program

Papers and Slides NEW

Required Reading

Registration

List of participants

Information for speakers

Practical Info

- How to get to CERN
- Workshop location
- Housing
- Dinner
- Paper submission
- Proceedings

FNAL CL Workshop

The proceedings are now available as CERN Yellow Report:
CERN 2000-005 30 May 2000 - Workshop on Confidence Limits

This workshop will be devoted to the problem of setting confidence limits in difficult cases: small or unobserved signal, background larger than signal, background not well known, and measurements near a physical boundary. Among the many examples in HEP are the Higgs mass, accelerator searches for neutrino oscillations, B_s mixing, SUSY, compositeness, neutrino masses, and dark matter. Several different methods are on the market - the CL_s methods used by the LEP Higgs searches; Bayesian; Feldman-Cousins and modifications thereof; etc .

Distinguished invited speakers will present the methods currently being proposed, and all participants are encouraged to submit abstracts for short contributed papers, of which a limited number will be selected for presentation. Proceedings will include summaries of all papers presented, as well as discussions.

Co-convenors

Fred James / CERN
Louis Lyons / Oxford

Local Organization

Fred James
Yves Perrin

Sorry, we cannot accept any more registrations, as we have reached the capacity of the conference room (the Council Chamber).

However, we have arranged for video transmission of all sessions to the nearby Room B. Those who are unable to register may follow the workshop from Room B, where a big screen and video projector will provide excellent viewing conditions.

A few exceptional registrations may be possible for people who can demonstrate they have made a significant contribution in the field. Requests should be made directly to one of the convenors.

There is no conference fee. The workshop is intended for high energy physicists and people in related fields concerned with the problem of setting statistical confidence limits. All participants will be expected to be familiar with the methods under discussion by having read the publications on the list of 'required reading'. Useful background material will also be presented in the Academic Training lectures of F. James at CERN, 10-14 January, where confidence limits will be discussed from different points of view.

Proceedings

The proceedings are now available as CERN Yellow Report:
CERN 2000-005 30 May 2000 - Workshop on Confidence Limits

We will produce proceedings which will contain summaries of all the invited talks and submited papers, as well as the important points brought up during the discussions, to be published as a CERN Yellow Report.

Comments to CLW workshop Web Administrator - Last modified on Friday January 12 2001

Workshop on Confidence Limits

27-28 March, 2000
Fermilab 1-West Conference Room

An increasingly common activity in our community is setting limits (Higgs mass, accelerator searches for neutrino oscillations, B_s mixing, SUSY, compositeness, neutrino masses, dark matter .....) There are several competing methods on the market - the CL_s methods used by the LEP Higgs searches; Bayesian; Feldman-Cousins and modifications thereof; etc.

This one-and-a-half day meeting at Fermilab will include:

1. talks by the proponents of the various methods; some comparison of their relative merits; methods of including uncertainties in background estimates; and
2. talks by the various experimental groups on the techniques they currently use for extracting limits.

The invited speakers to date include Bob Cousins, Fred James, Louis Lyons, Bill Murray, Harrison Prosper and Byron Roe.

Fred James will give a special series of three lectures on the present status in this field 22-24 March. A black and white copy of his complete talk transparencies is available in PDF (4 MB) and PostScript (5.5 MB) formats.

The dates for the meeting are Monday (all day) and Tuesday morning, March 27th and 28th. If you are interested in attending, please register on this web-site. There is no charge for attending, but if the CERN example is anything to go by, it may be necessary to limit participation because of space constraints.

The workshop is intended for high energy physicists and people in related fields concerned with the problem of setting statistical confidence limits, to help them decide on the approach to use. All participants should have some familiarity with the methods under discussion. A similar Workshop is being held at CERN on January 17th and 18th. There the participants are expected to be already fairly expert on the subject before the meeting - there is an extensive reading list to be studied before the meeting. It is evisaged that the Fermilab meeting will be at a more instructional level.

Convenor

Louis Lyons / Oxford, FNAL
(l.lyons@physics.ox.ac.uk)

Local Organization

Harry W.K. Cheung
(cheung@fnal.gov)
John Conway
(conway@pion.rutgers.edu)
Jim Linnemann
(linnemann@fnal.gov)
Cynthia Sazama

Registration

There is no registration fee. However the number of participants is limited so registration is required. Please see the registration page for information.

Proceedings

There will be no proceedings published for this meeting. However a CERN Yellow Report should be available that contains the summaries of all the invited talks and submited papers, as well as the important points brought up during the discussions at the earlier CERN Workshop on Confidence Limits.

Comments to Harry W.K. Cheung - Last modified on Monday, 10-Jan-2000

University of Durham

The Institute for Particle Physics Phenomenology

University of Durham

will host a Conference on

ADVANCED STATISTICAL TECHNIQUES IN PARTICLE PHYSICS

at

The University of Durham, UK, March 18 - 22, 2002

Topics to be covered include:

Setting Limits Signal Significance Systematics
Combining Results Unfolding Convolution Simulation Issues
Multivariate Event Classification Techniques for Blind Analysis
Statistical Issues to do with Parton Distributions

Organising Committee

Roger Barlow (Manchester)
Bob Cousins (UCLA)
Glen Cowen (RHUL)
Fred James (CERN)
Dean Karlen (Carleton)
Jim Linnemann (Michigan State)
Louis Lyons (Oxford)
Bill Murray (RAL)
Harrison Prosper (Florida State)
Pekka Sinervo (Toronto)

Local Organising Committee

James Stirling
Mike Whalley
Linda Wilkinson

Further information and registration procedures can be obtained via WWW at http://www.ippp.dur.ac.uk/statistics/

Issues

- Bayes versus Frequentism
- Limits, Significance, Future Experiments
- Blind Analyses
- Likelihood and Goodness of Fit
- Multivariate Analysis
- Unfolding
- At the pit-face
- Systematics and Frequentism

Talks at PHYSTAT 2003

2 Introductory Talks

8 Invited talks by Statisticians

8 Invited talks by Physicists

47 Contributed talks

Panel Discussion

Underlying much of the discussion:

Bayes and Frequentism

Invited Talks by Statisticians

Brad Efron	Bayesian, Frequentists & Physicists
Persi Diaconis	Bayes
Jerry Friedman	Machine Learning
Chris Genovese	Multiple Tests
Nancy Reid	Likelihood and Nuisance Parameters
Philip Stark	Inference with physical constraints
David VanDyk	Markov chain Monte Carlo
John Rice	Conference Summary

Invited Talks by Physicists

Eric Feigelson	Statistical issues for Astroparticles
Roger Barlow	Statistical issues in Particle Physics
Frank Porter	BaBar
Seth Digel	GLAST
Ben Wandelt	WMAP
Bob Nichol	Data mining
Fred James	Teaching Frequentism and Bayes
Pekka Sinervo	Systematic Errors
Harrison Prosper	Multivariate Analysis
Daniel Stump	Partons

Bayes versus Frequentism

Old controversy

Bayes 1763 Frequentism 1937

Both analyse data (x) ➔ statement about parameters (μ)

e.g. Prob ($\mu_l \leq \mu \leq \mu_u$) = 90%

but very different interpretation

Both use Prob (x; μ)

Bayesian $$P(A;B) = \frac{P(B;A) \text{ x } P(A)}{P(B)}$$ Bayes Theorem

$$P(param;data) \, \alpha \, P(data;param) \text{ x } P(param)$$

↑ posterior ↑ likelihood ↑ prior

Problems: P(param) True or False

"Degree of belief"

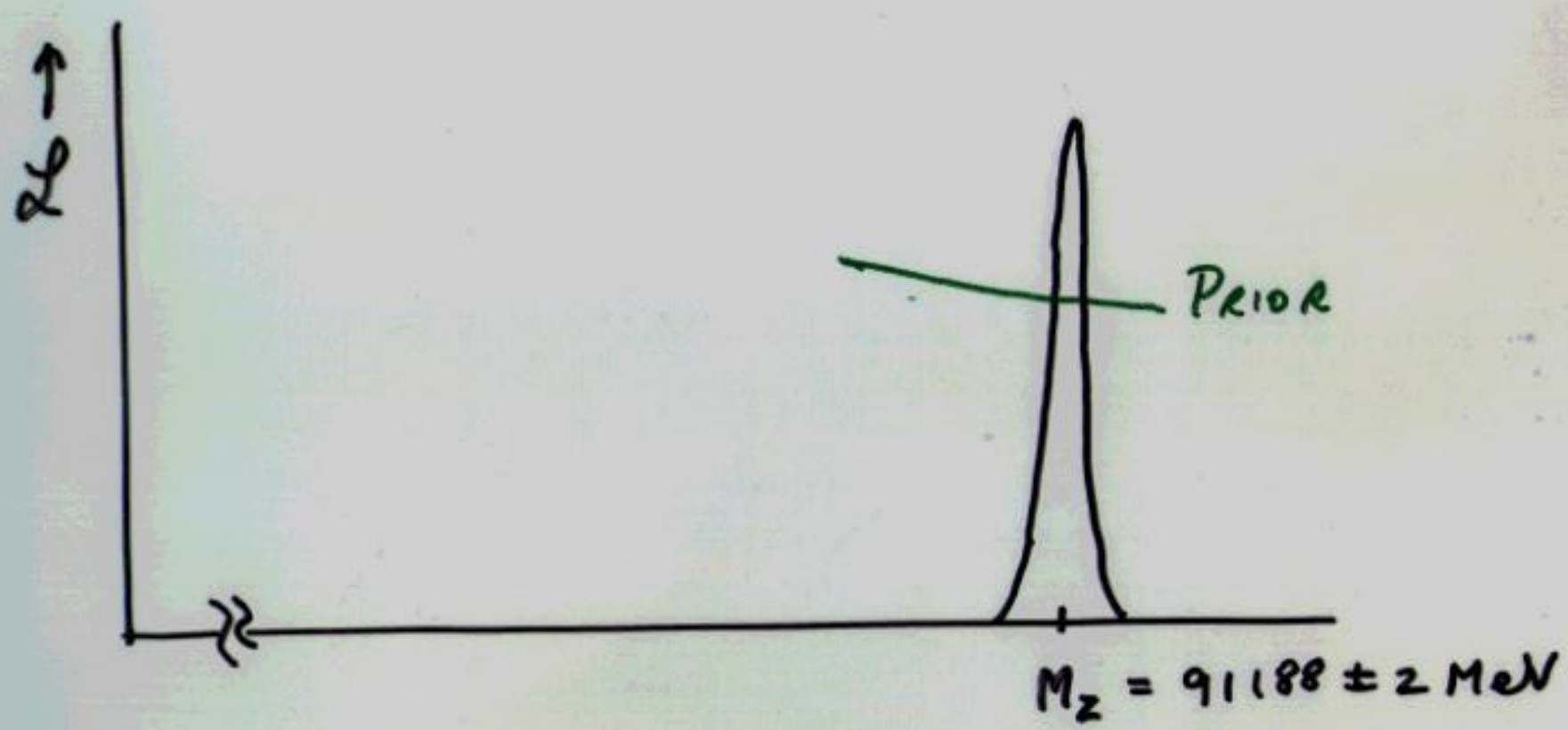
Prior What functional form?

Flat? Which variable?

Unimportant when "data overshadows prior"

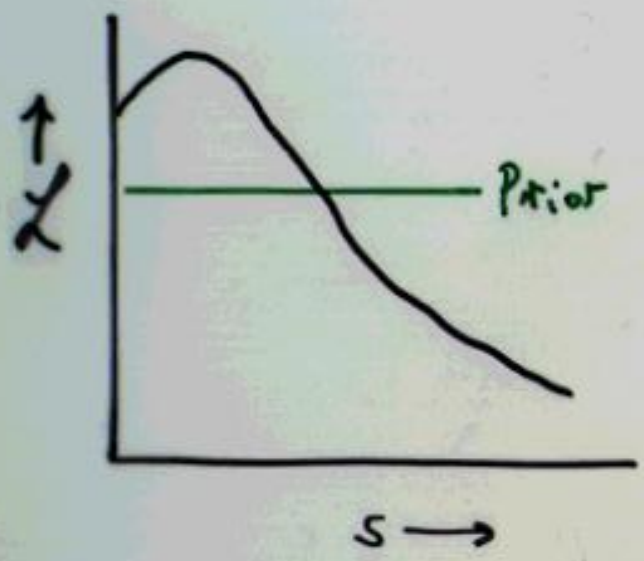
Important for limits

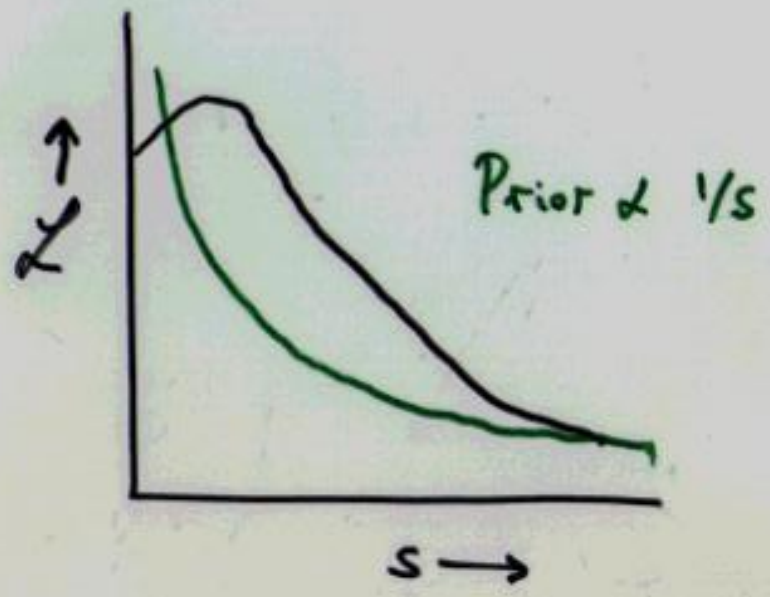
DATA OVERSHADOWS PRIOR



DATA ⇒ UPPER LIMIT ON SIGNAL s

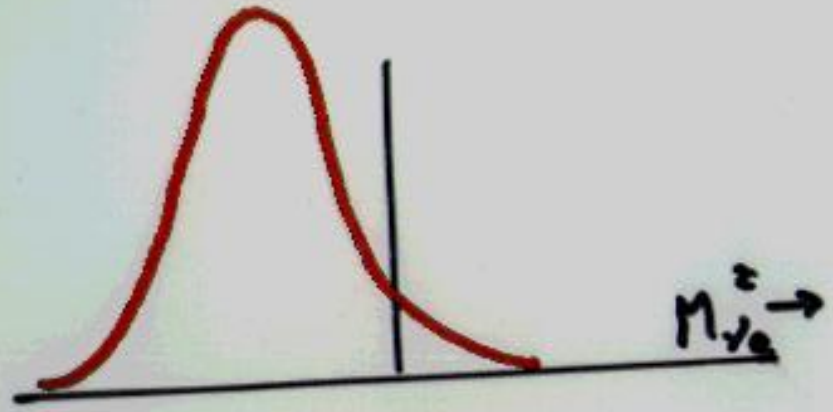
CHOICE OF PRIOR AFFECTS LIMIT





MASS-SQUARED OF ν_e

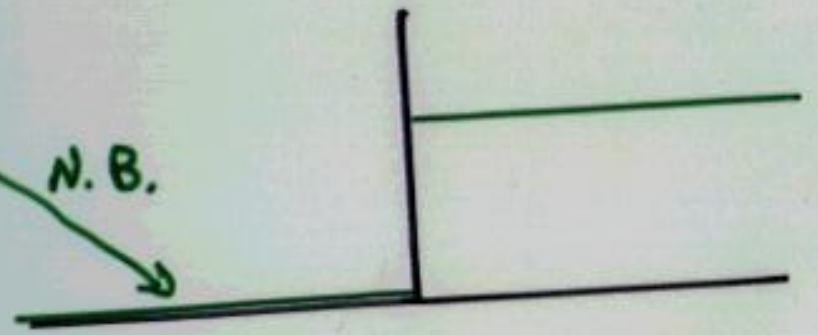
"$\mathcal{L}$"



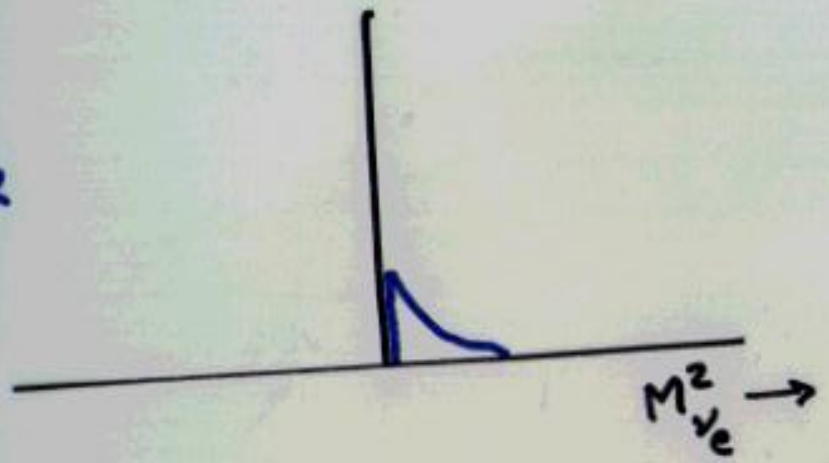
BAYESIAN PRIOR ALWAYS ENSURES POSTERIOR GIVES A PHYSICAL RANGE

PRIOR

N.B.



POSTERIOR



CLASSICAL (NEYMAN) CONFIDENCE INTERVALS

Uses only P(data | theory)

FIGURES

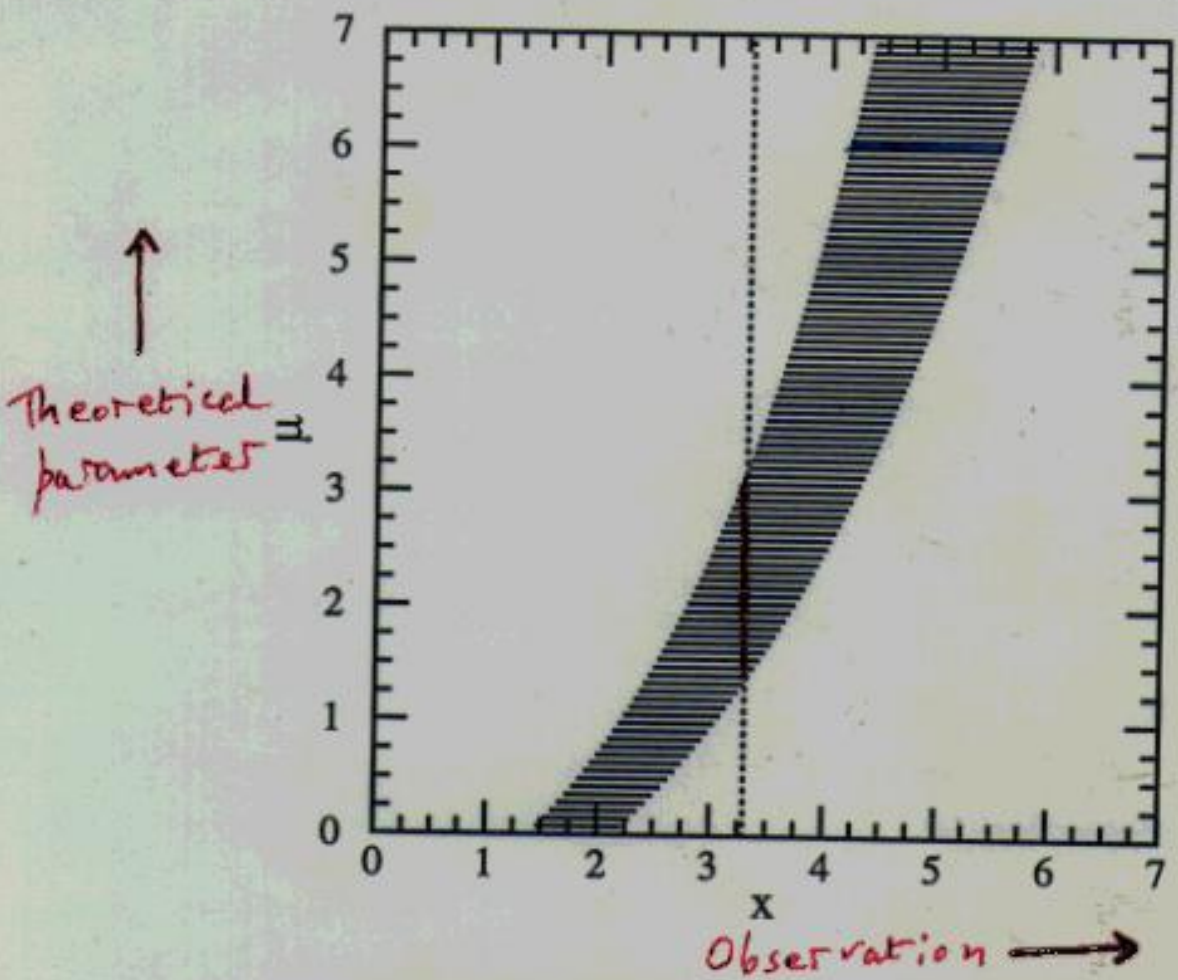


FIG. 1. A generic confidence belt construction and its use. For each value of μ, one draws a horizontal acceptance interval $[x_1, x_2]$ such that $P(x \in [x_1, x_2] | \mu) = \alpha$. Upon performing an experiment to measure x and obtaining the value x_0, one draws the dashed vertical line through x_0. The confidence interval $[\mu_1, \mu_2]$ is the union of all values of μ for which the corresponding acceptance interval is intercepted by the vertical line.

NO PRIOR INVOLVED

BF8 8

$\mu_l \leq \mu \leq \mu_u$ at 90% confidence

Frequentist

μ_l and μ_u known, but random

μ unknown, but fixed

Probability statement about μ_l and μ_u

Bayesian

μ_l and μ_u known, and fixed

μ unknown, and random

Probability/credible statement about μ

$$P\,(\text{Data};\text{Theory}) \neq P\,(\text{Theory};\text{Data})$$

Theory = male or female

Data = pregnant or not pregnant

P (pregnant ; female) ~ 3%

but

P (female ; pregnant) >>>3%

$$P(\text{Data};\text{Theory}) \neq P(\text{Theory};\text{Data})$$

HIGGS SEARCH at CERN

Is data consistent with Standard Model?

or with Standard Model + Higgs?

End of Sept 2000 Data not very consistent with S.M.

Prob (Data ; S.M.) < 1% valid frequentist statement

Turned by the press into: Prob (S.M. ; Data) < 1%

and therefore Prob (Higgs ; Data) > 99%

i.e. "It is almost certain that the Higgs has been seen"

WHY 5σ?

$$\text{Prob}[\text{Theory}; \text{Data}] \neq \text{Prob}[\text{Data}; \text{Theory}]$$

$$\underbrace{\frac{\text{Prob}[\text{Exotic}; \text{Data}]}{\text{Prob}[\text{S.M.}; \text{Data}]}}_{\text{Posterior}} = \underbrace{\frac{\text{Prob}(\text{Data}; \text{Exotic})}{\text{Prob}(\text{Data}; \text{S.M})}}_{\text{Likelihood}} \times \underbrace{\frac{\text{Prob}(\text{Exotic})}{\text{Prob}(\text{S.M.})}}_{\text{Prior}}$$

Want this ratio to be very large to compensate for

[How many σ for claiming Special Relativity is wrong?]

Bayesian versus Frequentism

	Bayesian	Frequentist
Basis of method	Bayes Theorem --> Posterior probability distribution	Uses pdf for data, for fixed parameters
Meaning of probability	Degree of belief	Frequentist defintion
Prob of parameters?	Yes	Anathema
Needs prior?	Yes	No
Choice of interval?	Yes	Yes (except F+C)
Data considered	Only data you have	....+ more extreme
Likelihood principle?	Yes	No

Bayesian versus Frequentism

	Bayesian	Frequentist
Ensemble of experiments	No	Yes (but often not explicit)
Final statement	Posterior probability distribution	Parameter values → Data is likely
Unphysical/ empty ranges	Excluded by prior	Can occur
Systematics	Integrate over prior	Extend dimensionality of frequentist construction
Coverage	Unimportant	Built-in
Decision making	Yes (uses cost function)	Not useful

Bayesianism versus Frequentism

"Bayesians address the question everyone is interested in, by using assumptions no-one believes"

"Frequentists use impeccable logic to deal with an issue of no interest to anyone"

Goodness of Fit

Basic problem:

χ^2 very general applicability, but

- Requires binning, with $n_i > 5.....20$ events per bin. Prohibitive with sparse data in several dimensions.
- Not sensitive to signs of deviations

O.K. by χ^2 $\Rightarrow$

K-S and related tests overcome these, but work in 1-D

So, need something else.

Goodness of Fit Talks

Zech	Energy test
Heinrich	$L_{\max}$?
Yabsley & Kinoshita	
Raja	
Narsky	What do we really know?
Pia	Software Toolkit for Data Analysis
Ribon	.....
Blobel	Comments on χ^2 minimisation

Goodness of Fit

Gunter Zech "Multivariate 2-sample test based on logarithmic distance function"

See also:

Aslan & Zech, Durham Conf., "Comparison of different goodness of fit tests"

R.B. D'Agostino & M.A. Stephens, "Goodness of fit techniques", Dekker (1986)

The statistic *energy*

The idea:

Compute potential energy of system of point charges

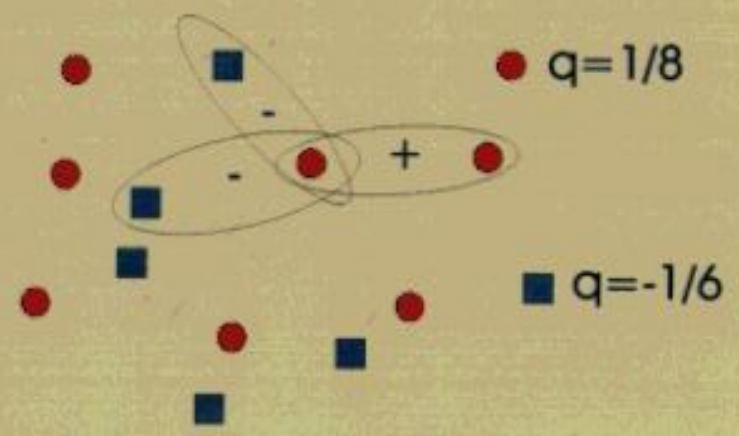
Sample A: positive $q_i = 1/n$
Sample B: negative $q_i = -1/m$
(total charge is zero)

$$\sum q_i = 0$$

Interaction energy is positive for charges of same sign and negative for charges of opposit sign.

Potential energy decreases with distance.

The energy is minimum if positive and negative charges follow the same distribution. (energy of vacuum)



Definition:

$$\phi_{nm} = \frac{1}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} R(|\mathbf{x}_i - \mathbf{x}_j|) - \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} R(|\mathbf{x}_i - \mathbf{y}_j|) + \frac{1}{m^2}\sum_{i=1}^{m-1}\sum_{j=1+1}^{m} R(|\mathbf{y}_i - \mathbf{y}_j|)$$

$$\phi_{\boldsymbol{A}} \qquad \phi_{\boldsymbol{AB}} \qquad \phi_{\boldsymbol{B}}$$

$$R(r) = -\log|r + \varepsilon|$$

x: position vector of sample *A* observations

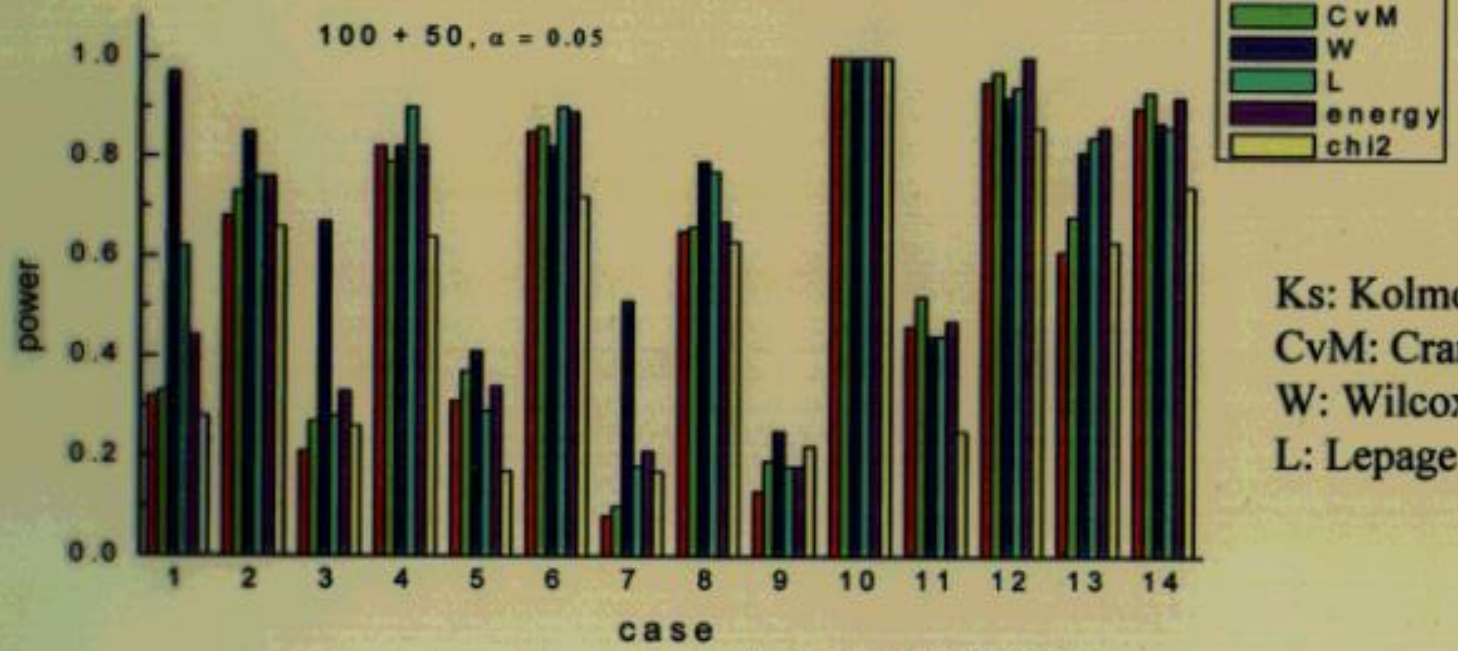
y: position vector of sample *B* observations

R: distance function

ε: cut-off parameter (uncritical)

($R(r)=1/r$ would corresponds to electrostatic energy of n point charges with charge –1/n each and m point charges of charge +1/m each.)

case	P_1	P_2	θ	τ
1	1	7	0.3	1.3
2	1	7	0.4	0.8
3	7	2	0.3	1.3
4	7	2	0.4	0.8
5	2	3	0.3	1.3
6	2	3	0.4	0.8
7	2	9	0.3	1.3
8	2	9	0.4	0.8
9	6	5	0.3	1.3
10	6	5	0.4	0.8
11	3	8	0.3	1.3
12	3	8	0.4	0.8
13	8	4	0.3	1.3
14	8	4	0.4	



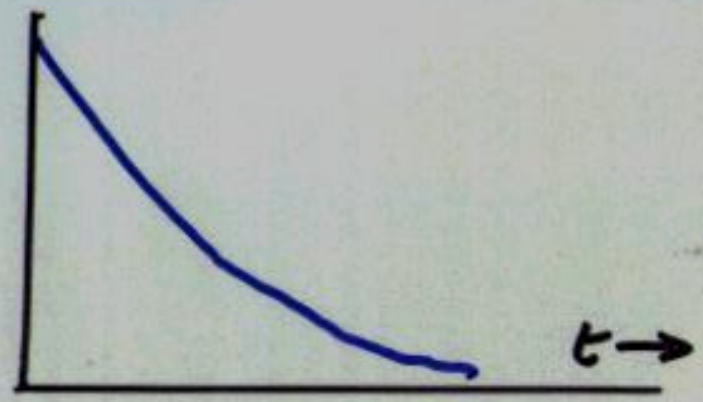
Ks: Kolmogorov-Smirnov
CvM: Cramer-van-Mises
W: Wilcox
L: Lepage

Summary

- Binning-free tests avoid arbitrary bin widths and boundaries.
- They are usually more powerfull than binned tests.
- The energy two-sample test with log. distance function is simple and has no free parameter.
- It samples all distances independent of the number of dimensions.
- The distribution of the test statistic is obtained from permutations drawn from the mixed sample.
- In 1d there exist many powerful two-sample tests. The energy test has comparable power as other common tests.
- In 2d and 4d it is superior to competing tests in the considered examples.
- The energy statistic provides also a powerful g.o.f. test.

$L_{\max}$ not very useful

e.g. Lifetime dist $\frac{1}{\tau} e^{-t/\tau}$



Fit for τ

$$\ln L_{max} = -N(1 - \ln \bar{t})$$

i.e. function only of $\bar{t}$, not their distribution

Therefore any data with the same $\bar{t}$ →same $L_{\max}$ ***

so $L_{\max}$ not useful for testing distribution

(Distribution of $L_{\max}$ due simply to different $\bar{t}$ in samples)

Likelihood & Goodness of Fit

Joel Heinrich CDF note #5639

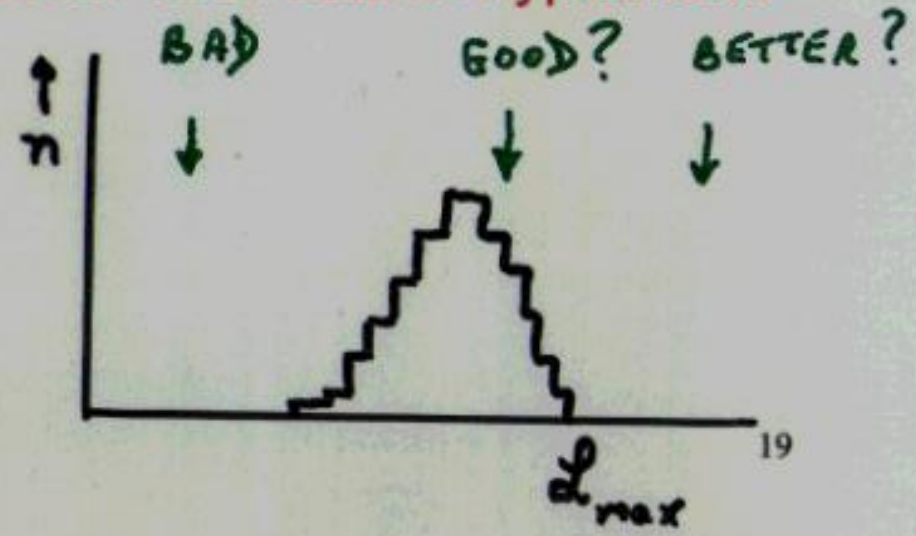
Faulty Logic:

Parameters determined by maximising L

So larger $L_{\max}$ is better

So larger $L_{\max}$ implies better fit of data to hypothesis

Monte Carlo dist of $L_{\max}$
for ensemble of expts



SYSTEMATICS

For example

$$N_{events} = \sigma LA + b$$

↑ Observed ↑ $N \pm \sqrt{N}$ for statistical errors

↑ Physics parameter

↑ ↑ we need to know these, probably from other measurements (and/or theory)

Uncertainties → error in σ

Some are arguably statistical errors

$$LA = LA_0 \pm \sigma_{LA}$$

$$b = b_0 \pm \sigma_b$$

Shift Central Value

Bayesian

Frequentist

Mixed

Shift Nuisance Parameters

$$N_{events} = \sigma \, LA + b$$

Simplest Method

Evaluate σ_0 using LA_0 and b_0

Move nuisance parameters (one at a time) by their errors → $\delta\sigma_{LA}$ & $\delta\sigma_b$

If nuisance parameters are uncorrelated,

combine these contributions in quadrature

→ total systematic

Bayesian

Without systematics

$$p(\sigma; N) \propto p(N; \sigma)\underset{\substack{\uparrow \\ \text{prior}}}{\Pi(\sigma)}$$

With systematics

$$p(\sigma, LA, b; N) \propto p(N; \sigma, LA, b)\underset{\substack{\uparrow \\ \sim \Pi_1(\sigma)\Pi_2(LA)\Pi_3(b)}}{\Pi(\sigma, LA, b)}$$

Then integrate over LA and b

$$p(\sigma; N) = \iint p(\sigma, LA, b; N)\, dLA\, db$$

$$p(\sigma; N) = \iint p(\sigma, LA, b; N)\, dLA\ db$$

If $\Pi_1(\sigma)$ = constant and $\Pi_2(\mathrm{LA})$ = truncated Gaussian TROUBLE!

Upper limit on σ from $\int p(\sigma; N)\, \mathrm{d}\sigma$

Significance from likelihood ratio for $\sigma = 0$ and $\sigma_{\max}$

Frequentist

Full Method

Imagine just 2 parameters σ and LA

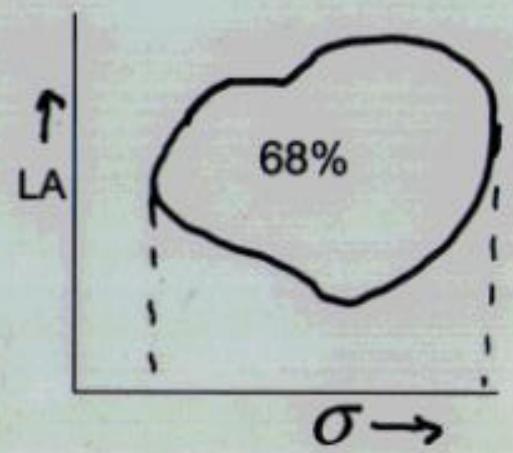
and 2 measurements N and M

↑ Physics ↑ Nuisance

Do Neyman construction in 4-D

Use observed N and M, to give

Confidence Region for LA and σ



Then project onto σ axis

This results in OVERCOVERAGE

Aim to get better shaped region, by suitable choice of ordering rule

Example: Profile likelihood ordering

$$\frac{\mathrm{L}(N_0, M_0; \sigma, LA_{best}(\sigma))}{\mathrm{L}(N_0, M_0; \sigma_{best}, LA_{best}(\sigma))}$$

Full frequentist method hard to apply in several dimensions

Used in ≤ 3 parameters

For example: Neutrino oscillations (CHOOZ)

$$\sin^2 2\theta\,,\ \Delta m^2$$

Normalisation of data

Use approximate frequentist methods that reduce dimensions to just physics parameters

e.g. <u>Profile</u> pdf

i.e. $pdf_{profile}(N;\sigma) = pdf(N, M_0; \sigma, LA_{best})$

Contrast Bayes marginalisation

Distinguish "profile <u>ordering</u>"

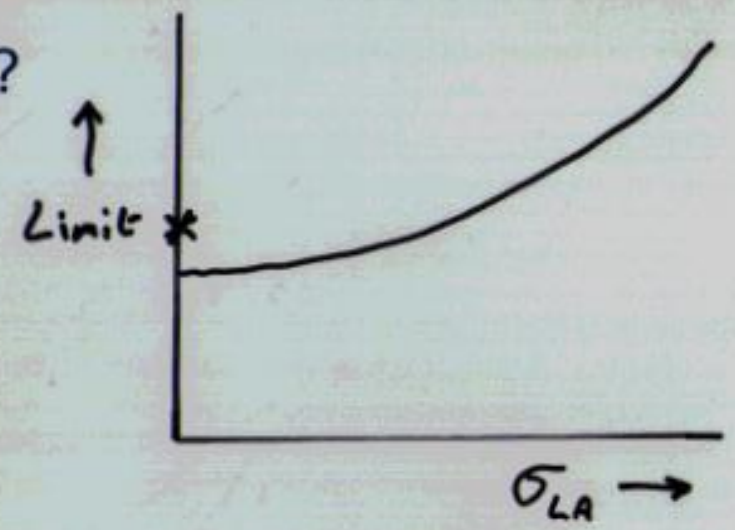
Properties being studied by Giovanni Punzi

Method: Mixed Frequentist - Bayesian

Bayesian for nuisance parameters and

Frequentist to extract range

Philosophical/aesthetic problems?



Highland and Cousins

NIM A320 (1992) 331

(Motivation was paradoxical behavior of Poisson limit when LA not known exactly)

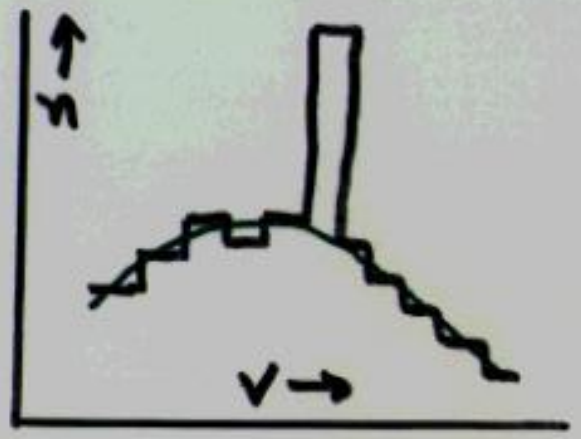
Systematics & Nuisance Parameters

Sinervo	Invited Talk (cf Barlow at Durham)
Barlow	Asymmetric Errors
Dubois-Felsmann	Theoretical errors, for BaBar CKM
Cranmer	Nuisance Param in Hypothesis Testing
	Higgs search at LHC with uncertain bgd
Rolke	Profile method
	see also: talk at FNAL Workshop
	and Feldman at FNAL

(N.B. Acceptance uncertainty worse than bgd uncertainty)

Demortier	Berger and Boos method

Significance



Significance = $S/\sqrt{B}$?

Potential Problems:

- Uncertainty in B
- Non-Gaussian behavior of Poisson
- Number of bins in histogram, no. of other histograms [FDR]
- Choice of cuts (Blind analyses
- Choice of bins Roodman and Knuteson)

For future experiments

- Optimising $S/\sqrt{B}$ could give S =0.1, B = 10^{-6}

Systematics: Tests

Do test (e.g. does result depend on day of week?) $\Rightarrow a \pm b$

Barlow: Are you (a) estimating effect, or (b) just checking?

- If (a), correct and add error
- If (b), ignore if OK, worry if not OK

BUT:

1) Quantify OK
2) What if still not OK after worrying?

My solution:

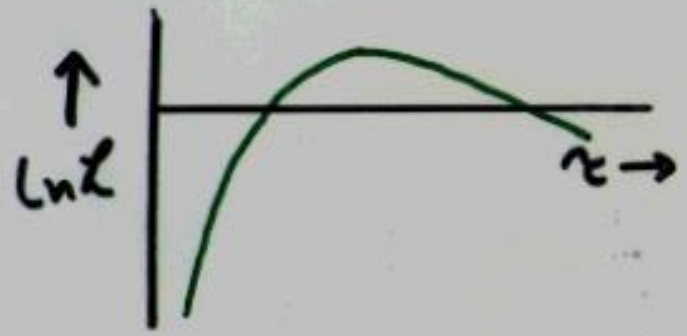
Contribution to systematic's variance is $a^2 - b^2$,

even if negative!

Talks on Significance

Genovese	Multiple Tests
Linnemann	Comparing Measures of Significance
Rolke	How to claim a discovery
Shawhan	Detecting a weak signal
Terranova	Scan statistics
Quayle	Higgs at LHC
Punzi	Sensitivity of future searches
Bityukov	Future exclusion/discovery limits

Barlow: Asymmetric Errors



e.g. $\tau = 5^{+4}_{-2}$

Either statistical or systematic

How to combine errors

(Combine upper errors in quadrature is clearly wrong)

How to calculate χ^2

How to combine results

CRANMER: TEST H_0 = No Higgs

The Basic Approach

H_1 = Higgs

IN PRESENCE OF UNCERTAINTY ON B&D = b

Perform the Neyman Construction in 3-D for $L(x, M|H_0, b)$

Use ordering rule that tests H_0 against H_1

Make a Measurement at M_0

Make a Measurement at x_0

If confidence interval for b (under H_0) empty $\Rightarrow$ Discovery

Comparison with Cousins-Highland

For a 5% systematic uncertainty on background
and a measurement at $M_0 = 100$
(with an expectation of 50 signal events)

Must observe 167 events to claim a discovery

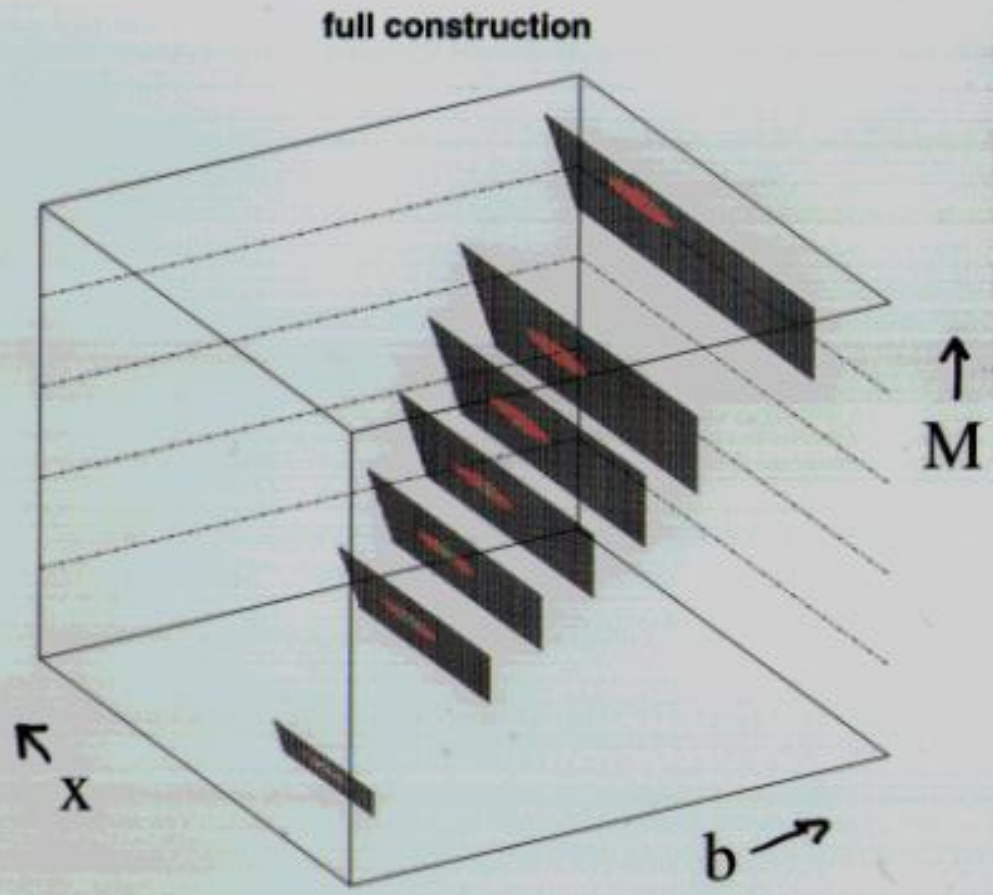
With the Cousins-Highland approach only need 161 events

Slight Modification

Modify Likelihood Ratio, which is independent of b

$$l = \frac{L(x, M|H_0, \hat{\hat{b}})}{L(x, M|H_1, \hat{b})}$$

Just put a window in M around $L(x, M|H_0, b)$



DEMORTIER

Nuisance Parameters (6)

A final comment on likelihood ratio p values is that they are two-sided: in the limit where $\Delta\nu \to 0$, the p value based on λ will be twice the p value based on n.

2. Confidence Interval Method

A standard p value that takes nuisance parameters into account is the so-called supremum p value:

$$p_{\text{sup}} = \sup_{\nu} \mathbb{P}(n \geq n_0 \,;\, \mu = 0, \nu)$$

This does not work for our example however, because the domain of ν is unbounded from above, so this p value is always 100%.

A solution was proposed independently by R.L. Berger and D.D. Boos (1994), and M.J. Silvapulle (1996). The idea is to construct a $1-\beta$ confidence interval C_β for the nuisance parameter ν, to maximize the p value over that interval, and to correct the result for the fact that $\beta \neq 0$:

$$p_\beta = \sup_{\nu \in C_\beta} \mathbb{P}(n \geq n_0 \,;\, \mu = 0, \nu) + \beta.$$

It can be shown that this is a *valid* p value, i.e.:

$$\mathbb{P}(p_\beta \leq \alpha) \leq \alpha, \quad \text{for each } \alpha \in [0, 1].$$

Berger + Boos JASA 89 (1994) 1012

Multivariate Analysis

Friedman	Machine learning
Prosper	Experimental review
Cranmer	A statistical view
Loudin	Comparing multi-dimensional distributions
Roe	Reducing the number of variables
	(Cf. Towers at Durham)
Hill	Optimising limits via Bayes posterior ratio
Etc.	

Blind Analyses

Potential problem: Experimenters' bias

Original suggestion?

Luis Alvarez concerning Fairbank's 'discovery' of quarks

Aaron Roodman's talk

Methods of blinding:

- Keep signal region box closed
- Add random numbers to data
- Keep Monte Carlo parameters blind
- Use part of data to define procedure

Don't modify result after unblinding, unless.....

Select between different analyses in pre-defined way

See also Bruce Knuteson:

QUAERO, SLEUTH, Optimal binning

From the Pit-face

Roger Barlow	Asymmetric errors
William Quayle	Higgs search at LHC
etc.	

From Durham:

Chris Parkes	Combining W masses and TGCs
Bruce Yabsley	Belle measurements

Where are we?

- Things that we learn from ourselves
  - Having to present our statistical analyses
- Learn from each other
  - Likelihood not pdf for parameter
    - Don't integrate L
  - Conf int not Prob(true value in interval; data)
  - Bayes' theorem needs prior
  - Flat prior in m or in m^2 are different
  - Max prob density is metric dependent
  - Prob (Data;Theory) not same as Prob(Theory;Data)
  - Difference of Frequentist and Bayes (and other) intervals wrt Coverage
  - Max Like not usually suitable for Goodness of Fit

Where are we?

- Learn from Statisticians
  - –Update of Current Statistical Techniques
  - –Bayes: Sensitivity to prior
  - –Multivariate analysis — Boosted decision trees
    - –Neural nets
    - –Kernel methods
    - –Support vector machines
  - –Hypothesis Testing : False discovery rate
  - –Goodness of Fit : Friedman at Panel Discussion
  - –Nuisance Parameters : Several suggestions

Conclusions

Very useful physicists/statisticians interaction

e.g. Upper Limit on Poisson parameter when:

observe n events

background has some uncertainty

acceptance has some uncertainty

For programs, transparencies, papers, etc. see:

http://www-conf.slac.stanford.edu/phystat2003

Future: Conference in spring/summer 2005?

Suggestions to: l.lyons@physics.ox.ac.uk