

---

# Asymmetric Errors

A statistician's perspective

Joint work with: Roger Barlow and Igor Volobouev

**PHYSTAT-Informal Reviews**

February 28, 2024

---

## Thanks!

March Hare: *"See all the trouble you've started?"*

Alice: *"Really, I didn't think..."*

March Hare: *"But that's the point. If you don't think, you shouldn't talk."*

**You made me think a lot!**



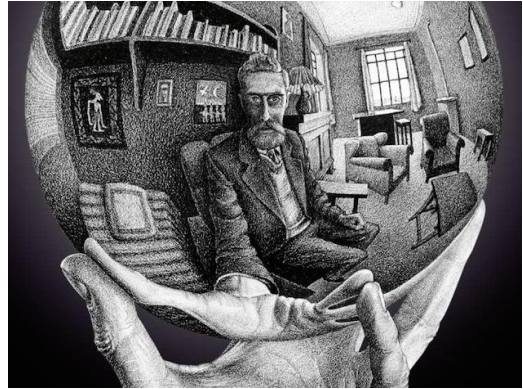
## Lessons learnt

Alice: "I give myself very good advice, but I very seldom follow it."

### Re-look at things!

- They may look differently.
- You may discover new things.

Never take anything for granted!



3

## What's an error?

Google different types of error in statistics

Images Videos Job sites News Books Finance All filters Tools

About 621,000,000 results (0.39 seconds)

From sources across the web

Error	Random error	Sampling error
Systematic errors	Gross errors	Total nonresponse error
Absolute error	Environmental	Instrumental
Observational	Regression error	Standard error

Show less Feedback

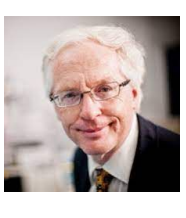

4

## What's an error?

The screenshot shows a Google search interface. The search bar contains the text "different types of error in physics searches". Below the search bar are several filters: Images, With examples, Videos, Pdf, Class 11, Places sites, News, Books, and Finance. Below the filters, it says "About 209,000,000 results (0.42 seconds)". The main text of the search results reads: "The three types of experimental error are systematic, random, and blunders. Systematic errors are errors of precision as all measurements will be off due to things such as miscalibration or background interference. Random errors occur due to happenstance, such as fluctuations in temperature or pH."

5

## Ask a physicist!

	<p><b>"Systematic"</b></p> <p>From an OPAT (One Parameter At a Time) error analysis when the response is not linear. Usually 3 points at <math>\pm\sigma</math> but can be more detailed.</p>	<p><b>"Statistical"</b></p> <p>From <math>\Delta \ln L = -\frac{1}{2}</math> errors when the likelihood is not parabolic</p>
	<p><b>BIAS</b></p> <p>Tendency to produce results that differ in a systematic manner from the 'true' value.</p>	<p><b>VARIANCE</b></p> <p>Scatter of values obtained from a data collection process.</p>

6

## What type of error?

$$x \pm \sigma$$

First option : (Wald type) **CONFIDENCE INTERVAL**

→ LIKs (statistical)

$$1 - \alpha = 0.68$$

confidence level

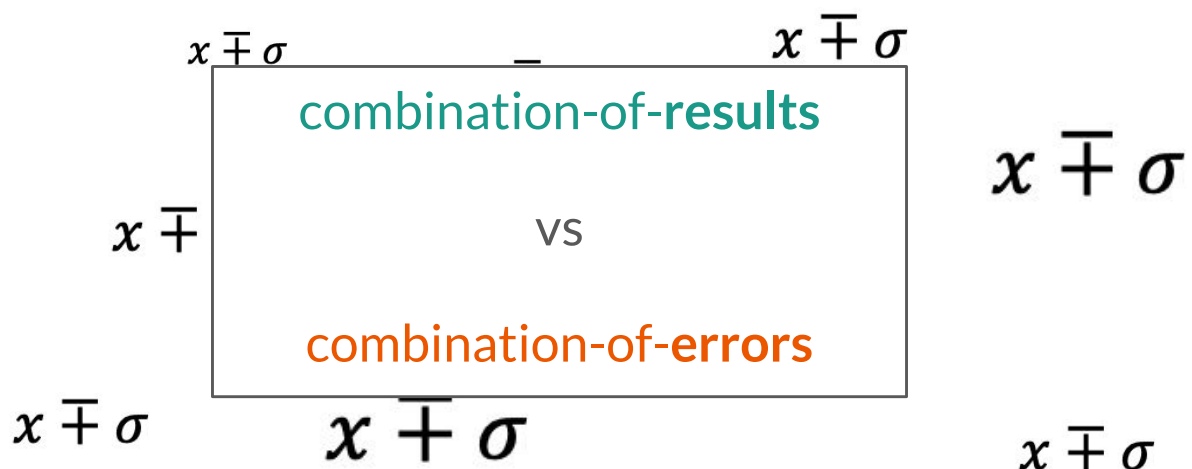
Second option : **NUMERICAL** (sample/population) **SUMMARIES**

→ PDFs (systematic)

probability

7

## What next?



8

## Combination of results (LIKs)

# Meta-analysis

Article [Talk](#)

From Wikipedia, the free encyclopedia

*For the process in historical linguistics known as metanalysis, see [Rebracketing](#).*

**Meta-analysis** is the statistical combination of the results of multiple studies addressing a similar research question. An important part of this method involves computing an effect size across all of the studies; this involves extracting effect sizes and variance measures from various studies. Meta-analyses are integral in

9

## Combination of errors (PDFs)



[Create account](#) [Log in](#)

# Convolution of probability distributions

[2 languages](#)

Article [Talk](#)

[Read](#) [Edit](#) [View history](#) [Tools](#)

From Wikipedia, the free encyclopedia

The **convolution/sum of probability distributions** arises in [probability theory](#) and [statistics](#) as the operation in terms of [probability distributions](#) that corresponds to the addition of [independent random variables](#) and, by extension, to forming [linear combinations of random variables](#). The operation here is a special case of [convolution](#) in the context of probability distributions.

10

Yes, but ...

$$x \pm \sigma \longrightarrow x \begin{matrix} +\sigma^+ \\ -\sigma^- \end{matrix}$$

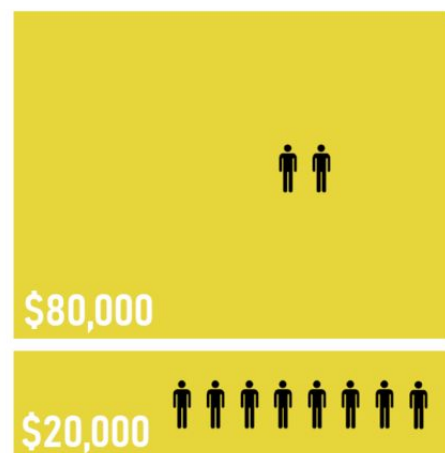
11

Not much around ... Why?

My fault !

“personalized” Pareto rule :

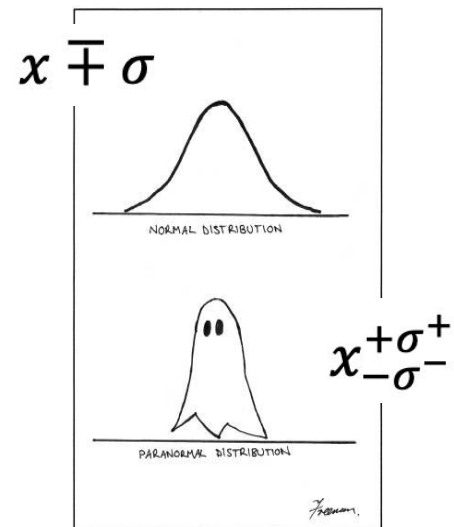
95 - 5



12

## Not much around... Why?

1. Great faith in **Central Limit Theorem**
  - Wald type confidence intervals
2. Or then, distribution given
  - LR (& Co.)
3. ...



13

## Asymmetric Errors

Roger Barlow  
The University of Huddersfield  
Huddersfield, UK  
[Roger.Barlow@cern.ch](mailto:Roger.Barlow@cern.ch)

Alessandra Rosalba Brazzale  
University of Padova  
Padova, Italy

Igor Volobouev  
Texas Tech University  
Lubbock, Texas, USA

and others

February 26, 2024

### Abstract

We present a procedure for handling asymmetric errors. Many results in particle physics are presented as values with different positive and negative errors, and there is no consistent procedure for handling them. We consider the difference between errors quoted using likelihood and using pdfs, and the difference between the rms spread of a measurement and the 68% central confidence region. We provide a full analysis of the possibilities, and software tools to enable their use.

14



## Combination of ...

### ERRORS

Identify asymmetric  
"surrogate" model



Convolution

### RESULTS

Identify non quadratic  
"surrogate" LIK



Combine LIKs

15

## Of course ...

- Combination of **results** with **PDFs**  
→ Weighted sum
- Combination of **errors** with **LIKs**  
→ Profiling
- ...

- ✓ File outline
- ✓ Introduction
  - ✓ Beyond the Gaussian
    - Combination of Errors
    - Combination of Results
- ✓ Pdf errors
  - > Modelling pdf errors
  - Combination of pdf errors
  - Combination of pdf results
- ✓ Likelihood Errors
  - > Modelling non-parabolic Likelihood...
  - Combination of Results using likelihood
  - Goodness of fit with likelihood
  - Combination of errors with likelihood
  - Example of combination of errors...
- > Some case studies
- > PDF or likelihood?
- Conclusions
- Acknowledgements
- Why the 'usual procedure' is wrong
- Implementation in C++
- Implementation in Python
- Implementation in R
- > Additional formulae

16



## “Surrogate” PDFs



skew Normal	log Normal	Edgeworth expansion
dimidiated Gaussian	distorted Gaussian	“railway” Gaussian
quantile variable width Gaussian	Johnson system	??

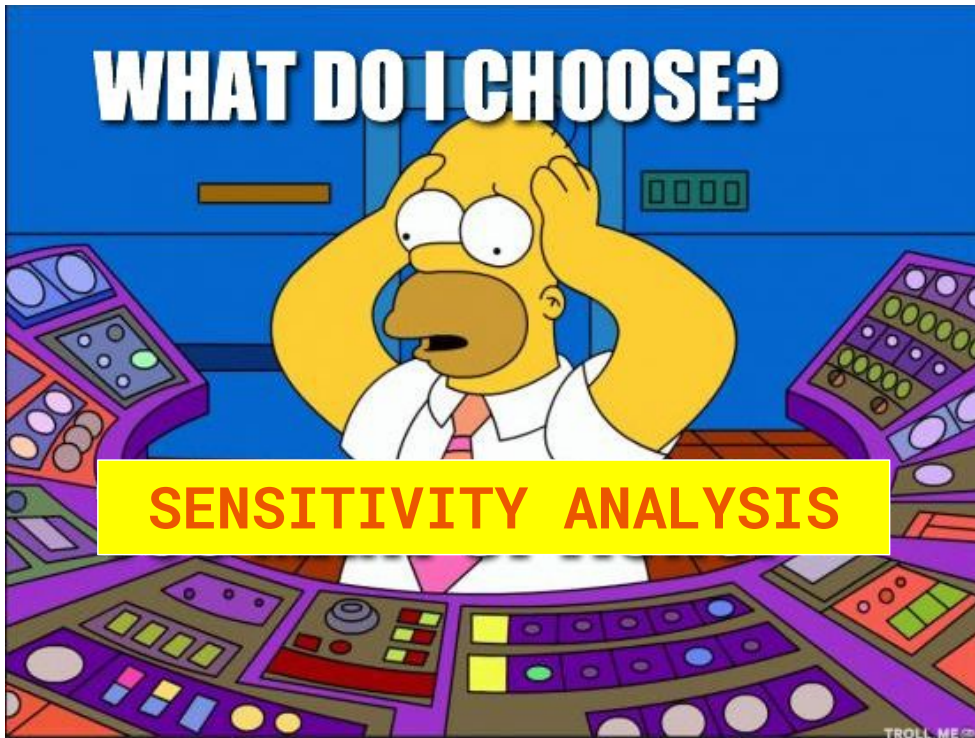
17

“It goes like this: / the fourth, the fifth, / the minor falls, and the major lifts» (L. Cohen)

## “Surrogate” LIKs

ln(PDF)	generalised Poisson	broken parabola	logarithmic
linear sigma	linear sigma in log space	double cubic sigma in log space	linear variance
PDG method	<b>cubic</b>	constrained <b>quartic</b>	molded <b>quartic</b>
simple double <b>quartic</b>	molded double <b>quartic</b>	matched <b>quintic</b>	simple double <b>quintic</b>
molded double <b>quintic</b>	<b>quintic</b> sigma in log space	interpolated <b>seventh order</b>	??

18



## Which LIK approximation?

- converge to quadratic LIK if asymmetry  $\rightarrow 0$
- interpretable
- "regular" model



hard boundary on / maximum "practicable" value for asymmetry

Submitted to the Statistical Science

### Likelihood Asymptotics in Nonregular Settings: A Review with Emphasis on the Likelihood Ratio

Alessandra R. Brazzale \* and Valentina Marni †

*Condition 1* The model functions defined by any two distinct values of  $\theta$  are distinct almost surely. Moreover, all components of  $\theta$  are identifiable.

*Condition 2* The support  $\mathcal{Y}$  of  $f(y; \theta)$  does not depend on  $\theta$ .

*Condition 3* The parameter space  $\Theta$  is a compact subset of  $\mathbb{R}^p$ , for a fixed positive integer  $p$ , and the true value  $\theta^0$  of  $\theta$  is an interior point of  $\Theta$ .

*Condition 4* The partial derivatives of the log-likelihood function  $l(\theta; y)$  with respect to  $\theta$  up to the order three exist in a neighbourhood of the true parameter value  $\theta^0$  almost surely. Furthermore, in such a neighbourhood,  $n^{-1}$  times the absolute value of the log-likelihood derivatives of order three are bounded above by a function of  $Y$  whose expectation is finite.

*Condition 5* The first two Bartlett identities hold, so

$$E[u(\theta; Y)] = 0, \quad i(\theta) = \text{Var}[u(\theta; Y)],$$

in addition to  $0 < \text{Var}[u(\theta; Y)] < \infty$ .

## Which PDF approximation?

closure under CONVOLUTION



### Stable distribution

[Article](#) [Talk](#)

From Wikipedia, the free encyclopedia

*Not to be confused with [Stationary distribution](#).*

In [probability theory](#), a [distribution](#) is said to be **stable** if a [linear combination](#) of two [independent random variables](#) with this distribution has the same distribution, up to [location](#) and [scale](#) parameters. A random variable is said to be **stable** if its distribution is stable. The stable



at least approximate...  
(work in progress)

... need suitable [distance metric](#)

21

## Ongoing and open questions

1. Beta testing
2. Model selection
3. Correlated errors?
4. Combining LIKs and PDFs
5. What if we want to be Bayesian?
6. ...

22

## The floor to you!

Questions?

