Asymmetric Uncertainties

Roger Barlow
The University of Huddersfield

with help from Alessandra Brazzale and Igor Volobouev

PHYSTAT informal review

28th February 2024

Asymmetric Errors

Many results in particle physics are presented with asymmetric errors. For instance current results on the Higgs width:

ıled at a p-value of 0.0003 (3.6 star
ɔn as $\Gamma_{\mathrm{H}} = 3.2^{+2.4}_{-1.7}$ MeV, in agr
ıddition, we set constraints on anc

CMS: Nature Physics **18** 1329

ıction has been observed in the H →
on-shell Higgs boson production the
$\Gamma_{H} = 4.5^{+3.3}_{-2.5}$ MeV

ATLAS (S Manzoni):
EPS-HEP conference 2023

- How do they arise?
- What do they mean?
- How should they be handled?

Some details...

Note 1

These questions are asked, and are to be answered, in frequentist (or,at least, agnostic) language.

Note 2

Distributions from transforming normally-distributed variables are not considered here as they're straightforward.

Note 3

We are working with slightly non-Gaussian distributions. As well as location and scale parameters, a third is needed to describe the asymmetry. (Cases needing more than 3 should not be ruled out).

How do they arise?

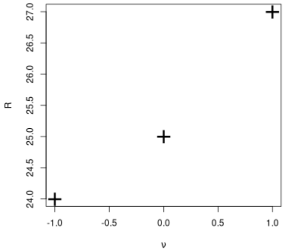
Through two sources

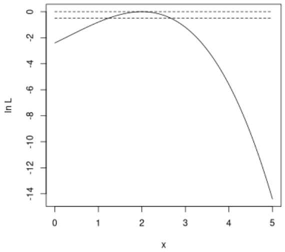
"Systematic"

From an OPAT (One Parameter At a Time) error analysis when the response is not linear. Usually 3 points at nominal and $\pm\sigma$ but can be more detailed.

"Statistical"

From $\Delta \ln L = -\frac{1}{2}$ errors when the likelihood is not parabolic

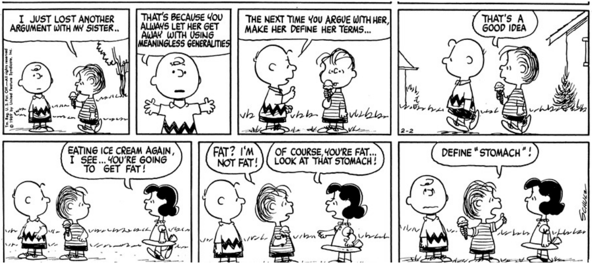




We need to define our terms very carefully
Even if they seem familiar

"How should I handle asymmetric errors?"

1. What do you mean by 'error'? $\sigma = \sqrt{\overline{x^2} - \overline{x}^2}$ or 68% central confidence region?
2. What is asymmetric? The pdf or the likelihood?
3. What do you mean by 'handle'? Combining errors or combining results?



PDFs and likelihoods

The Gaussian (Normal)

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

can be viewed (for a given σ) as a pdf $P(x, \mu)$ or as a likelihood $L(\mu, x)$.

Both are symmetric.

Asymmetric errors will involve a pdf which is not quite Gaussian or a ln L which is not quite parabolic. (Or both.)

Question 1

Are you working with an asymmetric pdf or an asymmetric likelihood?

What is an error?

What physicists call an "error" is not the statistician's ϵ *but the uncertainty, or 'probable error'.*

For the Gaussian, σ gives

1. The square root of the variance $\sigma = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$
2. The 68% central confidence region: $\int_{\mu-\sigma}^{\mu+\sigma} N(x, \mu, \sigma)\, dx = 0.68$ (other confidence regions are available)

For a non-Gaussian these definitions are no longer equivalent.

(2) is arguably more meaningful

(1) must be used if errors are to be added in quadrature. Variances add, even for non-Gaussian distributions.

Question 2

Are you working with σ as the rms spread or as a 68% central confidence region?

Combination of errors and combination of results

Contrast:

1: To measure the length of a rod you measure the positions of both ends: $x_1 \pm \sigma$ and $x_2 \pm \sigma$. The length is $L = |x_2 - x_1| \pm \sqrt{2}\sigma$

This comes from the famous combination-of-errors formula

$$\sigma_f^2 = \left(\frac{\partial f}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \sigma_y^2 + 2\rho \left(\frac{\partial f}{\partial x}\right) \left(\frac{\partial f}{\partial y}\right) \sigma_x \sigma_y$$

2: You measure a position twice, independently: $x_1 \pm \sigma$ and $x_2 \pm \sigma$. Combining the results, the position is $X = \frac{x_2 + x_1}{2} \pm \frac{\sigma}{\sqrt{2}}$

Combination of Results ("meta analysis") is a major activity of the PDG and HFLAV. Goodness-of-fit is vital.

Combination of Errors is a major activity as an experiment has many sources of uncertainty. Variances add and the CLT helps. Goodness-of-fit is meaningless.

Question 3

Are you using σ for combination-of-errors or combination-of-results?

Your 3 answers are linked

Likelihoods

If you are working with likelihoods then your σ quantities are the 68% CL bounds, as you can't get expectation values from likelihoods. You are probably combining results ('meta-analysis'), though combining errors is possible (profile likelihoods).

Pdfs

If you are working with pdfs you probably want to know about rms spreads, though your σ quantities may be given as 68% CL limits. You are probably combining errors, though combination of results is possible and can be viewed as a special case of combining errors, weighting to minimise the variance

All set!

Once you've answered the 3 (linked) questions, what next?

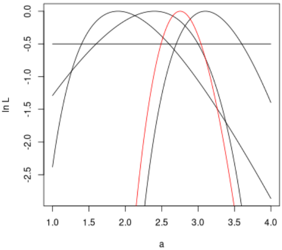
Dealing with Likelihoods

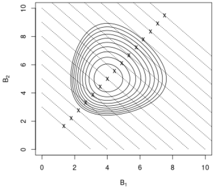
Given a set of results $\{a_i{}^{+\sigma_i^+}_{-\sigma_i^-}\}$ with asymmetric errors:

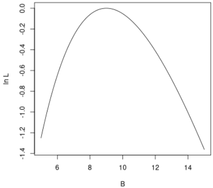
1. Choose near-parabolic 3 parameter model
2. Fit all results using this model

3a Combine results: Use total $\ln L$ to get best estimate, $\Delta \ln L = -\frac{1}{2}$ errors, and goodness of fit

3b Combine Errors: Use parameterised $\ln L$ functions to find profile likelihood, and extract $\Delta \ln L = -\frac{1}{2}$ errors





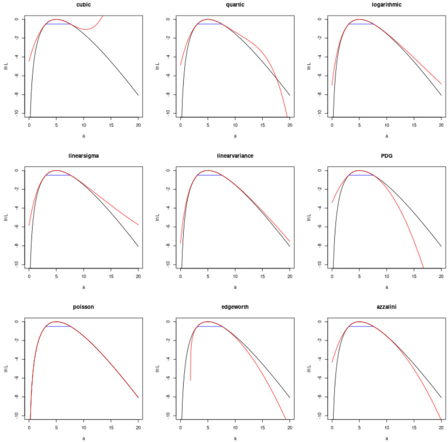


Many models

"In conics I can floor peculiarities parabolous" – W. S. Gilbert: The Major General's song

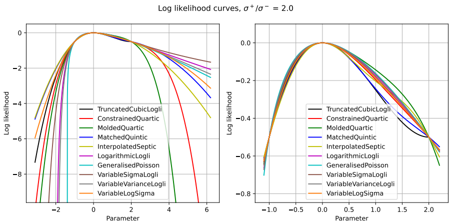
Many possible models

Software must translate between $\hat{a}, \sigma_+, \sigma_-$ and specific model parameters



Shows 9 models (in red, with true form in black)) approximating the Poisson likelihood for $n = 5$ using only the $\Delta \ln L = -\frac{1}{2}$ errors $5^{+2.58}_{-1.92}$

Many, many models...



For special consideration:

- Linear sigma: $\ln L(a) = -\frac{1}{2}\left(\frac{a-\hat{a}}{S+S'(a-\hat{a})}\right)^2$
- Linear Variance: $\ln L(a) = -\frac{1}{2}\frac{(a-\hat{a})^2}{V+V'(a-\hat{a})}$
- PDG: Like Linear sigma for $[\hat{a}-\sigma^-, \hat{a}+\sigma^+]$, but uses σ^+ above and σ^- below
- Molded double quintic
- Double cubic sigma in the log space

Dealing with pdfs

Given a set of results with asymmetric errors:

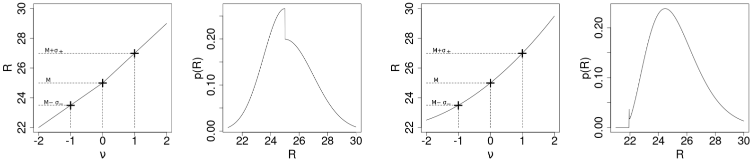
1. Choose near-Gaussian 3 parameter model
2. Fit all results using this model and for each find first 3 cumulants μ, V, γ

3a Combine Errors: add to get total cumulants. Then find parameters which give this total

3b Combine Results: Take mean result, weighted proprtional to $1/V_i$, get total cumulants, and extract parameters

Pdf models

Many possible models, e.g.
Dimidiated Gaussian: $P(R)$ given by two half-Gaussians
Distorted Gaussian: $P(R)$ given by unit Gaussian in ν with
$R = R_0 + \alpha\nu + \beta\nu^2$ going through 3 OPAT points



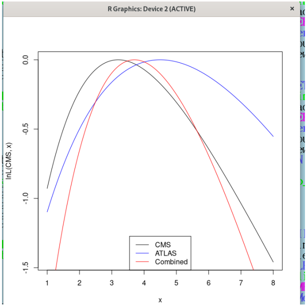
Software tools now have to translate between R_0, σ_+, σ_- and model parameters (e.g. R_0, α, β) and moments μ, V, γ

Also: Edgeworth, Azzalini skew normal, Johnson functions, railway Gaussian, lognormal, Quantile Variable Width...

Example

Let's combine the two Higgs width measurements

```
>
> CMS <- getlnLpars(c(val=3.2,sp=2.4,sm=1.7),"linearsigma")
> ATLAS <- getlnLpars(c(val=4.5,sp=3.3,sm=2.5),"linearsigma")
> combination <- combinelnLresults(list(ATLAS,CMS))
>
> print(combination)
      val        sp        sm     chisq       ndf         s    sprime
3.7000593 1.9095244 1.5092210 0.1436208 1.0000000 1.6859368 0.1170907
attr(,"class")
[1] "linearsigma"
>
> x <- seq(1,8,.01)
> plot(x,lnL(CMS,x),type='l')
> lines(x,lnL(ATLAS,x),col='blue')
> lines(x,lnL(combination,x),col='red')
>
> legend('bottom',col=c('black','blue','red'),legend=c('CMS','ATLAS','Combined'),lty=1)
> █
```



This just uses the quoted errors. If both experiments make their full likelihood function available to the PDG (and they presumably will) then they can do a better job

Another Example

Real analysis but in progress so anonymous

Dalitz plot fit to decays of the Λ_b^0 looking at production of the $\Lambda(1800)$ (and 21 other contributions)

Systematic uncertainties from

Source	σ^+	σ^-
real numbers fix res	+0.059	-0.029
amp model	+0.001	-0.008
res	+0.008	-0.015
finite acc	+0.003	-0.003
acc model	+0.001	-0.001
kin	+0.001	-0.001
sWt pg	+0.006	0.0
massfit comb	+0.004	0.0

plus 6 other sources that are evaluated as zero

Combined error using the dimidiated model: $\sigma^+ = 0.05965, \sigma^- = 0.03294$

Combined error using the distorted model: $\sigma^+ = 0.06098, \sigma^- = 0.03485$

Yet Another Example (sanity check)

Suppose a counting experiment sees 5 events in an hour. The result is quoted (using $\Delta \ln L = -\frac{1}{2}$ errors) as $5.000^{+2.581}_{-1.916}$.
This continues for another hour and again 5 events are seen.
The total gives a result $10.000^{+3.504}_{-2.838}$ and with the knowledge we have of the way the experiment has been done, we can estimate the number of events per hour by dividing this by 2 to get $5.000^{+1.752}_{-1.419}$.

But if this knowledge is suppressed we are just presented with two estimates $5.000^{+2.581}_{-1.916}$ to be combined,

With the linear variance method, the result is $5.000^{+1.747}_{-1.415}$. This is an excellent match to the ideal value, with the errors differing only in the 4th significant figure.

Using linear sigma we would get $5.000^{+1.737}_{-1.408}$ which is also very good.

Friendly software

Using C++ with python interface

In preparation (Igor Volobouev)

Using R: Package `AsymmetricErrors`

Install (once) from
`https://barlow.web.cern.ch/programs/AsymmetricErrors.tar.gz`
Thereafter load (when needed) with `library(AsymmetricErrors)`
Only 9 functions - help files provided:
`getPdfpars`, `getlnLpars`, `Pdf`, `lnL`, `combinePdferrors`, `combinelnLerrors`, `combinePdfresults`, `combinelnLresults`, `getflipPdfpars`

Code needs tidying, but ready for beta testing

Conclusions

1. This is a messy area with no 'right' answers (though plenty of 'wrong' ones). Avoid if possible. $12.34^{+0.46}_{-0.44} \to 12.34 \pm 0.45$
2. If not possible, need to be very clear about what you are doing
3. Choose model(s) and combine to get result or error. Then try another model as a consistency check.
4. Full details in a preprint to be released shortly
5. More input from the community (you!) would be really useful

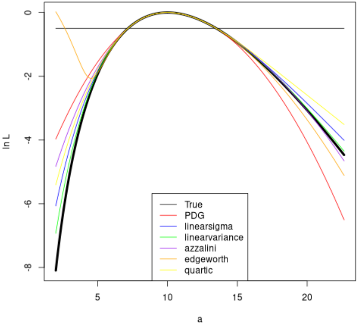
Backup Slides

The PDG method

When combining results the PDG uses a *lnL* parameterisation which is like linear-sigma in the central $[\hat{a} - \sigma^-, \hat{a} + \sigma^+]$ region, and parabolic with $\sigma = \sigma^-$ or $\sigma = \sigma^+$ for larger negative or positive values.

Thick black line is Poisson lnL distribution for $r = 10$. This gives result $10^{+3.50}_{-2.84}$

This is fitted with various models and the resulting *lnL* plotted.
All do well in the central region, varying success outside, but red PDG curve is clearly worst



For the 5+5=10 example, PDG gives same (OK) result as linear-sigma
But for 7+3=10 for which true result is still $5.000^{+1.752}_{-1.415}$, we get linear-variance $5.009^{+1.793}_{-1.456}$ linear-sigma $5.038^{+1.937}_{-1.529}$ PDG $5.009^{+1.334}_{-1.777}$

Why adding σ^+ and σ^- separately is wrong

What goes wrong

Suppose you combine n measurements, all with the same σ^+ and σ^-
This will give an error ${}^{+\sqrt{N}\sigma^+}_{-\sqrt{N}\sigma^-}$ which has the same shape as the original.
It does not become symmetric (Gaussian) at large N and breaks the Central Limit Theorem

Why it does wrong

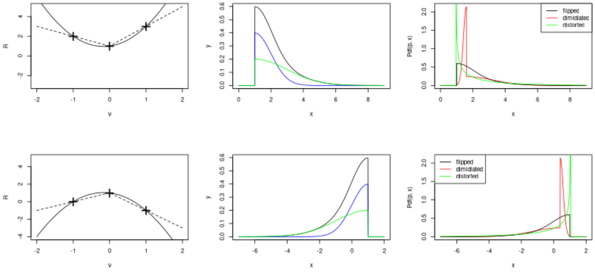
If two errors combine, there is a 25% chance that both fluctuate upwards, described by ${\sigma^1_+}^2 + {\sigma^2_+}^2$, and similarly 25% that both go downwards. But there is 50% chance that one goes up and one goes down, which reduces the asymmetry, and the method neglects this.

Combine ${}^{+2}_{-1}$ and ${}^{+2}_{-1}$ Wrong method: ${}^{+2.83}_{-1.41}$

Dimidiated model ${}^{+2.64}_{-1.65}$ Distorted model ${}^{+2.73}_{-1.76}$ Railway model ${}^{+2.72}_{-1.76}$

Flipped distributions

OPAT treatment where both differences have the same sign.
Probably due to numerical fluctuations in unimportant uncertainties and can be neglected. But maybe not. And need a consistent procedure



Distorted model copes quite naturally.
With dimidiated the central result is an absolute upper or lower limit, not the median. Suggest using a standard dimidiated model with the same moments.