Testing fundamental theories with global fits

Andrew Fowlie

26 April 2024

Suzhou University

Overview

1. Fundamental theories & global fits
2. Foundations
3. Pitfalls
4. Results from `GAMBIT`

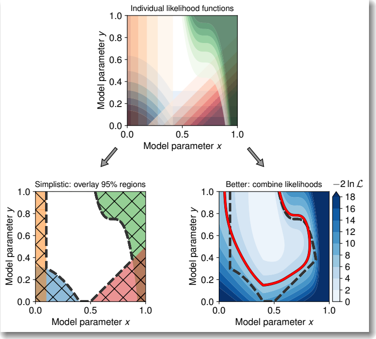
Fundamental theories & global fits

Fundamental theories — big models

- There are many motivations for physics beyond the Standard Model, e.g.,
  - Naturalness
  - Dark Matter
  - Gravity
  - Unification
  - Muon $g-2$
- These models are not simple, as they typically contain many unknown parameters and scales
- E.g., even the minimal supersymmetric model adds hundreds of new soft-breaking parameters

Big data

- We are searching everywhere for new physics
- `Inspire` lists more than 800 "search" papers since 2010
- Searches at LHC, astrophysics, cosmology, etc.
- Each one with an independent statistical analysis
- Enormously powerful data. We must make the most of it



Putting it all together

To make the most this data, we must

- combine searches and measurements coherently
- model statistical, systematic and theory uncertainties
- explore an entire model — the whole parameter space
- test and compare models

These activities are global fits

What do I mean by parameter spaces?

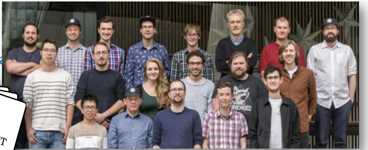
- In science we build models to describe data
- These models usually come with several unknown parameters — e.g., the masses and couplings of new particles
- Models make a range of predictions, depending on the values that the parameters take
- We must consider all these possibilities in our global fit

GAMBIT

GAMBIT: The Global And Modular BSM Inference Tool

gambit.hepforge.org github.com/GambitBSM EPJC 77 (2017) 784 arXiv:1705.07908

- Extensive model database, beyond SUSY
- Fast definition of new datasets, theories
- Extensive observable/data libraries
- Plug&play scanning/physics/likelihood packa
- Various statistical options (frequentist /Bayesian)
- Fast LHC likelihood calculator
- Massively parallel
- Fully open-source



Members of: ATLAS, Belle-II, CLiC, CMS, CTA, Fermi-LAT, DARWIN, IceCube, LHCb, SHiP, XENON

Authors of: BubbleProfiler, Capt'n General, Contur, DarkAges, DarkSUSY, DDCalc, DirectDM, Diver, EasyScanHEP, ExoCLASS, FlexibleSUSY, gamLike, GM2Calc, HEPLike, IsaTools, MARTY, nuLike, PhaseTracer, PolyChord, Rivet, SOFTSUSY, Superlso, SUSY-AI, xsec, Vevacious, WIMPSim

Recent collaborators: V Ananyev, P Athron, N Avis-Kozar, C Balázs, A Beniwal, LL Braseth, T Bringmann, A Buckley, J Butterworth, JE Camargo-Molina, C Chang, J Cornell, M Danninger, A Fowlie, T Gonzalo, W Handley, S Hoof, A Jueid, F Kahlhoefer, A Kvellestad, M Lecroq, C Lin, M Lucente, FN Mahmoudi, DJE Marsh, G Martinez, H Pacey, MT Prim, T Procter, F Rajec, A Raklev, R Ruiz, A Scaffidi, P Scott, W Shorrock, C Sierra, P Stöcker, W Su, J Van den Abeele, A Vincent, M White, A Woodcock, Y Zhang ++

70+ participants in many experiments and numerous major theory codes

What is GAMBIT?

Software framework

- Combines collider, flavor, astrophysics, comoslogical data
- Joint analysis of dark matter, neutrinos & BSM physics
- `MPI` + `OpenMP` parallelisation (record of 115,000 CPUs)
- Combines libraries & codes written in: `C++`, `Fortran`, `Python`, `Mathematica`...
- Highly modular `Bits`
- Often with several alternatives

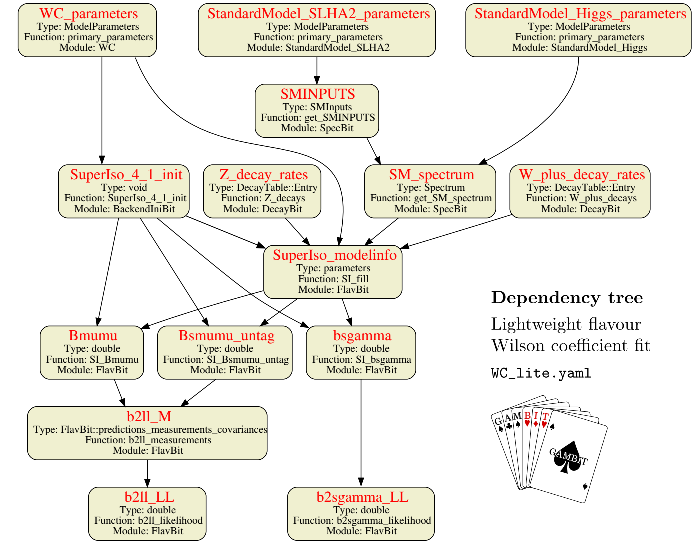
Community

- Particle physicists, cosmologists & statisticians (>80 members)
- Generates interdisciplinary expertise & inspires new techniques
- Develop and maintain open source software
- Access to massive computing resources — 40 M CPU hours / year
- In-person meetings every 9 months

The *Bits

- ColliderBit – Collider searches for new particles at LHC and LEP
- DarkBit – Direct and indirect searches for dark matter
- FlavBit – Measurements of flavour physics observables
- SpecBit – Precision calcuatlions of mass spectrum and RG running
- DecayBit – Precision calculations of particle decay widths and branching ratios
- PrecisionBit – Tests of precision observables and rare processes
- NeutrinoBit – Neutrino physics
- CosmoBit – Cosmology, e.g., Planck

GAMBIT structure



Dependency tree

Lightweight flavour Wilson coefficient fit

`WC_lite.yaml`

Foundations

Let the data speak for itself

Gould 1981

“ inanimate data can never speak for themselves, and we always bring to bear some conceptual framework, either intuitive and ill-formed, or tightly-formed and structured, to the task of investigation, analysis and interpretation ”

Tukey et al. 1977

“ No body of data tells us all we need to know about its own analysis ”

Jaynes 2003

“ The data cannot speak for themselves; and they never have, in any real problem of inference ”

Testing and estimation

Roughly speaking, statistical tasks separate into

1. Model testing or comparison
2. Estimating or inferring the model's parameters

I will focus on first. In my opinion, first we should establish whether a phenomena exists, and then infer its parameters or properties.

Testing

Jeffreys and Fisher agree!

Jeffreys 1939

“[I]n what circumstances do observations support a change of the form of the law itself? This question is really logically prior to the estimation of the parameters, since the estimation problem presupposes that the parameters are relevant”

Fisher 1925

“It is a useful preliminary before making a statistical estimate …to test if there is anything to justify estimation at all”

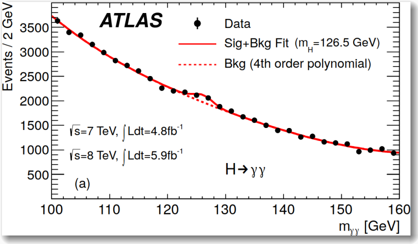
Discoveries!

Classic example. Higgs discovery in 2012.

How do we judge when the data indicates the presence of a new particle or phenomena?

Discoveries!

Di-photon spectrum contains a resonance (Aad et al. 2012).



Discovery was announced based on a particular choice of statistical methodology.

Methodology

We need a statistical methodology to judge evidence for a discovery. In the time available, let's consider

1. Frequentist; see e.g., Lyons 1989; Cowan 1998; James 2006; Behnke et al. 2013. Two schools
    - Error control
    - Evidential
2. Bayesian; see e.g., D'Agostini 2003; Gregory 2005; Sivia and Skilling 2006; Trotta 2008; Linden, Dose, and Toussaint 2014; Bailer-Jones 2017

Likelihood

Methods typically require at least the likelihood (see e.g., Cousins 2020)

$$\mathcal{L}(\boldsymbol{\Theta}) = p(D \mid M, \boldsymbol{\Theta})$$

This tells us the probability (density) of the observed data, D, given a particular model, M, and choice of parameters.

This is a function of the model's parameters, $\boldsymbol{\Theta}$, for fixed, observed data.

P-values

P-value (Wasserstein and Lazar 2016)

The p-value, p, is the probability of observing data as or more extreme than that observed, given the null hypothesis, H_0, i.e.,

$$p = P(\lambda \geq \lambda_{\text{Observed}} \mid H_0)$$

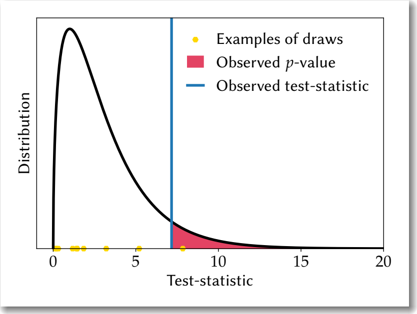
where λ is a test-statistic that summarises the data and defines extremeness, and H_0 specifies the distribution of λ

See Demortier 2008 for discussion about composite null hypotheses that don't uniquely specify the distribution of λ.

Test-statistic often based on (profiled) likelihood ratio (Neyman and Pearson 1933)

P-values

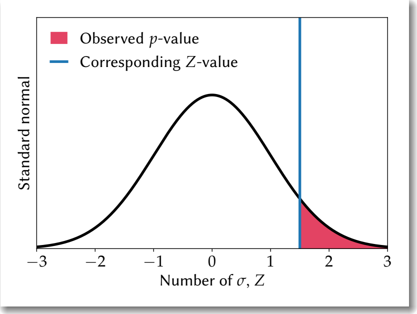
Thus p is a tail probability.



Thus p is uniformly distributed under H_0 (or dominated by uniform in discrete settings or composite null)

Z-values

In particle physics, it's common to translate p-values into Z-values. 5σ corresponds to about $p = 10^{-7}$. This is just a convention



through the equation

$$Z = \Phi^{-1}(1 - p)$$

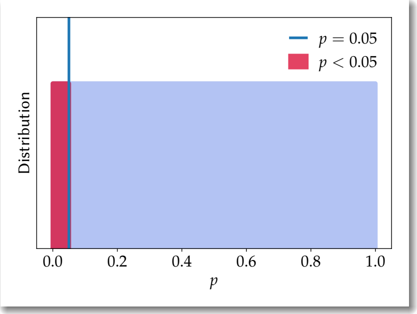
Interpreting p-values

P-values are popular in particle physics and elsewhere. Two possibly contradictory interpretations (Hubbard and Bayarri 2003):

- P is a measure of evidence against H_0 (Fisher 1925): small $p \Rightarrow$ H_0 implausible. See e.g., Hubbard and Lindsay 2008; Schervish 1996; Berger and Sellke 1987; Senn 2001; Murtaugh 2014
- P is a means to control error rate (Neyman and Pearson 1933): if we reject null when p-value ≤ 0.05, for example, becomes error theoretic approach with type 1 error rate $\alpha = 0.05$

Controlling type-1 error rate

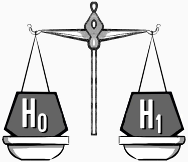
The p-value enables us to control type-1 error rate because it is uniformly distributed under the null



Placing a threshold $p < \alpha$ controls the type-one error rate to be α

Example from high-energy physics

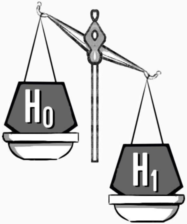
Original artwork Viktor Beekman and concepts Eric-Jan Wagenmakers



In high-energy physics, we want to discover new phenomena and new particles. Perform null hypothesis test:

- H_0 — Standard Model (SM) backgrounds only
- H_1 — SM + new physics, e.g. Higgs boson or supersymmetric particles

Example from high-energy physics

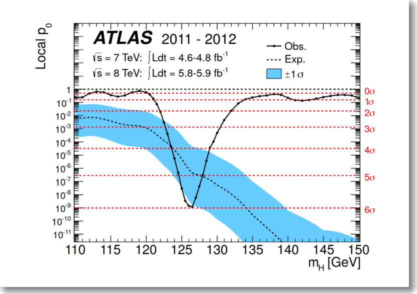


For a discovery we conventionally require a tiny global *p*-value of

$$p \lesssim 10^{-7} \, (5\sigma)$$

i.e., $\alpha \simeq 10^{-7}$ (Lyons 2013). Dual interpretation: threshold in evidence — extraordinary claims require extraordinary evidence — and imposes a 10^{-7} type-1 error rate.

Example from high-energy physics



Discovery of Higgs boson announced by ATLAS and CMS once significance greater than 5σ.

Some scares — e.g., 2015 diphoton excess (Strumia 2016) — but so far 5σ criterion prevented false discoveries (though think about flavor anomalies).

Misconceptions

Misconceptions galore by public and scientists, see e.g., Goodman 2008; Greenland et al. 2016

1. P is not probability of null hypothesis
2. P is not the probability that the data were produced by chance alone
3. P is not an error rate

Wagenmakers and al 2017

> *"The fact that academics don't know what p means is a symptom of the fact that p doesn't tell anything worth knowing"*

Though see Murtaugh 2014; Lakens et al. 2018; Cousins 2018; Lakens 2021; Mayo 2018.

Bayes factors

Forget long-run errors rates and data we don't have. Compute the change in plausibility of models in light of the data we have

- With this and priors for the models, we could compute the posterior plausibility of each model
- If you like, you can compute the probability that you are making an error in the case at hand (cf. long-run error rates that are independent of the observed data)
- We just apply probability theory to the problem (Jeffreys 1939). Simple in theory; in practice there are difficulties.

Bayes factors

The Bayes factor (Kass and Raftery 1995) relates the relative plausibility of two models after data to their relative plausibility before data;

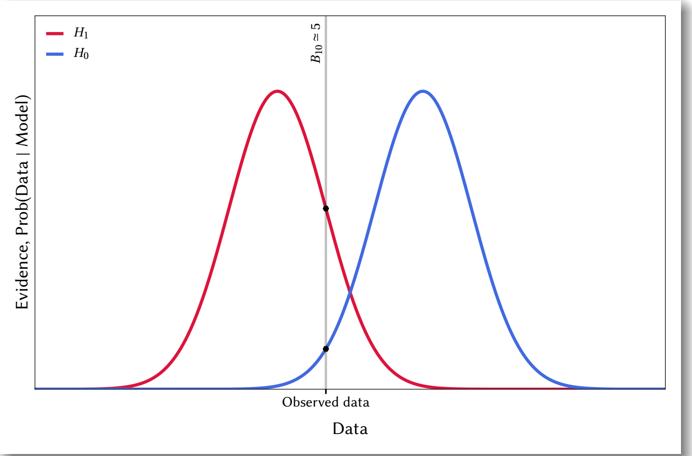
$$\text{Posterior odds} = \text{Bayes factor} \times \text{Prior odds}$$

where

$$\text{Bayes factor} = \frac{p(\text{Observed data} \mid \text{Model } a)}{p(\text{Observed data} \mid \text{Model } b)}$$

By applying laws of probability, we see that models should be compared by nothing other than their ability to predict the observed data.

Bayes factors



Bayesian evidence

The factors in the ratio are Bayesian evidences

$$\mathcal{Z} \equiv p(D \mid M) = \int_{\Omega_{\boldsymbol{\Theta}}} \mathcal{L}(\boldsymbol{\Theta})\, \pi(\boldsymbol{\Theta})\, \mathrm{d}\boldsymbol{\Theta},$$

where D is the observed data, $\mathcal{L}(\boldsymbol{\Theta}) = p(D \mid \boldsymbol{\Theta}, M)$ is the likelihood and $\pi(\boldsymbol{\Theta}) = P(\boldsymbol{\Theta} \mid M)$ is our prior, and $\boldsymbol{\Theta}$ are the model's parameters.

The prior describes what we knew about the parameters before seeing the data

The evidence is the likelihood averaged over the prior — the averaging penalises fine-tuned models

Sensitivity to priors

Evidences are the likelihoods averaged over priors.

Many consider the resulting dependence of the Bayes factor on the priors to be a major and perhaps fatal problem; see e.g., Berger and Pericchi 2001; Cousins 2008

- No priors, no predictions. I need to compare your model's predictions with data. If you don't tell the plausible parameters, how am I to know what it predicts?
- Sensitive to arbitrary choices. If the inference changes dramatically within a class of reasonable priors, we can't draw reliable conclusions.

Sensitivity to priors

Evidences are the likelihoods averaged over priors.

Many consider the resulting dependence of the Bayes factor on the priors to be a major and perhaps fatal problem; see e.g., Berger and Pericchi 2001; Cousins 2008

Paraphrasing Hill 1975

> *“ the lack of a concrete theory for choosing priors no more implies that one should not use Bayesian statistics than does the lack of a theory that tells us the right price to pay for groceries implies we should not use money ”*

Subjective & Objective

There are different approaches to constructing priors, leading to different flavors of Bayesian inference

Subjective

Priors reflect state of knowledge and could be constructed by e.g., consulting experts (see e.g., Goldstein 2006; Mikkola et al. 2021)

Dictated by state of knowledge

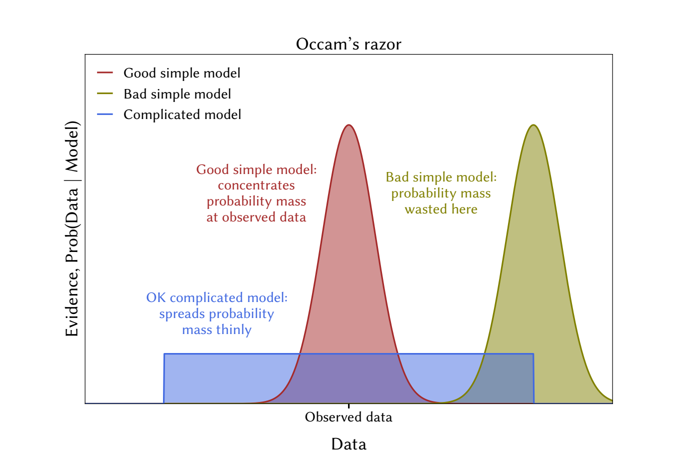
Priors could be dictated by e.g., a symmetry (Jaynes 1968)

Formal rules for selecting priors

Construct priors that e.g., maximise what we expect to learn about a model's parameters (Kass and Wasserman 1996; Consonni et al. 2018)

Occam's razor

Evidence automatic Occam razor (MacKay 1992; Jefferys and Berger 1992)



Likelihood principle

Originated by considering stopping rules (Barnard 1949). Proven by Birnbaum 1962

Berger and Wolpert 1988

> *“all evidence, which is obtained from an experiment, about an unknown quantity θ, is contained in the likelihood function of θ for the given data $[\mathcal{L}(\theta)]$”*

- Forbids evidential interpretations of frequentist statistics and *p*-values — since they depend on considering data other than that observed
- Implicitly obeyed by Bayesian statistics (though violated by reference priors)

Intuition for likelihood principle

Pratt 1962

> *"An engineer draws a random sample of electron tubes and measures the plate voltage under certain conditions with a very accurate volt-meter …*
>
> *A statistician examines the measurements, which look normally distributed and vary from 75 to 99 volts with a mean of 87 and a standard deviation of 4 …"*

Intuition for likelihood principle

Pratt 1962

“Later he visits the engineer’s laboratory, and notices that the volt meter used reads only as far as 100, so the population appears to be “censored.” This necessitates a new analysis, if the statistician is orthodox.”

Intuition for likelihood principle

Pratt 1962

> *“ However, the engineer says he has another meter, equally accurate and reading to 1000 volts, which he would have used if any voltage had been over 100. This is a relief to the orthodox statistician, because it means the population was effectively uncensored after all. ”*

Intuition for likelihood principle

Pratt 1962

> *“ But the next day the engineer telephones and says: “I just discovered my high-range volt-meter was not working the day I did the experiment you analyzed for me.” The statistician ascertains that the engineer would not have held up the experiment until the meter was fixed, and informs him that a new analysis will be required. The engineer is astounded. ”*

Intuition for likelihood principle

Pratt 1962

“ … “Next you’ll be asking me about my oscilloscope.” ”

Status of likelihood principle

At the time, considered profound by some though not universally accepted

Savage 1962

> *“ Without any intent to speak with exaggeration or rhetorically, it seems to me that this is really a historic occasion. ”*

Over time limited practical impact: ignored by Bayesian because it’s automatically satisfied; ignored by frequentists because it’s automatically violated.

Likelihood land

Key part of Bayesian analysis

$$\text{Likelihood} \times \text{Prior density} = \text{Evidence} \times \text{Posterior density}$$

in which we marginalise over unknown parameters by multi-dimensional integration over the likelihood function.

Frequentist analysis, on the other hand, usually involves finding the best-fit parameters by finding the maximum of the likelihood function.

Pitfalls

Pitfall — the likelihood is not enough?

- Frequentist analysis violates likelihood principle (Berger and Wolpert 1988) and in fact requires whole sampling distribution

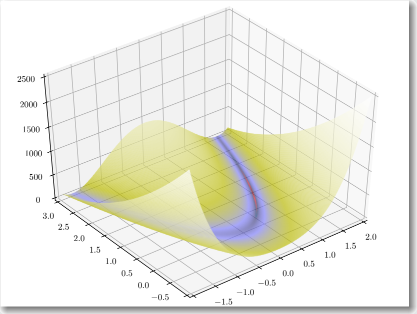
$$p(D \mid M, \boldsymbol{\Theta})$$

  as a distribution in the data, D.
- Often we make asymptotic approximations (Wilks 1938; Chernoff 1954) for the sampling distribution that only require the likelihood, e.g., $2 \times$ log likelihood ratio follows a χ^2 distribution. See e.g. Cowan et al. 2011.
- What if those assumptions don't apply? See e.g., Algeri et al. 2020. We need public sampling distributions.

Common pitfalls

Shape

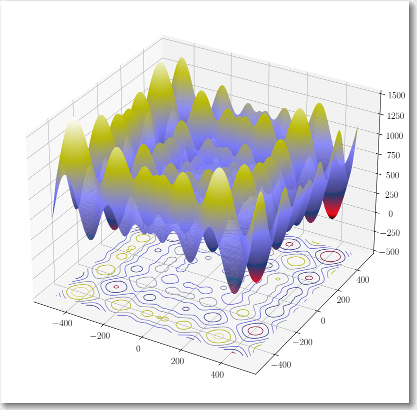
Likelihood could contain non-convex features, e.g., the classic Rosenbrock banana shape. Challenging for traditional optimisation and MCMC algorithms



Common pitfalls

Multi-modal

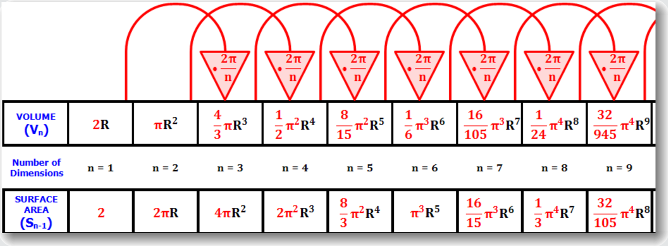
Likelihood could contain several distinct modes — how to find them all and the best one? (figure from Balázs et al. 2021)



Common pitfalls

Curse of dimensionality

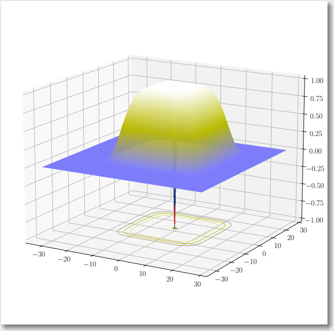
Performance of numerical algorithms deteriorate exponentially with dimension



Common pitfalls

Compression

If the best-fit is a flagpole in the Atlantic ocean, it could easily be missed (Balázs et al. 2021)

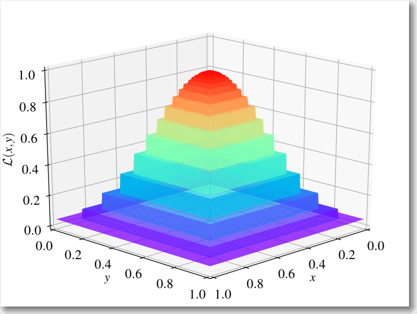


If sampling from the prior or if estimating the compression is important, convergence could be slow (**skilling2006**)

Common pitfalls

Plateaus

Likelihood functions involving plateaus may pose unexpected problems, see e.g., Schittenhelm and Wacker 2020; Fowlie, Handley, and Su 2021



ScannerBit

A suite of statistical computation and optimization tools to overcome these challenges

1. Scanning algorithms
   - Local — `MINUIT2` — and global optimization — `particle swarm`, `differential evolution`
   - Mark Chain Monte Carlo — Metropolis-Hastings and affine-invariance
   - Nested sampling — `PolyChord`, `MultiNest`
2. `Python` Scanners
   - Allows easy incorporation of new scanning algorithms into GAMBIT — e.g., `emcee`
   - Modern machine learning techniques
   - Embarassingly easy to incorporate

Pitfall — incomplete information

What if we don't know the form of the likelihood? What if an experiment only reports an asymmetric error (see e.g., Barlow 2004 for further discussion)

$$x = 45^{+5}_{-8}$$

or a bound

$$x < 62 \text{ @ } 95\% \text{ CL}$$

or an interval etc. Actually, what if they only report a symmetric error

$$x = 45 \pm 10$$

It's a Gaussian, $\mathcal{N}(45, 10^2)$, right? or is it?

Experimentalists know more, and they should tell us

Pitfall — intractable likelihood

What if we just can't work out the likelihood — it's intractable?

Fine — if we can draw pseudo data from the sampling distribution we can use likelihood-free inference methods, e.g., ABC (Diggle and Gratton 1984).

Public sampling distributions?

Pitfall — noisy estimate of likelihood

Consider case of LHC likelihoods. The likelihood might be a Poisson involving the expected number of events.

The expected number of events, though, can often only be computed through forward simulations of the whole experiment, involving MC simulations in e.g., `Pythia` and detector simulations

So we end up with a noisy estimate of the likelihood, $\hat{\mathcal{L}}$.

Pitfall — noisy estimate of likelihood

Careful! — if we run an optimiser, it will favour points with upward fluctuations in estimate of likelihood. Fitting noise.

If we run MCMC, we might get correct inferences if we have an unbiased estimator of the likelihood, $\langle\hat{\mathcal{L}}\rangle = \mathcal{L}$ — this is pseudo-marginal MCMC (Andrieu and Roberts 2009).

However, common MLE estimator of $\mathcal{L}$ is biased.

Pitfall — posterior of null versus p

Of course,

$$\text{Posterior of null} \neq p\text{-value}$$

However, well-known that typically

$$\text{Posterior of null} \gg p\text{-value}$$

for broad classes of priors. P typically overstates the evidence against the null

Pitfall — posterior of null versus p

Bounds

Famous bound (Vovk 1993; Sellke, Bayarri, and Berger 2001) that under mild assumptions

$$B_{10} \leq \frac{1}{-ep \ln p}$$

With equal priors for null and alternative, $p = 0.05$ corresponds to at least about 30% posterior probability of the null (see e.g., Berger and Sellke 1987; Berger and Delampady 1987; Benjamin et al. 2018)

Examples from high-energy physics

I see this in high-energy physics. B for some Z (read papers for discussion of priors etc)

- Higgs discovery — posterior of null about 100 times greater than p (Fowlie 2019)
- ATLAS 2015 diphoton — $Z = 2.1\sigma$ and $B \simeq 7$ (Fowlie 2017)
- DAMPE — $Z = 2.3\sigma$ and $B \simeq 2$ (Fowlie 2018)
- 2020 XENON — $Z = 3.5\sigma$ and $B \simeq 3$ (Athron et al. 2020)

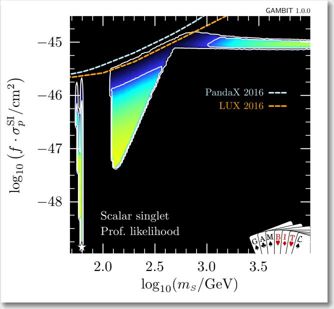
Results from GAMBIT

First results — 2017

Scalar singlet dark matter — 1705.07931

Mass and coupling + 13 nuisance parameters

LUX, XENON100, PandaX, SuperCDMS, IceCube, relic density, and Higgs width data

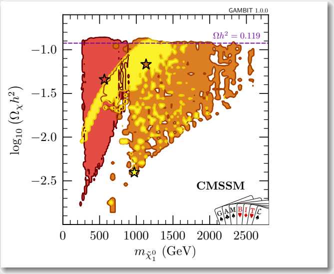


First results — 2017

Constrained MSSM — 1705.07935

4 GUT parameters + 5 nuisances

LUX, PandaX, and direct simulation of LHC Run 1 and early Run 2 searches

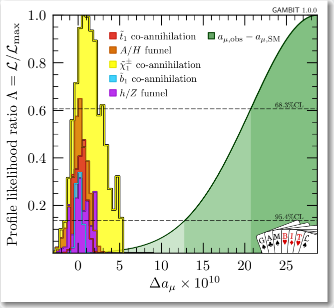


First results — 2017

Weak-scale MSSM — 1705.07917

Many parameters + 5 nuisances

LUX, PandaX, and direct simulation of LHC Run 1 and early Run 2 searches



Electroweakinos

Global collider physics fit of the electroweakino spectrum, using data from LEP, ATLAS and CMS — `1809.02097`

Combined collider constraints on neutralinos and charginos

The GAMBIT Collaboration: Peter Athron[1,2], Csaba Balázs[1,2], Andy Buckley[3], Jonathan M. Cornell[4], Matthias Danninger[5], Ben Farmer[6], Andrew Fowlie[1,2,9], Tomás E. Gonzalo[10], Julia Harz[11], Paul Jackson[12,2], Rose Kudzman-Blais[5], Anders Kvellestad[10,6,a], Gregory D. Martinez[13], Andreas Petridis[12,2], Are Raklev[10], Christopher Rogan[14], Pat Scott[6], Abhishek Sharma[12,2], Martin White[12,2,b], Yang Zhang[1,2]

Combine available leptonic SUSY search analysis likelihoods, and scan a larger phenomenological model space than in experimental papers

Decouple neutralinos and charginos

Name	Spin	P_R	Gauge Eigenstates	Mass Eigenstates
Higgs bosons	0	+1	$H_u^0\ H_d^0\ H_u^+\ H_d^-$	$h^0\ H^0\ A^0\ H^\pm$
squarks	0	-1	$\widetilde{u}_L\ \widetilde{u}_R\ \widetilde{d}_L\ \widetilde{d}_R$ $\widetilde{s}_L\ \widetilde{s}_R\ \widetilde{c}_L\ \widetilde{c}_R$ $\widetilde{t}_L\ \widetilde{t}_R\ \widetilde{b}_L\ \widetilde{b}_R$	(same) (same) $\widetilde{t}_1\ \widetilde{t}_2\ \widetilde{b}_1\ \widetilde{b}_2$
sleptons	0	-1	$\widetilde{e}_L\ \widetilde{e}_R\ \widetilde{\nu}_e$ $\widetilde{\mu}_L\ \widetilde{\mu}_R\ \widetilde{\nu}_\mu$ $\widetilde{\tau}_L\ \widetilde{\tau}_R\ \widetilde{\nu}_\tau$	(same) (same) $\widetilde{\tau}_1\ \widetilde{\tau}_2\ \widetilde{\nu}_\tau$
neutralinos	1/2	-1	$\widetilde{B}^0\ \widetilde{W}^0\ \widetilde{H}_u^0\ \widetilde{H}_d^0$	$\widetilde{\chi}_1^0\ \widetilde{\chi}_2^0\ \widetilde{\chi}_3^0\ \widetilde{\chi}_4^0$
charginos	1/2	-1	$\widetilde{W}^\pm\ \widetilde{H}_u^+\ \widetilde{H}_d^-$	$\widetilde{\chi}_1^\pm\ \widetilde{\chi}_2^\pm$
gluino	1/2	-1	$\widetilde{g}$	(same)

Decoupled

Thanks to Anders Kvellestad

Decouple neutralinos and charginos

Neutralinos

$$\psi^0 = (\tilde{B}, \tilde{W}^0, \tilde{H}_d^0, \tilde{H}_u^0)$$

$$M_N = \begin{pmatrix} M_1 & 0 & -\frac{1}{2}g'vc_\beta & \frac{1}{2}g'vs_\beta \\ 0 & M_2 & \frac{1}{2}gvc_\beta & -\frac{1}{2}gvs_\beta \\ -\frac{1}{2}g'vc_\beta & \frac{1}{2}gvc_\beta & 0 & -\mu \\ \frac{1}{2}g'vs_\beta & -\frac{1}{2}gvs_\beta & -\mu & 0 \end{pmatrix}$$

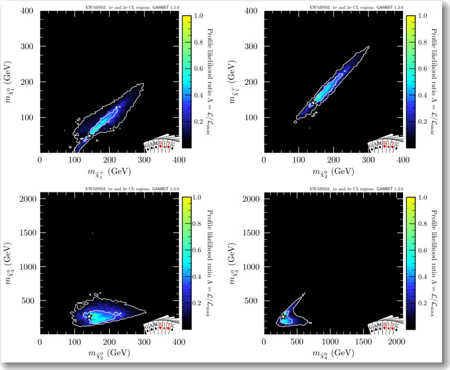
Charginos

$$\psi^\pm = (\tilde{W}^+, \tilde{H}_u^+, \tilde{W}^-, \tilde{H}_d^-)$$

$$M_C = \begin{pmatrix} 0 & X^T \\ X & 0 \end{pmatrix}, \quad \text{where} \quad X = \begin{pmatrix} M_2 & \frac{gvs_\beta}{\sqrt{2}} \\ \frac{gvc_\beta}{\sqrt{2}} & \mu \end{pmatrix}$$

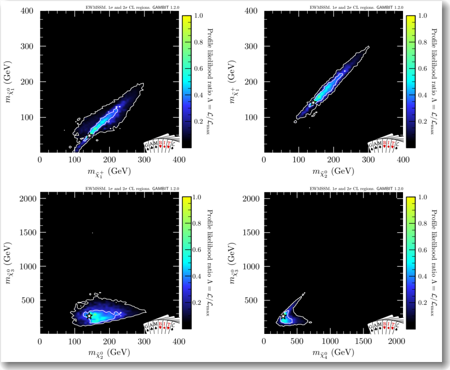
Thanks to Anders Kvellestad

- Combination of searches reveals mild excesses, local $\approx 3.3\sigma$
- Neutralinos and charginos $\lesssim 500$ GeV: is this light SUSY?
- Hidden from simplified models



Results

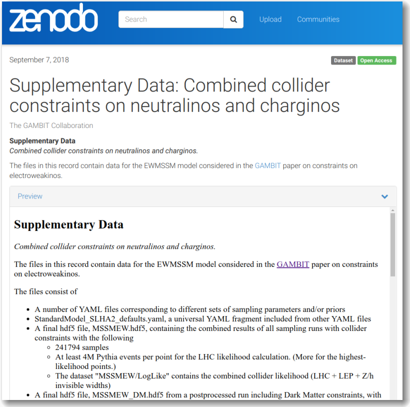
- Combination of searches reveals mild excesses, local $\approx 3.3\sigma$
- Neutralinos and charginos $\lesssim 500$ GeV: is this light SUSY?
- Hidden from simplified models



Anomaly gone in later datasets; update forthcoming

Open

Everything's open — code, data, plot scripts



For every GAMBIT paper, see https://gambit.hepforge.org/pubs

Summary

- There are many motivations for new physics
- Testing models of new physics requires complicated statistical and computational machinery — `GAMBIT`
- There are many subtleties in model testing
- Many public `GAMBIT` results and many more on the way!

References

References

- Aad, Georges et al. (2012). "**Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC.**" In: *Phys. Lett. B* 716, pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].
- Algeri, Sara et al. (Nov. 2020). "**Searching for new physics with profile likelihoods: Wilks and beyond.**" In: *Nat. Rev. Phys* 2.5, pp. 245–252. DOI: 10.1038/s42254-020-0169-5. arXiv: 1911.10237 [physics.data-an].

References ii

- Andrieu, Christophe and Gareth O. Roberts (Apr. 2009). **"The pseudo-marginal approach for efficient Monte Carlo computations."** In: *Ann. Statist.* 37.2, pp. 697–725. DOI: 10.1214/07-AOS574.
- Athron, Peter et al. (July 2020). **"Global fits of axion-like particles to XENON1T and astrophysical data."** In: arXiv: 2007.05517 [astro-ph.CO].
- Bailer-Jones, C.A.L. (2017). ***Practical Bayesian Inference: A Primer for Physical Scientists.*** Cambridge University Press. ISBN: 9781108127677.

References iii

- Balázs, Csaba et al. (Jan. 2021). **"A comparison of optimisation algorithms for high-dimensional particle and astrophysics applications."** In: JHEP 05, p. 108. DOI: 10.1007/JHEP05(2021)108. arXiv: 2101.04525 [hep-ph].
- Barlow, Roger (June 2004). **"Asymmetric statistical errors."** In: *Statistical Problems in Particle Physics, Astrophysics and Cosmology*, pp. 56–59. arXiv: physics/0406120.
- Barnard, G. A. (July 1949). **"Statistical Inference."** In: *Journal of the Royal Statistical Society: Series B (Methodological)* 11.2, pp. 115–139. DOI: 10.1111/j.2517-6161.1949.tb00028.x. URL: https://doi.org/10.1111/j.2517-6161.1949.tb00028.x.

- Behnke, O. et al. (2013). ***Data Analysis in High Energy Physics: A Practical Guide to Statistical Methods.*** Wiley. ISBN: 9783527653430.
- Benjamin, Daniel J. et al. (Jan. 2018). **"Redefine statistical significance."** In: *Nature Human Behaviour* 2.1, pp. 6–10. ISSN: 2397-3374. DOI: 10.1038/s41562-017-0189-z.
- Berger, J.O. and R.L. Wolpert (1988). ***The Likelihood Principle.*** second. Vol. 6. Lecture notes – monographs series. Institute of Mathematical Statistics. ISBN: 9780940600133.
- Berger, James O and Luis R Pericchi (2001). **"Objective Bayesian methods for model selection: Introduction and comparison."** In: *IMS Lecture Notes – Monograph Series* 38, pp. 135–207. DOI: 10.1214/lnms/1215540968.

References v

- Berger, James O. and Mohan Delampady (Aug. 1987). **"Testing Precise Hypotheses."** In: *Statist. Sci.* 2.3, pp. 317–335. DOI: 10.1214/ss/1177013238. URL: https://doi.org/10.1214/ss/1177013238.
- Berger, James O. and Thomas Sellke (1987). **"Testing a Point Null Hypothesis: The Irreconcilability of P Values and Evidence."** In: *J. Am. Stat. Assoc.* 82.397, pp. 112–122. DOI: 10.1080/01621459.1987.10478397.
- Birnbaum, Allan (1962). **"On the Foundations of Statistical Inference."** In: *J. Am. Stat. Assoc.* 57.298, pp. 269–306. DOI: 10.1080/01621459.1962.10480660.
- Chernoff, Herman (Sept. 1954). **"On the Distribution of the Likelihood Ratio."** In: *Ann. Math. Statist.* 25.3, pp. 573–578. DOI: 10.1214/aoms/1177728725.

Consonni, Guido et al. (June 2018). “**Prior Distributions for Objective Bayesian Analysis.**” In: *Bayesian Analysis* 13.2, pp. 627–679. DOI: 10.1214/18-BA1103.

Cousins, Robert D. (2008). “**Comment on ‘Bayesian Analysis of Pentaquark Signals from CLAS Data’, with Response to the Reply by Ireland and Protopopsecu.**” In: *Phys. Rev. Lett.* 101, p. 029101. DOI: 10.1103/PhysRevLett.101.029101. arXiv: 0807.1330 [hep-ph].

— (July 2018). “**Lectures on Statistics in Theory: Prelude to Statistics in Practice.**” In: *arXiv e-prints.* See Sec. 7.4. arXiv: 1807.05996 [physics.data-an].

- Cousins, Robert D. (Oct. 2020). **"What is the likelihood function, and how is it used in particle physics?"** In: *arXiv preprint*. CERN EP Newsletter. arXiv: 2010.00356 [physics.data-an].
- Cowan, G. (1998). ***Statistical Data Analysis*.** Clarendon Press. ISBN: 9780191583346.
- Cowan, Glen et al. (2011). **"Asymptotic formulae for likelihood-based tests of new physics."** In: *Eur. Phys. J. C* 71. [Erratum: *Eur. Phys. J. C* **73**, 2501 (2013), 10.1140/epjc/s10052-011-1554-0], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an].

- D'Agostini, Giulio (2003). ***Bayesian Reasoning In Data Analysis: A Critical Introduction.*** World Scientific Publishing Company. ISBN: 9789814486095.
- Demortier, Luv (2008). **"P Values and Nuisance Parameters."** In: DOI: 10.5170/CERN-2008-001.23. URL: https://cds.cern.ch/record/1099967.
- Diggle, Peter J. and Richard J. Gratton (1984). **"Monte Carlo Methods of Inference for Implicit Statistical Models."** In: *J. R. Stat. Soc.. Series B (Methodological)* 46.2, pp. 193–227. ISSN: 00359246.
- Fisher, R. A. (1925). ***Statistical Methods for Research Workers.*** Oliver and Boyd.

Fowlie, Andrew (2017). “**Bayes factor of the ATLAS diphoton excess: Using Bayes factors to understand anomalies at the LHC.**” In: *Eur. Phys. J. Plus* 132.1, p. 46. DOI: 10.1140/epjp/i2017-11340-1. arXiv: 1607.06608 [hep-ph].

— (2018). “**DAMPE squib? Significance of the 1.4 TeV DAMPE excess.**” In: *Phys. Lett. B* 780, pp. 181–184. DOI: 10.1016/j.physletb.2018.03.006. arXiv: 1712.05089 [hep-ph].

— (2019). “**Bayesian and frequentist approaches to resonance searches.**” In: *JINST* 14.10, P10031. DOI: 10.1088/1748-0221/14/10/P10031. arXiv: 1902.03243 [hep-ph].

References x

Fowlie, Andrew, Will Handley, and Liangliang Su (Oct. 2021). "**Nested sampling with plateaus.**" In: *Mon. Not. Roy. Astron. Soc.* 503.1, pp. 1199–1205. DOI: 10.1093/mnras/stab590. arXiv: 2010.13884 [stat.CO].

Goldstein, Michael (Sept. 2006). "**Subjective Bayesian Analysis: Principles and Practice.**" In: *Bayesian Analysis* 1.3. DOI: 10.1214/06-ba116. URL: https://doi.org/10.1214/06-ba116.

Goodman, Steven (2008). "**A Dirty Dozen: Twelve P-Value Misconceptions.**" In: *Seminars in Hematology* 45.3, pp. 135–140. ISSN: 0037-1963. DOI: 10.1053/j.seminhematol.2008.04.003.

Gould, Peter (1981). "**Letting the Data Speak for Themselves.**" In: *Annals of the Association of American Geographers* 71.2, pp. 166–176.

- Greenland, Sander et al. (Apr. 2016). **"Statistical tests, *P* values, confidence intervals, and power: a guide to misinterpretations."** In: *European Journal of Epidemiology* 31.4, pp. 337–350. ISSN: 1573-7284. DOI: 10.1007/s10654-016-0149-3.
- Gregory, P. (2005). ***Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica® Support*.** Cambridge University Press. ISBN: 9781139444286.
- Hubbard, Raymond and M. J Bayarri (2003). **"Confusion Over Measures of Evidence (*p*'s) Versus Errors (α's) in Classical Statistical Testing."** In: *Am. Stat.* 57.3, pp. 171–178. DOI: 10.1198/0003130031856.

- Hubbard, Raymond and R. Murray Lindsay (Feb. 2008). **"Why P Values Are Not a Useful Measure of Evidence in Statistical Significance Testing."** In: *Theory & Psychology* 18.1, pp. 69–88. DOI: 10.1177/0959354307086923.
- James, F. (2006). ***Statistical Methods in Experimental Physics.*** World Scientific. ISBN: 9789812567956.
- Jaynes, Edwin T (2003). ***Probability theory: The logic of science.*** Cambridge University Press.
- — (1968). **"Prior Probabilities."** In: *IEEE Transactions on Systems Science and Cybernetics* 4.3, pp. 227–241. DOI: 10.1109/TSSC.1968.300117.

- Jefferys, William H and James O Berger (1992). **"Ockham's razor and Bayesian analysis."** In: *American Scientist* 80.1, pp. 64–72.
- Jeffreys, Harold (1939). ***The Theory of Probability.*** Oxford Classic Texts in the Physical Sciences. Oxford University Press. ISBN: 9780198503682.
- Kass, Robert E. and Adrian E. Raftery (1995). **"Bayes Factors."** In: *J. Am. Statist. Assoc.* 90.430, pp. 773–795. DOI: 10.1080/01621459.1995.10476572.
- Kass, Robert E. and Larry Wasserman (1996). **"The Selection of Prior Distributions by Formal Rules."** In: *Journal of the American Statistical Association* 91.435, pp. 1343–1370. DOI: 10.1080/01621459.1996.10477003.

- Lakens, Daniël (2021). **"The Practical Alternative to the *p* Value Is the Correctly Used *p* Value."** In: *Perspectives on Psychological Science*. DOI: 10.1177/1745691620958012.
- Lakens, Daniel et al. (Mar. 2018). **"Justify your alpha."** In: *Nature Human Behaviour* 2.3, pp. 168–171. ISSN: 2397-3374. DOI: 10.1038/s41562-018-0311-x.
- Linden, W. von der, V. Dose, and U. von Toussaint (2014). ***Bayesian Probability Theory: Applications in the Physical Sciences.*** Cambridge University Press. ISBN: 9781107035904.
- Lyons, L. (1989). ***Statistics for Nuclear and Particle Physicists.*** Cambridge University Press. ISBN: 9781316101636.
- Lyons, Louis (Oct. 2013). **"Discovering the Significance of 5 sigma."** In: *arXiv preprint*. arXiv: 1310.1284 [physics.data-an].

MacKay, David J.C. (1992). **"Bayesian methods for adaptive models."** en. Doctoral dissertation. DOI: 10.7907/H3A1-WM07.

Mayo, Deborah G (2018). ***Statistical inference as severe testing.*** Cambridge University Press.

Mikkola, Petrus et al. (Dec. 2021). **"Prior knowledge elicitation: The past, present, and future."** In: arXiv: 2112.01380 [stat.ME].

Murtaugh, Paul A. (Mar. 2014). **"In defense of P values."** In: *Ecology* 95.3, pp. 611–617. DOI: 10.1890/13-0590.1.

Neyman, J. and E. S. Pearson (1933). **"On the Problem of the Most Efficient Tests of Statistical Hypotheses."** In: *Philos. Trans. Roy. Soc. London Ser. A* 231, pp. 289–337. ISSN: 02643952. DOI: 10.1098/rsta.1933.0009.

- Pratt, John W. (1962). **"On the foundations of statistical inference: Discussion."** In: *Journal of the American Statistical Association* 57.298, pp. 307–326. DOI: 10.1080/01621459.1962.10480661.
- Savage, L.J. (1962). **"On the foundations of statistical inference: Discussion."** In: *Journal of the American Statistical Association* 57.298, pp. 307–326. DOI: 10.1080/01621459.1962.10480661.
- Schervish, Mark J. (1996). **"P Values: What They are and What They are Not."** In: *Am. Stat.* 50.3, pp. 203–206. DOI: 10.1080/00031305.1996.10474380.
- Schittenhelm, Doris and Philipp Wacker (May 2020). **"Nested Sampling And Likelihood Plateaus."** In: arXiv: 2005.08602 [math.ST].

- Sellke, Thomas, M. J Bayarri, and James O Berger (2001). **"Calibration of p-values for Testing Precise Null Hypotheses."** In: *The American Statistician* 55.1, pp. 62–71. DOI: 10.1198/000313001300339950.
- Senn, S (Mar. 2001). **"Two cheers for P-values?"** In: *Journal of Epidemiology and Biostatistics* 6.2, pp. 193–204. DOI: 10.1080/135952201753172953.
- Sivia, D. and J. Skilling (2006). ***Data Analysis: A Bayesian Tutorial.*** Oxford University Press. ISBN: 9780198568322.
- Strumia, Alessandro (2016). **"Interpreting the 750 GeV digamma excess: a review."** In: *51st Rencontres de Moriond on EW Interactions and Unified Theories.* ARISF, pp. 407–426. arXiv: 1605.09401 [hep-ph].

References xviii

- Trotta, Roberto (2008). **“Bayes in the sky: Bayesian inference and model selection in cosmology.”** In: *Contemp. Phys.* 49, pp. 71–104. DOI: 10.1080/00107510802066753. arXiv: 0803.4089 [astro-ph].
- Tukey, John W et al. (1977). ***Exploratory data analysis.*** Vol. 2. Reading, MA.
- Vovk, V. G. (1993). **“A Logic of Probability, with Application to the Foundations of Statistics.”** In: *J. Royal Stat. Soc.* B55.2, pp. 317–341. DOI: 10.1111/j.2517-6161.1993.tb01904.x.
- Wagenmakers, E. and et al (2017). **“Bayesian Spectacles.”** In: https://www.bayesianspectacles.org/.

Wasserstein, Ronald L. and Nicole A. Lazar (Apr. 2016). **"The ASA Statement on p-Values: Context, Process, and Purpose."** In: *The American Statistician* 70.2, pp. 129–133. DOI: 10.1080/00031305.2016.1154108. URL: https://doi.org/10.1080/00031305.2016.1154108.

Wilks, S.S. (1938). **"The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses."** In: *Annals Math. Statist.* 9.1, pp. 60–62. DOI: 10.1214/aoms/1177732360.