Bayesian approach to naturalness

Andrew Fowlie

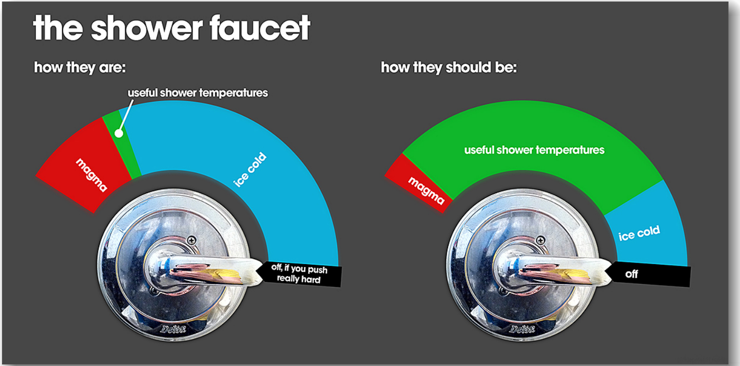
November 25 2016. Fine-tuning, the Multiverse and Life

Monash University

Table of contents

Nutshell

We've all experienced fine-tuning.



Fine-tuning in the 21st century

We know that showers that require fine-tuning are bad showers.

I will show you that models that require fine-tuning are bad models.

I will use natural = absence of fine-tuning (though be rigorous in context of Bayesian statistics).

In high-energy physics, a theory is considered fine-tuned/unnatural if small variations in its parameters result in dramatic changes in its predictions.

> *We criticise a proposed theory as fine-tuned when its featured are adjusted to make some things equal... [F]ine-tuning in a scientific theory is like a cry of distress from nature, complaining that something needs to be better explained.*

Weinberg [1]

Fine-tuning in the 14th century

You've surely heard this before — Occam's razor [2]:

> *When you have two competing theories that make exactly the same predictions, the simpler one is the better.*

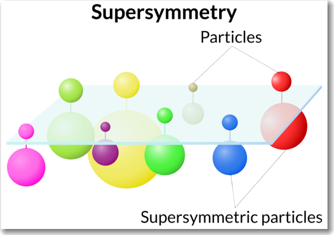
Fine-tuning in contemporary high-energy physics

Hierarchy problem

Higgs mass is fine-tuned — hierarchy problem [3–7]:

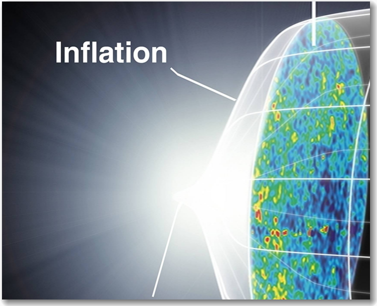
$$m^2 \simeq {m_0}^2 + \kappa M_P^2.$$

Require fine-tuning of bare mass and loop correction $\sim M_P^2$ such that $m^2 \lll M_P^2$.



Decades searching for solutions — supersymmetry [8–10], large extra dimensions [11] and technicolor [3].

Flatness and horizon problems — ordinary big-bang cosmology requires fine-tuning such that Universe is flat and causally disconnected patches have similar temperatures.



With cosmological inflation [12], flatness and homogeneity are generic.

Cosmological constant

Cosmological constant requires fine-tuning [13] of bare and corrections such that

$$\rho \lesssim 10^{-121}$$

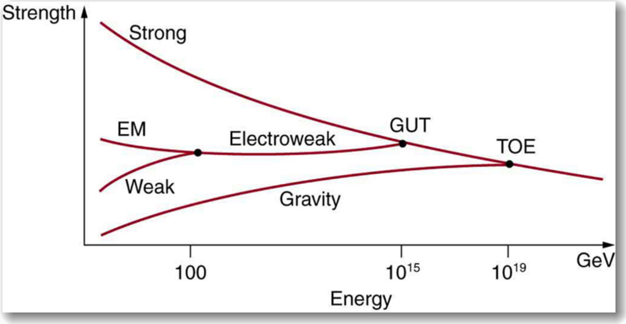
Decades searching for solutions. Weinberg [1] states

> *This level of fine-tuning is intolerable, and theorists have been working hard to find a better way to explain why the amount of dark energy is so much smaller than that suggested by our calculations.*

Anthropics invoked [14] along the way.

Unification

Grand unification: wouldn't it be simpler if forces/gauge couplings and even gravity were unified [15, 16]? Decades spent motivating new physics/constructing models.

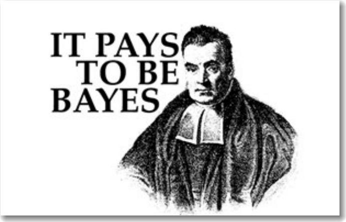


Bayesian automatic Occam's razor

Problems with Occam's razor/fine-tuning/naturalness

- What is meant by "better"?
- What is meant by "simpler"?
- What if the competing theories make slightly *different* predictions?
- Can we combine Occam's razor with experimental data and judge competing theories?
- Why is Occam's razor true? Is it true?
- Why are fine-tuned theories bad? Are they bad?

Fortunately, Bayes' theorem answers our questions.



Bayesian probability is a measure of plausibility [17, 18]. $0/1$ represent absolute certainty. Bayes' theorem automatically penalises fine-tuning and rewards agreement with data.

The measure of plausibility and the rules are unique. No choices/arbitrariness (Cox's theorem).

Inductive reasoning

Bayes' theorem permits inductive as well as deductive reasoning.

Deductive

Conclusion follows from premises. Premise: I saw a black swan. Conclusion: not all swans are white.

Inductive

Conclusion appears more/less plausible in light of premises. Premise: the sun has risen every day of my life. Conclusion: the sun will rise tomorrow.

Hume's problem of induction: is induction reliable? It worked in the past, but that's an inductive argument...

Inductive reasoning

Bayes' theorem permits inductive as well as deductive reasoning.

In a stroke, unify arguments about scientific theories (fine-tuning, agreement with data), and (arguably) solves problem of induction [21].

Bayes' theorem

Plausibility is updated by data via Bayes' theorem, e.g., plausibility of A given B:

$$p(A \mid B) = \frac{p(B \mid A)}{p(B)} \cdot p(A)$$

Contrasts with frequentist probability — frequency with which repeatable event occur in repeat trials.

If $p(B \mid A) / p(B) > 1$, A is more plausible in light of B.

What do you calculate?

We want to find which model is most plausible in light of data.
The relative plausibility is called the posterior odds:

$$\frac{p(M_b \,|\, D)}{p(M_a \,|\, D)} = \frac{p(D \,|\, M_b)}{p(D \,|\, M_a)} \times \frac{p(M_b)}{p(M_a)}$$

Posterior odds $=$ Bayes factor $\times$ Prior odds

This requires more than one model. With a single model, $p(M \,|\, D) = p(M) = 1$. Plausibility/fine-tuning is relative.

Furthermore, a single evidence $p(D \,|\, M)$ is meaningless, as it has $[1/D]$ dimension and dependent on parameterisation of data. Makes no sense to say "data improbable in a model, thus model is fine-tuned."

What do you calculate?

Calculate the Bayesian evidence for each model under consideration

$$p(D \mid M) = \int p(D \mid M, \boldsymbol{p}) \cdot p(\boldsymbol{p} \mid M) \prod \mathrm{d}\boldsymbol{p}$$

Probability of data given point in model (likelihood).
Probability of point given model (prior). *Somewhat* subjective, though should reflect knowledge or ignorance about parameters.

Compare the evidences in a so-called Bayes-factor:

$$p(D \mid M_b) \,/\, p(D \mid M_a) \propto p(M_b \mid D) \,/\, p(M_a \mid D)$$

which is proportional to the posterior odds. May not agree with frequentist methods, even with *informative* data [22, 23].

Automatic Occam's razor

We calculated the change in relative plausibility of models in light of data. The framework was unique.

There are caveats about priors.

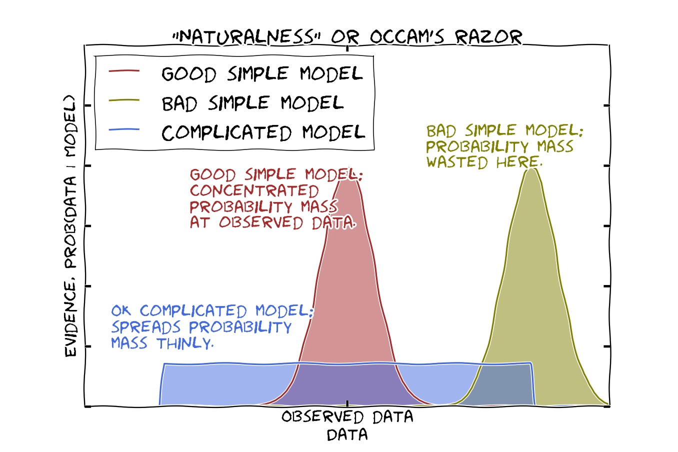
What more is there?! It would be re-assuring if it included an Occam's razor.

Note that the Bayesian evidence is a pdf of the data such that

$$\int p(D \mid M) \, \mathrm{d}D = 1$$

Each model has a finite amount of probability mass to spend/squander on its predictions for data.

Why Bayesian evidence captures fine-tuning in one slide [24]



Bayesian evidence captures old-fashioned ideas about FT.

`<digression>` Anthropics

Just data

The fact that you are alive, or that the Earth is habitable, is just data that could be included when judging plausibility by Bayesian evidences.

But it's typically useless as more convenient to use quantitative experimental data.

Weinberg's celebrated "anthropic" measurement of the cosmological constant in general relativity was just a measurement much like any other:

> *knowing only that the world is older than 5000 years and larger than Belgium would suffice to tell us that* $|\Lambda| \ll 1$ *[13]*

No ad hoc "anthropic reasoning" ever required/justified

</digression>

Caveats

Is it everything meant by Occam's razor/fine-tuning?

Many people have ideas about fine-tuning/Occam's razor/naturalness that aren't visible in Bayesian statistics.

That's their problem — not a fault of Bayesian statistics.

Whatever they were thinking of, if it doesn't appear in calculations in Bayesian statistics, it wasn't relevant to the plausibility of a model.

Priors

What you believe *after* seeing data depends on what you believed *before* seeing data. Priors should reflect knowledge/ignorance prior to experimental data.

Priors split into two pieces: prior of model, $p(M)$ and prior for parameters, $p(\boldsymbol{p} \mid M)$.

There are rules for quantifying ignorance (e.g., maximum entropy/invariance under symmetry groups [26]).

Some open issues, but nothing fatal. Bayesian statistics "too big to fail" — if you give up on it, the consequences (no induction) are disastrous.

Something else?

Could just stick with deductive logic. Don't talk about plausibility/fine-tuning. Ignore fine-tuning problems/solutions — limited.

Could allow fine-tuning arguments at qualitative level — vague, unreliable.

Could invent "pseudo" measures of plausibility. E.g., Barbieri-Giudice [27, 28] measure for fine-tuning in a SUSY theory. A theory is "good" if

$$\Delta \equiv \frac{\partial \ln M_Z}{\partial \ln p}$$

is "small." No logical framework. Completely arbitrary. Cannot be combined with experimental data. How is this reliable knowledge/scientific?

Calculations with real data, real models of interest, real conclusions about physics

Hierarchy problem [29] I

Supersymmetry motivated by fact that it “solves” fine-tuning problem associated with the weak scale in the Standard Model.

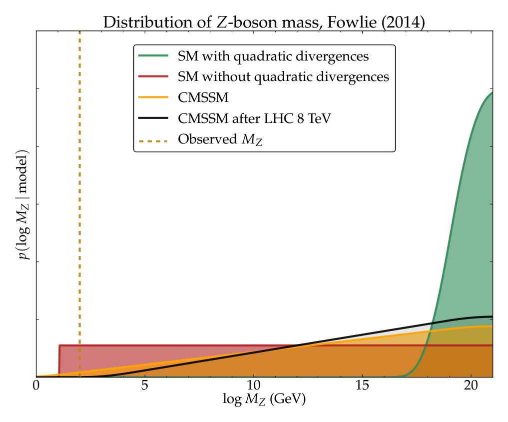
Weak scale is measurements of e.g., Z-boson mass, $M_Z \simeq 90\,\text{GeV}$.

I calculated the Bayes factor,

$$\frac{p(M_Z \,|\, \text{SUSY})}{p(M_Z \,|\, \text{SM})}$$

I found that SUSY (in this case the CMSSM) is favoured by a colossal Bayes factor in light of measurement of Z-boson mass.

Hierarchy problem [29] II



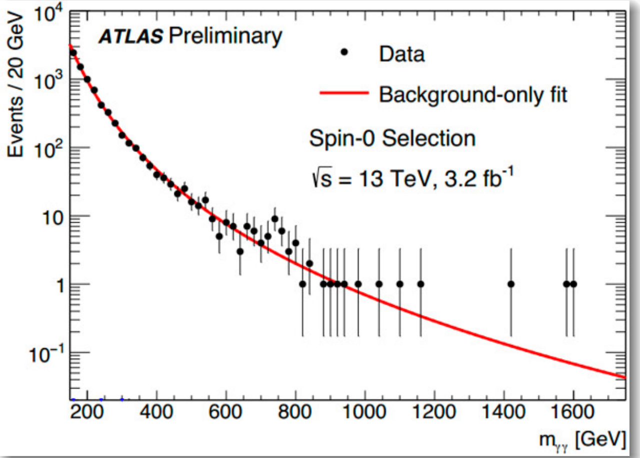
The SM squanders probability mass at $M_Z \sim M_P$ and is punished.
Physicists' intuition about fine-tuning/hierarchy problem was correct.

Digamma excess [32] I

Bayes factor needn't only be applied in cases in which fine-tuning is suspected — it is a general rule for judging models in light of data.

A recent "anomaly" from the LHC was the infamous digamma excess [31]. It resulted in about 500 theoretical studies between December 2015 and August 2016 (when it was refuted).

Digamma excess [32] II



An excess of events at 750 GeV that ostensibly indicated a new particle $m \approx 750$ GeV.

Digamma excess [32] III

Many people liked to speculate about the "odds" of a digamma particle. No one calculated. Once again, I calculated a Bayes factor,

$$\frac{p(D \,|\, \textsf{SM + digamma})}{p(D \,|\, \textsf{SM})}$$

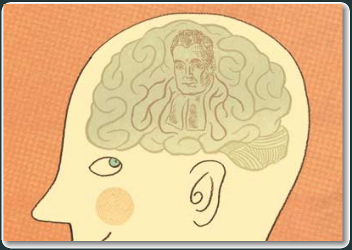
Found that it is $\lesssim 10$ at the peak of the excitement. Given that it wasn't that interesting prior to the data, an increase of plausibility of 10 isn't much.

Summary

Summary

- Fine-tuning/naturalness extremely important in high-energy physics.
- Theories should be judged by calculating their relative plausibility in light of data in Bayesian statistics.
- It automatically includes an effect that can identified with an Occam's razor/penalty for fine-tuning/unnaturalness.
- Applied it in several cases, including supersymmetric, non-supersymmetric models, and the diphoton anomaly.
- Intuition about fine-tuning was correct, but the formalism was absent/faulty. Bayesian statistics corrects that.

Questions?



References

[1] S. Weinberg, *To explain the world: the discovery of modern science*, (Penguin Books Limited, 2015).

[2] P. V. Spade, and C. Panaccio, "William of Ockham," in *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta, Winter 2016 (2016).

[3] S. Weinberg, "Implications of Dynamical Symmetry Breaking," Phys. Rev. D13, 974–996 (1976).

[4] S. Weinberg, "Implications of Dynamical Symmetry Breaking: An Addendum," Phys. Rev. D19, 1277–1280 (1979).

References II

[5] L. Susskind, "Dynamics of Spontaneous Symmetry Breaking in the Weinberg-Salam Theory," Phys. Rev. D20, 2619–2625 (1979).

[6] E. Gildener, "Gauge Symmetry Hierarchies," Phys. Rev. D14, 1667 (1976).

[7] E. Witten, "Dynamical Breaking of Supersymmetry," Nucl. Phys. B188, 513 (1981).

[8] A. Salam, and J. A. Strathdee, "Supergauge Transformations," Nucl. Phys. B76, 477–482 (1974).

[9] H. P. Nilles, "Supersymmetry, Supergravity and Particle Physics," Phys. Rept. 110, 1–162 (1984).

References III

[10] H. E. Haber, and G. L. Kane, "The Search for Supersymmetry: Probing Physics Beyond the Standard Model," Phys. Rept. 117, 75–263 (1985).

[11] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, "The Hierarchy problem and new dimensions at a millimeter," Phys. Lett. B429, 263–272 (1998), arXiv:hep-ph/9803315 [hep-ph].

[12] A. H. Guth, "The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems," Phys. Rev. D23, 347–356 (1981).

[13] R. Bousso, "TASI Lectures on the Cosmological Constant," Gen. Rel. Grav. 40, 607–637 (2008), arXiv:0708.4231 [hep-th].

References IV

[14] S. Weinberg, “Anthropic Bound on the Cosmological Constant,” Phys. Rev. Lett. 59, 2607 (1987).

[15] H. Georgi, and S. L. Glashow, “Unity of All Elementary Particle Forces,” Phys. Rev. Lett. 32, 438–441 (1974).

[16] S. Weinberg, *Dreams of a final theory: The Search for the fundamental laws of nature*, (Pantheon Books, 1992).

[17] H. Jeffreys, *The Theory of Probability*, (Oxford University Press, 1939).

[18] E. T. Jaynes, *Probability theory: the logic of science*, (Cambridge University Press, 2003).

References V

[19] T. Bayes, and R. Price, "An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S.," Philosophical Transactions 53, 370–418 (1763).

[20] J. O'Connor, and E. Robertson, "Thomas Bayes," in *MacTutor History of Mathematics archive* (2004).

[21] J. Vickers, "The Problem of Induction," in *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta, Spring 2016 (2016).

[22] D. V. Lindley, "A statistical paradox," Biometrika 44, 187–192 (1957).

References VI

[23] M. S. Bartlett, "A comment on D. V. Lindley's statistical paradox," Biometrika 44, 533–534 (1957).

[24] D. Mackay, *Bayesian Methods for Adaptive Models*, (Caltech thesis, 1992).

[25] W. Jefferys, and J. Berger, "Sharpening Occam's Razor on a Bayesian strop," in Bulletin of the American Astronomical Society, Vol. 23 (1991), p. 1259.

[26] G. L. Bretthorst, "An Introduction to Model Selection Using Probability Theory as Logic," in *Maximum Entropy and Bayesian Methods*, edited by G. R. Heidbreder, (Springer Netherlands, Dordrecht, 1996), pp. 1–42.

References VII

[27] J. R. Ellis, K. Enqvist, D. V. Nanopoulos, and F. Zwirner, “Observables in Low-Energy Superstring Models,” Mod. Phys. Lett. A1, 57 (1986).

[28] R. Barbieri, and G. F. Giudice, “Upper Bounds on Supersymmetric Particle Masses,” Nucl. Phys. B306, 63 (1988).

[29] A. Fowlie, “CMSSM, naturalness and the ”fine-tuning price” of the Very Large Hadron Collider,” Phys. Rev. D90, 015010 (2014), arXiv:1403.3407 [hep-ph].

[30] A. Fowlie, “Is the CNMSSM more credible than the CMSSM?” Eur. Phys. J. C74, 3105 (2014), arXiv:1407.7534 [hep-ph].

[31] A. Strumia, “Interpreting the 750 GeV digamma excess: a review,” (2016), arXiv:1605.09401 [hep-ph].

[32] A. Fowlie, “Bayes-factor of the ATLAS diphoton excess,” (2016), arXiv:1607.06608 [hep-ph].