# Model Specification and Efficient Truth Convergence

Rafael Fuchs

Munich Center of Mathematical Philosophy
LMU Munich

CAS Workshop 2026: Model Choice in Bayesian Inference
24th March 2026

# Motivation: The Ideal Story

Standard Truth Convergence results:

- if the true hypothesis $H^*$ is in the set of alternatives, the Bayesian agent converges on it in the limit (e.g. Blackwell & Dubins 1962)
- even if the truth is not in the set of alternatives, the posterior distribution can still converge on the *best alternative* (e.g. in the sense of minimal KL-divergence, see Barron 1998)

# Motivation: The Ideal Story

Standard Truth Convergence results:

- if the true hypothesis $H^*$ is in the set of alternatives, the Bayesian agent converges on it in the limit (e.g. Blackwell & Dubins 1962)
- even if the truth is not in the set of alternatives, the posterior distribution can still converge on the *best alternative* (e.g. in the sense of minimal KL-divergence, see Barron 1998)

Problems:

- The set of *all* possible hypotheses may be too large/intractable (and requires lots of data to infer anything)
- Model misspecification is prevalent in practice (see upcoming examples)
- We need convergence (with sufficiently high prob.) *before* the infinite limit ($\rightarrow$ e.g. calibrated error rates)

# Motivation: The Non-Ideal Story

- Grünwald & van Ommen (2017): models that are '*badly misspecified*' exhibit inconsistency: posterior distribution doesn't converge on 'closest alternative' (though posterior *predictive* distribution can)
  - *Example*: linear regression model that presupposes homoscedasticity, while true DGP is heteroscedastic.
- Gelman & Shalizi (2013): in practice we already know that our models are often wrong and misspecified (quite unavoidably so).
  - *Example/Case Study:* U.S. Voter Distributions (Dem vs. Rep), as a function of income (stratified by states)
  - varying-intercept logistic regression model: $P(Vote\ R \mid x, s) = \text{logit}^{-1}(a_s + b \cdot x)$ (income coded as $x \in \{-2, -1, 0, 1, 2\}$)
  - repeated adaptations in response to non-fit ($\rightarrow$ variable intercept, variable slope, non-monotonic effects,...)

# Plan

- Given $X_1, \ldots, X_n$ (*all* observable variables) the *full* Hypothesis space $\mathcal{H}^*$ is the set of all joint distributions $P(X_1, \ldots, X_n)$

  $\rightarrow$ huge, often intractable

- Any convex subset $\rightarrow$ convergence on the 'best', but often still huge (and may still require much data for informative conclusions)

- **Idea**: first, determine *roughly* the direction/area of the true model/effect, *then* refine model for more precise estimates

  $\Rightarrow$ *strategically distribute* **point-** (or small composite) **hypotheses** over $\mathcal{H}^*$ ('purposeful misspecification', perhaps more flexible than using parameters).

- *efficient* convergence on best alternative (with limited data): **stopping rule/threshold** with **calibrated error rates** (e.g. 95%); controllable risk of failure.

  $\Rightarrow$ The currency of success is not just accuracy but *efficiency* (accuracy gained per resource used, risk-benefit balance)

# Plan

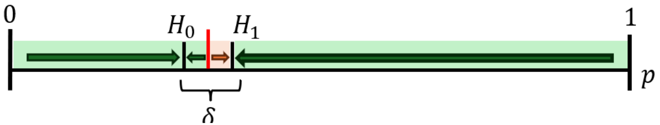
→ Plan for now:

1. Start with a simple example (binomial experiment)
2. Generalise to higher dimensions
3. Return to regression models
4. End with some Caveats

# Starting Point: Efficient Convergence in Sequential Testing

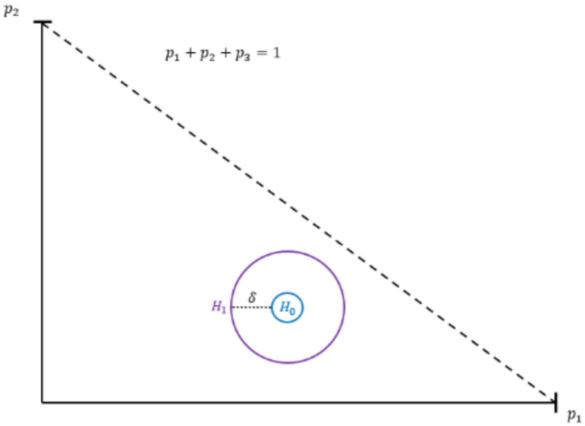
- **Example**: repeated independent observations of a *binary* random variable $X$, binomially distributed according to $P^*$ ($\rightarrow |\Omega| = 2$).
- Control group follows known distribution $P_0$
- Task: determine whether there is a relevant effect $\delta \geq \delta_{min}$ (e.g. effectiveness of treatment vs. rate of spontaneous recovery)
- Approach:
  - two *point hypotheses*: $H_0 : P^* = P_0$ and $H_1 : P^* = P_0 + \delta_{min}$,
  - start with uniform prior, stop at $P(H_i \mid X_n) \geq \tau$ (e.g. $\tau = 0.95$)

$\Rightarrow$ efficient convergence in $\tau$ percent, calibrated to error rate of $\tau$, (see also Schönbrodt et al. 2017); much more efficient on avg. than frequentist testing; convergence *time* points towards effect-size ($\delta$).



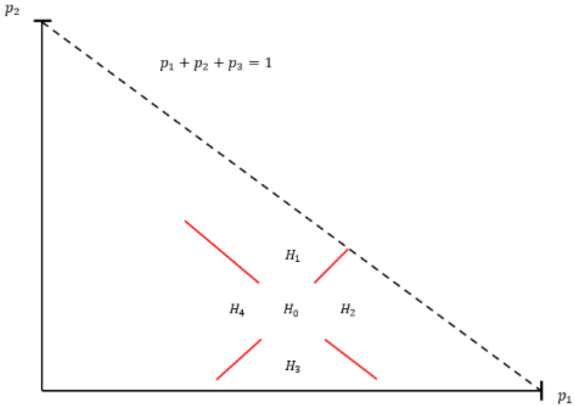
# Setting up well-structured hypothesis-spaces: Generalisation

We can generalise this to an outcome space with $|\Omega| = 3, \ldots$ and higher (still finite):



# Setting up well-structured hypothesis-spaces: Generalisation

We can generalise this to an outcome space with $|\Omega| = 3, \ldots$ and higher (still finite):



# Discretising Continuous Outcome Spaces

What if $\Omega$ is *continuous* (or at least countably infinite)?

- In practice, any empirical measurement has a *smallest meaningful unit* $\rightarrow$ *discrete* outcome space (for data/measurements)
- If there are also practical lower/upper bounds, the outcome space also becomes finite (height, weight, psychometrics, income etc.)
- $\rightarrow$ reverting to the finite case (even if just as an approximation for infinite/continuous cases)
- might be harder/more relevant in the context of *estimation*

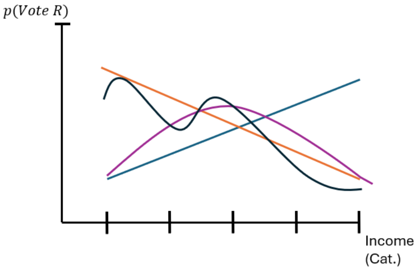
# Application: Regression Models

- Going back to regression models (initial examples): $P(Y) = f(X_1, \ldots, X_n)$
- how do we set up a well-structured hypothesis space here?
- $\rightarrow$ *Reference/starting point (again):* the set of all possible distributions over $\Omega$ (i.e. *all* joint distributions over $Y, X_1, \ldots, X_n$)
- $P(Y, X_1, \ldots, X_n) = P(Y \mid X_1, \ldots, X_n) \cdot P(X_1, \ldots, X_n)$, but we are interested specifically in $P(Y \mid X_1, \ldots, X_n)$ (which allows us to segment/batch and reduce dimensionality).

# Application: Regression Models

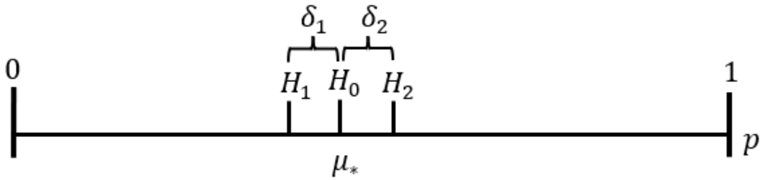
How to divide up the set of possible conditional distributions $\rightarrow$ many possible heuristics:

- Qualitative ordering (ranking, (non-)monotonicity, number of peaks)
- Quantitative specifications (varying levels of fine-grainedness):
  - strength of each influence,
  - interaction/aggregation of influences ($\rightarrow$ e.g. DAG),
  - change of strength (acceleration)
- **And:** specific choices informed by background theory (keeping in mind the complement)



# Application: Regression Models

**Or**: given enough data, we look at every value of the regressor variable(s), and update on the class-specific data ($\rightarrow$ the overall model emerges from that):

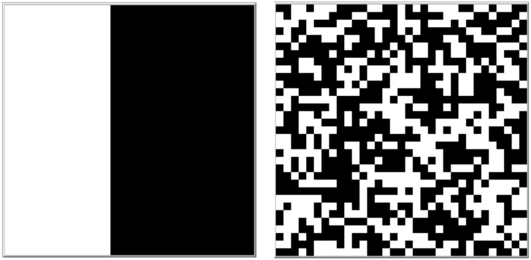



($\mu_*$: overall population mean/voter distribution; $H_0$: this subpopulation is not different from the overall population; $H_1/H_2$: this subpopulation is more left/right leaning).

# Caveat: the sampling process (observation vs. experiment)

Results may still turn out non-robust ('replication failure'); a set of related caveats:

- Statistical heterogeneity
- Hidden variables and changes over time, non-stationarity
- Biased Sampling from population (simple random sampling/controlled exp. not always possible)



These can be included in the model (but make things more complicated/less informative) $\Rightarrow$ *nested/conditional* analysis (hierarchical).

# Summary

- Since considering the space of *all* possible models is intractable, we need to make strategic choices
- These can be selected as points within the hypothesis space that cover a broad set of alternatives (as the 'best' fit)
- Idea: first do a 'broad scan' (limit the subregion of true/model effect), then zoom in.
- Efficient convergence with calibrated error rates (keeping in mind potential issues with sampling, heterogeneity, hidden-variables/changes over time)

# Thank You!

## References:


1. Blackwell, D. & Dubins, L. (1962). Merging of opinion with increasing information. *The Annals of Mathematical Statistics*, 33, 882–886.
2. Barron, A. R. (1998). 'Information-Theoretic Characterization of Bayes Performance and the Choice of Priors in Parametric and Nonparametric Problems.' In Bernardo, J. M., Berger, J. O., Dawid, A. P., and Smith, A. F. M. (eds.), *Bayesian Statistics*, vol. 6, 27–52. Oxford University Press.
3. Grünwald, P., & van Ommen, T. (2017). Inconsistency of Bayesian Inference for Misspecified Linear Models, and a Proposal for Repairing It. *Bayesian Analysis*, 12(4).
4. Gelman, A., & Shalizi, C. R. (2013). Philosophy and the practice of Bayesian statistics. *British Journal of Mathematical and Statistical Psychology*, 66(1), 8-38.
5. Schönbrodt, F. D., Wagenmakers, E. J., Zehetleitner, M., & Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological methods*, 22(2), 322.