Statistical Quantification of Discovery

Bayesian and Frequentist Perspectives

David A. van Dyk

Statistics Section, Imperial College London

INFN School on Statistics 2024

Search and Discovery

2012-13 Higgs Discovery

The New York Times

Science

Physicists Find Elusive Particle Seen as Key to Universe

Scientific and Statistical Themes

- High-stakes science: discovery vs. estimation.
- Model selection is much harder than estimation.
- Frequentist and Bayesian methods: different conclusions.
- Is a non-partisan approach possible?

Outline

Outline

Motivating Problem: Neutrino Oscillation

Neutrino Oscillation

- Neutrino created as electron, muon or tau may later be measured with different flavor.
- Flavor probability varies periodically as neutrino travels through space and *depends on several parameters*.

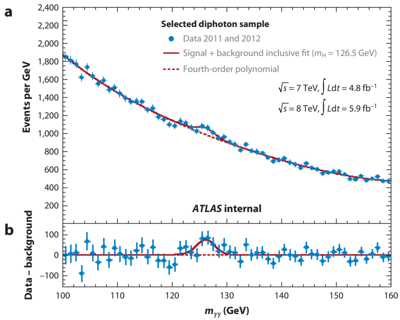
Mass Hierarchy *....ordering of the mass eigenstates*

- normal ($\Delta m^2_{32} > 0$) *vs* inverted hierarchy ($\Delta m^2_{32} < 0$)
- $|\Delta m^2_{32}|$ well constrained, degeneracy of sign with θ_{23} or δ_{CP}.

CP-violation

- Is there evidence to counter $\delta_{\mathrm{CP}} \in \{0, \pi\}$?
- Current data is limited.

Motivating Problem: Higgs Search



Searching for a Bump above Background

- Expect excess counts at invariant mass of Higgs boson.
- Statistically: no bump *vs* bump.
- The Location of possible bump unknown.
- What is the bump location if there is no bump?

Outline

Statistical Framework for Discovery

Model / Hypothesis Testing

H_0: The null hypothesis (e.g., no CP-violoation, $\delta_{\mathrm{CP}} = 0$)

H_1: The alternative hypothesis (e.g., CP-violation)

- Without further evidence, H_0 is presumed true.
- "Deciding" on H_1 means scientific discovery: new physics.
- Model Selection: No presumed model. (normal/inverted hierarchy)

Appropriate Statistical Approach Depends on

- Is H_0 the *presumed* model? or more than 2 possible models?
- Is H_0 a special case of H_1, "nested models"
- Parameters: (i) Unknown values under H_0?
  (ii) No "true value" under H_0?, (iii) Boundary concerns.
- Bayesian vs. Frequentist methods

Statistical Criterion for Discovery

The most common criterion is the p-value,

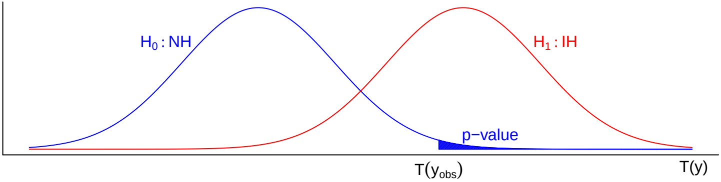
$$\text{p-value} = \Pr\left(T(y) \geq T(y_{\text{obs}}) \mid H_0\right)$$

- $T(\cdot)$ is a *Test Statistic*, e.g., $\Delta\chi^2$ or likelihood ratio statistic

$$\text{Likelihood Ratio Test} = -2\log\frac{\max_\theta p_0(y \mid \theta)}{\max_\theta p_1(y \mid \theta)}$$

Likelihood under H_0 (numerator)

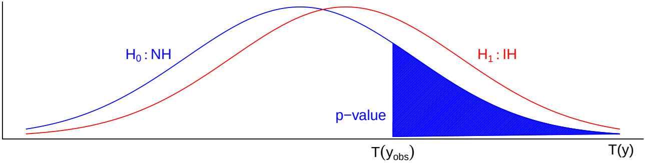
Likelihood under H_1 (denominator)



Computing p-values

The most common criterion is the p-value,

$$\text{p-value} = \Pr\left(T(y) \geq T(y_{\text{obs}}) \mid H_0\right)$$



Requires distribution of $T(y)$ under H_0

- Distributions depend on unknown parameters *(e.g., δ_{CP}, θ_{23})*
- Standard Theory:
  - estimates of unknown parameters converge to true values
  - models nested, parameter values under H_0, "large" data.

... often violated in physics

- Monte Carlo toys infeasible with 5σ criterion.

Misuse of P-values

The most common criterion is the p-value,

$$\text{p-value} = \Pr\left(T(y) \geq T(y_{\text{obs}}) \mid H_0\right) \ \text{ with } T = \text{test statistic}$$

But....

Misuse of P-values

The most common criterion is the p-value,

$$\text{p-value} = \Pr\left(T(y) \geq T(y_{\text{obs}}) \mid H_0 \right) \quad \text{with } T = \text{test statistic}$$

But....

nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive

Archive > Volume 519 > Issue 7541 > Research Highlights: Social Selection

NATURE | RESEARCH HIGHLIGHTS: SOCIAL SELECTION

Psychology journal bans *P* values

Test for reliability of results 'too easy to pass', say editors.

Chris Woolston

26 February 2015 | Clarified: 09 March 2015

Misuse of P-values

The most common criterion is the p-value,

$$\text{p-value} = \Pr\left(T(y) \geq T(y_{\text{obs}}) \mid H_0 \right) \quad \text{with } T = \text{test statistic}$$

But....

nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive

Archive > Volume 519 > Issue 7541 > Research Highlights: Social Selection

NATURE | RESEARCH HIGHLIGHTS: SOCIAL SELECTION

Psychology journal bans *P* values

Test for reliability of results 'too easy to pass', say editors.

Chris Woolston

26 February 2015 | Clarified: 09 March 2015

nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For

Archive > Volume 531 > Issue 7593 > News > Article

NATURE | NEWS

Statisticians issue warning over misuse of *P* values

Policy statement aims to halt missteps in the quest for certainty.

Monya Baker

07 March 2016

(ASA Statement on Statistical Significance and P-values)
February 5, 2016

The Problem with P-values

The misuse of P-values:

- **Do not measure relative likelihood of hypotheses.**
- Large p-values do not validate H_0.
- May depend on bits of H_0 that are of no interest.
- **Single filter** for publication / judging quality of research.
- **Should be viewed as a data summary, not the summary**

Reviewers, Editors, and Readers want a simple black-and-white rule: $p < 0.05$, or $> 5\sigma$.

But, statistics is about quantifying uncertainty, not expressing certainty.

A Bayesian Criterion for Discovery

To determine mass hierarchy, suppose we find

$$\text{p-value} = \Pr\left(T(y) \geq T(y_{\text{obs}}) \mid \text{NH}\right) = 0.0001$$

Questions

- Can we conclude NH is unlikely?
- Does $\Pr(\text{data} \mid \text{NH})$ small imply $\Pr(\text{NH} \mid \text{data})$ is small?

Order of conditioning matters!

Consider $\Pr(A \mid B)$ and $\Pr(B \mid A)$ with

A: A person is a woman.

B: A person is pregnant.

Bayesian Methods

Bayes Theorem

$$\Pr(\text{NH} \mid \text{data}) = \frac{\Pr(\text{data} \mid \text{NH})\,\Pr(\text{NH})}{\Pr(\text{data} \mid \text{NH})\,\Pr(\text{NH}) + \Pr(\text{data} \mid \text{IH})\,\Pr(\text{IH})}$$

Bayesian methods

- have cleaner mathematical foundations
- more directly answer scientific questions

... but they depend on ***prior distributions***

- $\Pr(\text{NH}) =$ probability of NH before seeing data.

Prior distributions must also be specified for model parameters.

The Problem with Priors

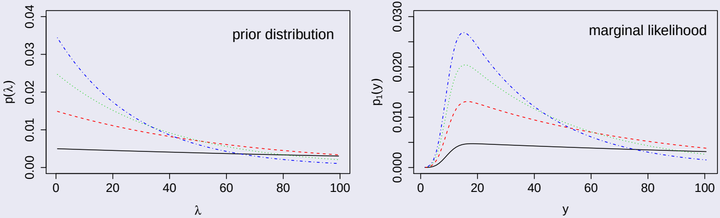
Bayesian Criteria for Discovery:

$$\text{Bayes Factor} = \frac{p_0(y)}{p_1(y)} \text{ with } p_i(y) = \int p_i(y|\theta)p_i(\theta)d\theta.$$

$$\Pr(H_0 \mid y) = \frac{p_0(y)\pi_0}{p_0(y)\pi_0 + p_1(y)\pi_1} = \frac{\pi_0}{\pi_0 + \pi_1(\text{Bayes Factor})^{-1}}$$

Example: (simplified) Higgs search

Likelihood: $y|\lambda \sim \text{Poisson}(10+\lambda)$ Test: $\lambda = 0$ vs $\lambda > 0$



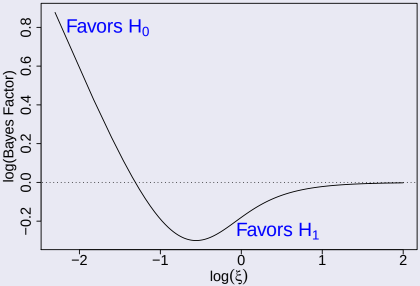
Value of $p_1(y)$ depends on prior!

Choice of Prior Matters!

Bayes Factor

$$H_0: \quad y \sim \text{Poisson}(10).$$
$$H_1: \quad y \sim \text{Poisson}(10+\lambda).$$
$$\text{with } \lambda \sim \exp(\xi)$$

- Observe $y = 15$
- log(Bayes Factor)



Must think hard about choice of prior and report!

Bayes Factors vs Likelihood Ratios

Likelihood Ratio optimizes parameters, whereas Bayes Factor marginalizes.

$$\text{Likelihood Ratio} = \frac{\max_{\theta_0} p_0(y \mid \theta_0)}{\max_{\theta_1} p_1(y \mid \theta_1)} \neq \text{Bayes Factor} = \frac{\int p_0(y \mid \theta_0)\, p(\theta_0)\, \mathrm{d}\theta_0}{\int p_1(y \mid \theta_1)\, p(\theta_1)\, \mathrm{d}\theta_1}$$

....unless there are no parameters under either model.

A Bayesian Occam's Razor

- Suppose $p(\theta_i)$ are both essentially flat over range where corresponding likelihoods are non-negligible.

$$\text{Bayes Factor} = \frac{\int p_0(y \mid \theta_0)\, p(\theta_0)\, \mathrm{d}\theta_0}{\int p_1(y \mid \theta_1)\, p(\theta_1)\, \mathrm{d}\theta_1} \approx \frac{p(\hat{\theta}_0) \int p_0(y \mid \theta_0)\, \mathrm{d}\theta_0}{p(\hat{\theta}_1) \int p_1(y \mid \theta_1)\, \mathrm{d}\theta_1}$$

- The term $p(\hat{\theta}_0)/p(\hat{\theta}_1)$ is sensitive to dimension and scale.
  - At mode, multivariate normal prior $\propto 1/|\Sigma|^{d/2}$.
- Bayes Factor penalizes larger models. *...and depends strongly on choice of prior.*
- The degree we penalize complex models is a subjective choice.
- Don't hide your priors!

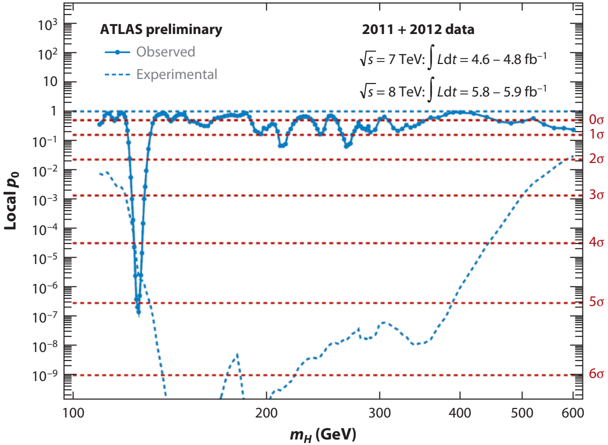
Frequentist vs Bayesian: Does it Matter?

Model Testing and Model Selection

- Frequency and Bayesian methods **may not agree.**
  - Bayes automatically penalizes larger models *(Occam's Razor)*
  - and adjusts for trial factors / look elsewhere effect.
- Choice of prior distribution **is often critical.**
- **Problem cases:** Dimension of model parameters differ.
  - CP-violation: $H_0 : \delta_{\mathrm{CP}} \in \{0, \pi\}$ vs. $H_1 :\notin \{0, \pi\}$.
  - Higgs search: location and intensity of bump above bkgd.
- Anti-conservative: p-value $\ll \Pr(H_0 \mid y)$.
- *Remember:*
  p-value and $\Pr(H_0 \mid y)$ *quantify different things!*

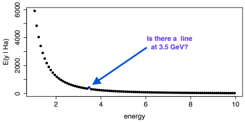
Interpreting p-value as $\Pr(H_0 \mid y)$ *may significantly overstate evidence for new physics.*

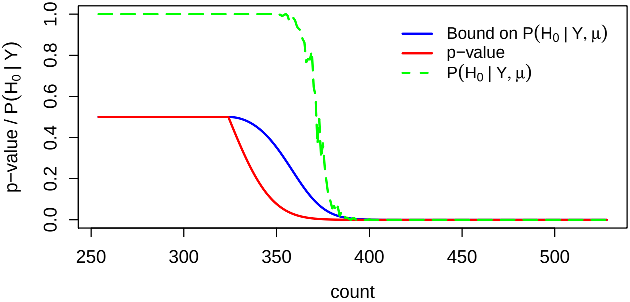
Trial Factors, Local, and Global p-values.



Reporting the minimum (local) p-value is cheating.

Example: Searching for a Bump above Background.





.... but researchers interpret p-value as $\Pr(H_0 \mid y)$*.*

Solution: Report both.

5σ Discovery Threshold

5σ is required for "discovery"

- High profile false discoveries led to conservative threshold
- Treat Higgs mass as known (multiple-testing)
- "What would you have done had you had different data"
- **Calibration, systematic errors, and model misspecification**
- But **cranking up required σ doesn't address these issues**

"In particle physics, this criterion has become a convention ... but should not be interpreted literally [1]."

At PhyStat-nu....

Cousins: *Two* 3.5σ *results are better than one* 5σ *result.*

van Dyk: *Calibrated* 3.5σ *result better than uncalibrated* 5σ.

[1] Glossary in the *Science* review of the 2012 CMS and ATLAS discoveries.

Outline

Normal Hierarchy versus Inverted Hierarchy

Non-nested parameterized models

$$H_0 : \text{normal hierarchy} \qquad \text{i.e., } \Delta m^2_{32} \leq 0$$
$$H_1 : \text{inverted hierarchy} \qquad \text{i.e., } \Delta m^2_{32} > 0$$

Computing a p-value using LRT

- Non-nested models: If no unknown parameters in either model.
  - LRT follows a Gaussian distribution under H_0 or H_1.
- With unknown parameters (e.g., Δm^2_{32}, δ_{CP}, θ_{23}):
  - Std theory (Wilks, Chernoff) does not apply: dist'n of LRT unknown.
  - Some results, but strong assumptions *(Blennow, et al. arXiv:1311.1822)*
    Apply with reactor neutrino experiments, not accelerator experiments which involve δ_{CP} (E. Ciuffoli).
  - What about uncertainty in $|\Delta m^2_{32}|$?

Are we back to Monte Carlo (toys)? at 5σ??

Is There an Easier Solution?

Two paradigms for statistical inference:

Likelihood: inference based on $p(y \mid \theta)$. *... and LRT, p-value, etc.*

Bayesian: inference based on $p(\theta \mid y) \propto p(y \mid \theta)p(\theta)$.

Model Fitting

- Specify one model, fit parameters, estimate uncertainty.
- Frequency and Bayesian methods tend to agree.
- Choice of prior distribution is often not critical.

Some "model selection" can be accomplished via model fitting, e.g., confidence intervals.

Normal versus Inverted Hierarchy: Easier Way?

Non-nested parameterized models

H_0 : normal hierarchy i.e., $\Delta m^2_{32} \leq 0$

H_1 : inverted hierarchy i.e., $\Delta m^2_{32} > 0$

Is there an easier solution??

Why not just compute $\Pr(H_0 \mid y) = \Pr(\Delta m^2_{32} \leq 0 \mid y)$?

In this case Bayes Criterion is particularly easy:

$$\text{Posterior Odds} = \frac{\Pr(\Delta m^2_{32} \leq 0 \mid y)}{\Pr(\Delta m^2_{32} > 0 \mid y)}$$

...model fitting with Δm^2_{32} a free parameter.

One model and one prior, easy to compute, not sensitive to prior... what's not to like?

Bayesian solution is easier in this case.

CP-violation

Test: $H_0 : \delta_{\mathrm{CP}} \in \{0, \pi\}$ versus $H_1 : \delta_{\mathrm{CP}} \notin \{0, \pi\}$

p-value

- Standard theory (Wilks, Chernoff) applies...
  but insufficient data for asymptotics.
- Monte Carlo (toys) required to assess p-value.
- More data required! (For 5σ??)

Posterior Odds or Bayes Factor

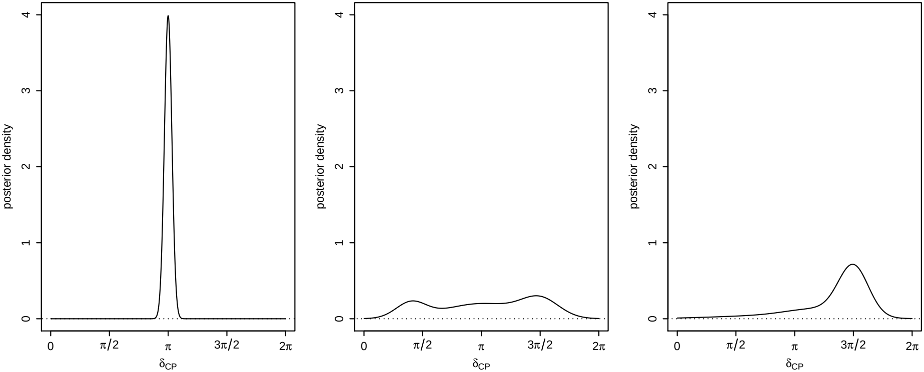
- Sensitive to prior on δ, but finite support.

Again, Bayesian solution is easier (with limited data).

Still Easier:

- Report a confidence interval for δ_{CP}.
- Employ model fitting rather than model selection.

Assessing CP-violation via Model Fitting



Is data consistent with $\delta_{\text{CP}} \in \{0, \pi\}$*??*

Outline

Higgs Search: Statistical Framework

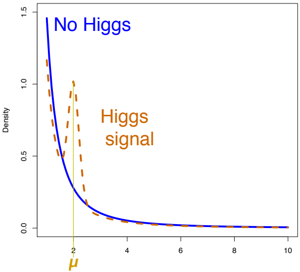
A Mixture Model:

$$f(y_i|\theta) = (1-\lambda)f_0(y_i|\alpha) + \lambda f_1(y_i|\mu)$$
$$= \text{background} + \text{Higgs}$$

Compare

$$H_0 : \lambda = 0 \quad \text{(no discovery)}$$
$$H_1 : \lambda > 0 \quad \text{(discovery)}$$

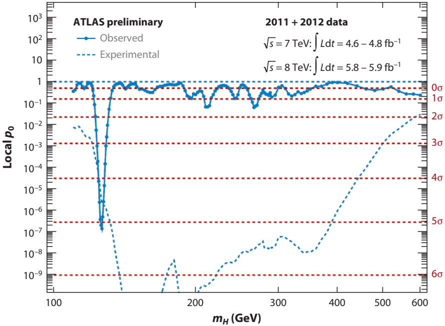
(And $\lambda < 1$, there will always be background!)



Types of Parameters:

1. α: (nuisance) parameter for H_0
2. λ: parameter determining hypothesis
3. μ: bump location, *has not value under H_0*.

Trial Factors, Local, and Global p-values.



- For fixed μ: Chernoff's Theorem applies, asymptotic null distribution known, and we can compute local p-values.
- *But, reporting the minimum (local) p-value is cheating!!*
- Global p-values correct for multiple looks.

Bounding the Global P-value

Consider the stochastic process $\{T_\mu(y), \mu \in M\}$ indexed by μ.

- Statistic: $T^+(y) = \max_{\mu \in M_R} T_\mu(y)$, maximize over grid of size R.
- Global P-value:

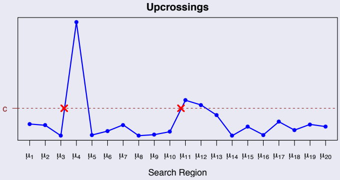
$$p_G = \Pr\left(\max_{\mu \in M_R} T_\mu(y) \geq \max_{\mu \in M_R} T_\mu(y_{\text{obs}}) \mid H_0\right)$$

Bounds on global p-value

1. Bonferroni: $p_G \leq R p_L$.
2. Markov (Davies, 1987):

$$\begin{aligned} p_G &= \Pr\left(\max_\mu T_\mu(y) \geq c \mid H_0\right) \\ &\leq p_L + \Pr(N_c \geq 1 \mid H_0) \\ &\leq p_L + \mathrm{E}(N_c \mid H_0). \end{aligned}$$

N_c = number of upcrossings.



Background on Bounds

Bonferroni Bound

Suppose we conduct two tests, with $\Pr(T_i \geq c) = \epsilon$,

$$\begin{aligned} \Pr(T_1 \geq c \text{ or } T_2 \geq c) &= \Pr(T_1 \geq c) + \Pr(T_2 \geq c) - \Pr(T_1 \geq c \text{ and } T_2 \geq c) \\ &\leq \Pr(T_1 \geq c) + \Pr(T_2 \geq c) = 2\epsilon. \end{aligned}$$

Thus, bound on global p-value is twice local p-value.

Markov Bound

Let X be a random variable that can take on values $0, 1, 2, \ldots$.

$$\begin{aligned} E(X) &= \sum_{x=0}^{\infty} x \Pr(X = x) \geq \sum_{x=1}^{\infty} x \Pr(X = x) \\ &\geq \sum_{x=1}^{\infty} \Pr(X = x) = \Pr(X \geq 1). \end{aligned}$$

Evaluating the Bounds

Questions:

1. Which bound is sharper?
2. Which bound is easier to compute?

The method of Gross and Vitells (2010)

- To avoid MC evaluation of $\mathrm{E}(N_c|H_0)$

$$\mathrm{E}(N_c \mid H_0) = \mathrm{E}(N_{c_0} \mid H_0) \left(\frac{c}{c_0}\right)^{(s-1)/2} \exp\left(-\frac{(c-c_0)}{2}\right), \quad c_0 \ll c$$

- 6σ / 5σ significances reduce to 5.1σ /4.6σ (ATLAS/CMS)

Higgs Search: Is a Bayes Factor Possible?

Types of Parameters:

1. α: parameter for H_0
2. λ: determines hypothesis
3. μ: no value under H_0.

Basic Model:

$$\begin{aligned} p(y_i|\theta) &= (1-\lambda)f_0(y_i|\alpha) + \lambda f_1(y_i|\mu) \\ &= \text{background} + \text{Higgs} \end{aligned}$$

P-values are "biased toward discovery." How about $\Pr(H_0 \mid y)$?

Strategies for Setting Prior Distributions

- Easiest case: Bkgd parameters common to both models.
- Diffuse prior: flat over region where $p_i(y|\alpha)$ non-negligible.
- Fixing λ and μ,

$$\text{BF} = \frac{\int \prod_i f_0(y_i|\alpha)p(\alpha)d\alpha}{\int \prod_i [(1-\lambda)f_0(y_i|\alpha) + \lambda f_1(y_i|\mu)]p(\alpha)d\alpha} = \frac{p(\hat{\alpha}_0)\int p_0(y|\alpha)d\alpha}{p(\hat{\alpha}_1)\int p_1(y|\alpha)d\alpha}$$

- The choice of prior on α is not critical.

Hypothesis Indexing Parameter: λ

Lower Bound on Bayesian evidence for H_0

- P-values tend to favor H_1 more strongly than $\Pr(H_0 \mid y)$.

  [At least when H_0 is "precise".]

- Using a parameterized prior $\lambda \sim p(\lambda \mid \beta)$,

$$\bar{p}_1(y \mid \mu) = \sup_{\beta} \int p_1(y \mid \lambda, \mu) p(\lambda \mid \beta) \mathrm{d}\lambda$$

$$\Pr(H_0 \mid y, \mu) = \frac{\pi_0 p_0(y)}{\pi_0 p_0(y) + \pi_1 p_1(y \mid \mu)} \geq \frac{\pi_0 p_0(y)}{\pi_0 p_0(y) + \pi_1 \bar{p}_1(y \mid \mu)}$$

Example

$y_i \overset{\text{indep}}{\sim} \textsc{Poisson}\big(f_0(\alpha, i) + \lambda f_1(\mu, i)\big)$

Test: $H_0 : \lambda = 0$ vs $H_0 : \lambda > 0$

- $\lambda \sim \textsc{Gamma}(\alpha, \beta)$
- Prior should peak at zero: we set $\alpha = 1$.

Parameters Not Identifiable Under H_0: μ

Local $p(H_0|y)$: $\inf_\mu p(H_0 \mid y, \mu)$

Global $p(H_0|y)$: properly average over $p(\mu)$

Like global p-value, averaging over $p(\mu)$ penalizes wide search

$$
\begin{aligned}
p_1(y) &= \int p_1(y \mid \mu) p(\mu) \mathrm{d}\mu \leq \sup_\mu p_1(y \mid \mu) \\
\Pr(H_0 \mid y) &= \frac{\pi_0 p_0(y)}{\pi_0 p_0(y) + \pi_1 p_1(y)} \geq \frac{\pi_0 p_0(y)}{\pi_0 p_0(y) + \pi_1 \sup_\mu p_1(y \mid \mu)} \\
&= \inf_\mu p(H_0 \mid y, \mu) = \text{Local probability of } H_0
\end{aligned}
$$

- Simplest choice of $p(\mu)$ is uniform over the search region.
  - Look-elsewhere correction similar to frequency methods.

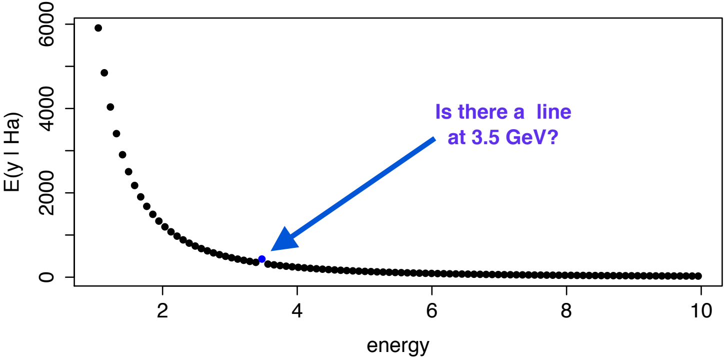
Example: Are P-values Biased in Favor H_1?

Model:

$y_i \overset{\text{indep}}{\sim} \textsc{Poisson}\Big(f_0(\alpha, i) + \lambda f_1(\mu, i)\Big)$

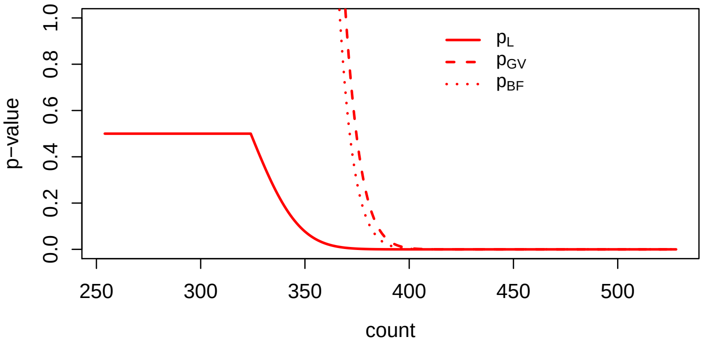
Test: $H_0 : \lambda = 0$ vs $H_0 : \lambda > 0$

- f_0 = power law
- $f_1 = \mathcal{I}\{i = \mu\}$
- 100 bins



Example: Local vs Global P-values

- Varying the count in the line bin (3.5 GeV).
- The expected count in this bin under H_0: 330.

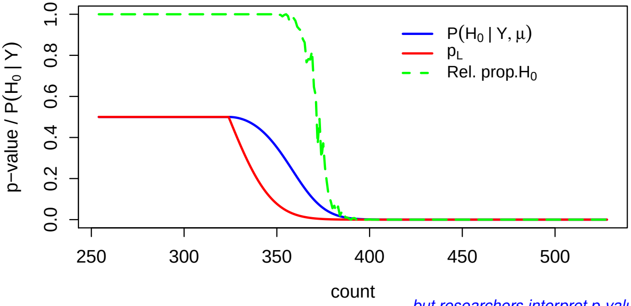


Example: Comparing $\Pr(H_0 \mid y)$ with p-value

Consider physicists who repeatedly conducts hypothesis tests

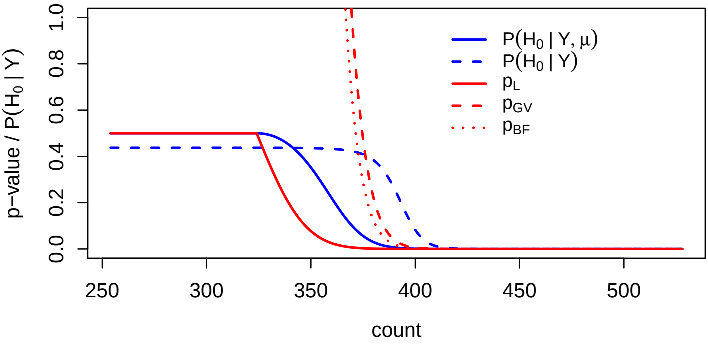
- Half the time H_0 is true; when H_1 is true, $\mu = 3.5 GeV$.
- Dashed green line: relative frequency of H_0.

We compute lower bound on $\Pr(H_0 \mid y)$.... *[Recall prior on λ.]*



.... but researchers interpret p-value as $\Pr(H_0 \mid y)$.

Natural Bayesian correction for multiple testing



Prior on μ naturally and simply corrects for the "look elsewhere effect"

For Bayesians the challenges are different... setting the prior.

Outline

Frequentist or Bayesian?

Do you have to choose??

- Bayes prescribes methodology.
- Frequentists evaluate methods.
- Frequency evaluation of Bayesian methods.
- Model fitting: often little difference in fits and errors.
- Why not control rate of false detection
  and assess probability of new physics?
- Why throw away half of your tool box?

Neutrino physicists open to both Bayesian / Frequency methods

- Lots of Bayesian and Frequentist proposals at PhyStat-ν.
- My experience with cosmologists and particle physicists.

Strategies

What is a physicists to do?

- Controlling false discovery is critical in physical sciences.
- Comparing p-values with a predetermined significant level can control false discovery.... *if used with care, e.g., no cherry picking!*
- When confronted with small p-values researchers *...even statisticians!!...* may believe H_0 is unlikely.
- Bayesian solutions can better quantify likelihood of H_0 / H_1.
- **Solution:** Compute both *global* p-value *and* Bayes Factor.

But be Careful...

1. *quantification of p-values in non-standard problems*
2. *choice and validation of prior distributions*

remain challenging!

Resources

PhyStat-ν Tokyo

- http://indico.ipmu.jp/indico/conferenceDisplay.py?confId=82
- Summary Document in preperation

PhyStat-ν Fermilab

- Continuation of meeting in Japan.
- https://indico.fnal.gov/conferenceDisplay.py?confId=11906

PhyStat Repository

- Links to ten PhyStat meetings, with slides, papers, and proceedings.
- Some software packages and tools
- http://www.phystat.org

References

van Dyk, D. A. (2014).
The Role of Statistics in the Discovery of a Higgs Boson.
Annual Review of Statistics and Its Application, **1**, 41–59.

Stein, N. M., van Dyk, D. A., Kashyap, V. L., and Siemiginowska, A. (2015).
Detecting Unspecified Structure in Low-Count Images.
The Astrophysical Journal, **813**, 66 (15pp).

Algeri, S., Conrad, J., and van Dyk, D. A. (2016).
Comparing Non-Nested Models in the Search for New Physics.
Monthly Notices of the Royal Astronomical Society: Letters, **458 (1)**, L84-L88.

Algeri, S., van Dyk, D. A., Conrad, J., and Anderson, B. (2016).
Methods for Correcting the Look-Elsewhere Effect in Searches for New Physics.
Journal of Instrumentation, **11**, P12010.

Algeri, S. and van Dyk, D. A., and Conrad, J. (2017+).
Testing one Hypothesis Multiple Times.
Submitted.

Workshop Participants (2017+).
PhyStat-ν 2016 at the IPMU: A Summary.
In preparation.