Bayesian and frequentist approaches to resonance searches

Andrew Fowlie

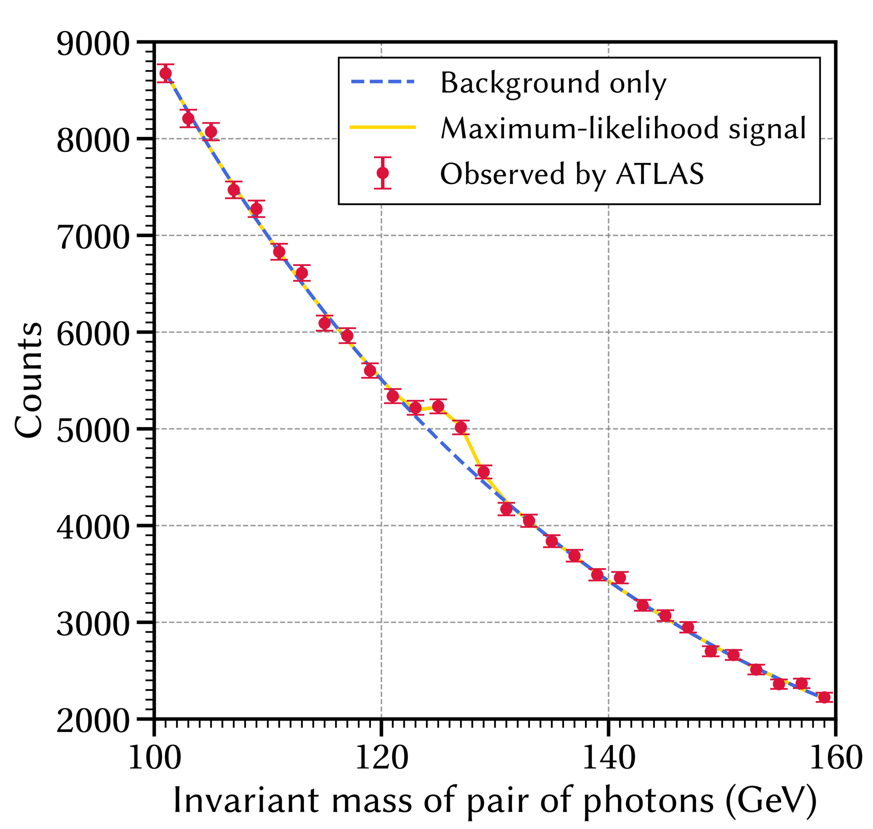
Department of Physics and Institute of Theoretical Physics, Nanjing Normal University, China

Summary

- Investigated Bayesian and frequentist approaches to discoveries in resonance searches at the LHC
- Compared behavior of global p-values and posterior of the background only hypothesis in toy experiments
- Posterior converged as we accumulated data, whereas the p-value made random walk under the background only model
- P-values were typically one or two orders of magnitude smaller than the posterior

New physics at the LHC

New physics could appear as a resonance in an invariant mass distribution. This is how the Higgs boson was discovered.[1,2]



This shows a result from an ATLAS search[3] for the Higgs in the diphoton channel.

Each data point is a counting experiment — we observe a particular number of counts which we compare with the number predicted by the background and a model with a new particle.

To make many toy experiments, we neglect systematic uncertainties in the background model.

Statistical approaches

Bayesian

We calculated the Bayes factor[4,5] between the background only and background + signal models,

$$\text{Bayes factor} \equiv B = \frac{P(\text{data} \,|\, H_1)}{P(\text{data} \,|\, H_0)}$$

To facilitate a comparison with p-values, we found the posterior of the background model,

$$\text{Posterior of background} = \frac{P(H_0)}{P(H_0) + B P(H_1)}$$

The signal model contained two a priori unknown parameters: the Higgs mass and the signal strength. We marginalised them,

$$P(\text{data} \,|\, H_1) = \int P(\text{data} \,|\, \mathbf{x})\, p(\mathbf{x} \,|\, H_1)\, d^2x$$

The second factor is our choice of prior density for the unknown parameters, $\mathbf{x}$.

Frequentist

We constructed a log likelihood ratio test-statistic,[6]

$$q \equiv -2 \ln \left(\frac{P(\text{data} \,|\, \hat{\mathbf{x}}, H_1)}{P(\text{data} \,|\, H_0)} \right)$$

where the hat indicates the maximum-likelihood parameters.

We found the global p-value — the probability of obtaining a test-statistic at least as extreme as that observed, assuming that the background only model is true,

$$p\text{-value} = P(q \geq q_{\text{Observed}} \,|\, H_0)$$

We found it using a semi-analytic Monte-Carlo method called Gross-Vitells,[7]

$$p\text{-value} \approx \frac{1}{2} P\left(\chi^2_1 > q\right) + N e^{-q/2}$$

This includes a look-elsewhere effect.

Evolution as we collect data in toy experiments

Under the background model

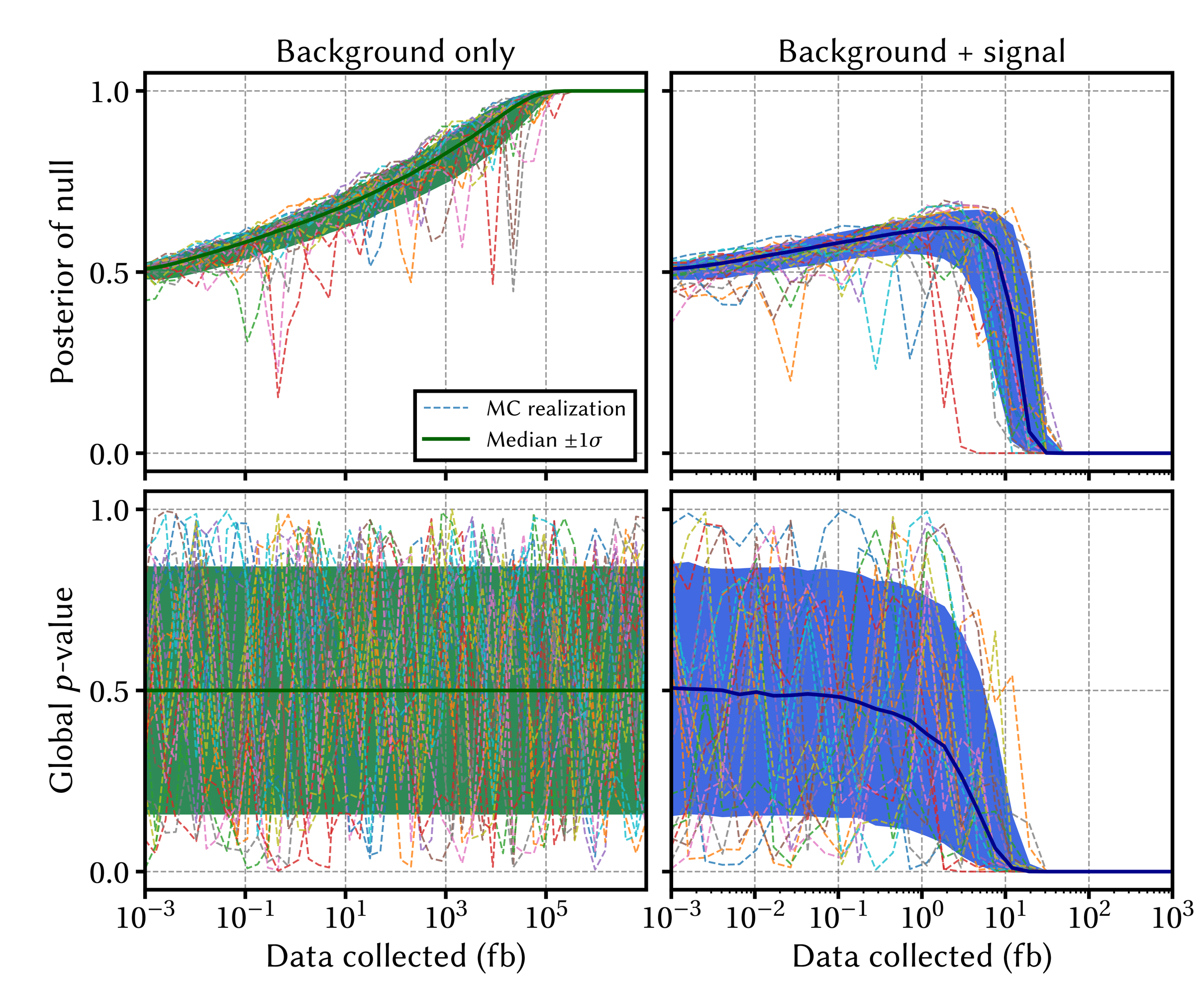
Posterior of the background slowly converged from a half to one, within a narrow interval.

The p-value made a random walk between zero and one.

Under the signal model

Surprisingly, the posterior first increased to about 0.7 but ultimately rapidly decreased towards zero.

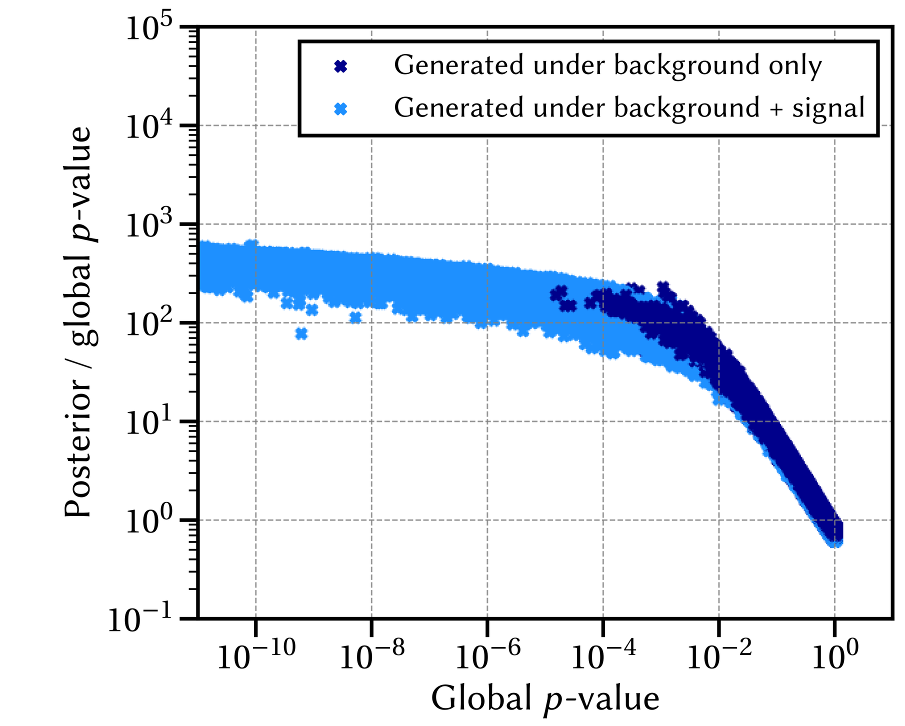
The p-value moved noisily but monotonically towards zero.



Median posterior of the background and global p-value as we collect data.

Comparison

We directly compared the posterior of the background only model and the p-value obtained in our toy experiments.



The global p-value was typically one or two orders of magnitude smaller than the posterior of the null. Similar to findings in other contexts.[8,9]

References

1. S. Chatrchyan *et al.* (CMS), Phys. Lett. **B716**, 30–61 (2012).
2. G. Aad *et al.* (ATLAS), Phys. Lett. **B716**, 1–29 (2012).
3. G. Aad *et al.* (ATLAS), Phys. Lett. **B726**, 88–119 (2013), [Erratum: Phys. Lett. **B734**, 406 (2014)].
4. H. Jeffreys, *The Theory of Probability*, Oxford Classic Texts in the Physical Sciences (Oxford University Press, 1939).
5. R. E. Kass and A. E. Raftery, J. Am. Stat. Assoc. **90**, 773–795 (1995).
6. G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Eur. Phys. J. **C71**, 1554 (2011), [Erratum: Eur. Phys. J. **C73**, 2501 (2013)].
7. E. Gross and O. Vitells, Eur. Phys. J. **C70**, 525–530 (2010).
8. J. O. Berger and M. Delampady, Statist. Sci. **2**, 317–335 (1987).
9. T. Sellke, M. J. Bayarri, and J. O. Berger, The American Statistician **55**, 62–71 (2001).

Further Information

andrew.j.fowlie@nnu.edu.cn

Preprint: arxiv.org/abs/1902.03243