Comparing non-nested models
in the search for new physics

Sara Algeri\textsuperscript{1,2}
In collaboration with: Prof. David van Dyk\textsuperscript{1} and Prof. Jan Conrad\textsuperscript{2}

\textsuperscript{1}Department of Mathematics, Imperial College London, London, UK
\textsuperscript{2}Department of Physics, Stockholm University, Stockholm, Sweden.

SCMA-VI, 2016
Detection of new physics - The scientific problem

Detection of a new particle

- E.g., Higgs boson, quark, neutrino.
- We want to detect a bump (the signal of the new particle) on top of a background flux.

Distinguish known astrophysics from new signals

- E.g., Dark Matter.
- We can even have a fake signal, i.e., something mimicking Dark Matter, but not a background to it.
Detection of a new particle - The statistical problem

The model of interest is proportional to

\[ f(y, \alpha) + \mu g(y, \beta) \]

and we test

\[ H_0 : \mu = 0 \quad \text{vs.} \quad \mu > 0. \]  \hspace{1cm} (2)

Problems

\( \mu \) is on the boundary of its parameter space + \( \beta \) is not defined under \( H_0 \).

Solutions


Theoretical solutions + Practical solution
Testing on the boundary of the parameter space

Model:

\[ \propto f(y, \alpha) + \mu g(y, \beta) \quad \mu \geq 0 \quad (3) \]

For now, let \( \beta \) be fixed, the model in (3) is identifiable.

Test

\[ H_0 : \mu = 0 \quad \text{versus} \quad H_1 : \mu > 0 \]

Test statistics*:

\[ LRT = -2 \log[L(0, \hat{\alpha}_0, -)] - L(\hat{\mu}, \hat{\alpha}, \beta)] \quad (4) \]

* for the specific case of \( \beta \) fixed.

\( \mu \) is on the boundary \( \Rightarrow \) WE CAN USE Chernoff, 1954 i.e.:

\[ LRT \xrightarrow{n \to \infty} \frac{1}{2} \chi^2_1 + \frac{1}{2} \delta(0) \quad \text{under } H_0 \quad (5) \]
Testing with non-identifiable parameters

- If $\beta$ fixed, under $H_0$ the LRT is asymptotically $\frac{1}{2} \chi^2_1 + \frac{1}{2} \delta(0)$.
- If we let $\beta$ vary $\Rightarrow$ Under $H_0$, $\{LRT(\beta), \beta \in B\}$ is asymptotically a $\frac{1}{2} \chi^2_1 + \frac{1}{2} \delta(0)$ random process indexed by $\beta$.
- In practice:
  - Define a grid $B_R$ of $R \beta_r$ values over the energy spectrum $B$.
  - $\forall \beta_r \in B_R$ calculate $LRT(\beta_r)$.

We combine the $R$ $LRT(\beta_r)$ values in a unique test statistics...

$$c = \max_{\beta_r \in B_R} LRT(\beta_r)$$

... and we produce a global p-value...

$$P(\sup_{\beta \in B} LRT(\beta) > c) \quad (6)$$

...which we must calculate/approximate somehow!
Approximation of $P(\sup_{\beta \in B} LRT(\beta) > c)$

From [Davies, 1987] we have

$$P(\sup LRT(\beta) > c) \leq \frac{P(\chi^2_1 > c)}{2} + E[N(c)|H_0]$$

where $\approx$ holds if $c \to +\infty$. In Davies, 1987

$$E[N(c)|H_0] = \frac{e^{-\frac{c}{2}}}{\sqrt{2\pi}} \int_L^U \kappa(\beta) d\beta$$

(not easy to deal with).

$\Rightarrow$ use the "empirical" version of [Gross and Vitells, 2010]

$$P(\sup LRT(\beta) > c) \leq \frac{P(\chi^2_1 > c)}{2} + e^{-\frac{c-c_0}{2}} E[N(c_0)|H_0]$$

where $c_0 << c$ and $E[N(c_0)|H_0]$ is estimated using (few) Bootstrap simulations.
Distinguish known astrophysics from new signals - The statistical problem

The model for the known cosmic source is $f(y, \alpha)$; the model for the new source is $g(y, \beta)$; $f \neq g$ for any $\alpha$ and $\beta$.

Is $f$ sufficient to explain the data, or does $g$ provide a better fit?

**Problem**

$f$ and $g$ are non-nested.

**Solutions**

Cox, 1961-1962; Atkinson, 1970; etc., Bootstrap, **next slides**.

**Theoretical solutions**  **Practical solutions**

**Note**

In High Energy Physics (HEP) a discovery is claimed at $5\sigma$ significance. Simulating $O(10^8)$ from a detector might get quite prohibitive.
A new formulation of the problem

Consider a comprehensive model which includes $f(y, \alpha)$ and $g(y, \beta)$ as special cases:

$$(1 - \eta)f(y, \alpha) + \eta g(y, \beta)$$

Thus, considering the model in (9) we test

$H_0 : \eta = 0$ versus $H_1 : \eta > 0$

To exclude intermediate values of $\eta$ we can interchange the roles of the hypotheses and test

$H_0 : \eta = 1$ versus $H_1 : \eta < 1$. 
From a new formulation to a well known problem

Model:

\[(1 - \eta) f(y, \alpha) + \eta g(y, \beta) \] with \[0 \leq \eta \leq 1\] (10)

Test:

\[H_0 : \eta = 0 \quad \text{versus} \quad H_1 : \eta > 0\]

similar argument for \[H_0 : \eta = 1 \quad \text{versus} \quad H_1 : \eta < 1\]

Note!

These are precisely the same issues we encounter when detecting new particles \(\implies\) we already have a solution!
Does it actually work? Let’s see an example...

**Null model:** Power law (Pareto Type I)

\[ f(E, \phi) \propto \phi E^{-(\phi+1)} \]

**Alternative model:** Dark Matter

(from Bergström et al., 1998)

\[ g(E, M_\chi) \propto E^{-1.5} \exp\left\{-7.8 \frac{E}{M_\chi}\right\} \]

**Comprehensive model:**

\[
(1 - \eta) \frac{\phi E^{-(\phi+1)}}{k(\phi)} + \eta \frac{E^{-1.5}}{k(M_\chi)} \exp\left\{-7.8 \frac{E}{M_\chi}\right\}
\]

where \( E, M_\chi \in [1, 100] \) and \( \phi > 0 \).

**Test:** \( H_0 : \eta = 0 \) versus \( H_1 : \eta > 0 \)

What if we have just few events?

Simulation with “not-that-large” $N$

Realistic data analysis

- We simulated 200 events a 5 years observation of putative dark matter source from the Fermi Large Area Telescope (LAT).
- The Fermi LAT is a $\gamma$-ray telescope on board the earth-orbiting Fermi satellite.

**Results**

**Power-law vs. dark matter**

$p$-value $= 2.7 \cdot 10^{-5} \ (\text{sig.} \ 4.038 \sigma)$

**Dark matter vs. power-law**

$p$-value $= 0.528$

- To improve the power of the test one could take into account:
  - $\gamma$-ray directions.
  - Instrumental error.

---


Conclusions and future works

We have presented a two-step solution to compare competing non-nested models:

- **Step 1** - Extend the parameter space of the models to be compared through an additive comprehensive model.
- **Step 2** - Apply Gross and Vitells, 2010 on the model in Step 1.

### Advantages of the procedure

- (Extremely) easy to implement.
- No extensive calculations on a case-by-case basis.
- Computationally more efficient than standard Bootstrap simulations.

### Limitations and future works

- It does not handle multi-dimensional nuisance parameter under $H_1$.
- The nuisance parameter under $H_0$ is required to lie in the interior of its parameter space.
- Improvement of the GV bound w.r.t the dependence structure of the LRT process.
References