Statistical Issues in the Search for Particle Dark Matter

Sara Algeri[1,2]
Supervisors: Prof. David van Dyk[1] and Prof. Jan Conrad[2]

[1]Department of Mathematics, Imperial College London, London, UK
[2]Department of Physics, Stockholm University, Stockholm, Sweden.

CHASC May 19, 2015

Contents

Outline

A shortcut on Dark Matter

- Dark Matter is the substance postulated in the 30s by Fritz Zwicky to explain the evidence of missing mass in the universe.
- It is hypothesized to constitute 85% of the total matter in the universe.
- It has never been observed.
- We do not know what it is made of.
- The best candidate are WIMPs.

How do we detect WIMPs?

- **LHC** $\Rightarrow$ We look for discrepancies in terms of momentum and energy.
- **Direct detection** $\Rightarrow$ We look for a WIMPs-atoms collisions.
- **Indirect detection** $\Rightarrow$ We look for their decay by-products.

The claim of a discovery

- When we look for a new particle (e.g. Higgs boson, quark etc.), we look from the presence of a (5σ) **line/bump** (the signal of the particle) on top of a background flux (what we know).

In the case of the search for Dark Matter, we might have something more complicated than a bump, and we can even have a fake signal (i.e something mimicking WIMPs, but not a background to them)
$\Rightarrow$ **We might have to deal with an entirely new distribution!**

- **E.g.:** In the indirect detection scenario we want to make sure that the gamma rays we observed are due to Dark Matter and not to a different cosmic source (Pulsars for example).

What statistical tool should we use to do so?

Open question that we aim to address with this talk.

Outline

Brief review on LRT

- Suppose we have

$$f(E_i, \mu) \propto 3E_i^{-2} + \mu 1_{\{i=10\}} \text{ with } \mu \geq 0 \tag{1}$$

and we want to test

$$H_0 : \mu = 1 \qquad vs. \quad H_a : \mu \neq 1$$

- **Problem:** I want to know if I have a line of intensity 1 in the 10^{th} energy bin.
- **Solution:** Likelihood Ratio Test (LRT)

$$-2\log\frac{\prod_{i=1}^{n} f(E_i, 1)}{\prod_{i=1}^{n} f(E_i, \hat{\mu}_{MLE})} \overset{H_0}{\sim} \chi^2_1 \tag{2}$$

Important

We can do this just because some regularity conditions hold. Between these, we have:

1. Under H_0, μ is on the interior of the parameter space.
2. The model is identifiable.
3. The models under H_0 and H_a are nested.

What if condition 1. falls?

1. The parameter of interest lies on the boundary of the parameter space.

   E.g:

   - We have

$$f(E_i,\mu) \propto 3E_i^{-2} + \mu 1_{\{i=10\}}, \qquad \mu \geq 0 \tag{3}$$

   - we want to test

$$H_0 : \mu = 0 \qquad vs. \quad H_a : \mu > 0$$

Practical problem:

If there is a line, we know that it occurs at the 10^{th} energy bin. So, how do I check if I have a line there or not?

Theoretical/practical solutions:

Chernoff, 1954† ; Bootstrap.

What if condition 2. falls?

2. There exists a nuisance parameters which is defined just under the alternative model.

E.g:

- We have

$$f(E_i, \mu, \omega) \propto 3E_i^{-2} + \mu 1_{\{i=\omega\}} \quad (4)$$

- we want to test

$$H_0 : \mu = 0 \quad vs. \quad H_a : \mu > 0$$

Practical problem:

Do I have a line somewhere?

Theoretical solution:

Davies, 1987[♮].

Practical solution:

Gross and Vitells, 2010[§].

What if condition 3. falls?

3. The plausible models to be tested are non-nested.

E.g:

- We have

$$f(E_i,\phi) \propto \frac{\phi E_0^{\phi}}{E^{\phi+1}} \quad \text{and} \quad g(E_i, M_\chi) \propto \frac{0.73}{\left(\frac{E}{M_\chi}\right)^{1.5}} \exp\left\{-7.8\frac{E}{M_\chi}\right\} \qquad (5)$$

- we want to test: $H_0 : f(E_i,\phi)$ *vs.* $H_a : g(E_i, M_\chi)$

Practical problem:

Are my particles coming from a power law distributed cosmic source or are they coming from Dark Matter?

Theoretical solution:

Cox, 1961-1962; Atkinson, 1970; etc.

Practical solution:

Hopefully, this talk (using †, ♮ and §) ; Pilla et al., 2005-2006.

The problem in statistical terms

Let $f(y,\alpha)$ and $g(y,\beta)$ such that $f \not\equiv g$ for any α and β

$\Rightarrow$f(y,α) are $g(y,\beta)$ **non-nested models**.

The goal is to develop a test for the hypotheses:

$$H_0 : f(y,\alpha) \text{ is the correct model}$$

versus

$$H_a : g(y,\beta) \text{ is the correct model}$$

Formulation of the problem

- Consider a comprehensive model which includes $f(y, \boldsymbol{\alpha})$ and $g(y, \boldsymbol{\beta})$ as special cases. We have two possibilities:
  - **Multiplicative form**

$$k\{f(y,\boldsymbol{\alpha})\}^{1-\eta}\{g(y,\boldsymbol{\beta})\}^{\eta} \tag{6}$$

where

$$k = \left(\int \{f(y,\boldsymbol{\alpha})\}^{1-\eta}\{g(y,\boldsymbol{\beta})\}^{\eta} dy\right)^{-1} \tag{7}$$

  - **Additive form**

$$(1-\eta)f(y,\boldsymbol{\alpha}) + \eta g(y,\boldsymbol{\beta}) \tag{8}$$

- We prefer the formulation in (8), so that we do not have to worry about dealing with the normalizing constant k.

Thus, considering the model in (8) the test reduces to

$$H_0 : \eta = 0 \quad \text{versus} \quad H_a : \eta > 0 \tag{9}$$

The Likelihood Ratio Statistics

Assume that:

- $\boldsymbol{\alpha}$ lies on the interior of its parametric space
- $\boldsymbol{\beta}$ to be one-dimensional i.e., $\boldsymbol{\beta} = \beta$.

Notice that for β fixed, the model

$$(1-\eta)f(y,\boldsymbol{\alpha}) + \eta g(y,\beta) \qquad 0 \leq \eta \leq 1 \tag{10}$$

is identifiable and thus the only remaining problem when testing $H_0 : \eta = 0$ versus $H_a : \eta > 0$ would be η being on the boundary

$\Rightarrow$ **WE CAN USE** **Chernoff, 1954** i.e.:

$$\text{LRT} = -2\log[L(0,\hat{\boldsymbol{\alpha}}_0,0) - L(\hat{\eta},\hat{\boldsymbol{\alpha}},\beta)] \xrightarrow[n\to\infty]{d} \frac{1}{2}\chi^2_1 + \frac{1}{2}\delta(0) \tag{11}$$

With $L(\cdot)$ being the likelihood function of the model in (10), $\hat{\eta}, \hat{\boldsymbol{\alpha}}$ are the respective ML estimates for $\eta, \boldsymbol{\alpha}$, whereas $\hat{\boldsymbol{\alpha}}_0$ is the MLE for $\boldsymbol{\alpha}$ under H_0.

The Likelihood Ratio Statistics process

- Notice that if β is not fixed then $(\hat{\eta}, \hat{\boldsymbol{\alpha}}, \hat{\beta}) \not\xrightarrow{p} (\eta, \boldsymbol{\alpha}, \beta)$.
- So, the LRT statistics is asymptotically $\frac{1}{2}\chi_1^2 + \frac{1}{2}\delta(0)$ distributed for β fixed.

This means that if we let β vary

$\Rightarrow \{LRT(\beta), \beta \in \mathbf{B}\}$ corresponds asymptotically to a $\frac{1}{2}\chi_1^2 + \frac{1}{2}\delta(0)$ random process indexed by β.

- Thus the p-value of our test $H_0 : \eta = 0$ versus $H_a : \eta > 0$ will correspond to the excursion probability

$$P\left(\sup_{\beta \in \mathbf{B}} LRT(\beta) > c\right) \tag{12}$$

- **How do we calculate/approximate this?**

Approximation of $P\left(\sup_{\beta\in\mathbf{B}} T(\beta) > c\right)$

- From **Davies, 1987** we have that if $\{T(\beta), \beta \in \mathbf{B}\}$ is a χ^2_ν process, then as $c \to +\infty$

$$P(\sup T(\beta) > c) \approx P(\chi^2_\nu > c) + \underbrace{\frac{c^{\frac{\nu-1}{2}} e^{\frac{c}{2}}}{\sqrt{\pi} 2^{\frac{\nu}{2}} \Gamma(\frac{\nu}{2} + \frac{1}{2})} \int_L^U \kappa(\beta) d\beta}_{\text{Expected \# of upcrossings over } c} \qquad (13)$$

- if $c \not\to +\infty \Rightarrow$ we have an upper bound for $P(\sup T(\beta) > c)$.
- $\kappa(\beta)$ is complicated $\Rightarrow$ use the "empirical" version of (13) proposed in **Gross and Vitells, 2010**

$$P(\sup T(\beta) > c) \approx P(\chi^2_\nu > c) + E[N(c_0)|H_0] e^{-\frac{c-c_0}{2}} \left(\frac{c}{c_0}\right)^{\frac{\nu-1}{2}} \qquad (14)$$

where $c_0 << c$, $E[N(c_0)|H_0]$ is the number of upcrossings over c_0 under the null model (to be estimated via Monte Carlo simulation).

Approximation of $P\left(\sup_{\beta \in \mathbf{B}} LRT(\beta) > c\right)$

- So, how do we adjust such results for the case of a $\frac{1}{2}\chi^2_1 + \frac{1}{2}\delta(0)$ random process?

We have that

$$P(\sup LRT(\beta) > c) \approx \frac{p_{Davies}}{2} \approx \frac{p_{GV}}{2} \qquad (15)$$

where p_{Davies} is the approximation in (13) with $\nu = 1$ whereas p_{GV} is the correspective empirical version in (14).

- This holds because of the following result.

Approximation of $P\left(\sup_{\beta\in\mathbf{B}} LRT(\beta) > c\right)$

Result 1.

Let $\{Y_t, t \in \mathbf{T}\}$ *be a stochastic process such that*

$$\forall t \in \mathbf{T},\, Y_t = \begin{cases} W_t, & \text{if } E \text{ occurs} \\ V_t, & \text{if } E^c \text{ occurs} \end{cases}$$

with $W_t \sim \chi^2_{(\nu)}$, $V_t \sim \delta(0)$ *and* $P(E) = P(E^c) = 0.5$. *Then, for* $c \in \mathbb{R}^+$ *we have*

$$P(\sup\{Y_t\} > c) \approx \frac{p_{Davies}}{2}$$

Note: In the case of the our LRT process indexed by β, the event E corresponds to $\frac{d\log L(\eta,\hat{\alpha},\beta)}{d\eta}\Big|_{\eta=0} < 0$, and thus for $n \to \infty$ we have that $P\left(\frac{d\log L(\eta,\hat{\alpha},\beta)}{d\eta}\Big|_{\eta=0} < 0\right) = \frac{1}{2}$.

Approximation of $P\left(\sup_{\beta \in \mathbf{B}} LRT(\beta) > c\right)$

Proof of Result 1.

$$P(\sup\{Y_t\} > c) = P(\sup\{Y_t\} > c|E)P(E) + P(\sup\{Y_t\} > c|E^c)P(E^c)$$

because of total probabilities

$$= P(\sup\{W_t\} > c)\frac{1}{2} + P(\sup\{V_t\} > c)\frac{1}{2}$$

because $c > 0$ always ($c \in \mathbb{R}^+$)

$$= \frac{1}{2}P(\sup\{W_t\} > c)$$

$$\approx \frac{p_{Davies}}{2}$$

because $W_t \sim \chi^2_{(\nu)}$ and c is large.

Note: All we need is the law of total probabilities!

Alternative method, Pilla et al. 2005-2006

The model of reference is again

$$(1-\eta)f(y,\boldsymbol{\alpha})+\eta g(y,\beta) \qquad 0\leq\eta\leq 1 \tag{16}$$

we want to test $H_0 : \eta = 0$ versus $H_a : \eta > 0$.
Now, we focus on the normalized Score function

$$S^{\star}(\boldsymbol{\beta})=\frac{S(\boldsymbol{\beta})}{\sqrt{nC(\boldsymbol{\beta},\boldsymbol{\beta})}} \quad \text{The } \textit{sup} \text{ of this will be our test statistics} \tag{17}$$

and the associated Score process $\{S^{\star}(\boldsymbol{\beta}) \in \mathbf{B}\}$, where

$$S(\boldsymbol{\beta})=\sum_{i=1}^{n}\left[\frac{g(y_i,\boldsymbol{\beta})}{f(y_i,\boldsymbol{\alpha})}-1\right] \tag{18}$$

is the Score function of the model in (16) under H_0 and

$$C(\boldsymbol{\beta},\boldsymbol{\beta}^{\dagger})=\int\frac{g(y_i,\boldsymbol{\beta})g(y_i,\boldsymbol{\beta}^{\dagger})}{f(y_i,\boldsymbol{\alpha})}dy_i-1 \tag{19}$$

is the respective covariance function.

Approximation of $P\left(\sup_{\beta\in\mathbf{B}} S^\star(\beta) > c\right)$, α known (Pilla et al. 2005)

- The p-value of the test is of the form $P(\sup S^\star(\beta) > c)$
- Pilla et al. show that for $n \to +\infty$,

$$P(\sup S^\star(\beta) > c) \to P(\sup Z(\beta) > c) \tag{20}$$

where $Z(\beta)$ is a mean zero Gaussian process.

- Using tube formulae they show that for $c \to \infty$

$$P(\sup Z(\beta) > c) \approx \frac{\xi_0}{A_{d+1}} P(\chi^2_{d+1} \geq c^2) + \sum_{k=1}^{d} \frac{\xi_k}{A_k A_{d+1-k}} P(\chi^2_{d+1-k} \geq c^2) \tag{21}$$

In words: Ratio between the volume of the tube of radius r (function of c) built around the manifold associated to $\sup Z(\beta)$ on a unit sphere, and the volume of the unit sphere itself.

where d is the dimension of β, ξ_j with $j = 0, \dots, d$ are the geometric constants depending on specific model to be tested and $A_w = \frac{2\pi^{w/2}}{\Gamma(w/2)}$.

Approximation of $P\left(\sup_{\beta\in\mathbf{B}} S^{\star}(\beta) > c\right)$, $\boldsymbol{\alpha}$ unknown (Pilla et al. 2006)

- If the nuisance parameter under the null $\boldsymbol{\alpha}$ is unknown, the covariance function of the Score process becomes

$$\underbrace{C^{\star}(\boldsymbol{\beta},\boldsymbol{\beta}^{\dagger}) = C(\boldsymbol{\beta},\boldsymbol{\beta}^{\dagger}) - C(\boldsymbol{\beta}|\boldsymbol{\alpha})'\mathbf{I}^{-1}(\boldsymbol{\alpha})C(\boldsymbol{\beta}^{\dagger}|\boldsymbol{\alpha})}_{\text{This is essentially what is changing!}} \tag{22}$$

  where $\boldsymbol{\alpha}$ is required to lie on the interior of the parameter space, $\mathbf{I}^{-1}(\boldsymbol{\alpha})$ is the inverse of the Fisher information matrix whereas

$$C(\boldsymbol{\beta}|\boldsymbol{\alpha}) = \int g(y_i,\boldsymbol{\beta})\nabla \log f(y_i,\boldsymbol{\alpha})dy.$$

- $\boldsymbol{\alpha}$ is unknown, in the application: $C^{\star}(\boldsymbol{\beta},\boldsymbol{\beta}^{\dagger})$ can be consistely estimated by $C^{\star}(\boldsymbol{\beta},\boldsymbol{\beta}^{\dagger})|_{\boldsymbol{\alpha}=\hat{\boldsymbol{\alpha}}}$.

Outline

Motivating Examples

General Problem

We have n of incoming particles in our detector. We want to know if these n particles come from Dark Matter or from a different cosmic source.

- We assume that the n particles are distributed as a Marked Poisson random process with marks corresponding to the energies, i.e,

$$n \sim Pois(n, \zeta) \qquad E_i \sim h(E_i, \boldsymbol{\theta}) \quad i = 1, \ldots, n; \tag{23}$$

- LIkelihood:

$$L(\theta) = Pois(n, \zeta) \prod_{i=1}^{n} h(E_i, \boldsymbol{\theta}) \propto \prod_{i=1}^{n} h(E_i, \boldsymbol{\theta}) \tag{24}$$

$\Rightarrow$ **WE CAN FOCUS JUST ON THE DISTRIBUTION OF THE MARKS!**

Example 1: Power law vs. Dark Matter

In this example, the goal is to distinguish the Dark Matter signal from a power law distributed cosmic source. The models for the marks are

Power law (Pareto Type I)

$$\frac{\phi E_0^{\phi}}{E^{\phi+1}} \tag{25}$$

Dark Matter (from Bergström et al., 1998)

$$\frac{0.73}{\psi(M_\chi)}\left(\frac{E}{M_\chi}\right)^{-1.5}\exp\left\{-7.8\frac{E}{M_\chi}\right\} \tag{26}$$

$$\text{with} \quad \psi(M_\chi) = \int \frac{0.73}{\psi(M_\chi)}\left(\frac{E}{M_\chi}\right)^{-1.5}\exp\left\{-7.8\frac{E}{M_\chi}\right\}dE$$

where $E \geq E_0$, $E_0 > 0$, $M_\chi \geq E_0$ and $\phi > 0$. In our specific case: $E_0 = 1$, $E, M_\chi \in [1; 100]$.

Example 2: Power law + power law vs. power law + Dark Matter.

We consider a generalized version of Example 1 which includes a background source also distributed as a power law. The signal can be either power law or Dark Matter. The models for the marks become

Power law + power law

$$(1-\lambda)\frac{\delta E_0^{\delta}}{E^{\delta+1}} + \lambda\frac{\phi E_0^{\phi}}{E^{\phi+1}} \tag{27}$$

Power law + Dark Matter

$$(1-\gamma)\frac{\delta E_0^{\delta}}{E^{\delta+1}} + \gamma\frac{0.73}{\psi(M_\chi)}\left(\frac{E}{M_\chi}\right)^{-1.5}\exp\left\{-7.8\frac{E}{M_\chi}\right\} \tag{28}$$

where $E \geq E_0$, $E_0 > 0$, $M_\chi \geq E_0$, $\phi, \delta > 0$ and $0 < \lambda, \eta < 1$. Also in this case E_0 is chosen equal to 1, $E \in [1; 100]$ and $M_\chi \in [1; 100]$.

Example 3: Pulsar Spectrum vs. Dark Matter.

The aim of considering such example is to extend to the statistical framework the difficulty of distinguishing between Dark Matter and pulsar origins discussed in Baltz, 2007. The models in analysis are

Pulsar Spectrum (from Baltz, 2007)

$$\frac{E^{-\rho} \exp\{-(\frac{E}{E_0})^{\tau}\}}{\int E^{-\rho} \exp\{-(\frac{E}{E_0})^{\tau}\} dE} \tag{29}$$

Dark Matter

$$\frac{0.73}{\psi(M_\chi)} \left(\frac{E}{M_\chi}\right)^{-1.5} \exp\left\{-7.8\frac{E}{M_\chi}\right\} \tag{30}$$

where $E > E_0$, $M_\chi \geq E_0$ and $\rho, \tau > 0$.

E_0 is chosen to be equal to 0.1, $E \in [0.1; 5]$, $M_\chi \in [0.1; 5]$ and α will be fixed to 2 to guarantee the two models to be non-nested.

Specify the comprehensive additive model

Let η be the parameter of interest to be tested and let M_χ be the nuisance parameter defined just under the alternative model. Then we have:

- **Power law vs. Dark Matter**

$$(1-\eta)\frac{\phi E_0^{\phi}}{E^{\phi+1}} + \eta\frac{0.73}{\psi(M_\chi)}\left(\frac{E}{M_\chi}\right)^{-1.5}\exp\left\{-7.8\frac{E}{M_\chi}\right\}$$

where $E_0 = 1$, $M_\chi, E \in [1; 100]$, $0 \le \eta \le 1$. ϕ is unknown.

- **Power law + power law vs. power law + Dark Matter**

$$(1-\eta)\left\{(1-\lambda)\frac{\delta E_0^{\delta}}{E^{\delta+1}}+\lambda\frac{\phi E_0^{\phi}}{E^{\phi+1}}\right\}+\eta\left\{(1-\gamma)\frac{\delta E_0^{\delta}}{E^{\delta+1}}+\gamma\frac{0.73}{\psi(M_\chi)}\left(\frac{E}{M_\chi}\right)^{-1.5}\exp\left\{-7.8\frac{E}{M_\chi}\right\}\right\}$$

where $E_0 = 1$, $M_\chi, E \in [1; 100]$, $0 \le \eta \le 1$, $\lambda = \gamma = 0.2$, $\delta = 2$. ϕ is unknown.

- **Pulsar Spectrum vs. Dark Matter**

$$(1-\eta)\frac{E^{-\rho}\exp\{-(\frac{E}{E_0})^{\tau}\}}{\int E^{-\rho}\exp\{-(\frac{E}{E_0})^{\tau}\}dE} + \eta\frac{0.73}{\psi(M_\chi)}\left(\frac{E}{M_\chi}\right)^{-1.5}\exp\left\{-7.8\frac{E}{M_\chi}\right\}$$

where $E_0 = 0.1$, $M_\chi, E \in [0.1; 5]$, $0 \le \eta \le 1$, $\rho = 4/3$, $\tau = 2$.

Verify that Chernoff, 1954 applies

- We have seen above that if M_χ is fixed

$$LRT \xrightarrow{d} \frac{1}{2}\chi^2_1 + \frac{1}{2}\delta(0) \qquad \text{when} \qquad n \to +\infty$$

- How big must n be to guarantee that this result holds?
- Simulate from some values for LRT and from $\frac{1}{2}\chi^2_1 + \frac{1}{2}\delta(0)$. Compare them using:
  - Wilcoxon rank sum test
  - qq-plots
  - histograms
  - empirical cdf graphs
  - etc.

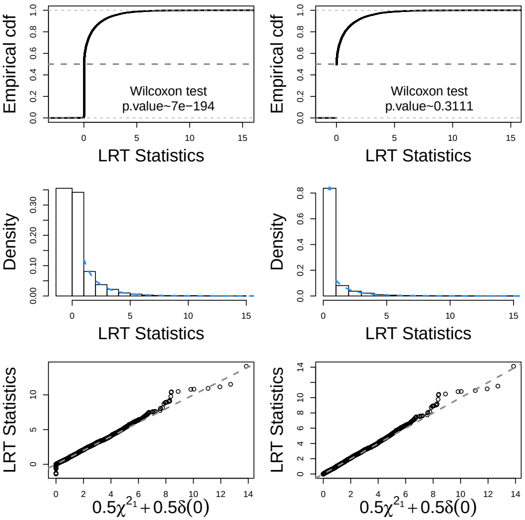


Figure 1: The left panels refer to the the original simulated values for the LRT statistics. The right panels refer to the adjusted LRT statistics obtained imputing the negative values to be equal to 0 in order to correct for the effect of floating points.

Compute the approximate p-values

- LRT-based method

$$P(\sup LRT(M_\chi) > c) \approx \frac{P(\chi^2_1 > c)}{2} + \underbrace{E[N(c_0)|H_0]}_{\text{we need this}} e^{-\frac{c-c_0}{2}} \left(\frac{c}{c_0}\right)^{\frac{\nu-1}{2}} \quad (31)$$

- Pilla et al., 2005-2006

$$P(\sup Z(M_\chi) > c) \approx \frac{\overbrace{\xi_0}^{\text{we need this}}}{2\pi} P(\chi^2_2 \geq c^2) + \frac{1}{2} P(\chi^2_1 \geq c^2) \quad (32)$$

Computing $E[N(c_0)|H_0]$

Model	Fixed parameters	Unknown parameters	Sample size	$E[N(c_0)|H_0]$ $\mathbf{c_0} = 0.1$
Power law vs. Dark Matter	$E_0 = 1$	ϕ, M_χ	5000	0.906
Power law + power law vs. power law + Dark Matter	$E_0 = 1$ $\lambda = \gamma = 0.2$ $\delta = 2$	ϕ, M_χ	5000	0.867
Pulsar Spectrum vs Dark Matter	$E_0 = 0.1$ $\rho = \frac{4}{3}$ $\tau = 2$	M_χ	1000	0.219

Table 1: Estimated number of upcrossings of the process $LRT(M_\chi)$ assuming the null model to be true (i.e., $E[N(c_0)|H_0]$). For all the models 1000 Monte Carlo simulations has been generated. A grid of resolution 100 for the parameter M_χ over the range [1; 100] has been considered for the first two models; whereas a grid of size 20 over the range [0.1;5] was selected for the model in our third example. For all the three cases, the threshold c_0 has been set to 0.1. The sample size has been chosen large enough to guaranteed the $\frac{1}{2}\chi^2_1 + \frac{1}{2}\delta(0)$ for fixed values of M_χ as discussed in the previous slide.

Formulae for ξ_0

- If the null model is completely specified

$$\xi_0 = \int_{\mathbf{M}_\chi} \frac{\sqrt{\left| C(M_\chi, M_\chi^\dagger) \frac{d^2 C(M_\chi, M_\chi^\dagger)}{dM_\chi dM_\chi^\dagger} - \frac{dC(M_\chi, M_\chi^\dagger)}{dM_\chi} \frac{dC(M_\chi, M_\chi^\dagger)}{dM_\chi^\dagger} \right|_{M_\chi^\dagger = M_\chi}}}{C(M_\chi, M_\chi^\dagger)} dM_\chi \tag{33}$$

- If it is not

$$\hat{\xi}_0 = \int_{\mathbf{M}_\chi} \sqrt{\left. \frac{d^2 \rho^\star(M_\chi, M_\chi^\dagger)}{dM_\chi dM_\chi^\dagger} \right|_{M_\chi^\dagger = M_\chi, \boldsymbol{\alpha} = \hat{\boldsymbol{\alpha}}}} dM_\chi \tag{34}$$

with $\rho^\star(M_\chi, M_\chi^\dagger)$ being

$$\rho^\star(M_\chi, M_\chi^\dagger) = \frac{C^\star(M_\chi, M_\chi^\dagger)}{\sqrt{C^\star(M_\chi, M_\chi) C^\star(M_\chi^\dagger, M_\chi^\dagger)}}.$$

They look pretty complicated!

But notice that to compute them we only need the covariance function of the Score process and a good numerical algorithm to solve the integrals.

Computing ξ_0

Model	**Fixed parameters**	**Unknown parameters**	ξ_0
Power law vs. Dark Matter	$E_0 = 1$	ϕ, M_χ	5.3379
Power law + power law vs. power law + Dark Matter	$E_0 = 1$ $\lambda = \gamma = 0.2$ $\delta = 2$	ϕ, M_χ	5.3635
Pulsar Spectrum vs Dark Matter	$E_0 = 0.1$ $\rho = \frac{4}{3}$ $\tau = 2$	M_χ	2.7397

Table 2: Geometric constants ξ_0. When the null model ($\eta = 0$) is fully specified, ξ_0 is calculated according to equation (33). When a nuisance parameter is present under the null model, its estimate is provided via MLE and ξ_0 is calculated according to equation (34).

Outline

Power law vs. Dark Matter

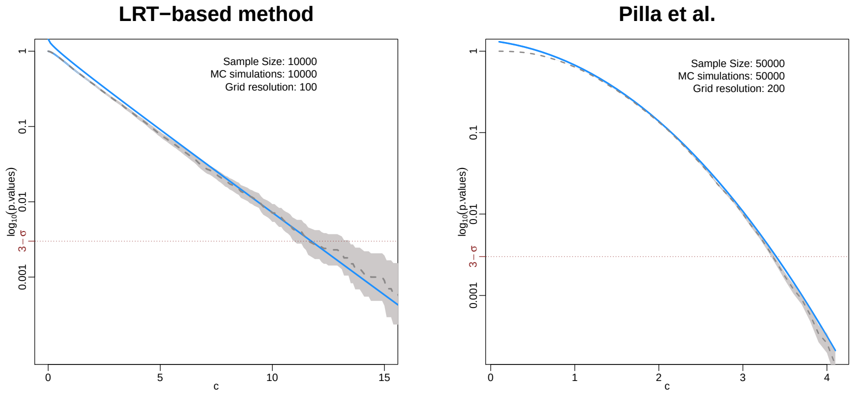


Figure 2: Blue curves: Approximation for $P(\sup LRT(M_\chi) > c)$ (left panel) and $P(\sup S^\star(M_\chi) > c)$ (right panel). Gray dotted curve: Monte Carlo p-values. Gray area: Monte Carlo errors.

Power law + power law vs. Power law + Dark Matter

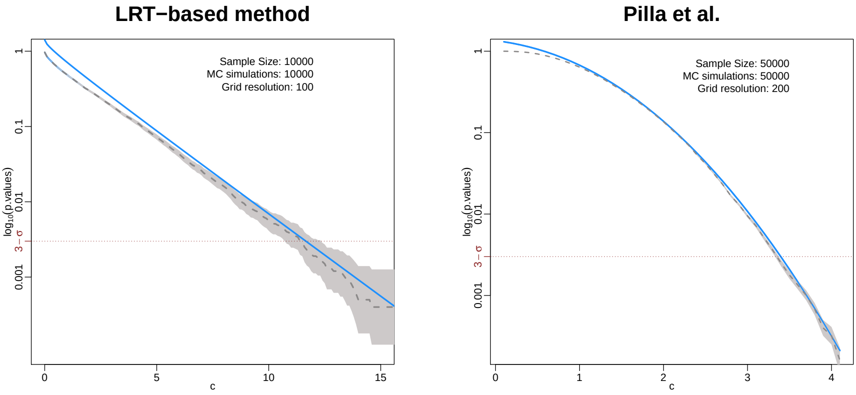


Figure 3: Blue curves: Approximation for $P(\sup LRT(M_\chi) > c)$ (left panel) and $P(\sup S^\star(M_\chi) > c)$ (right panel). Gray dotted curve: Monte Carlo p-values. Gray area: Monte Carlo errors.

Pulsar Spectrum vs. Dark Matter

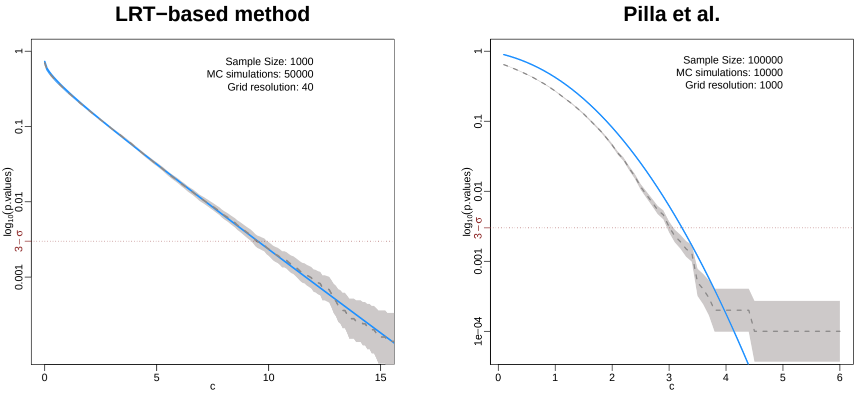
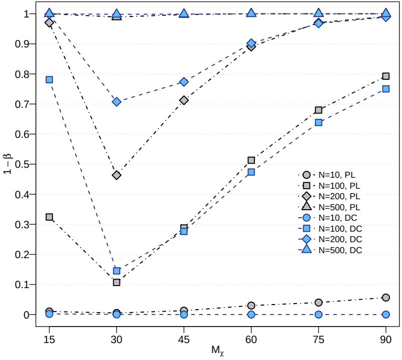


Figure 4: Blue curves: Approximation for $P(\sup LRT(M_\chi) > c)$ (left panel) and $P(\sup S^\star(M_\chi) > c)$ (right panel). Gray dotted curve: Monte Carlo p-values. Gray area: Monte Carlo errors.

Power/Type I error comparison



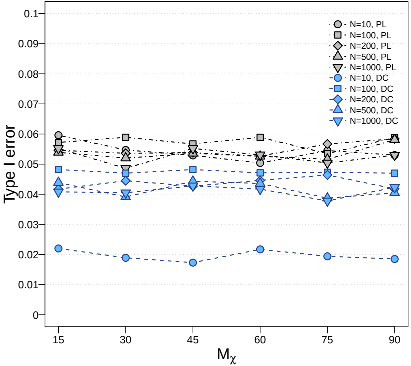


Figure 5: Power function at 3σ (left panel) and Type I error at 2σ (right panel) at different values of M_χ for the model power law vs. Dark Matter.

Outline

Which method is better when?

We identified two methods to pursue a test for non-nested models. Which one is better when?

Pros LRT-based method	Cons Pilla et al. 2005-2006
-If α is unknown the theory does not change.	-If α is unknown the theory changes.
-We can do intermediate checks .	-Intermediate checks cannot be done easily.
-The theory is fairly simple.	-The theory is quite complicated.
-It appears more powerful.	-It appears less powerful.
-Lower Type I error.	-Higher Type I error.
-It requires smaller n to reach the asymptotic.	-It requires larger n to reach the asymptotic.
-It requires Monte Carlo simulations.	-It requires numerical integrations.
-It works for Pulsar Spectrum vs. Dark Matter.	-It does not work for Pulsar Spectrum vs. Dark Matter.
-If c small, we still have an upper bound.	

Cons LRT-based method	Pros Pilla et al. 2005-2006
-It cannot be applied (yet) if β is multidimensional.	- if β is multidimensional the theory does not change.
-To simulate from the null model might not be easy.	-Numerical integrations might be simpler.

Outline

Future developments

- Refinements of the LRT-based method.
  - Loosen the assumption on β being one-dimensional.
  - Loosen the assumption on α being in the interior of the parameter space.
- Evaluate the effect of the resolution of the grid for the nuisance parameter under the alternative model (i.e., M_χ in our three examples).
- Apply both the LRT-based and the Score-based methods to real data taking in account the measurement of the error.
- Identify a Bayesian solution for testing non-nested models and compare it the LRT-based approach proposed.
- Build an R package to implement the procedures presented in this talk.

References

- A. C. Atkinson. *"A Method For Discriminating Between Models"*. In: Journal of the Royal Statistical Society. Series B (Methodological) 32.3 (1970).
- E. A. Baltz, J. E. Taylor, and L. L. Wai. *"Can Astrophysical Gamma-Ray Sources Mimic Dark Matter Annihilation in Galactic Satellites?"*. In: The Astrophysical Journal Letters 659.2 (2007).
- L. Bergström, p. Ullio, and J.H. Buckley. *"Observability of rays from dark matter neutralino annihilations in the Milky Way halo"*. In: Astroparticle Physics 9.2 (1998).
- H. Chernoff. *"On the Distribution of the Likelihood Ratio"*. In: The Annals of Mathematical Statistics 25.3 (1954).
- D. R. Cox. *"Tests of Separate Families of Hypotheses"*. In: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics. (1961).
- D. R. Cox. *"Further Results on Tests of Separate Families of Hypotheses"*. In: Journal of the Royal Statistical Society. Series B (Methodological) 24.2 (1962).
- R. B. Davies. *"Hypothesis Testing when a Nuisance Parameter is Present Only Under the Alternatives"*. In: Biometrika 74.1 (1987).
- E. Gross and O. Vitells. *"Trial factors for the look elsewhere effect in high energy physics"*. In: The European Physical Journal C 70.1-2 (2010), pp. 525-530.
- R.S. Pilla and C. Loader. *"Inference in Perturbation Models, Finite Mixtures and Scan Statistics: The Volume-of-Tube Formula"*. In: ArXiv Mathematics e-prints (Nov. 2005).
- R. S. Pilla, C. Loader, and C.C. Taylor. *"New Technique for Finding Needles in Haystacks: Geometric Approach to Distinguishing between a New Source and Random Fluctuations"*. In: Phys. Rev. Lett. 95 (23) (2006).


Thank you