Multiple hypothesis testing and testing one hypothesis multiple times: two sides of the same coin?

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

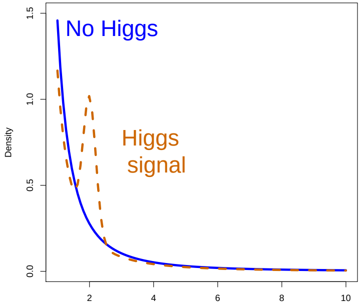
[1]Department of Mathematics, Imperial College London, London, UK
[2]Department of Physics, Stockholm University, Stockholm, Sweden.

CHASC - May 10, 2016

General framework

Goal of statistical signal detection in physics

We would like to distinguish signals of new physics phenomena from the random fluctuations of the data.



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

How does statistics tackle this problem?

- Approach 1:
  Multiple hypotesis testing $\Rightarrow$ Bonferroni's correction.
- Approach 2:
  Simulations $\Rightarrow$ Monte Carlo, Bootstrap.
- Approach 3:
  Hypothesis testing when a nuisance parameter is present only under the alternative $\Rightarrow$ Davies (1977, 1987), Gross and Vitells (2010).

$\downarrow$

We refer to this as **Testing one hypothesis multiple times**.

Note!

In High Energy Physics a discovery is claimed at $\mathbf{5}\sigma$ significance
$\Rightarrow$ in Approach 2 we need to simulate $O(10^8)$, **can we avoid that?**
Yes! Use (responsibly) Approach 1 and/or Approach 3.

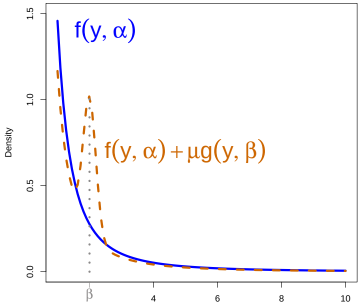
Questions I would like to address with this talk

1. What does it mean exactly to "test one hypothesis multiple times", and in what sense is it equivalent to a testing problem when a nuisance parameter is present only under the alternative?
2. Can we tackle both nested and non-nested models with this approach?
3. What is the difference between testing one hypothesis multiple times and multiple hypothesis testing?
4. When do multiple hypothesis testing and testing one hypothesis multiple times coincide in some sense?
5. What else can we do, and what is the potential of working in this direction?

Outline

1. What does it mean exactly to "test one hypothesis multiple times", and in what sense is it equivalent to a testing problem when a nuisance parameter is present only under the alternative?
2. Can we tackle both nested and non-nested models with this approach?
3. What is the difference between testing one hypothesis multiple times and multiple hypothesis testing?
4. When do multiple hypothesis testing and testing one hypothesis multiple times coincide in some sense?
5. What else can we do and what is the potential of working in this direction?

A statistical framework for a physics problem



The model of interest is proportional to

$$\underbrace{f(y,\alpha)}_{\text{background}} + \underbrace{\mu}_{\text{signal strength}} \underbrace{g(y, \overbrace{\beta}^{\text{signal location}})}_{\text{bump}} \qquad (1)$$

and we test

$$H_0 : \mu = 0 \quad \text{vs.} \quad \mu > 0. \qquad (2)$$

Problems

μ is on the boundary of its parameter space + β is not defined under H_0.

Solutions

$\underbrace{\text{Chernoff, 1954 + Davies, 1987}}_{\text{Theoretical solutions}},\underbrace{\text{Gross and Vitells, 2010}}_{\text{Practical solution}}.$

Testing on the boundary of the parameter space

- **Model:**

$$\propto f(y,\alpha)+\mu g(y,\beta) \qquad \mu\geq 0 \tag{3}$$

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

- **Test**

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

- **Test statistics*:**

$$LRT = -2\log[\underbrace{L(0,\hat{\alpha}_0,\text{-})}_{\text{Likelihood under } H_0} - \underbrace{L(\hat{\mu},\hat{\alpha},\beta)}_{\text{Likelihood under } H_1}] \tag{4}$$

* for the specific case of β be fixed.

- η is on the boundary $\Rightarrow$ WE CAN USE Chernoff, 1954 i.e.:

$$LRT \xrightarrow[n\to\infty]{d} \frac{1}{2}\chi^2_1 + \frac{1}{2}\delta(0) \quad \text{under } H_0 \tag{5}$$

Testing one hypothesis multiple times (1)

- If β fixed, under H_0 the LRT is asymptotically $\frac{1}{2}\chi_1^2 + \frac{1}{2}\delta(0)$.
- If we let β vary$\Rightarrow$ Under H_0, $\{LRT(\beta), \beta \in \mathbf{B}\}$ is asymptotically a $\frac{1}{2}\chi_1^2 + \frac{1}{2}\delta(0)$ random process indexed by β.
- In practice:
  - Define a grid $\mathbf{B}_R$ of R β_r values over the energy spectrum $\mathbf{B}$.
  - $\forall \beta_r \in \mathbf{B}_R$ calculate $LRT(\beta_r)$.

Many "sub"-alternatives...

It is like if we had many alternative hypothesis $H_{11}, \ldots, H_{1r}, \ldots, H_{1R}$, one for each value $\beta_r \in \mathbf{B}_R$, and for each of them we have one value $LRT(\beta_r)$.

...but yet just one test statistic...

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

$$\max_{\beta_r \in \mathbf{B}_R} LRT(\beta_r)$$

Testing one hypothesis multiple times (2)

... and one global p-value...

The **p-value** of our test $H_0 : \eta = 0$ versus $H_a : \eta > 0$ is in the form

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

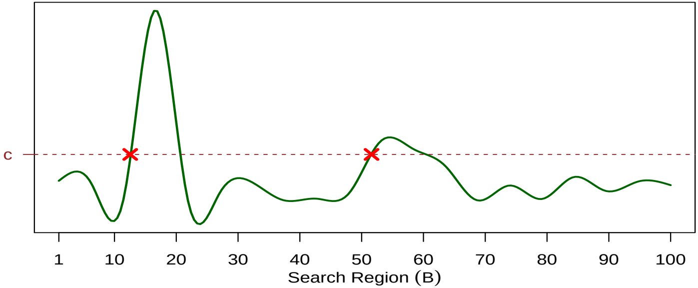
with $c = \max_{\beta_r \in \mathbf{B}_R} LRT(\beta)$.

...which we must calculate/approximate somehow!

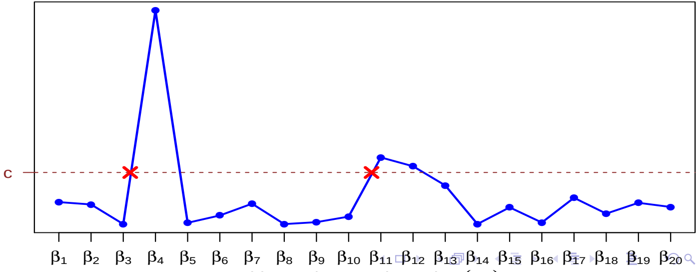
To do so, we first need to introduce the concept of **upcrossings** of the LRT-process $\{LRT(\beta), \beta \in \mathbf{B}\}$.

What do we mean by "upcrossings"?

True LRT-process under H_0



Discretized version we deal with in practice



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

- From **Davies, 1987** we have that if $\{LRT(\beta), \beta \in \mathbf{B}\}$ is a "regular" χ^2_1 process, then as $c \to +\infty$

$$P(\sup LRT(\beta) > c) \approx \frac{P(\chi^2_1 > c)}{2} + \frac{e^{\frac{c}{2}}}{\sqrt{2\pi}} \int_L^U \kappa(\beta) d\beta \quad (7)$$

Expected # of upcrossings over c of the LRT process under H_0

- if $c \not\to +\infty \Rightarrow$ we have an upper bound for $P(\sup LRT(\beta) > c)$.
- $\kappa(\beta)$ is complicated $\Rightarrow$ use the "empirical" version of (7) proposed in **Gross and Vitells, 2010**

$$P(\sup LRT(\beta) > c) \approx \frac{P(\chi^2_1 > c)}{2} + \underbrace{e^{-\frac{c-c_0}{2}} E[N(c_0)|H_0]}_{=E[N(c)|H_0]} \quad (8)$$

Expected # of upcrossings over c_0 of the LRT process under H_0

- where $c_0 << c$ and $E[N(c_0)|H_0]$ is estimated using (few) Bootstrap simulations.

For more details and an alternative approach to the problem, check out:

- Algeri S., van Dyk D.A., Conrad J., Brandon, A. *Looking for a Needle in a Haystack? Look Elsewhere! A statistical comparison of approximate global p-values.* Submitted, 2016.

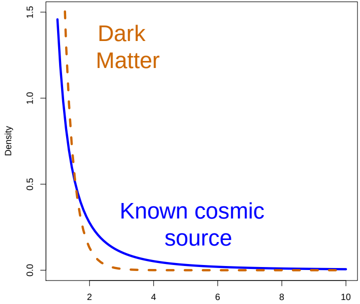
Outline

1. What does it mean exactly to "test one hypothesis multiple times", and in what sense is it equivalent to a testing problem when a nuisance parameter is present only under the alternative?
2. Can we tackle both nested and non-nested models with this approach?
3. What is the difference between testing one hypothesis multiple times and multiple hypothesis testing?
4. When do multiple hypothesis testing and testing one hypothesis multiple times coincide in some sense?
5. What else can we do and what is the potential of working in this direction?

Non-nested models comparison in physics

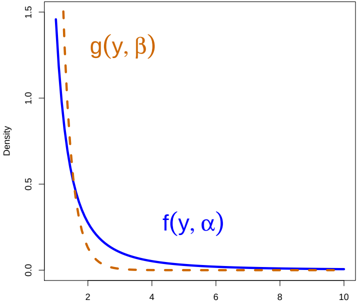
Goal

We would like to distinguish known astrophysics from new signals.



- **E.g.,** Dark Matter.
- We wish to distinguish a dark matter signal from a "fake" signal that mimics it.

The statistical problem



- The model for the know cosmic source is $f(y, \alpha)$;
- The model for the new source is $g(y, \beta)$;
- $f \not\equiv g$ for any α and β.

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 two slides.**.

Theoretical solutions (Cox, 1961-1962, Atkinson, 1970; etc.) — Practical solutions (Bootstrap, next two slides)

Formulation of the problem

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

$$\propto \{f(y, \alpha)\}^{1-\eta}\{g(y, \beta)\}^{\eta} \tag{9}$$

  - **Additive form**

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

- We prefer (10), it avoids the need to deal with the normalizing constant.

Thus, considering the model in (10) we test

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

- To exclude intermediate values of η we can interchange the roles of the hypotheses and test

$$H_0 : \eta = 1 \quad \text{versus} \quad H_1 : \eta < 1.$$

From a new formulation to a well known problem

Model:

$$(1 - \underbrace{\eta}_{\text{Tested on the boundary}})f(y,\alpha) + \eta g(y, \underbrace{\beta}_{\text{Not defined under } H_0}) \quad \text{with} \quad 0 \leq \eta \leq 1 \qquad (11)$$

Test:

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

similar argument for $H_0 : \eta = 1$ versus $H_1 : \eta < 1$

Note!

These are precisely the same issues we encounter when detecting new particles, i.e., when testing one hypothesis multiple times

$\Longrightarrow$ **we already have a solution!**

For more details, check out:

- Algeri S., Conrad J., van Dyk D.A. *A method for comparing non-nested models with application to astrophysical searches for new physics.* MNRAS: Letters, 2016.
- Algeri S., R package 'NONnest', 2015.

Outline

1. What does it mean exactly to “test one hypothesis multiple times”, and in what sense is it equivalent to a testing problem when a nuisance parameter is present only under the alternative?
2. Can we tackle both nested and non-nested models with this approach?
3. What is the difference between testing one hypothesis multiple times and multiple hypothesis testing?
4. When do multiple hypothesis testing and testing one hypothesis multiple times coincide in some sense?
5. What else can we do and what is the potential of working in this direction?

Multiple hypothesis testing - Framework

Also in this case:

- We define a grid $\mathbf{B}_R$ of R β_r values over the energy spectrum $\mathbf{B}$.
- $\forall \beta_r \in \mathbf{B}_R$ calculate $LRT(\beta_r)$.

However, now we have:

Many ~~sub-~~alternatives...

We have many alternative hypothesis $H_{11}, \ldots, H_{1r}, \ldots, H_{1R}$, one for each value $\beta_r \in \mathbf{B}_R$.

...many test statistics...

$\forall \beta_r \in \mathbf{B}_R$ we have one test statistics $LRT(\beta_r)$, and such that $LRT(\beta_r) \sim \frac{1}{2}\chi^2_1 + \frac{1}{2}\delta(0)$ asymptotically.

...many p-values!

$\forall \beta_r \in \mathbf{B}_R$ we have $p_r = \frac{P(\chi^2_1 > LRT(\beta_r))}{2}$.

Local p-values and type I error

- We have an ensemble of R local p-values $p_1, \ldots, p_r, \ldots, p_R$.
- The smallest, names p_{L} is then compared with the target probability of type I error α_{L}.
- But what is α_{L} if we want to claim a discovery at 5σ?

Global and local probability of false detection

$\alpha_{\mathrm{L}} =$ specific probability of false detection for each of the R

$\neq$

$\alpha_{\mathrm{G}} =$ probability of having at least one false detection over the whole ensemble of R tests.

$\Rightarrow$ **we must correct p_{L} accordingly**

Local p-values corrections

- **If the R tests were independent**

$$\alpha_{\mathrm{G}} = 1 - (1 - \alpha_{\mathrm{L}})^R \quad \Rightarrow \quad p_{\mathrm{G}} = 1 - (1 - p_{\mathrm{L}})^R \tag{12}$$

E.g.: Suppose we are conducting $R = 50$ simultaneous test, each of them at 5σ

$$\alpha_{\mathrm{L}} = 1 - \Phi(5) \quad \Rightarrow \text{ by (11): } \quad \alpha_{\mathrm{G}} = 1 - \Phi(4.18)$$

i.e., $\frac{\alpha_{\mathrm{G}}}{\alpha_{\mathrm{L}}} \approx 50$.

- **If the R tests were dependent** (which is generally the case)

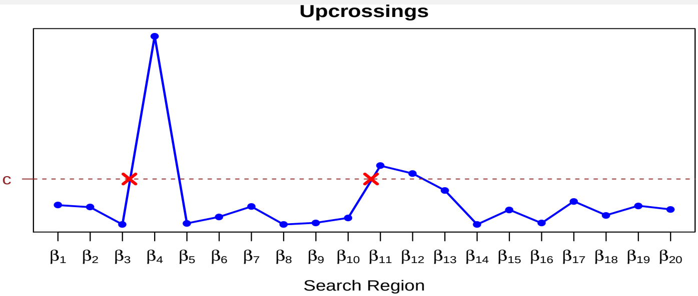
$$\alpha_{\mathrm{G}} \leq R\alpha_{\mathrm{L}} \quad \Rightarrow \quad \boxed{p_{\mathrm{B}F} = Rp_{\mathrm{L}} \quad \text{Bonferroni's correction}} \tag{13}$$

Outline

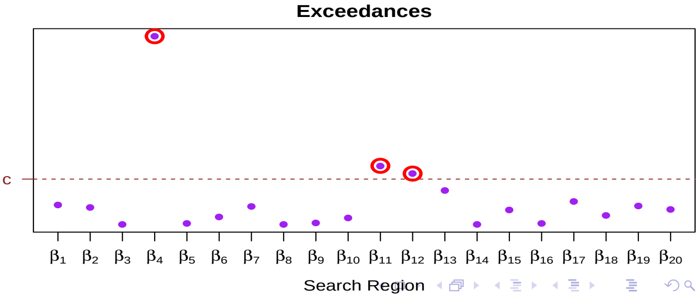
1. What does it mean exactly to "test one hypothesis multiple times", and in what sense is it equivalent to a testing problem when a nuisance parameter is present only under the alternative?
2. Can we tackle both nested and non-nested models with this approach?
3. What is the difference between testing one hypothesis multiple times and multiple hypothesis testing?
4. When do multiple hypothesis testing and testing one hypothesis multiple times coincide in some sense?
5. What else can we do and what is the potential of working in this direction?

Upcrossings and Exceedances

Discretized LRT-process under H_0



Multiple LRTs under H_0



Why are we interested in the Exceedances?

We can identify situations where the average number of exceedances under H_0, namely $E[N_c^\star|H_0]$, and the average number of upcrossings under H_0, $E[N_c|H_0]$ are approximately equal.

- We will soon see two conditions we need for this to happen.
- For now let's focus on $E[N_c^\star|H_0]$:

$$E[N_c^\star|H_0] = \sum_{r=1}^{R} 1 \cdot P(LRT(\beta_r) > c)$$

$$\text{under } H_0,\ \forall \beta_r \in \mathbf{B}_R,$$

$$LRT(\beta_r) \sim \frac{1}{2}\chi_1^2 + \frac{1}{2}\delta(0) \text{ asymptotically}$$

$$= \sum_{r=1}^{R} \frac{P(\chi_1^2 > c)}{2} = R\frac{P(\chi_1^2 > c)}{2} = Rp_{\mathrm{L}} = p_{\mathrm{B}F}$$

Bonferroni's correction!

Two sides of the same coin

What the previous slide is telling us is that, if

$$\underbrace{E[N_c|H_0]}_{\substack{\text{Expected}\\ \#\text{ upcrossings}\\ \text{under } H_0}} \approx \underbrace{E[N_c^\star|H_0]}_{\substack{\text{Expected}\\ \#\text{ exceedances}\\ \text{under } H_0}} = \underbrace{p_{\mathrm{B}F}}_{\substack{\text{Bonferroni's}\\ \text{correction}}} \tag{14}$$

and $\exists\lambda$ s.t as $c \to +\infty \quad (p_{\mathrm{L}} \to 0)$ and $R \to +\infty$

$$E[N_c|H_0] \approx E[N_c^\star|H_0] = p_{\mathrm{B}F} \to \lambda$$

then, for R and c large we have

$$\begin{aligned}\underbrace{P(\sup LRT(\beta) > c)}_{\text{Global p-value}} &\approx \underbrace{\frac{P(\chi^2_1 > c)}{2}}_{\substack{\to 0\\ \text{as } c\to+\infty}} + E[N_c|H_0] \\ &\approx E[N_c|H_0] \approx E[N_c^\star|H_0] \\ &\approx p_{\mathrm{B}F} \quad \text{Bonferroni adjusted local p-value}\end{aligned} \tag{15}$$

This means that if $E[N_c|H_0] \approx E[N_c^\star|H_0]$, then testing one hypothesis multiple times and multiple hypothesis testing will lead to approximately the same inference. (But, since the latter is much quicker than the former, I might gain in computing time.)

When do we have $E[N_c|H_0] \approx E[N_c^\star|H_0]$?

To guarantee $E[N_c|H_0] \approx E[N_c^\star|H_0]$ (as $c \to +\infty$), we need the following two conditions to be satisfied:

1. **Long range independence**

$$|F_{1,\dots,r,r+1,\dots,r+k} - F_{1,\dots,r}F_{r+1,\dots,r+k}| \leq q(r) \tag{16}$$

where $F(\cdot)$ is the cdf of $LRT(\beta_r), \forall \beta_r \in \mathbf{B}_R$, and $q(r)$ is a function such that $q(r) \to 0$ as $r \to \infty$.

This condition implies that independence is achieved for distant points β_r of the (discretized) energy/mass spectrum.

2. **Local dependence**

$$\limsup R \sum_{r=2}^{[R/l]} P(LRT(\beta_1) > c, LRT(\beta_r) > c) \to 0 \quad \text{as} \quad l \to +\infty \tag{17}$$

where $F(\cdot)$ be the cdf of $LRT(\beta_r), \forall \beta_r \in \mathbf{B}_R$,

This condition excludes the chance of clustering of the upcrossings of the LRT-process.

How to assess if these two conditions hold?

- Let the model of reference be $(1-\eta)f(y,\alpha)+\eta g(y,\beta_r)$, and let $l(\eta|\alpha,\beta_r,y)$ be its log-likelihood.
- $\forall\beta_r$ the score function evaluated at H_0 is $S(\beta_r)=\frac{\partial l(\eta|\alpha,\beta_r,y)}{\partial\eta}\Big|_{\eta=0}$
  $\Rightarrow$ the score process under H_0 is $\{S(\beta_r),\beta_r\in\mathbf{B}_r\}$
- with covariance function is $cov(S(\beta_r),S(\beta_t))=\int\frac{g(y,\beta_r)g(y,\beta_t)}{f(y,\alpha)}\partial y-1$

$$S^\star(\beta_r)=\frac{S(\beta_r)}{\sqrt{cov(S(\beta_r),S(\beta_r))}} \tag{18}$$

A sufficient condition on $S^\star(\beta_r)$ (Berman's condition)

If the covariance function of $S^\star(\beta_r)$ satisfies

$$\sup_{|\beta_r-\beta_t|>\tau}|cov(S^\star(\beta_r),S^\star(\beta_t))|\log(\tau)\to 0 \quad \text{as} \quad \tau\to+\infty \tag{19}$$

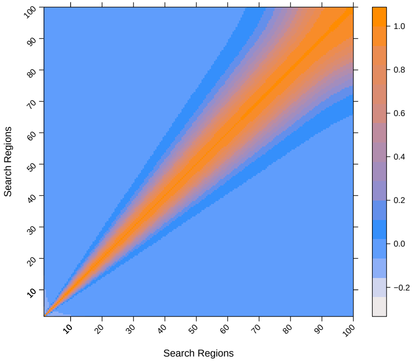
then **long range independence** and **local independence** hold on both the normalized score and the LRT processes.

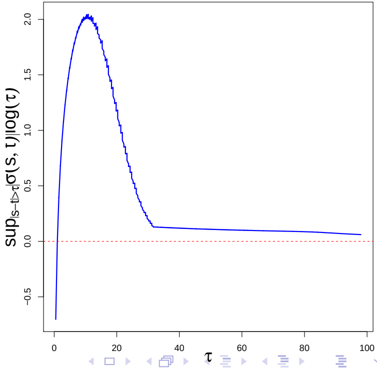
Example

Consider a power-law distributed background with index ψ and a Gaussian signal with dispersion proportional to the signal location.
The full model is

$$(1-\eta)\frac{1}{k_{\psi}y^{\psi+1}}+\frac{\eta}{k_{M_{\chi}}}\exp\left\{-\frac{(y-M_{\chi})^2}{0.02M_{\chi}^2}\right\} \tag{20}$$

with k_{ψ} and $k_{M_{\chi}}$ normalizing constants, $y \in [1;100]$, $\psi = 1.4$ and $M_{\chi} \in [1;100]$.





Realistic data analysis

We simulated observation of monochromatic feature by the Fermi Large Area Telescope (LAT).

- 2391 events from an astrophysical background corresponding to isotropic emission following a spectral power-law with index 2.4, i.e., $\psi = 1.4$.
- 64 events from a Gaussian signal with mass of 35 GeV.
- 80 energy bins, spaced equally from 10-350 GeV.

Method	Signal Location	Signal Strength	Sig.
Unadjusted local	35.82	0.042	5.920σ
Bonferroni adj. local	35.82	0.042	5.152σ
Gross & Vitells	35.82	0.042	5.192σ

Outline

1. What does it mean exactly to "test one hypothesis multiple times", and in what sense is it equivalent to a testing problem when a nuisance parameter is present only under the alternative?
2. Can we tackle both nested and non-nested models with this approach?
3. What is the difference between testing one hypothesis multiple times and multiple hypothesis testing?
4. When do multiple hypothesis testing and testing one hypothesis multiple times coincide in some sense?
5. What else can we do and what is the potential of working in this direction?

What can we do more?

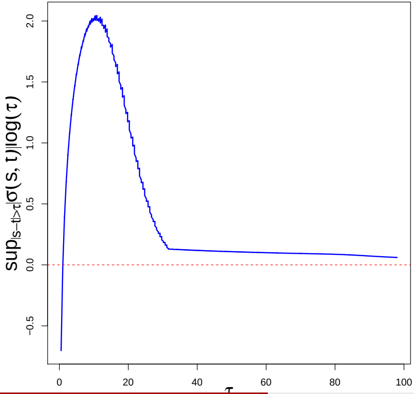
- Berman's condition is not only a sufficient condition to guarantee asymptotic equivalence between testing one hypothesis multiple times and multiple hypothesis testing.
- Indeed, it can be used as diagnostic tool to assess the validity of the Davies (1987) and Gross and Vitells (2010) approximations for the global p-value $P(\sup LRT(\beta) > c)$.
- Several cases can be identified and additional conditions, in addition to long range independence and local independence, are needed.
- But we still have to refine the details...
- ...however, we already can take a look at some examples.

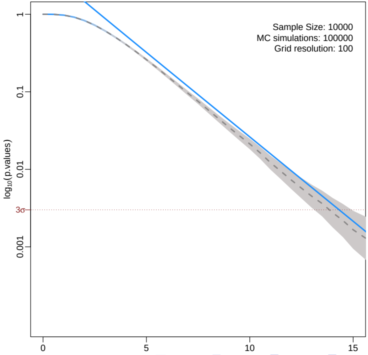
A case where everything works nicely

Considering again the Power Law background + Gaussian signal example:

$$(1-\eta)\frac{1}{k_\psi y^{\psi+1}} + \frac{\eta}{k_{M_\chi}} \exp\left\{-\frac{(y-M_\chi)^2}{0.02M_\chi^2}\right\} \tag{21}$$

with k_ψ and k_{M_χ} normalizing constants, $y \in [1;100]$, $\psi = 1.4$ and $M_\chi \in [1;100]$.



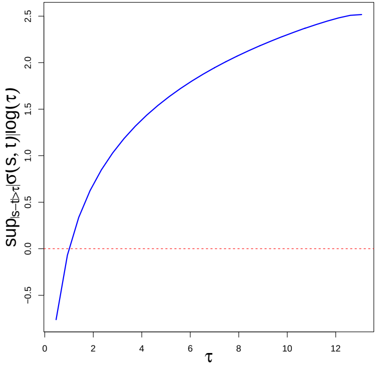


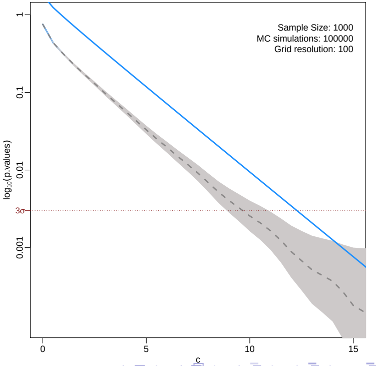
A non-ideal case

Suppose we want to distinguish between Pulsar Spectrum and Dark Matter. The full model is:

$$(1-\eta)\frac{\exp\{-y^2\}}{k_\rho y^\rho} + \frac{\eta\exp\{-7.8\frac{y}{\phi}\}}{k_\phi y^{1.5}} \tag{22}$$

with k_ρ and k_ϕ normalizing constants, $y \in [1;15]$, $\rho = 4/3$ and $\phi \in [1;15]$.



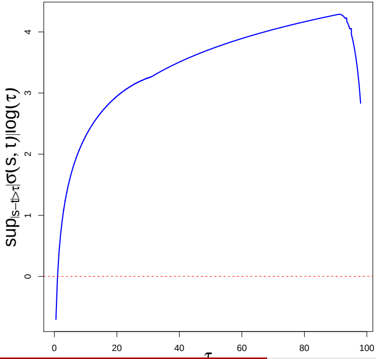


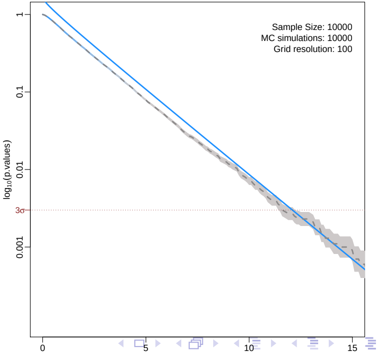
A case somewhere in between

Suppose we want to distinguish between a Power Law distributed cosmic source and and Dark Matter. The full model is:

$$(1-\eta)\frac{1}{k_\psi y^{\psi+1}}+\frac{\eta\exp\{-7.8\frac{y}{\phi}\}}{k_\phi y^{1.5}} \tag{23}$$

with k_ψ and k_ϕ normalizing constants, $y \in [1;100]$, $\psi = 1.4$ and $\phi \in [1;100]$.





...in the "next episode"...

Work in progress and (immediate) future goals:

- We would like to provide a formal explanation of cases where the global p-value approximations do and do not work.
- We would like to provide precise indications on how to spot these cases.
- We would like to exploit the information on the dependence structure of the underlying processes to improve, if possible, the global p-value approximations discussed in this talk.

All this will be discussed in:
Algeri S., van Dyk D.A., Conrad J. *Testing one hypothesis multiple times.* In preparation, 2016.
(Hopefully, available on ArXiv by the end of the summer.)

Thank you for listening!

References

- Algeri S., van Dyk D.A., Conrad J., Brandon, A. *Looking for a Needle in a Haystack? Look Elsewhere! A statistical comparison of approximate global p-values.* Submitted, 2016.
- Algeri S., Conrad J., van Dyk D.A. *A method for comparing non-nested models with application to astrophysical searches for new physics.* In: MNRAS: Letters, 2016.
- S. Algeri, *R package 'NONnest'* (2015). Download: `http://wwwf.imperial.ac.uk/~sa2514/Research.html`.
- Algeri S., van Dyk D.A., Conrad J. *Testing one hypothesis multiple times.* In preparation, 2016.
- A. C. Atkinson. "*A Method For Discriminating Between Models*". In: Journal of the Royal Statistical Society. Series B (Methodological) 32.3 (1970).
- 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).
- R. B. Davies. "*Hypothesis Testing when a Nuisance Parameter is Present Only Under the Alternatives*". In: Biometrika 64.2 (1977).
- B. Efron. *Large-Scale Inference.* Cambridge Books Online. Cambridge University Press, 2010.
- 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.
- M.R. Leadbetter, G. Lindgren, and H. Rootzén. *Extremes and Related Properties of Random Sequences and Processes.* Springer Series in Statistics. Springer-Verlag New York Inc., 1983.