Upper Limits

David A. van Dyk[1]
Joint work with Vinay Kashyap[2] and
Members of the California-Boston AstroStatistics Collaboration

[1]Department of Statistics, University of California, Irvine
[2]Smithsonian Astrophysical Observatory

BIRS, July 2010

Outline

Outline

Goals

Goals

- Clear up the terminology of “upper bounds” and “upper limits” among (high energy?) astronomers.
- Clarify the probability calculations of “upper limits”.
- Illustrate a difficulty with frequency coverage of selected confidence intervals.

I am not an Astronomer....

Detection Problems

For simplicity, consider a simple Poisson model

$$\begin{aligned} n_B|(\lambda_B, r, \tau_B) &\sim \text{Poisson}(r\tau_B\lambda_B) \\ n_S|(\lambda_S, \lambda_B, \tau_S) &\sim \text{Poisson}\Big(\tau_S(\lambda_S + \lambda_B)\Big) \end{aligned}$$

For simplicity we assume λ_B is known.
We use a standard hypothesis testing framework:

H_0 There is no source: $\lambda_S = 0$

H_A There is a source: $\lambda_S > 0$.

Detection Threshold

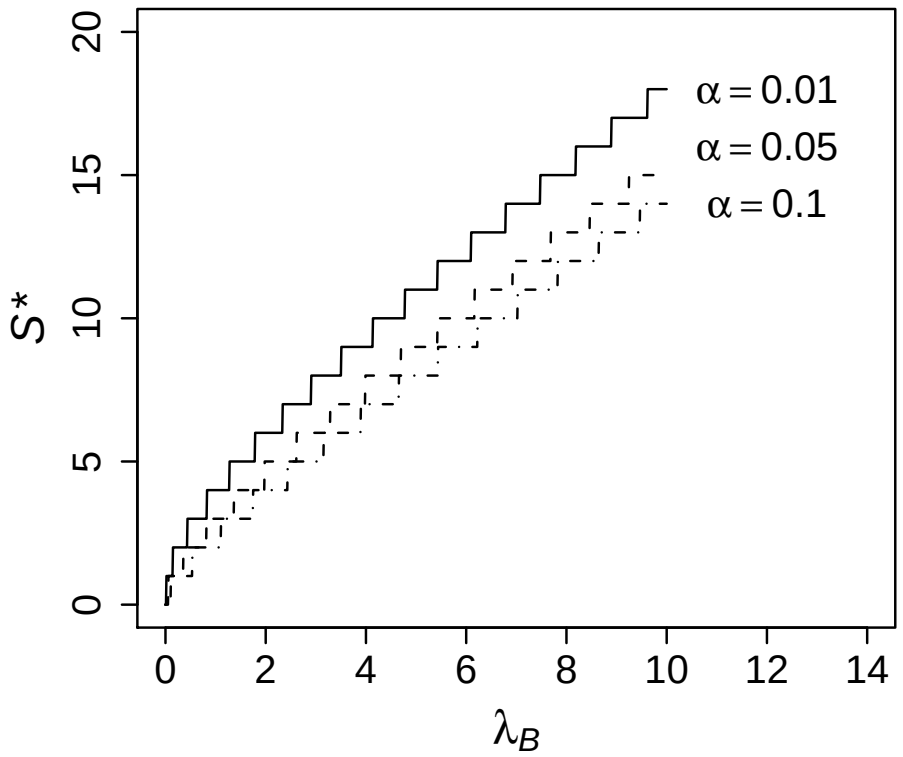
The detection threshold $n_S^\star$ is the smallest value such that

$$\Pr(n_S > n_S^\star | \lambda_S = 0, \lambda_B, \tau_S, \tau_B, r) \leq \alpha,$$

If $n_S \leq n_S^\star$ we conclude there is insufficient evidence to declare a source detection.

If $n_S > n_S^\star$ we conclude there is sufficient evidence to declare a source detection.

Detection Threshold



α-level detection threshold $n_S{}^\star$ as a function of the background intensity λ_B.

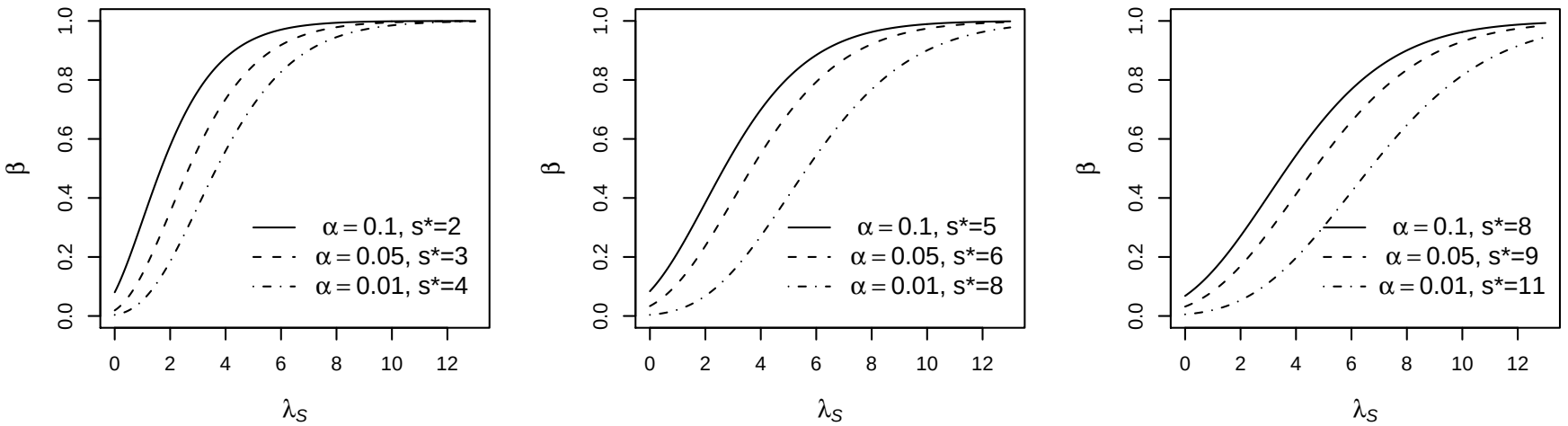
Power

The *power* of the test to detect a source as a function of its intensity is

$$\beta(\lambda_S) = \Pr(n_S > n_S^\star | \lambda_S, \lambda_B, \tau_S, \tau_B, r)\,.$$

Note $\beta(\lambda_S = 0) \leq \alpha$.

Power



Power for $\lambda_B = 1, 3, 5$ and given α

Outline

Typical Detection Prodedure

When there is a detection astronomers often

1. Report a detection
2. Report a confidence interval for λ_S

When there is not a detection astronomers often

1. Report no detection
2. Report an “Upper Limit” for λ_S

What is the difference?

Upper Limits

What is an “upper limit”?

In astronomy upper limits are inextricably bound to source detection: by an upper limit, an astronomer means

> *The maximum intensity that a source can have without having at least a probability of $\beta_{\min}$ of being detected under an α-level detection threshold.*

or conversely,

> *The smallest intensity that a source can have with at least a probability of $\beta_{\min}$ of being detected under an α-level detection threshold.*

Requires two probability calculations.

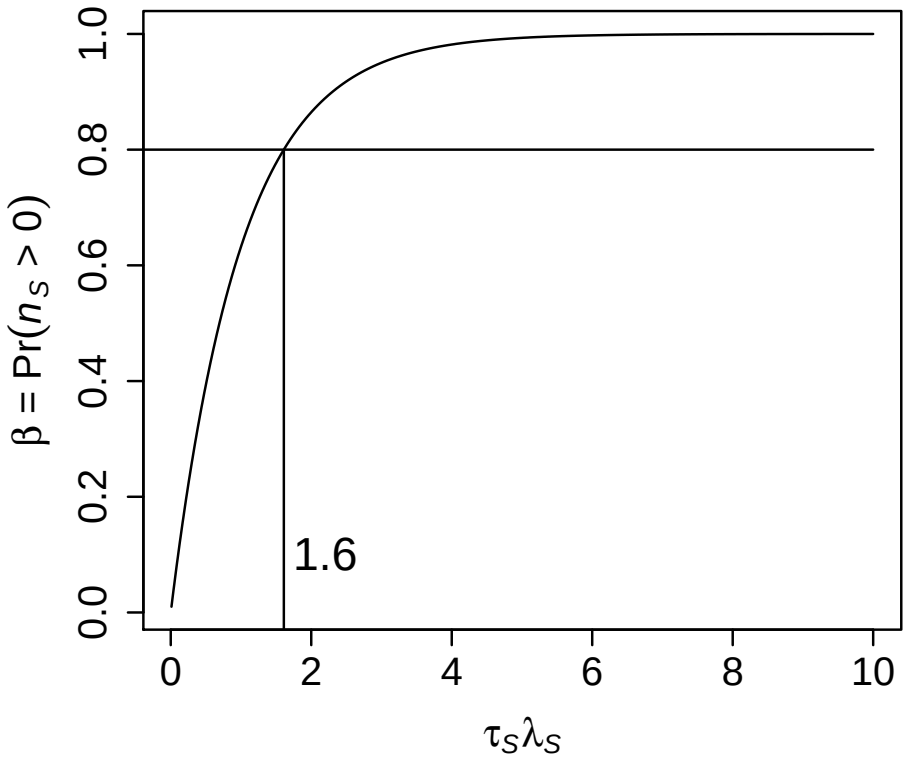
Upper Limits

Upper Limits are analogous to sample sizes as follows:

If you don't have a detection, the sample size indicates how much you should worry.

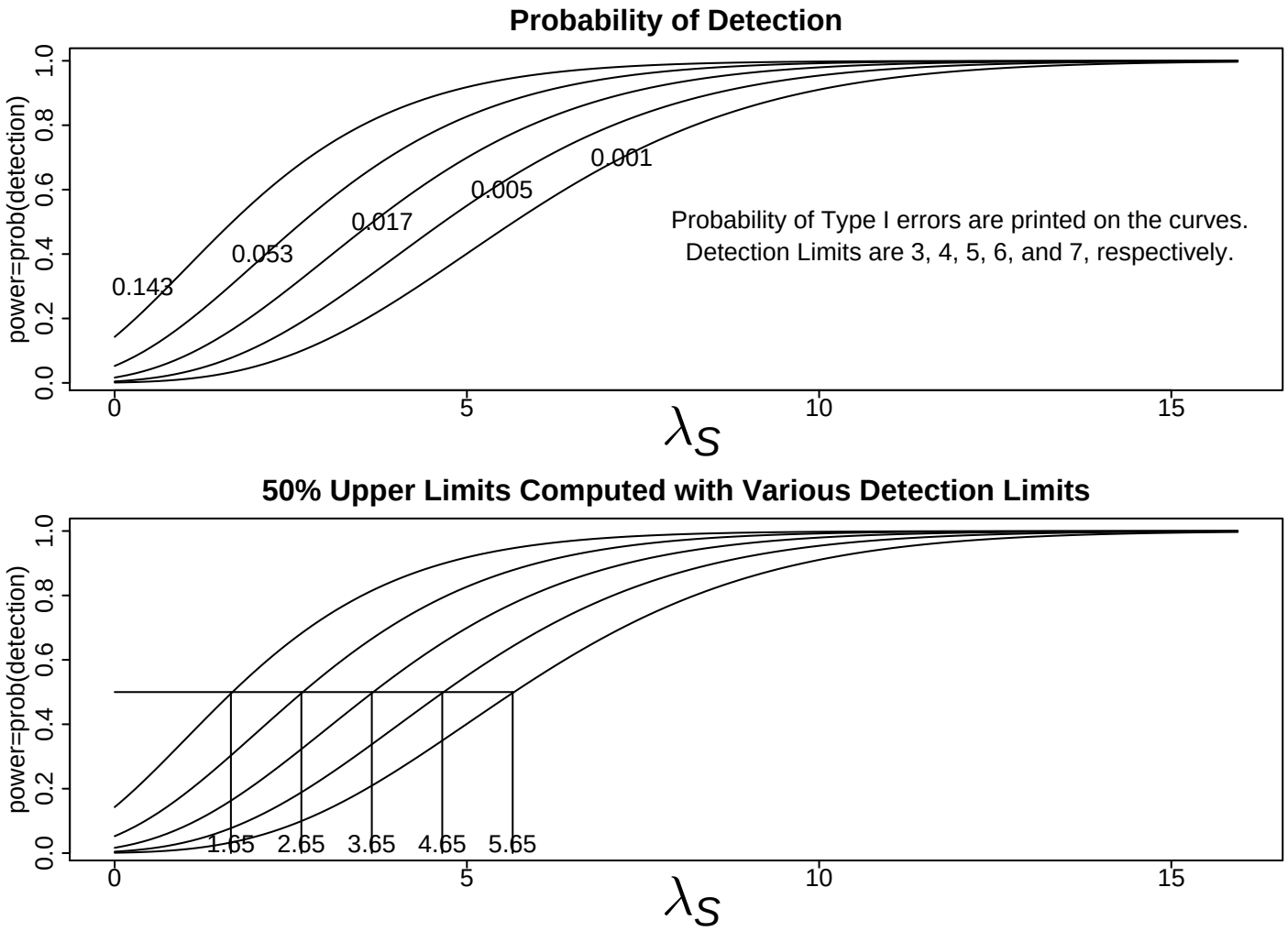
The Upper Limit aims to directly calibrate this.

Illustrating Upper Limits

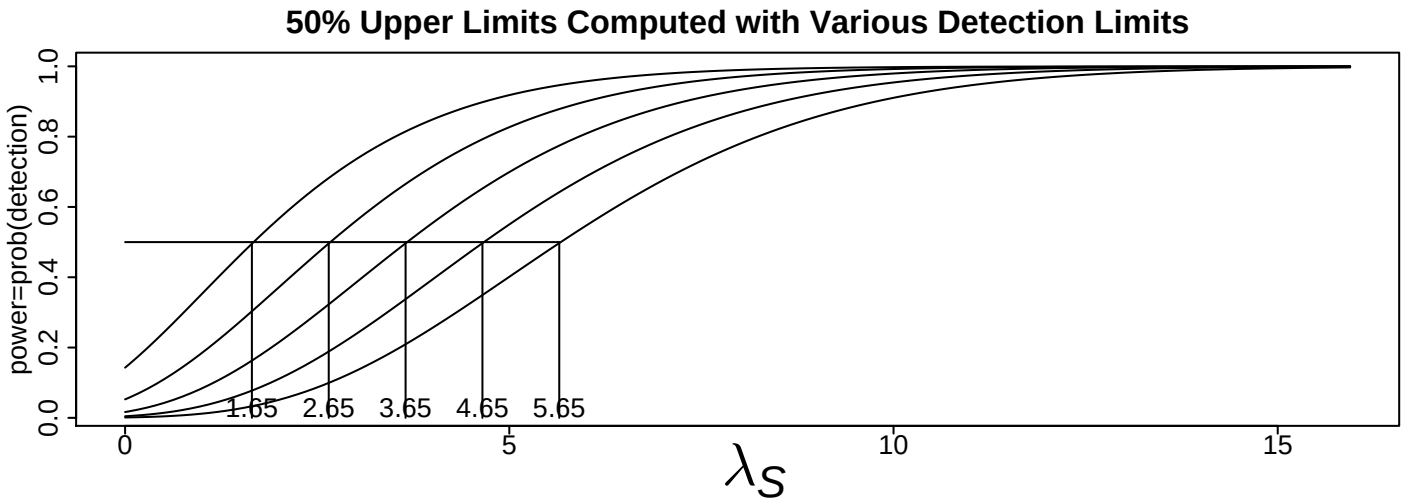


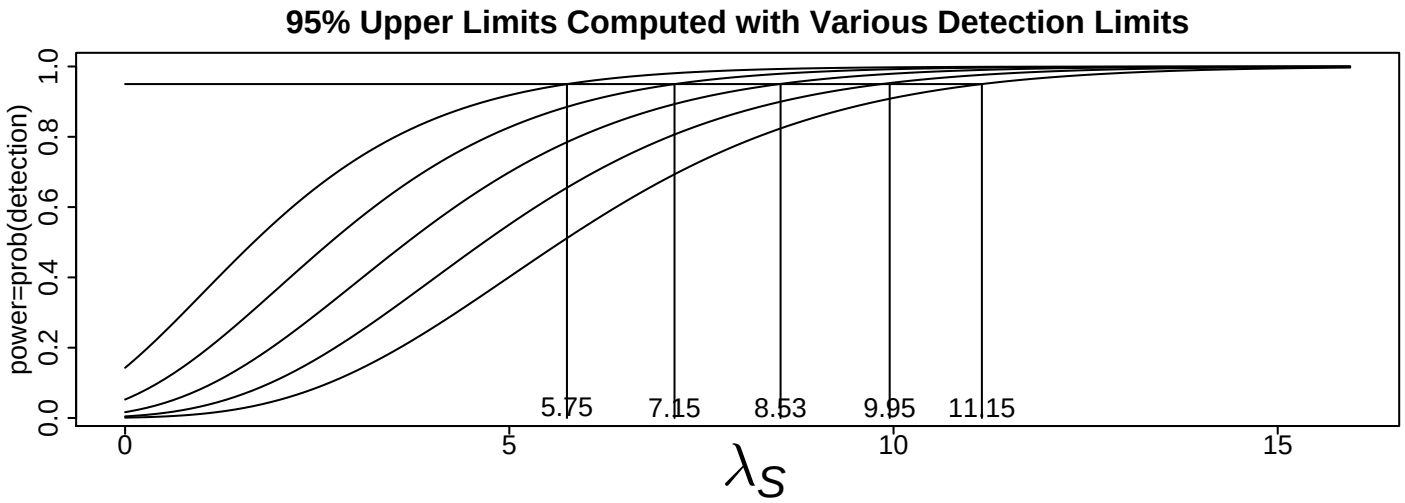
Upper limit with no background contamination.

Effect of Detection Threshold on UL



Effect of UL probability on UL





Upper Limits and Power

- In a typical power calculation, we would find the minimum τ_S so

$$\beta(\lambda_S) = \Pr(n_S > n_S^\star | \lambda_S, \lambda_B, \tau_S, \tau_B, r)$$

  achieves a given value for a given λ_S. Say 90% for $\lambda_S = 2$.
- For an upper limit we solve the same equation, but fixing τ_S and solving for λ_S.

Like power, an upper limit does not depend on the data and can be computed in advance.

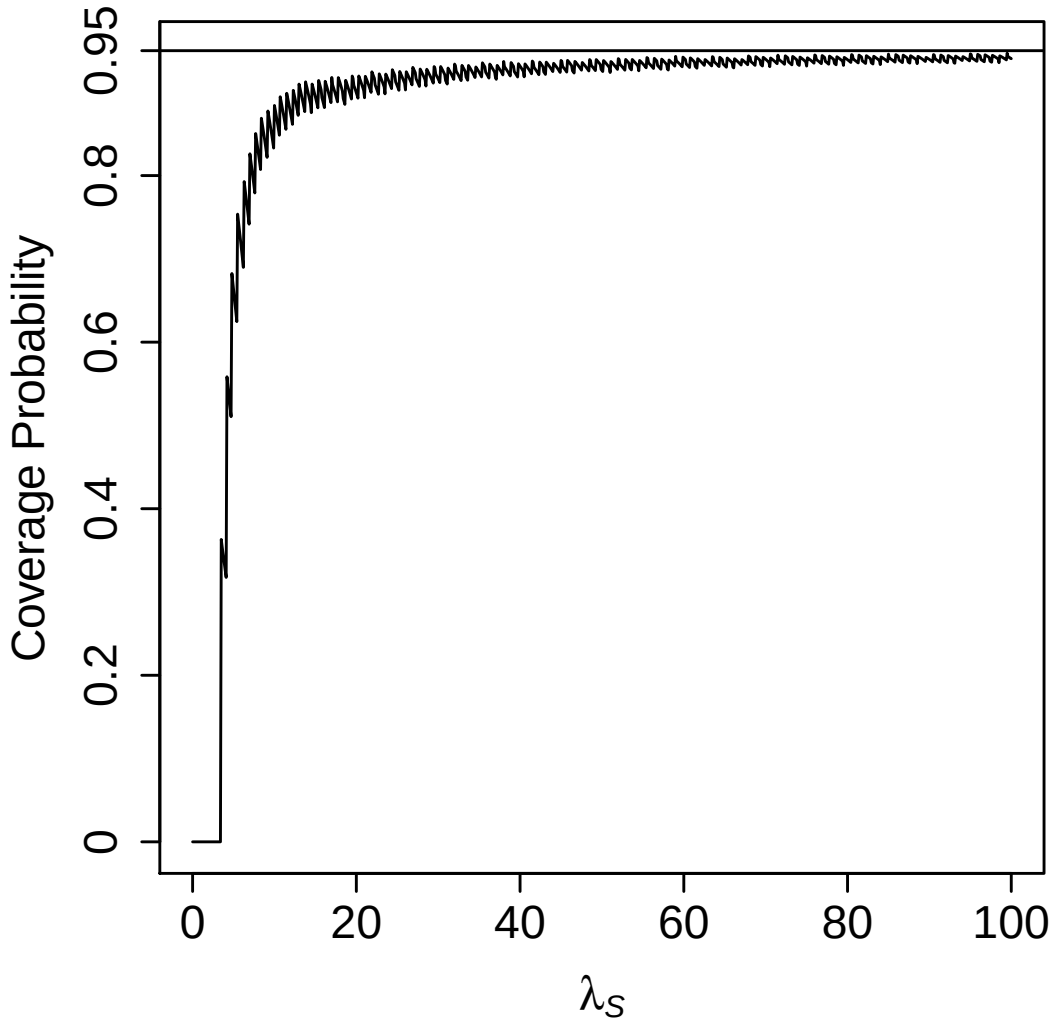
Outline

The Typical Procedure

- In the typical procedure, the confidence interval is only reported if a source is detected.
- But deciding whether to report the CI *based on the data* alters it frequency properties.
- This is similar to the problem reported in Feldman and Cousins (1998).

Unfortunately, frequency properties depend on what you would have done, had you had a different data set.

Under Coverage



Proposed Procedure

Always report

1. Whether the source was detected.
2. A Confidence Interval for the source intensity.
3. An upper limit, to quantify the strength of the experiment.

Proposed Procedure

Always report

1. Whether the source was detected.
2. A Confidence Interval for the source intensity.
3. An upper limit, to quantify the strength of the experiment.

But NEVER a p-value!!