Upper Limits

David A. van Dyk\textsuperscript{1}
Joint work with Vinay Kashyap\textsuperscript{2} and
Members of the California-Boston AstroStatistics Collaboration

\textsuperscript{1}Department of Statistics, University of California, Irvine
\textsuperscript{2}Smithsonian Astrophysical Observatory

BIRS, July 2010
Outline

1. Detection Problems
2. Upper Limits
3. Reporting Confidence Intervals
Outline

1. Detection Problems
2. Upper Limits
3. Reporting Confidence Intervals
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....
For simplicity, consider a simple Poisson model

\[ n_B | (\lambda_B, r, \tau_B) \sim \text{Poisson}(r \tau_B \lambda_B) \]
\[ n_S | (\lambda_S, \lambda_B, \tau_S) \sim \text{Poisson}(\tau_S(\lambda_S + \lambda_B)) \]

For simplicity we assume $\lambda_B$ is known.

We use a standard hypothesis testing framework:

\[ H_0 \quad \text{There is no source: } \lambda_S = 0 \]
\[ H_A \quad \text{There is a source: } \lambda_S > 0. \]
The detection threshold $n_S^*$ is the smallest value such that

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

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

If $n_S > n_S^*$ we conclude there is sufficient evidence to declare a source detection.
\( \alpha \)-level detection threshold \( n^{*}_{S} \) as a function of the background intensity \( \lambda_{B} \).
The *power* of the test to detect a source as a function of its intensity is

\[ \beta(\lambda_S) = \Pr(n_S > n_S^*|\lambda_S, \lambda_B, \tau_S, \tau_B, r) . \]

Note \( \beta(\lambda_S = 0) \leq \alpha \).
Power for $\lambda_B = 1, 3, 5$ and given $\alpha$
Outline

1. Detection Problems
2. Upper Limits
3. Reporting Confidence Intervals
Typical Detection Procedures

When there is a detection astronomers often
1. Report a detection
2. Report a confidence interval for $\lambda_S$

When there is not a detection astronomers often
1. Report no detection
2. Report an “Upper Limit” for $\lambda_S$

What is the difference?
What is an “upper limit”?  

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

\[ \text{The maximum intensity that a source can have without having at least a probability of } \beta_{\text{min}} \text{ of being detected under an } \alpha\text{-level detection threshold.} \]

or conversely,

\[ \text{The smallest intensity that a source can have with at least a probability of } \beta_{\text{min}} \text{ of being detected under an } \alpha\text{-level detection threshold.} \]

Requires two probability calculations.
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.

\[ \beta = \Pr(n_S > 0) \]

\[ \tau S \lambda S \]

David A. van Dyk
Upper Limits
Detection Problems
Upper Limits
Reporting Confidence Intervals

Effect of Detection Threshold on UL

Probability of Detection

Probability of Type I errors are printed on the curves. Detection Limits are 3, 4, 5, 6, and 7, respectively.

50% Upper Limits Computed with Various Detection Limits
Effect of UL probability on UL

50% Upper Limits Computed with Various Detection Limits

95% Upper Limits Computed with Various Detection Limits
In a typical power calculation, we would find the minimum \( \tau_S \) so

\[
\beta(\lambda_S) = \Pr(n_S > n_S^* | \lambda_S, \lambda_B, \tau_S, \tau_B, r)
\]

achieves a given value for a given \( \lambda_S \). Say 90% for \( \lambda_S = 2 \).

For an upper limit we solve the same equation, but fixing \( \tau_S \) and solving for \( \lambda_S \).

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

1. Detection Problems
2. Upper Limits
3. Reporting Confidence Intervals
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 its 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

![Graph showing coverage probability vs. \( \lambda_S \)]
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!!