

SAMSI Spring 2006 Program on Astrostatistics

Workshop on Source and Feature Detection

Upper Limits: The Simplest Example

David van Dyk

Consider the model $X \sim \text{Poisson}(\psi + \beta)$, where β is the known expected count from background and ψ is the unknown expected count from source. In source detection, we are interested in the basic question of whether $\psi > 0$.

A Basic Power-Based Method.

1. Suppose $\psi = 0$ and $X \sim \text{Poisson}(\beta)$. Under this assumption, let x^* be the smallest value such that

$$\Pr(X \leq x^* | \beta) \geq 95\%.$$

Figure 1 illustrates x^* as a function of β .

2. If $X > x^*$, we conclude that $\psi > 0$ and that we have detected a source.
3. If $X \leq x^*$ we do not detect a source but may compute an upper limit for ψ . For a given x^* , we can compute

$$\Pr(X > x^* | \psi + \beta)$$

as a function of ψ .

4. Let ψ^* be the value of ψ so that

$$\Pr(X > x^* | \psi^* + \beta) = 95\%.$$

Notice that ψ^* does not depend on the observed counts. We are only using the fact that $X \leq x^*$.

5. For $\psi > \psi^*$ we are likely to detect the source and conclude that $\psi > 0$. For a non-detection, we can call ψ^* an upper limit.

Example: Suppose $\beta = 2$, then the 95% detection threshold is $x^* = 5$, see Figure 1. Thus, we would conclude there is a source if we observe more than 5 photons. Figure 2 plots the power of the test as a function of the source intensity, ψ . We have a 95% or greater chance of detecting a source with $\psi \geq 8.55$. Suppose, we observe $X = 3$ photons, and cannot conclude there is a source. Our power-based upper limit for ψ is 8.55. *Notice we would have the same upper limit for any $X \leq 5$.*

Significance Tests and Confidence Intervals. Generalizing the above discussion, we might in principle be interested in accepting or dismissing any value of ψ , rather than just the value $\psi = 0$. Suppose we are particularly interested in the possibility that $\psi = \psi_0$. If we assume $\psi = \psi_0$ and $X \sim \text{Poisson}(\psi_0 + \beta)$, we can compute $x_U^*(\psi_0)$, the smallest value such that

$$\Pr(X \leq x_U^*(\psi_0) | \psi_0 + \beta) \geq 95\%.$$

We use the subscript ‘U’ because $x_U^*(\psi_0)$ is a 95% upper bound for X . We can decide between the possibilities that $\psi \leq \psi_0$ and $\psi > \psi_0$ as follows:

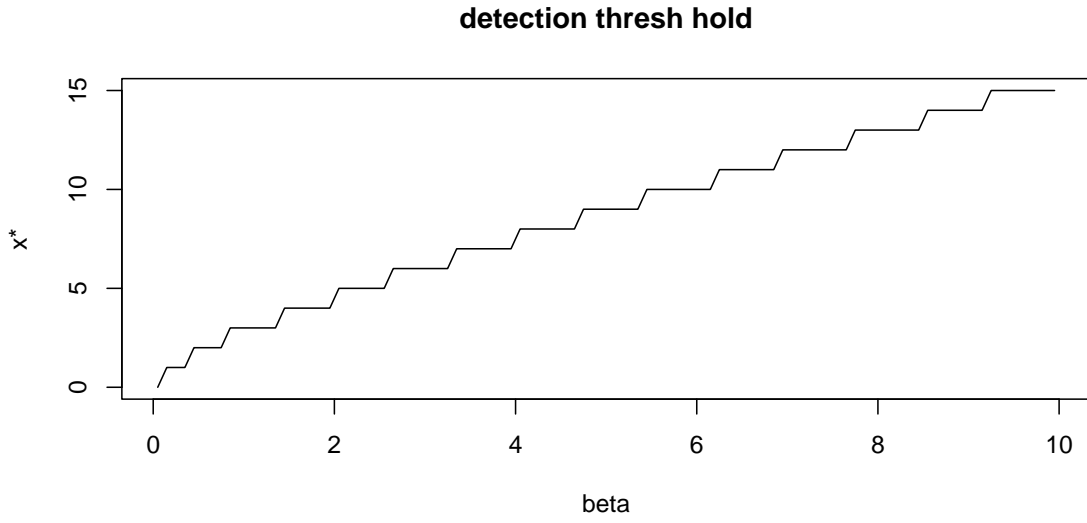


Figure 1:

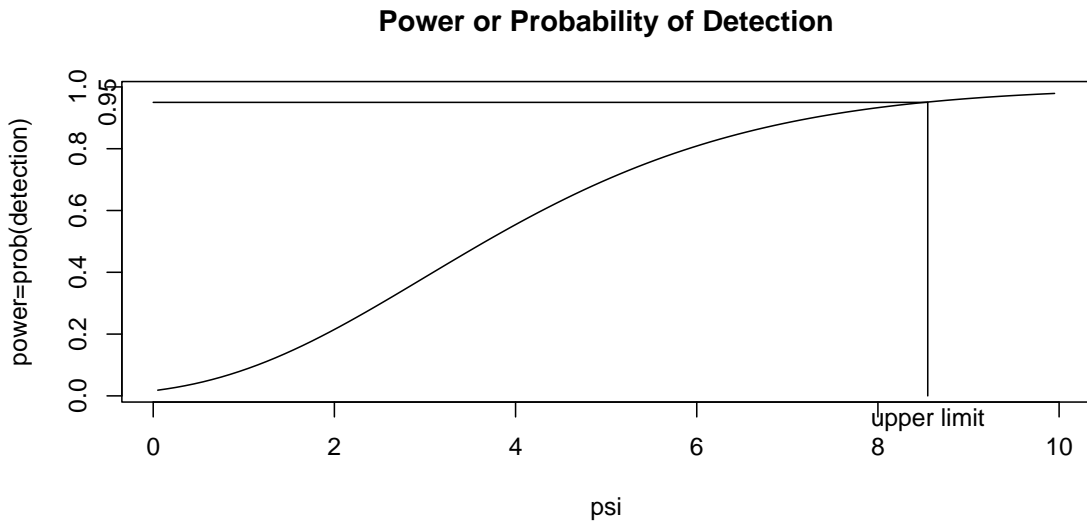


Figure 2:

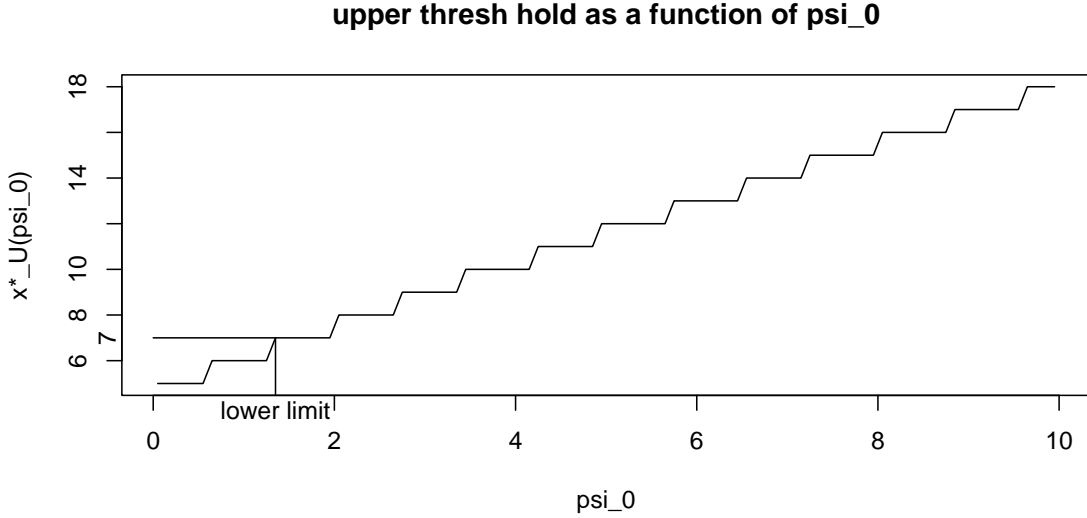


Figure 3:

- If $X > x^*(\psi_0)$, we conclude that $\psi > \psi_0$.
- If $X \leq x^*(\psi_0)$ we cannot dismiss the possibility that $\psi = \psi_0$ (or less than ψ_0).

This is a significance test for testing the “null hypothesis” that $\psi = \psi_0$. We can use this to generate a confidence interval by setting $I_L(X) = \{\psi : X \leq x^*(\psi)\}$. This is the set of values of ψ that we cannot dismiss, the plausible values of ψ . WE now use the subscript ‘L’, because as we shall see, this interval generates a lower bound for ψ .

Example: Again, suppose $\beta = 2$, then the 95% detection threshold is $x^* = 5$. Figure 3 gives the upper threshold hold as a function of ψ_0 , i.e., $x^*_U(\psi_0)$. Suppose, we observe $X = 7$ photons, and conclude there is a source. As Figure 3 illustrates, $X = 7 \leq x^*_U(\psi)$ for and $\psi \geq 1.35$., the resulting confidence interval is $[1.35, \infty)$. On the other hand, if we had observed only $X = 3$ photons, then $X \leq x^*(\psi)$ for all non-negative ψ , and the resulting confidence interval is simply $[0, \infty)$.

This is a “one-sided” confidence interval: it provides a lower limit, but no upper limit for ψ . This type of interval is ideal for source detection. If the lower-limit is greater than zero, we have found a source. If we have not found a source, however, this interval will always be of the form $[0, \infty)$ and an upper limit or a “two-sided” interval is more informative. Next we describe how to compute upper limits based on the “other” one-sided interval.

A Basic Significance-Test Based Method.

1. Suppose $\psi = \psi_0$ and $X \sim \text{Poisson}(\beta + \psi_0)$. Under this assumption, let $x^*_L(\psi_0)$ be the largest value such that

$$\Pr(X \geq x^*_L(\psi_0) | \psi_0 + \beta) \geq 95\%.$$

2. If $X < x^*_L(\psi_0)$, we conclude that $\psi < \psi_0$.
3. If $X \geq x^*_L(\psi_0)$ we cannot dismiss the possibility that $\psi = \psi_0$.

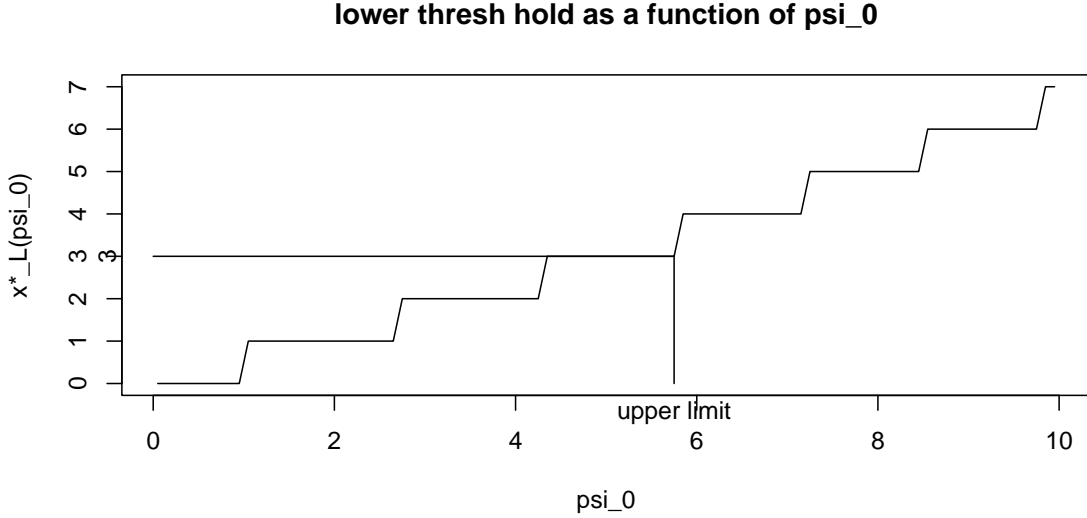


Figure 4:

4. We can construct a confidence interval as $I_U(X) = \{\psi : X \geq x_L^*(\psi)\}$. These are the values of ψ that we cannot dismiss, the plausible values of ψ .
5. If $X > x_U^*(0)$, we conclude $\psi > 0$ and that we have detected a source. In this case zero is not in $I_L(X)$ and $I_L(X)$ provides a lower limit for ψ , that is greater than zero.
6. If $X \leq x_U^*(0)$, we have not detected a source, zero is contained in $I_L(X)$ and the largest value in $I_U(X)$ is an upper limit for ψ .

Example: Again, suppose $\beta = 2$ and that the 95% detection threshold is $x^* = 5$. Figure 4 gives the lower threshold as a function of ψ_0 , i.e., $x_L^*(\psi_0)$. Suppose, we observe $X = 3$ photons, and cannot conclude there is a source. Figure 4 illustrates the values of ψ such that $3 \geq x_L^*(\psi)$, this is the confidence interval that provides the upper limit for ψ , i.e., $[0, 5.75]$.

A Better Strategy: Two-Sided Tests and Intervals. The basic strategy outlined above over-estimate confidence in the upper bounds. First computing a lower bound and then computing an upper bound only if the lower bound is zero in a manner that does not account for conditional nature of the computation will result in upper bounds that are too low. There are two possible ways to avoid this.

1. A single two-sided interval can be computed. These intervals compute an upper bound and a lower bound coherently and simultaneously. It should be emphasized, however, that these intervals will have less power to detect weak sources. This power exchanged for an upper bound.
2. An upper bound can be computed accounting for the fact that an uninformative lower bound has already been computed. In this way we avoid sacrificing detection power, but will have higher upper limits. (Indeed, the obvious upper limit in this case is $+\infty$, but perhaps we can do better.)

Here, I'll just describe the second strategy.

A Basic Two-Sided Significance-Test Based Method.

1. Suppose $\psi = \psi_0$ and $X \sim \text{Poisson}(\beta + \psi_0)$. Under this assumption, let $(x_L^\dagger(\psi_0), x_U^\dagger(\psi_0))$ be the shortest interval such that

$$\Pr(x_L^\dagger(\psi_0) \leq X \leq x_U^\dagger(\psi_0) | \psi_0 + \beta) \geq 95\%.$$

2. If $X < x_L^\dagger(\psi_0)$, we conclude that $\psi < \psi_0$.
3. If $X > x_U^\dagger(\psi_0)$, we conclude that $\psi > \psi_0$.
4. If $x_L^\dagger(\psi_0) \leq X \leq x_U^\dagger(\psi_0)$, we cannot dismiss the possibility that $\psi = \psi_0$.
5. We can construct a confidence interval as $I_2(X) = \{\psi : x_L^\dagger(\psi) \leq X \leq x_U^\dagger(\psi)\}$. These are the values of ψ that we cannot dismiss, the plausible values of ψ .
6. If $X > x_U^\dagger(0)$, we conclude $\psi > 0$ and that we have detected a source. In this case zero is not in $I_2(X)$ and $I_2(X)$ provides a lower and upper limits for ψ , that are both greater than zero. (Because $x_U^\dagger(0) \geq x_U^*(0)$, we have less power to detect sources than when we compute a simple lower bound.)
7. If $X \leq x_U^\dagger(0)$, we have not detected a source, zero is contained in $I_2(X)$ and the largest value in $I_2(X)$ is an upper limit for ψ .

Using the shortest interval in Step 1 gives us the shortest two-sided Confidence Interval and balances in some sense the power for the upper and lower limits. If we want more power for the lower limit (i.e., the detection) than for the upper limit, we may want to use different interval.

The confidence interval computed in Step 5 can be replaced with any other interval (e.g., a Bayesian interval computed with a Jeffrey's prior). Such a substitution may effect the frequency properties of the procedure, but may be necessary in the presence of nuisance parameters. Likewise the basic significance test that we use here to generate a confidence interval can be replaced by any appropriate test (e.g., a properly calibrated likelihood ratio test). Again, this substitution can effect the frequency properties of the procedure, but may be necessary in more complex problems.