Class 10: Confidence intervals

Miguel F. Morales
Bryna Hazelton

Frequentist vs. Bayesian

Sigh…

Common statistical questions

- [x] Am I confident this is a real signal?
- [ ] I am confident I saw something, what was the real signal strength/level?
- [ ] I didn’t see anything, how faint/small must the signal have been for me not to see it?

Uncertainty on a Measurement

Uncertainty on a measurement

- What if you have a high σ detection, and now you want to know what the accuracy of your measurement is?

Uncertainty on a measurement

Example statistical question:

- If I performed the same measurement many times, what range of signal values would I observe?

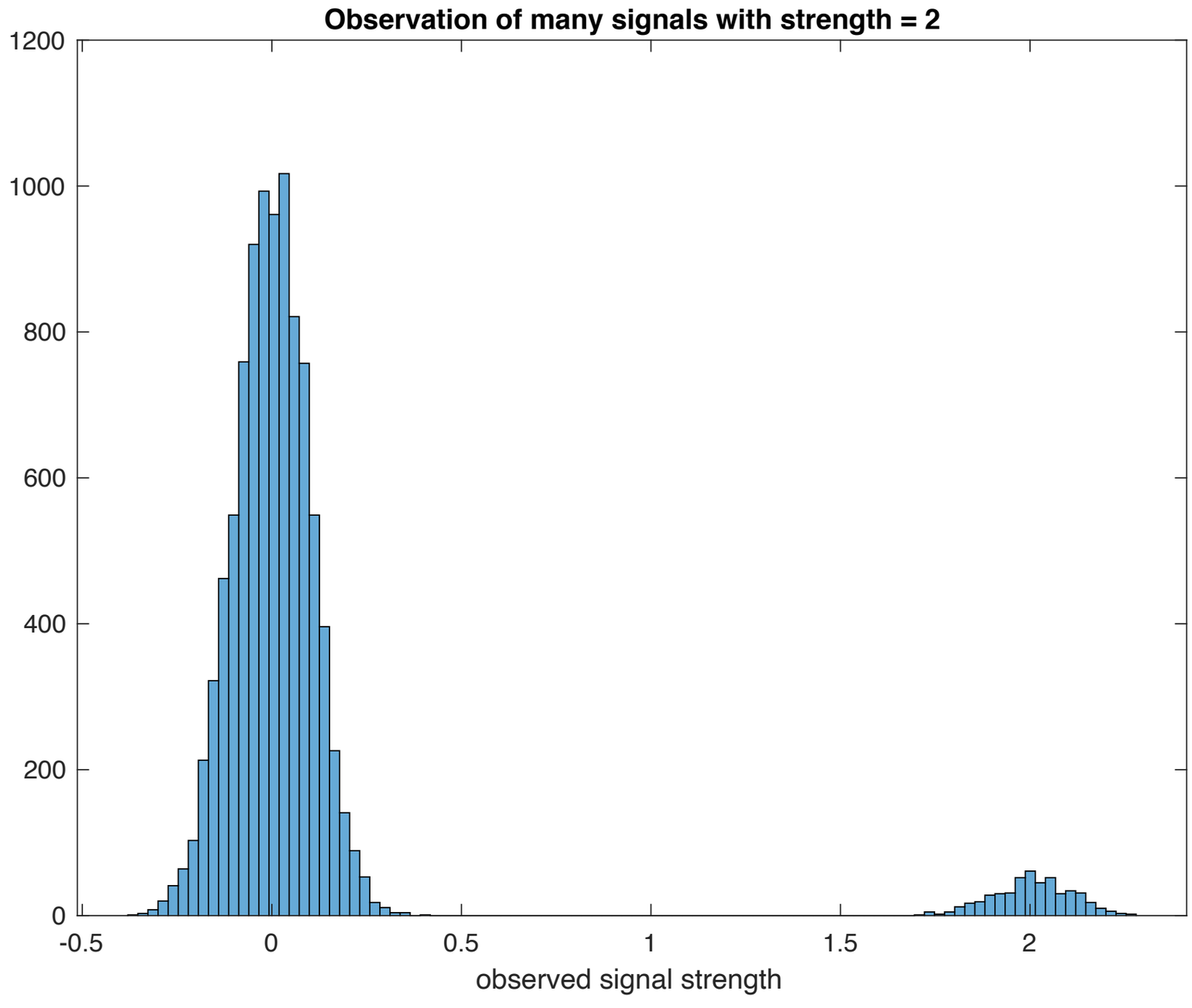
Uncertainty on a measurement

Simulation approach

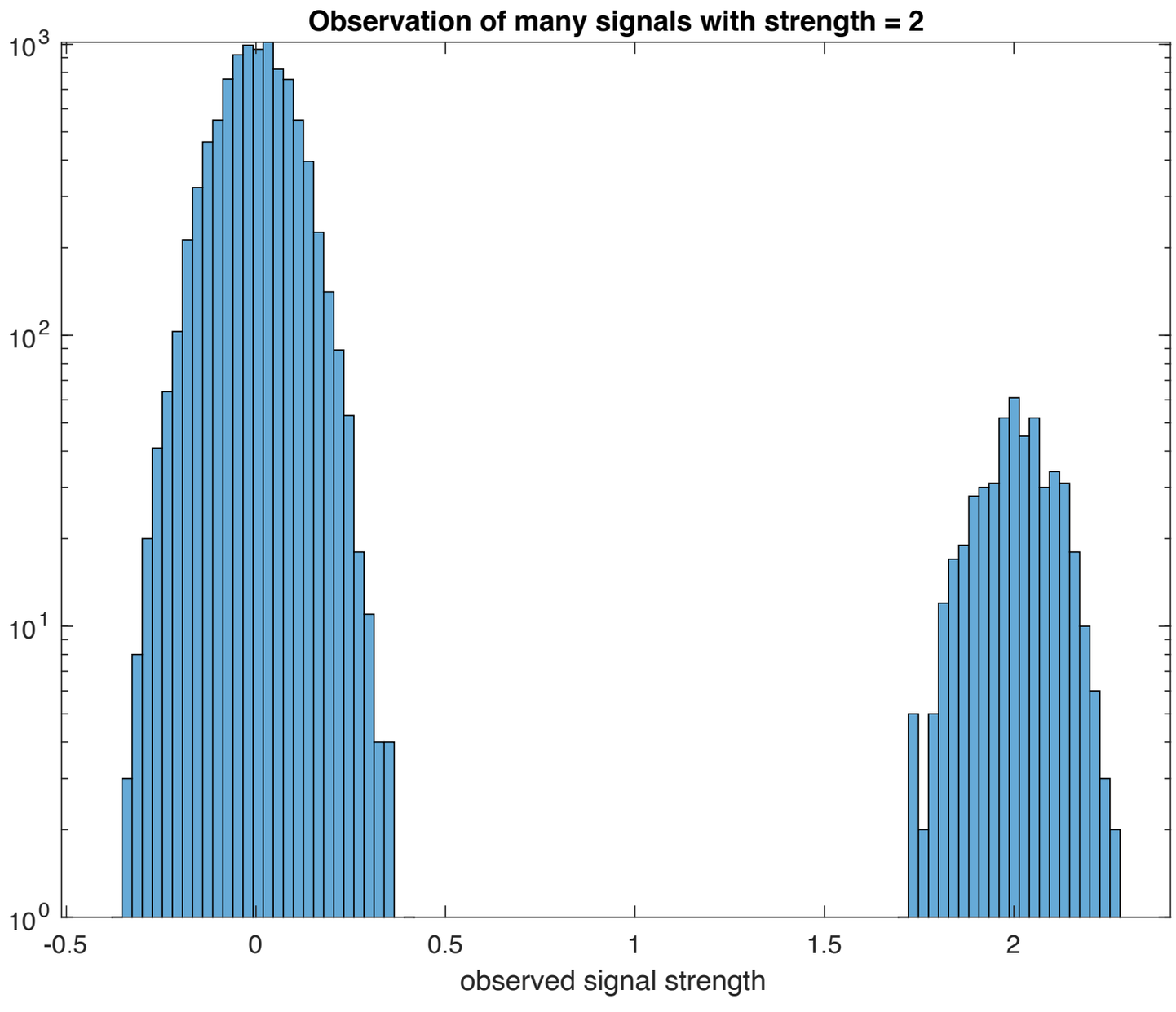
```
size = 10000;
noise = randn(1,size)*0.1;
signal = zeros(1,size);
signal(randi(numel(signal),[1,500])) = 2;

obssiganl = noise + signal;
```

Uncertainty on a measurement



Uncertainty on a measurement



Uncertainty on a measurement

Statistical question:

- If I performed the same measurement many times, what range of signal values would I observe?

Even if the signal strength is constant, we observe a range of measurements

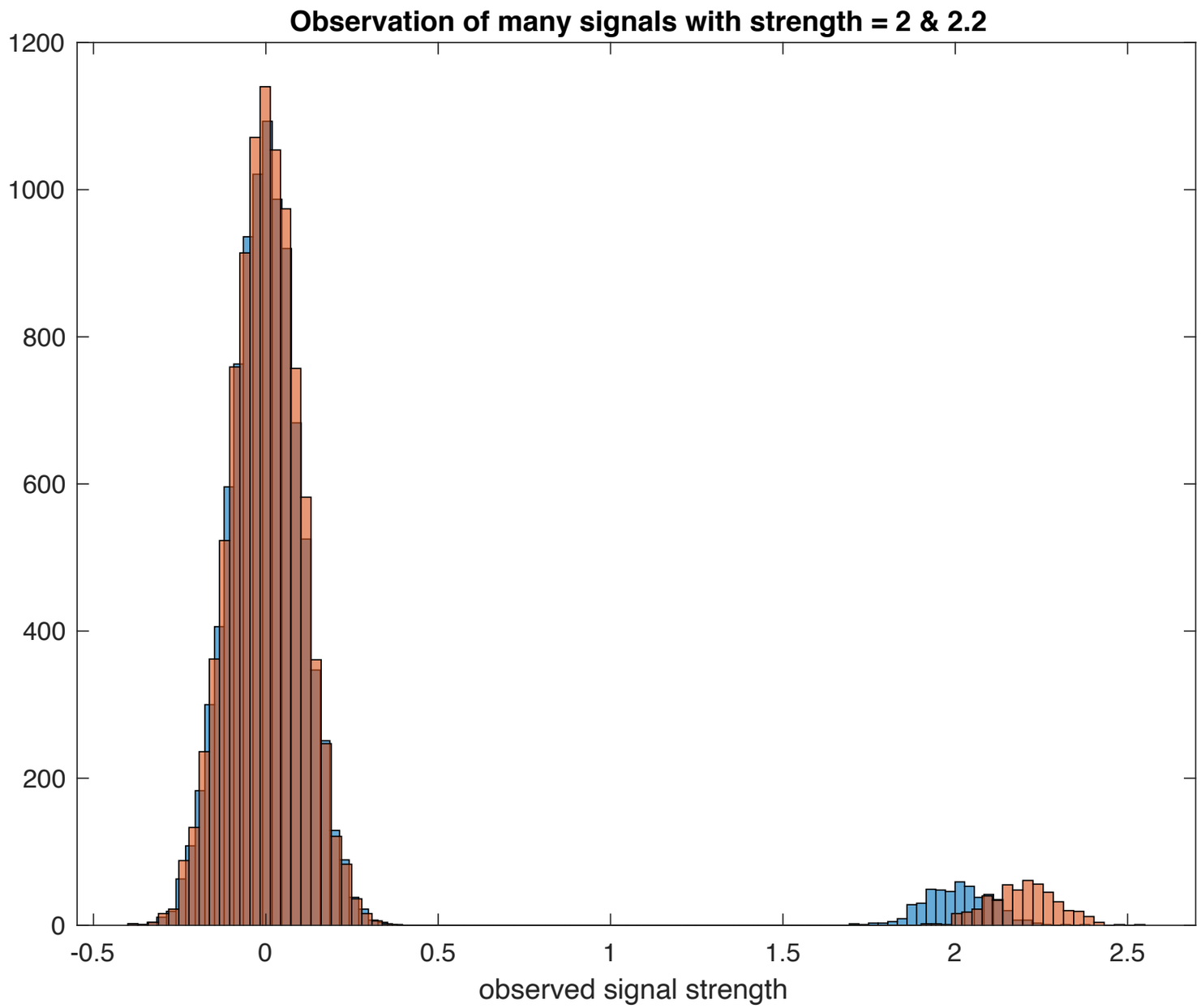
Uncertainty on a measurement

Turn statistical question around:

- If I measure a signal once, what range of true signal strengths could have given me the same observation?

Test with simulation of two input signals (2.0 & 2.2)

Uncertainty on a measurement



An asside in math notation

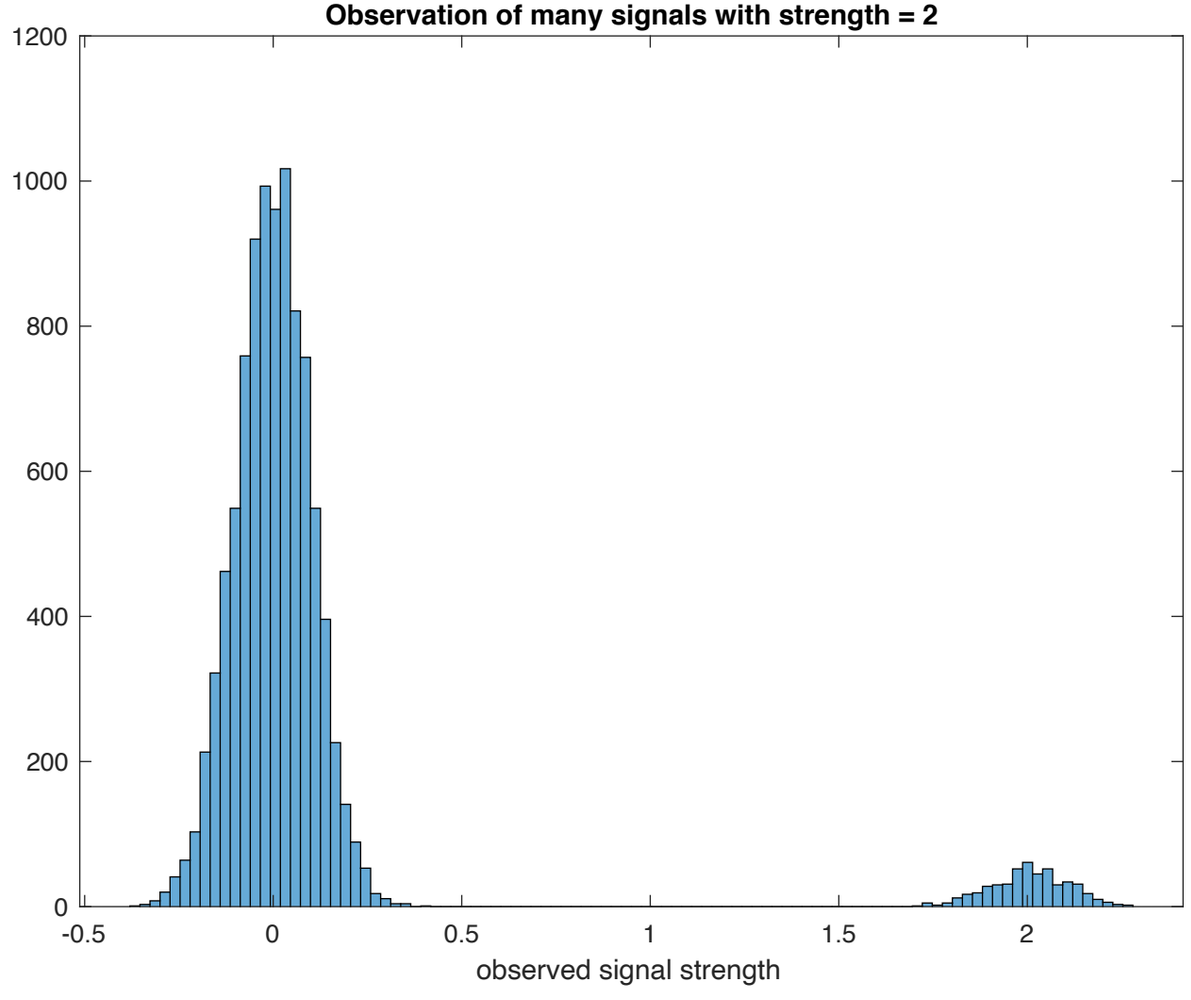
Read as:

- Given a particular true signal, what is the probability of getting a particular data value?

$$P(\text{data} \mid \text{signal}_{\text{T}})$$

An asside in math notation

$$P(\text{data} \mid \text{signal}_{\text{T}})$$



Two separate questions

A. If I performed the same measurement many times, what range of observed signal values (data) would I observe?
$P(\mathrm{data}\,|\,\mathrm{signal_T})$

B. If I measure a signal once (data), what is the probability of the true signal strengths? $P(\mathrm{signal_T}\,|\,\mathrm{data})$

Bayes' theorem

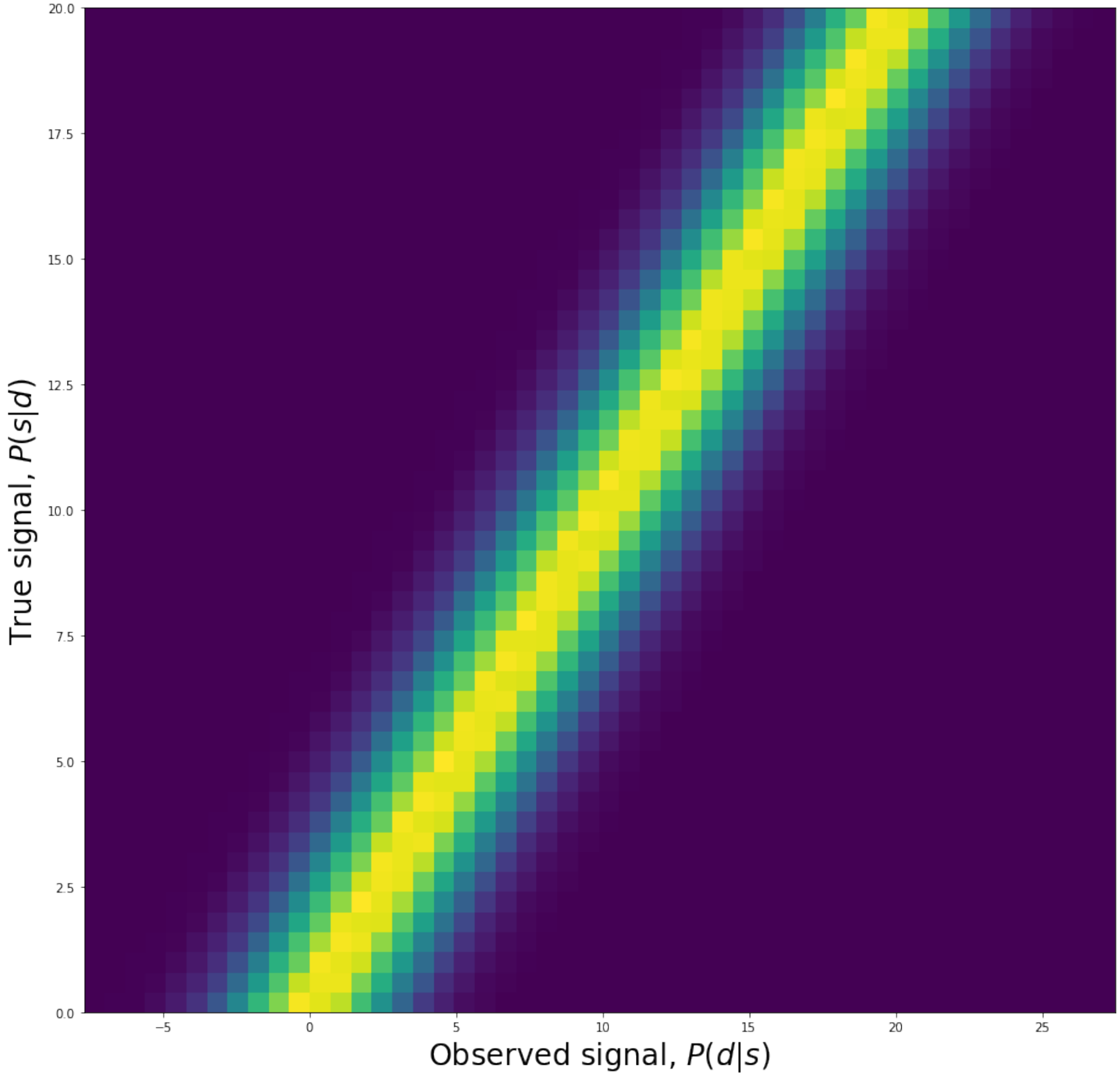
- Formally can change questions and calculate the desired $P(\mathrm{signal_T} \,|\, \mathrm{data})$

- In practice must be used with great care

$$P(s \mid d) = \frac{P(d \mid s)P(s)}{P(d)}$$

Confidence interval

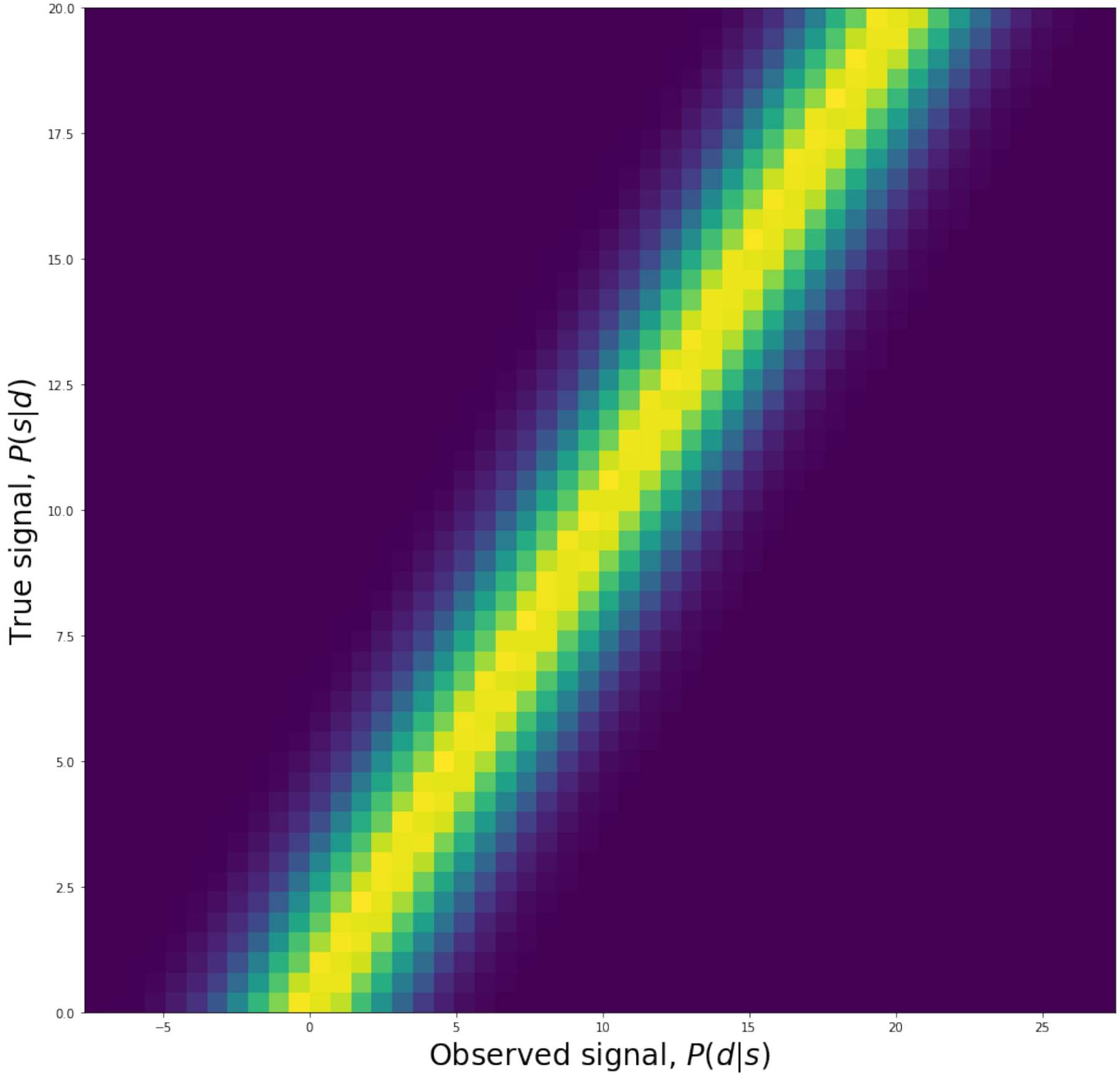
Simulated observations



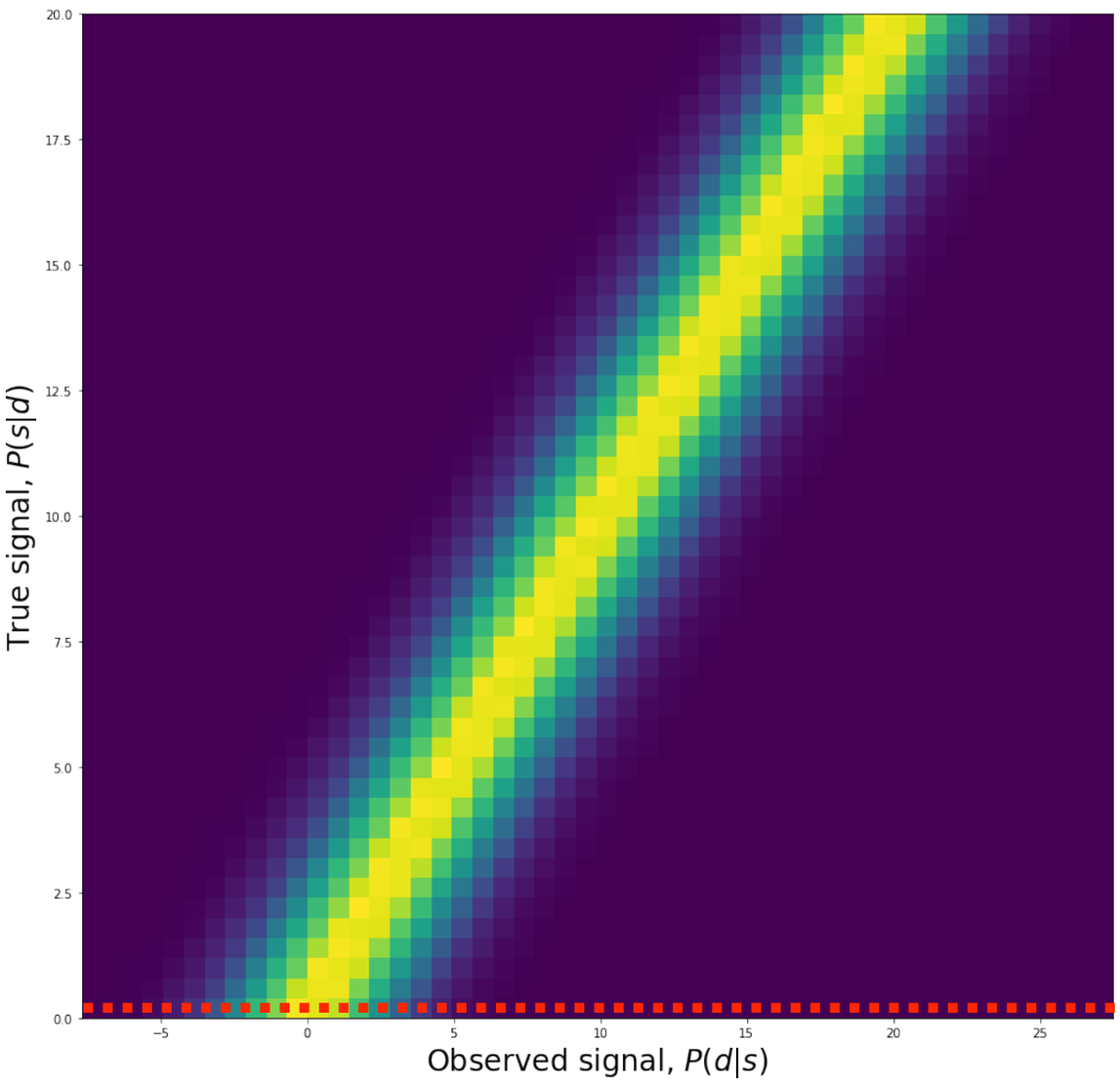
How to make

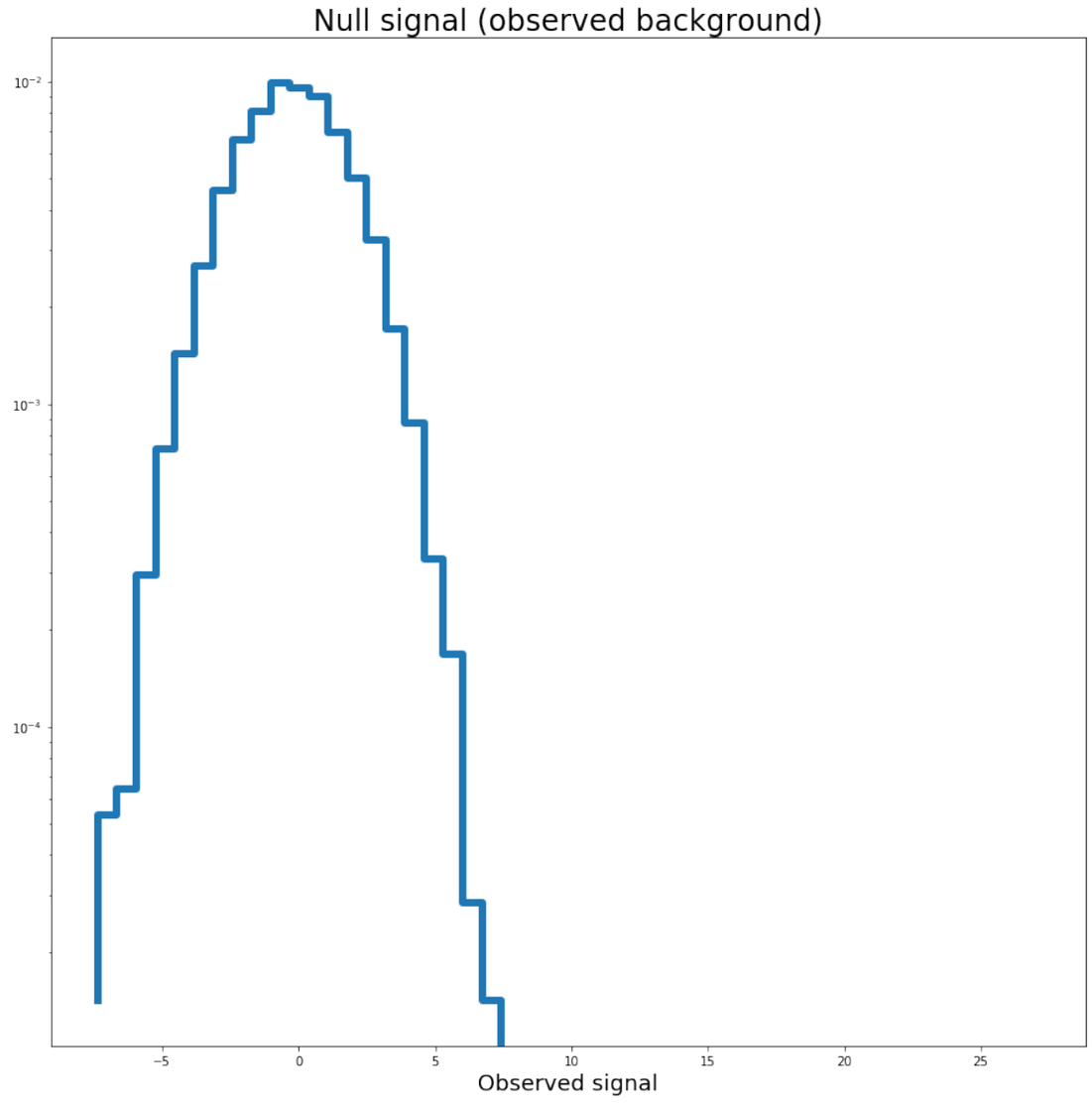
- Start with background (model or data)
- Inject fake signals of varying strength
- Measure observed signal
- Histogram true signal vs. observed signal

Simulated observations



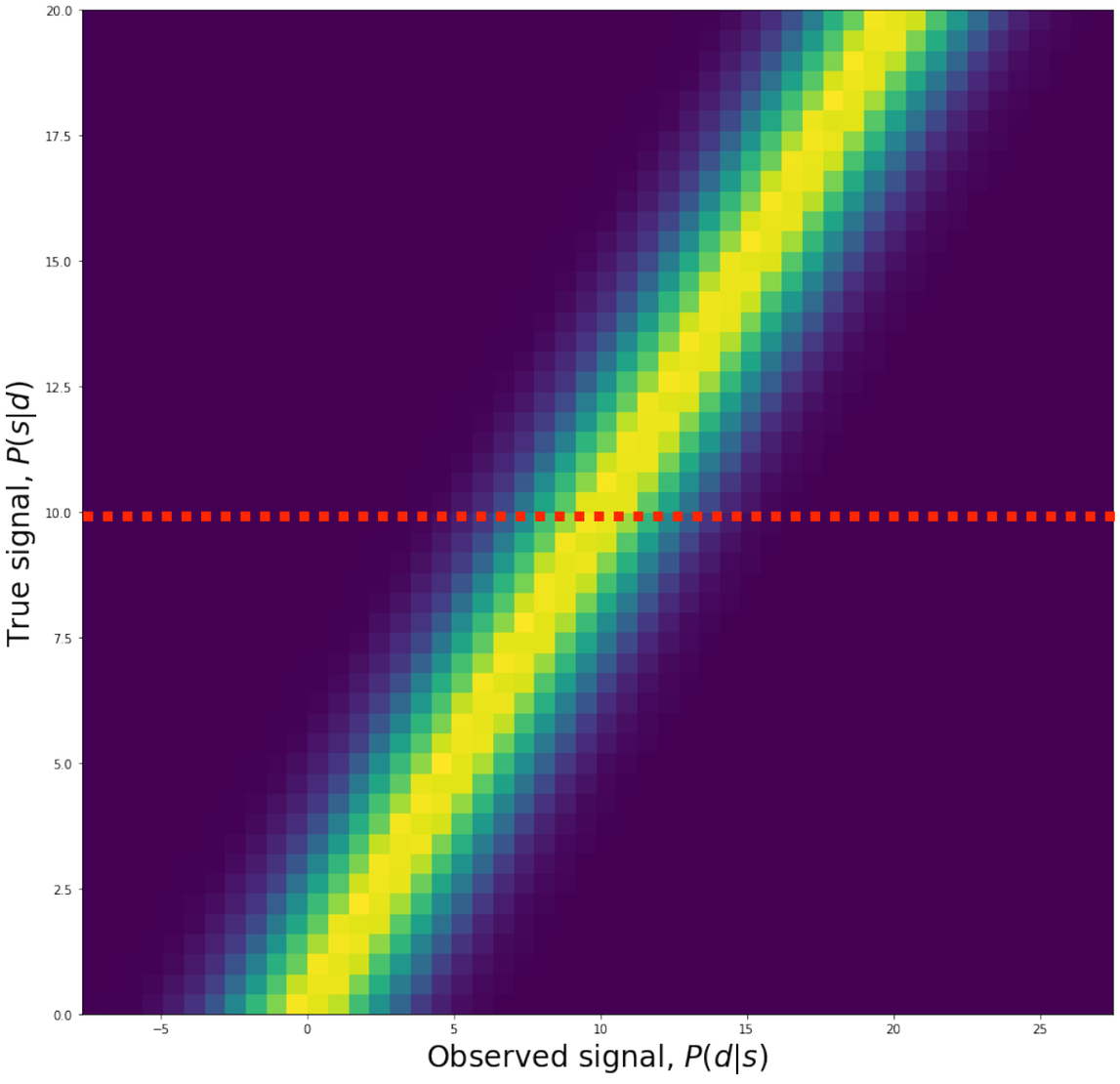
Slices

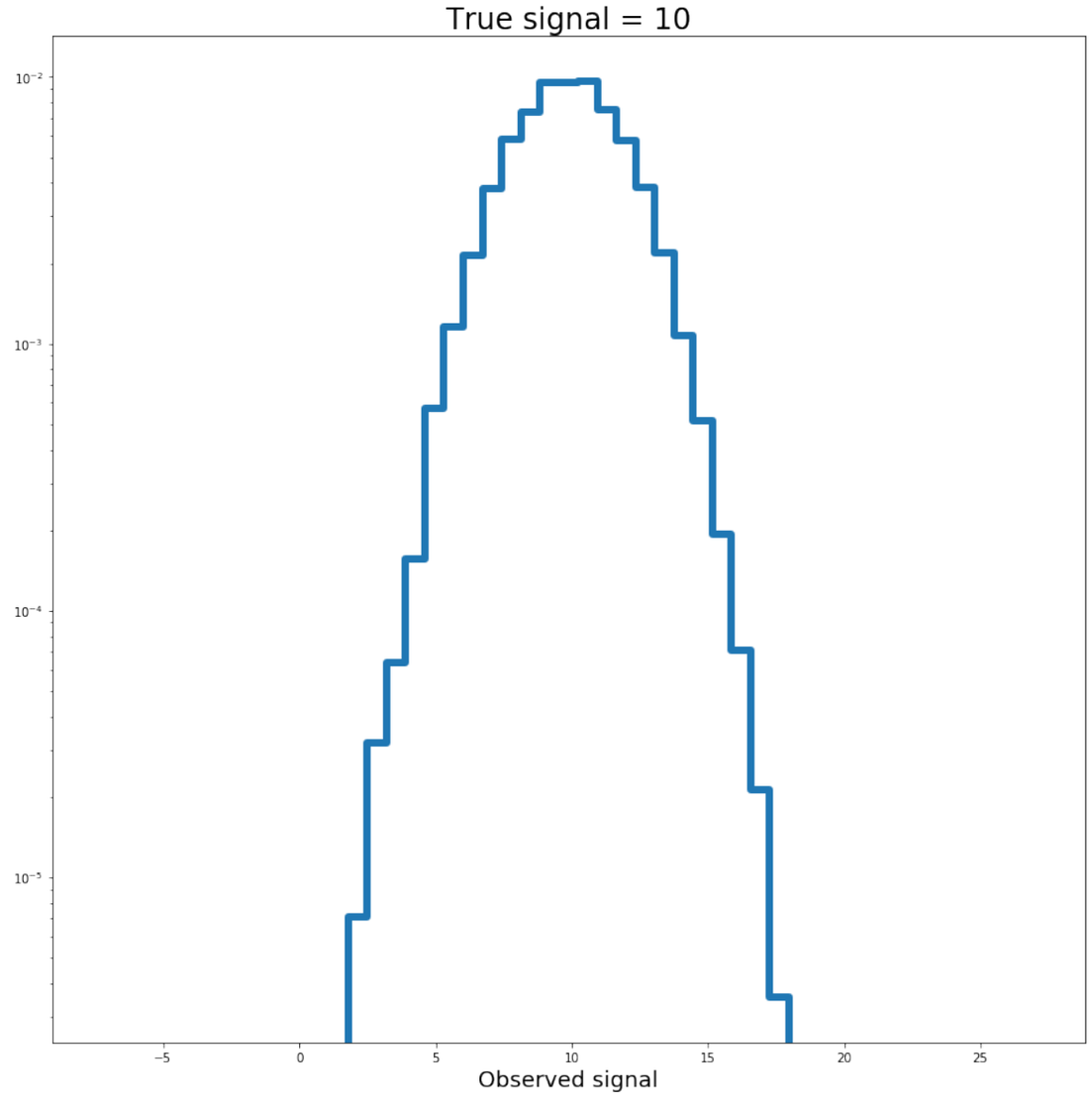




$$P(d \mid s)$$

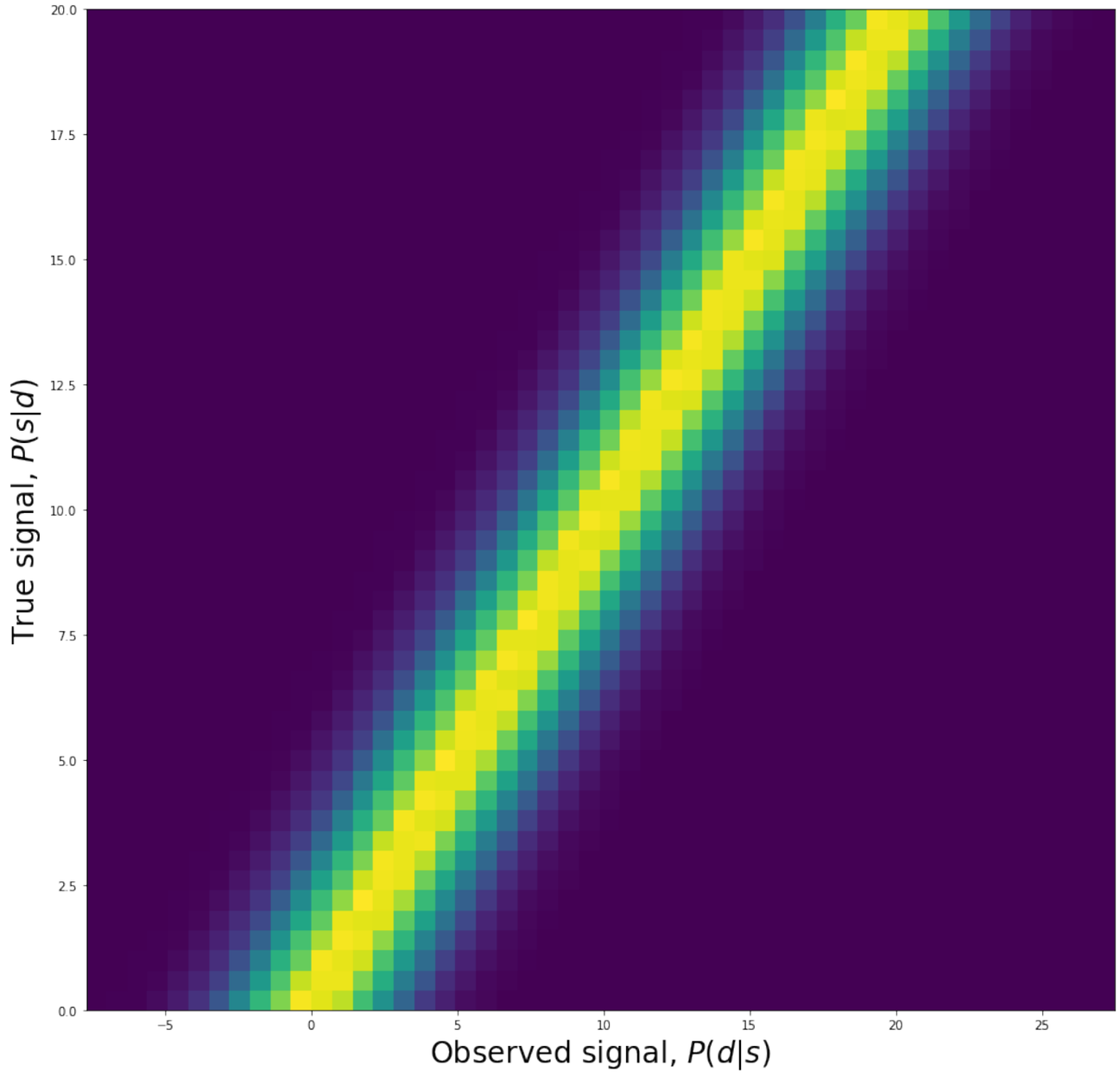
Slices



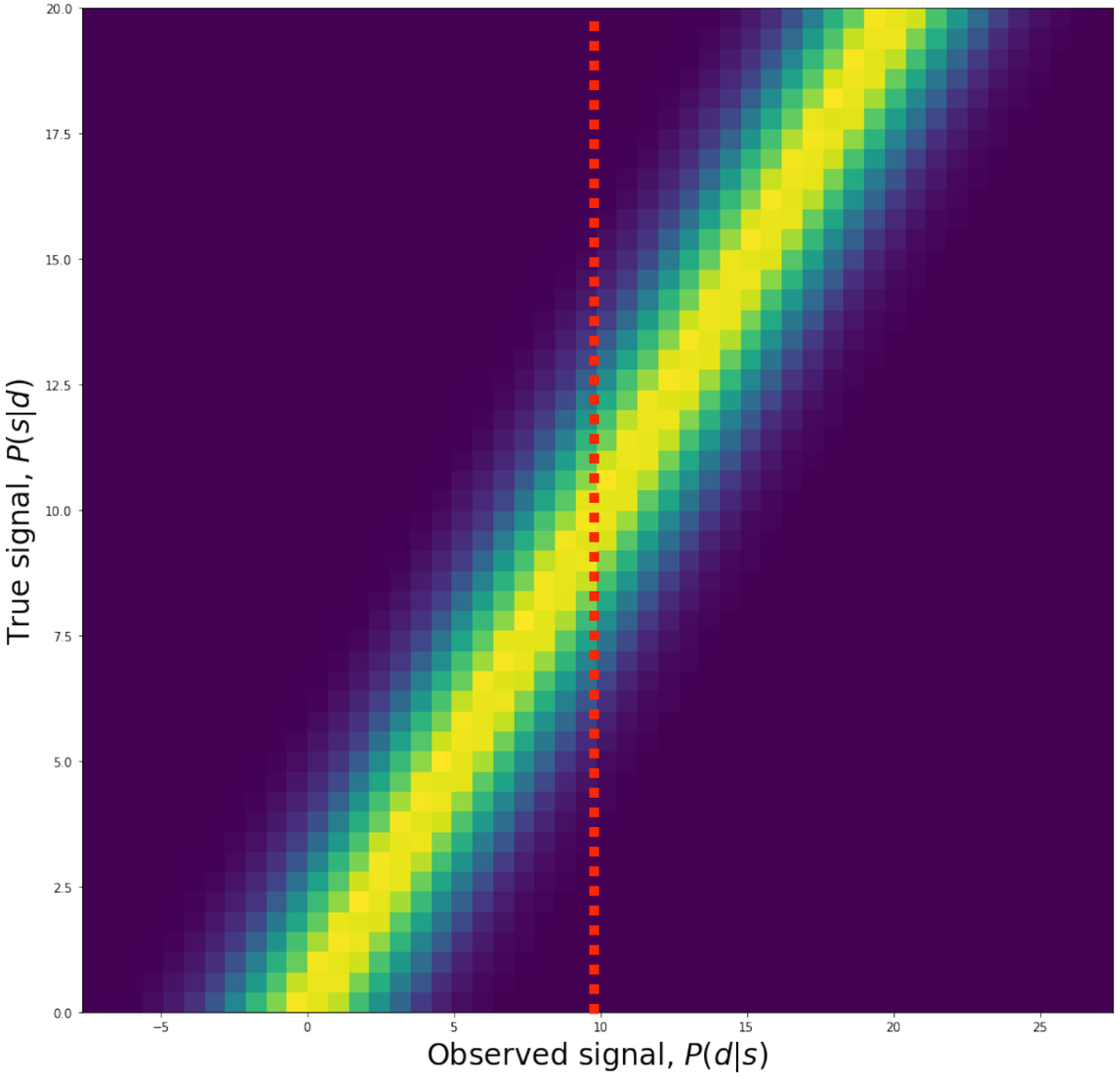


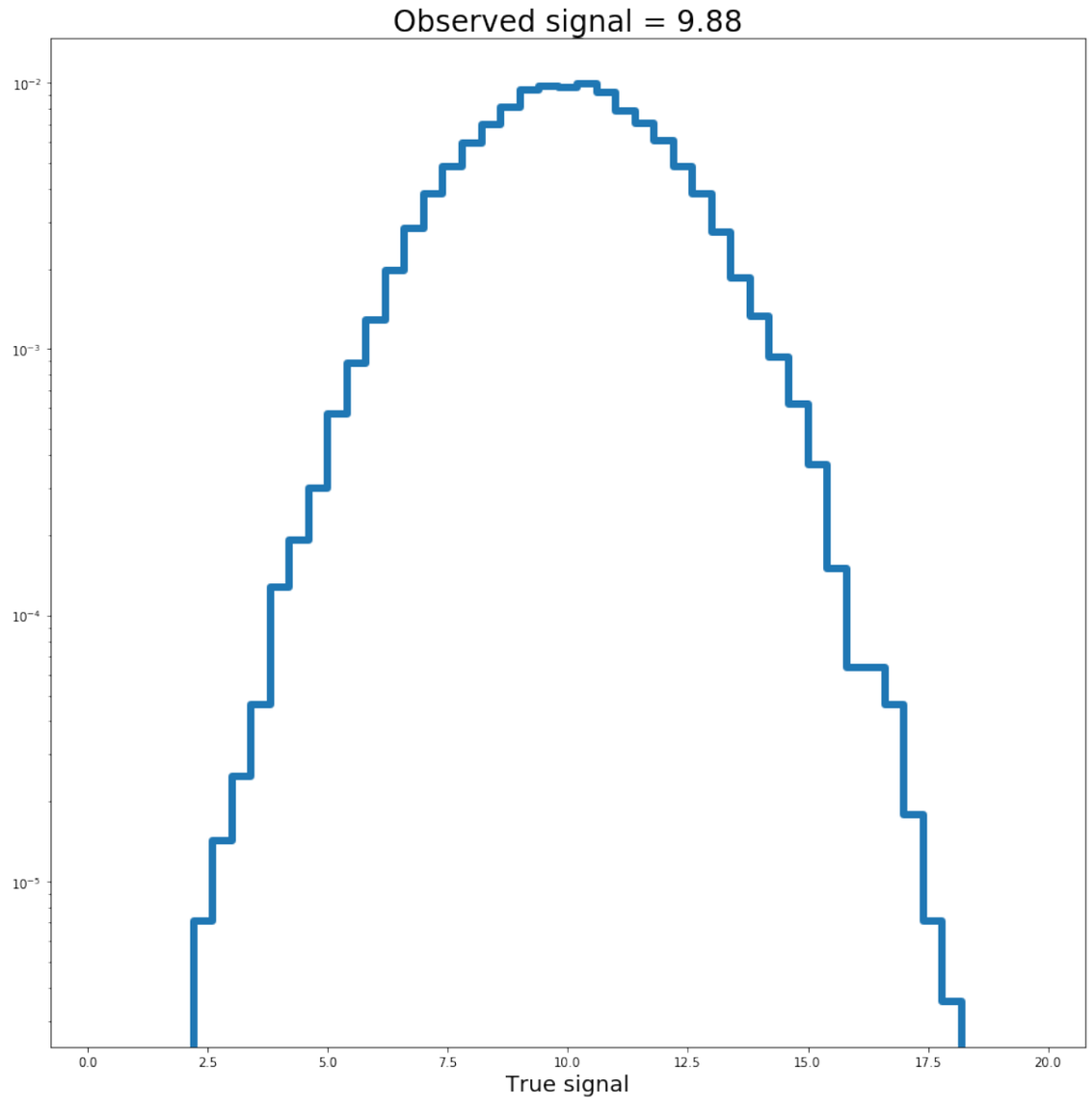
$$P(d \mid s)$$

What if I measure data value X and what to know what the range of true signals might be?



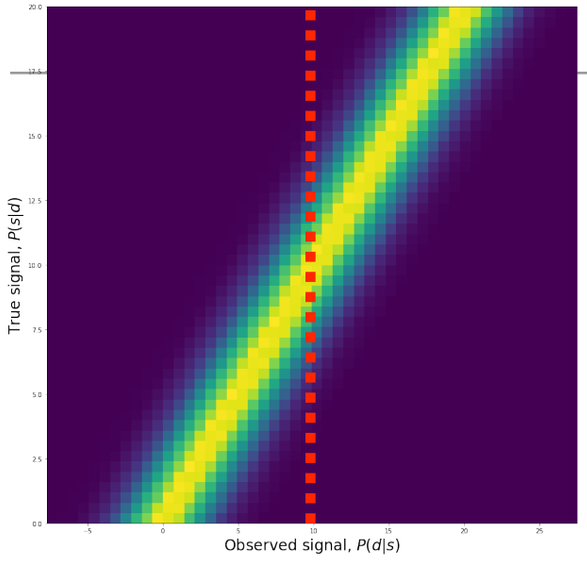
Slices



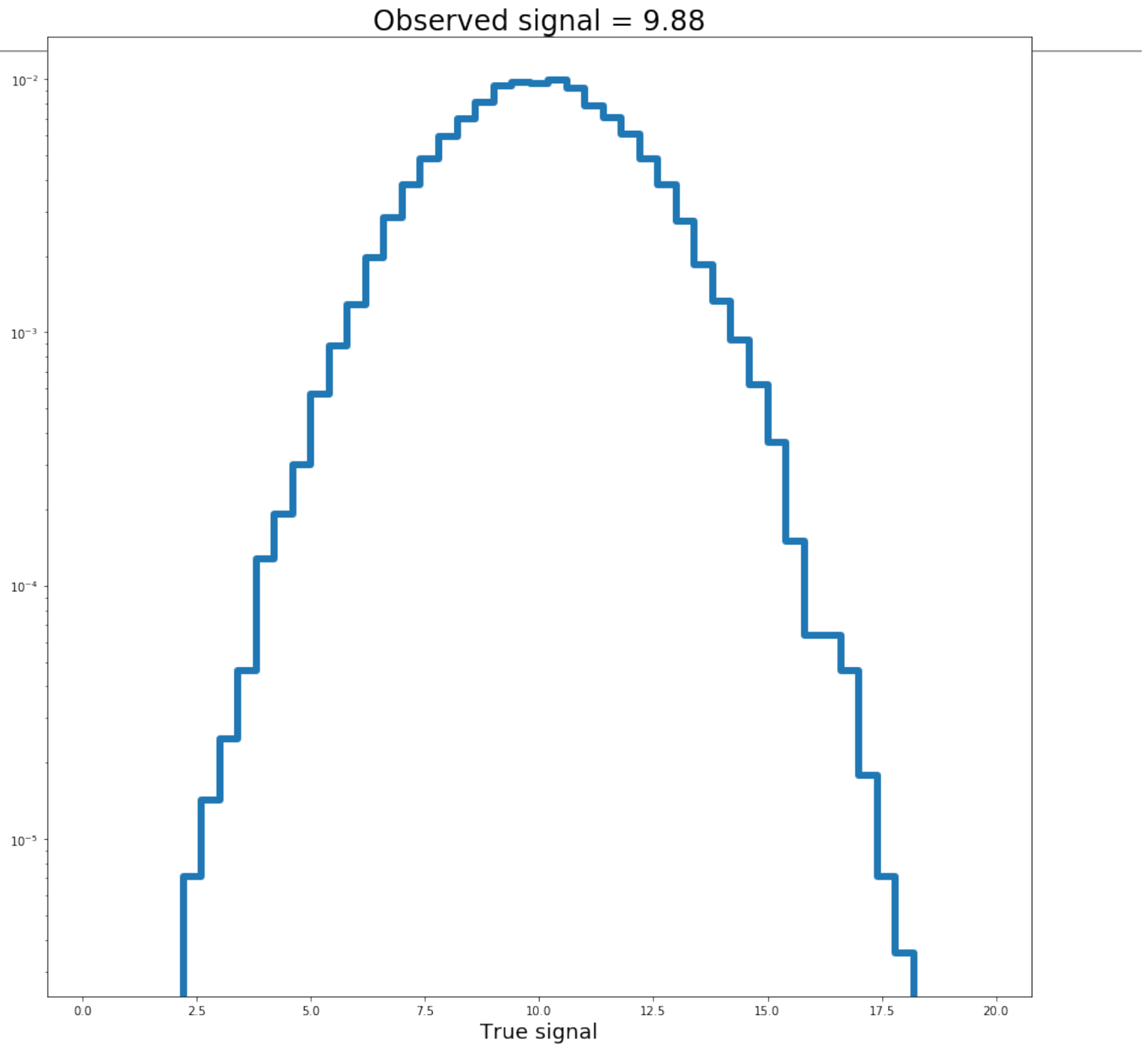


$$P(s \mid d)$$

Confidence interval



9.88 ± 2.0



Matlab & upper limits

Examples of mistakes

Careful with priors

- Useful when adding information to previous results
- Usually want ‘flat’ or ‘uninformative’ priors

$$P(s \mid d) = \frac{P(d \mid s)P(s)}{P(d)}$$