Spectral analysis: decompose a signal into its periodic constituents

Given basis signals $s_1(t), s_2(t), s_3(t), s_4(t) ... s_n(t)$,

$$W(t) = a_1 s_1(t) + a_2 s_2(t) + a_3 s_3(t) + a_4 s_4(t) + \ldots + a_n s_n(t)$$

Spectral analysis as lossy data compression

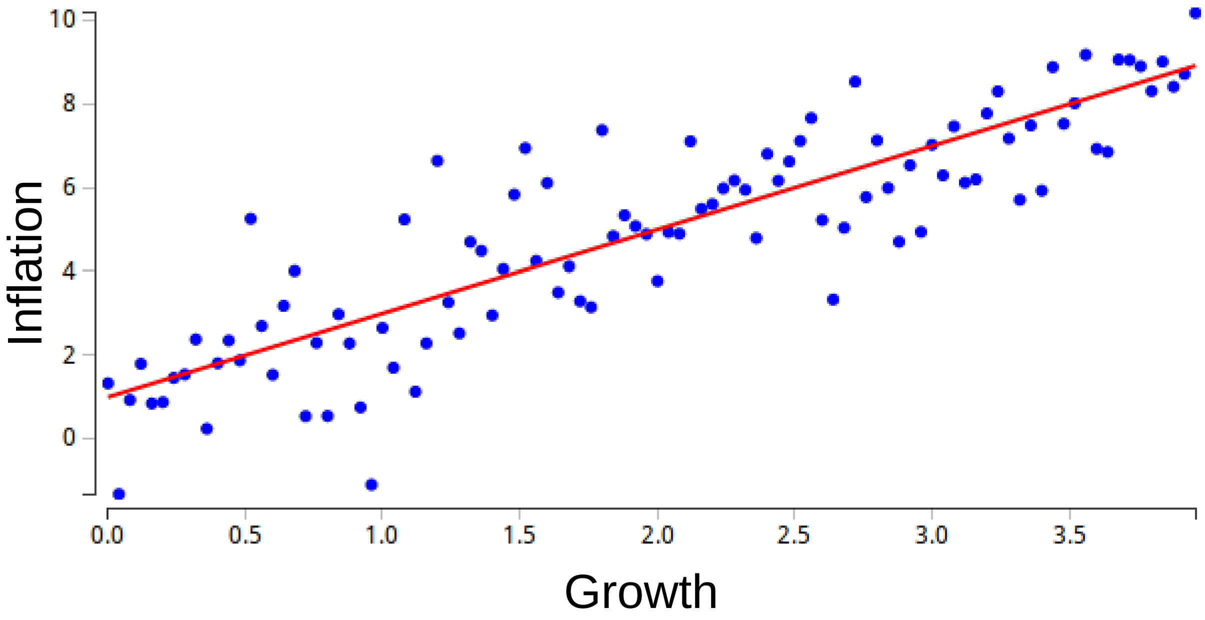
Assuming basis functions have intrinsically identical amplitudes, represent a signal as the sum of the largest-coefficient components:

$$W(t) = a_1 s_1(t) + a_2 s_2(t) + a_3 s_3(t) + a_4 s_4(t) + \ldots + a_n s_n(t)$$

$$\downarrow$$

$$W(t) = a_1 s_1(t) + a_3 s_3(t) + a_8 s_8(t)$$

Statistical inference as lossy compression



The Slow Fourier Transform

Tom Nielsen
OpenBrain
http://openbrain.co.uk

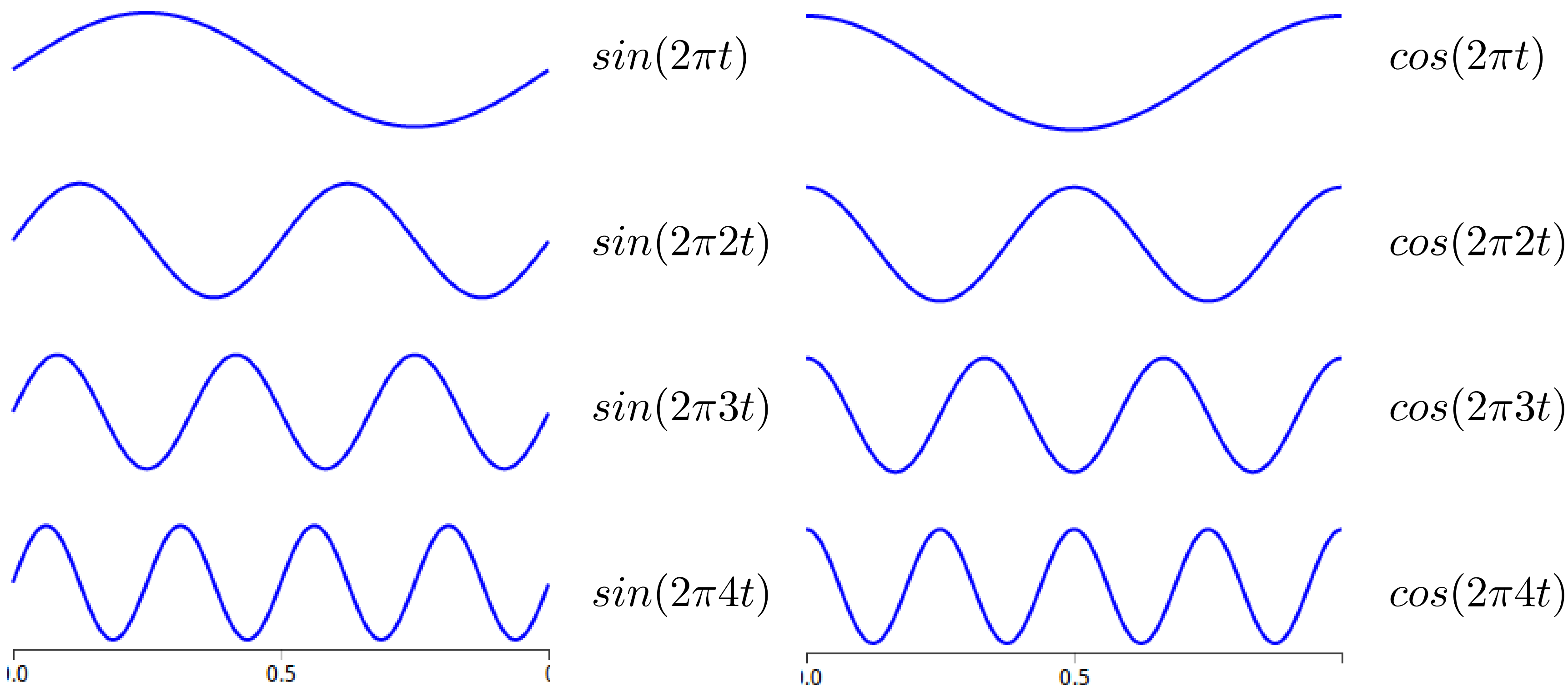
PHIL GREGORY

Bayesian Logical Data Analysis for the Physical Sciences

A Comparative Approach with *Mathematica*® Support

CAMBRIDGE

Fourier Spectral Analysis



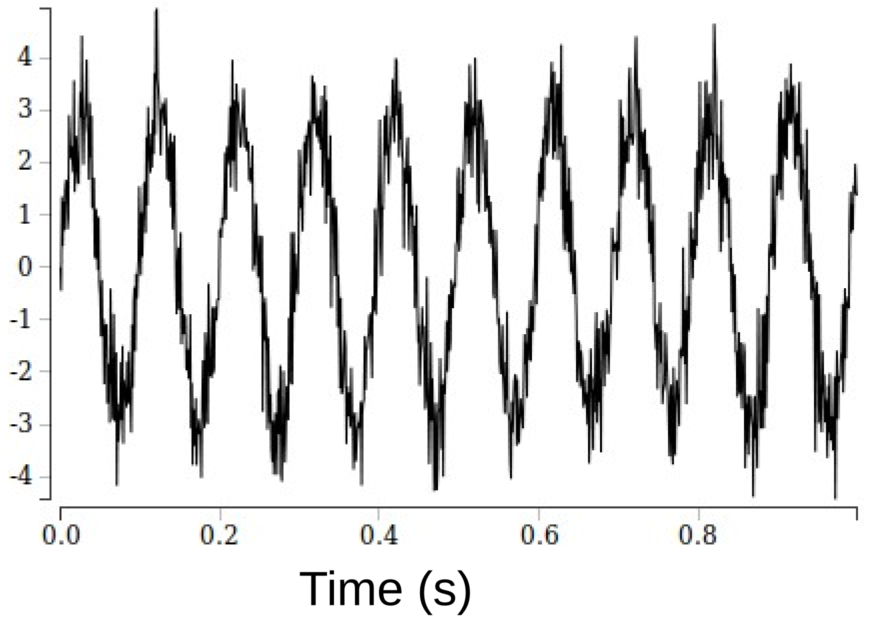
$$s(x) = \frac{a_0}{2} + \sum_{n=1}^{N} \left(a_n \cos(\tfrac{2\pi nx}{P}) + b_n \sin(\tfrac{2\pi nx}{P})\right)$$

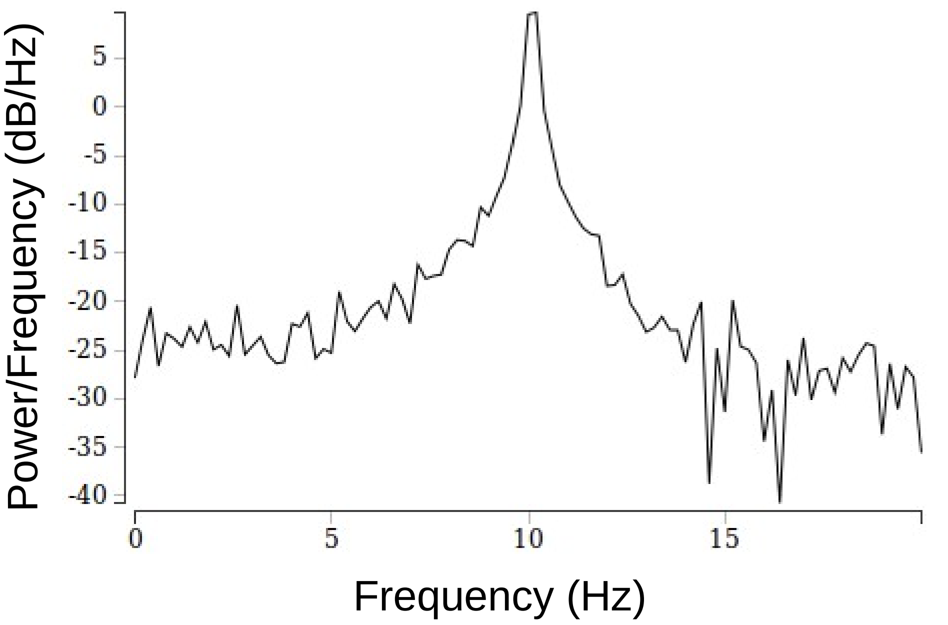
$$s(x) = \frac{a_0}{2} + \sum_{n=1}^{N} \left(a_n \cos(\tfrac{2\pi n x}{P}) + b_n \sin(\tfrac{2\pi n x}{P})\right)$$
$$= \sum_{n=-N}^{N} c_n \cdot e^{i\frac{2\pi n x}{P}}$$

$$c_n = \sum_{k=0}^{N-1} s_k \cdot e^{-i2\pi kn/N}$$

Periodogram $C(n) = \frac{1}{N}|c_n|^2$

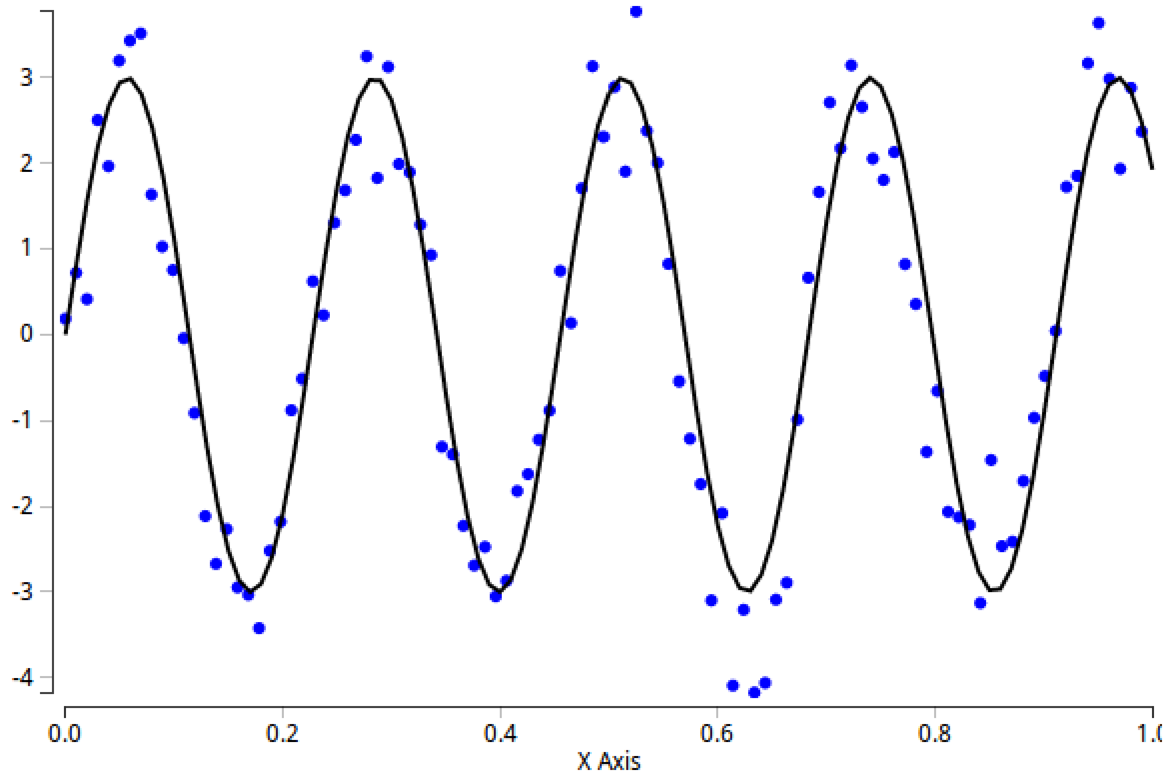
DFT periodogram





A statistical model for periodic data

$$y(t) = A \cdot sin(p(t - t_0)) + \epsilon_t$$
$$\epsilon_t \sim \mathcal{N}(0, \sigma)$$



A Bayesian statistical model for periodic data

$$A \sim Gamma(1, 10)$$

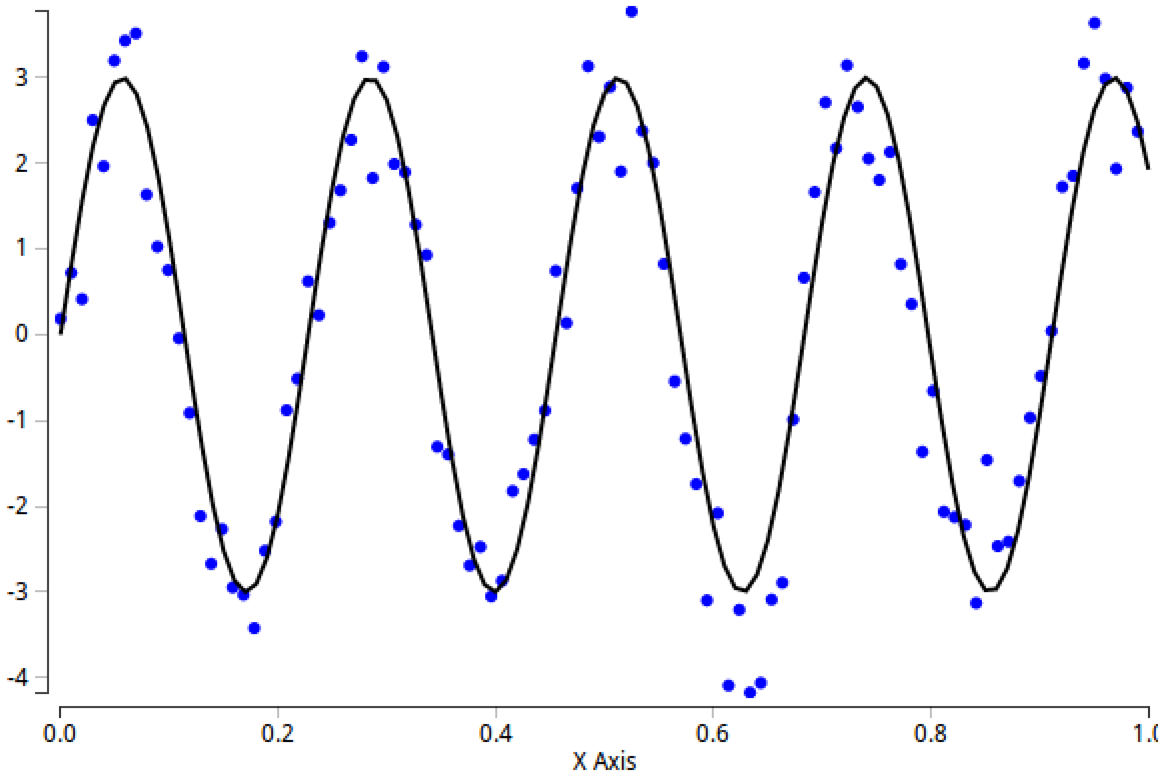
$$p \sim Gamma(1, 50)$$

$$\sigma \sim Gamma(1, 1)$$

$$t_0 \sim \mathcal{U}(0, 1)$$

$$\epsilon_t \sim \mathcal{N}(0, \sigma)$$

$$y_t = A \cdot sin(p(t - t_0)) + \epsilon_t$$



Bayes' Theorem:

$$P(\theta|\mathcal{D}) \propto \mathcal{L}(\mathcal{D}|\theta) \cdot P(\theta)$$

$$P(\omega|D,\sigma,I) \propto \exp\left\{\frac{C(\omega)}{\sigma^2}\right\}$$

HOWTO Bayesian inference

$$A \sim Gamma(1, 10)$$

$$p \sim Gamma(1, 50)$$

$$\sigma \sim Gamma(1, 1)$$

$$t_0 \sim \mathcal{U}(0, 1)$$

$$y_t \sim \mathcal{N}(A \cdot sin(p(t - t_0)), \sigma)$$

```
posterior_log_pdf (A,p,sigma,t_0,y[],t[])=
    gamma_log_pdf(1,10,A)
  + gamma_log_pdf(1,50,p)
  + gamma_log_pdf(1,1)
  + uniform_log_pdf(0,1,t_0)
  + sum(i in 1 to y.length;
      normal_log_pdf ( A*sin(p*(t[i]-t0)),
                       sigma,
                       y[i]) )

parameters := MCMC(posterior_log_pdf)
```

Bayes' Theorem:

$$P(\theta|\mathcal{D}) \propto \mathcal{L}(\mathcal{D}|\theta) \cdot P(\theta)$$

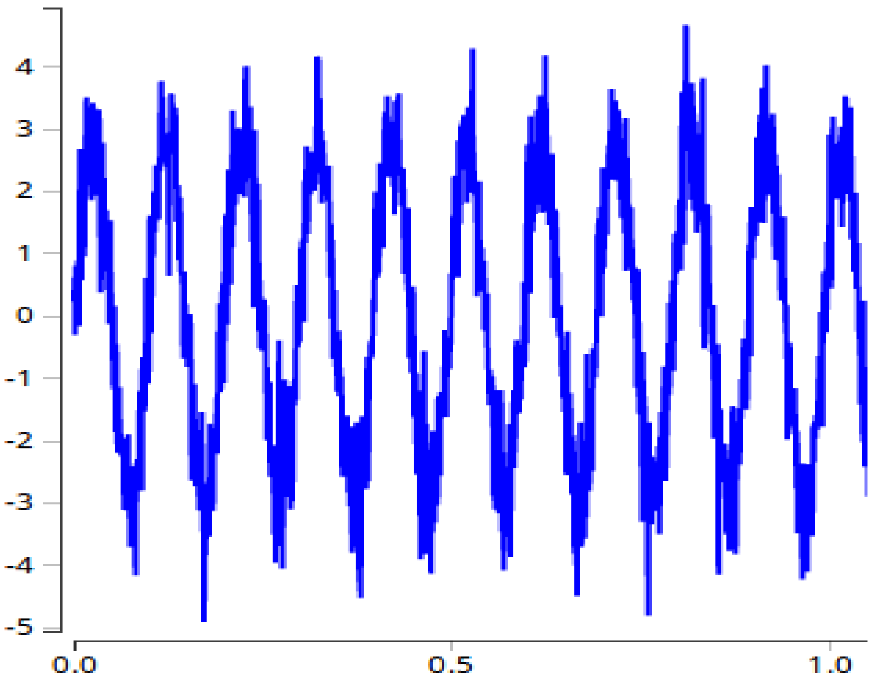
Starting values?
Adaptation parameters?
Simulating data?

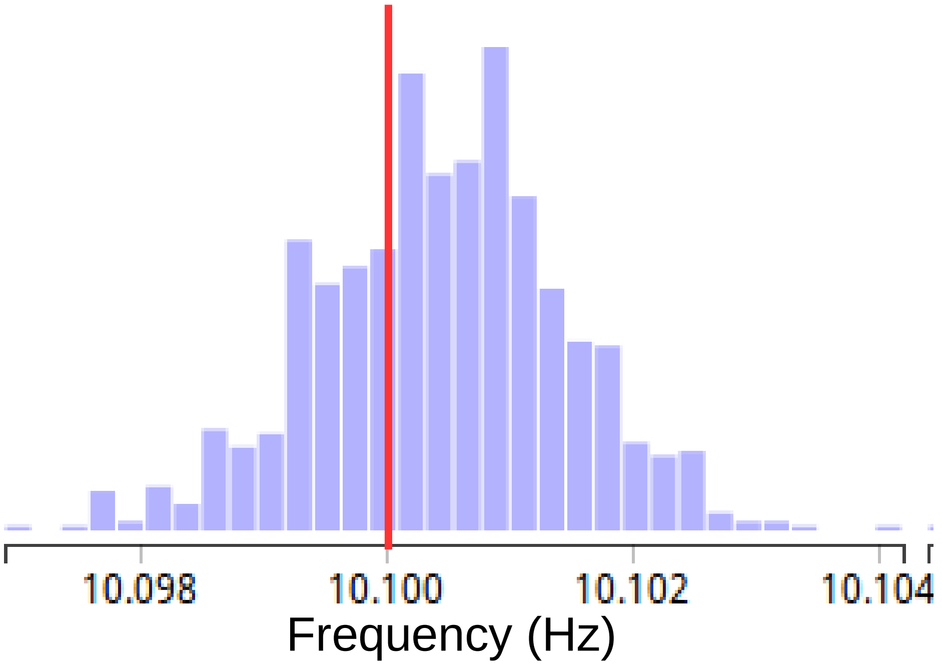
Stan implementation (http://mc-stan.org)

```
model {
   per    ~ gamma(1, 0.1);
   amp    ~ gamma(1, 0.1);
   noise  ~ gamma(1, 1.0);
   t0     ~ uniform(0, 6.28);
   for (i in 1:N) {
      y[i] ~ normal(amp*sin(per*(t[i]*(2*pi)-t0)), noise);
   }
}
data {
   int N;
   real t[N];
   real y[N];
}
parameters {
   real<lower=0> per;
   real<lower=0> amp;
   real<lower=0> noise;
   real<lower=0, upper=6.280> t0;
}
```

Baysig implementation (https://bayeshive.com)

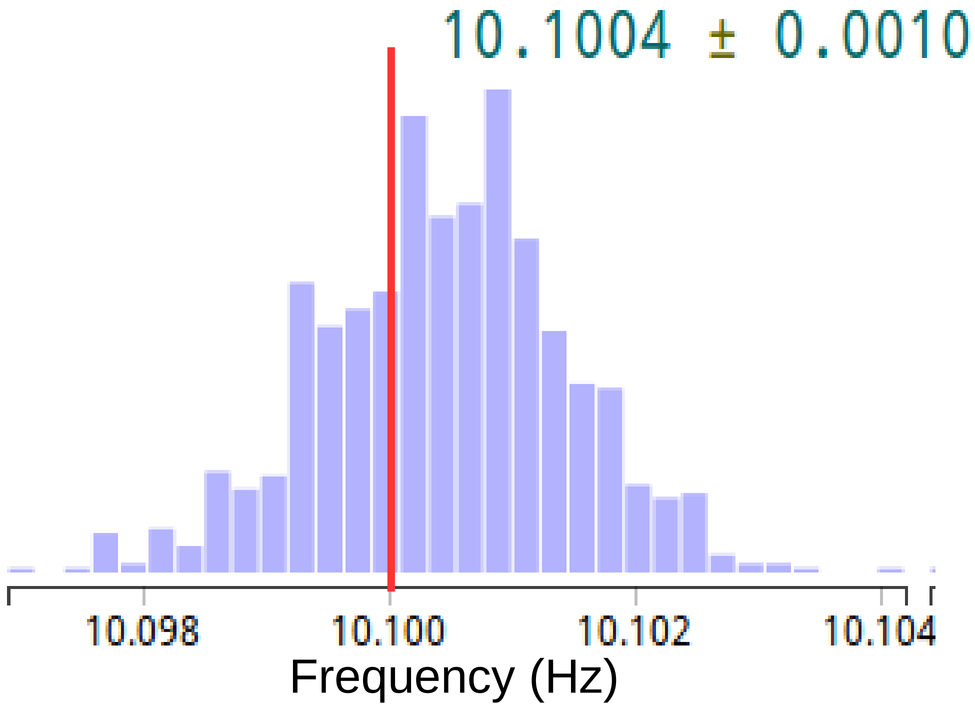
```
model = prob
  per ~ gamma 1 200
  amp ~ gamma 1 10
  v   ~ gamma 1 1.1
  t0  ~ uniform 0 6.28
  repeat 100 $ prob
    t ~ uniform 0 tmax
    y ~ normal (amp*sin (per*(t*2*pi - t0))) v
    return (t,y)

model1 <- update model $ return { per => 175;
                                  amp => 3;
                                  v => 0.5;
                                  start => 1 }

noisy0 <- sample model1
pars <- estimate model noisy0

?> scatterPlot noisy0
?> pars
```

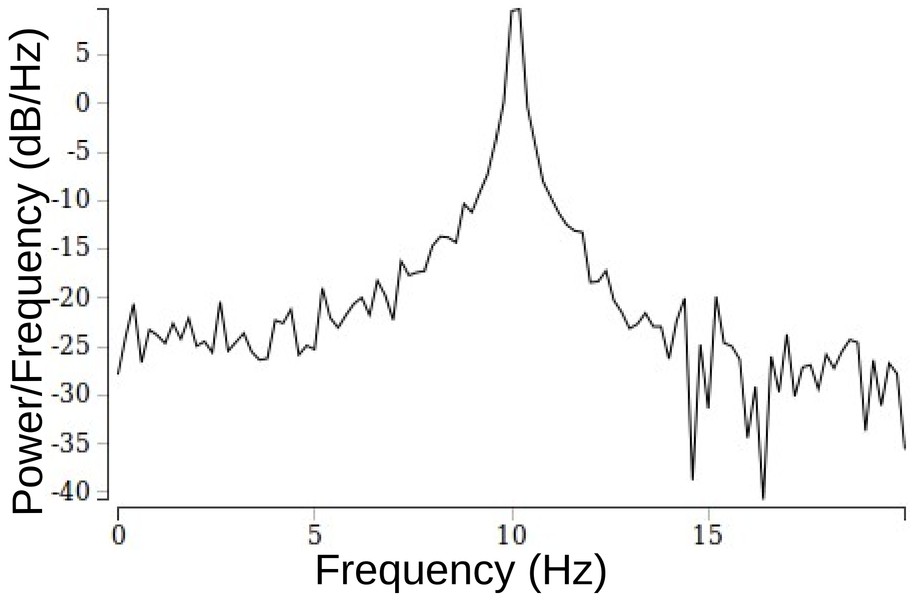




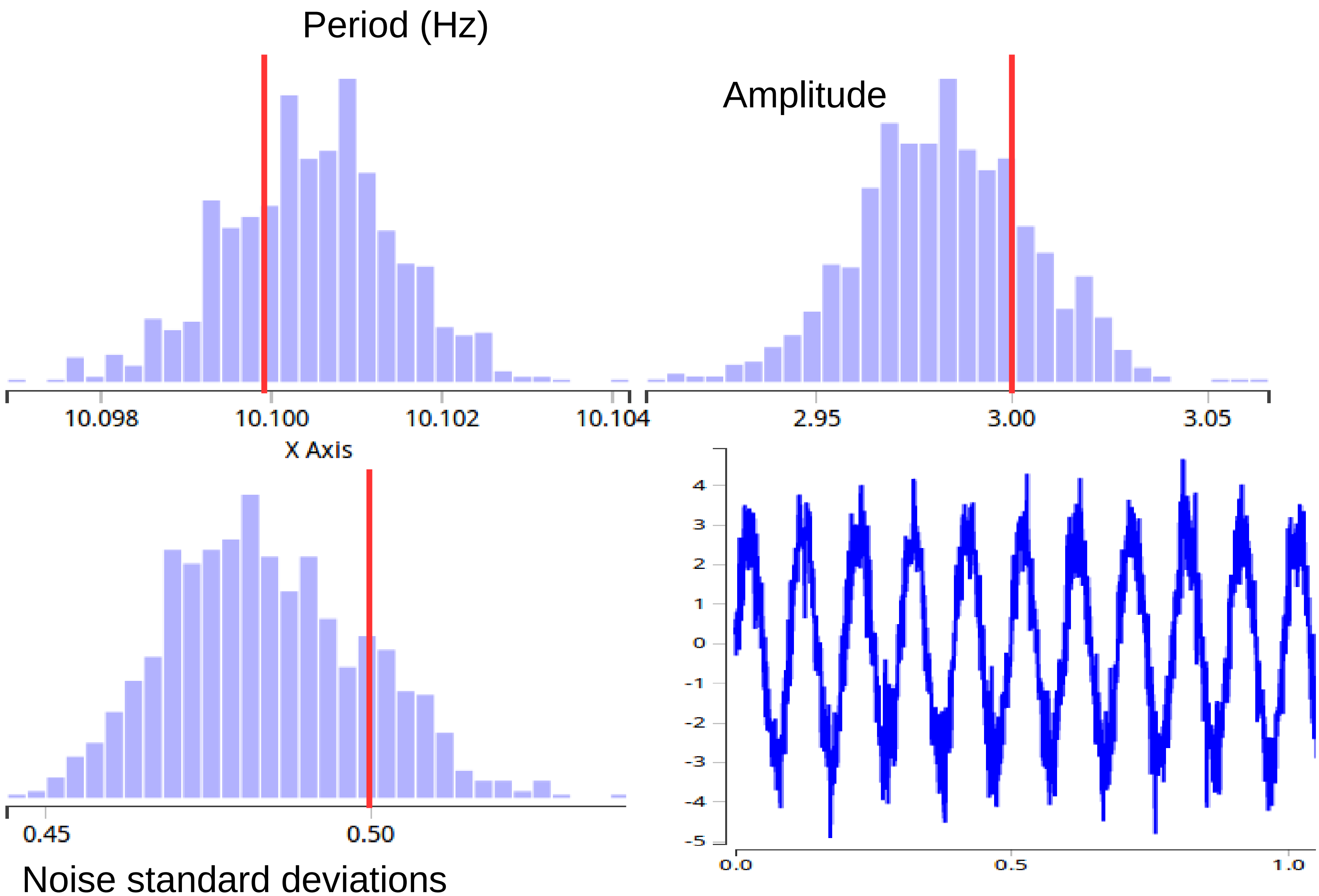
Three orders of magnitude slower!

Sub-bin resolution

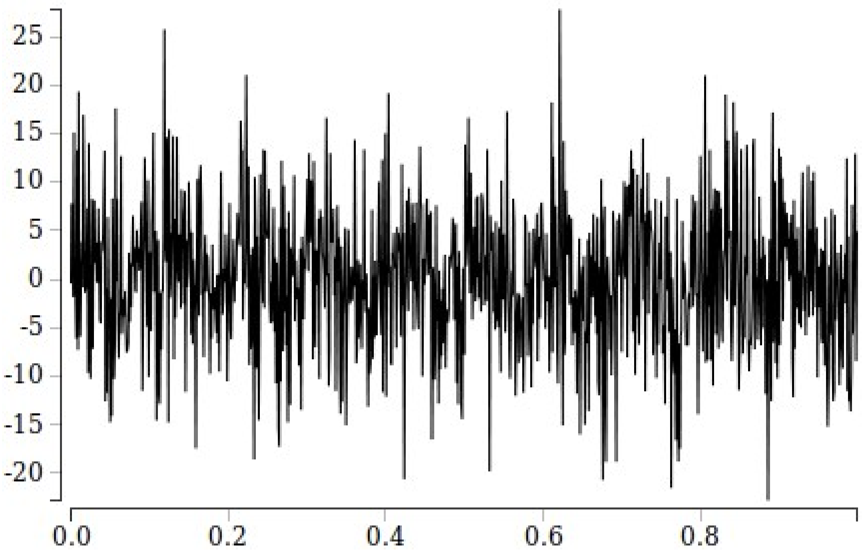


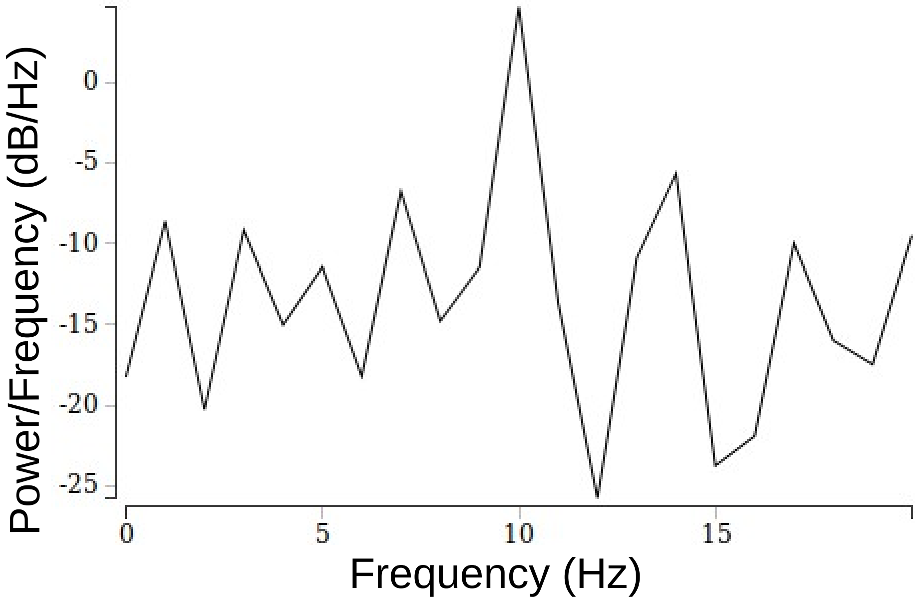


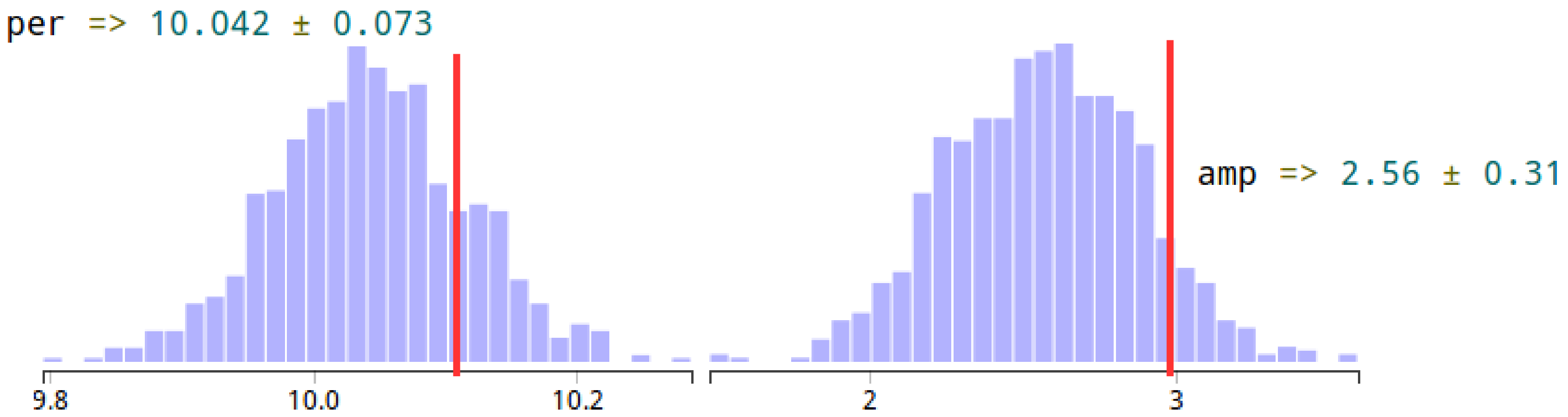
More information



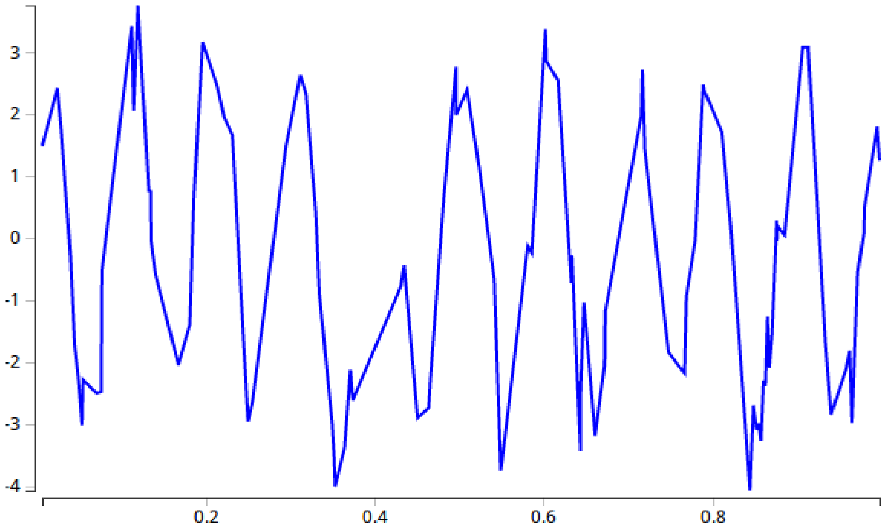
Noise sensitivity

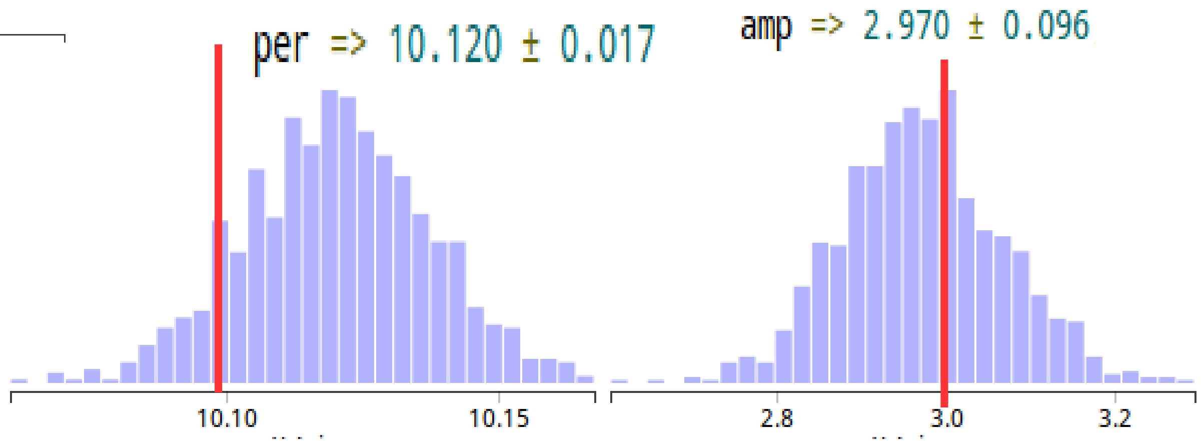




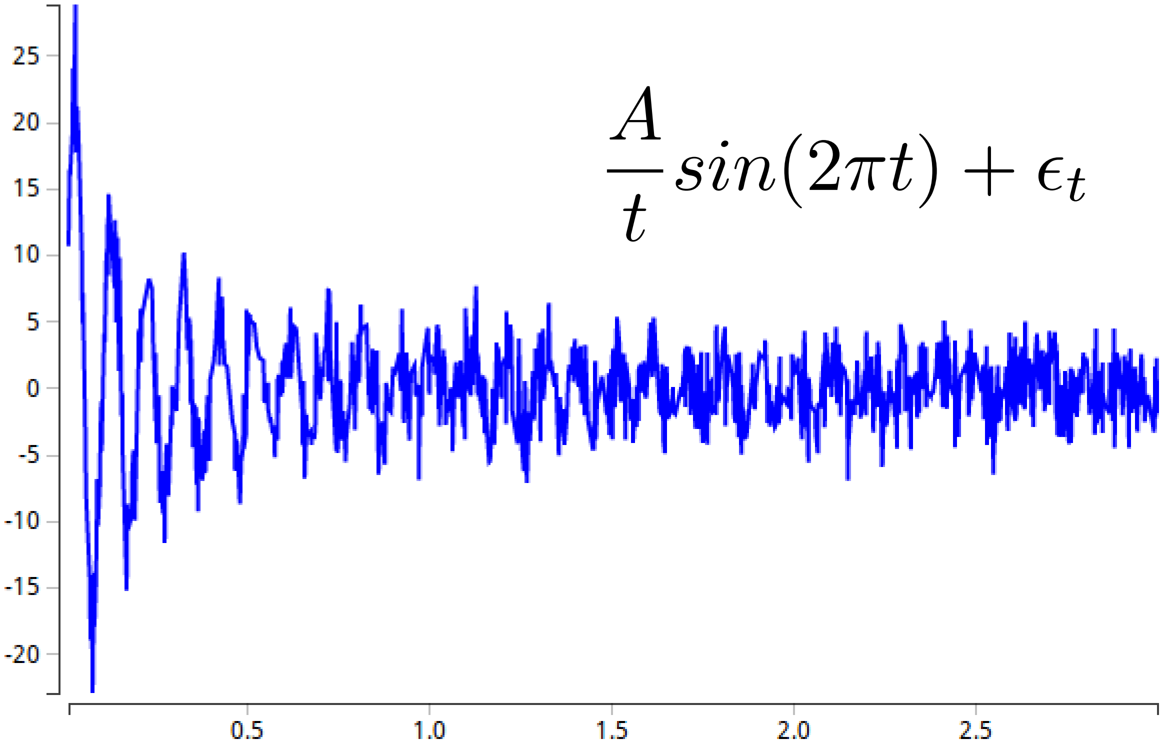


Missing / irregularly sampled data

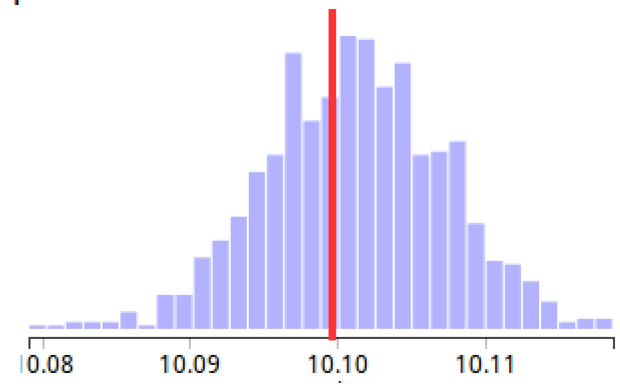


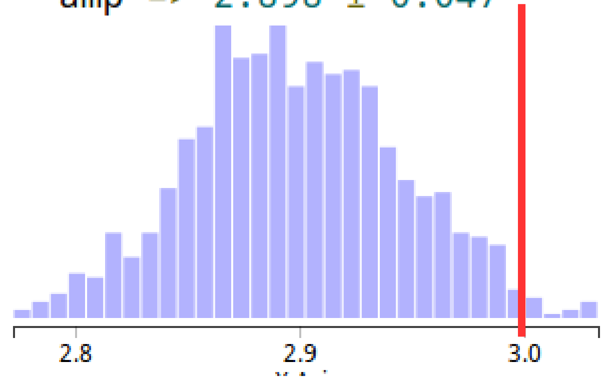


Vanishing periodicity

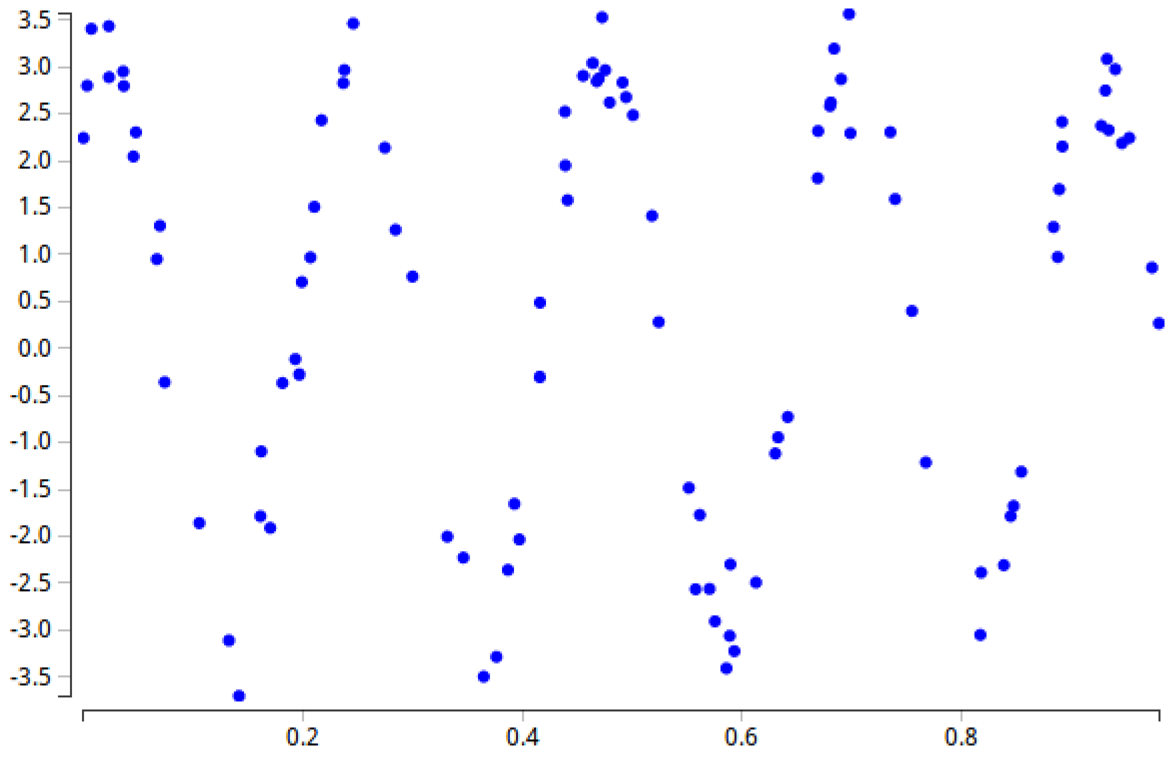


$$\frac{A}{t} sin(2\pi t) + \epsilon_t$$

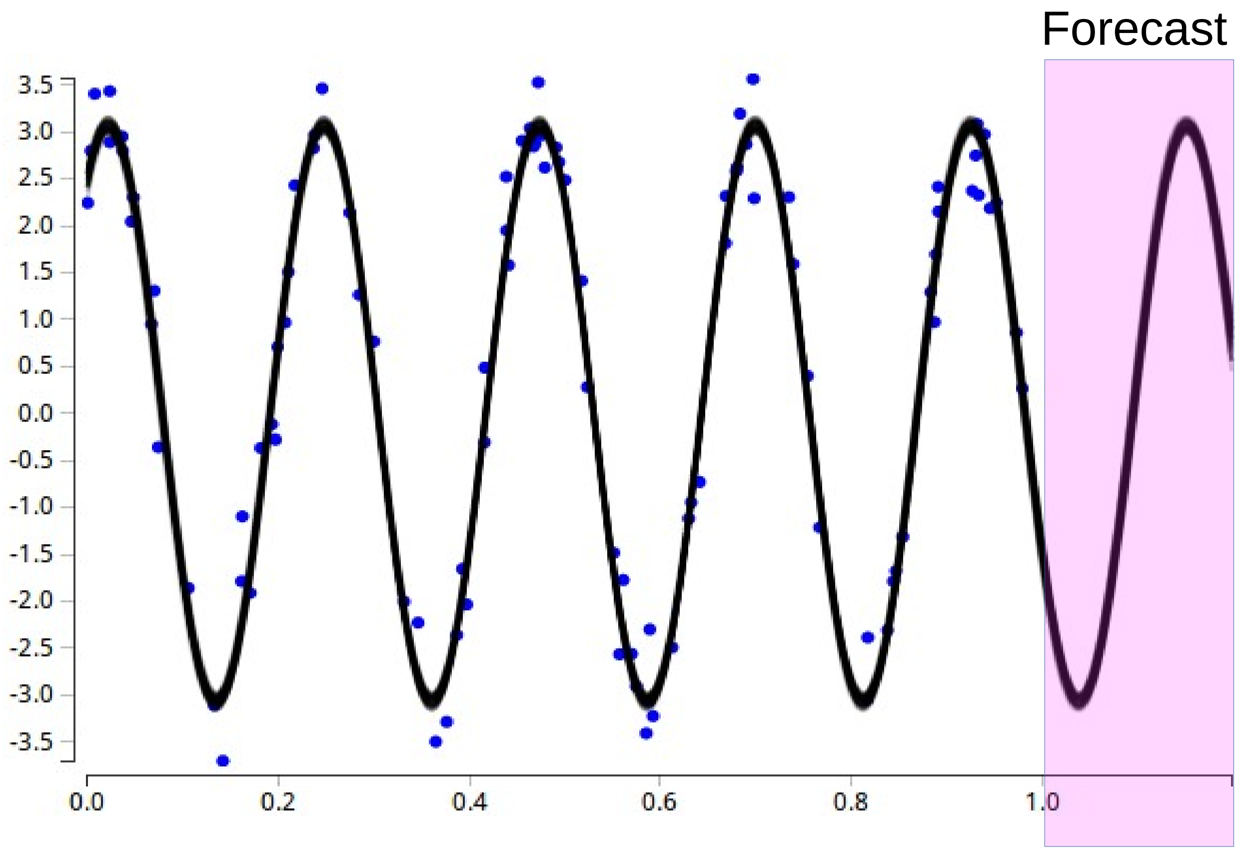




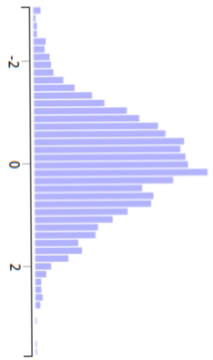
HOWTO Forecasting

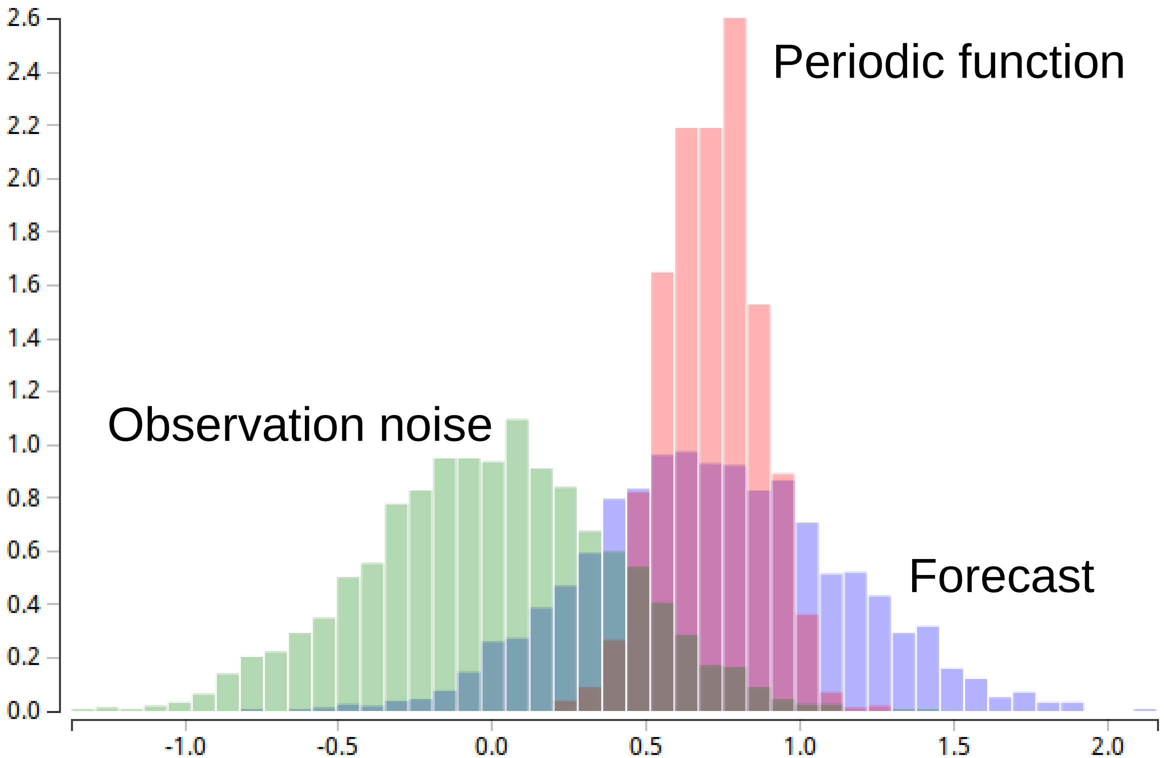


HOWTO Forecasting

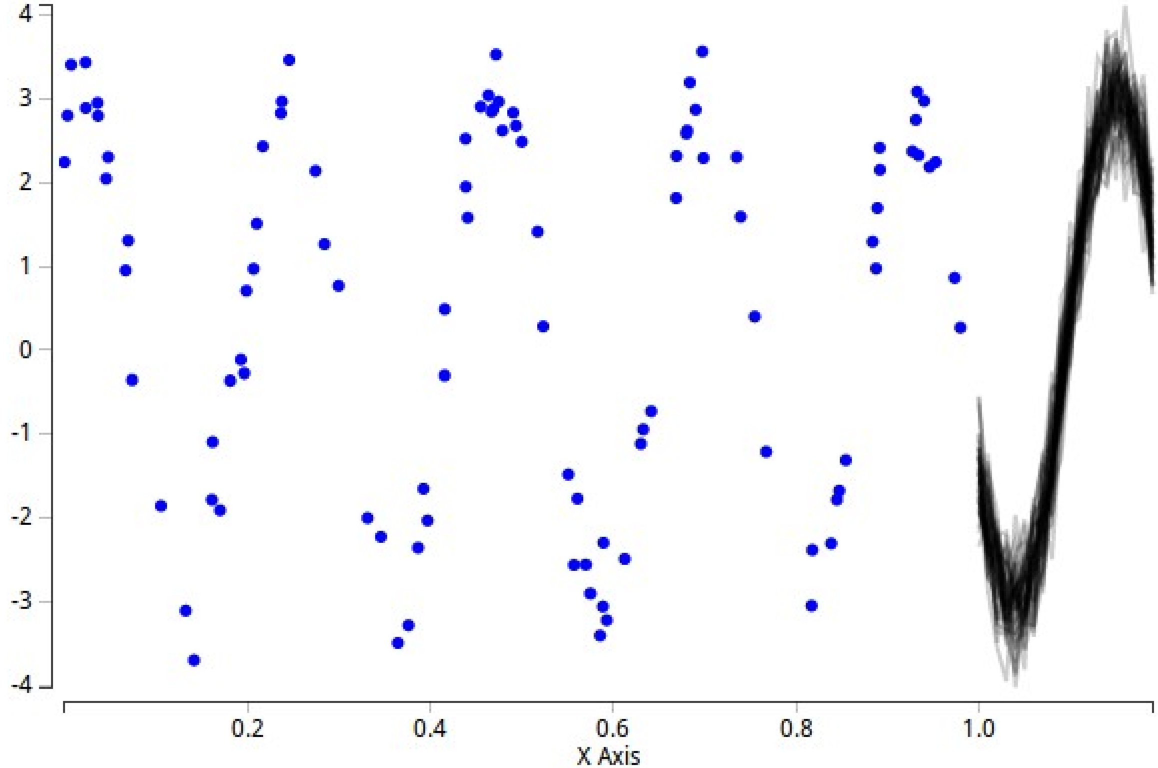


+



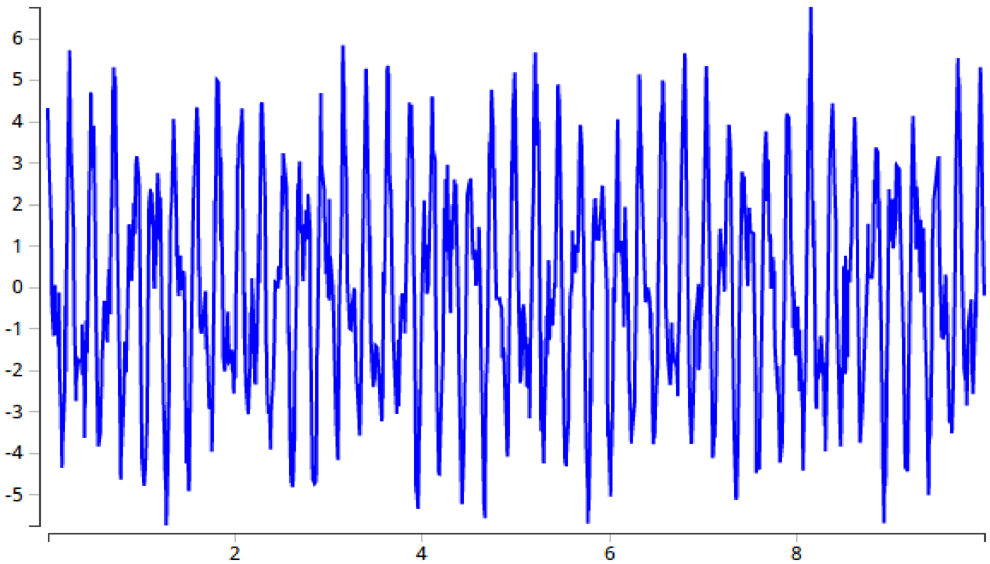


Periodic function
Observation noise
Forecast
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0

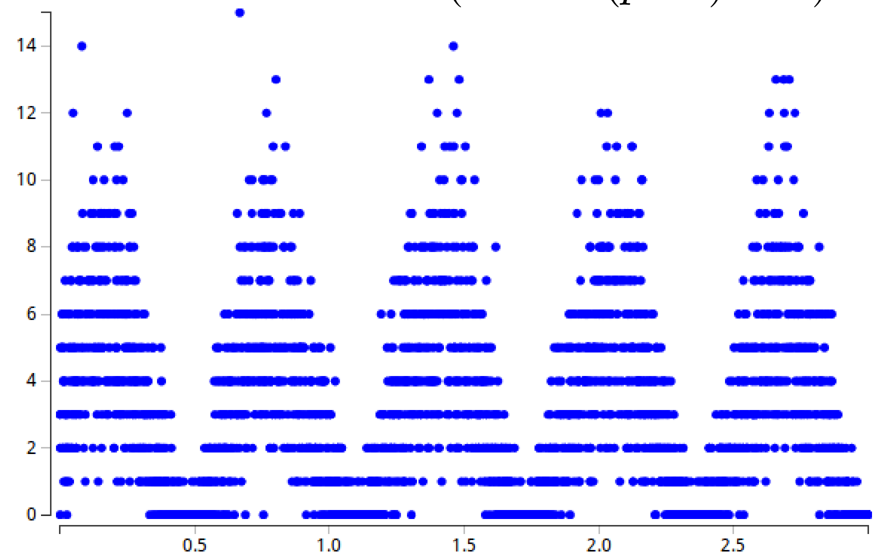

4
3
2
1
0
-1
-2
-3
-4
0.2
0.4
0.6
0.8
1.0
X Axis

Robustness – Inference failure

Two components



$$count \sim Poisson(k * sin(p * t) + k)$$



I hope to have convinced you that probabilistic programming gives you

- Flexibility
- Applicability
- Accuracy
- Correct uncertainty

While writing this talk I have convinced myself that

- Multi-modal posteriors can be a serious problem for black-box Bayesian inference
- Need to look at parallel tempering
- Spectral analysis good benchmark for multi-modal inference

References+Links

- Gregory: Bayesian Logical Data Analysis for the Physical Sciences (2005)
- Gregory: A Bayesian revolution in spectral analysis (AIP 2000)
- Sivia: Data Analysis – a Bayesian Tutorial (2006)
- https://bayeshive.com
- http://mc-stan.org
- http://www.skybluetrades.net/haskell-fft-index.html