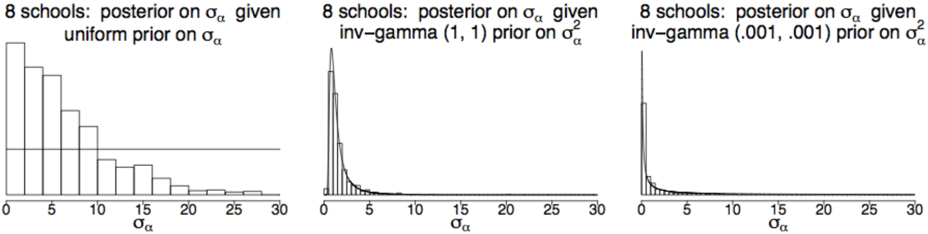
7. Open problems in Bayesian data analysis

- Modeling
- Computing
- Model checking and workflow

Open problems in Bayesian modeling

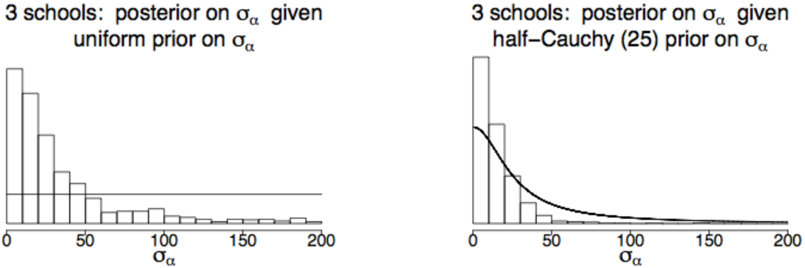
- The 3-schools problem
- A surprisingly tricky nonlinear model
- Formulating big models

For the 8 schools, uniform prior is ok



- Inv-gamma prior is *not* "noninformative"
- What is the effect of using a bad prior here?

For the 3 schools, we need a stronger prior



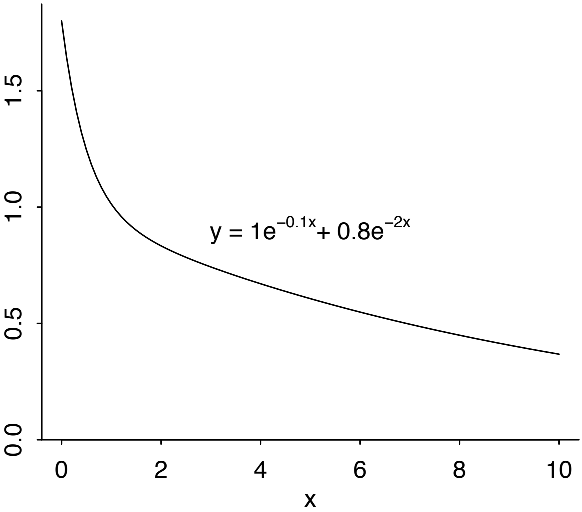
- Uniform prior doesn't cut off the long tail
- Must add real prior information

Priors!

- General guidelines
- What happens if your prior is too weak? Too strong?

A surprisingly tricky model

- Sum of declining exponentials: $y = a_1 e^{-b_1 x} + a_2 e^{-b_2 x}$
- Statistical version: $y_i = \left(a_1 e^{-b_1 x_i} + a_2 e^{-b_2 x_i}\right) \cdot \epsilon_i$



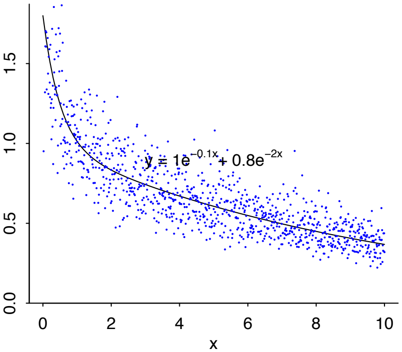
Stan code

```
data {
  int N;
  vector[N] x;
  vector[N] y;}
parameters {
  vector[2] log_a;
  ordered[2] log_b;
  real<lower=0> sigma;}
transformed parameters {
  vector<lower=0>[2] a;
  vector<lower=0>[2] b;
  a = exp(log_a);
  b = exp(log_b);}
model {
  vector[N] ypred;
  ypred = a[1]*exp(-b[1]*x) + a[2]*exp(-b[2]*x);
  y ~ lognormal(log(ypred), sigma);
}
```

Simulate fake data in R

```
a <- c(1, 0.8)
b <- c(0.1, 2)
sigma <- 0.2

x <- (1:1000)/100
N <- length(x)
ypred <- a[1]*exp(-b[1]*x) + a[2]*exp(-b[2]*x)
y <- ypred*exp(rnorm(N, 0, sigma))
```



Fit the model in Stan

- Remember true values:

```
a <- c(1, 0.8)
b <- c(0.1, 2)
sigma <- .2
```

```
Inference for Stan model: exponentials.
4 chains, each with iter=1000; warmup=500; thin=1;
post-warmup draws per chain=500, total post-warmup draws=2000.

      mean se_mean   sd  25%  50%  75% n_eff Rhat
a[1]  1.00    0.00 0.03 0.99 1.00 1.02   494    1
a[2]  0.70    0.00 0.08 0.65 0.69 0.75   620    1
b[1]  0.10    0.00 0.00 0.10 0.10 0.10   532    1
b[2]  1.71    0.02 0.34 1.48 1.67 1.90   498    1
sigma 0.19    0.00 0.00 0.19 0.19 0.20   952    1
```

Try again with new parameter values

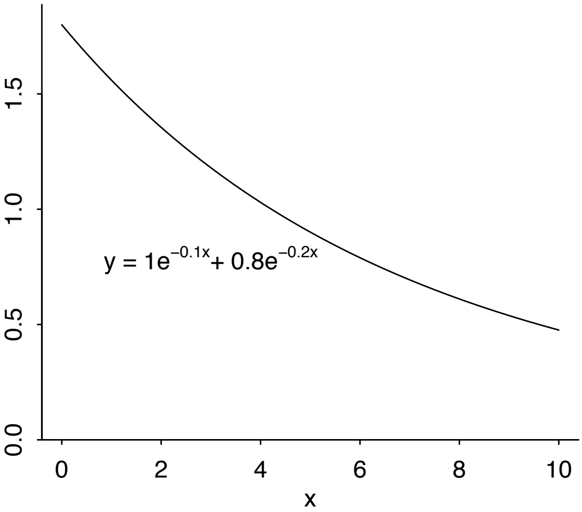
- Simulate new data using these new parameter values:

```
a <- c(1, 0.8)
b <- c(0.1, 0.2)
```

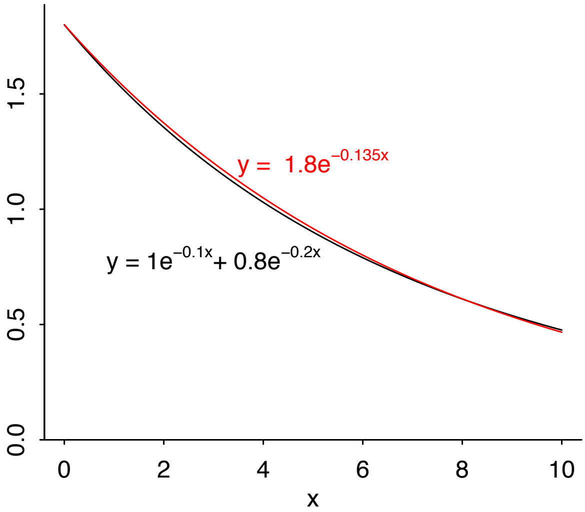
- Then fit the model:

```
            mean se_mean   sd  25%       50%       75% n_eff Rhat
a[1]    1.33e+00    0.54 0.77 1.28  1.77e+00  1.79e+00     2 44.2
a[2]   2.46e+294     Inf  Inf 0.00  0.00e+00  1.77e+00  2000  NaN
b[1]    1.00e-01    0.04 0.06 0.10  1.30e-01  1.30e-01     2 33.6
b[2]   3.09e+305     Inf  Inf 0.50 1.15e+109 4.77e+212  2000  NaN
sigma   2.00e-01    0.00 0.00 0.19  2.00e-01  2.00e-01    65  1.0
```

What went wrong?



What went wrong?



Informative prior distribution

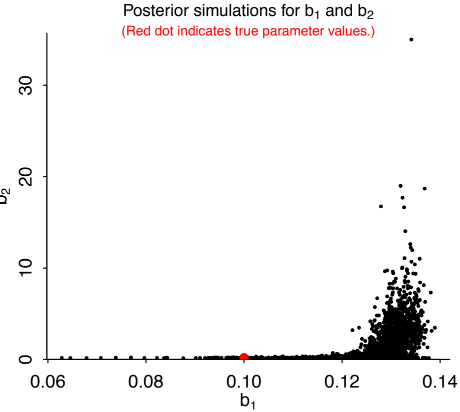
```
log_a ~ normal(0, 1);
log_b ~ normal(0, 1);
```

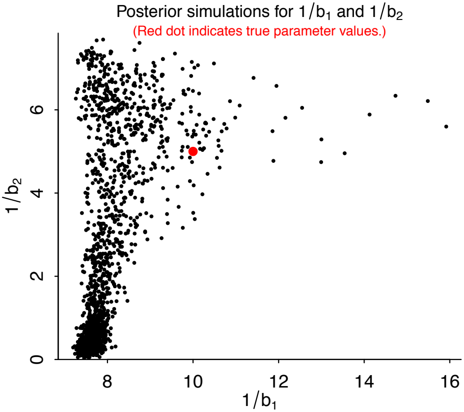
Happy ending

```
a <- c(1, 0.8)
b <- c(0.1, 0.2)
sigma <- 0.2

      mean se_mean   sd  25%  50%  75% n_eff Rhat
a[1]  1.56    0.09 0.32 1.52 1.72 1.75    13 1.25
a[2]  0.32    0.08 0.28 0.14 0.22 0.37    13 1.20
b[1]  0.13    0.00 0.01 0.12 0.13 0.13    22 1.14
b[2]  1.94    0.20 2.29 0.22 1.26 3.00   127 1.05
sigma 0.20    0.00 0.00 0.19 0.20 0.20   656 1.00
```

Skewed posterior distribution





Open problems in Bayesian computing

- Covariance matrices
- Multiple modes
- Discrete parameters
- Approximate algorithms

Open problems in Bayesian workflow

- Fake-data checks
- Posterior predictive checks
- Model comparison
- The network of models

Summary

- Big data . . . messy data
- Clean up messy data . . . Big model
- Big model . . . Bayesian inference
- Bayesian inference . . . Stan
- It's not about "fitting a model," it's about the modeling, fitting, checking, improvement *process*

Let us have
the serenity to embrace the variation that we cannot reduce,
the courage to reduce the variation we cannot embrace,
and the wisdom to distinguish one from the other.