

(R)Stan live session

Måns Magnusson

Avd. Statistik, LiU

- BUGS (**B**ayesian inference **U**sing **G**ibbs **S**ampling)
- JAGS (**J**ust **A**nother **G**ibbs **S**ampler)
- Stan (**S**tanislav Ulam)

Why Stan?

- Easy to use (BUGS type)
- Effective samplers/fast
- Wrappers from Python and R (and MatLab in the future)
- Linux / OS X / Windows
- Active development (at GitHub)
- (Really) good documentation

mc-stan.org

- Six parts in a Stan model:
 - data
 - transformed data
 - parameters
 - transformed parameters
 - model*
 - generated quantities

- My example:

$$\text{weight}_i = \alpha + \beta \cdot \text{stHeight}_i + \epsilon_i$$

where

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

- We want to estimate α , β and σ (these need priors)

- Read in data (for example from R or Python) once
- Only variable declarations
- A lot of different data types
 - `int`, `real`, `vector`, `arrays`, `matrix` and more Stan specific (?) data types as `cholesky_factor_cov` and `unit_vector`

Example of the data block

```
data {  
  int<lower=0> n; # The number of observations  
  real<lower=0> my_height;  
  vector[n] height;  
  vector[n] weight;  
}
```

- Variable declarations **and** statements (done once)
- See chapter V in the documentation for all functions that can be used.

Example of the transformed data block

```
transformed data {  
  vector[n] SHheight;  
  real MySHheight;  
  SHheight <- (height - mean(height)) / sd(height);  
  MySHheight <- (my_height - mean(height)) / sd(height);  
}
```

- Parameters (that should be sampled)
- Parameter declarations only.

Example of the parameters block

```
parameters {  
  real alpha;  
  real beta;  
  real<lower=0> tau;  
}
```


- Parameter declarations **and** statements
- The transformations is done in each sampling step

Example of the transformed parameters block

```
transformed parameters {  
  real<lower=0> sigma;  
  sigma <- 1.0 / sqrt(tau);  
}
```

- Declare the priors and data with sampling statements ~
- Distributions can be found in chapter VI and VII in the documentation

Example of the model block

```
model {  
  // Priors  
  alpha ~ normal(0.0,1000.0);  
  beta ~ normal(0.0,1000.0);  
  tau ~ gamma(0.001,0.001);  
  # Model  
  for (i in 1:n)  
    weight[i] ~ normal(alpha + beta * SHeight[i], sigma);  
}
```

Model in Stan: generated quantities

- Computations after the sampling has been done
- Is used for:
 - model checking
 - predictive distributions for new data
 - applying full Bayesian decision theory
 - transforming parameters for reporting, etc

Example of the generated quantities block

```
generated quantities {  
  real my_weight_pred; // Can be negative.  
  my_weight_pred <- alpha + beta * MySHeight +  
  normal_rng(0, sigma);  
}
```

(R)Stan demonstration