

BAYESIAN MODELING
Homework Assignment 2
Instructor: Sascha Göbel
April 2025

Task 1 -----

Task 1 (1.5 points):

Your outcome variable of interest y_i is a judge's sentence (**sentence**), which takes the value 1 if a defendant was incarcerated and the value 0 if a defendant was put on probation. Given this outcome, you want to specify a varying-intercepts model with intercepts varying across (1) judges (**j_id**) and (2) years (**year**). Additionally, you want to add varying intercepts that capture (3) the judge-year-specific probability of incarceration/percentage incarcerated. Write this model down formally, fully specifying the likelihood and all priors. Is this a nested or non-nested model? Explain your decision. Upload the answers to this task in a .pdf file.

```

varying_intercepts_model <- "
data {
  int<lower=1> N;           // Number of observations
  int<lower=1> J;           // Number of judges
  int<lower=1> nT;          // Number of years (nT instead of T because R reads T as TRUE)
  int<lower=1> JT;          // Number of judge year pairs
  array[N] int<lower=1, upper=J> jj; // Judge ID
  array[N] int<lower=1, upper=nT> tt; // Year ID
  array[N] int<lower=1, upper=JT> jt; // Judge Year ID
  array[N] int<lower=0, upper=1> y;  // Outcome, sentence (0/1)
}

parameters {
  real mu_alpha;           // Mean across judges (only judges get the mu!!)
  real<lower=0> sigma_alpha; // Standard deviation across judges
  real<lower=0> sigma_gamma; // Standard deviation across years
  real<lower=0> sigma_delta; // Standard deviation across judge years

  vector[J] alpha;         // Judge-level intercepts
  vector[nT] gamma;        // Year-level intercepts
  vector[JT] delta;        // Judge-Year-level intercepts
}

model {
  // Hyperpriors
  mu_alpha ~ student_t(3, 0, 1); // only judges get mu
  sigma_alpha ~ student_t(3, 0, 1) T[0.01, ]; // our sigmas were creating errors
  sigma_gamma ~ student_t(3, 0, 1) T[0.01, ]; // 'Scale parameter is 0, but must be positive!'
  sigma_delta ~ student_t(3, 0, 1) T[0.01, ]; // so we add a tiny lower bound to stop that

  // Priors (for the actual intercepts)
  alpha ~ normal(mu_alpha, sigma_alpha); // distributed around its hyperparameters
  gamma ~ normal(0, sigma_gamma);        // distributed around 0 and its st dev
  delta ~ normal(0, sigma_delta);         // distributed around 0 and its st dev

  // Likelihood
  for (n in 1:N)
    y[n] ~ bernoulli_logit(alpha[jj[n]] + gamma[tt[n]] + delta[jt[n]]);
}
"

```

Nested or Non-Nested?

This is a NON NESTED model, because there is no strict hierarchy across levels. There are many-to-many relationships throughout the base variables we're looking at here (sentence, judge, year), meaning that there are many ways of arranging the data into groups.

While it's true that each sentence only appears in each year once, and only in one year, and each sentence is given only once by one judge, we must note that judges give sentences across many different years, breaking a strict hierarchy and creating CROSSING RELATIONSHIPS.

Thus, our model is non-nested.