

# Bayesian Approaches to Statistics and Modeling Case Study • Part II

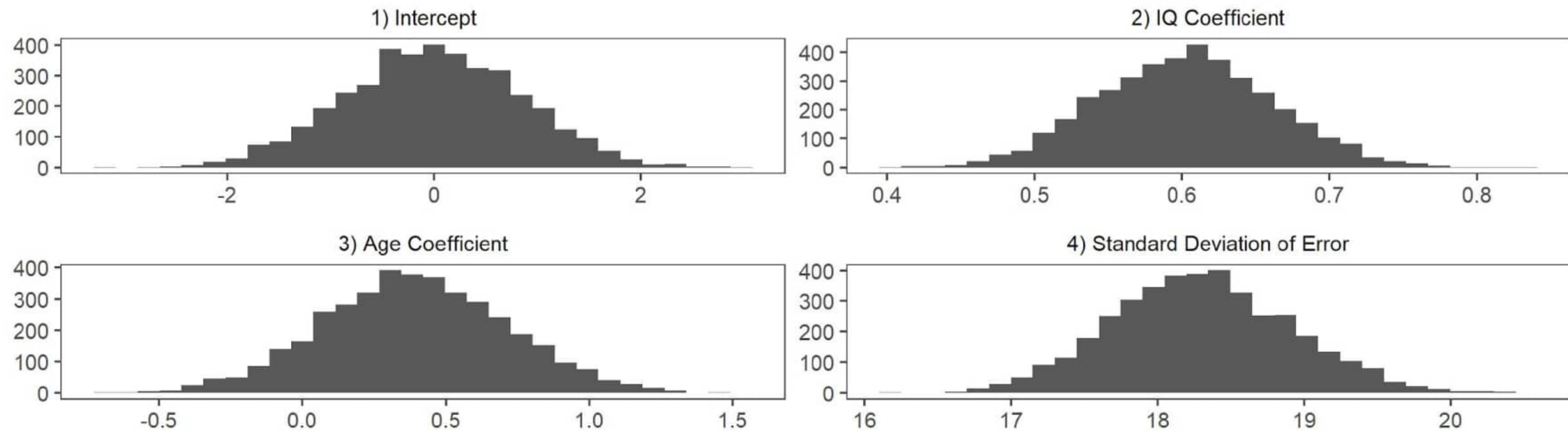
*Mark Kurzeja*

# Case Layout

- We are going to walk through a multi-level regression problem using a Bayesian framework
- Bayesian frameworks are flexible enough to do variable selection, regularization, modeling of dependence, fit models where the number of parameters exceed the number of observations, and model dependence *all within the same model*
- We will barely scratch the surface of these methods in this case, but we will explore the Bayesian workflow for modeling

# The Model – The Posteriors

Posterior Distributions of Parameters



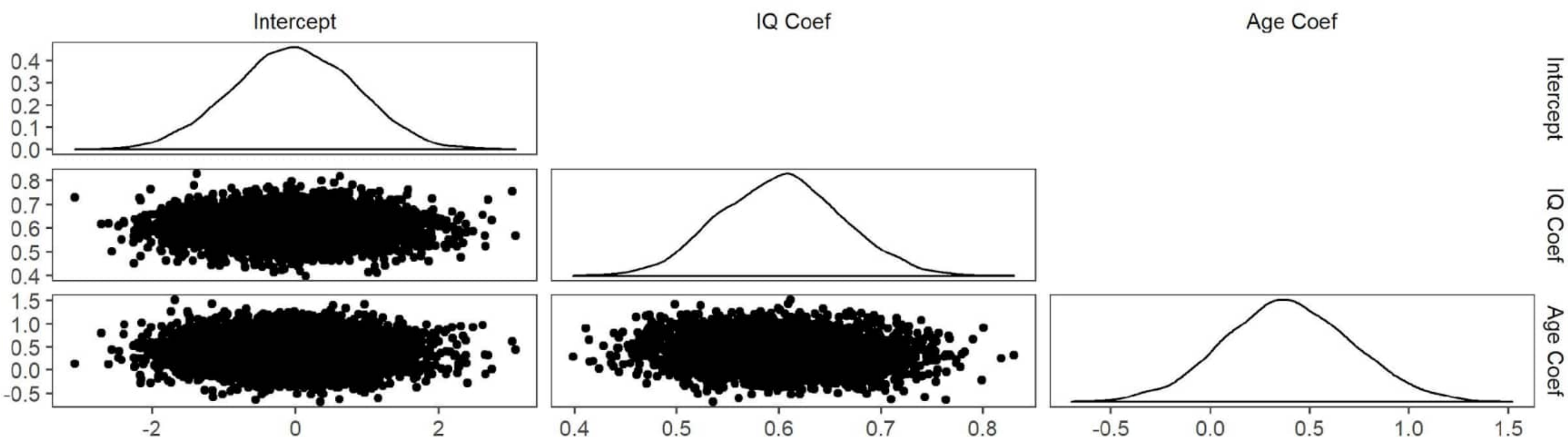
Belief about the World

Collect Data

Bayesian Update

# The Model – The Posteriors

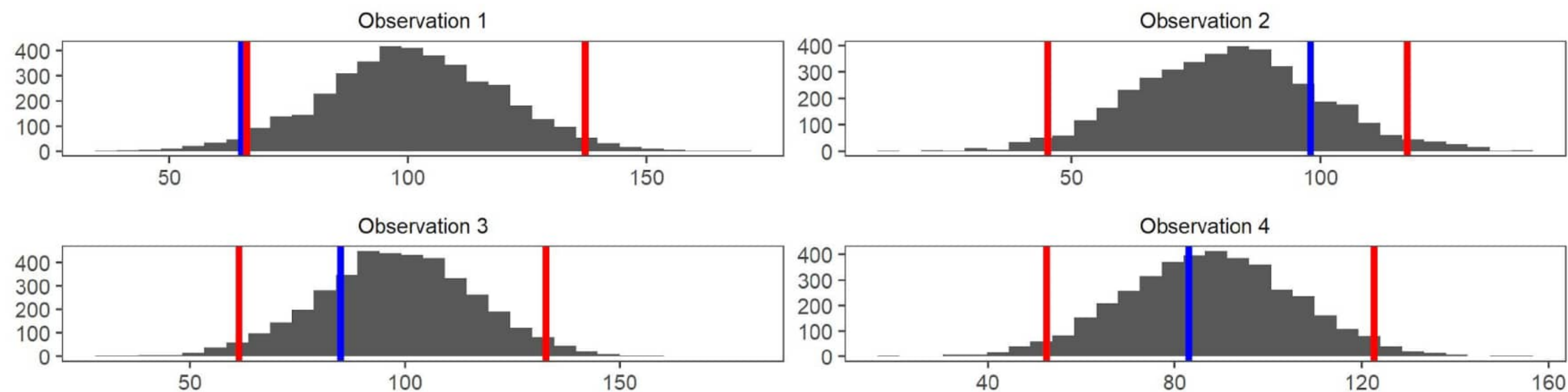
Joint Distributions of Parameters



# The Model – The Posteriors

## Posterior Predictive Intervals for First Four Observations

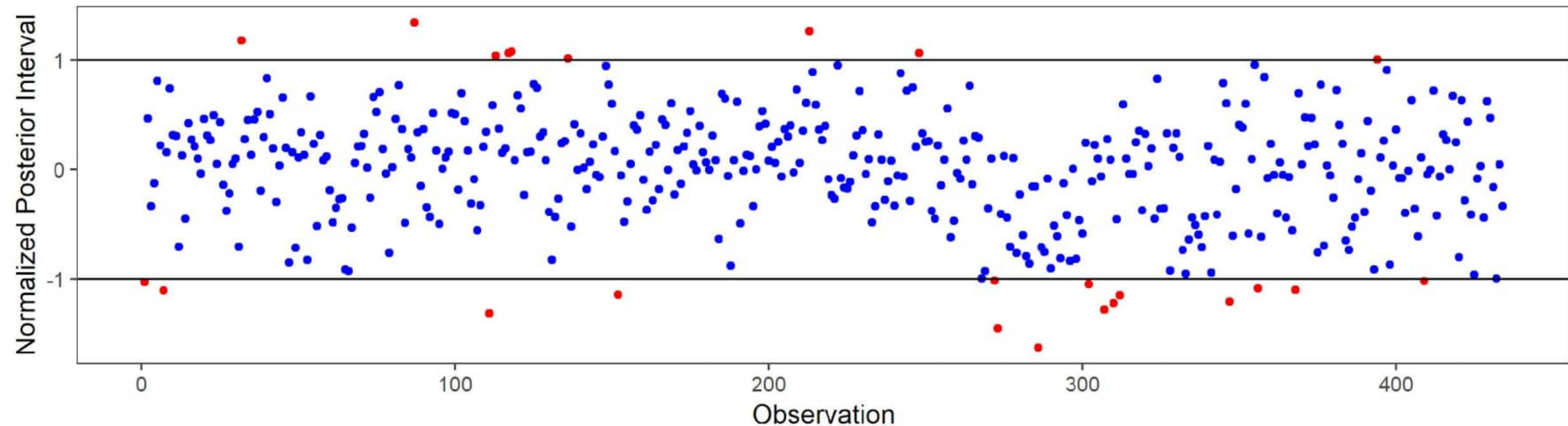
The blue line is the observed value for Childs IQ. The red lines are the 95% predictive interval



# The Model – The Posteriors

Normalized Posterior Predictive Intervals

If a dot is in  $[-1, 1]$  then the posterior predictive interval contained the true child's IQ



Belief about the World

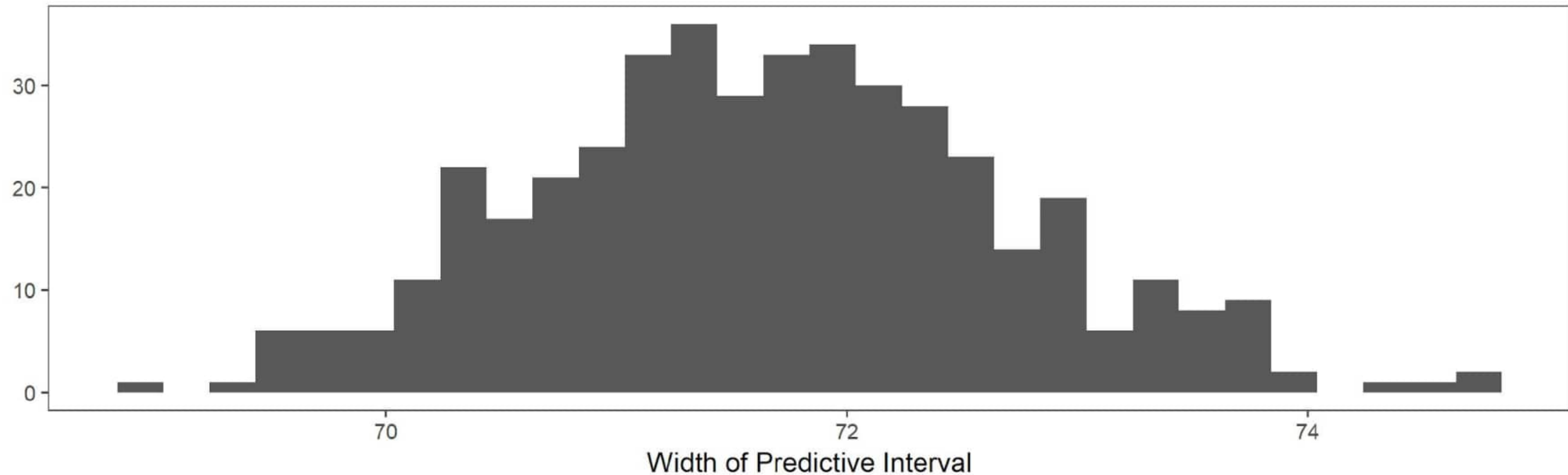
Collect Data

Bayesian Update



# The Model – The Posteriors

Histogram of Width of Predictive Intervals



Belief about the World

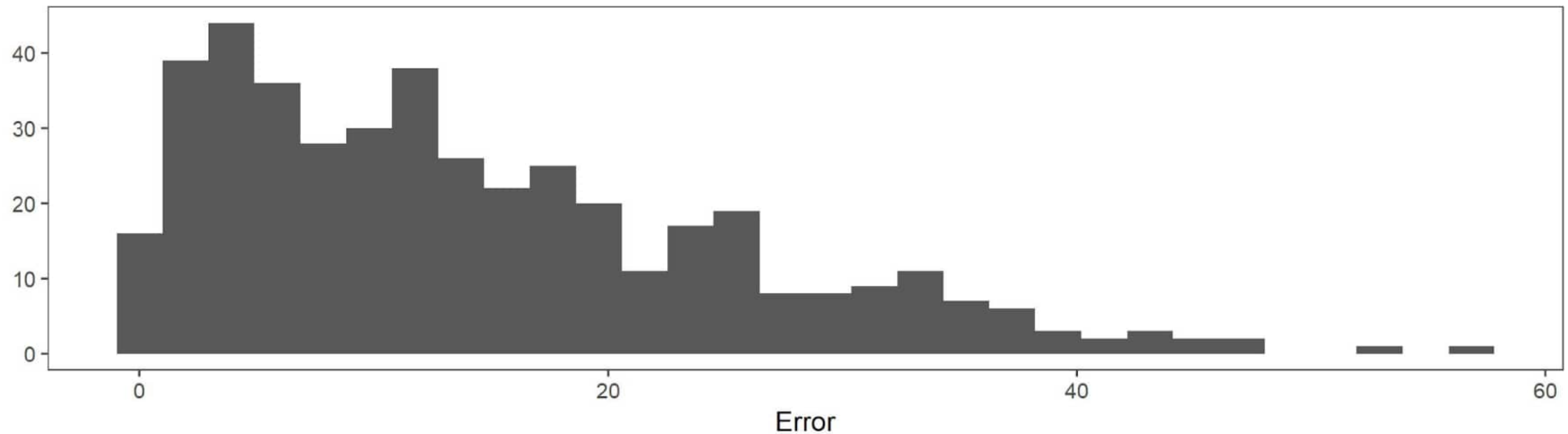
Collect Data

Bayesian Update

# The Model – The Posteriors

Histogram of Errors - Normal Model

Computed as Mean Prediction - Actual IQ



Belief about the World

Collect Data

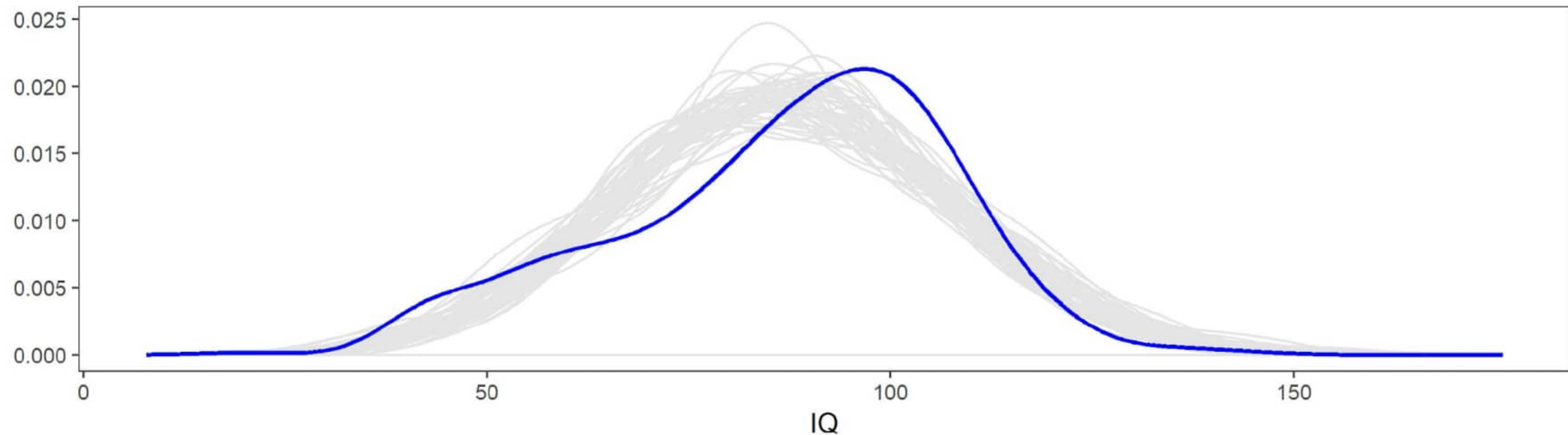
Bayesian Update



# The Model – The Posteriors

## Posterior Predictive Check

The true density, of Child IQs, is in blue. The generated posteriors are in grey



Belief about the World

Collect Data

Bayesian Update

# The Model - Observations

- We have not accounted for the high school education variable
- We have a reasonable number of observations that fall outside of their 95% predictive intervals ( $\approx 5\%$ )
- The width of the predictive intervals is pretty large. We should see if we can do better
- The estimates systematically overestimate IQ's of around 65-80 and underestimate those around 90-115

Belief about the World

Collect Data

Bayesian Update

# The Model - Observations

- Next time, we are going to explore updates that we can make to this model
- We will look to incorporating the high school status of the mother and working with a hierarchal structure will improve our model
- We will need to come up with some way of addressing the skew of the data

Belief about the World

Collect Data

Bayesian Update