Chapter 12

WAIC balances how well a model predicts, with flexibility of the model

- How well does a model predict the data?

$$\sum_{j=1}^{n} \log(p(y_j|y_1,\ldots,y_n)) = \sum_{j=1}^{n} \log(\mathbb{E}[p(y_j|\alpha,\beta,\sigma,y_1,\ldots,y_n)])$$

$$= \sum_{j=1}^{n} \log\left(\int p(y_j|\tilde{\alpha},\tilde{\beta},\tilde{\sigma})\cdot \right.$$

$$\left. \cdot p(\tilde{\alpha},\tilde{\beta},\tilde{\sigma}|y_1,\ldots,y_n)\, d\tilde{\alpha}\, d\tilde{\beta}\, d\tilde{\sigma}\right)$$

- How flexible is the model?

$$p_{WAIC} = \sum_{j=1}^{n} \mathbb{V}[\log(p(y_j|\alpha,\beta,\sigma,y_1,\ldots,y_n))]$$

WAIF in STAN

How to get WAIF in STAN, use generate quantities to get log likelihood

```
data{
...
}
parameters{
...
}
model{
...

}
generated quantities {
        vector[N_obs] log_lik;
        for (i in 1 : N_obs)
                log_lik[i] = binomial_logit_lpmf(y[i] | n[i],
                 alpha + beta_P * P[i] + beta_V * V[i] + beta_A * A[i] + beta_PA * PA[i]);
}
```

WAIF in STAN

Then in R:

```
colVars <- function(a) {
#helpful function exists in resample I belive
n <- dim(a)[[1]]
c <- dim(a)[[2]]
return(.colMeans(((a - matrix(.colMeans(a, n, c), nrow = n, ncol = c, byrow = TRUE))^2),
n, c) * n/(n - 1))

}
list_param1 <- extract(stan_model1)
log_lik      <- list_param1.1$log_lik #collect the log lik created in generated quantities
lpd           <- log_sum_exp(log_lik) # prediction part
p_waic       <- sum(colVars(log_lik)) # flexiablity part
WAIC         <- -2 * lpd + 2 * p_waic.
```

Generate quantities can be used to much more:

```
data{
        ...
}
parameters{
        ...
}
model{
        ...
}
generated quantities {
        vector[N_obs] predict_y;
        for (i in 1 : N_obs)
        predict_y[i] = exp( alpha + beta * x[i]);
}
```

Multilevel model

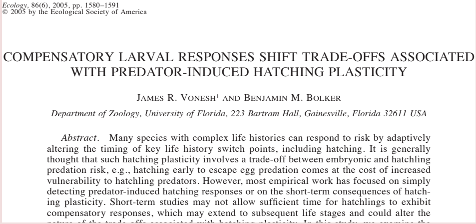
Several equivalent names:

- Multilevel models
- Hierarchical models
- Random effect models

Frog data

Data survival of hatching of reed frogs.

Figure: African reed frog, Hyperolius spinigularis


Ecology, 86(6), 2005, pp. 1580–1591

COMPENSATORY LARVAL RESPONSES SHIFT TRADE-OFFS ASSOCIATED WITH PREDATOR-INDUCED HATCHING PLASTICITY

JAMES R. VONESH[1] AND BENJAMIN M. BOLKER

Department of Zoology, University of Florida, 223 Bartram Hall, Gainesville, Florida 32611 USA

Abstract. Many species with complex life histories can respond to risk by adaptively altering the timing of key life history switch points, including hatching. It is generally thought that such hatching plasticity involves a trade-off between embryonic and hatchling predation risk, e.g., hatching early to escape egg predation comes at the cost of increased vulnerability to hatchling predators. However, most empirical work has focused on simply detecting predator-induced hatching responses or on the short-term consequences of hatching plasticity. Short-term studies may not allow sufficient time for hatchlings to exhibit compensatory responses, which may extend to subsequent life stages and could alter the

Figure: Article

THEY ARE ALL THE SAME!

Frogs hatched in different tanks (using same probability):

$$\begin{aligned} y_i &\sim Bin(n_i, p_i), \\ g(p_i) &= \alpha, \\ \alpha &\sim N(0, 10). \end{aligned}$$

Hierarchicals model DAG

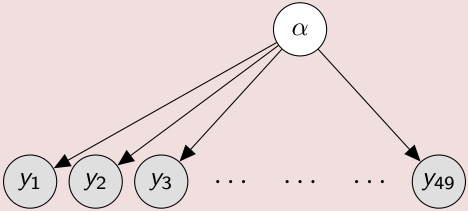


Figure: A DAG (directed acyclic graph) describing the model.

STAN simple

```
data{
        int<lower=0> N;
        int<lower=0> y[N];
        int<lower=1> n[N];
}
parameters{
        real alpha;
}
model{
        alpha ~ normal(0, 10);
        y ~ binomial_logit(n, alpha);
}
generated quantities {
        real predict_p;
        predict_p = exp( alpha)/(1 + exp(alpha));
}
```

THEY ARE ALL UNIQUE

Frogs hatched in different tanks:

$$\begin{aligned} y_i &\sim Bin(n_i, p_i), \\ g(p_i) &= \alpha_i, \\ \alpha_i &\sim N(0, 10). \end{aligned}$$

STAN simple

```
data{
        int<lower=0> N;
        int<lower=0> y[N];
        int<lower=1> n[N];
        int<lower=1> ntank;
        int<lower=0> tank[N];
}
parameters{
        real a[ntank];
}
model{
        vector[N] mu;
        for(i in 1:N)
                mu[i] = a[tank[i]];

        a ~ normal(0, 10);
        y ~ binomial_logit(n, mu);
}
generated quantities {
        real predict_p[ntank];
        for(i in 1:ntank)
                predict_p[i] = exp( a[i])/(1 + exp(a[i]));
}
```

DAG

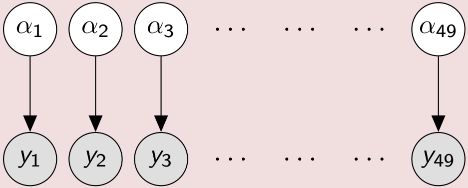


Figure: A DAG (directed acyclic graph) describing the model.

Simple binomial model

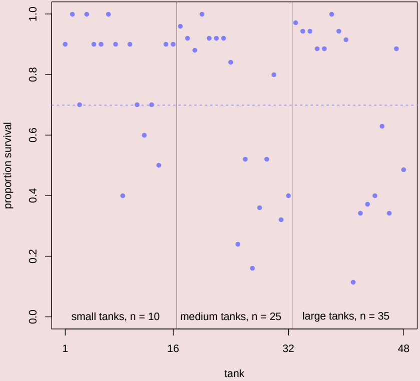


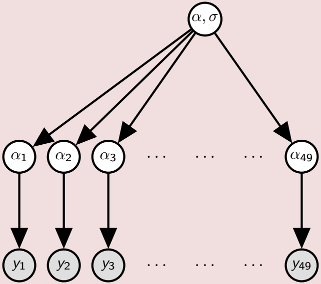
Figure: Posterior mean for each tank

Simple binomial model

Multilevel model:

$$\begin{aligned} y_i &\sim Bin(n_i, p_i), \\ g(p_i) &= \alpha_i, \\ \alpha_i &\sim N(\alpha, \sigma), \\ \alpha &\sim N(0, 10), \\ \sigma &\sim HC(0, 1). \end{aligned}$$

Hierarchicals model DAG



STAN multilevel

```
data{
        int<lower=0> N;
        int<lower=0> y[N];
        int<lower=1> n[N];
        int<lower=1> ntank;
        int<lower=0> tank[N];
}
parameters{
        real a[ntank];
        real alpha;
        real<lower=0> sigma;
}
model{
        vector[N] mu;
        for(i in 1:N)
        mu[i] = a[tank[i]];

        sigma ~ cauchy(0,1);
        alpha ~ normal(0, 20);
        a     ~ normal(alpha, sigma);
        y ~ binomial_logit(n, mu);
}
generated quantities {
        real predict_p[ntank];
        for(i in 1:ntank)
        predict_p[i] = exp( a[i])/(1 + exp(a[i]));
}
```

Multilevel binomial model

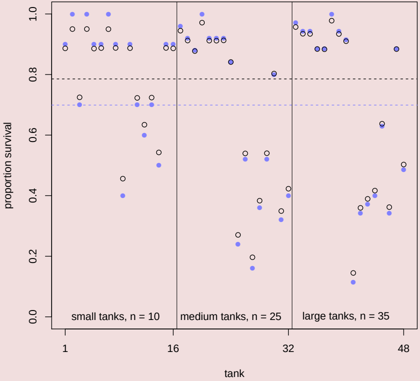


Figure: Posterior mean for each tank

Lets examine, $y_2 = 10$, $n_2 = 10$

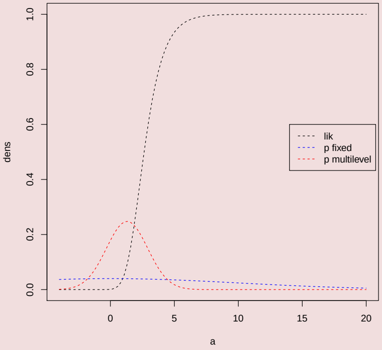


Figure: Prior + likelihood for a.

Lets examine, $y_2 = 10, n_2 = 10$

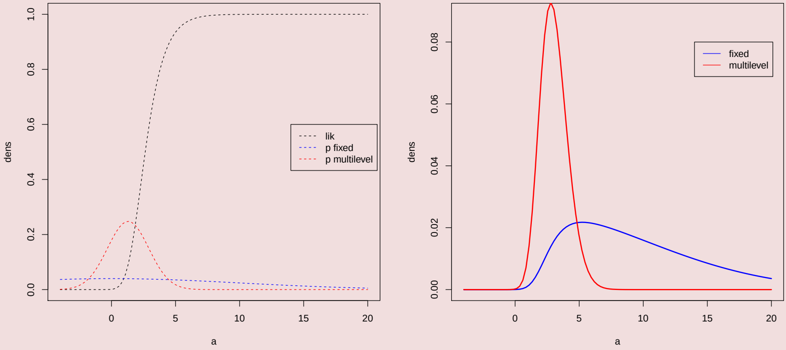
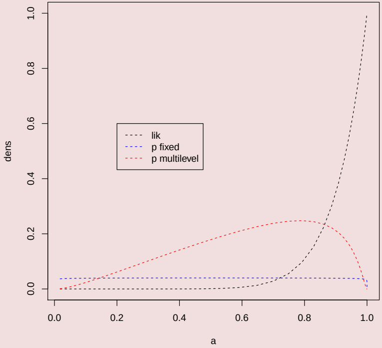


Figure: Posterior for a.

Lets examine, $y_2 = 10, n_2 = 10$



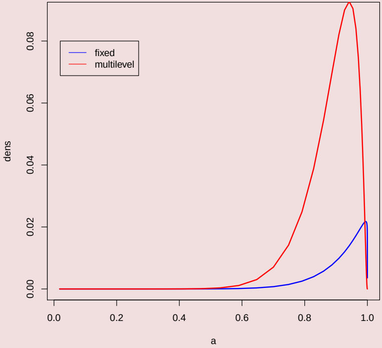
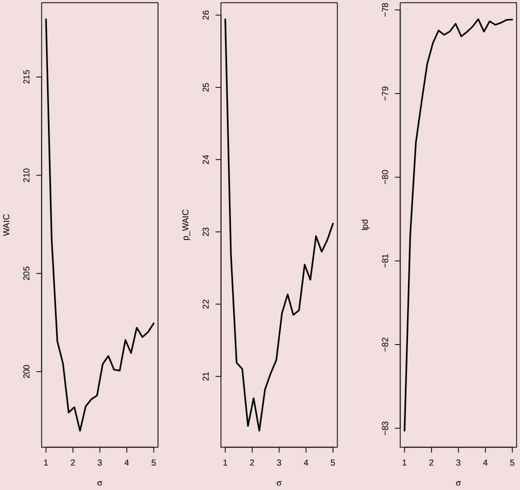
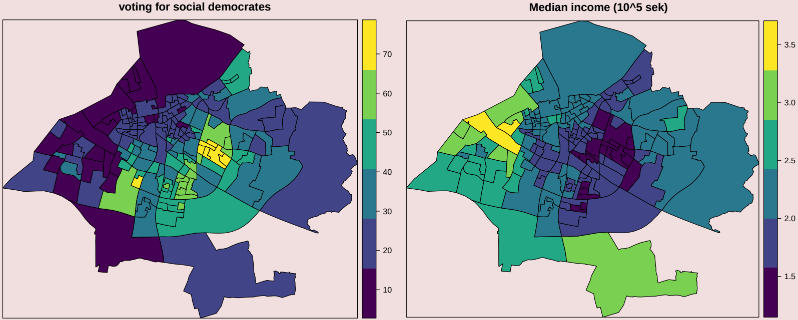


Figure: Figure on the left prior + likelihood for p. Figure on the right posterior p.

WAIC for varying σ



Voting in Malmö, data



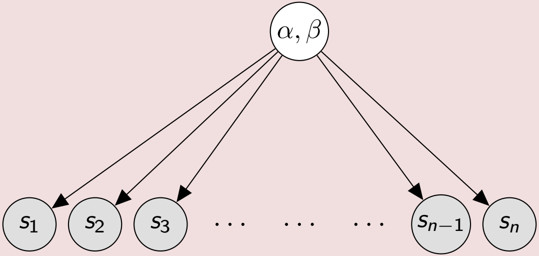
THEY ARE ALL THE UNIQUE!

$$
\begin{aligned}
s_i &\sim bin(n_i, p_i), \\
g(p_i) &= \alpha_i + med_i \beta_i, \\
\alpha_i &\sim N(0, 10) \\
\beta_i &\sim N(0, 10)
\end{aligned}
$$

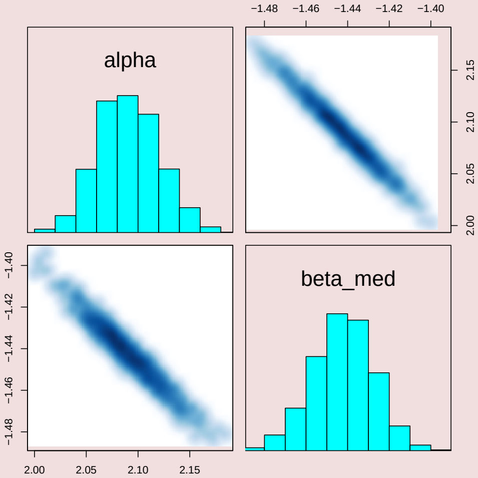
THEY ARE ALL THE SAME!

$$
\begin{aligned}
s_i &\sim bin(n_i, p_i),\\
g(p_i) &= \alpha + med_i \beta,\\
\alpha &\sim N(0, 10)\\
\beta &\sim N(0, 10)
\end{aligned}
$$

Independent model DAG



Posterior parameter



Predictions

- By the model the prediction given the data is

$$\hat{Y}_i \sim Bin(n_i, p_i),$$
$$p_i \sim p(\cdot|y_1, y_2, \ldots, y_n)$$

Predictions

- By the model the prediction given the data is

$$\hat{Y}_i \sim Bin(n_i, p_i),$$
$$p_i \sim p(\cdot|y_1, y_2, \ldots, y_n)$$

- The variance is:

$$V[\hat{Y}|p_i, n_i] = n_i(1 - p_i)p_i$$
$$V[\frac{\hat{Y}}{n_i}|p_i, n_i] = \frac{(1 - p_i)p_i}{n_i}$$

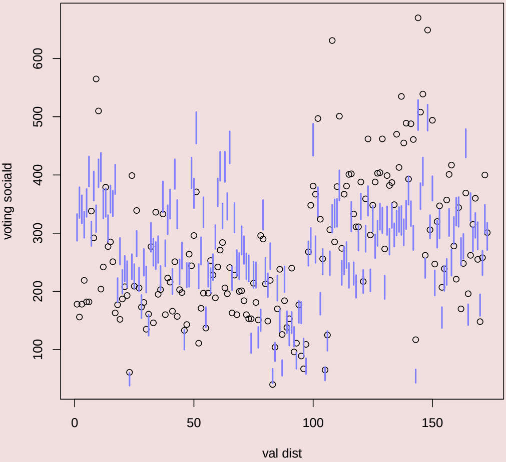


Figure: Predicition by district

Overdispertion

- Both binomial and Poisson has only one parameter.
- These models are extremely sensitivity to incorrect parameter.
- They can not adjust it variance to the data.

Solution

- This is typically solved by overdispersion model. Like Beta-binomial.

Solution

- This is typically solved by overdispersion model. Like Beta-binomial.
- For each observation one adds a random non-negative parameter:

$$p(y_i|n_i) = \int Bin(y_i|n_i, p_i)h(p_i|p, \theta)p(p, \theta)dp_i dp d\theta,$$

  Then one puts covariates on p not p_i.

Solution

- This is typically solved by overdispersion model. Like Beta-binomial.
- For each observation one adds a random non-negative parameter:

$$p(y_i|n_i) = \int Bin(y_i|n_i, p_i)h(p_i|p, \theta)p(p, \theta)dp_i dp d\theta,$$

  Then one puts covariates on p not p_i.
- overdisperation is typically a multilevel model.

multilevel

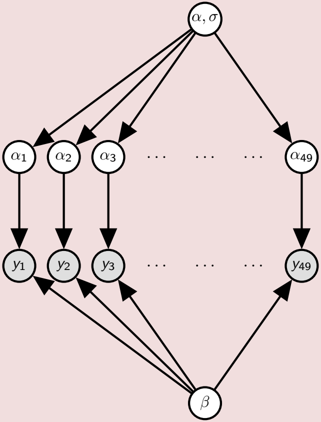
$$
\begin{aligned}
y_i &\sim Bin(n_i, p_i) \\
g(p_i) &\sim \alpha_0 + med_i \beta + Z_i \\
Z_i &\sim N(0, \sigma) \\
\alpha_0 &\sim N(0, 10) \\
\sigma &\sim HC(0, 5).
\end{aligned}
$$

$$
\begin{aligned}
y_i &\sim Bin(n_i, p_i) \\
g(p_i) &\sim \alpha_0 + med_i\beta + Z_i \\
Z_i &\sim N(0, \sigma) \\
\alpha_0 &\sim N(0, 10) \\
\sigma &\sim HC(0, 5).
\end{aligned}
$$

or equivalently

$$
\begin{aligned}
y_i &\sim Bin(n_i, p_i) \\
g(p_i) &\sim \alpha_i + med_i\beta \\
\alpha_i &\sim N(\alpha_0, \sigma) \\
\alpha_0 &\sim N(0, 10) \\
\sigma &\sim HC(0, 5).
\end{aligned}
$$

Hierarchicals model DAG



```
data {
        int<lower = 1> N;
        int<lower = 1> voters[N];
        int<lower = 0> soc[N];
        vector[N] med;
}

parameters{
        real alphas[N];
        real alpha0;
        real<lower=0> sigma;
        real beta_med;
}

model{
        vector[N] mu;
        sigma     ~ cauchy(0, 1);
        alpha0    ~ normal(0, 10);
        beta_med  ~ normal(0, 10);
        alphas    ~ normal(alpha0, sigma);
        for(i in 1:N)
                mu[i] = alphas[i] + med[i] * beta_med;
        soc ~ binomial_logit(voters, mu);
}
```

PI for model

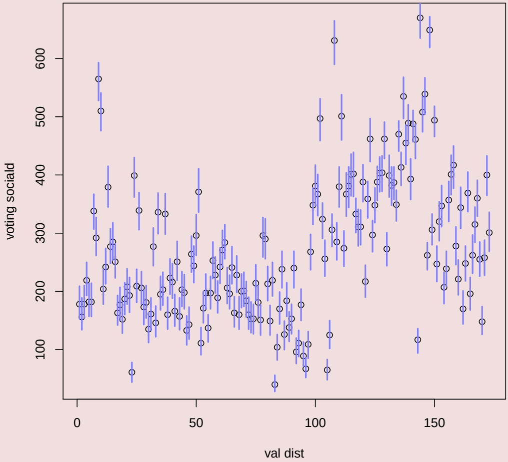


Figure: Prediction by district multilevel

PI for multilevel

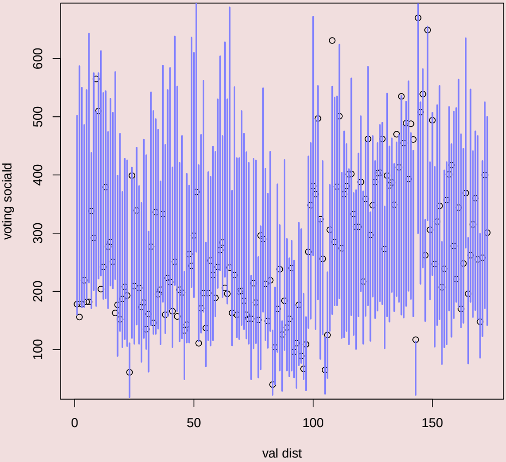


Figure: Prediction unconditional by district multilevel

- The variance is:

$$V[\hat{Y}|n_i] \approx n_i(1-\hat{p}_i)\hat{p}_i + n_i^2\tilde{\sigma}$$
$$V[\frac{\hat{Y}}{n_i}|n_i] \approx \frac{(1-\hat{p}_i)\hat{p}_i}{n_i} + \tilde{\sigma}$$

Where $\tilde{\sigma}$ is the variation from

PI for multilevel without cheating

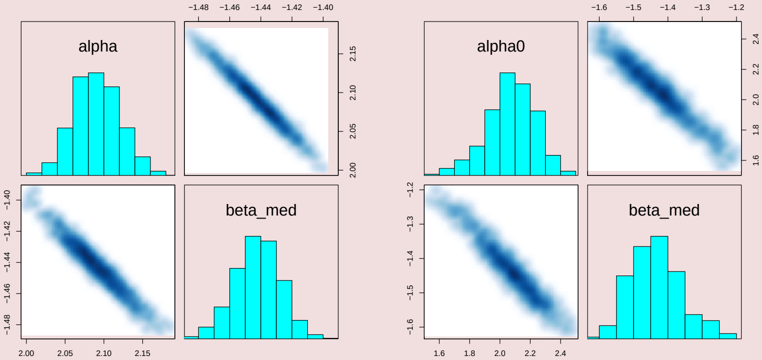


Figure: Look at parameter certainty

Multilevel binomial model

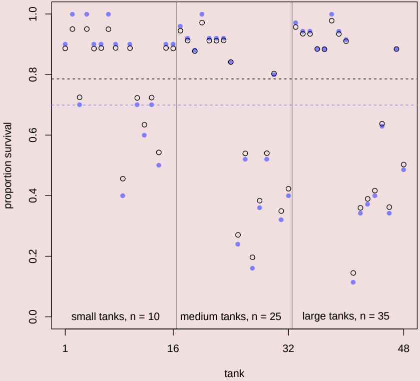


Figure: Posterior mean for each tank

Histogram of the α

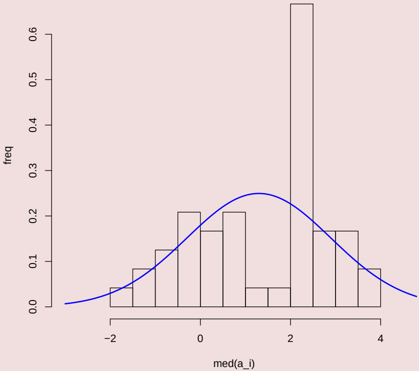


Figure: Histogram of the α_i

Multilevel mixture model:

$$
\begin{aligned}
y_i &\sim Bin(n_i, p_i), \\
g(p_i) &= \alpha_i, \\
\alpha_i &\sim \theta N(\mu_1, \sigma_1) + (1-\theta) N(\mu_2, \sigma_2), \\
\mu_1 &\sim N(0, 10), \\
\mu_2 &\sim N(0, 10), \\
\sigma_1 &\sim HC(0, 1). \\
\sigma_2 &\sim HC(0, 1). \\
\theta &\sim B(2, 2).
\end{aligned}
$$

Stan model

```
parameters{
  ordered[2] a0;
  real a[ntank];
  real<lower=0> sigma_0;
  real<lower=0> sigma_1;
  real<lower=0, upper=1> theta;
}
model{
  vector[N] mu;
  theta   ~ beta(2, 2);
  a0      ~ normal(0,10);
  sigma_0 ~ cauchy(0,1);
  sigma_1 ~ cauchy(0,1);
  for(i in 1:ntank){
   target += log_sum_exp(
     bernoulli_lpmf(1|theta) + normal_lpdf(a[i]| a0[1],sigma_0),
     bernoulli_lpmf(0|theta) + normal_lpdf(a[i]| a0[2],sigma_1)
     );
  }
  for(i in 1:N)
    mu[i] = a[tank[i]];

  y ~ binomial_logit(n, mu);
}
```

Multilevel mixture binomial model

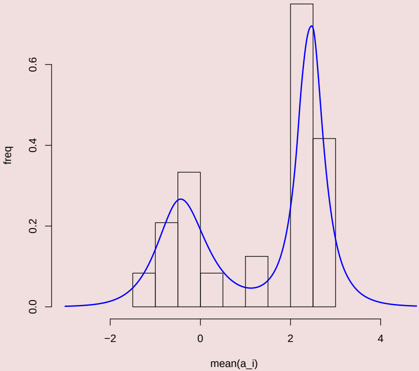


Figure: Histogram of the α_i

Multilevel mixture binomial model

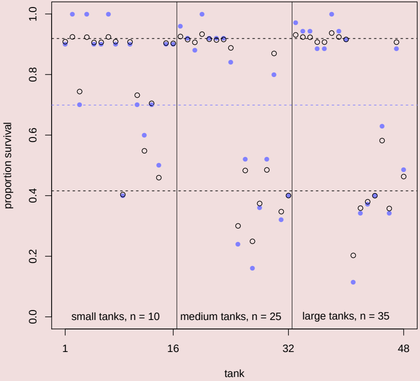


Figure: Posterior mean for each tank