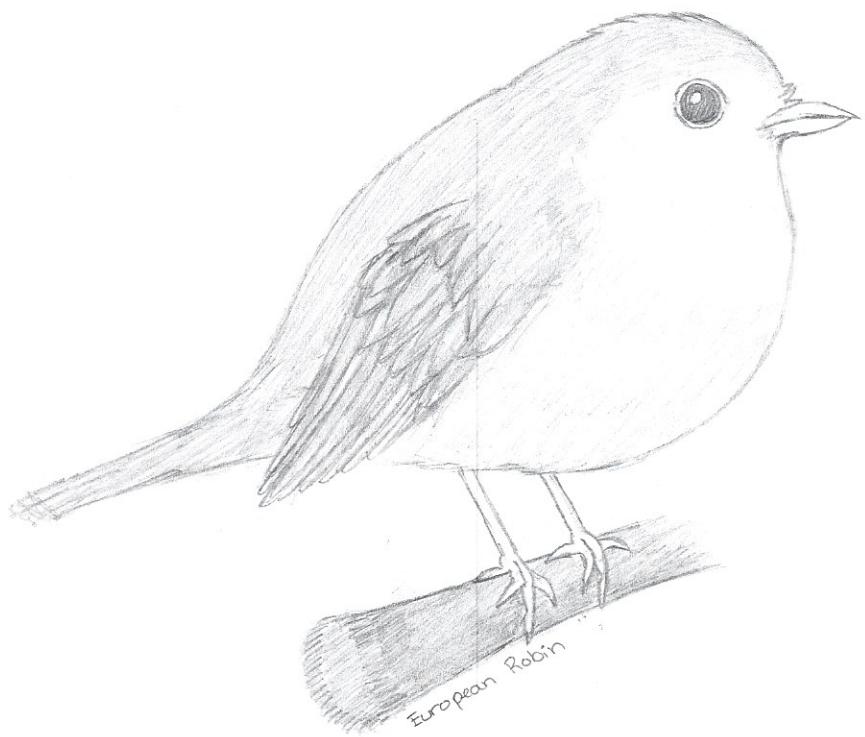


AMS 274

Homework #4



Diana Geraudo



20/20
A.K.

AMS 274 Homework 4

Diana Gerardo

December 2, 2018

- (10)
1. Consider a binary response ($y = 0, 1$) that corresponds to the choice between two options (say, two product brands), a choice that can be explained by a covariate x . Let U_0 denote the utility of choice $y = 0$, and U_1 the utility of choice $y = 1$. For $y = 0, 1$, suppose that $U_y = a_y + b_y x + \varepsilon_y$, using a scale such that ε_y follows a distribution in standard form. (Assume that the coefficients a_0, a_1, b_0 and b_1 are all known.) A subject selects $y = 1$ if $U_1 > U_0$ for that subject.
 - (a) If ε_0 and ε_1 are independent $N(0, 1)$ random variables, show that $\Pr(Y = 1)$ satisfies the probit regression model structure, and write the regression coefficients in terms of a_0, a_1, b_0 and b_1 .

Solution:

$$\begin{aligned}\Pr(Y = 1) &= \Pr(U_1 > U_0) \\ &= \Pr(a_1 + b_1 x + \varepsilon_1 > a_0 + b_0 x + \varepsilon_0) \\ &= \Pr(a - a_0 + (b_1 - b_0)x > \varepsilon_0 - \varepsilon_1)\end{aligned}$$

Since ε_0 and ε_1 are independent $N(\mu_y = 0, \sigma_y^2 = 1)$, then the difference would be $N(\mu_0 - \mu_1, \sigma_0^2 + \sigma_1^2)$. So let $\varepsilon_0 - \varepsilon_1 \sim N(0, 2)$ then

$$\begin{aligned}\Pr(Y = 1) &= \Pr(a - a_0 + (b_1 - b_0)x > \varepsilon_0 - \varepsilon_1) \\ &= \Pr\left(\frac{a_1 - a_0}{\sqrt{2}} + \frac{b_1 - b_0}{\sqrt{2}}x > \frac{\varepsilon_0 - \varepsilon_1}{\sqrt{2}}\right) \\ &= \Pr\left(\frac{a_1 - a_0}{\sqrt{2}} + \frac{b_1 - b_0}{\sqrt{2}}x > \varepsilon\right) \\ &= \Phi(\beta_0 + \beta_1 x)\end{aligned}$$

where $\beta_0 = \frac{a_1 - a_0}{\sqrt{2}}$, $\beta_1 = \frac{b_1 - b_0}{\sqrt{2}}$, and $\varepsilon = \frac{\varepsilon_0 - \varepsilon_1}{\sqrt{2}} \sim N(0, 1)$. Thus, $\Pr(Y = 1)$ satisfies the probit regression model structure.

- (b) If ε_0 and ε_1 are independent random variables with c.d.f. $F(\varepsilon) = \exp\{-\exp(-\varepsilon)\}$, show that $\Pr(Y = 1)$ satisfies the logistic regression model structure (again, express the regression coefficients in terms of a_0, a_1, b_0 and b_1).

Solution:

Note $F(\varepsilon) = \exp\{-\exp(-\varepsilon)\}$ is the cdf of a standard Gumbel distribution, $Gum(\mu = 0, \beta = 1)$.

$$\begin{aligned}\Pr(Y = 1) &= \Pr(U_1 > U_0) \\ &= \Pr(a_1 + b_1 x + \varepsilon_1 > a_0 + b_0 x + \varepsilon_0) \\ &= \Pr(a - a_0 + (b_1 - b_0)x > \varepsilon_0 - \varepsilon_1)\end{aligned}$$

We want to show that the difference of two Gumbel distributions follows a logistic distribution and we can do this by using the Gumbel mgf.

$$\begin{aligned}E(e^{t(\varepsilon_0 - \varepsilon_1)}) &= E(e^{t\varepsilon_0}) E(e^{-t\varepsilon_1}), \text{ since } \varepsilon_0, \varepsilon_1 \text{ are independent} \\ &= \Gamma(1-t) \cdot \Gamma(1+t) \\ &= \frac{\Gamma(1-t) \cdot \Gamma(1+t)}{\Gamma(2)} \\ &= \frac{\Gamma(1-t) \cdot \Gamma(1+t)}{\Gamma(1-t+1+t)} \\ &= B(1-t, 1+t)\end{aligned}$$

Where $B(1-t, 1+t)$ is the mgf of the Logistic(0,1). Let $\varepsilon = \varepsilon_0 - \varepsilon_1 \sim Logistic(0, 1)$, then

$$\begin{aligned}\Pr(Y = 1) &= \Pr(a - a_0 + (b_1 - b_0)x > \varepsilon) \\ &= F(\beta_0 + \beta_1 x)\end{aligned}$$

where $\beta_0 = a - a_0$ and $\beta_1 = b_1 - b_0$. Thus $\Pr(Y = 1)$ satisfies the logistic regression model structure.

2. Consider the “alligator food choice” data example, the full version of which is discussed in Section 7.1 of Agresti (2002), *Categorical Data Analysis*, Second Edition. Here, consider the subset of the data reported in Table 7.16 (page 304) of the above book. This data set involves observations on the primary food choice for $n = 63$ alligators caught in Lake George, Florida. The nominal response variable is the primary food type (in volume) found in each alligator’s stomach, with three categories: “fish”, “invertebrate”, and “other”. The invertebrates were mainly apple snails, aquatic insects, and crayfish. The “other” category included amphibian, mammal, bird, reptile, and plant material. Also available for each alligator is covariate information on its length (in meters) and gender.
 - (a) Focus first on length as the single covariate to explain the response probabilities for the “fish”, “invertebrate” and “other” food choice categories. Develop a Bayesian multinomial regression model, using the baseline-category logits formulation with “fish” as the baseline category, to estimate (with point and interval estimates) the response probabilities as a function of length. (Note that in this data example, we have $m_i = 1$, for $i = 1, \dots, n$.) Discuss your prior choice and approach to MCMC posterior simulation.

Solution:

We use “Fish” as the baseline category and then we have two logistic regression

$$\begin{aligned}\log \left(\frac{\pi_I}{\pi_F} \right) &= \alpha_1 + \beta_1 x_i \\ \log \left(\frac{\pi_O}{\pi_F} \right) &= \alpha_2 + \beta_2 x_i\end{aligned}$$

where π_I is the probability of choosing invertebrates, π_O is the probability of choosing other food, π_F is the probability of choosing fish, and x is the covariate length (in meters). Here, we would like to estimate the response curves

$$\begin{aligned}\pi_1(x_i) = \pi_I(x_i) &= \frac{\exp(\alpha_1 + \beta_1 x_i)}{1 + \sum_{l=1}^2 \exp(\alpha_l + \beta_l x_i)} \\ \pi_2(x_i) = \pi_O(x_i) &= \frac{\exp(\alpha_2 + \beta_2 x_i)}{1 + \sum_{l=1}^2 \exp(\alpha_l + \beta_l x_i)} \\ \pi_3(x_i) = \pi_F(x_i) &= \frac{1}{1 + \sum_{l=1}^2 \exp(\alpha_l + \beta_l x_i)}\end{aligned}$$

Using the *multinom()* function in R, we obtain the mle estimates of $\alpha_1, \beta_1, \alpha_2, \beta_2$ as show in Table 1.

	α_1	β_1	α_2	β_2
MLE	4.1815515	-2.472596	-0.9951974	-0.086008

Table 1: MLE estimates of $\alpha_1, \beta_1, \alpha_2, \beta_2$

Now we wish to develop a Bayesian multinomial regression model and we’ll use Table 1 to compare our Bayesian parameter estimates. So, we use a flat prior $p(\alpha_1, \beta_1, \alpha_2, \beta_2) \propto 1$ since we have no prior information about how aligators choose their food, then the posterior distribution is

$$\begin{aligned}p(\alpha_1, \beta_1, \alpha_2, \beta_2 | data) &\propto p(\alpha_1, \beta_1, \alpha_2, \beta_2) \prod_{i=1}^{63} Multi(y_{iI}, y_{iO}, y_{iF} | m_i = 1, \pi_I(x_i), \pi_O(x_i), \pi_F(x_i)) \\ &= \prod_{i=1}^{63} \left\{ \left(\frac{\exp(\alpha_1 + \beta_1 x_i)}{1 + \sum_{l=1}^2 \exp(\alpha_l + \beta_l x_i)} \right)^{y_{iI}} \left(\frac{\exp(\alpha_2 + \beta_2 x_i)}{1 + \sum_{l=1}^2 \exp(\alpha_l + \beta_l x_i)} \right)^{y_{iO}} \left(\frac{1}{1 + \sum_{l=1}^2 \exp(\alpha_l + \beta_l x_i)} \right)^{y_{iF}} \right\} \\ &:= k(\alpha_1, \beta_1, \alpha_2, \beta_2)\end{aligned}$$

Where $k(\cdot)$ is the exponentiated log of the posterior. Note y_{iI} , y_{iO} , and y_{iF} is either 0 or 1. We have an MCMC algorithm that simulates the posterior samples of all four parameters shown in Algorithm 1.

Algorithm 1: Metropolis-Hastings for Multinomial GLM using baseline-category logits

```

1. initialize  $(\alpha_1^{(1)}, \beta_1^{(1)}, \alpha_2^{(1)}, \beta_2^{(1)}) = (0, 0, 0, 0)$ 
for    $i$  to 2 : 200000
    2. draw  $(\alpha_1^*, \beta_1^*, \alpha_2^*, \beta_2^*)_i \sim N_4 \left( (\alpha_1^{(t-1)}, \beta_1^{(t-1)}, \alpha_2^{(t-1)}, \beta_2^{(t-1)}), d \cdot J^{-1}(\hat{\alpha}_1, \hat{\beta}_1, \hat{\alpha}_2, \hat{\beta}_2) \right)$ 
       where  $d$  is a tuning parameter,  $J^{-1}(\hat{\alpha}_1, \hat{\beta}_1, \hat{\alpha}_2, \hat{\beta}_2)$  is the fisher information and
        $\hat{\alpha}_1, \hat{\beta}_1, \hat{\alpha}_2, \hat{\beta}_2$  are the mle's
    3. If  $U \sim Unif(0, 1) < q = \min \left\{ 1, \frac{k(\alpha_1^*, \beta_1^*, \alpha_2^*, \beta_2^*)_i}{k(\alpha_1^{(t-1)}, \beta_1^{(t-1)}, \alpha_2^{(t-1)}, \beta_2^{(t-1)})} \right\}$ . Set
        $(\alpha_1^{(t)}, \beta_1^{(t)}, \alpha_2^{(t)}, \beta_2^{(t)}) = (\alpha_1^*, \beta_1^*, \alpha_2^*, \beta_2^*)_i$ 
    4. Else  $(\alpha_1^{(t)}, \beta_1^{(t)}, \alpha_2^{(t)}, \beta_2^{(t)}) = (\alpha_1^{(t-1)}, \beta_1^{(t-1)}, \alpha_2^{(t-1)}, \beta_2^{(t-1)})$ 
end

```

When implementing Algorithm 1 we set $d = 0.25$. The total number of iterations is 200000 and we burn the first 100000. Leaving us 100000 draws left for analysis. Figure 1 shows the trace plots and posterior distributions of $\alpha_1, \beta_1, \alpha_2$, and β_2 . These trace plot exhibits rapid up-and-down variation with no long-term trends or drifts. This indicates convergence in distribution. All of the distributions seem to follow a normal distribution.

	Posterior Mean	2.5%	97.5%	MLE
α_1	4.508041	1.773915	7.905415	4.181551
β_1	-2.665339	-4.628856	-1.188540	-2.472596
α_2	-0.9661486	-3.4037910	1.4917552	-0.995197
β_2	-0.1128403	-1.2048569	0.8413531	-0.086008

Table 2: posterior means, 95% quantile, mle

Table 2 shows posterior mean and its corresponding 95% quantile, along with the mles we saw in table 1. From the table we see posterior mean $\beta_1 < 0$ with high probability, meaning that bigger alligators tend to eat fish instead of invertebrates. Although the posterior mean $\beta_2 < 0$, it is hard to say that the odds of eating other food against eating fish is increasing or decreasing as the length of alligators increases since the 95% confidence interval is $(-1.2048569, 0.8413531)$. Our posterior mean estimates are close to the mle values which further supports our bayesian mcmc algorithm.

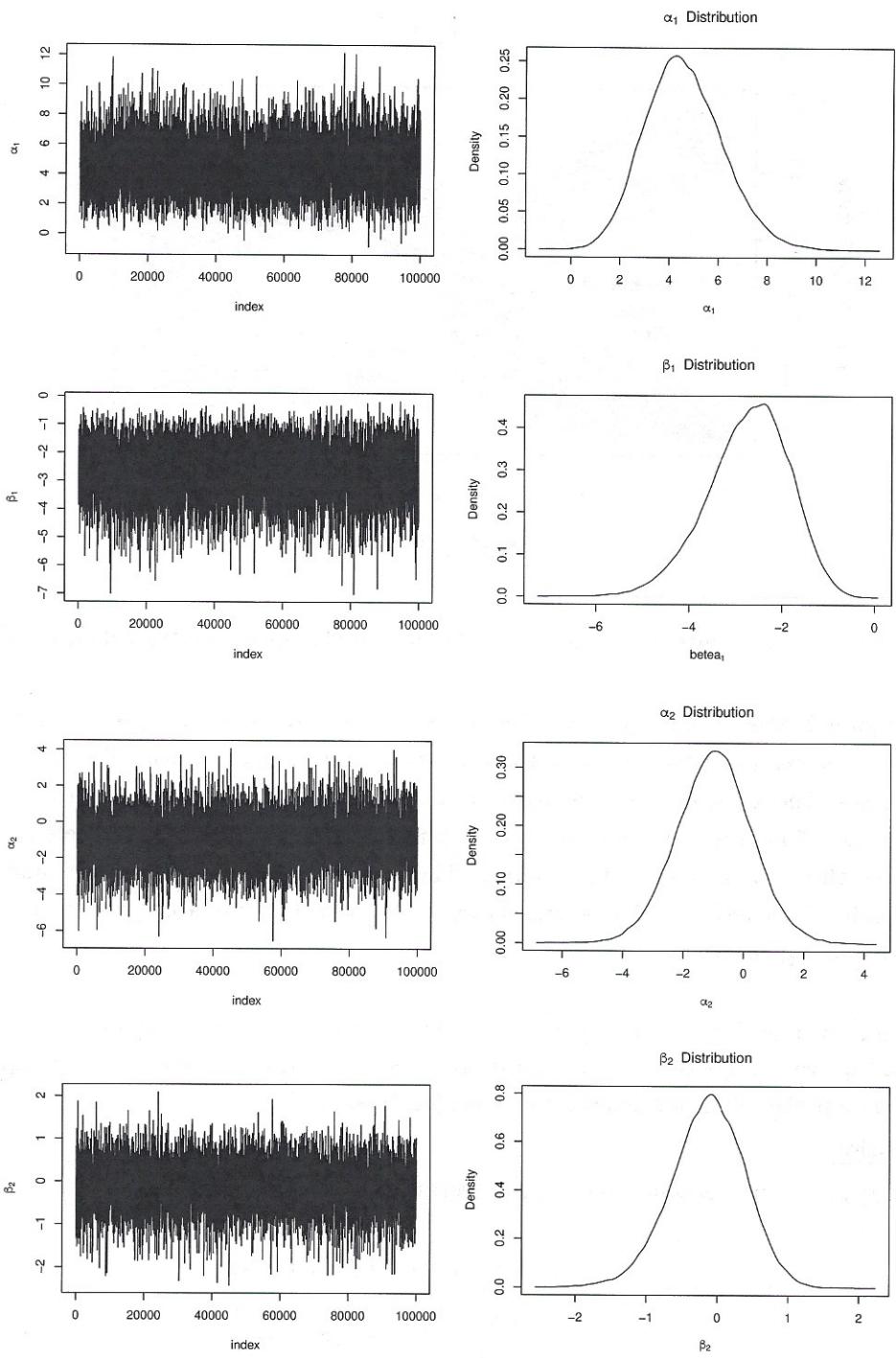


Figure 1: Trace and density plots to show healthy convergence of parameters

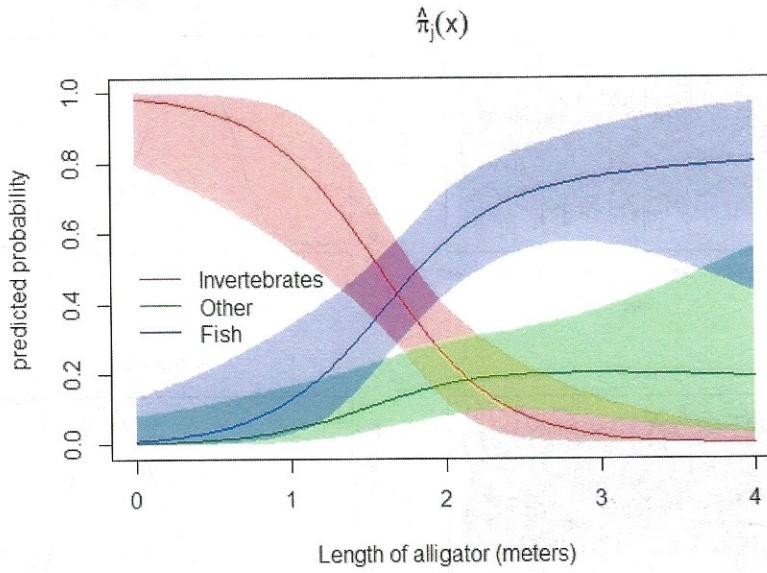


Figure 2: point & 95% interval estimates for the estimated response curves $\pi_j(x)$; $j = I, O, F$

In Figure 2, the probability of an alligator eating invertebrates decreases as the length of the alligator increases. This is because the probability of an alligator eating fish increases as length increases. The probability of an alligator eating other food slightly increases as the length increases from 1 to 2.5 meters, but seems to keep constant when the length is over 2.5 meters. The intervals for the three categories are thick/large. The width of intervals for “Other” and “Fish” keep increasing for length over 2.5 meters, showing a lot of uncertainty about the choice of “Other” or “Fish” for big alligators.

- (b) Extend the model from part (a) to describe the effects of both length and gender on food choice. Based on your proposed model, provide point and interval estimates for the length-dependent response probabilities for male and female alligators.

Solution:

Taking gender into account, our logistic regressions become

$$\begin{aligned} \log\left(\frac{\pi_I}{\pi_F}\right) &= \alpha_1 + \beta_1 x_i + \lambda_1 z_i \\ \log\left(\frac{\pi_O}{\pi_F}\right) &= \alpha_2 + \beta_2 x_i + \lambda_2 z_i \end{aligned}$$

where z_i , for $i = 1, \dots, 63$, is the gender dummy variable with $z_i = 1$ for male and $z_i = 0$ for female. And λ_1 and λ_2 are the coefficients for gender. Now we have a total of six parameters

$\alpha_1, \beta_1, \lambda_1, \alpha_2, \beta_2, \lambda_2$. Then response probabilities become

$$\begin{aligned}\pi_I(x) &= \frac{\exp(\alpha_1 + \beta_1 x + \lambda_1 z)}{1 + \sum_{l=1}^2 \exp(\alpha_l + \beta_l x + \lambda_l z)} \\ \pi_O(x) &= \frac{\exp(\alpha_2 + \beta_2 x + \lambda_2 z)}{1 + \sum_{l=1}^2 \exp(\alpha_l + \beta_l x + \lambda_l z)} \\ \pi_F(x) &= \frac{1}{1 + \sum_{l=1}^2 \exp(\alpha_l + \beta_l x + \lambda_l z)}\end{aligned}$$

Again we choose a flat prior and implement the same algorithm as in part (a) but this time including the gender coefficient parameters. We then get the following healthy convergence results shown in Figure 3; a few posterior summary statistics in Table 3; and the point and interval estimates for the response probabilities for both male and female alligators shown in Figure 4.

	Posterior Mean	2.5%	97.5%	MLE
α_1	6.229759	2.762183	10.576666	5.719262
β_1	-3.189714	-5.374722	-1.489817	-2.930679
λ_1	-1.311672	-2.812340	0.055907	-1.207351
α_2	-1.138617	-3.841813	1.519311	-1.102027
β_2	-0.1168479	-1.166605	0.861814	-0.086955
λ_2	0.174841	-1.394463	1.914359	0.158640

Table 3: posterior means and 95% quantile

From Table 3 we see that the posterior mean of $\alpha_1, \beta_1, \alpha_2$, and β_2 increased slightly in value from the posterior estimates in part (a), but still hold the same inference. Recall that λ_1, λ_2 are the gender coefficients for z_i where $z = 1$ is for male and $z = 0$ is for female. Although the posterior mean $\lambda_1 < 0$, it is hard to say that the odds of a male alligator eating invertebrates against a male eating fish is increasing or decreasing as the length of the male alligators increases. Similar inference goes for λ_2 . Using *multinom()* in R, we get the MLEs for $\alpha_1, \beta_1, \alpha_2$, and β_2 . We see that our posterior mean estimates are close to mle values, thus supporting our bayesian mcmc algorithm.

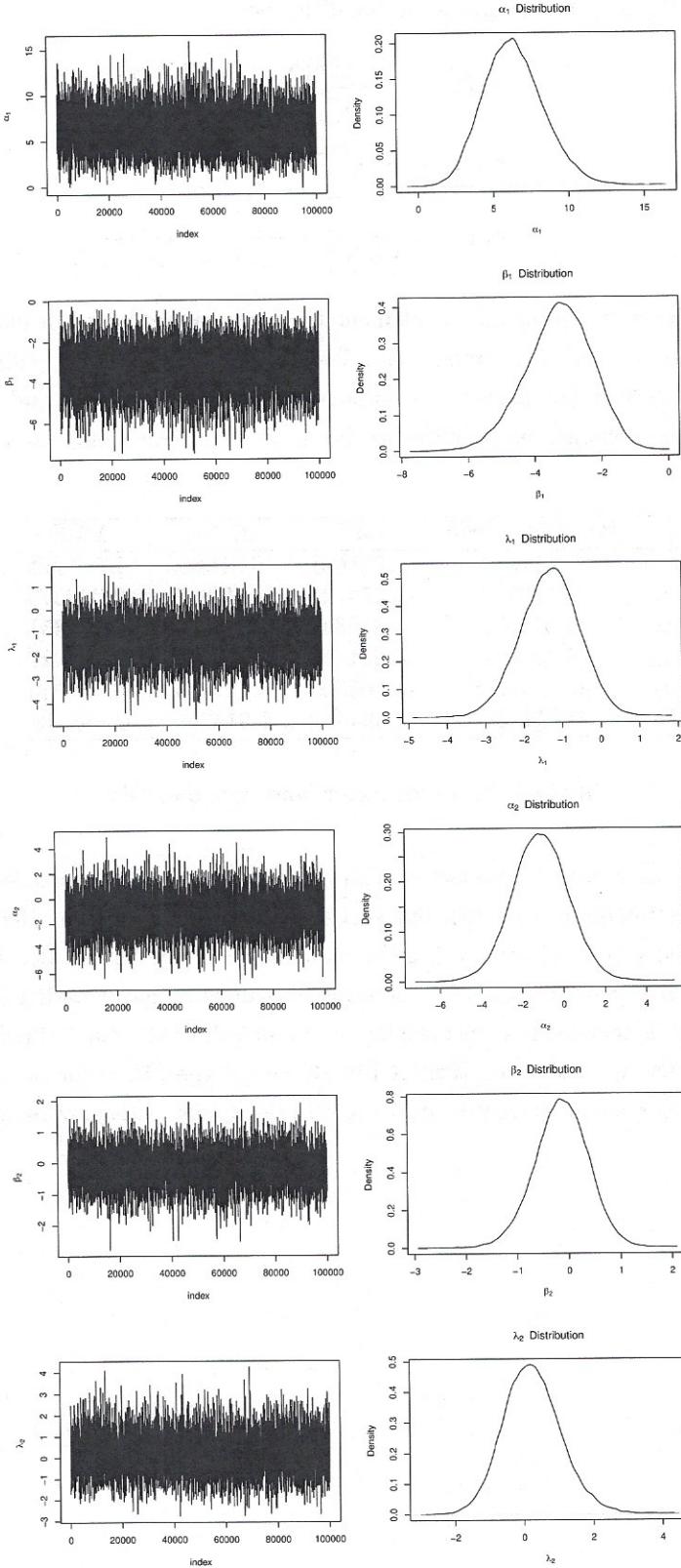


Figure 3: Healthy Convergence of parameters $\alpha_1, \beta_1, \lambda_1, \alpha_2, \beta_2, \lambda_2$

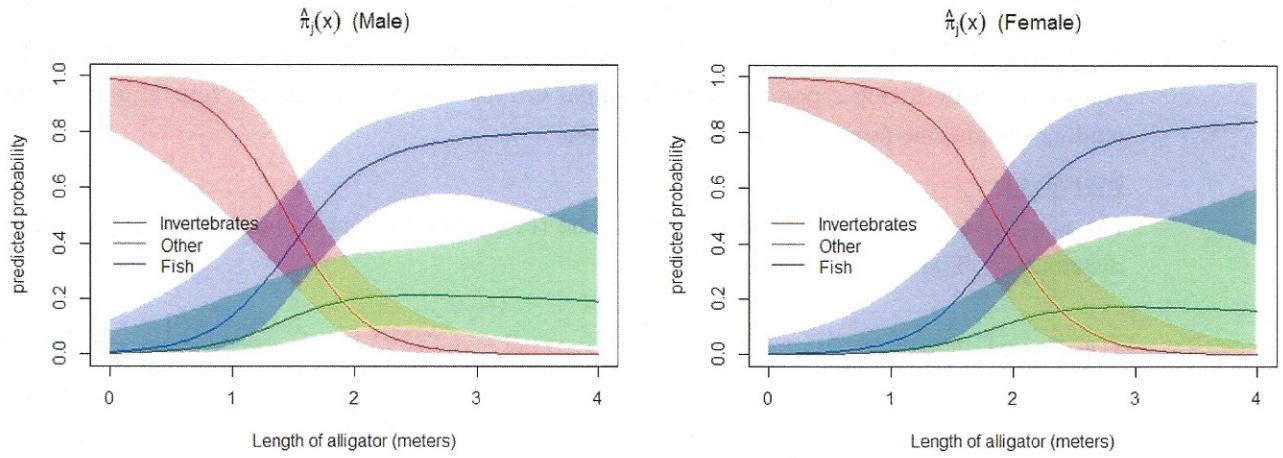


Figure 4: point and interval estimates for the estimated response curves $\pi_j(x)$; $j = I, O, F$ for male and female alligators

In Figure 4, we can see that male alligators tend to eat fish instead of invertebrates or other food when they reach past a length of 1 meter in size. Whereas females tend to eat fish instead of invertebrates or other food when they reach past a length of 2 meters. The width of intervals for “Other” and “Fish” keep increasing for a male alligator length over 2.5 meters, showing a lot of uncertainty about the choice of “Other” or “Fish” for male big alligators. The intervals for female alligators are wider than the intervals for male alligators. This may be due to a small female sample size of 24 compared to the male sample size of 39.

3. The table below reports results from a developmental toxicity study involving ordinal categorical outcomes. This study administered diethylene glycol dimethyl ether (an industrial solvent used in the manufacture of protective coatings) to pregnant mice. Each mouse was exposed to one of five concentration levels for ten days early in the pregnancy (with concentration 0 corresponding to controls). Two days later, the uterine contents of the pregnant mice were examined for defects. One of three (ordered) outcomes (“Dead”, “Malformation”, “Normal”) was recorded for each fetus.

Concentration (mg/kg per day) (x_i)	Response			Total Number of subjects (m_i)
	Dead (y_{i1})	Malformation (y_{i2})	Normal (y_{i3})	
0	15	1	281	297
62.5	17	0	225	242
125	22	7	283	312
250	38	59	202	299
500	144	132	9	285

Build a multinomial regression model for these data using continuation-ratio logits for the response probabilities $\pi_j(x)$, $j = 1, 2, 3$, as a function of concentration level, x . Specifically, consider the following model

$$L_1^{(cr)} = \log \left(\frac{\pi_1}{\pi_2 + \pi_3} \right) = \alpha_1 + \beta_1 x; \quad L_2^{(cr)} = \log \left(\frac{\pi_2}{\pi_3} \right) = \alpha_2 + \beta_2 x$$

for the multinomial response probabilities $\pi_j \equiv \pi_j(x)$, $j = 1, 2, 3$.

- (a) Show that the model, involving the multinomial likelihood for the data $= \{(y_{i1}, y_{i2}, y_{i3}, x_i) : i = 1, \dots, 5\}$, can be fitted by fitting separately two Binomial GLMs. Provide details for your argument, including the specific form of the Binomial GLMs.

Solution:

Let $\rho_1 = \pi_1$ and $\rho_2 = \frac{\pi_2}{\pi_2 + \pi_3}$. Also, let $f(\cdot)$ be the probability mass function for the binomial distribution. Then,

$$\begin{aligned} f(y_{i1}|m, \rho_1) \cdot f(y_{i2}|m - y_{i1}, \rho_2) &= \frac{m!}{y_{i1}!(m - y_{i1})!} \rho_1^{y_{i1}} (1 - \rho_1)^{m - y_{i1}} \times \frac{(m - y_{i1})!}{y_{i2}!(m - y_{i1} - y_{i2})!} \rho_2^{y_{i2}} (1 - \rho_2)^{m - y_{i1} - y_{i2}} \\ &= \frac{m!}{y_{i1}!y_{i2}!y_{i3}!} \pi_1^{y_{i1}} (1 - \pi_1)^{m - y_{i1}} \times \left(\frac{\pi_2}{\pi_2 + \pi_3} \right)^{y_{i2}} \left(\frac{\pi_3}{\pi_2 + \pi_3} \right)^{m - y_{i1} - y_{i2}} \\ &= \frac{m!}{y_{i1}!y_{i2}!y_{i3}!} \pi_1^{y_{i1}} (\pi_2 + \pi_3)^{m - y_{i1}} \times \left(\frac{\pi_2}{\pi_2 + \pi_3} \right)^{y_{i2}} \left(\frac{\pi_3}{\pi_2 + \pi_3} \right)^{m - y_{i1} - y_{i2}} \\ &= \frac{m!}{y_{i1}!y_{i2}!y_{i3}!} \pi_1^{y_{i1}} \pi_2^{y_{i2}} \pi_3^{y_{i3}} (\pi_2 + \pi_3)^{m - y_{i1} - y_{i2} - m + y_{i1} + y_{i2}} \\ &= \frac{m!}{y_{i1}!y_{i2}!y_{i3}!} \pi_1^{y_{i1}} \pi_2^{y_{i2}} \pi_3^{y_{i3}} \end{aligned}$$

which is the pmf for the multinomial distribution. By the multinomial property $\pi_3 = 1 - \pi_1 - \pi_2$. Then $L_1^{(cr)} = \log \left(\frac{\pi_1}{1 - \pi_1} \right) = \alpha_1 + \beta_1 x$ and $L_2^{(cr)} = \log \left(\frac{\pi_2}{1 - \pi_1 - \pi_2} \right) = \alpha_2 + \beta_2 x$, and so our model can be fitted with two separate binomial glms, $Bin(y_{i1}|m_i, \rho_1)$ for $L_1^{(cr)}$ and $Bin(y_{i2}|m_i - y_{i1}, \rho_2)$ for $L_2^{(cr)}$.

- (b) Use the result from part (a) to obtain the MLE estimates and corresponding standard errors for parameters $(\alpha_1, \alpha_2, \beta_1, \beta_2)$. Plot the estimated response curves $\hat{\pi}_j(x)$, for $j = 1, 2, 3$, and discuss the results.

Solution:

Using the $glm(\cdot)$ function in R, we obtain the mle and standard errors for the parameters $(\alpha_1, \alpha_2, \beta_1, \beta_2)$ and shown in Table 4.

	α_1	β_1	α_2	β_2
MLE	-3.247934	0.006389	-5.70190	0.01737
standard error	0.1576602	0.0004348	0.332248	0.001227

Table 4: MLE and standard error estimates for the parameters

Figure 5 shows the plotted estimated response curves $\hat{\pi}_j(x)$, for $j = 1, 2, 3$, where

$$\begin{aligned}\hat{\pi}_1(x) &= \frac{\exp(\hat{\alpha}_1 + \hat{\beta}_1 x)}{1 + \exp(\hat{\alpha}_1 + \hat{\beta}_1 x)} \\ \hat{\pi}_2(x) &= \frac{\exp(\hat{\alpha}_2 + \hat{\beta}_2 x)}{1 + \exp(\hat{\alpha}_2 + \hat{\beta}_2 x)} \times \frac{1}{1 + \exp(\hat{\alpha}_1 + \hat{\beta}_1 x)} \\ \hat{\pi}_3(x) &= 1 - \hat{\pi}_1(x) - \hat{\pi}_2(x)\end{aligned}$$



They correspond to “Dead”, “Malformation”, and “Normal” respectively. We can obtain those equations by solving for π_1 in $L_1^{(cr)}$ and solving for π_2 in $L_2^{(cr)}$. Observing the curves in Figure 5, we see that they fit the data well. As the concentration level of diethylene glycol dimethyl ether increases, the probability of the mice dieing and the probability of mice becoming malformed go up, whereas the probability of mice staying normal goes down. However, the probability of mice becoming malformed reaches its high peak around a concentration level of 400 and seems to slowly decreases afterwards. From this we know $\hat{\pi}_2(x)$ is a logistic curve.

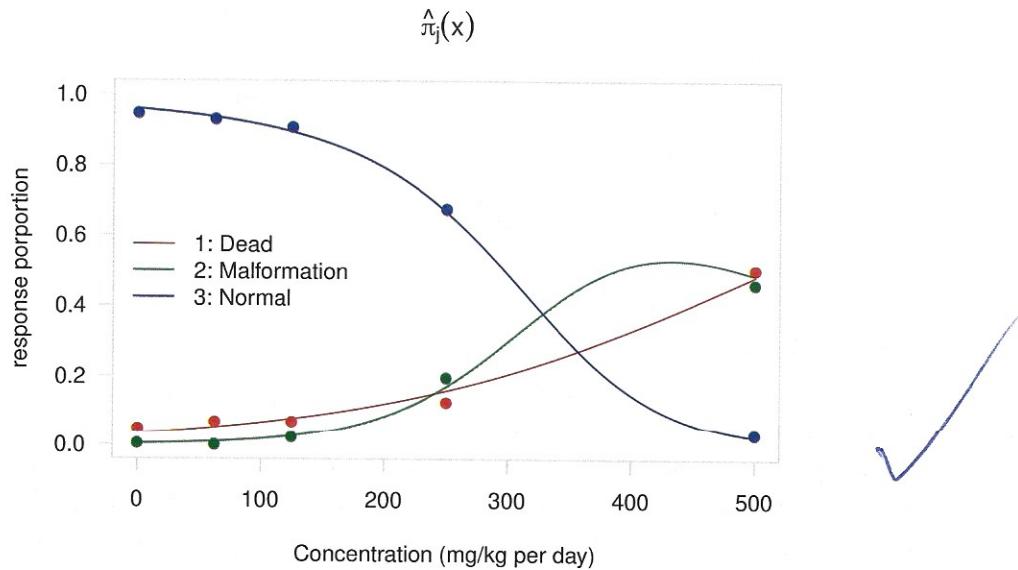


Figure 5: Estimated response curves $\hat{\pi}_j(x)$, for $j = 1, 2, 3$

- (c) Develop and implement a Bayesian version of the model above. Discuss your prior choice, and provide details for the posterior simulation method. Provide point and interval estimates for the response curves $\pi_j(x)$, for $j = 1, 2, 3$.

Solution:

Suppose $p(\alpha_1, \alpha_2, \beta_1, \beta_2)$ are independent priori. That is $p(\alpha_1, \alpha_2, \beta_1, \beta_2) = p(\alpha_1, \beta_1)p(\alpha_2, \beta_2)$.

Then our posterior distribution is

$$\begin{aligned} p(\alpha_1, \alpha_2, \beta_1, \beta_2 | data) &\propto \left[p(\alpha_1, \beta_1) \prod_{i=1}^5 Bin(y_{i1} | m_i, \rho_1) \right] \left[p(\alpha_2, \beta_2) \prod_{i=1}^5 Bin(y_{i2} | m_i - y_{i1}, \rho_2) \right] \\ &\propto k_1(\alpha_1, \beta_1) k_2(\alpha_2, \beta_2, \alpha_1, \beta_1) \end{aligned}$$

Where $k_1(\cdot)$ and $k_2(\cdot)$ is the exponentiated log of the respective posterior components. Note that $k_2(\cdot)$ will contain α_1 and β_1 also because they are parameters in ρ_2 . We choose our priors to be $p(\alpha_1, \beta_1) \propto 1$ and $p(\alpha_2, \beta_2) \propto 1$ to show that we have now prior information about this study. Looking at the posterior we notice that we can build our model by fitting two separate binomial glms. Hence, the Bayesian version of our model will arise from two sets of MCMC. The MCMC algorithms are shown in Algorithm 1 and Algorithm 2.

Algorithm 1: Metropolis-Hastings

```

1. initialize  $(\alpha_1^{(1)}, \beta_1^{(1)}) = (0, 0)$ 
for  $i$  to 2 : 200000
  2. draw  $(\alpha_1^*, \beta_1^*)_i \sim N_2 \left( (\alpha_1^{(t-1)}, \beta_1^{(t-1)}), d \cdot J^{-1} (\hat{\alpha}_1, \hat{\beta}_1) \right)$  where  $d$  is a
     tuning parameter,  $J^{-1} (\hat{\alpha}_1, \hat{\beta}_1)$  is the fisher information and  $\hat{\alpha}_1, \hat{\beta}_1$  are the
     mle's
  3. If  $U \sim Unif(0, 1) < q = \min \left\{ 1, \frac{k_1(\alpha_1^*, \beta_1^*)_i}{k_2(\alpha_1^{(t-1)}, \beta_1^{(t-1)})} \right\}$ . Set
      $(\alpha_1^{(t)}, \beta_1^{(t)}) = (\alpha_1^*, \beta_1^*)_i$ 
  4. Else  $(\alpha_1^{(t)}, \beta_1^{(t)}) = (\alpha_1^{(t-1)}, \beta_1^{(t-1)})$ 
end

```

When implementing Algorithm 1 we set $d = 39$. The total number of iterations is 200000 and we burn the first 100000. Leaving us 100000 draws left for analysis. Figure 6 (a) and (b) are trace plots of posterior parameters α_1 and β_1 , respectively. These trace plot exhibits rapid up-and-down variation with no long-term trends or drifts. This indicates convergence in distribution. In Figure 6 (c) and (d) we see that α_1 and β_1 both follow a normal distribution. In R, the posterior mean estimates are $\hat{\alpha}_1 = -3.249984$ which is close to the true parameter estimate of -3.247934 ; and $\hat{\beta}_1 = 0.006384$ which is also close to the true parameter estimate of 0.006389 . Thus Algorithm 1 has performed well.

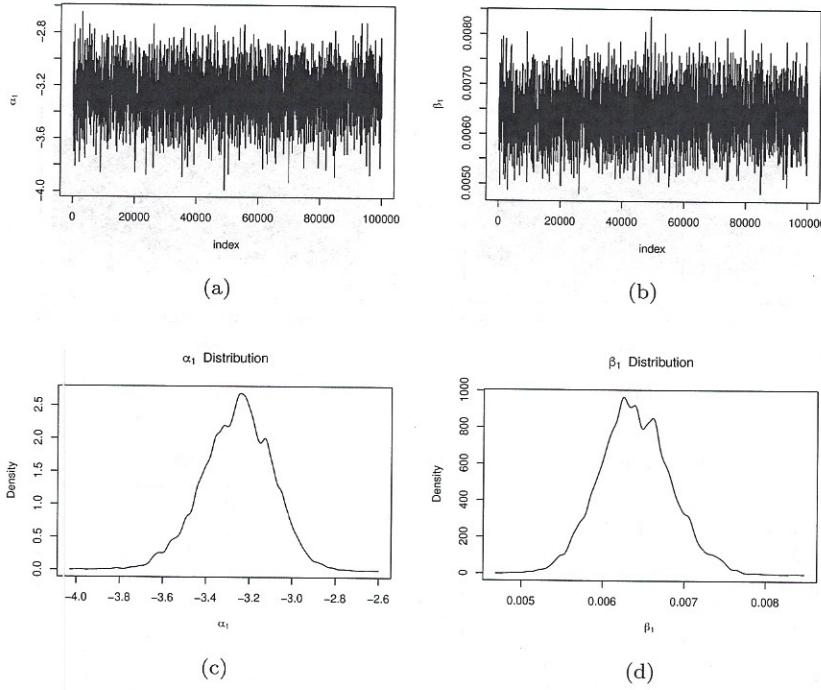


Figure 6: Healthy Convergence of α_1, β_1

Next, we implement Algorithm 2 in a similar fashion as Algorithm 1 in order to obtain posterior estimates of α_2 and β_2 . We set $d = 4$. The total number of iterations is 200000 and we burn the first 100000. Leaving us 100000 draws left for analysis.

Algorithm 2: Metropolis-Hastings

```

1. initialize  $(\alpha_2^{(1)}, \beta_2^{(1)}) = (0, 0)$ 
for  $i$  to  $2 : 200000$ 
    2. draw  $(\alpha_2^*, \beta_2^*)_i \sim N_2 \left( (\alpha_2^{(t-1)}, \beta_2^{(t-1)}), d \cdot J^{-1}(\hat{\alpha}_2, \hat{\beta}_2) \right)$  where  $d$  is a
       tuning parameter,  $J^{-1}(\hat{\alpha}_2, \hat{\beta}_2)$  is the fisher information and  $\hat{\alpha}_2, \hat{\beta}_2$  are the
       mle's
    3. If  $U \sim Unif(0, 1) < q = \min \left\{ 1, \frac{k_1(\alpha_2^*, \beta_2^*, \hat{\alpha}_1, \hat{\beta}_1)_i}{k_2(\alpha_2^{(t-1)}, \beta_2^{(t-1)}, \hat{\alpha}_1, \hat{\beta}_1)} \right\}$ . Set
        $(\alpha_2^{(t)}, \beta_2^{(t)}) = (\alpha_2^*, \beta_2^*)_i$ 
    4. Else  $(\alpha_2^{(t)}, \beta_2^{(t)}) = (\alpha_2^{(t-1)}, \beta_2^{(t-1)})$ 
end

```

Figure 7 (a) and (b) are trace plots of posterior parameters α_2 and β_2 , respectively. These trace plot exhibits rapid up-and-down variation with no long-term trends or drifts. This indicates convergence in distribution. In Figure 3 (c) and (d) we see that α_2 and β_2 both follow a normal distribution. In R, the posterior mean estimates are $\hat{\alpha}_2 = -5.753089$ which is close to the true

parameter estimate of -5.70190 and $\hat{\beta}_2 = 0.017578$ which is close to the true parameter estimate of 0.01737 . Thus Algorithm 2 has also performed well.

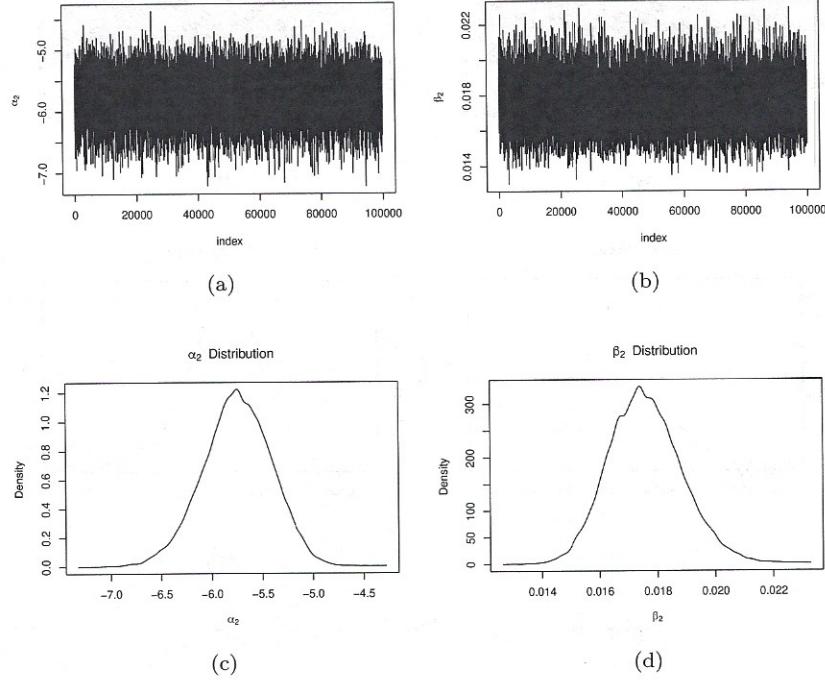


Figure 7: Healthy Convergence of α_2, β_2

	Posterior Mean	Posterior SD	2.5%	97.5%	Table4 mean	Table 4 sd
α_1	-3.249984	0.161354	-3.583850	-2.958953	-3.247934	0.1576602
β_1	0.006384	0.000443	0.005568	0.007324	0.006389	0.0004348
α_2	-5.745362	0.339905	-6.446320	-5.121535	-5.70190	0.332248
β_2	0.017522	0.001261	0.015254	0.020155	0.017370	0.001227

Table 5: postreior means and their 95% quantiles, and mle from table 4.

Figure 8 shows the point and 95% interval estimates for the response curves $\hat{\pi}_j(x)$, for $j = 1, 2, 3$. Similar inference can be made in Figure 8 as in Figure 5, so we focus on the interval estimates. For the Dead and Malformation category, the interval is getting wider as the concentration level increases. This shows that there is an uncertainty that a porportion of mice will either die or malform after being exposed to high concentration levels. The interval for Normal category is also wider when the concentration level is higher, but becomes narrower when the concentration goes above 400. Meaning that there is an uncertainty that a porportion of mice will either die, malform, or remain normal after being exposed to high concentration levels roughly between 250 and 400.

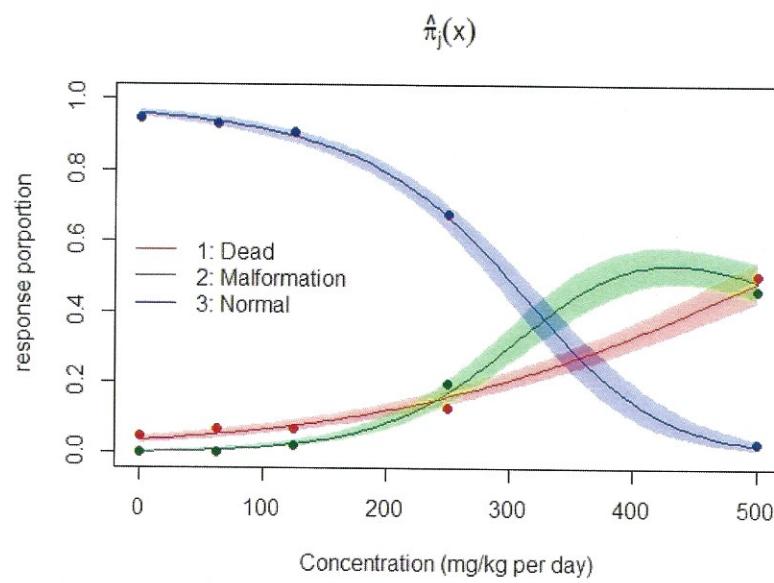


Figure 8: point and interval estimates for response curves $\hat{\pi}_j(x)$, for $j = 1, 2, 3$

nice work