AMS 274 Homework #4 **Mary Silva**
UCSC 16 November 2018

1. (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.

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

Since $\varepsilon_0, \varepsilon_1 \sim N(0,1)$, $\varepsilon_0 - \varepsilon_1 \sim N(0,2)$. It follows then that

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

where $\beta_0 = \frac{a_1 - a_0}{\sqrt{2}}$, $\beta_1 = \frac{b_1 - b_0}{\sqrt{2}}$, and $\varepsilon = \frac{\varepsilon_0 - \varepsilon_1}{\sqrt{2}} \sim N(0,1)$

(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.

Solution: Here we have again that

$$P(Y=1) = P(a_1 - a_0 + (b_1 - b_0)x > \varepsilon_0 - \varepsilon_1)$$

However, here the error terms have the c.d.f. $F(\varepsilon) = \exp\{-\exp(-\varepsilon)\}$, which is the standardized Gumbel distribution. If $\varepsilon_0, \varepsilon_1 \sim \text{Gumbel}(0,1)$, then $\varepsilon_1 - \varepsilon_0 \sim \text{Logistic}(0,1)$ by lemma 1 below.

Lemma 1. *Difference of two Gumbel random variables is a Logistic random variable, i.e.,*

$$X, Y \sim \textit{Gumbel}(0,1) \implies W = X - Y \sim \textit{Logistic}(0,1).$$

In other words,

$$P(W < w) = \frac{1}{1 + e^{-w}}.$$

Proof. This fact can be proved using convolution, which is messy. We shall use moment generating functions to prove it instead. Note that

$$\begin{aligned} \mathrm{E}[\exp(\theta W)] &= \mathrm{E}[\exp(\theta X)\ \exp(-\theta Y)] \\ &= \mathrm{E}[\exp(\theta X)]\mathrm{E}[\exp(-\theta Y)] \\ \implies M_W(\theta) &= M_X(\theta) M_Y(-\theta). \end{aligned}$$

where $M_W(\theta)$ is the moment generating function of W. It is known that

$$\mathrm{E}[M_X(\theta)] = \Gamma(1-\theta)\exp(\theta).$$

where $\Gamma(n) = (n-1)!$ if n is an integer. Thus,

$$\begin{aligned}
\mathrm{E}[\exp(\theta W)] &= \Gamma(1-\theta)\exp(\theta)\Gamma(1+\theta)\exp(-\theta) \\
&= \Gamma(1-\theta)\Gamma(1+\theta) \\
&= \frac{\Gamma(1-\theta)\Gamma(1+\theta)}{\Gamma(2)} \\
&= \mathrm{Beta}(1-\theta, 1+\theta).
\end{aligned}$$

where $\mathrm{Beta}(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ is the Beta distribution which is also the MGF of Logistic$(0,1)$ □

2 (a) Focusing on length as a single covariate, we develop a Bayesian multinomial regression model using baseline category logits formulation with "fish" as the baseline category. I use a slice sampling algorithm, which I described in detail in homework 3. I found that it converged much more nicely than Metropolis Hastings and the code wasn't difficult to change.

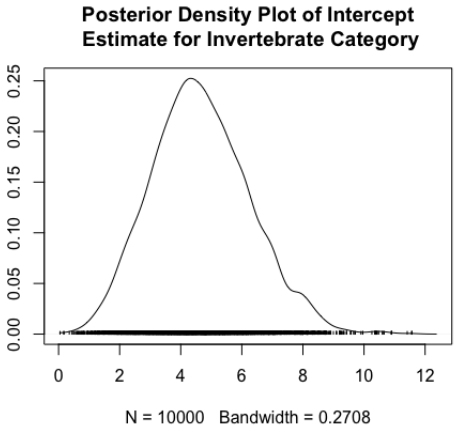
In order to reflect the prior uncertainty of the parameters, since we have no propr knowledge of alligators or their eating habbits, we assign flat priors. The response probabilities as a function of length are as follows

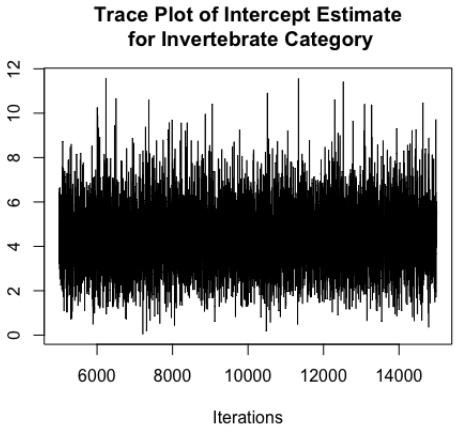
$$\hat{\pi}_{\mathrm{I}} = \frac{\exp\{4.66 - 2.76x_1\}}{1 + \exp\{4.66 - 2.76x_1\} + \exp\{-0.96 - 0.13x_1\}}$$

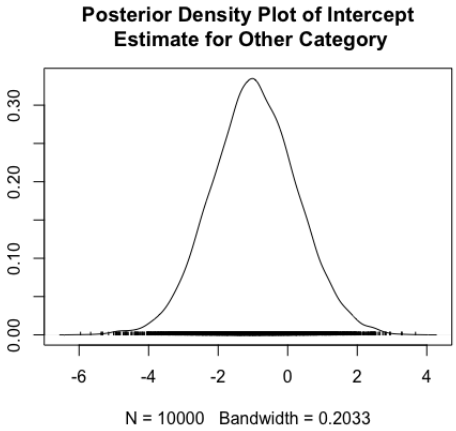
$$\hat{\pi}_{\mathrm{O}} = \frac{\exp\{-0.96 - 0.13x_1\}}{1 + \exp\{4.66 - 2.76x_1\} + \exp\{-0.96 - 0.13x_1\}}$$

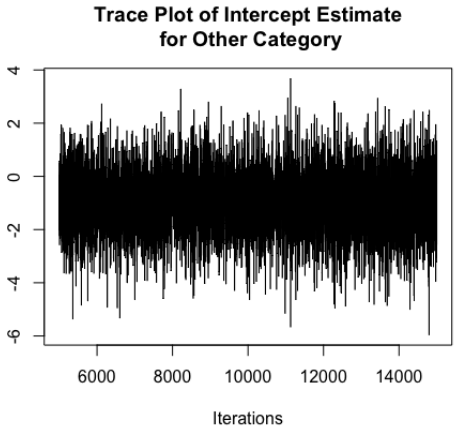
$$\hat{\pi}_{\mathrm{F}} = \frac{1}{1 + \exp\{4.66 - 2.76x_1\} + \exp\{-0.96 - 0.13x_1\}}$$

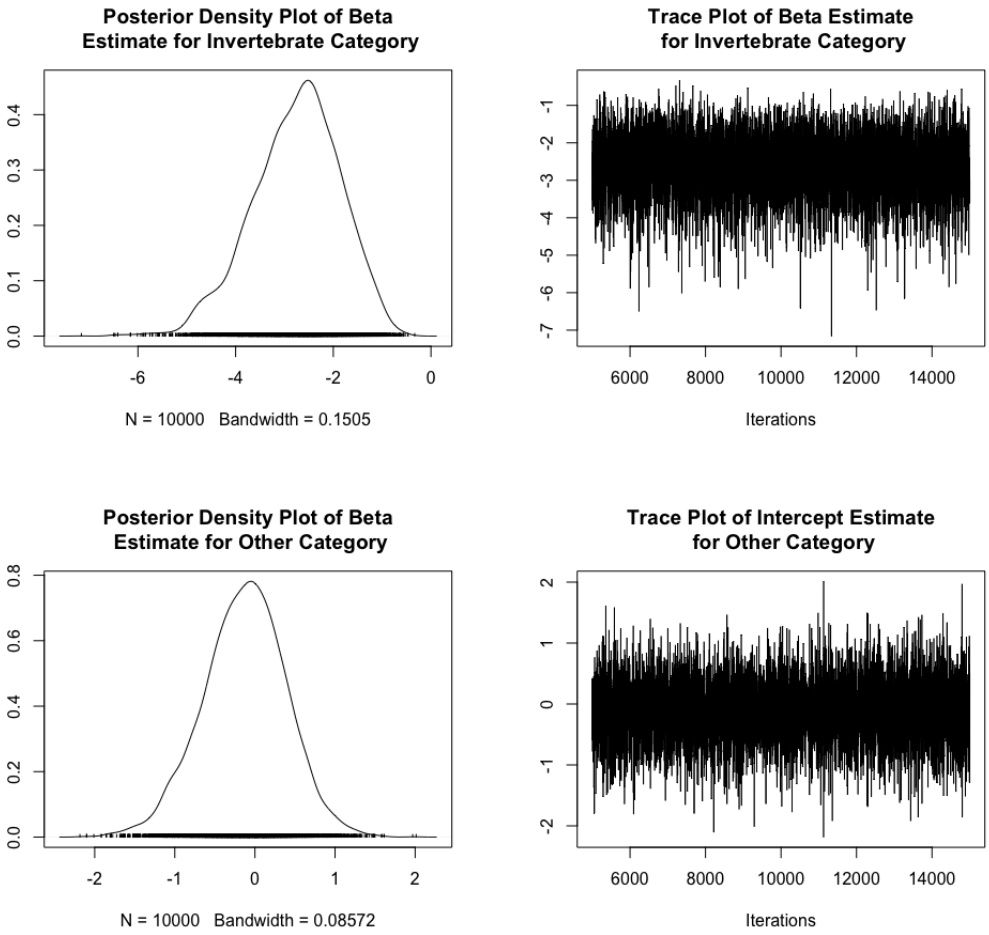
As a sanity check, and to check convergence, we examine the following plots







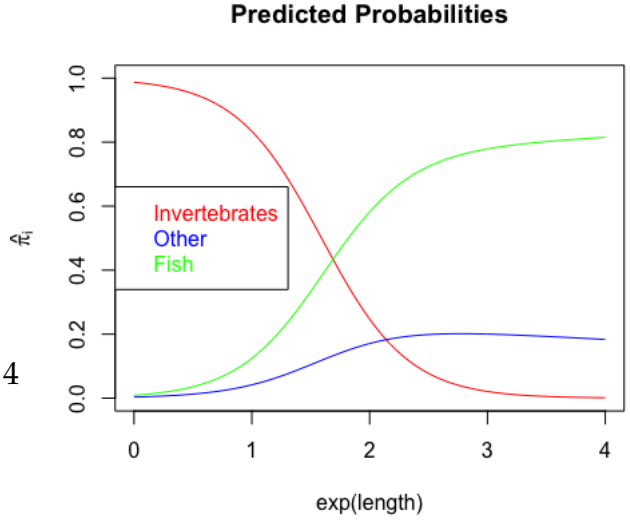




From the above plots, we see that there are no issues with convergence, and density plot indicates that the sampler has converged on the maxima for each parameter. The summary table below provides much more detail for the parameter estimates

	Mean	SD	2.5%	25%	50%	75%	97.5%
$\hat{\alpha}_I$	4.66	1.61	1.77	3.54	4.57	5.70	8.01
$\hat{\alpha}_O$	-0.96	1.23	-3.36	-1.77	-0.97	-0.15	1.46
$\hat{\beta}_I$	-2.76	0.90	-4.68	-3.33	-2.69	-2.12	-1.18
$\hat{\beta}_O$	-0.13	0.52	-1.18	-0.47	-0.11	0.22	0.86

Lastly, we produce a plot of our prediction as a function of length.



2 (b) Now, we extend the model from part (a) to describe the effects of both length and gender on food choice. Just like before, the response probabilities were shown to be

$$\hat{\pi}_{ij} = \frac{\exp\{\hat{\alpha}_j + x_i^T \hat{\beta}_j\}}{1 + \sum_{k \neq J} \exp\{\hat{\alpha}_j + x_i^T \hat{\beta}_k\}}$$

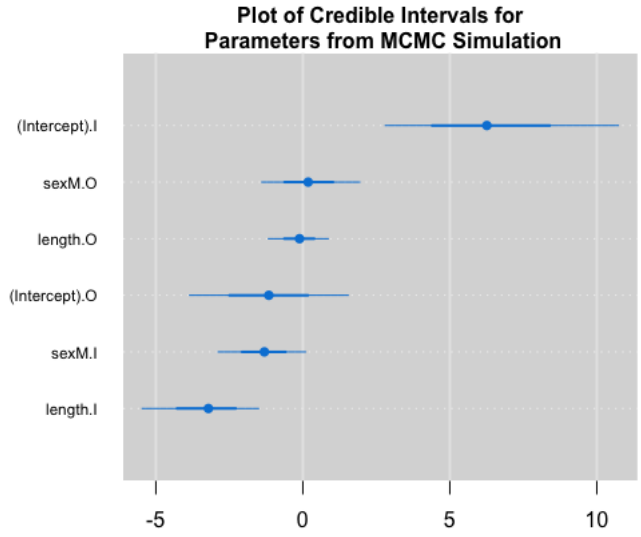
for the non-baseline category, where $j \neq J$ is again the category for "fish". And for the baseline category, we have

$$\hat{\pi}_{ij} = \frac{1}{1 + \sum_{k \neq J} \exp\{\hat{\alpha}_j + x_i^T \hat{\beta}_k\}}$$

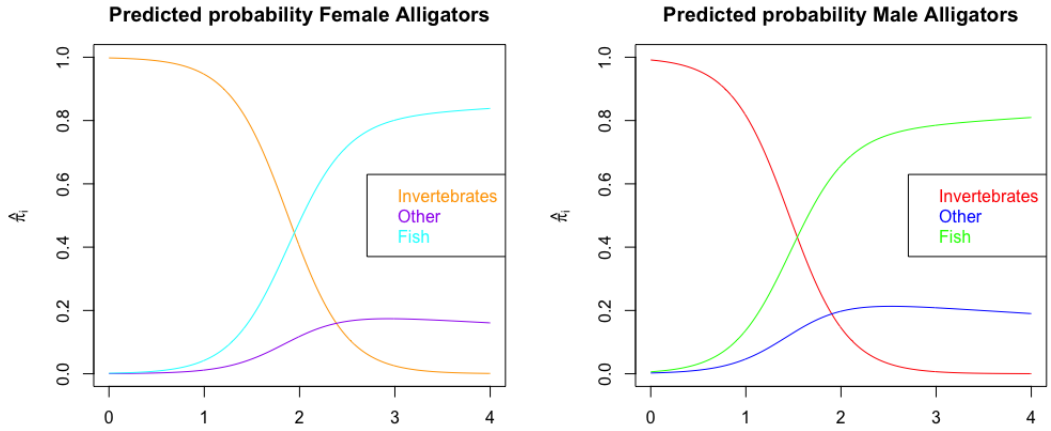
The estimates for the above equations are are summarized in the table below where $\hat{\beta}_{1,j}$ is for gender = male, and $\hat{\beta}_{2,j}$ is for gender = female.

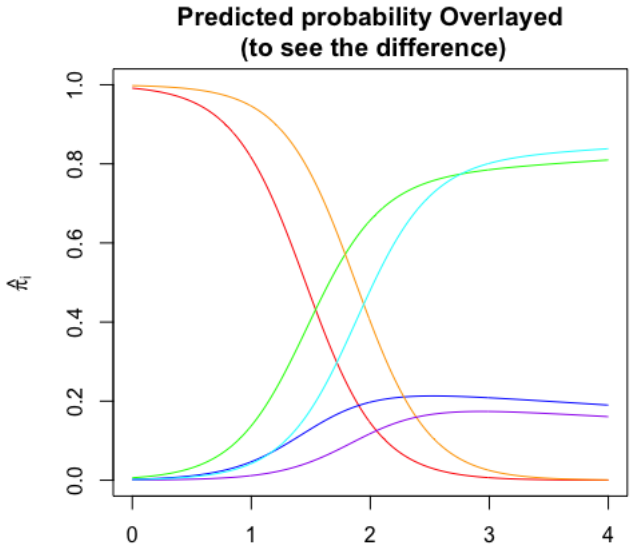
	Mean	SD	2.5%	25%	50%	75%	97.5%
$\hat{\alpha}_I$	6.40	2.03	2.81	4.98	6.26	7.69	10.73
$\hat{\alpha}_O$	-1.16	1.37	-3.85	-2.06	-1.16	-0.26	1.54
$\hat{\beta}_{1,I}$	-1.33	0.75	-2.88	-1.81	-1.31	-0.82	0.09
$\hat{\beta}_{1,O}$	0.20	0.84	-1.39	-0.37	0.18	0.74	1.93
$\hat{\beta}_{2,I}$	-3.29	1.01	-5.46	-3.92	-3.21	-2.57	-1.51
$\hat{\beta}_{2,O}$	-0.12	0.52	-1.17	-0.46	-0.11	0.22	0.86

The following image shows a plot of credible intervals for parameters estimates from MCMC simulation



Lastly, plots of the response curves as a function of length and fixed genders are shown below.





3 (a) Let $\rho_j = \dfrac{\pi_j}{\pi_j + ... + \pi_J}$ and let $f(y|m, \rho)$ be the p.m.f. for the binomial distribution. Also, let $\boldsymbol{y}_i = (y_{i1}, y_{i2}, y_{i3})$. Then,

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

which is the p.m.f for the multinomial distribution. Therefore, the continuous-ratio logits model model can be fit by fitting independent binomial GLMs.

3 (b) Using the result from part (a), we obtain the MLE estimates and corresponding standard errors for parameters $(\alpha_1, \alpha_2, \beta_1, \beta 2)$. These estimates are summarized in table 1.

Parameter	Estimate	SE
$\hat{\alpha}_1$	-3.248	0.1577
$\hat{\alpha}_2$	-5.7019	0.3322
$\hat{\beta}_1$	-0.0063	0.0004
$\hat{\beta}_2$	0.0174	0.0012

Table 1: MLE estimates for the model described in part (a).

Based on the figure 1, it appears that as the concentration of diethylene glycodimethyl ether increases, the proportion of normal mice decreases. Similarly, the proportion of malformed and dead mice increased with concentration.

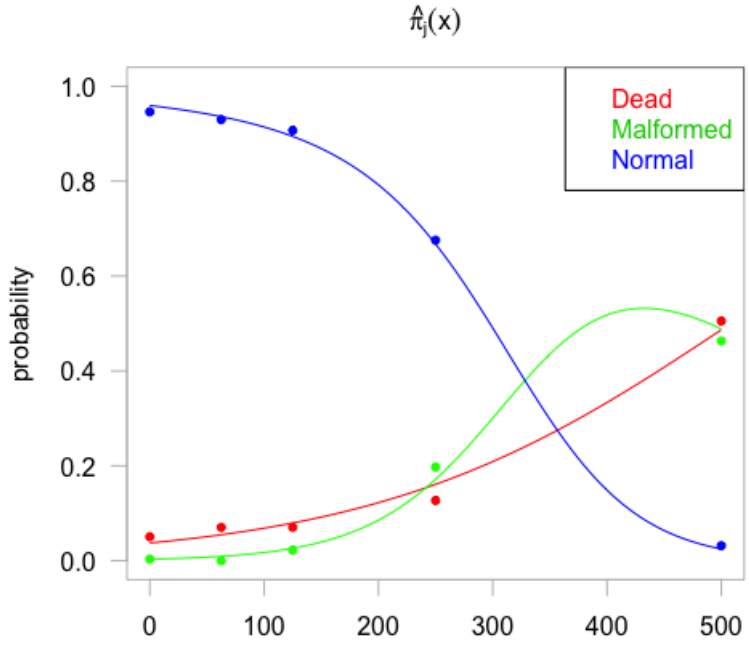


Figure 1: The estimated response curves $\hat{\pi}_j(x)$, for $j = 1, 2, 3$.

3 (c) In order to reflect the uncertainties in the parameters, the posterior distribution for the covariates (α_i, β_j) were chosen to be $N_2(\mathbf{0}, 100\boldsymbol{I}_2)$ for $j = 1, 2$. A Metropilis within Gibbs sampling algorithm is implemented to simulate from the joint posterior. Following the style of part (a), two separate MCMC sampling algorithms were implemented to obtain the posterior distributions of (α_1, β_1) and (α_2, β_2). The results are summarized in table 2.

	Mean	SD	CI 2.5%	CI 97.5%
α_1	-3.2554	0.1504	-3.5515	-2.9696
α_2	-5.7636	0.3437	-6.5308	-5.1614
β_1	0.0064	0.0004	0.0055	0.0073
β_2	0.0176	0.0013	0.0155	0.0206

Table 2: Summary of posterior distributions of (α_1, β_1) and (α_2, β_2).