1. The table below reports 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. 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 the concentration level, $x$. Specifically, consider the following model

$$L^{(cr)}_1 = \log \left( \frac{\pi_1}{\pi_2 + \pi_3} \right) = \alpha_1 + \beta_1 x; \quad L^{(cr)}_2 = \log \left( \frac{\pi_2}{\pi_3} \right) = \alpha_2 + \beta_2 x$$

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 $\mathcal{D} = \{(y_{i1}, y_{i2}, y_{i3}, x_i) : i = 1, ..., 5\}$, can be fitted by fitting separately two Binomial GLMs. Provide details for your argument, including the specific form of the Binomial GLMs.

**Solution:**

Note that for a multinomial distribution $\text{Multi}(y_1, y_2, y_3 | m, \pi_1, \pi_2, \pi_3)$,

$$\begin{align*}
\text{Multi}(y_1, y_2, y_3 | m, \pi_1, \pi_2, \pi_3) &= \frac{m!}{y_1!y_2!y_3!} \pi_1^{y_1} \pi_2^{y_2} \pi_3^{y_3} \\
&= \frac{m!}{y_1!(m-y_1)!} \pi_1^{y_1} (1-\pi_1)^{m-y_1} (1-\pi_1)^{-(m-y_1)} \pi_2^{y_2} \pi_3^{y_3} \\
&= \frac{m!}{y_1!(m-y_1)!} \pi_1^{y_1} (1-\pi_1)^{m-y_1} \frac{(m-y_1)!}{y_2!(m-y_1-y_2)!} \left( \frac{\pi_2}{1-\pi_1} \right)^{y_2} \left( \frac{\pi_3}{1-\pi_1} \right)^{y_3} \\
&= \frac{m!}{y_1!(m-y_1)!} \pi_1^{y_1} (1-\pi_1)^{m-y_1} \frac{(m-y_1)!}{y_2!(m-y_1-y_2)!} \left( \frac{\pi_2}{\pi_2 + \pi_3} \right)^{y_2} \left( \frac{\pi_3}{\pi_2 + \pi_3} \right)^{m-y_1-y_2} \\
&= \text{Bino}(y_1 | m, \rho_1) \times \text{Bino}(y_2 | m - y_1, \rho_2),
\end{align*}$$

where $\rho_1 = \frac{\pi_2}{1-\pi_1}$ and $\rho_2 = \frac{\pi_3}{1-\pi_1}$.
where \( \rho_1 = \pi_1, \rho_2 = \frac{\pi_2}{\pi_2 + \pi_3} \).

Hence for each \( i = 1, \ldots, 5 \), \( \text{Multi}(y_{i1}, y_{i2}, y_{i3} \mid m_i, \pi_1(x_i), \pi_2(x_i), \pi_3(x_i)) \) can be expressed by the product of two Binomial distributions \( \text{Bino}(y_{i1} \mid m_i, \rho_{i1}) \times \text{Bino}(y_{i2} \mid m_i - y_{i1}, \rho_{i2}) \), and the likelihood can be written as

\[
L(\text{data} \mid \alpha_1, \alpha_2, \beta_1, \beta_2) = \prod_{i=1}^{5} \text{Multi}(y_{i1}, y_{i2}, y_{i3} \mid m_i, \pi_1(x_i), \pi_2(x_i), \pi_3(x_i)) \\
= \prod_{i=1}^{5} \text{Bino}(y_{i1} \mid m_i, \rho_{i1}) \prod_{i=1}^{5} \text{Bino}(y_{i2} \mid m_i - y_{i1}, \rho_{i2}).
\]

Since \( L^{(cr)}_1 = \log \left( \frac{\pi_1}{1 - \pi_1} \right) = \alpha_1 + \beta_1 x \) and \( L^{(cr)}_2 = \log \left( \frac{\pi_2/(\pi_2 + \pi_3)}{\pi_3/(\pi_2 + \pi_3)} \right) = \alpha_2 + \beta_2 x \), the model can be fitted by fitting separately two Binomial GLMs, \( \text{Bino}(y_{i1} \mid m_i, \rho_{i1}) \) for \( L^{(cr)}_1 \) and \( \text{Bino}(y_{i2} \mid m_i - y_{i1}, \rho_{i2}) \) for \( L^{(cr)}_2 \).

(b) Use the result from part (a) to obtain the MLE estimates and corresponding standard errors for the 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:**

The MLE estimates and corresponding standard errors for the parameters \((\alpha_1, \alpha_2, \beta_1, \beta_2)\) are shown in Table 1.

<table>
<thead>
<tr>
<th>( \alpha_1 )</th>
<th>( \beta_1 )</th>
<th>( \alpha_2 )</th>
<th>( \beta_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLE</td>
<td>-3.24793</td>
<td>0.00639</td>
<td>-5.70190</td>
</tr>
<tr>
<td>standard error</td>
<td>0.15766</td>
<td>0.00043</td>
<td>0.33225</td>
</tr>
</tbody>
</table>

Table 1: MLE estimates and corresponding standard errors for the parameters.

Figure 1 shows the estimated response curves \( \hat{\pi}_j(x), j = 1, 2, 3 \). Here,

\[
\hat{\pi}_1(x) = \frac{\exp(\hat{\alpha}_1 + \hat{\beta}_1 x)}{1 + \exp(\hat{\alpha}_1 + \hat{\beta}_1 x)}, \\
\hat{\pi}_2(x) = \frac{\exp(\hat{\alpha}_2 + \hat{\beta}_2 x)}{1 + \exp(\hat{\alpha}_2 + \hat{\beta}_2 x) + \exp(\hat{\alpha}_1 + \hat{\beta}_1 x)}, \\
\hat{\pi}_3(x) = 1 - \hat{\pi}_1(x) - \hat{\pi}_2(x).
\]

The curves fit the data quite well. As the concentration level increases, the probability of Dead and the probability of Malformation go up, and the probability of Normal goes down. There is no data information for the concentration level between 250 and 500. With the constraint \( \sum_{j=1}^{3} \hat{\pi}_j(x) = 1 \), \( \hat{\pi}_2(x) \) is shaped a bit like a logistic curve and it is higher at \( x = 450 \) than \( x = 500 \).

(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 \( \hat{\pi}_j(x) \) for \( j = 1, 2, 3 \).

**Solution:**
Figure 1: Estimated response curves $\hat{\pi}_j(x), j = 1, 2, 3$. Numbers 1, 2, 3 are observed proportions for category Dead (red), Malformation (blue) and Normal (darkgreen)

Suppose the parameters for $L_1^{(cr)}$ and $L_2^{(cr)}$ are independent $a priori$. That is, $p(\alpha_1, \alpha_2, \beta_1, \beta_2) = p(\alpha_1, \beta_1)p(\alpha_2, \beta_2)$. Then the posterior distribution is

$$p(\alpha_1, \alpha_2, \beta_1, \beta_2 \mid \text{data}) \propto p(\alpha_1, \alpha_2, \beta_1, \beta_2) \prod_{i=1}^5 \text{Multi}(y_{i1}, y_{i2}, y_{i3} \mid m_i, \pi_1(x_i), \pi_2(x_i), \pi_3(x_i))$$

$$= \left[ p(\alpha_1, \beta_1) \prod_{i=1}^5 \text{Bino}(y_{i1} \mid m_i, \rho_{i1}) \right] \left[ p(\alpha_2, \beta_2) \prod_{i=1}^5 \text{Bino}(y_{i2} \mid m_i - y_{i1}, \rho_{i2}) \right]$$

$$:= S_1(\alpha_1, \beta_1)S_2(\alpha_2, \beta_2)$$

We expect $\beta_1$ and $\beta_2$ to be positive because higher concentration of diethylene glycol dimethyl ether hurts human body more. Hence I use normal prior with mean zero for the intercept $\alpha_1$ and $\alpha_2$ and mean 0.001 for the slope $\beta_1$ and $\beta_2$. $\alpha_1, \alpha_2, \beta_1$ and $\beta_2$ are assumed independent $a priori$ and prior variance is assumed to be 100 for each parameter to reflect weak prior information about this study.

Because the model can be fitted by two separate binomial GLMs, the Bayesian version of the model has two sets of MCMC and each has two parameters. The MCMC algorithm is shown in Algorithm 1. The two sets of MCMC can be implemented in parallel.

The initial values for both $(\alpha_1, \beta_1)$ and $(\alpha_2, \beta_2)$ are $(0, 0)$. The total number of iterations is 110000. I burn the first 10000 draws and keep every 100th draw and discard the rest. The 1000 draws are left for analysis. The tuning parameters $d_i$ are chosen so that the acceptance rate is around 40%.

Figure 4 shows a good convergent behavior of the parameters $\alpha_1$ and $\beta_1$. Although not shown here, the convergent behavior of $\alpha_2$ and $\beta_2$ is also healthy. Posterior summary is in Table 2. The posterior means are close to the MLEs. On average, the estimated odds...
Figure 2: Convergence of parameters in Problem 1(c).

Figure 3: Point and interval estimates for the estimated response curves $\hat{\pi}_j(x), j = 1, 2, 3$. Solid lines are for posterior means. Shaded areas are 95% interval estimates. Numbers 1, 2, 3 are observed proportions for category Dead (red), Malformation (blue) and Normal (darkgreen).
Algorithm 1: Metropolis-Hastings Algorithm for multinomial GLM using continuation-ratio logits

for iter = 1, 2, . . . until convergence do
  for i = 1, 2 (Can do it in parallel) do
    1. Draw \((\alpha_i, \beta_i)^* \sim N_2((\alpha_i, \beta_i)^{(m-1)}, d_i J_i^{-1}(\hat{\alpha}_i, \hat{\beta}_i))\), where \(J_i^{-1}\) is the Fisher information, \(d_i\) is the tuning parameter and \(\hat{\alpha}_i, \hat{\beta}_i\) are the MLEs;
    2. Set \((\alpha_i, \beta_i)^{(m)} = (\alpha_i, \beta_i)^*\) if \(U \sim Unif(0, 1) < \rho_i = \min \left\{ \frac{S_i((\alpha_i, \beta_i)^*)}{S_i((\alpha_i, \beta_i)^{(m-1)})}, 1 \right\}; \)
       \((\alpha_i, \beta_i)^{(m)} = (\alpha_i, \beta_i)^{(m-1)}\) otherwise
  end
end

<table>
<thead>
<tr>
<th></th>
<th>PostMean</th>
<th>2.5%Quantile</th>
<th>97.5%Quantile</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha_1)</td>
<td>-3.2531</td>
<td>-3.5731</td>
<td>-2.9357</td>
</tr>
<tr>
<td>(\beta_1)</td>
<td>0.0064</td>
<td>0.0055</td>
<td>0.0072</td>
</tr>
<tr>
<td>(\alpha_2)</td>
<td>-5.7394</td>
<td>-6.4492</td>
<td>-5.1026</td>
</tr>
<tr>
<td>(\beta_2)</td>
<td>0.0175</td>
<td>0.0154</td>
<td>0.0203</td>
</tr>
</tbody>
</table>

Table 2: Posterior summary for \((\alpha_1, \alpha_2, \beta_1, \beta_2)\) in Problem 1.

that a fetus was dead rather than alive multiplies by 1.896 for every 100-unit increase
in the concentration. Given that a fetus was not dead, the estimated odds that it was
malformed was 5.754 times higher than the odds that it is normal for every 100-unit
increase in the concentration.

Figure 3 shows point and 95% interval estimates for the response curves \(\hat{\pi}_j(x), j = 1, 2, 3\). For the Dead category, the interval is getting wider as the concentration level
increases. The Malformation category also has wider interval when the concentration
level is high. The interval for Normal category is also wider when the concentration
level is higher, but becomes narrower when the concentration goes above 400.

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 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 \(i = 1, ..., n\).) Discuss your prior choice and approach
to MCMC posterior simulation.

Solution:
Using “fish” as the baseline category, we have two logistic regression

\[
\log \left( \frac{\pi_I}{\pi_F} \right) = \alpha_1 + \beta_1 x, \\
\log \left( \frac{\pi_O}{\pi_F} \right) = \alpha_2 + \beta_2 x,
\]

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

\[
\pi_1(x) = \pi_I(x) = \frac{\exp(\alpha_1 + \beta_1 x)}{1 + \sum_{l=1}^{2} \exp(\alpha_l + \beta_l x)} \\
\pi_2(x) = \pi_O(x) = \frac{\exp(\alpha_2 + \beta_2 x)}{1 + \sum_{l=1}^{2} \exp(\alpha_l + \beta_l x)} \\
\pi_3(x) = \pi_F(x) = \frac{1}{1 + \sum_{l=1}^{2} \exp(\alpha_l + \beta_l x)}
\]

With prior distribution \( p(\alpha_1, \alpha_2, \beta_1, \beta_2 \mid \text{data}) \), the posterior distribution is

\[
p(\alpha_1, \alpha_2, \beta_1, \beta_2 \mid \text{data}) \\
\propto p(\alpha_1, \alpha_2, \beta_1, \beta_2) \prod_{i=1}^{63} \text{Multi}(y_{iI}, y_{iO}, y_{iF} \mid \pi_I(x_i), \pi_O(x_i), \pi_F(x_i)) \\
= p(\alpha_1, \alpha_2, \beta_1, \beta_2) \prod_{i=1}^{63} \left\{ \left( \frac{\exp(\alpha_1 + \beta_1 x_i)}{1 + \sum_{l=1}^{2} \exp(\alpha_l + \beta_l x_i)} \right)^{y_{iI}} \left( \frac{\exp(\alpha_2 + \beta_2 x_i)}{1 + \sum_{l=1}^{2} \exp(\alpha_l + \beta_l x_i)} \right)^{y_{iO}} \left( \frac{1}{1 + \sum_{l=1}^{2} \exp(\alpha_l + \beta_l x_i)} \right)^{y_{iF}} \right\} \\
:= S(\alpha_1, \alpha_2, \beta_1, \beta_2)
\]

\( y_{iI}, y_{iO}, \) and \( y_{iF} \) is either 0 or 1. Unlike problem 1, here we cannot fit the model by fitting separately two binomial GLMs because each \( \pi_j(x), j = I, O, F \) is a function of all parameters \( \alpha_1, \alpha_2, \beta_1 \) and \( \beta_2 \). We will have one MCMC algorithm that simulates posterior samples of the all four parameters.

Flat prior \( p(\alpha_1, \alpha_2, \beta_1, \beta_2) \propto 1 \) is used as I have no prior information about how alligators choose their food. The Metropolis-Hastings algorithm for he inference on \( (\alpha_1, \alpha_2, \beta_1, \beta_2) \) is shown in Algorithm 2.

The algorithm starts at initial values \( (\alpha_1, \alpha_2, \beta_1, \beta_2) = (0, 0, 0, 0) \). The total number of iterations is 1100000. 100000 draws were burn out and the chain was thinned by keeping every 200th draw to have final 5000 draws to be analyzed. The tuning parameter \( d \) is chosen so that the acceptance rate is around 38%.

**Algorithm 2:** Metropolis-Hastings Algorithm for multinomial GLM using baseline-category logits

for \( \text{iter} = 1, 2, \ldots \), until convergence do

1. Draw \( (\alpha_1, \alpha_2, \beta_1, \beta_2)^* \sim N_4( (\alpha_1, \alpha_2, \beta_1, \beta_2)^{(m-1)}; dJ^{-1}(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}_1, \hat{\beta}_2)) \), where \( J^{-1} \) is the Fisher information, \( d \) is the tuning parameter and \( \hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}_1, \hat{\beta}_2 \) are the MLEs;
2. Set \( (\alpha_1, \alpha_2, \beta_1, \beta_2)^{(m)} = (\alpha_1, \alpha_2, \beta_1, \beta_2)^* \) if \( U \sim Unif(0, 1) < \rho = \min \left\{ \frac{S((\alpha_1, \alpha_2, \beta_1, \beta_2)^*)}{S((\alpha_1, \alpha_2, \beta_1, \beta_2)^{(m-1)})}, 1 \right\}; \)

end
Figure 4: Convergence of parameters in Problem 2(a).

Figure 5: Point and interval estimates for the estimated response curves $\hat{\pi}_j(x), j = I, O, F$ in Problem 2(a). Solid lines are for posterior means. Shaded areas are 95% interval estimates. Category Invertebrate is colored red, Other colored blue and Fish colored darkgreen.
Table 3: Posterior summary for \((\alpha_1, \alpha_2, \beta_1, \beta_2)\) in Problem 2(a).

<table>
<thead>
<tr>
<th></th>
<th>PostMean</th>
<th>2.5% Quantile</th>
<th>97.5% Quantile</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha_1)</td>
<td>4.6451</td>
<td>1.7974</td>
<td>8.0811</td>
</tr>
<tr>
<td>(\alpha_2)</td>
<td>-0.9924</td>
<td>-3.4283</td>
<td>1.3547</td>
</tr>
<tr>
<td>(\beta_1)</td>
<td>-2.7500</td>
<td>-4.7126</td>
<td>-1.2109</td>
</tr>
<tr>
<td>(\beta_2)</td>
<td>-0.1180</td>
<td>-1.1630</td>
<td>0.8565</td>
</tr>
</tbody>
</table>

Figure 4 shows a good convergent behavior of the parameters. Point and interval estimates of the four parameters can be found in Table 3. The parameters have large variation, but \(\beta_1 < 0\) with high probability, meaning that bigger alligators tend to eat fish instead of invertebrates. Although \(\beta_2 < 0\), it is hard to say that the odds of eat other food against eating fish is increasing or decreasing as the length of alligators increases.

Figure 5 shows the point and interval estimates of response probabilities. There is an obvious trade-off between invertebrates and fish as the length increases. The probability of eating other food slightly goes up as the length increases from 1 to 2.5 meters, and keep constant when the length is over 2.5 meters. The intervals for the three categories are quite large. The width of intervals for “Other” and “Fish” keep increasing for the length over 2.5 meters, showing a lots of uncertainty about the choice of food 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:

When taking gender into account, our two logistic regressions become

\[
\log \left( \frac{\pi_I}{\pi_F} \right) = \alpha_1 + \beta_1 x + \gamma_1 z, \\
\log \left( \frac{\pi_O}{\pi_F} \right) = \alpha_2 + \beta_2 x + \gamma_2 z,
\]

where \(z\) is the gender dummy variable with \(z = 1\) for male and \(z = 0\) for female.

Now we have totally six parameters \(\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \) and \(\gamma_2\). The response probabilities we’d like to estimate become

\[
\pi_1(x) = \pi_I(x) = \frac{\exp(\alpha_1 + \beta_1 x + \gamma_1 z)}{1 + \sum_{l=1}^{2} \exp(\alpha_l + \beta_l x + \gamma_l z)}, \\
\pi_2(x) = \pi_O(x) = \frac{\exp(\alpha_2 + \beta_2 x + \gamma_2 z)}{1 + \sum_{l=1}^{2} \exp(\alpha_l + \beta_l x + \gamma_l z)}, \\
\pi_3(x) = \pi_F(x) = \frac{1}{1 + \sum_{l=1}^{2} \exp(\alpha_l + \beta_l x + \gamma_l z)}.
\]

With the flat prior and the same algorithm as in (a), the convergence result of the six parameters is shown in Figure 6. Point and interval estimates for the parameters are shown in Table 4. As in (a), these parameters are of lots of variation. Figure 7 provides point and interval estimates for the response probabilities for male (left)
Figure 6: Convergence of parameters in Problem 2(b).
Table 4: Posterior summary for \((\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2)\) in Problem 2(b).

<table>
<thead>
<tr>
<th></th>
<th>PostMean</th>
<th>2.5%Quantile</th>
<th>97.5%Quantile</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha_1)</td>
<td>4.6451</td>
<td>1.7974</td>
<td>8.0811</td>
</tr>
<tr>
<td>(\alpha_2)</td>
<td>-0.9924</td>
<td>-3.4283</td>
<td>1.3547</td>
</tr>
<tr>
<td>(\beta_1)</td>
<td>-2.7500</td>
<td>-4.7126</td>
<td>-1.2109</td>
</tr>
<tr>
<td>(\beta_2)</td>
<td>-0.1180</td>
<td>-1.1630</td>
<td>0.8565</td>
</tr>
</tbody>
</table>

Estimated response curves (Male)  Estimated response curves (Female)

Figure 7: Point and interval estimates for the estimated response curves \(\hat{\pi}_j(x), j = I, O, F\) for male and female alligators in Problem 2(b). Solid lines are for posterior means. Shaded areas are 95% interval estimates. Category Invertebrate is colored red, Other colored blue and Fish colored darkgreen.

and female (right) alligators. One can see that male alligators tend to eat fish instead of invertebrates or other food, especially for male alligators with short length. The intervals for female alligators are wider than the intervals for male alligators due to smaller sample sizes.

3. Consider the inverse Gaussian distribution with density function

\[
f(y \mid \mu, \phi) = (2\pi\phi y^3)^{-1/2} \exp \left\{ -\frac{(y - \mu)^2}{2\mu^2 y} \right\}, \quad y > 0; \quad \mu > 0, \phi > 0.
\]

Denote the inverse Gaussian distribution with parameters \(\mu, \phi\) by \(IG(\mu, \phi)\).

(a) Show that the inverse Gaussian distribution is a member of the exponential dispersion family. Show that \(\mu\) is the mean of the distribution and obtain the variance function.

Solution:
Since
\[
f(y \mid \mu, \phi) = (2\pi \phi y^3)^{-1/2} \exp \left\{ -\frac{(y - \mu)^2}{2\phi \mu^2 y} \right\} \\
= (2\pi \phi y^3)^{-1/2} \exp \left\{ -\frac{1}{2} \frac{y}{\phi \mu^2} + \frac{1}{\phi \mu} - \frac{1}{2\phi} \right\} \\
= \exp \left\{ \frac{y \left(\frac{-1}{2\mu^2}\right)}{\phi} - \left(\frac{-1}{\mu}\right) + \left(\frac{-1}{\phi y} - \frac{1}{2} \log(2\pi \phi y^3)\right) \right\},
\]

IG(\mu, \phi) is a member of the exponential dispersion family with \( \theta = -\frac{1}{2\mu^2} \), \( b(\theta) = -\frac{1}{\mu} = -(2\theta)^{1/2} \), \( a(\phi) = \phi \), and \( c(y, \phi) = -\frac{1}{2\phi y} - \frac{1}{2} \log(2\pi \phi y^3) \).

The mean of the distribution is
\[
E(Y \mid \theta, \phi) = b'(\theta) = \frac{1}{2} (-2\theta)^{1/2} (-2) = -(2\theta)^{-1/2} = \mu
\]
from the fact that \( \theta = -\frac{1}{2\mu^2} \). The variance function is
\[
V(\mu) = b''(\theta) = -\frac{1}{2} (-2\theta)^{-3/2} (-2) = \left(-2 \left(\frac{-1}{2\mu^2}\right)\right)^{-3/2} = \mu^3,
\]
and so \( \text{Var}(Y \mid \theta, \phi) = a(\phi)b''(\theta) = \phi \mu^3 \).

(b) Consider a GLM with random component defined by the inverse Gaussian distribution. That is, assume that \( y_i \) are realizations of independent random variables \( Y_i \) with IG(\( \mu_i, \phi \)) distributions, for \( i = 1, \ldots, n \). Here \( g(\mu_i) = x_i^T \beta \), where \( \beta = (\beta_1, \ldots, \beta_p) \) \((p < n)\) is the vector of coefficients, and \( x_i = (x_{i1}, \ldots, x_{ip})^T \) is the covariate vector for the \( i \)th response, \( i = 1, \ldots, n \). Define the full model so that the \( y_i \) are realizations of independent IG(\( \mu, \phi \)) distributed random variables \( Y_i \), with distinct \( \mu_i \) for each \( y_i \). Obtain the scaled deviance for the comparison of the full model with the inverse Gaussian GLM.

Solution:

For the full model, \( \bar{\mu}_i = y_i \), and so \( \bar{\theta}_i = -\frac{1}{2\mu_i^2} = -\frac{1}{2\bar{\mu}_i^2} \). For the inverse Gaussian GLM, \( \hat{\mu}_i = g^{-1}(x_i^T \beta) \), and \( \hat{\theta}_i = -\frac{1}{2\hat{\mu}_i^2} \).

\( b(\bar{\theta}_i) = -1/\bar{\mu}_i = -1/y_i = -(2\bar{\theta}_i)^{1/2} \) and \( b(\hat{\theta}_i) = -1/\hat{\mu}_i = -(2\hat{\theta}_i)^{1/2} \).

Hence, the scaled deviance is
\[
D^* = \frac{2}{\phi} \sum_{i=1}^{n} \left\{ y_i (\bar{\theta}_i - \hat{\theta}_i) - b(\bar{\theta}_i) + b(\hat{\theta}_i) \right\} \\
= \frac{2}{\phi} \sum_{i=1}^{n} \left\{ y_i \left(-\frac{1}{2\mu_i^2} + \frac{1}{2\bar{\mu}_i^2}\right) - \left(-\frac{1}{\mu_i}\right) + \left(-\frac{1}{\bar{\mu}_i}\right) \right\} \\
= \frac{2}{\phi} \sum_{i=1}^{n} \left\{ \frac{1}{2 y_i} + \frac{1}{2\bar{\mu}_i^2 y_i} - \frac{1}{\mu_i} \right\} \\
= \frac{1}{\phi} \sum_{i=1}^{n} \left\{ \frac{\bar{\mu}_i^2 + y_i^2 - 2\bar{\mu}_i y_i}{\bar{\mu}_i^2 y_i} \right\} \\
= \frac{1}{\phi} \sum_{i=1}^{n} \left\{ \frac{(y_i - \bar{\mu}_i)^2}{\bar{\mu}_i^2 y_i} \right\}
\]