1 Exercises: quasilikelihood

1.1 Exercise

Men and women in a particular sample were asked whether or not they believe in the afterlife.

<table>
<thead>
<tr>
<th>Yes</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>435</td>
<td>375</td>
</tr>
<tr>
<td>No or Undecided</td>
<td>147</td>
<td>134</td>
</tr>
</tbody>
</table>

1. Apply Lindsey’s to these \((X_B, X_G) \in \{Y, N\} \times \{M, F\}\) valued random variables and fit a Poisson model to this data under the null hypothesis that \(X_B\) is independent of \(X_G\).

2. Compare the residual deviance to the usual Pearson’s \(\chi^2\) test of independence.

1.2 Exercise

```r
bernoulli_cases = as.numeric(bernoulli_cases)
expanded_rain = as.numeric(expanded_rain)
city = as.factor(city)
expanded.glm = glm(bernoulli_cases ~ poly(expanded_rain,3), family=binomial())
print(summary(expanded.glm))
```

Why does the deviance decrease for adding `city` as a factor equal the residual deviance from the original model?
1.3 Exercise
Suppose that, marginally, \( Y_i \sim \text{Bernoulli}(\pi(X_i)) \), \( 1 \leq i \leq n \) but the joint distribution is unknown and
\[
\pi(x_i) = \frac{e^{x_i^T \beta}}{1 + e^{x_i^T \beta}}.
\]
Show that
\[
\mathbb{E}(\nabla \ell_{\text{quasi}}(\beta)) = 0.
\]

1.4 Exercise
What is the domain of the exponential family \( Q_{\zeta, \theta} \)?

1.5 Exercise
1. Plot the quasidensity \( f_{\zeta, \theta, n}(\bar{y}) \) for the Poisson family with \( n = 10, \zeta = 1 \) for various values of \( \theta \).
2. Vary \( n \) over some range. Is the justification \( C(\zeta, \theta, n) \approx 1 \) reasonable?

1.6 Exercise
1. Compute the Fisher information of \((\eta, \theta) = (\zeta/\theta, \theta)\) in the double exponential family. (You can ignore the normalization constant).
2. Argue that this implies that \((\hat{\eta}, \hat{\theta})\) are asymptotically independent as \( n \to \infty \).

1.7 Exercise
1. Compute the score equations of the above linear model.
2. Show that \( \hat{\beta} \) is exactly the same as the quasilikelihood estimator of \( \beta \).
3. Show that \( \hat{\theta} \) is almost the same as the quasilikelihood estimator of \( \theta \). What’s different?
4. Compute the Fisher information of \((\beta, \theta)\) in this linear model. What advantage does this model have over the quasilikelihood model? (Hint: can you compute the Fisher information of the pair \((\beta, \theta)\) in the quasilikelihood setup?)

1.8 Exercise
1. Show the linear model we’ve formed for repeated measurements is not an exponential family (in the sense we’ve been using the term exponential family)?
2. Modify the model so that it is an exponential family.
1.9 Exercise

In the linear model we introduced above we made the assumption that $\theta_i = \theta$, but we might want to model $\theta$ as a function of $x$ as well. Suppose that $\eta_i = x_i^T \alpha$.

1. Form a regression model for $Y_i \mid X$ that allows both $(\zeta, \theta)$ to depend on some covariates. Choose your model so that it is a genuine exponential family.

2. Write out the score and Fisher information for this model.

3. Fit this model to the toxoplasmosis data where the effect on $\zeta$ is assumed to be cubic while the effect on $\theta$ is assumed to be linear (plus a constant). (Don’t forget the constraint $\theta \geq 0$.)

4. This model is some sort of random effects model. Suppose that the original family we started with was the normal means model (i.e. reference measure $e^{-x^2/2}$ on $\mathbb{R}$ and sufficient statistic $x$). Describe this model of the variance for the random means model.

1.10 Exercise

Suppose the original family we start with is Bernoulli($\pi$).

1. What is the quasidensity $f_{\zeta, \theta, \tilde{n}}$ for $\bar{y}$? (You can ignore the normalizing constant.)

2. Make a plot of the quasidensity for $n = 16, \pi = 0.4, \theta = 0.5$.

3. Compare this quasidensity to the density of $\bar{y}$ if we assume the reduced sample size $\tilde{n} = n\theta = 8$. 