Chapter 5 Part 1

05/06/2021

Chapter Packages

```
# Packages required for Chapter 3
library(knitr)
```

Learning Objectives

- Determine if a probability distribution can be expressed in one-parameter exponential family form.
- Identify canonical links for distributions of one-parameter exponential family form.

One-Parameter Exponential Families

- Thus far, we have expanded our repertoire of models from linear least squares regression to include Poisson regression.
- In the early 1970s, @Nelder1972 identified a broader class of models that generalizes the multiple linear regression we considered in the introductory chapter and are referred to as **generalized linear models (GLMs)**.

On Class to Rule them All

- All GLMs have similar forms for their likelihoods, MLEs, and variances.
- This makes it easier to find model estimates and their corresponding uncertainty.
- To determine whether a model based on a single parameter θ is a GLM, we consider the following properties.

GLM Properties

- When a probability formula can be written in the form below

$$f(y;\theta) = e^{[a(y)b(\theta)+c(\theta)+d(y)]} \tag{1}$$

- if the support (the set of possible input values) does not depend upon θ,
- Then is is said to have a **one-parameter exponential family form**.
- We will see that the Poisson distribution is a member of the one-parameter exponential family by writing its probability mass function (pmf) in the form of the above equation and assessing its support.

One-Parameter Exponential Family: Poisson

- Recall we begin with

$$P(Y = y) = \frac{e^{-\lambda}\lambda^y}{y!} \quad \text{where} \quad y = 0, 1, 2 \ldots \infty$$

- and consider the following useful identities for establishing exponential form:

$$a = e^{\log(a)}$$

$$a^x = e^{x \log(a)}$$

$$\log(ab) = \log(a) + \log(b)$$

$$\log\left(\frac{a}{b}\right) = \log(a) - \log(b)$$

One-Parameter Exponential Family: Poisson

- Determining whether the Poisson model is a member of the one-parameter exponential family is a matter of
  - writing the Poisson pmf in the form of the format of Equation 1
  - checking that the support does not depend upon λ.
- First, consider the condition concerning the support of the distribution.
- The set of possible values for any Poisson random variable is $y = 0, 1, 2 \ldots \infty$ which does not depend on λ.

One-Parameter Exponential Family: Poisson

- The support condition is met. Now we see if we can rewrite the probability mass function in one-parameter exponential family form.

$$
\begin{aligned}
P(Y = y) &= e^{-\lambda} e^{y \log \lambda} e^{-\log(y!)} \\
&= e^{y \log \lambda - \lambda - \log(y!)}
\end{aligned}
$$

- The first term in the exponent for equation 1 must be the product of two factors, one solely a function of y, $a(y)$, and another, $b(\lambda)$, a function of λ only.
- The middle term in the exponent must be a function of λ only; no $y's$ should appear.
- The last term has only $y's$ and no λ.

One-Parameter Exponential Family: Poisson

- We can identify the different functions in this form:

$$
\begin{aligned}
a(y) &= y \\
b(\lambda) &= \log(\lambda) \\
c(\lambda) &= -\lambda \\
d(y) &= -\log(y!)
\end{aligned}
$$

Canonical Links

- These functions have useful interpretations in statistical theory.
- We won't be going into this in detail, but we will note that function $b(\lambda)$, or more generally $b(\theta)$, will be particularly helpful in GLMs.
- The function $b(\theta)$ is referred to as the **canonical link**.
    - The canonical link is often a good choice to model as a linear function of the explanatory variables. That is, Poisson regression should be set up as $\log(\lambda) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots$.
- There is a distinct advantage to modeling the canonical link as opposed to other functions of θ, but it is also worth noting that other choices are possible, and at times preferred, depending upon the context of the application.

More Benifits

- By creating a unifying theory for regression modeling, Nelder and Wedderburn made possible a common and efficient method for finding estimates of model parameters using iteratively reweighted least squares (IWLS).
- In addition, we can use the one-parameter exponential family form to determine the expected value and standard deviation of Y. With statistical theory you can show that

$$\mathsf{E}(Y) = -\frac{c'(\theta)}{b'(\theta)} \quad \text{and} \quad \mathsf{Var}(Y) = \frac{b''(\theta)c'(\theta) - c''(\theta)b'(\theta)}{[b'(\theta)]^3}$$

where differentiation is with respect to θ. Verifying these results for the Poisson response:

$$\mathsf{E}(Y) = -\frac{-1}{1/\lambda} = \lambda \quad \text{and} \quad \mathsf{Var}(Y) = \frac{1/\lambda^2}{(1/\lambda^3)} = \lambda$$

A BIG Family

- We'll find that other distributions are members of the one-parameter exponential family by writing their pdf or pmf in this manner and verifying the support condition.
- For example, we'll see that the binomial distribution meets these conditions, so it is also a member of the one-parameter exponential family.
- The normal distribution is a special case where we have two parameters, a mean μ and standard deviation σ. If we assume, however, that one of the parameters is known, then we can show that a normal random variable is also from a one-parameter exponential family.

One-Parameter Exponential Family: Normal

- Here we determine whether a normal distribution is a one-parameter exponential family member.
- First we will need to assume that σ is known.
- Next, possible values for a normal random variable range from $-\infty$ to ∞, so the support does not depend on μ.

One-Parameter Exponential Family: Normal

- We need to write the probability density function (pdf) in the one-parameter exponential family form. We start with the pdf of the normal:

$$f(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2/(2\sigma^2)}$$

Even writing $1/\sqrt{2\pi\sigma^2}$ as $e^{-\log\sigma - \log(2\pi)/2}$ we still do not have the pdf written in one-parameter exponential family form. We will first need to expand the exponent so that we have

$$f(y) = e^{[-\log\sigma - \log(2\pi)/2]} e^{[-(y^2 - 2y\mu + \mu^2)/(2\sigma^2)]}$$

One-Parameter Exponential Family: Normal

- Without loss of generality, we can assume $\sigma = 1$, so that

$$f(y) \propto e^{y\mu - \frac{1}{2}\mu^2 - \frac{1}{2}y^2}$$

and $a(y) = y$, $b(\mu) = \mu$, $c(\mu) = -\frac{1}{2}\mu^2$, and $d(y) = -\frac{1}{2}y^2$.

- From this result, we can see that the canonical link for a normal response is μ
- This is consistent with what we've been doing with LLSR, since the simple linear regression model has the form:

$$\mu_{Y|X} = \beta_0 + \beta_1 X.$$

Generalized Linear Modeling

- GLM theory suggests that the canonical link can be modeled as a linear combination of the explanatory variable(s).
- This approach unifies a number of modeling results used throughout the text. For example, likelihoods can be used to compare models in the same way for any member of the one-parameter exponential family.(deviance)

Generalized Linear Modeling

- We have now **generalized** our modeling to handle non-normal responses.
- In addition to normally distributed responses, we are able to handle Poisson responses, binomial responses, and more.
- Writing a pmf or pdf for a response in one-parameter exponential family form reveals the canonical link which can be modeled as a linear function of the predictors.
- This linear function of the predictors is the last piece of the puzzle for performing generalized linear modeling. But, in fact, it is really nothing new.
- We already use linear combinations and the canonical link when modeling normally distributed data.

Three components of a GLM

1. Distribution of Y (e.g., Poisson)
2. Link Function (a function of the parameter, e.g., $\log(\lambda)$ for Poisson)
3. Linear Predictor (choice of predictors, e.g., $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots$)

All our Regression Methods

Table 1: One-parameter exponential family form and canonical links.

Distribution	One-Parameter Exponential Family Form	Canonical Link
Binary		
Binomial		$\text{logit}(p)$
Poisson	$P(Y=y) = e^{y \log \lambda - \lambda - y!}$	$\log(\lambda)$
Normal	$f(y) \propto e^{y\mu - \frac{1}{2}\mu^2 - \frac{1}{2}y^2}$	μ
Exponential		
Gamma		
Geometric		

Time Permitting

- We will work through some of these other ones in class.
- Since E(X) is not in our content, let us just work on showing each is in the GLM family and identifying the canonical link.
- Exercises 1a,e,f,i
- When using GLMS for modeling, our linearity assumption is verified by checking if there is linearity between our continuous predictors and the canonical link.