MA478 - Lesson 4

Clark

Recall that a Generalized Linear Model consists of three components:

Today, we're going to talk about what distributions are allowed for the random (or stochastic) component of a GLM. In our study of Statistics, we've inevitably come across a ton of different distributions:

However, not every distribution we come across can be used for the distribution of our random component for our GLMs. In fact, we're still limited. Our random component of a GLM must come from the *exponential dispersion family*. That is, we must be able to put the distribution in to the following form:

$$f(y_i|\theta_i, \phi) = \exp\{[y_i\theta_i - b(\theta_i)]/a(\phi) + c(y_i, \phi)\}$$

While this does seem rather strict, many of our common distributions can be put into this framework.

Let's start with the normal distribution:

Let's do one more. Let's assume $y_i \sim Binom(n, p_i)$

Now, your turn. Let's take $y_i \sim Pois(\lambda_i)$. As a reminder, the typical Poisson is written as:

$$f(y_i|\lambda_i) = \frac{\exp(-y_i)y_i^{\lambda_i}}{y_i!}$$

So, why? Why can't we use distributions like the Weibull, or the Beta distribution? Well, as we will see in future classes, the structure of the model allows us to create a general formula for fitting the model which allows functions in R like `glm` to exist.

There's a few other really nice properties of exponential dispersion models. While proving this is beyond the scope of what we are doing here, it's nothing more than a calculous excursion to show:

However, this leads to something rather important for GLMs:

Why is this important? Well, it allows us to come up with a sort of natural link function to use. If we recall, we want to link a linear predictor to $\mu_i = E(y_i)$. The parameter θ_i, as we can see, links to μ_i through:

One reason this is useful is it allows us to find $E(y_i)$ without integrating. Similarly we can find $var(y_i)$ again just through differentiation.

But, perhaps more importantly, it gives us a way to pick a link function. If we allow our linear predictor to put structure directly on the natural parameter, our link function must be $b'^{-1}(\mu_i)$

Taking a look at our three examples above, what would the link functions be?

Homework

Quiz