What are GLMs?

The Three Components of a GLM

- GLMs extend the ideas of **linear regression** to all types of outcomes!

The Three Components of a GLM

- GLMs extend the ideas of **linear regression** to all types of outcomes!
- There are **three main components** of a GLM:

The Three Components of a GLM

- GLMs extend the ideas of **linear regression** to all types of outcomes!
- There are **three main components** of a GLM:
  1. A random component, where $Y_i \sim f(y; \theta)$ is **exponential family**.

The Three Components of a GLM

- GLMs extend the ideas of **linear regression** to all types of outcomes!
- There are **three main components** of a GLM:
  1. A random component, where $Y_i \sim f(y;\theta)$ is **exponential family**.
  2. The linear predictor, $\eta_i = X_i\beta$, based on the **variates of interest**.

The Three Components of a GLM

- GLMs extend the ideas of **linear regression** to all types of outcomes!
- There are **three main components** of a GLM:
  1. A random component, where $Y_i \sim f(y; \theta)$ is **exponential family**.
  2. The linear predictor, $\eta_i = X_i\beta$, based on the **variates of interest**.
  3. A **link function**, $g(\cdot)$, such that $g(E[Y_i|X_i]) = \eta_i$.

The Three Components of a GLM

- GLMs extend the ideas of **linear regression** to all types of outcomes!
- There are **three main components** of a GLM:
  1. A random component, where $Y_i \sim f(y;\theta)$ is **exponential family**.
  2. The linear predictor, $\eta_i = X_i\beta$, based on the **variates of interest**.
  3. A **link function**, $g(\cdot)$, such that $g(E[Y_i|X_i]) = \eta_i$.
- Can estimate the parameter values using **maximum likelihood estimation**, numerically through *Fisher-Scoring*.

Exponential Family

For a random variable, Y, we say that it follows an **exponential family** distribution if it has a density that can be expressed as

$$f(y;\theta,\phi) = \exp\left\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)\right\}.$$

We call θ the **canonical parameter** and ϕ the **scale** or **dispersion parameter**.

The exponential family provide nice **score functions**, as well as **expected information**. These results give us a relationship between the mean and the variance, specifically, $E[Y] = b'(\theta)$ and $\mathrm{var}(Y) = a(\phi)b''(\theta)$.

Common Exponential Families

Distribution	Link Function	Link Form
Normal	Identity	$g(\mu) = \mu$
Exponential	Inverse	$g(\mu) = \mu^{-1}$
Binomial	Logit	$g(\mu) = \log\left(\frac{\mu}{1-\mu}\right)$
Poisson	Log	$g(\mu) = \log(\mu)$

These are the **canonical link** functions for various exponential family distributions. In theory, *any* link function can be used, but these have the nice property that $g(\mu) = \theta$, where μ is the mean of the random variable.

The *primary* limitation of a GLM, as implemented through **maximum likelihood estimation**, is that it **requires** specification of the distribution for Y.

This is not *strictly* necessary!

Quasi-(log)likelihood Review

- Instead of specifying the **exponential family**, we take:

Quasi-(log)likelihood Review

- Instead of specifying the **exponential family**, we take:
  1. $E[Y_i|X_i] = g(\eta_i) = \mu_i$

Quasi-(log)likelihood Review

- Instead of specifying the **exponential family**, we take:
  1. $E[Y_i|X_i] = g(\eta_i) = \mu_i$
  2. $\text{var}(Y_i|X_i) = \phi V(\mu_i)$

Quasi-(log)likelihood Review

- Instead of specifying the **exponential family**, we take:
  1. $E[Y_i|X_i] = g(\eta_i) = \mu_i$
  2. $\text{var}(Y_i|X_i) = \phi V(\mu_i)$
- Then, we define $U(\mu_i; Y_i) = \frac{Y_i - \mu_i}{\phi V(\mu_i)}$ to be the **quasi-score** function.

Quasi-(log)likelihood Review

- Instead of specifying the **exponential family**, we take:
  1. $E[Y_i|X_i] = g(\eta_i) = \mu_i$
  2. $\text{var}(Y_i|X_i) = \phi V(\mu_i)$
- Then, we define $U(\mu_i; Y_i) = \frac{Y_i - \mu_i}{\phi V(\mu_i)}$ to be the **quasi-score** function.
- If we solve,

$$U(\beta) = \sum_{i=1}^{n} \frac{\partial \mu_i}{\partial \beta} U(\mu_i; Y_i) \overset{!}{=} 0,$$

  then this gives us a CAN estimator for β.

For a *correctly* assumed exponential family distribution, **quasi-likelihood** is exactly **likelihood**. However, we got to this point **without any** distributional assumptions!

Quasi-likelihood Properties

- Whenever μ_i is *correctly* specified, $\widehat{\beta}$ is consistent.

Quasi-likelihood Properties

- Whenever μ_i is *correctly* specified, $\widehat{\beta}$ is consistent.
- We can estimate the variance of $\widehat{\beta}$, **even if** $V(\mu_i)$ is incorrect!

Quasi-likelihood Properties

- Whenever μ_i is *correctly* specified, $\widehat{\beta}$ is consistent.
- We can estimate the variance of $\widehat{\beta}$, **even if** $V(\mu_i)$ is incorrect!
- The value of ϕ can be estimated using a modified method of moments approach.

Quasi-likelihood Properties

- Whenever μ_i is *correctly* specified, $\widehat{\beta}$ is consistent.
- We can estimate the variance of $\widehat{\beta}$, **even if** $V(\mu_i)$ is incorrect!
- The value of ϕ can be estimated using a modified method of moments approach.
- This will generally be *less* efficient than MLE, but it is **more robust!**

Summary

- GLMs extend the ideas of linear regression to **other types of outcome data**.

Summary

- GLMs extend the ideas of linear regression to **other types of outcome data**.
- GLMs can be fit using **MLE** by specifying a random (exponential family) distribution, a linear predictor, and a link function.

Summary

- GLMs extend the ideas of linear regression to **other types of outcome data**.
- GLMs can be fit using **MLE** by specifying a random (exponential family) distribution, a linear predictor, and a link function.
- GLMs can be fit **without** the need for a distributional assumption, by leveraging quasi-likelihood estimation.

Summary

- GLMs extend the ideas of linear regression to **other types of outcome data**.
- GLMs can be fit using **MLE** by specifying a random (exponential family) distribution, a linear predictor, and a link function.
- GLMs can be fit **without** the need for a distributional assumption, by leveraging quasi-likelihood estimation.
- **However**, like linear regression models, GLMs assume **IID data**. Oh no.