Generalized Linear Models

- Are the conglomeration and extension of familiar regression models
- The generalized Linear Models consists of three components:

a) Random component: Specifying the conditional distribution of the response variable. Originally Y_i was a member of exponential family like Gaussian, binomial, Poisson, gamma and inverse Gaussian distribution
Now the distribution extended to multivariate exponential (multinomial distribution) as well as non exponential (two parameter negative binomial distribution) and for distribution we do not specify the distribution

b) Linear Predictor: Consists of linear function of regressors

$$\eta_i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} \ldots\ldots + \beta_k X_{ik}$$

The X_{ij} are prespecified explanatory variables that can be quantitative, transformation of quantitative, dummy regressors, interactions etc.

c) The smooth and invertible linearizing function(.) which transforms the expectation of response variable $\mu_i = E(Y_i)$ to the Linear predictor

$$g(\mu_i) = \eta_i = \alpha + \beta_1 X_{i1} + \beta_2 X_{i2} \ldots\ldots + \beta_k X_{ik}$$

Since the link function is invertible we can write:

$$\mu_i = g^{-1}(\eta_i) = g^{-1}(\alpha + \beta_1 X_{i1} + \beta_2 X_{i2} \ldots\ldots + \beta_k X_{ik})$$

GLM can be defined as the Linear model for a transformation of the expected response variable or a non linear regression for the model.

The inverse link function is $g^{-1}(.)$ is called mean function.

The identity link simple returns the argument unaltered.

$\eta_i = g(\eta_i) = \mu_i$ and thus $\mu_i = g^{-1}(\eta_i) = \eta_i$

Some other link functions are as follows:

Table 15.1 Some Common Link Functions and Their Inverses

Link	$\eta_i = g(\mu_i)$	$\mu_i = g^{-1}(\eta_i)$
Identity	μ_i	η_i
Log	$\log_e \mu_i$	e^{η_i}
Inverse	μ_i^{-1}	η_i^{-1}
Inverse-square	μ_i^{-2}	$\eta_i^{-1/2}$
Square-root	$\sqrt{\mu_i}$	η_i^2
Logit	$\log_e \frac{\mu_i}{1-\mu_i}$	$\frac{1}{1+e^{-\eta_i}}$
Probit	$\Phi^{-1}(\mu_i)$	$\Phi(\eta_i)$
Log-log	$-\log_e[-\log_e(\mu_i)]$	$\exp[-\exp(-\eta_i)]$
Complementary log-log	$\log_e[-\log_e(1-\mu_i)]$	$1-\exp[-\exp(\eta_i)]$

NOTE: μ_i is the expected value of the response, η_i is the linear predictor, and $\Phi(\cdot)$ is the cumulative distribution function of the standard-normal distribution.

https://stat.ethz.ch/R-manual/R-devel/library/stats/html/glm.html

https://www.r-bloggers.com/generalised-linear-models-in-r/

http://www.statmethods.net/advstats/glm.html