Generalized Linear Models (GLMs) extend the traditional linear regression model to accommodate a wider range of response variables beyond continuous data. A key component of a GLM is the **link function**, which establishes a relationship between the linear predictor (a linear combination of predictor variables) and the expected value of the response variable.

Here's a list of common link functions and their purposes:

**1. Identity Link:**

* **Purpose:** Models the expected value of the response variable directly as the linear predictor.
* **Common Use:** Used for continuous response variables with a normal distribution, essentially making the GLM equivalent to ordinary least squares regression.
* **Mathematical Form:** g(μ) = μ

**2. Log Link:**

* **Purpose:** Models the logarithm of the expected value of the response variable. This ensures the predicted values are always positive.
* **Common Use:** Often used for count data (Poisson regression) or continuous positive data with right-skewed distributions.
* **Mathematical Form:** g(μ) = log(μ)

**3. Logit Link:**

* **Purpose:** Models the log-odds (logit) of the probability of an event. This constrains the predicted probabilities to lie between 0 and 1.
* **Common Use:** The standard link function for binary response variables (logistic regression).
* **Mathematical Form:** g(μ) = log(μ / (1 - μ))

**4. Probit Link:**

* **Purpose:** Similar to the logit link, it models the inverse of the cumulative distribution function (CDF) of the standard normal distribution.
* **Common Use:** Also used for binary response variables, often when there's an underlying latent variable with a normal distribution.
* **Mathematical Form:** g(μ) = Φ⁻¹(μ), where Φ⁻¹ is the inverse of the standard normal CDF.

**5. Complementary Log-Log Link:**

* **Purpose:** Models the log of the negative log of the complementary probability (1 - probability of event).
* **Common Use:** Suitable for binary response data, particularly when the probability of an event is very low or very high. Also used in survival analysis.
* **Mathematical Form:** g(μ) = log(-log(1 - μ))

**6. Inverse Link:**

* **Purpose:** Models the inverse of the expected value of the response variable.
* **Common Use:** Often used for response variables with a gamma distribution.
* **Mathematical Form:** g(μ) = 1 / μ

**7. Square Root Link:**

* **Purpose:** Models the square root of the expected value of the response variable.
* **Common Use:** Can be used for count data with a Poisson distribution when variance increases faster than the mean.
* **Mathematical Form:** g(μ) = √μ

**Choosing the Right Link Function:**

The choice of link function depends on:

* **The nature of the response variable:** Whether it's continuous, binary, count, etc.
* **The distribution of the response variable:** Normal, Poisson, binomial, gamma, etc.
* **The desired interpretation of the model coefficients:** How changes in the predictor variables affect the response variable.

It's important to note that the link function is a crucial part of the GLM specification. Choosing an inappropriate link function can lead to biased estimates and incorrect inferences.