**Generalized Estimating Equations Type of Model**

The Type of Model tab allows you to specify the distribution and link function for your model, providing shortcuts for several common models that are categorized by response type.

Model Types

**Scale Response.** The following options are available:

* **Linear.** Specifies Normal as the distribution and Identity as the link function.
* **Gamma with log link.** Specifies Gamma as the distribution and Log as the link function.

**Ordinal Response.** The following options are available:

* **Ordinal logistic.** Specifies Multinomial (ordinal) as the distribution and Cumulative logit as the link function.
* **Ordinal probit.** Specifies Multinomial (ordinal) as the distribution and Cumulative probit as the link function.

**Counts.** The following options are available:

* **Poisson loglinear.** Specifies Poisson as the distribution and Log as the link function.
* **Negative binomial with log link.** Specifies Negative binomial (with a value of 1 for the ancillary parameter) as the distribution and Log as the link function. To have the procedure estimate the value of the ancillary parameter, specify a custom model with Negative binomial distribution and select Estimate value in the Parameter group.

**Binary Response or Events/Trials Data.** The following options are available:

* **Binary logistic.** Specifies Binomial as the distribution and Logit as the link function.
* **Binary probit.** Specifies Binomial as the distribution and Probit as the link function.
* **Interval censored survival.** Specifies Binomial as the distribution and Complementary log-log as the link function.

**Mixture.** The following options are available:

* **Tweedie with log link.** Specifies Tweedie as the distribution and Log as the link function.
* **Tweedie with identity link.** Specifies Tweedie as the distribution and Identity as the link function.

**Custom.** Specify your own combination of distribution and link function.

Distribution

This selection specifies the distribution of the dependent variable. The ability to specify a non-normal distribution and non-identity link function is the essential improvement of the generalized linear model over the general linear model. There are many possible distribution-link function combinations, and several may be appropriate for any given dataset, so your choice can be guided by a priori theoretical considerations or which combination seems to fit best.

* **Binomial.** This distribution is appropriate only for variables that represent a binary response or number of events.
* **Gamma.** This distribution is appropriate for variables with positive scale values that are skewed toward larger positive values. If a data value is less than or equal to 0 or is missing, then the corresponding case is not used in the analysis.
* **Inverse Gaussian.** This distribution is appropriate for variables with positive scale values that are skewed toward larger positive values. If a data value is less than or equal to 0 or is missing, then the corresponding case is not used in the analysis.
* **Negative binomial.** This distribution can be thought of as the number of trials required to observe *k* successes and is appropriate for variables with non-negative integer values. If a data value is non-integer, less than 0, or missing, then the corresponding case is not used in the analysis. The value of the negative binomial distribution's ancillary parameter can be any number greater than or equal to 0; you can set it to a fixed value or allow it to be estimated by the procedure. When the ancillary parameter is set to 0, using this distribution is equivalent to using the Poisson distribution.
* **Normal.** This is appropriate for scale variables whose values take a symmetric, bell-shaped distribution about a central (mean) value. The dependent variable must be numeric.
* **Poisson.** This distribution can be thought of as the number of occurrences of an event of interest in a fixed period of time and is appropriate for variables with non-negative integer values. If a data value is non-integer, less than 0, or missing, then the corresponding case is not used in the analysis.
* **Tweedie.** This distribution is appropriate for variables that can be represented by Poisson mixtures of gamma distributions; the distribution is "mixed" in the sense that it combines properties of continuous (takes non-negative real values) and discrete distributions (positive probability mass at a single value, 0). The dependent variable must be numeric, with data values greater than or equal to zero. If a data value is less than zero or missing, then the corresponding case is not used in the analysis. The fixed value of the Tweedie distribution's parameter can be any number greater than one and less than two.
* **Multinomial.** This distribution is appropriate for variables that represent an ordinal response. The dependent variable can be numeric or string, and it must have at least two distinct valid data values.

Link Function

The link function is a transformation of the dependent variable that allows estimation of the model. The following functions are available:

* **Identity.** *f*(*x*)=*x*. The dependent variable is not transformed. This link can be used with any distribution.
* **Complementary log-log.** *f*(*x*)=log(−log(1−*x*)). This is appropriate only with the binomial distribution.
* **Cumulative Cauchit.** *f*(*x*) = tan(π (*x* – 0.5)), applied to the cumulative probability of each category of the response. This is appropriate only with the multinomial distribution.
* **Cumulative complementary log-log.** *f*(*x*)=ln(−ln(1−*x*)), applied to the cumulative probability of each category of the response. This is appropriate only with the multinomial distribution.
* **Cumulative logit.** *f*(*x*)=ln(*x* / (1−*x*)), applied to the cumulative probability of each category of the response. This is appropriate only with the multinomial distribution.
* **Cumulative negative log-log.** *f*(*x*)=−ln(−ln(*x*)), applied to the cumulative probability of each category of the response. This is appropriate only with the multinomial distribution.
* **Cumulative probit.** *f*(*x*)=Φ−1(*x*), applied to the cumulative probability of each category of the response, where Φ−1 is the inverse standard normal cumulative distribution function. This is appropriate only with the multinomial distribution.
* **Log.** *f*(*x*)=log(*x*). This link can be used with any distribution.
* **Log complement.** *f*(*x*)=log(1−*x*). This is appropriate only with the binomial distribution.
* **Logit.** *f*(*x*)=log(*x* / (1−*x*)). This is appropriate only with the binomial distribution.
* **Negative binomial.** *f*(*x*)=log(*x* / (*x*+*k* −1)), where *k* is the ancillary parameter of the negative binomial distribution. This is appropriate only with the negative binomial distribution.
* **Negative log-log.** *f*(*x*)=−log(−log(*x*)). This is appropriate only with the binomial distribution.
* **Odds power.** *f*(*x*)=[(*x*/(1−*x*))α−1]/α, if α ≠ 0. *f*(*x*)=log(*x*), if α=0. α is the required number specification and must be a real number. This is appropriate only with the binomial distribution.
* **Probit.** *f*(*x*)=Φ−1(*x*), where Φ−1 is the inverse standard normal cumulative distribution function. This is appropriate only with the binomial distribution.
* **Power.** *f*(*x*)=*x* α, if α ≠ 0. *f*(*x*)=log(*x*), if α=0. α is the required number specification and must be a real number. This link can be used with any distribution.