Print

Quiz, Lesson 5: Introduction to Generalized Linear Models

Your Score: 100%

Congratulations! Your score of 100% indicates that you've mastered the topics in this lesson. If you'd like, you can review the feedback for each question.

When you're ready to start the next lesson, exit this lesson and begin the next one.

1. While fitting the generalized linear models, on what basis would you determine the type of link function to use?

   a. type of response variable
   b. distribution of response variable
   c. values of mean and variance
   d. both *a* and *b*

   Your answer: d
   Correct answer: d

   In generalized linear models, you determine which link function to use based on the type of response variable and its distribution. For each distribution, there is a link function that can be determined algebraically from the distribution equation. This link is called the canonical link. However, you can choose to use other links.

---

2. The log link function is often applied to count data that have nonnegative integer values for which of the following reasons?

   a. to remove the lower bound and create a linear model
   b. to remove the upper bound and create a linear model
   c. to make interpretations on a log scale
   d. It is easier to apply in comparison to other transformations.

   Your answer: a
   Correct answer: a

   When the response variable is discrete and consists of nonnegative integer values, then you often apply the log link function. The log transformation removes the lower bound and creates a linear model.

---

3. General regression models for continuous responses are special cases of generalized linear models. Which link function would you use to model the mean of a continuous response variable?

   a. inverse link function
   b. identify link function
   c. log link function
   d. logit link function

   Your answer: b
   Correct answer: b

General regression models for continuous responses are special cases of generalized linear models. They assume a normal distribution for the response variable and model its mean directly, using the identity link function, g (μ) = μ. It has the form: $\mu = \beta 0 + \beta_1 X_1 + \beta_2 X_2 + ...+ \beta_k X_k$.

---

4. Below is the partial output from the GENMOD procedure used to fit a Poisson regression model. Which concern about the data might be indicated from the values in the Value/DF column?

Criteria for Assessing Goodness of Fit			
Criterion	**DF**	**Value**	**Value/DF**
Deviance	165	549.5856	3.3308
Scaled Deviance	165	549.5856	3.3308
Pearson Chi-Square	165	533.8165	3.2353
Scaled Pearson X2	165	533.8165	3.2353

a. overdispersion
b. unrelated mean and variance
c. negatively skewed data
d. underdispersion

Your answer: a
Correct answer: a

The Value/DF values are computed by dividing the goodness of fit statistics by the degrees of freedom. Values for scaled deviance or the scaled Pearson chi-square are useful for assessing the goodness of model fit, and values close to 1 indicate good model fit. However, the table indicates values of these statistics around 3, clearly not close to 1. Poisson regression models assume that the variance is equal to the mean. This might not be the case here. These values might indicate overdispersed data.

---

5. What makes overdispersion a cause of concern in Poisson regression?

a. It affects parameter estimates.
b. It affects standard error of the parameter estimates.
c. It affects the predicted value of response.
d. It affects fit statistics such as AIC and AICC.

Your answer: b
Correct answer: b

Overdispersion occurs when the observed variance is larger than the nominal variance for a particular distribution, which can occur frequently in Poisson regression. It does not affect the parameter estimates, but it causes the estimates of the standard error o the parameter estimates to be underestimated.

---

6. Assume for a single explanatory variable X that the Poisson loglinear model has the form Log(μ) = $\beta 0 + \beta_X$. How will you interpret the relationship between X and the mean μ if $\hat{\beta} < 0$?

a. The mean decreases as X increases.

b. The mean increases as X increases.
c. The mean remains the same for any change in X.
d. Nothing can be said.

Your answer: a
Correct answer: a

To see the effect of the change in the original scale, you must exponentiate the parameter estimates $\hat{\beta}$. The effect on the mean (μ) will be a multiplicative effect. The exponentiation of the parameter estimate yields a percent increase or decrease on the mean. Regardless of the value of the parameter estimate, the exponentiation will be positive. Hence, if $\hat{\beta} < 0$ then $e^{\hat{\beta}} < 1$. This means that the mean of Y decreases as X increases.

---

7. What are the possible ways to account for overdispersion when you are fitting the generalized linear model using the GENMOD procedure?

a. Use the DIST=NEGBIN option in the MODEL statement.
b. Fit a Poisson model and use the PSCALE or DSCALE option in the MODEL statement.
c. Overdispersion cannot be accounted for in the GENMOD procedure.
d. either *a* or *b*

Your answer: d
Correct answer: d

One way to deal with overdispersion is to introduce a more flexible distribution than the Poisson distribution in the model. A related distribution for count data that allows the variance to exceed the mean is the negative binomial distribution. Hence, the negative binomial distribution can be used to model the overdispersion by using the DIST=NEGBIN option in the MODEL statement in PROC GENMOD. Another way of accounting for overdispersion is to fit a Poisson model and apply a multiplicative adjustment factor to adjust the standard errors accordingly by using the PSCALE or DSCALE option in the MODEL statement in PROC GENMOD.

---

8. In a negative binomial distribution, the limiting case when dispersion parameter *k* is equal to zero corresponds to which of the following distributions?

a. Poisson
b. normal
c. binomial
d. gamma

Your answer: a
Correct answer: a

The dispersion parameter, *k*, enables the variance to exceed the mean and enables the negative binomial distribution to account for overdispersion. The dispersion parameter *k* is not allowed to vary over observations. The limiting case when the parameter *k* is equal to 0 corresponds to a Poisson regression model. The negative binomial distribution has variance function given by : $Var(Y) = \mu + k\mu^2$.

As k-> 0, Var(Y) -> μ, and the negative binomial distribution converges to the Poisson distribution.

---

9. Which SAS procedure, in addition to PROC GENMOD, can be used to fit a generalized linear

model for nonnormal responses?

a. PROC GLM
b. PROC MIXED
c. PROC GLIMMIX
d. PROC GLMSELECT

Your answer: c
Correct answer: c

In addition to PROC GENMOD, PROC GLIMMIX is a modeling procedure that is used to fit generalized models. You can use PROC GLIMMIX to fit generalized linear models for nonnormal responses such as logistic regression for binary responses, Poisson or negative binomial regression for counts, and gamma regression for continuous skewed, positive values.

10. Which of the following distributions has variance that is proportional to the squared mean?

a. gamma distribution
b. normal distribution
c. Poisson distribution
d. Negative binomial distribution

Your answer: a
Correct answer: a

Gamma distribution is used when the response variable has continuous positive values, is highly skewed to the right, and has variances that are proportional to the squared mean.

Close