8

POISSON REGRESSION MODELS

OBJECTIVES OF THIS CHAPTER

This chapter introduces Poisson regression models. It first starts with an introduction to the Poisson regression model followed by a discussion of the incidence rates and incidence rate ratios in the model, goodness-of-fit statistics, and how to interpret parameter estimates. After a description of the research example, the data, and the sample, a one-predictor Poisson regression model and a multiple-predictor Poisson regression model are illustrated with the `glm()` function in R. The `vglm()` function in the `VGAM` package is also used to fit the multiple-predictor model. R commands and output are explained in detail. This chapter focuses on fitting the Poisson regression models with R, as well as on interpreting and presenting the results. After reading this chapter, you should be able to:

- Identify when Poisson regression models are used.
- Fit a Poisson regression model using R.
- Interpret the output.
- Interpret the incidence rate ratios and marginal effects.
- Compute, plot, and interpret the predicted counts.
- Compare models using the likelihood ratio test.
- Present results in publication-quality tables using R.
- Write the results for publication.

8.1 POISSON REGRESSION MODELS: AN INTRODUCTION

The Poisson regression model is used to estimate a count response variable. The count response variable can be a count of events, or the number of events during a time period or in a location. For example, researchers are interested in the number of AP courses which students have taken in high school, the number of absences in a college class, the number of publications by new faculty in a year, the number of visits to a doctor in a year, and the number of hospitalized patients. These count response variables are nonnegative integers and follow a Poisson distribution.

The Poisson regression model (Cameron & Trivedi, 2013; Hilbe, 2014; Long & Freese, 2014) can be expressed as follows:

$$\ln(\mu) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p \quad (8.1)$$

where μ is the mean or the expected number of events; α is the intercept; and $\beta_1, \beta_2, \ldots, \beta_p$ are the Poisson coefficients for the predictors. The left side of the equation, $\ln(\mu)$, is the log link function. The right side of the equation is the linear predictor. An important assumption of the Poisson regression model is that the mean of the count response variable is equal to the variance of the variable. $E(Y) = \text{Variance }(Y) = \mu$. By exponentiating both sides of the equation, we get the predicted mean of the count response variable:

$$\mu = \exp\left(\alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p\right) \quad (8.2)$$

When a count response variable is the number of events during a time period or in a location, a count of events can also referred to as an incidence rate. If we define the incidence rate as the expected number of events per unit time or location, μ/t, the Poisson model can also be expressed as follows:

$$\ln(\mu/t) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p \quad (8.3)$$

where μ is the mean count; t is a period of time; μ/t is the incidence rate; and $\ln(\mu/t)$ is the log of the incidence rate or log incidence rate. Since $\ln(\mu/t) = \ln(\mu) - \ln(t)$, the equation can be rewritten as:

$$\ln(\mu) = \ln(t) + \alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p \quad (8.4)$$

where $\ln(t)$ is the offset in the model equation. When $t = 1$, the offset $\ln(t) = 0$, which can be omitted from the equation. For example, if the count response variable (Y) is the number of visits to a zoo in a year in the model, then the mean count (μ) is the average number of visits to a zoo, and the time period (t) is 1. Therefore, the expected number of visits to a zoo is the same as the incidence rate.

8.1.1 The Poisson Distribution

The Poisson probability distribution for a count response variable is expressed as:

$$P(y) = \frac{e^{-u} u^{y}}{y!} \tag{8.5}$$

where y is a count response variable, μ is the expected or the average number of events, and $y!$ is the y factorial or the factorial of y. $y! = y*(y-1) \ldots 2*1$. In literature, μ is often symbolized as λ. In the Poisson distribution, the mean of a count variable is equal to the variance of the variable.

$$\mu = \mathrm{E}(y) = \mathrm{Variance}(y).$$

The log likelihood function for the Poisson distribution or the Poisson log likelihood function is expressed as:

$$l(u;\ y) = \sum_{i=1}^{n} \{y_i \ln u_i - u_i - \ln(y_i!)\} \tag{8.6}$$

where $l(u; y)$ is the log likelihood function of μ given the values of the count variable y.

8.1.2 Incidence Rate Ratios in Poisson Regression Models

The Poisson regression model estimates the log expected counts of an event or the log incidence rate of the response variable. The incidence rate is defined as the expected number of events during a period of time or in a location. In a simple Poisson regression model with one predictor, $\ln(\mu) = \alpha + \beta X$, where μ is the expected counts of an event or the incidence rate. The estimated coefficient is the Poisson coefficient, which is the coefficient on the scale of the natural logarithm. It can be also referred to as the log coefficient. We estimate the relationship between the predictor variable and the log function of the expected counts of an event or the log incidence rate.

By exponentiating both sides of the equation, we get the expected counts of an event or the incidence rate:

$$\mu = \exp(\alpha + \beta X) \tag{8.7}$$

If the independent variable X is a categorical variable with the values of 0 and 1, the incidence rates of the response variable can be computed as follows.

When $X = 0$, the incidence rate $= \exp(\alpha)$, which is the exponentiated intercept.

When $X = 1$, the incidence rate $= \exp(\alpha + \beta)$, which is the exponentiated sum of the intercept and the Poisson coefficient.

The incidence rate ratio of the group 1 ($X = 1$) to the group 2 ($X = 0$):

$$\text{IRR} = \frac{\exp(\alpha + \beta)}{\exp(\alpha)} = \frac{\exp(\alpha) \times \exp(\beta)}{\exp(\alpha)} = \exp(\beta) \tag{8.8}$$

For a one-unit increase in an independent variable (e.g., from 0 to 1 in the previous example) the change in the incidence rate is the incidence rate ratio, which is the exponentiated Poisson coefficient. When the independent variable is continuous, for a one-unit increase from any value of x to the value of $(x + 1)$, the change in the incidence rate is still the exponentiated Poisson coefficient.

8.1.3 Model Fit Statistics

Same as those discussed in the previous chapters, model fit statistics, such as the log likelihood statistic, the residual deviance, the model chi-square statistic, the AIC and BIC statistics, and the pseudo R^2 statistics, can be computed for the Poisson regression model. The likelihood ratio test and the AIC and BIC statistics can also be used for model comparisons.

8.1.4 Interpretation of Model Parameter Estimates

When the Poisson coefficient is positive, it indicates the relationship between the predictor variable and the log expected counts of an event or the log incidence rate ratio is positive. By exponentiating the log coefficient, we get the incidence rate ratio, which is larger than 1. This means that the expected number of events or the incidence rate of a response variable increases for a one-unit increase in the predictor variable.

When the Poisson coefficient is negative, it indicates that the relationship between the predictor variable and the log expected counts of an event or the log incidence rate is negative. The exponentiated coefficient, the incidence rate ratio, is less than 1. It means that the expected number of events or the incidence rate of a response variable decreases for a one-unit increase in the predictor variable.

When the Poisson coefficient equals 0, the incidence rate ratio equals 1. It indicates that a one-unit increase in the predictor variable does not impact the expected number of events or the incidence rate of a response variable.

When there are multiple predictors in the model, the incidence rate ratio for a predictor can be interpreted as the change in the expected number of events or the incidence rate of a response variable for a one-unit change in a predictor variable when holding other predictor variables constant.

8.1.5 Interpreting an Incidence Rate Ratio as a Percentage Change in an Incidence Rate

Another way of interpreting an incidence rate ratio is the percentage change in an incidence rate. It can be calculated by using (Incidence rate − 1) × 100%. A positive percentage change in an incidence rate indicates there is an increase in the incidence rate, whereas a negative percentage change corresponds to a decrease in the rate. A zero percentage change indicates no change in the rate at all. In other words, the predictor variable does not influence the incidence rate of a response variable.

For example, if an incidence rate ratio for a predictor variable equals 1.7, then the percentage change in the incidence rate can be computed as follows: $(1.7 - 1) \times 100\% = 70\%$. This indicates that each one-unit increase in the predictor variable corresponds to an increase of 70% in the incidence rate of a response variable.

In another example, if an incidence rate ratio = .80, then the percentage change in the incidence rate is $(.80 - 1) \times 100\% = -20\%$. Since the percentage change is negative, it indicates that for each one-unit increase in the predictor variable, there is a decrease of 20% in the incidence rate.

8.1.6 Interpreting Marginal Effects as Changes in Predicted Counts

In Poisson regression, a marginal effect is a change in the expected counts of a response variable related to the change in an independent variable. Mathematically, it is the product of a Poisson coefficient for a particular predictor variable (β_p) and the predicted mean or the expected number of events (μ). Recall that in Equation 8.2 the predicted mean is the exponential of the linear predictor, $\exp(\alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p)$. So a marginal effect $= \beta_p \times \mu = \beta_p \times \exp(\alpha + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p)$. It varies across the values of a predictor variable. In other words, for each predictor variable, we can calculate the marginal effect at each value of that variable. An average marginal effect (AME) is the mean value of the marginal effects at these values.

8.2 RESEARCH EXAMPLE AND DESCRIPTION OF THE DATA AND SAMPLE

We will investigate the relationships between the count response variable, the number of zoo visits in a year, and four predictor variables. Unlike other chapters, however, here the research interest focuses on using Poisson regression to predict the count response variable. The GSS 2016 data are used for the following analyses. The following are the variables used for data analysis in this chapter:

- `vistzoo`: the recoded variable of the number of zoo visits in a year
- `maritals`: the recoded variable of marital (marital status) with 1 = currently married and 0 = not currently married
- `educ`: the highest education completed
- `female`: recoded variable of sex with 1 = female and 0 = male
- `wrkfull`: working full time or not

8.3 FITTING A ONE-PREDICTOR POISSON REGRESSION MODEL WITH R

8.3.1 Packages and Functions for Poisson Regression Models in R

Several packages in R can be used for fitting Poisson regression models. This chapter focuses on the `glm()` function in R and the `vglm()` function in the `VGAM` package (Yee, 2010). The `glm()` function can be used directly since it is included in the `stats` package with the installation of R. However, you need to load the `VGAM` package with `library(VGAM)` if it is installed. The `glm()` function is introduced first.

8.3.2 The `glm()` Function

As introduced in the earlier chapters, the `glm()` function is normally used to fit generalized linear models. The model formula in `glm()` specifies the dependent variable and the predictor variable(s), which are separated by the tilde (~). The plus (+) symbol is used to connect multiple predictor variables. We also need to specify the probability distribution of the outcome variable with the `family` argument. We specify `family = poisson` for a count outcome variable in Poisson regression. For more details on how to use this function, type `help(glm)` in the command prompt.

8.3.3 The Poisson Regression Model: One-Predictor Model

In the following example, the command `PR.1 <- glm(vistzoo ~ educ, family = poisson, data = count)` tells R to predict the count outcome variable `vistzoo` from the independent variable `educ` with Poisson regression by specifying the Poisson family (`family = poisson`). The fitted model is named `PR.1`. The `summary(PR.1)` command displays the output of the fitted model.

```
> # One-predictor Poisson regression model with glm()
> PR.1 <- glm(vistzoo ~ educ, family = poisson, data = count)
> summary(PR.1)

Call:
glm(formula = vistzoo ~ educ, family = poisson, data = count)

Deviance Residuals:
    Min        1Q     Median        3Q       Max
-1.5545   -1.2289   -1.0389    0.3291    4.1534

Coefficients:
             Estimate   Std. Error   z value    Pr(>|z|)
(Intercept)  -1.15403     0.18915    -6.101    1.05e-09 ***
educ          0.06716     0.01295     5.187    2.13e-07 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 1364.6  on 901  degrees of freedom
Residual deviance: 1337.6  on 900  degrees of freedom
AIC: 2342.6

Number of Fisher Scoring iterations: 5
```

8.3.4 Interpreting the Output

In the R output for the one-predictor Poisson regression model, the first part is the call, which shows the R command for the model. The second part shows the minimum, first quarter, median, third quarter, and maximum values of the deviance residuals. The third part shows the coefficients table including the parameter estimates for the predictor variable and the intercept, their standard errors, the Wald z statistics, and the associated p values. The null hypothesis for the Wald test is that the coefficient of the predictor variable is zero, and the alternative hypothesis is that the coefficient of the predictor variable is significantly different from zero.

The Wald z statistic equals the parameter estimate divided by its standard error.

$$\text{Wald } z = \frac{\hat{\beta}}{\text{SE}(\hat{\beta})} \tag{8.9}$$

where $\hat{\beta}$ is the estimated log coefficient and $\text{SE}(\hat{\beta})$ is the standard error.

$$\text{Confidence intervals for the parameter } \beta = \hat{\beta} \pm z{*}\text{SE}(\hat{\beta})$$

where $\hat{\beta}$ is the estimated log coefficient for a predictor, z is the z-score from a normal distribution for the chosen confidence interval, and $\text{SE}(\hat{\beta})$ is the standard error. For the 95% confidence intervals, $z = 1.96$. So 95% confidence intervals for the parameter $\beta = \hat{\beta} \pm 1.96{*}\text{SE}(\hat{\beta})$.

For the predictor variable `educ`, Wald $z = 5.187$. The associated p value, `Pr(>|z|)` < `.001`, so we reject the null hypothesis. The rejection of the null hypothesis indicates that the predictor variable `educ` is a significant predictor of the count response variable `vistzoo`. For a one-unit increase in education the log expected number of visits to a zoo increases by a factor of .067.

Finally, the fourth part of the output shows the fit statistics including the null deviance, the residual deviance, and the AIC. The null deviance is the deviance for the null model with the intercept only. The residual deviance is the deviance for the fitted model, which is defined as −2(log likelihood of the current model − log likelihood of the saturated model). The difference between the residual deviance and the null deviance can be used to evaluate the significance of the fitted model.

We can extract the coefficients with `coef(PR.1)` and obtain the profiled confidence intervals with the `confint(PR.1)` command. The Wald confidence intervals are extracted with `confint.default(PR.1)`. The profiled confidence intervals are based on the profile likelihood while the Wald confidence intervals are based on the Wald test. Although the results of both confidence intervals are similar in this example, the former is preferred when the sample size is small.

```
> coef(PR.1)
(Intercept)        educ
-1.15403149  0.06715748
> confint(PR.1)
Waiting for profiling to be done...
                  2.5 %       97.5 %
(Intercept) -1.52744224  -0.78592700
educ         0.04178749   0.09254033
> confint.default(PR.1)
                  2.5 %       97.5 %
(Intercept)  -1.52475464  -0.78330834
educ          0.04178338   0.09253158
```

8.3.5 Interpreting the Incidence Rate Ratios in the One-Predictor Poisson Regression Model

The Poisson regression model estimates the log expected counts of an event. Recall that the exponentiated (β_j) is the incidence rate ratio (IRR) for a one-unit change in a predictor variable. In this model, the IRR for educ is 1.069, which indicates that for a one-unit increase in education the incidence rate or the expected number of visits to a zoo increases by a factor of 1.069. In other words, for a one-unit increase in education the expected number of visits to a zoo increases by 6.9%.

The above results can be obtained using the exp(coef(PR.1)) command. We also use the exp(confint(PR.1)) command to obtain the corresponding confidence intervals. Both results are combined with the cbind(exp(PR.1)), exp(confint(PR.1))) command.

```
> exp(coef(PR.1))
(Intercept)       educ
  0.3153628  1.0694629
> exp(confint(PR.1))
Waiting for profiling to be done...
                2.5 %     97.5 %
(Intercept) 0.2170902  0.4556971
educ        1.0426729  1.0969574
> cbind(exp(coef(PR.1)), exp(confint(PR.1)))
Waiting for profiling to be done...
                             2.5 %      97.5 %
(Intercept)  0.3153628  0.2170902  0.4556971
educ         1.0694639  1.0426729  1.0969574
```

The standard errors of IRRs can be obtained with `exp(coef(PR.1))*sqrt(diag(vcov(PR.1)))`.

```
> exp(coef(PR.1))*sqrt(diag(vcov(PR.1)))
(Intercept)          educ
 0.05965023    0.01384550
```

8.3.6 Model Fit Statistics

Testing the Overall Model Using the Likelihood Ratio Test

To test if the overall model is significant, we fit a null model with the intercept only and compare the single-predictor model with the null model using the `anova()` function. The null model is fitted using the `glm()` function with 1 as the intercept term in the model formula. The command and the output are displayed below.

```
> # Testing the overall model using the likelihood ratio test
> PR.0 <- glm(vistzoo ~ 1, family = poisson, data = count)
> summary(PR.0)

Call:
glm(formula = vistzoo ~ 1, family = poisson, data = count)

Deviance Residuals:
    Min        1Q    Median      3Q      Max
-1.2731   -1.2731   -1.2731   0.2031   3.6938

Coefficients:
            Estimate  Std. Error  z value   Pr(>|z|)
(Intercept) -0.21020     0.03699   -5.683   1.32e-08 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 1364.6  on 901  degrees of freedom
Residual deviance: 1364.6  on 901  degrees of freedom
AIC: 2367.6

Number of Fisher Scoring iterations: 6
```

The `anova(PR.0, PR.1, test = "Chisq"))` command compares the deviance statistics of the fitted model `PR.1` and the null model `PR.0` using the likelihood ratio test.

```
> anova(PR.0, PR.1, test = "Chisq")
Analysis of Deviance Table

Model 1: vistzoo ~ 1
Model 2: vistzoo ~ educ
  Resid. Df  Resid. Dev  Df  Deviance   Pr(>Chi)
1       901      1364.5
2       900      1337.6   1    26.935  2.104e-07 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Here we first fit the null model and then compare the single-predictor model with the null model. This two-step process can also be simplified to the one-line command with the `update()` function within the `anova()` function. In the `anova(PR.1, update(PR.1, ~1), test = "Chisq")` command, we use `update(PR.1, ~ 1)` to fit the null model.

```
> anova(PR.1, update(PR.1, ~1), test = "Chisq")
Analysis of Deviance Table

Model 1: vistzoo ~ educ
Model 2: vistzoo ~ 1
  Resid. Df  Resid. Dev  Df  Deviance   Pr(>Chi)
1       900      1337.6
2       901      1364.5  -1   -26.935  2.104e-07 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The null hypothesis of the test for the overall model is that the predictor variable does not contribute to the model, and the alternative hypothesis is that the one-predictor model is better than the null model with no independent variables. The likelihood ratio test statistic $LR\ \chi^2_{(1)} = 26.935, p < .001$, which indicates that the overall model with one predictor `educ` is significantly different from zero. Therefore, the one-predictor model provides a better fit than the null model in predicting the log number of visits to a zoo in a year. The output displays a negative value for the likelihood ratio test statistic since the residual deviance for the one-predictor model is smaller than that for the null model.

Pseudo R^2

We use the `nagelkerke()` function in the `rcompanion` package (Mangiafico, 2021) to compute the pseudo R^2 statistics for the single-predictor model. We load the package first with `library(rcompanion)` and then execute the `nagelkerke(PR.1)` command.

```
> # Pseudo R2 with nagelkerke()
> library(rcompanion)
> nagelkerke(PR.1)
$`Models`

Model: "glm, vistzoo ~ educ, poisson, count"
Null:  "glm, vistzoo ~ 1, poisson, count"

$Pseudo.R.squared.for.model.vs.null
                             Pseudo.R.squared
McFadden                            0.0113862
Cox and Snell (ML)                  0.0294198
Nagelkerke (Cragg and Uhler)        0.0317234

$Likelihood.ratio.test
Df.diff    LogLik.diff     Chisq        p.value
     -1        -13.467    26.935    2.1043e-07

$Number.of.observations

Model: 902
Null:  902

$Messages
[1] "Note: For models fit with REML, these statistics are based on refitting with ML"

$Warnings
[1] "None"
```

The McFadden R^2 is .011, the Cox and Snell R^2 is .029, and the Nagelkerke R^2 is .032. The same results can be computed using the equations for these three pseudo R^2 statistics. In the R command below, `LLM1` is the log-likelihood value for the single-predictor model and `LL0` is the log-likelihood value for the null model. The number of the observations is 902. In addition, `McFadden1` is the object name for the McFadden R^2, `CS1` for the Cox and Snell R^2, and `NG1` for the Nagelkerke R^2.

```
> # Pseudo R2 with equations
> LLM1 <- logLik(PR.1)
> LL0 <- logLik(PR.0)
> McFadden1 <- 1-(LLM1/LL0)
> McFadden1
'log Lik.' 0.01138621 (df=2)
> CS1 <- 1-exp(2*(LL0-LLM1)/902)
> CS1
'log Lik.' 0.02941984 (df=1)
> NG1 <- CS1/(1-exp(2*LL0/902))
> NG1
'log Lik.' 0.03172344 (df=1)
```

AIC and BIC Statistics

The AIC and BIC statistics can also be computed from the `AIC()` and `BIC()` functions. The output is shown as follows.

```
> # AIC and BIC Statistics
> AIC(PR.1)
[1] 2342.636
> BIC(PR.1)
[1] 2352.245
```

8.4 FITTING A MULTIPLE-PREDICTOR POISSON REGRESSION MODEL WITH R

8.4.1 The Poisson Regression Model: Multiple-Predictor Model

We still use the `glm()` function for multiple Poisson regression. The `glm(vistzoo ~ educ + maritals + female + wrkfull, family = poisson, data = count)` command tells R to predict the count response variable `vistzoo` from the four independent variables. In the `glm()` function, the model equation is specified as `vistzoo ~ educ + maritals + female + wrkfull`. The `family = poisson` argument specifies that the Poisson family is used to fit the model. The data argument specifies `data = count`. The fitted model is named `PR.2`. The output is shown by the `summary(PR.2)` function.

```
> # Multiple-predictor Poisson regression model with glm()
> PR.2<-glm(vistzoo ~ educ + maritals + female + wrkfull, family = poisson,
data = count)
> summary(PR.2)

Call:
glm(formula = vistzoo ~ educ + maritals + female + wrkfull, family = poisson,
    data = count)

Deviance Residuals:
    Min        1Q     Median        3Q       Max
-1.7880   -1.1659   -0.9280    0.4835    4.4634

Coefficients:
              Estimate   Std. Error   z value   Pr(>|z|)
(Intercept)   -1.34993      0.20084    -6.721   1.80e-11 ***
educ           0.05086      0.01343     3.787   0.000152 ***
maritals       0.21382      0.07494     2.853   0.004328 **
female         0.04837      0.07535     0.642   0.520896
wrkfull        0.53962      0.07792     6.925   4.35e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 1364.6  on 901  degrees of freedom
Residual deviance: 1275.8  on 897  degrees of freedom
AIC: 2286.8

Number of Fisher Scoring iterations: 6
```

8.4.2 Interpreting R Output

In the R output for the multiple Poisson regression model, the first part is the call, which shows the R command for the model. The second part shows the minimum, first quarter, median, third quarter, and maximum values of the deviance residuals. The third part shows the coefficients table including the parameter estimates for the four predictor variables and the intercept, their standard errors, the Wald z statistics, and the associated p values.

For the predictor variable `educ`, Wald $z = 3.787$. The associated p value, `Pr(>|z|) < .001`, so we reject the null hypothesis. The rejection of the null hypothesis indicates that the predictor variable `educ` is a significant predictor of the count response variable `vistzoo`.

For the predictor variable `maritals`, Wald $z = 2.853$. The associated p value, `Pr(>|z|) < .001`, so we reject the null hypothesis. For the predictor variable `wrkfull`, Wald $z = 6.925$. The associated p value, `Pr(>|z|) < .001`, so we also reject the null hypothesis. Therefore, `maritals` and `wrkfull` are significant predictors of the count response variable.

For the predictor variable `female`, the Wald $z = .642$. The associated p value `Pr(>|z|) = .521`, so we fail to reject the null hypothesis and conclude that there is no significant effect of female on the outcome variable. In other words, whether a person is a female or male does not significantly predict the log expected number of visits to a zoo.

Finally, the fourth part of the output shows the fit statistics including the null deviance, the residual deviance, and the AIC.

We use the `coef(PR.2)` command to extract the coefficients. Then we use the `confint(PR.2)` command to compute the corresponding confidence intervals.

```
> coef(PR.2)
(Intercept)          educ      maritals        female       wrkfull
-1.34992531    0.05085899    0.21382421    0.04837019    0.53962440
> confint(PR.2)
Waiting for profiling to be done...
                   2.5 %        97.5 %
(Intercept)  -1.74637649   -0.95905948
educ          0.02452348    0.07716665
maritals      0.06694611    0.36083727
female       -0.09900149    0.19647922
wrkfull       0.38773418    0.69331480
```

We request the IRRs with `exp(coef(PR.2))` and `exp(confint(PR.2))`, respectively. The results are combined with the `cbind()` function.

```
> exp(coef(PR.2))
(Intercept)        educ     maritals       female      wrkfull
0.2592596     1.0521745    1.2384049    1.0495591    1.7153624
> exp(confint(PR.2))
Waiting for profiling to be done...
                 2.5 %      97.5 %
(Intercept)  0.1744048   0.3832532
educ         1.0248267   1.0802221
maritals     1.0692379   1.4345300
female       0.9057414   1.2171100
wrkfull      1.4736380   2.0003353
> cbind(exp(coef(PR.2)), exp(confint(PR.2)))
Waiting for profiling to be done...
                             2.5 %      97.5 %
(Intercept)  0.2592596   0.1744048   0.3832532
educ         1.0521745   1.0248267   1.0802221
maritals     1.2384049   1.0692379   1.4345300
female       1.0495591   0.9057414   1.2171100
wrkfull      1.7153624   1.4736380   2.0003353
```

To compute the standard errors of the IRRs, we multiply the IRRs by the standard errors of the Poisson coefficients with the `exp(coef(PR.2))*sqrt(diag(vcov(PR.2)))` command. In the syntax, `exp(coef(PR.2))` exponentiates the coefficients and `sqrt(diag(vcov(PR.2)))` computes the square root of the variances of the coefficients which are the diagonal elements in the variance-covariance matrix to obtain the standard errors of the coefficients.

```
> exp(coef(PR.2))*sqrt(diag(vcov(PR.2)))
(Intercept)          educ       maritals         female        wrkfull
0.05207020     0.01412980     0.09280886     0.07908132     0.13366331
```

8.4.3 Interpreting the Incidence Rate Ratios (IRRs) in the Multiple-Predictor Poisson Model

The incidence rate ratio is the exponentiated coefficients in a Poisson regression model. It is the ratio of two incidence rates. A positive Poisson regression coefficient corresponds to an incidence rate ratio greater than 1, whereas a negative coefficient is associated with an incidence rate ratio less than 1.

The incidence rate ratio for a predictor can be interpreted as the change in the incidence rate or the number of events for a one-unit increase in the predictor variable when holding other predictors constant.

For educ, the incidence rate ratio is 1.052. The result indicates that the incidence rate increases by a factor of 1.052 for a one-unit increase in education when holding all other predictors constant. It can also be interpreted as the change in the number of events as follows. The expected number of visits to a zoo increases by 1.052 for a one-unit increase in education when holding all other predictors constant. In other words, for a one-unit increase in education the expected number of visits to a zoo increases by 5.2%.

For maritals, the incidence rate ratio is 1.238, which indicates that expected number of visits to a zoo for the married is 1.238 times as large as the expected number of visits for the unmarried when holding the other predictors constant. In other words, the expected number of visits to a zoo for the married is 23.8% higher than that for the unmarried.

The incidence rate ratio for wrkfull can be interpreted in the similar way. The incidence rate ratio is 1.715, which indicates that expected number of visits to a zoo for those working full time is 1.715 times as large as that for those not working full time when holding the other predictors constant. In other words, the expected number of visits to a zoo for those working full time is 71.5% higher than that for those not working full time.

With regard to female, the incidence rate ratio is 1.050, which is not significant (see the associated *p* value in the coefficients table). It indicates that being female does not impact the expected number of visits to a zoo.

8.4.4 Model Fit Statistics

Testing the Overall Model Using the Likelihood Ratio Test

To test if the overall model is significant, we compare the multiple-predictor model with the null model using the anova() function. The anova(PR.0, PR.2, test = "Chisq")) command compares the log-likelihood statistics of the fitted model PR.2 and the null model PR.0 using the likelihood ratio test. The resulting output is displayed below.

```
> anova(PR.2, update(PR.2, ~1), test = "Chisq")
Analysis of Deviance Table

Model 1: vistzoo ~ educ + maritals + female + wrkfull
Model 2: vistzoo ~ 1
    Resid. Df   Resid. Dev   Df   Deviance   Pr(>Chi)
1         897       1275.8
2         901       1364.5   -4    -88.772   < 2.2e-16 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The same results can be obtained with the anova(PR.2, update(PR.2, ~ 1), test = "Chisq")) command. In the command, we use update(PR.2, ~ 1) to fit the null model.

```
> anova(PR.0, PR.2, test = "Chisq")
Analysis of Deviance Table

Model 1: vistzoo ~ 1
Model 2: vistzoo ~ educ + maritals + female + wrkfull
     Resid. Df  Resid. Dev  Df   Deviance    Pr(>Chi)
1        901       1364.5
2        897       1275.8    4    88.772    < 2.2e-16 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The likelihood ratio chi-square test $\chi^2_{(4)} = 88.772$, $p < .001$ indicates that the full model with the four predictors provides a better fit than the null model with no independent variables in predicting the count response variable.

Pseudo R^2

The `nagelkerke(PR.2)` command computes the three types of pseudo R^2 statistics and the likelihood ratio test statistic for the overall multiple-predictor model.

```
> # Pseudo R2 with nagelkerke()
> nagelkerke(PR.2)
$`Models`

Model: "glm, vistzoo ~ educ + maritals + female + wrkfull, poisson, count"
Null:  "glm, vistzoo ~ 1, poisson, count"

$Pseudo.R.squared.for.model.vs.null
                              Pseudo.R.squared
McFadden                             0.0375265
Cox and Snell (ML)                   0.0937285
Nagelkerke (Cragg and Uhler)         0.1010680

$Likelihood.ratio.test
Df.diff    LogLik.diff     Chisq        p.value
     -4        -44.386    88.772    2.4012e-18

$Number.of.observations

Model: 902
Null:  902

$Messages
[1] "Note: For models fit with REML, these statistics are based on refitting with ML"

$Warnings
[1] "None"
```

We can also compute the three types of pseudo R^2 statistics with their equations for the multiple-predictor model as follows.

```
> # Pseudo R2 with equations
> LLM2 <- logLik(PR.2)
> McFadden2 <- 1-(LLM2/LL0)
> McFadden2
'log Lik.' 0.03752648 (df=5)
> CS2 <- 1-exp(2*(LL0-LLM2)/902)
> CS2
'log Lik.' 0.0937285 (df=1)
> NG2 <- CS2/(1-exp(2*LL0/902))
> NG2
'log Lik.' 0.1010675 (df=1)
```

In the output, LLM2 is the log-likelihood value for the multiple-predictor model and LL0 is the log-likelihood value for the null model. In addition, the number of the observations is 902. The three types of pseudo R^2 statistics are as follows. The McFadden R^2 is .038, the Cox and Snell R^2 is .094, and the Nagelkerke R^2 is .101.

AIC and BIC Statistics

The AIC(PR.2) and BIC(PR.2) commands produce the AIC and BIC statistics. We also use the AIC(PR.1, PR.2) and BIC(PR.1, PR.2) commands to compare the AIC and BIC statistics between the two models, respectively.

```
> # AIC and BIC Statistics
> AIC(PR.2)
[1] 2286.799
> BIC(PR.2)
[1] 2310.822
> AIC(PR.1, PR.2)
     df      AIC
PR.1  2 2342.636
PR.2  5 2286.799
> BIC(PR.1, PR.2)
     df      BIC
PR.1  2 2352.245
PR.2  5 2310.822
```

The AIC and BIC statistics for the multiple-predictor model are 2,286.799 and 2,310.822, respectively. Compared with the single-variable model, both AIC and BIC indicate that the multiple-predictor model fits the data better.

8.4.5 Interpreting the Marginal Effects in the Poisson Regression Model

We load the `margins` package (Leeper, 2021) with the `library(margins)` command, compute the average marginal effects with the `margins(PR.2)` command, and name the results `marg.pr`. The summary results are obtained with `summary(marg.pr)` as follows.

```
> # marginal effects
> library(margins)
> marg.pr <- margins(PR.2)
> summary(marg.pr)
   factor     AME      SE       z       p    lower   upper
     educ  0.0412  0.0110  3.7506  0.0002   0.0197  0.0628
   female  0.0392  0.0611  0.6418  0.5210  -0.0805  0.1589
 maritals  0.1733  0.0611  2.8374  0.0045   0.0536  0.2930
  wrkfull  0.4373  0.0652  6.7087  0.0000   0.3096  0.5651
```

The average marginal effect for `educ` is .041. The result indicates that there are on average .041 more visits to a zoo for a one-unit increase in education when holding all other predictors constant.

The average marginal effect for `maritals` is .173. This indicates that the married have on average .173 more visits to the zoo than the unmarried when holding the other predictors constant. The marginal effects of the other two predictor variable can be interpreted in a similar way.

8.4.6 Interpreting the Predicted Counts With the `ggpredict()` Function in the `ggeffects` Package

By using the `ggpredict()` function in the `ggeffects` package (Lüdecke, 2018b), we can compute the predicted number of events of the count response variable at specified values of the predictor variables. We first load the package with `library(ggeffects)` since it has been installed in the previous chapters. The command `pr.ed <- ggpredict(PR.2, terms = "educ[12, 14, 16]")` tells R to compute the predicted counts of the response variable using the `ggpredict()` function. The argument inside the function includes the estimated model, `PR.2`, the `terms = "educ[12, 14, 16]"` option, which specifies the predictor variable `educ` at the values of 12, 14, and 16 when holding other predictor variables at their means. The `terms` option can specify up to four variables, including the second to fourth grouping variables. The output is assigned to the object named `pr.ed`. The `as.data.frame()` function or the `sqrt(diag(vcov()))` function can be used to request the standard errors. The output is omitted here.

```
> library(ggeffects)
> pr.ed <- ggpredict(PR.2, terms = "educ[12, 14, 16]")
> pr.ed
# Predicted counts of vistzoo

educ    |    Predicted    |             95% CI
--------------------------------------------
  12    |         0.70    |    [0.63, 0.77]
  14    |         0.77    |    [0.71, 0.83]
  16    |         0.85    |    [0.78, 0.93]

Adjusted for:
* maritals = 0.44
*   female = 0.56
*  wrkfull = 0.47
> plot(pr.ed)
```

The results table displays the values of `educ`, the predicted counts, and the confidence intervals. When `educ` = 12, and the other predictor variables are held at their means (`maritals` = .44, `female` = .56, and `wrkfull` = .47), the predicted number of visits to a zoo is .70.

When `educ` = 14, and the other three predictor variables are held at their means, the predicted number of visits to a zoo is .77.

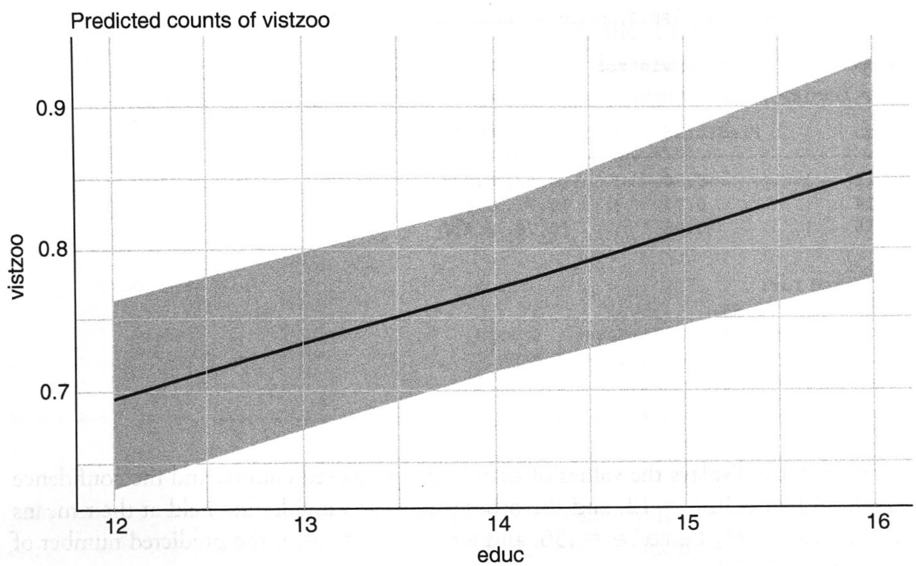
When `educ` = 16, and the other predictor variables are held at their means, the predicted number of visits to a zoo is .85.

The predicted counts are plotted using `plot(pr.ed)`. Figure 8.1 shows the predicted number of visits to a zoo when `educ` is at the values of 12, 14, and 16.

The graph shows that with the increase of the years of education, the predicted number of visits to a zoo increases. In other words, people with higher levels of education are associated with having more visits to a zoo.

In the next example, we compute the predicted counts for a continuous variable at given values by different groups. In the following example, we compute the predicted number of visits to a zoo for `educ` at the values of 12, 14, and 16 by the two groups in `wrkfull` when holding other variables at their means. The command is as follows: `pr.ew <- ggpredict(PR.2, terms = c("educ[12, 14, 16]", "wrkfull"))`. In the `ggpredict()` function, the `terms = c("educ[12, 14, 16]", "wrkfull")` option specifies both `educ` and `wrkfull`, with the latter as the grouping variable. The output is assigned to an object named `pr.ew` and is plotted with the `plot(pr.ew)` function.

FIGURE 8.1 ● Predicted Counts for educ at 12, 14, and 16




```
> pr.ew <- ggpredict(PR.2, terms = c("educ[12, 14, 16]", "wrkfull"))
> pr.ew
# Predicted counts of vistzoo

# wrkfull = 0

educ  |  Predicted  |         95% CI
--------------------------------------
  12  |       0.54  |  [0.47, 0.61]
  14  |       0.60  |  [0.53, 0.67]
  16  |       0.66  |  [0.58, 0.75]

# wrkfull = 1

educ  |  Predicted  |         95% CI
--------------------------------------
  12  |       0.92  |  [0.82, 1.04]
  14  |       1.02  |  [0.93, 1.12]
  16  |       1.13  |  [1.02, 1.25]

Adjusted for:
* maritals = 0.44
*   female = 0.56
> plot(pr.ew)
```

FIGURE 8.2 ● Predicted Counts for `educ` at 12, 14, and 16 by `maritals`

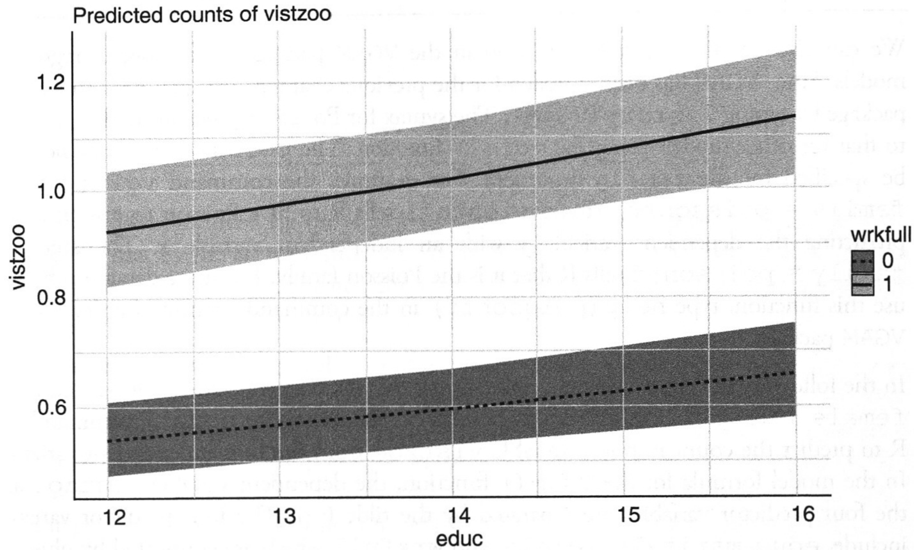


Figure 8.2 shows the predicted counts for `educ` at 12, 14, and 16 by the grouping variable `wrkfull`. As shown in the graph, the predicted number of visits to a zoo increases with the increase of years of education and the predicted counts for the people who work full-time are higher than the predicted counts for those not working full-time.

8.4.7 Model Comparisons Using the Likelihood Ratio Test

The likelihood ratio test, or the deviance difference test, is used to compare the multiple-predictor model and the one-predictor model. In the `anova(PR.1, PR.2, test = "Chisq")` command, `PR.1` and `PR.2` are the two models being compared. The following output is displayed.

```
> # Model comparison with the likelihood ratio test
> anova(PR.1, PR.2, test = "Chisq")
Analysis of Deviance Table

Model 1: vistzoo ~ educ
Model 2: vistzoo ~ educ + maritals + female + wrkfull
    Resid. Df    Resid. Dev    Df    Deviance     Pr(>Chi)
1         900        1337.6
2         897        1275.8     3      61.837    2.381e-13 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The likelihood ratio test, $\chi^2_{(3)} = 61.837, p < .001$, which indicates that the full model with the four predictor variables fits the data better than the single-predictor model.

8.5 POISSON REGRESSION WITH THE vglm() FUNCTION IN THE VGAM PACKAGE

We can also use the vglm() function in the VGAM package to fit Poisson regression models. Since VGAM has been installed for the previous chapters, you just need to load the package by typing library(VGAM). The syntax for Poisson regression models is similar to that for other models using the vglm() function. The poissonff family needs to be specified for the family argument. For example, the command vglm(y ~ x, family = poissonff, data = data1) tells R to fit a Poisson regression model predicting the dependent variable *y* with an independent variable *x*. The argument family = poissonff tells R that it is the Poisson family. For more details on how to use this function, type help(poissonff) in the command prompt after loading the VGAM package.

In the following example, the pr.v <- vglm(vistzoo ~ educ + maritals + female + wrkfull, family = poissonff, data = count) command tells R to predict the count response variable vistzoo from the four independent variables. In the model formula for the vglm() function, the dependent variable vistzoo and the four predictor variables are separated by the tilde (~). The four predictor variables include, educ, maritals, female, and wrkfull which are connected by plus (+) symbols. We also specify the data arguments data = count. The fitted model is named pr.v. The following output is shown by the summary(pr.v) command.

```
> # Multiple-predictor Poisson regression model with vglm() in VGAM
> library(VGAM)
Loading required package: stats4
Loading required package: splines
> pr.v <- vglm(vistzoo ~ educ + maritals + female + wrkfull, family = poissonff,
data=count)
> summary(pr.v)

Call:
vglm(formula = vistzoo ~ educ + maritals + female + wrkfull,
     family = poissonff, data = count)


Pearson residuals:
                      Min        1Q    Median      3Q     Max
loglink(lambda)    -1.264   -0.8244   -0.6562  0.5319   8.247


Coefficients:
                Estimate   Std. Error   z value   Pr(>|z|)
(Intercept)     -1.34993     0.20084    -6.721   1.80e-11 ***
educ             0.05086     0.01343     3.787   0.000152 ***
maritals         0.21382     0.07494     2.853   0.004328 **
female           0.04837     0.07535     0.642   0.520896
wrkfull          0.53962     0.07792     6.925   4.35e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Name of linear predictor: loglink(lambda)
```

```
Residual deviance: 1275.779 on 897 degrees of freedom

Log-likelihood: -1138.399 on 897 degrees of freedom

Number of Fisher scoring iterations: 6

No Hauck-Donner effect found in any of the estimates
```

The R output produced by the `vglm()` function for the Poisson regression model is similar to that by the `glm()` function. It includes the call of the model command, the Pearson residuals, the coefficients, the name of linear predictor, the residual deviance, the log-likelihood, and the number of Fisher scoring iterations.

The first section shows the call, which is the R command for the model. The second section shows the minimum, first quarter, median, third quarter, and maximum values of the Pearson residuals. The third section shows the coefficients table including the parameter estimates for the predictor variable and the intercept, their standard errors, the Wald z statistics, and the associated p values. The fourth section shows the name of linear predictor or the link function, which is the log link for the expected count. The fifth section shows the residual deviance and the degrees of freedom. The sixth section provides the log-likelihood value and the degrees of freedom. Finally, the number of Fisher scoring iterations is displayed at the end.

The Poisson coefficients in the coefficients section (labeled `Coefficients:`) are the same as those produced from the `glm()` function. See the preceding section on the interpretation of the coefficients.

The Poisson coefficients of the predictor variables can be extracted by using `coef(pr.v, matrix = TRUE)`. The confidence intervals are obtained with the `confint(pr.v, matrix = TRUE)` command.

```
> coef(pr.v, matrix = TRUE)
              loglink(lambda)
(Intercept)       -1.34992531
educ               0.05085899
maritals           0.21382421
female             0.04837019
wrkfull            0.53962440
> confint(pr.v, matrix = TRUE)
                    2.5 %        97.5 %
(Intercept)   -1.74356822   -0.95628240
educ           0.02453836    0.07717963
maritals       0.06694009    0.36070834
female        -0.09930758    0.19604796
wrkfull        0.38690141    0.69234738
```

We use the `exp(coef(pr.v, matrix = TRUE))` and `exp(confint(pr.v, matrix = TRUE))` commands to compute the IRRs and the corresponding confidence intervals, respectively. Both results are combined with the `cbind()` function.

```
> exp(coef(pr.v, matrix = TRUE))
              loglink(lambda)
(Intercept)        0.2592596
educ               1.0521745
maritals           1.2384049
female             1.0495591
wrkfull            1.7153624
> exp(confint(pr.v, matrix = TRUE))
                  2.5 %     97.5 %
(Intercept)    0.1748952  0.384319
educ           1.0248419  1.080236
maritals       1.0692314  1.434345
female         0.9054642  1.216585
wrkfull        1.4724113  1.998401
> cbind(exp(coef(pr.v, matrix = TRUE)), exp(confint(pr.v, matrix = TRUE)))
             loglink(lambda)       2.5 %     97.5 %
(Intercept)  0.2592596  0.1748952  0.384319
educ         1.0521745  1.0248419  1.080236
maritals     1.2384049  1.0692314  1.434345
female       1.0495591  0.9054642  1.216585
wrkfull      1.7153624  1.4724113  1.998401
```

The standard errors of the IRRs are computed with the `exp(coef(pr.v))*sqrt(diag(vcov(pr.v)))` command.

```
> exp(coef(pr.v))*sqrt(diag(vcov(pr.v)))
(Intercept)        educ      maritals       female      wrkfull
 0.05207020  0.01412980  0.09280886  0.07908133  0.13366331
```

We use the `nagelkerke(pr.v)` command to compute the three types of pseudo R^2 statistics and the likelihood ratio test statistic for the multiple-predictor model.

```
> # Pseudo R2 with nagelkerke()
> library(rcompanion)
> nagelkerke(pr.v)
$`Models`


Model: "vglm, vistzoo ~ educ + maritals + female + wrkfull, poissonff, count"
Null:  "vglm, vistzoo ~ 1, poissonff, count"


$Pseudo.R.squared.for.model.vs.null
                             Pseudo.R.squared
McFadden                            0.0375265
Cox and Snell (ML)                  0.0937285
Nagelkerke (Cragg and Uhler)        0.1010680


$Likelihood.ratio.test
Df.diff   LogLik.diff    Chisq      p.value
      4       -44.386   88.772   2.4012e-18
```

```
$Number.of.observations

Model: 902
Null:  902

$Messages
[1] "Note: For models fit with REML, these statistics are based on refitting with ML"

$Warnings
[1] "None"
```

8.6 MAKING PUBLICATION-QUALITY TABLES

8.6.1 Presenting the Results Using the stargazer Package

We can use the `stargazer` package (Hlavac, 2018) to make a table containing the results of the fitted models with the `glm()` function. Since the package has been installed in the preceding chapters, we only need to load the package by typing `library(stargazer)`. After fitting the single-predictor model `PR.1` and the multiple-predictor model `PR.2`, we use the command as follows: `stargazer(PR.1, PR.2, type = "text", align = TRUE, out = "pr2mod.txt")`. In the `stargazer()` function, we first specify the two model objects to be presented and then the type of the table. The option `type = "text"` specifies the table type and the `align = TRUE` option aligns the results of the two models. The `out = "pr2mod.txt"` argument saves the output named `pr2mod.txt`.

```
> library(stargazer)
> stargazer(PR.1, PR.2, type = "text", align = TRUE, out = "pr2mod.txt")

==============================================
                    Dependent variable:
               ----------------------------
                         vistzoo
                    (1)            (2)
 ------------------------------------------------------
educ             0.067***       0.051***
                 (0.013)        (0.013)

maritals                        0.214***
                                (0.075)

female                           0.048
                                (0.075)

wrkfull                         0.540***
                                (0.078)

Constant        -1.154***      -1.350***
                 (0.189)        (0.201)

-------------------------------------------------------
Observations       902             902
Log Likelihood  -1,169.318      -1,138.400
Akaike Inf. Crit. 2,342.636      2,286.799
=============================================
Note:           *p<0.1;   **p<0.05;   ***p<0.01
```

TABLE 8.1 ● Results of the Poisson Regression Models: Single-Predictor and Multiple-Predictor Models (Shown in Original Format Generated by R)

	Dependent variable:	
	vistzoo	
	(1)	(2)
educ	0.067***	0.051***
	(0.013)	(0.013)
maritals		0.214***
		(0.075)
female		0.048
		(0.075)
wrkfull		0.540***
		(0.078)
Constant	−1.154***	−1.350***
	(0.189)	(0.201)
Observations	902	902
Log Likelihood	−1,169.318	−1,138.400
Akaike Inf. Crit.	2,342.636	2,286.799

**p* < .1
***p* < .05
****p* < .01

We can also create the table in the HTML format and copy it into Microsoft Word. The command is as follows: `stargazer(PR.1, PR.2, type = "html", align = TRUE, out = "pr2mod.htm")`. It produces Table 8.1, as shown here in its original format, presenting the results of both the single-predictor and the multiple-predictor Poisson regression models.

Presenting the Results of the `vglm` Models Using the `texreg` Package

We can also use the `screenreg()` and `htmlreg()` functions from the `texreg` package (Leifeld, 2013) to present the results. If you have not done so, you need to install the package first by typing `install.packages("texreg")` and then load the package by typing `library(texreg)`. Since the package has been installed in previous chapters, we only need to load the package by typing `library(texreg)`. We use the: `screenreg(list(PR.1, PR.2))` command to display the results. In the `screenreg()` function, we specify the two model objects to be presented with the `list()` function. The output for the resulting plain text table is omitted here.

8.7 REPORTING THE RESULTS

Reporting the results for Poisson regression is similar to that used for binary logistic regression. The following are the generic guidelines for reporting the results. You may need to adjust your writing since your discipline or journals may have different requirements.

First, describe the Poisson regression model, the count response variable and the independent variables, and your research hypothesis or the purpose of your study. Include a couple of sentences justifying your use of this model for the analysis.

Second, report the likelihood ratio test statistic for the model and the associated *p* value, followed by the interpretation on whether the fitted model is better than the null model. If more than one model is developed, then compare models using likelihood ratio test statistics and/or the AIC and BIC statistics.

Third, report the parameter estimates for the predictor variables, their standard errors, and the associated *p* values in a table. In addition, report the incidence rate ratio for each predictor in the table or text and interpret the results. The following is an example of summarizing the results for the Poisson regression model illustrated previously.

> The Poisson regression analysis was conducted to predict the count outcome variable, the number of zoo visits in a year, from a set of predictor variables, such as marital status, years of education, gender, and working status. The Poisson regression model was fitted since the response variable was a count of the number of visits to a zoo in a year.
>
> The likelihood ratio test for the fitted model $\chi^2_{(4)} = 88.772$, $p < .001$, indicates that the full model with the four predictors provides a better fit than the null model with no independent variables in predicting the count response variable.
>
> Table 8.1 displays the parameter estimates for the multiple-predictor Poisson regression model. The results can be interpreted in terms of the incidence rate ratios which are the exponentiated coefficients.
>
> For `educ`, the incidence rate ratio is 1.052. The result indicates that the incidence rate increases by a factor of 1.052 for a one-unit increase in education when holding all other predictors constant.
>
> For `maritals`, the incidence rate ratio is 1.238, which indicates that expected number of visits to a zoo for the married is 1.238 times as large as the expected number of visits for the unmarried when holding other

predictors constant. In other words, the expected number of visits to a zoo for the married is 23.8% higher than that for the unmarried.

The incidence rate ratio for `wrkfull` can be interpreted in the similar way. The incidence rate ratio is 1.715, which indicates that the expected number of visits to a zoo for those working full time is 71.5% higher than that for those not working full time.

With regard to `female`, the incidence rate ratio is 1.050, $p = .521$, which is not significant. It indicates that being female does not impact the expected number of visits to a zoo.

8.8 SUMMARY OF R COMMANDS IN THIS CHAPTER

```
# Chap 8 R Script

# Remove all objects
rm(list = ls(all = TRUE))

# The following user-written packages need to be installed first by using
install.packages(" ") and then by loading it with library()
# library(VGAM)          # It is already installed for Chapter 4
# library(rcompanion)    # It is already installed for Chapter 3
# library(margins)       # It is already installed for Chapter 3
# library(ggeffects)     # It is already installed for Chapter 2
# library(stargazer)     # It is already installed for Chapter 2

# Import the count dataset
library(foreign)
count <- read.dta("C:/CDA/count.dta")

# Convert variables from integer to numeric so they will work well with ggpredict()
count$educ <- as.numeric(count$educ)
count$wrkfull <- as.numeric(count$wrkfull)
count$maritals <- as.numeric(count$maritals)

attach(count)

# One-predictor Poisson regression model with glm()
PR.1 <- glm(vistzoo ~ educ, family = poisson, data = count)
summary(PR.1)
coef(PR.1)
confint(PR.1)
confint.default(PR.1)
exp(coef(PR.1))
```

```
exp(coef(PR.1))*sqrt(diag(vcov(PR.1)))
exp(confint(PR.1))
cbind(exp(coef(PR.1)), exp(confint(PR.1)))

# Testing the overall model using the likelihood ratio test
PR.0 <- glm(vistzoo ~ 1, family = poisson, data = count)
summary(PR.0)
anova(PR.0, PR.1, test = "Chisq")
anova(PR.1, update(PR.1, ~1), test = "Chisq")

# Pseudo R2 with nagelkerke()
library(rcompanion)
nagelkerke(PR.1)

# Pseudo R2 with equations
LLM1 <- logLik(PR.1)
LL0 <- logLik(PR.0)
McFadden1 <- 1-(LLM1/LL0)
McFadden1
CS1 <- 1-exp(2*(LL0-LLM1)/902)
CS1
NG1 <- CS1/(1-exp(2*LL0/902))
NG1

# AIC and BIC Statistics
AIC(PR.1)
BIC(PR.1)

# Multiple-predictor Poisson regression model with glm()
PR.2 <-glm(vistzoo ~ educ + maritals + female + wrkfull, family = poisson, data = count)
summary(PR.2)
coef(PR.2)
confint(PR.2)
exp(coef(PR.2))
exp(confint(PR.2))
cbind(exp(coef(PR.2)), exp(confint(PR.2)) )
exp(coef(PR.2))*sqrt(diag(vcov(PR.2)))

# Testing the overall model using the likelihood ratio test
anova(PR.2, update(PR.2, ~1), test = "Chisq")
anova(PR.0, PR.2, test = "Chisq")

# Pseudo R2 with nagelkerke()
nagelkerke(PR.2)

# Pseudo R2 with equations
LLM2 <- logLik(PR.2)
McFadden2 <- 1-(LLM2/LL0)
McFadden2
CS2 <- 1-exp(2*(LL0-LLM2)/902)
CS2
NG2 <- CS2/(1-exp(2*LL0/902))
NG2
```

```
# AIC and BIC Statistics
AIC(PR.2)
BIC(PR.2)
AIC(PR.1, PR.2)
BIC(PR.1, PR.2)

# Model comparison with the likelihood ratio test
anova(PR.1, PR.2, test = "Chisq")

library(stargazer)
stargazer(PR.1, PR.2, type = "text", align = TRUE, out = "pr2mod.txt")
stargazer(PR.1, PR.2, type = "html", align = TRUE, out = "pr2mod.htm")

# Marginal effects: AME
library(margins)
marg.pr <- margins(PR.2)
summary(marg.pr)

# Predicted counts with ggpredict() in ggeffects
library(ggeffects)
pr.ed <- ggpredict(PR.2, terms = "educ[12, 14, 16]")
pr.ed
plot(pr.ed)
pr.ew <- ggpredict(PR.2, terms = c("educ[12, 14, 16]", "wrkfull"))
pr.ew
plot(pr.ew)

# Multiple-predictor Poisson regression model with vglm() in VGAM
library(VGAM)
pr.v <- vglm(vistzoo ~ educ + maritals + female + wrkfull, family = poissonff,
data=count)
summary(pr.v)
coef(pr.v, matrix = TRUE)
confint(pr.v, matrix = TRUE)
exp(coef(pr.v, matrix = TRUE))
exp(confint(pr.v, matrix = TRUE))
cbind(exp(coef(pr.v, matrix = TRUE)), exp(confint(pr.v, matrix = TRUE)))
exp(coef(pr.v))*sqrt(diag(vcov(pr.v)))

# Pseudo R2 with nagelkerke()
library(rcompanion)
nagelkerke(pr.v)

detach(count)
```

Glossary

A marginal effect in Poisson regression is a change in the expected counts of a response variable related to the change in an independent variable.

Count response variables are nonnegative integers and follow a Poisson distribution.

In the **Poisson distribution**, the mean of a count variable is equal to the variance of the variable.

The incidence rate is defined as the expected number of events during a period of time or in a location.

The Poisson regression model is used to estimate a count response variable. It estimates the log expected counts of an event or the log incidence rate of the response variable.

Exercises

Use the `rwm1984` data (Hilbe, 2014) available at **https://edge.sagepub.com/liu1e** for the following problems.

1. Conduct an analysis for a Poisson regression model and estimate the count response variable `docvis` (the number of visits to a doctor in a year) from the two predictor variables, `outwork` (1 = not working and 0 = working) and `female` (1 = female and 0 = male).
2. Interpret the likelihood ratio test for the overall model.
3. List three measures of pseudo R^2 and the AIC and BIC statistics.
4. In the regression table, identify the coefficients for the predictor variable `outwork` and `female`. Are they both statistically significant?
5. Compute the IRRs and interpret the IRR for the predictor variable `outwork`.
6. Make a publication-quality table containing the estimated coefficients.
7. Write a report to summarize the results from the output.