Poisson Regression

Dr. Kiah Wah Ong

Introduction

By now, we are very familiar with the situation when the response variable Y is continuous, taking values in $\mathbb{R}$.

In that situation, the regression model we used were the usual linear regression.

In the previous module, we investigated the case when our response variable Y is binary, taking values 0 or 1.

In that case the regression model we used were Logistic regression.

Introduction

What do we do when our response variable Y is a count variable, taking values like $0, 1, 2, 3, \cdots$?

Here are some examples of such variable.

Example

- The number of persons killed in a car accident per year.
- The number of childbirths in a specific hospital per month.
- The number of damages inflicted on a Lancaster bomber over 30 missions.

Poisson Distribution

The variable, y, that we saw earlier is a type of random variable that said to follow a Poisson distribution.

That is, this random variable Y take values in $\{0, 1, 2, \cdots\}$ and

$$\Pr(Y = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

for $k = 0, 1, 2, \cdots$.

Here, $\lambda > 0$ is the expected value of Y, i.e. $\lambda = \mathrm{E}(Y)$.

Poisson Regression Model

Suppose Y is the number of damages inflicted on a Lancaster bomber over 30 missions.

We expect the mean of Y, $\lambda = \mathrm{E}(Y)$, to vary as a function of X_1, X_2 and X_3 shown below:

- X_1: Bomb load (Quantitative)
- X_2: Total months of aircrew experience (Quantitative)
- X_3: Day or Night mission (Qualitative)

Let us write

$$\lambda(X_1, X_2, X_3)$$

to indicate that λ is a function of the predictor variables X_1, X_2 and X_3.

Poisson Regression Model

Instead of model the number of damages, Y, as a Poisson distribution with a fixed mean value λ_0, we want to consider the following model for the mean $\lambda(X_1, X_2, X_3)$:

$$\lambda(X_1, X_2, X_3) = e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3}$$

or equivalently

$$\ln\left(\lambda(X_1, X_2, X_3)\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

where $\beta_0, \beta_1, \beta_2$ and β_3 are parameters to be estimated.

Poisson Regression Model

Note: We write

$$\ln\left(\lambda(X_1, X_2, X_3)\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

instead of

$$\lambda\left(X_1, X_2, X_3\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

in order to ensure that $\lambda\left(X_1, X_2, X_3\right)$ takes on nonnegative values for all values of the predictor variables.

Estimation in Poisson Regression

To estimate the coefficients $\beta_0, \beta_1, \beta_2$ and β_3, we will employ the same maximum likelihood approach seen in the section on logistic regression.

Estimation in Poisson Regression

In order to use the maximum likelihood method. We would like to calculate the total probability of observing all of the data

$$(x_{i1}, x_{i2}, x_{i3}, y_i), \ i = 1, \cdots, n$$

namely

$$\Pr(Y_i = y_i, i = 1, \cdots, n)$$

This can get very complicated, hence we need the assumption that each data point is generated independently of each others. With that we can write

$$\Pr(Y_i = y_i, i = 1, \cdots, n) = \prod_{i=1}^{n} \Pr(Y_i = y_i)$$

Estimation in Poisson Regression

With the assumption that the observations are independent, we get

$$\begin{aligned}\Pr(Y_i = y_i, i = 1, \cdots, n) &= \prod_{i=1}^{n} \Pr(Y_i = y_i) \\ &= \prod_{i=1}^{n} \frac{e^{-\lambda(x_i)} \lambda(x_i)^{y_i}}{y_i!}\end{aligned}$$

Where $\lambda(x_i) = e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3}}$.

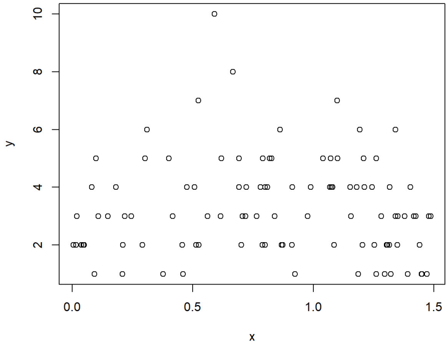
We estimate the coefficients by maximizing the likelihood function above.

Poisson Regression in *R*

We now look at how to perform Poisson regression in *R*: Let us import the simulated data from data set Poisson1.csv and plot the data.

```
PoissonR<-read.csv("Poisson1.CSV", header=TRUE, sep=",")
x<-PoissonR$x
y<-PoissonR$y
plot(x,y)


Model1<-glm(y~x, family=poisson)
summary(Model1)
```



Poisson Regression in *R*

The glm() and summary() functions give the following output.

```
PoissonR<-read.csv("Poisson1.CSV", header=TRUE, sep=",")
x<-PoissonR$x
y<-PoissonR$y
plot(x,y)


Model1<-glm(y~x, family=poisson)
summary(Model1)
```

```
Call:
glm(formula = y ~ x, family = poisson)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5350  -0.7651  -0.1358   0.4176   2.9488

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  1.22914    0.11558  10.635   <2e-16 ***
x           -0.04981    0.12320  -0.404    0.686
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 83.180  on 99  degrees of freedom
Residual deviance: 83.017  on 98  degrees of freedom
AIC: 382.47

Number of Fisher Scoring iterations: 4
```

Making Prediction

The $\text{predict}()$ function can be used to obtain the fitted value from this Poisson regression model.

We need to use the argument $\text{type} = \text{“response”}$ to specify that we want the output to be

$$\lambda(X) = \exp(\hat{\beta}_0 + \hat{\beta}_1 X)$$

Without that argument, the output $\hat{\beta}_0 + \hat{\beta}_1 X$ will be given by default. See the example below:

```
newdata1 <- data.frame(x=c(0.01, 1.013, 1.21, 0.145))
predict(Model1, newdata=newdata1, type="response")
```

The following output are for $X = 0.01, 1.013, 1.21$ and 0.145.

```
> newdata1 <- data.frame(x=c(0.01, 1.013, 1.21, 0.145))
> predict(Model1, newdata=newdata1, type="response")
       1        2        3        4 
3.416598 3.250106 3.218371 3.393701 
```

Interpret the Model Parameters

From the *R*-output, we see that $\hat{\beta}_0 = 1.22914$ and $\hat{\beta}_1 = -0.04981$.

Hence the fitted model is given by

$$\ln(\mathrm{E}(\mathrm{Y}|\mathrm{X} = \mathrm{x})) = \ln(\lambda(x)) = 1.22914 - 0.04981x$$

Notice that from the *p*-value of 0.686, $\hat{\beta}_1 = -0.04981$ is not statistically significant.

Interpret the Model Parameters

In order to understand the meaning of these parameters, let us write

$$\ln(\lambda) = \hat{\beta}_0 + \hat{\beta}_1 x$$

and consider $x = x_0$ and $x = x_0 + 1$.

Let the expected value λ, when $x = x_0$ be

$$\lambda_0 = e^{\hat{\beta}_0} e^{\hat{\beta}_1 x_0}$$

Then the expected value λ, when $x = x_0 + 1$ is

$$\begin{aligned}
\lambda_1 &= e^{\hat{\beta}_0} e^{\hat{\beta}_1 (x_0 + 1)} \\
&= e^{\hat{\beta}_0} e^{\hat{\beta}_1 x_0} e^{\hat{\beta}_1} \\
&= e^{\hat{\beta}_1} e^{\hat{\beta}_0} e^{\hat{\beta}_1 x_0} \\
&= e^{\hat{\beta}_1} \lambda_0
\end{aligned}$$

Interpret the Model Parameters

Now, since

$$\ln(\mathrm{E}(\mathrm{Y}|\mathrm{X}=\mathrm{x})) = \ln(\lambda(x)) = 1.22914 - 0.04981x$$

and

$$\begin{aligned}\frac{\lambda_1 - \lambda_0}{\lambda_0} \times 100\% &= \left(e^{\hat{\beta}_1} - 1\right) \times 100\% \\ &= (e^{-0.04981} - 1) \times 100\% \\ &= (0.95141 - 1) \times 100\% \\ &= -4.859\%\end{aligned}$$

we can say that the mean of Y is 4.859% lower for every one-unit increase in X.

Note that this difference is not statistically significant (p-value $=$0.686).