Binomial distribution: logit, probit and CLL models for the death penalty data

Contents

1. Introduction

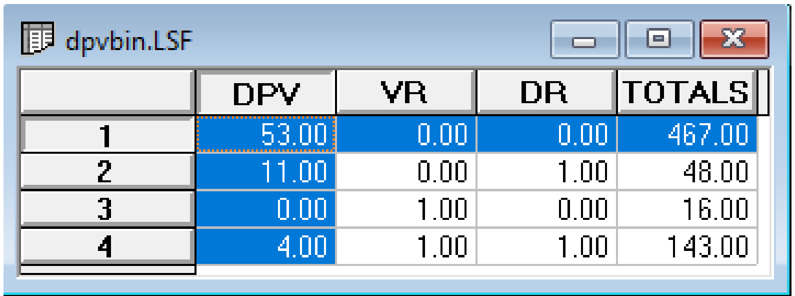
In this example we consider a model with binomial counts as the outcome variable. It quite frequently happens that the predictors of interest are categorical in nature. An example may be the response of residents in favor or against the building of a facility in their neighborhood. In such a case, results are frequently tabulated in a format such as the table below, with success defined as being in favor of the development.

Subgroup	In favor of	Against	Total
Males under 40	59 (y_1)	32 ($n_1 - y_1$)	91 (n_1)
Males above 40	44 (y_2)	47 ($n_2 - y_2$)	91(n_2)
Females under 40	66 (y_3)	20 ($n_3 - y_3$)	86 (n_3)
Females above 40	40 (y_4)	53 ($n_4 - y_4$)	93 (n_4)

Assuming that the random variables y_1 to y_4 are independently distributed with the same π, the proportion of successes in the subgroups can be expressed as $p_i = y_i / n_i$ and $E(y_i) = \pi_i$. These data may be viewed as frequencies for *N* binomial distributions (*N* = 4 in this case).

In this example, we use tabulated data from Radelet and Pierce (1991) that report the number of death penalty verdicts for cases involving multiple murders in Florida during the time period 1976 to 1987. The number of death penalties is given by two additional categorical predictors, ethnicity of the defendant and ethnicity of the victim. We also have information on the total number of verdicts handed down.

These data are given in **Dpvbin.lsf.** The contents of this file are shown below. Data and syntax files can be found in the **MVABOOK\Chapter3** folder.



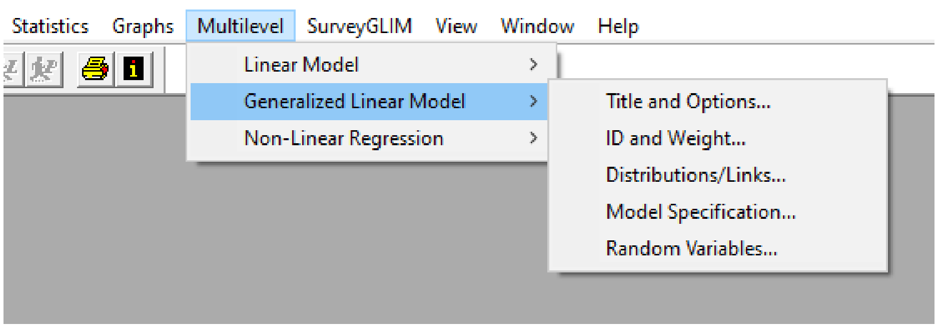
dpvbin.LSF

	DPV	VR	DR	TOTALS
1	53.00	0.00	0.00	467.00
2	11.00	0.00	1.00	48.00
3	0.00	1.00	0.00	16.00
4	4.00	1.00	1.00	143.00

The outcome variable of interest is the variable DPV. The variable VR indicates the victim ethnicity and DR the defendant ethnicity, both assuming the value 1 if white, 0 otherwise. TOTALS represent the total number of verdicts handed down. The focus is on exploring whether the ethnicity of defendant or victim influences the probability of a death penalty being handed down.

2. Logit model

We fit a binomial-logit model to the data using the GLIM (Generalized LInear Model) module in LISREL to do so. This module is accessed via the **Multilevel** option on the main toolbar.



When the options on this menu is compared to those for the **Linear Model** option, we note the additional **Distributions/Links** option. This is because a generalized linear model not only includes the response variable and a linear part consisting of the explanatory variable(s), but also a link function which transforms the mean of the response variable to linear form.

On the **Titles and Options** dialog box, we enter a title and proceed to the **Distributions and Links** tab by using the **Next** button. Here we select Binomial as the **distribution type** and the **link function** field at Logit.

On the **Dependent and Independent Variables** tab DPV is selected as outcome. We also set the **NTRials Variable** to TOTALS. Click next to move to the final tab and then **Finish** to generate the syntax file shown below.

```
GlimOptions;
Title=Binomial Logit Model fo Death Penalty Data;
SY=dpvbin.LSF;
Distribution=BIN;
Link=LOGIT;
Intercept=Yes;
Scale=None;
DepVar=DPV;
CoVars=VR DR;
NTrials=TOTALS;
```

Results are given below.

Goodness of Fit Statistics

Statistic	Value	DF	Ratio
Likelihood Ratio Chi-square	0.3794	1	0.3794
Pearson Chi-square	0.1978	1	0.1978
-2 Log Likelihood Function	418.9565		
Akaike Information Criterion	424.9565		
Schwarz Criterion	423.1154		

Estimated Regression Weights

Parameter	Estimate	Standard Error	z Value	P Value
intcept	-2.0595	0.1458	-14.1208	0.0000
VR	-2.4044	0.6006	-4.0033	0.0001
DR	0.8678	0.3671	2.3641	0.0181

Results are quite similar to that obtained for the same data when fitting a Poisson model: the victim's ethnicity seems to have a bigger influence on the probability of a death penalty being handed down than does the defendant's ethnicity.

3. Probit model

Another model to consider fitting to these data is the probit model. Whereas the logit model uses the cumulative distribution function of the logistic distribution, the probit model uses the cumulative distribution function of the standard normal distribution to define $f(*)$. Both of these functions will take any number and rescale it to fall between 0 and 1.

Subsequently, it can be transformed by the function to yield a predicted probability. It is very similar to logit regression in that they both maximize the same log-likelihood function varying only in terms of the obtained partial derivatives.

Switching from the binomial logit model to the probit is accomplished by changing one line in the previously used syntax.

We set

```
Link=PROBIT;
```

instead of the

```
Link=LOGIT;
```

used previously (see **dvpbin2.prl**). Results for the probit model are given below.

```
                     Goodness of Fit Statistics

    Statistic                                     Value          DF          Ratio
    ---------                                     -----          --          -----
    Likelihood Ratio Chi-square                  0.2669           1         0.2669
    Pearson Chi-square                           0.1399           1         0.1399
    -2 Log Likelihood Function                 418.8440
    Akaike Information Criterion               424.8440
    Schwarz Criterion                          423.0029


                 Estimated Regression Weights

                                      Standard
    Parameter          Estimate         Error       z Value    P Value
    ---------          --------      --------       -------    -------
    intcept             -1.2102        0.0762      -15.8802     0.0000
    VR                  -1.2004        0.2836       -4.2332     0.0000
    DR                   0.4827        0.2092        2.3074     0.0210
```

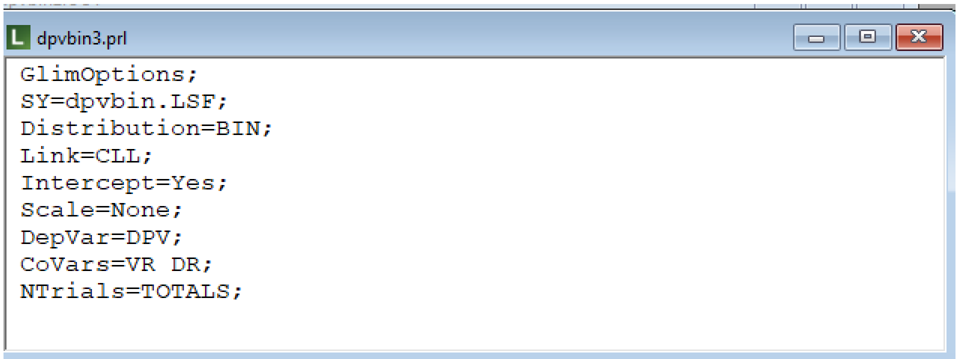
Judging by the -2 $\ln(L)$, AIC and Schwarz criterion, these models are very similar in fit. Again, we see that the victim's ethnicity plays a statistically significant role in the probability of a death penalty being handed down. While the estimated coefficients are different between the model, they differ by a scale factor.

4. Complimentary log-log model

Another model frequently used in the analysis of data such as these is the complimentary log-log link function. The difference between the probit and CLL link functions for the Bernoulli and Binomial distributions are shown in the table below.

Probit	$\eta = \Phi^{-1}(\pi)$	$\pi = \Phi(\eta)$
CLL	$\eta = \log[-\log(1-\pi)]$	$\pi = 1 - e^{-\eta}$

Syntax for the CLL model is given in **dpvbin3.prl**:



```
GlimOptions;
SY=dpvbin.LSF;
Distribution=BIN;
Link=CLL;
Intercept=Yes;
Scale=None;
DepVar=DPV;
CoVars=VR DR;
NTrials=TOTALS;
```

to consider fitting to these data is the probit model. Whereas the logit model uses the cumulative distribution function

Goodness of Fit Statistics

Statistic	Value	DF	Ratio
Likelihood Ratio Chi-square	0.4048	1	0.4048
Pearson Chi-square	0.2110	1	0.2110
-2 Log Likelihood Function	418.9819		
Akaike Information Criterion	424.9819		
Schwarz Criterion	423.1408		

Estimated Regression Weights

Parameter	Estimate	Standard Error	z Value	P Value
intcept	-2.1201	0.1374	-15.4251	0.0000
VR	-2.2843	0.5746	-3.9754	0.0001
DR	0.7919	0.3269	2.4226	0.0154

The estimated coefficients are similar to those obtained for the binomial-logit model. All three the models seem to fit the data well. To decided which of the three is the best, we evaluate the -2 $\ln(L)$ values and conclude that the probit model fits best.

5. Calculating expected probabilities

We can use the results obtained for these three models to calculate the expected probabilities for a death penalty being handed down for each model using the results in the previous sections.

Recall that the estimated regression results for the logit model were:

Parameter	Estimate	Standard Error	z Value	P Value
intcept	-2.0595	0.1458	-14.1208	0.0000
VR	-2.4044	0.6006	-4.0033	0.0001
DR	0.8678	0.3671	2.3641	0.0181

Let

$$a_i = -2.0595 - 2.4044(VR_i) + 0.8678(DR_i)$$

Then the expected probabilities are calculated as

$$\hat{\pi}_i = \frac{e^{a_i}}{1+e^{a_i}}$$

Calculations for the probit regression results are done in a similar way using the results

```
                                  Standard
Parameter          Estimate        Error       z Value    P Value
---------          --------     --------       -------    -------
intcept             -1.2102       0.0762      -15.8802     0.0000
VR                  -1.2004       0.2836       -4.2332     0.0000
DR                   0.4827       0.2092        2.3074     0.0210
```

so that

$$a_i = -1.2102 - 1.2004(VR_i) + 0.4827(DR_i)\,.$$

In the case of the CLL link function, we use the results

```
                                  Standard
Parameter          Estimate        Error       z Value    P Value
---------          --------     --------       -------    -------
intcept             -2.1201       0.1374      -15.4251     0.0000
VR                  -2.2843       0.5746       -3.9754     0.0001
DR                   0.7919       0.3269        2.4226     0.0154
```

to obtain

$$a_i = -2.1201 - 2.2843(VR_i) + 0.7919(DR_i)$$

and subsequently

$$\hat{\pi}_i = 1 - e^{-e^{a_i}}\,.$$

Results are summarized in the table below. The observed probabilities are calculated as the ratio between the death penalties (DPV) and the total number of verdicts handed down (TOTALS) for each of the four cells formed by the cross-tabulation of victim and defendant ethnicity. We note that both the logit and CLL link function results are closer to the observed probabilities than is the case for the probit link.

	observed prob.	**expected prob: logit**	**expected prob: probit**	**expected prob: CLL**
VR = 0	0.1135	0.1131	0.2300	0.1131
VR = 0	0.2292	0.2392	0.1628	0.2328
VR = 1	0	0.0114	0.0349	0.0122
VR = 1	0.0280	0.0267	0.0553	0.0266