November 1, 2016

Converting a logistic model odds ratio to a risk ratio

Joseph M Hilbe

CONVERTING AN ODDS RATIO TO RISK RATIO IN A SINGLE PREDICTOR MODEL.

Joseph M. Hilbe
Jet Propulsion Laboratory, California Institute of Technology
and
Arizona State University
hilbe@asu.edu : 6 May, 2012

In *Logistic Regression Models* (Hilbe, 2009) the text demonstrates that a risk ratio (if *y* or the response is binary), or rate ratio (if *y* are counts) can be modeled using a Poisson or log-binomial regression. Logistic regression produces odds ratios, not risk ratios, although some statisticians tend to interpret odds ratios as if they were risk or rate ratios. Sometimes the respective values are very close to one another, sometimes not. For an example of them differing in a logistic model, we use the *Titanic* data from the book, with *survived* (1/0) as the response (*y*) and *age* (1/0) as the predictor. *age*==1 are adults; *age*==0 are children. The odds ratio of *age* is

LOGISTIC REGRESSION -- ODDS RATIO

```
------------------------------------------------------------------
survived | Odds Ratio Std. Err.    z  P>|z|  [95% Conf. Interval]
---------+--------------------------------------------------------
     age | .4149037    .081936  -4.45  0.000    .2817413  .6110042
   _cons | 1.096154   .2102063   0.48  0.632    .7527297  1.596261
------------------------------------------------------------------
```

Risk ratios, however, may be calculated from the Poisson and log-binomial models as:

POISSON REGRESSION -- RISK RATIO (INCIDENCE RATE RATIO)

```
------------------------------------------------------------------
         |             Robust
survived |       IRR Std. Err.     z    P>|z| [95% Conf. Interval]
---------+--------------------------------------------------------
     age | .5978162  .0580371  -5.30  0.000    .4942328  .7231092
   _cons | .5229358  .0478518  -7.08  0.000    .4370773    .62566
------------------------------------------------------------------
```

LOG-BINOMIAL -- RISK RATIO

```
-------------------------------------------------------------------
         |                EIM
survived | Risk Ratio Std. Err.    z    P>|z| [95% Conf. Interval]
---------+---------------------------------------------------------
     age |  .5978162  .0580239  -5.30  0.000    .4942542  .7230779
   _cons |  .5229358  .0478409  -7.09  0.000    .4370951  .6256346
-------------------------------------------------------------------
```

Children had a 2.4 times greater odds [1/.4149037 = 2.4020] of surviving than did adults. If this were a risk ratio, we could say that children had 1.67 (one and two-thirds) greater probability or likelihood [1/.5978162=1.6702] of surviving than did adults.

To obtain the risk ratio from the odds ratio given in logistic regression output, calculate the predicted value, which is a probability, for *age*==1 and *age*==0. Take the ratio of *mu* if *age*==1 to *mu* if age==0. The result is .5978 -- exactly the value that is given in the Poisson and log-binomial models.

PREDICTED VALUES: SURVIVED==1

```
. predict mu
. mean mu if age==1                          /*  AGE ==1  */

Mean estimation Number of obs = 2092
-----------------------------------------------------
   | Mean Std. Err. [95% Conf. Interval]
---+-------------------------------------------------
mu | .3126195    0        .       .
-----------------------------------------------------

. mean mu if age==0      /*  AGE == 0 ; the Poisson intercept */

Mean estimation Number of obs = 109
-----------------------------------------------------
   | Mean Std. Err. [95% Conf. Interval]
---+-------------------------------------------------
mu | .5229358   0         .        .
-----------------------------------------------------
```

RISK RATIO -- AGE

```
. di .3126195 / .5229358
.59781621                           /* AGE: Rate or Risk Ratio */
```

You may **calculate it by hand** from the table of logistic coefficients

LOGISTIC REGRESSION

```
-------------------------------------------------------------
survived |     Coef. Std. Err.   z  P>|z| [95% Conf. Interval]
---------+---------------------------------------------------
     age | -.8797087 .1974819 -4.45 0.000  -1.266766 -.4926514
   _cons |  .0918075 .1917671  0.48 0.632  -.2840491  .4676642
-------------------------------------------------------------
```

PREDICTED PROBABILITY OF SURVIVED==1 IF AGE==1

```
. di 1/(1+exp(-(_b[_cons] + _b[age]*1)))  = 1/(1+exp(-(.0918075 + (-.8797087) *1)))
.3126195
```

PREDICTED PROBABILITY OF SURVIVED==1 IF AGE==0

```
. di 1/(1+exp(-(_b[_cons] + _b[age]*0)))  = 1/(1+exp(-.0918075))
.52293578
```

which are the same values as before. Dividing we have .5978, which is the probability or risk of adults surviving compared to children. No adjustments are made here for gender of passenger class.

```
. di .3126195 / .5229358
.59781621
```

The risk ratio (0.598) of survival given age (adult vs. child) is quite different from the odds ratio (0.415). It is unwise, therefore, to always interpret logistic model odds ratios as if they were risk ratios