

Models In Epidemiology And Biostatistics

Gordon Hilton Fick

Session 7 : Nonlinearity Example

Consider a study of participants with arthritis. [part of this study data is in Session_7_Example.dta] They were assigned to one of 2 different treatments (treat: A=0, B=1) and then, after an appropriate time, they were asked whether or not their symptoms improved (improve: no=0 yes=1). In addition, the participant's actual age was considered. For our purposes here, we will not consider a comprehensive analysis but rather focus attention on the model below.

$$p = P(\text{Improve})$$

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 T + \beta_2 A + \beta_3 A^2 + \beta_4 TA + \beta_5 TA^2$$

$$\log(\text{OR}) = \beta_1 + \beta_4 A + \beta_5 A^2$$

$$\log(\hat{\text{OR}}) = b_1 + b_4 A + b_5 A^2$$

Since b_5 is negative, we know the parabola will be concave down [an upside down U].

The maximum log OR is estimated to be $b_1 - \frac{b_4^2}{4b_5} = 2.82$

The estimated age with this maximum log OR is $-\frac{b_4}{2b_5} = 46.69$ years

```
. gen a2=age*age
. gen ta =treat*age
. gen ta2=treat*a2
```

```
. logit improve treat age a2 ta ta2
Logistic regression
```

```
Number of obs   =      84
LR chi2(5)      =     22.03
Prob > chi2     =     0.0005
Pseudo R2      =     0.1892
```

```
Log likelihood = -47.209403
```

improve	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
treat	-11.47808	7.332568	-1.57	0.117	-25.84964	2.893493
age	-.2961694	.2187996	-1.35	0.176	-.7250087	.1326698
a2	.0026751	.0022315	1.20	0.231	-.0016985	.0070487
ta	.6128015	.3190886	1.92	0.055	-.0126006	1.238204
ta2	-.0065631	.003285	-2.00	0.046	-.0130016	-.0001247
_cons	6.916652	5.008255	1.38	0.167	-2.899349	16.73265

```
. nlcom -_b[ta]/2/_b[ta2]
```

```
      _nl_1:  -_b[ta]/2/_b[ta2]
```

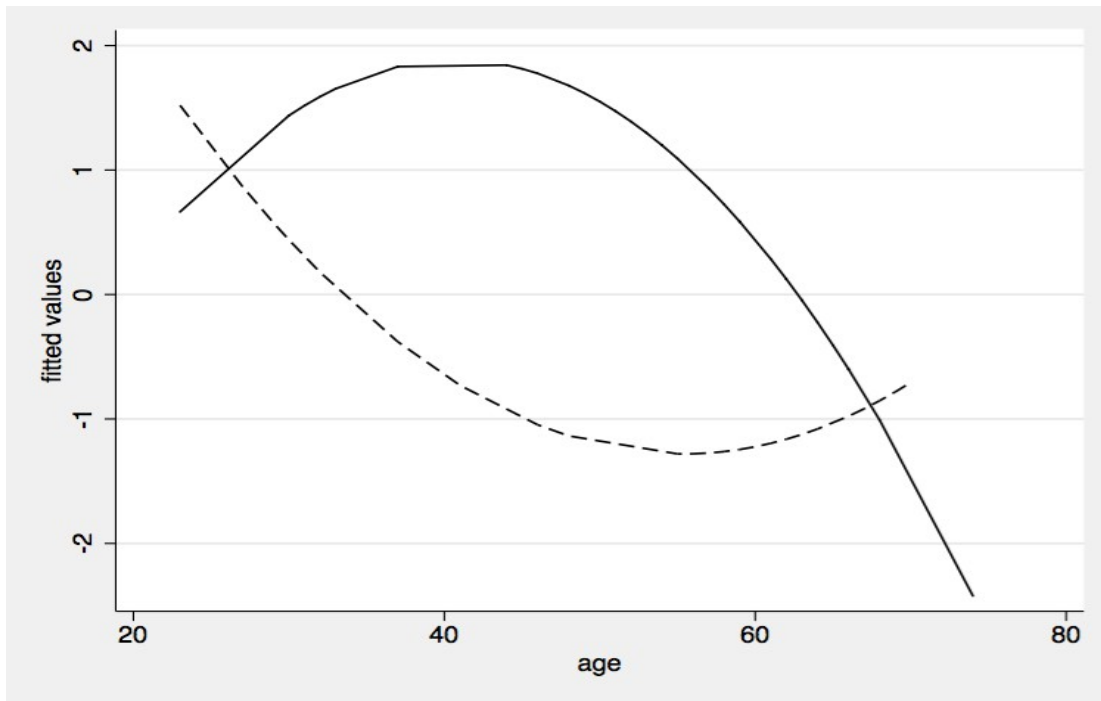
improve	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
_nl_1	46.6852	3.147509	14.83	0.000	40.5162	52.85421

```
. nlcom _b[treat]-_b[ta]*_b[ta]/4/_b[ta2]
```

```
      _nl_1:  _b[treat]-_b[ta]*_b[ta]/4/_b[ta2]
```

improve	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
_nl_1	2.826306	.859669	3.29	0.001	1.141386	4.511227

The estimated log odds versus age :



The estimated log OR versus age

