**Section 4.2.2 – Parameter estimation and inference (continued)**

Model interpretation

We are no longer modeling the log-odds anymore so we will not use odds ratios to interpret the effect of an explanatory variable on the response.

Consider a model with one explanatory variable again:

μ(x) = exp(β0 + β1x)

where I am using “μ(x)” here to emphasize we are evaluating the model as particular numerical value of x. The model evaluated at a c-unit increase in the explanatory variable is

μ(x+c) = exp(β0 + β1(x + c))

If we examine the ratio of the two cases, we obtain:



This leads to a convenient way to interpret the effect of x:

The percentage change in the mean response that results from a c-unit change in x is PC = 

For example, suppose . This leads to PC = 10%.

Comments:

* This interpretation is not dependent on the original value of x!
* Choose a value of c appropriate for the data.
* The estimate of  is

%

* Wald and LR confidence intervals can be found using the usual methods
* If there is more than one explanatory variable in the model, the same result holds (make sure you can show it). You also need to add “holding the other variables in the model constant” to the interpretation.
* If there are interactions or transformations of explanatory variables or categorical explanatory variables, similar types of adjustments need to be made as in the past for odds ratios.

Example: Horseshoe crabs and satellites (Horseshoe.R, Horseshoe.csv)

A c = 1 cm increase in width results in:

> exp(mod.fit$coefficients[2])

 width

1.178267

> 100\*(exp(mod.fit$coefficients[2]) - 1)

 width

17.82674

The percentage change in the estimated mean number of satellites that results from a 1-unit change in width is 17.83%.

To help emphasize the type of change, we could phrase our interpretation as:

* A 1-unit INCREASE in width leads to an estimated 17.83% INCREASE in the mean number of satellites.
* The mean number of satellites is estimated to increase by 17.83% for every 1 cm increase in the width of the shell.

Is c = 1 appropriate? This is difficult to answer without knowing more about the crabs. When there is not an obvious choice for c, you can resort to using one standard deviation:

> c.unit <- sd(crab$width)

> c.unit

[1] 2.109061

> 100\*exp(c.unit\*mod.fit$coefficients[2]) - 1

 width

41.33759

A 2.11 cm increase in width leads to an estimated 41.34% increase in the mean number of satellites.

Profile LR confidence interval with c = 1:

> #Profile likelihood interval

> beta.ci <- confint(object = mod.fit, parm = "width", level = 0.95)

Waiting for profiling to be done...

> 100\*(exp(beta.ci) - 1)

 2.5 % 97.5 %

13.28362 22.50566

> #Profile LR interval using mcprofile

> library(package = mcprofile)

> K <- matrix(data = c(0, 1), nrow = 1, ncol = 2)

> #Calculate -2log(Lambda)

> linear.combo <- mcprofile(object = mod.fit, CM = K)

> ci.beta <- confint(object = linear.combo, level = 0.95)

> 100\*(exp(ci.beta$confint) - 1)

 lower upper

C1 13.28365 22.50525

With 95% confidence, a 1-unit INCREASE in width leads to a 13.3% to 22.5% INCREASE in the mean number of satellites.

Wald confidence interval with c = 1:

> beta.ci <- confint.default(object = mod.fit, parm = "width", level = 0.95)

> beta.ci

 2.5 % 97.5 %

width 0.1249137 0.2031764

> exp(beta.ci)

 2.5 % 97.5 %

width 1.133051 1.225289

> 100\*(exp(beta.ci) - 1)

 2.5 % 97.5 %

width 13.30507 22.52887

> # Calculation without confint.default

> vcov(mod.fit) #Var^(beta^\_1) is in the (2,2) element of

 the matrix

 (Intercept) width

(Intercept) 0.29402590 -0.0107895239

width -0.01078952 0.0003986151

> beta.ci <- mod.fit$coefficients[2] + qnorm(p = c(0.025,

 0.975))\*sqrt(vcov(mod.fit)[2,2])

> 100\*(exp(beta.ci) - 1)

[1] 13.30507 22.52887

> library(package = emmeans)

> calc.est <- emmeans(object = mod.fit, specs = ~ width, at = list(width = c(31, 30)), type = "response")

> test.info <- contrast(object = calc.est, method = "pairwise")

> confint(object = test.info, adjust = "none", level = 0.95)

 contrast ratio SE df asymp.LCL asymp.UCL

 width31 / width30 1.18 0.0235 Inf 1.13 1.23

Confidence level used: 0.95

Intervals are back-transformed from the log scale

Comments:

* The emmeans package does not provide an easy way to get PC exactly. Instead, we can take the information provided and state  with the 95% confidence interval as (13%, 23%).
* I needed a c = 1 unit increase in width for emmeans so I chose two different values of width and included them in the at argument. Because x is not in PC, I could have included other values for width as long as they were 1 unit apart.

**Section 4.2.3 – Categorical explanatory variables**

You can handle these variables in a similar manner as shown in the past! Suppose an explanatory variable has levels of A, B, C, and D. Three indicator variables can be used to represent the explanatory variable in a model:

|  |  |
| --- | --- |
|  | Indicator variables |
| Levels | x1 | x2 | x3 |
| A | 0 | 0 | 0 |
| B | 1 | 0 | 0 |
| C | 0 | 1 | 0 |
| D | 0 | 0 | 1 |

A Poisson regression model with these indicator variables is



Focusing on β1, we have PC =  representing the percentage change in the mean response for level B when compared to level A. Similar comparisons can be made for C to A and D to A. To compare level B to C, we have PC = .