Cumulative Logistic Regression Prediction

J. Elsner, Aug 6, 2019

- I use all cases 1,523 (training & testing)
- I scale the height & velocity variables to speed convergence
- The resulting coefficients are:

$\alpha_0 = .4409$

$\alpha_1 = 3.165$

$\alpha_2 = 4.8816$

$\alpha_3 = 6.766$

$$\boxed{\text{based on } \underline{1,523} \text{ cases}}$$

$\beta_{Height} = .1369$

$\beta_{velocity} = .9182$

$\beta_{STP} = .1052$

$\beta_{TDS} = 1.5345$

- scaled variables in red

$$\text{Heights} = \frac{\text{Height} - \text{mean}(\text{Height})}{\text{sd}(\text{Height})}$$

standard deviation

$\rightarrow$ sd(Height) = scale, mean(Height) = center

$$\text{Heights} \times \text{scale} + \text{center} = \text{Height}$$

Velocitys is done similarly.

- Prediction for a new case i

$$\log\left[\frac{\Pr(EF \le 0)}{1 - \Pr(EF \le 0)}\right] = \alpha_0 - \phi_i$$

(see Eq. 2 Elsner/Schroder)

logit function

where $\phi_i = \beta_{\text{Heights}} \text{Heights}_i + \beta_{\text{Velocitys}} \text{Velocitys}_i + \beta_{STP} \cdot STP_i + \beta_{TDS} \cdot TDS_i$

so $\phi_i = .1369 \text{ Heights}_i +$

$.9182 \text{ Velocitys}_i +$

$.1052 \text{ STP}_i +$

1.5345 TDS_i

$\alpha_0 = .4409$ (also from the model fit)

$$\phi_i = .1369\left(\frac{\text{Height}_i - 3875.443}{2294.394}\right) +$$
$$.9182\left(\frac{\text{Velocity}_i - 37.99127}{14.6302}\right) +$$
$$.1052\ STP_i + 1.5345\ TDS_i$$

$\text{Height}_i = 2000$

$\text{Velocity}_i = 100$

$STP_i = 10$

$TDS_i = 1$

As an example prediction, let

then $\phi_i = 6.3663$

so

$$z_0 = \log\left[\frac{\Pr(EF \leq 0)}{1 - \Pr(EF \leq 0)}\right] = .4409 - 6.3663$$
$$= -5.9254$$

logistic function

and

$$Pr(EF=0) = 1/[1+\exp(-z_0)]$$

$$= 1/[1+\exp(5.9254)]$$

$$= .00266 \qquad .266\%$$

$$z_1 = \underset{(\alpha_1)}{3.165} - 6.3663 = -3.2013$$

$$Pr(EF=1) = 1/[1+\exp(3.2013)]$$

$$= .03912 \qquad 3.9\%$$

$$z_2 = \underset{(\alpha_2)}{4.8816} - 6.3663 = -1.4847$$

$$Pr(EF=2) = .1847 \qquad 18.5\%$$

$$z_3 = \underset{(\alpha_3)}{6.7662} - 6.3663 = +.4$$

$$Pr(EF=3) = .5987 \qquad 60\%$$

$$Pr(EF \geq 4) = 1 - .5987 - .1847 - .03912 - .00266$$

what remains

$$= .175 \qquad 17.5\%$$

Another example:

Let $Velocity_i = 10$ and all else the same as before, then

$$\phi_i = .7178$$

$$z_0 = .4409 - .7178 = -.2769$$

$$Pr(EF=0) = 1/[1+\exp(.2769)] = 43\%$$

$$\vdots$$