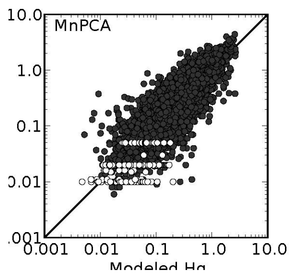
Inverse Mills Ratio

November 9, 2011

When plotting modeled versus measured results in censored regression, we suffer from lines on the plot when "measured" results for censored values are plotted as their detection limit. For example, see the plot of modeled vs. measured results from the initial LOO analysis on the national Fish Mercury model.



Mathematically, we can state an expected value for an uncensored result as

$$E\left[x|x=\alpha\right]=\alpha \tag{1}$$

where x is the observation and α is a reference value we, in this case, *a priori* know x is equal to. In a censored case, we are implicitly assuming

$$E\left[x|x<\alpha\right]=\alpha \tag{2}$$

where x is the observation and α is the detection limit.

We can take advantage of a correction made through the truncated normal distribution in which the expected value is expressed as fluctuation about the mean where the fluctuation is expressed as σ multiplied by the inverse Mills ratio

$$E\left[x|x<\alpha\right]=\mu+\sigma\frac{-\phi\left(\frac{\alpha-\mu}{\sigma}\right)}{\Phi\left(\frac{\alpha-\mu}{\sigma}\right)} \tag{3}$$

where ϕ is the normal PDF and Φ is the normal CDF, calculated as

$$\phi=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\frac{(\alpha-\mu)^2}{\sigma^2}\right) \tag{4}$$

$$\Phi=\frac{1}{2}\left[1+\operatorname{erf}\left(\frac{\alpha-\mu}{\sigma\sqrt{2}}\right)\right] \tag{5}$$

So, our goal is to replace the detection limit value used for plotting the modeled result with a more appropriate expected value. To do this, we can first define the residual as

$$\mathrm{res}=x-y \tag{6}$$

where x is the measured value and y is modeled. Now, we can adapt eq. 3 to correct the residual instead of the value, taking advantage of an expected value of the mean as $\mu=0$ and σ being reported from the model estimation. Then,

$$E\left[\mathrm{res}\right]=E\left[x-y|x-y<\alpha\right]=\sigma\frac{-\phi\left(\frac{\alpha}{\sigma}\right)}{\Phi\left(\frac{\alpha}{\sigma}\right)}=\sigma\frac{-\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\frac{(\alpha)^2}{\sigma^2}\right)}{\frac{1}{2}\left[1+\operatorname{erf}\left(\frac{\alpha}{\sigma\sqrt{2}}\right)\right]} \tag{7}$$

Substituting $E\left[\mathrm{res}\right]$ in eq. 7 for res in eq. 6, we obtain

$$x=E\left[\mathrm{res}\right]+y \tag{8}$$