Ecological Archives E086-041-A4

Brian S. Cade, Barry R. Noon, and Curtis H. Flather. 2005. Quantile regression reveals hidden bias and uncertainty in habitat models. Ecology 86:786-800.

Appendix D: Estimating Local Quantile Weights

We used the Hall and Sheather (1988) bandwidth selection rule recommended by Koenker and Machado (1999) but did not use their approach of taking differences between estimates for the highest and lowest quantile within the bandwidth. Instead, weights were computed by taking the average pairwise difference between all unweighted quantile estimates for $b_0(\tau)$, $b_1(\tau)$, and $b_2(\tau)$ within the interval $\tau \pm h(\tau)$, where $h(\tau)$ was the bandwidth for a specified quantile. This reduced the number of negative weights due to crossing of regression quantile estimates at extreme regions of the design matrix that occurred with the method used by Koenker and Machado (1999). Still, small constants had to be added to the average pairwise differences for $b_0(\tau)$ to assure positive weights for a couple of quantiles.

An example of computations for the quantile interval weights based on a modification of the method proposed by Koenker and Machado (1999) is provided for the 0.90 quantile for the model including bed elevation and bed elevation^2. The Hall and Sheather (1988) bandwidth rule assuming a normal distribution (for convenience) is $h(\tau) = n^{-1/3} z_\alpha^{2/3}[1.5\phi^2(\Phi^{-1}(\tau))/2(\Phi^{-1}(\tau))^2 + 1)]^{1/3}$, where z_α satisfies $\Phi(z_\alpha) = 1 - \alpha/2$, Φ is the cdf and ϕ is the pdf of the standard normal distribution; and for the 0.90 quantile, $\alpha = 0.10$, and $n = 200$ yielded a recommended bandwidth of $h(0.90) = 0.05264$. The estimates $b_0(\tau)$, $b_1(\tau)$, and $b_2(\tau)$ were obtained for all quantiles in the interval $0.90 \pm h(0.90) \in [0.84736, 0.95264]$. This interval contained 22 regression quantile estimates, and the average pairwise difference between them was 81.0003 for $b_0(\tau)$, 56.5343 for $b_1(\tau)$, and 9.98316 for $b_2(\tau)$. Plots of $b_1(\tau)$ and $b_2(\tau)$ by τ were examined to determine the sign of the rates of change to assign to the estimated difference coefficients. For this quantile the weights were $w(0.90) = (2 \times 0.05264)/(81.0031 - 56.5343 \times \text{bed elevation} + 9.9832 \times \text{bed elevation}^2)$. Plots of the weights as a function of bed elevation were examined to check for any negative weights; none occurred for $w(0.90)$. When negative weights were encountered a small constant was added to the denominator of the function to shift them all to positive values while preserving their relative value. The weights were then multiplied by *Macomona* >15 mm counts (y), bed elevation (X_1) and bed elevation^2 (X_2) to estimate the 0.90 quantile regression for the model $w(0.90)y = w(0.90)\beta_0(0.90) + w(0.90)\beta_1(0.90)X_1 + w(0.90)\beta_2(0.90)X_2$ and to compute confidence intervals based on inverting the quantile rank score tests.

Literature cited

Hall, P., and S. Sheather. 1988. On the distribution of a studentized quantile. Journal of the Royal Statistical Society, Ser. B **50**:381-391.

Koenker, R., and J. A. F. Machado. 1999. Goodness of fit and related inference processes for quantile regression. Journal of the American Statistical Association **94**:1296-1310.