# Supporting Material for: Quantile Benchmark Dose Estimation for Continuous Endpoints

Matthew W. Wheeler, Kan Shao,
A. John Bailer

March 9, 2015

## 1 Quantile Smoothing Regression

As discussed in Koenker (2005), the quantile regression estimate minimizes the solution to the check loss function

$$\rho_\tau(c) = \begin{cases} c \times (\tau - 1), & c < 0 \\ c \times \tau, & \text{otherwise}, \end{cases} \tag{1}$$

and this function exists for any $0 < \tau < 1$. As an example, when $\tau = \frac{1}{2}$ the check loss function is proportional to the absolute value function $|c|$. For the absolute value loss, the value that minimizes this function is well known to be the median. For an arbitrary distribution represented by the CDF $F(Y)$, it can be shown that $E\left[\rho_\tau(Y - \theta)\right]$ is minimized when $F(\theta) = \tau$. The minimization problem is a linear programming problem subject to equality constraints. Geometrically, the solution to this linear programming problem is found on the surface of a polyhedron whose vertices are data points; consequently, the solution to the minimization problem can exactly interpolate the observed data.

In our problem, $\theta = \omega_\tau(d) = \alpha_0^\tau + \sum_{q=1}^{Q} \beta_q^\tau b_q(d)$, which is a quantile exposure response function given exposure $d$, and is based upon monotone splines $b_q(d)$. Given responses at various doses $y_1, \ldots, y_n$, the solution, with no smoothing, is found by minimizing

$$\arg\min_{\boldsymbol{\beta}^{\tau}} \sum_{i=1}^{n} \rho_{\tau} \left[ y_i - \omega_{\tau}(d_i) \right]. \tag{2}$$

We add a smoothness penalty of $\lambda \sum_{q=1}^{Q} |\beta_q^{\tau_j} - \beta_{q-1}^{\tau_j}|$, which can be seen as adding additional equality constraints to the minimization problem.

As mentioned in the manuscript, the minimization approach above computes each quantile separately, and it is possible that these estimated quantiles may intersect. We compute the minimum subject to non-crossing constraints (Bondell et al., 2010), which adds more constraints to the linear program. Using this approach, given $\tau_0$ and $\tau_1$, one finds $\omega_{\tau_0}$, and $\omega_{\tau_1}$ with $\omega_{\tau_0}(d) \geq \omega_{\tau_1}(d)$ that satisfies

$$\arg\min_{\boldsymbol{\beta}^{\tau_0}, \boldsymbol{\beta}^{\tau_1}} \sum_{j=0}^{1} \left\{ \sum_{i=1}^{n} \rho_{\tau_j} \left[ y_i - \omega_{\tau_j}(d_i) \right] \right\} \tag{3}$$

Here the doses are rescaled to be on $[0, 1]$ and constraints, which are described in Bondell et al. (2010), are added that force non-crossing of the quantile response curve. These constraints force the quantiles to not to cross; however, as the solution is found on a polyhedron whose verticies are observed data, it is possible that the quantiles are equal, especially when $\tau_0$ and $\tau_1$ are defined to be in the tails of the distribution. From a practical perspective, this behavior is rare and most often occurs during the bootstrap resamples when there are few replications per dose group. As mentioned in the manuscript, it can cause issues in BMD being estimated to be zero especially when $\lambda$ is chosen to be small.

The choice of the smoothing bandwidth parameter $\lambda$ is important when estimating the smoothing spline. For large values of $\lambda$, the estimated quantiles will be flat across the domain, and for small values of $\lambda$, the estimated quantiles will approach that of the unsmoothed estimate. We follow Koenker et al. (1994) and Bondell et al. (2010) and use the Schwartz-type information criterion (SIC) for choosing $\lambda$.

That is, we choose the parameter that minimizes

$$SIC(\lambda) = \sum_{j=0}^{1} log \left\{ n^{-1} \sum_{i=1}^{n} \rho_{\tau_j} \left[ y_i - \hat{\omega}_{\tau_j}^{\lambda}(d) \right] \right\} + (2n)^{-1} log(n) \sum_{j=0}^{1} p_{\tau_j}^{\lambda},$$

where $\hat{\omega}_{\tau_j}^{\lambda}(d)$ is the estimated function given $\lambda$ and $p_{\tau_j}^{\lambda}$ is the number of points interpolated by $\hat{\omega}_{\tau_j}^{\lambda}(d)$. This is the number of points such that the estimate $\hat{\omega}_{\tau_j}^{\lambda}(d)$ correspond to an observed data point.

As mentioned in Koenker (2005), this quantity is essentially an *ad hoc* method for determining the appropriate level of smoothness, and some care is needed for its use. This function is frequently jagged with widely different values of $\lambda$ that produce almost identical SIC values with the method having a tendency to over smooth the data. In our problem, we search for an optimal $\lambda$, over a range of values that does not favor overly smooth curves. This is done by limiting $\lambda$ to be between 0 and 1.5.

## References


Bondell, H. D., Reich, B. J., and Wang, H. (2010). Noncrossing quantile regression curve estimation. *Biometrika* **97,** 825–838.

Koenker, R. (2005). *Quantile regression*. Number 38. Cambridge university press.

Koenker, R., Ng, P., and Portnoy, S. (1994). Quantile smoothing splines. *Biometrika* **81,** 673–680.