Project

March 8, 2014

Problem 1

As suggested by the hint, in order to find the bias at $t = 0.5$, we're going to be examining the value of β as n goes to ∞. The idea behind this is that the proposed model will converge to some β as the number of samples increases, to the point that it fits as closely as possible to the actual function. Of course, it isn't possible for the model to fit precisely, because the actual relationship isn't linear like the model is, so there will be some bias in the resulting model.

The first thing to consider is model construction. The simplest way to construct the model would be to use a least-squares method, so in the construction we would minimize the following with respect to $g(X)$, which is the set of all possible functions for our model.

$$E[(EY - g(X))^2] \tag{1}$$

For this problem, $g(X) = \beta X$ for all possible β, and the actual distribution of EY is $X^{0.75}$, so we can rewrite this to minimizing the following with respect to β

$$E[(X^{0.75} - \beta X)^2] \tag{2}$$

To actually find this value, we will first make a new variable $Q = (X^{0.75} - \beta X)^2$, and then use the continuous case of iterated expectations along X (eq 5.33) to find EQ

$$EQ = \int_{-\infty}^{\infty} f_X(t)\, E(Q|X = t)\, dt \tag{3}$$

$$E(Q|X = t) = E[(t^{0.75} - \beta t)^2] = (t^{0.75} - \beta t)^2 \tag{4}$$

X is $U(0, 1)$, so $f_X(t)$ is 1 between the values of 0 and 1, and is zero elsewhere. With this and the value of $E(Q|X = t)$, we can rewrite EQ as the following.

$$EQ = \int_0^1 (t^{0.75} - \beta t)^2\, dt \tag{5}$$

From here it's a simple matter to minimize with respect to β. First, finish computing the integral to get the final function of β to minimize.

$$EQ = \frac{1}{3}(\beta^2 - \frac{24}{11}\beta + \frac{6}{5}) \tag{6}$$

And then take the derivative of the function with respect to β.

$$\frac{d\,EQ}{d\beta} = \frac{2}{3}\beta - \frac{8}{11} \tag{7}$$

And finally set that equal to zero and solve for β, which yields $\beta = \frac{12}{11}$ as the value that minimizes our function.

Now we can move on and compute the bias. The bias is simply $\hat{m}_{Y;X}(t) - m_{Y;X}(t)$, which for the selected β and $t = 0.5$ is

$$Bias(0.5) = \beta 0.5 - 0.5^{0.75} \approx \boxed{\text{-0.0491}} \tag{8}$$

Problem 2