Tutorial 12
December 4/5, 2008

1. (a) We observe that once the data are given, the sum of the squared residuals is a quadratic function of θ_0 and θ_1. To perform the minimization, we set to zero the partial derivatives with respect to θ_0 and θ_1. We obtain two linear equations in θ_0 and θ_1, which can be solved explicitly, yielding the following regression formulas:

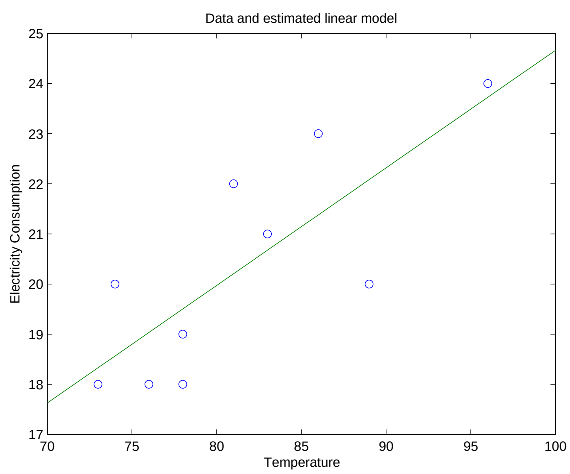
$$\hat{\theta}_1 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2}, \qquad \hat{\theta}_0 = \overline{y} - \hat{\theta}_1 \overline{x},$$

where

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 81.4, \qquad \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = 20.3.$$

The linear regression model is

$$y = 0.23x + 1.21.$$



(b) Using the estimation model above with $x = 90$ we obtain

$$y = 0.23x + 1.21 = 21.91.$$

2. (a) This follows from the direct application of The Weak Law of Large Numbers. Note that the Y_i's are independent.

(b)

$$E[(M_k - \theta)^2] \quad = \quad E[(\frac{1}{k}\sum_{i}^{k} N_i)^2],$$

$$\begin{aligned} &= \frac{\mathrm{var(N_i)}}{k}, \\ &= \frac{\sigma^2}{k}. \end{aligned}$$

(c) The MSE of M_k for this problem is given by,

$$E[(M_k - \theta)^2] = \frac{\sigma^2}{k} \text{ from part(b) }.$$

In this instance we have a uniform distribution over $[-\delta, \delta]$, therefore the variance is $\frac{\delta^2}{3}$.

(d)

$$\theta \in [y_1 - \delta, y_1 + \delta].$$

$$\theta \in [\max\{y_1, y_2\} - \delta, \min\{y_1, y_2\} + \delta].$$

$$\theta \in \bigcap_{i=1}^{K} [y_i - \delta, y_i + \delta].$$

$$\begin{aligned} E_\theta[\hat{\Theta}_k] &= \theta + E[\min_i\{N_i\} + \max_i\{N_i\}] \\ &= \theta \end{aligned}$$

This follows from the fact that $\min_i\{N_i\} = -\max_i\{N_i\}$.

(e)

$$\hat{\Theta}_k - \theta = \frac{1}{2}(\min_i\{N_i\} + \max_i\{N_i\})$$

Let $Z = \min_i\{N_i\}$ and $W = \max_i\{N_i\}$. Then $f_{W,Z}(w,z) = f_{W|Z}(w|z)f_Z(z)$. We calculate the two parts on the right hand side. Since $f_{W|Z}(w|z) = \frac{dF_{W|Z}(w|z)}{dw}$ we first calculate $F_{W|Z}(w|z)$.

$$\begin{aligned} F_{W|Z}(w|z) &= P(W < w | W > z), \\ &= P(z \le N_2 \le w, \cdots, z \le N_k \le w), \\ &= (\frac{w-z}{\delta - z})^{k-1} \end{aligned}$$

Differentiating the above gives $f_{W,Z}(w,z) = (k-1)\frac{(w-z)^{k-2}}{(\delta-z)^{k-1}}$. We now find $F_Z(z)$.

$$\begin{aligned} F_Z(z) &= P(Z < z), \\ &= 1 - P(Z > z), \\ &= P(N_1 \ge z, \cdots, N_k \ge z), \\ &= 1 - (\frac{\delta - z}{2\delta})^k \end{aligned}$$

By differentiating the above CDF we get, $f_Z(z) = k\frac{(\delta-z)^{k-1}}{2\delta^k}$. Finally we have,

$$f_{W,Z}(w,z) = f_{W|Z}(w|z)f_Z(z) = k(k-1)\frac{(w-z)^{k-2}}{(2\delta)^k}.$$

(f) MSE for $\hat{\Theta}_k$, is given by,

$$
\begin{aligned}
E[(\hat{\Theta}_k - \theta)^2] &= \frac{1}{4}E[(w+z)^2] \\
&= \frac{k(k-1)}{4(2\delta)^k}\int_{-\delta}^{\delta}\int_{-\delta}^{w}(w+z)^2(w-z)^{k-2}dzdw \\
&= \frac{2\delta^2}{(k+1)(k+2)}.
\end{aligned}
$$

The MSE for the $\hat{\Theta}_k$ estimator is smaller.