STA 108 HW 1

Author: Dylan Ang
Page 33: Q 1.2, 1.3, 1.4, 1.5, 1.7, 1.8

1.2

The relation is a functional one, since the annual cost points map perfectly on a line.

$$Y = 300 + 2X$$

1.3

Although hardening is a result of a natural chemical process, real world conditions may affect the predicted time and the actual time. In real life, there are many factors that may contribute and affect the hardening in ways that are difficult to account for, such as the exact temperature of the room, external substances getting in the plastic during the molding process, the temperature of the surface it is resting on, and so on. Due to these factors, regression analysis is important since it allows us to develop a working model without having to understand the intricacies of the plastic at a chemical level.

1.4

Here we are seeing the error in the model, represented by epsilon. While the model predicts one value, a point may be off by some amount; this amount is epsilon. In this case we have two epsilons.

1.5

The simple linear regression model is actually

$$Y_i = \beta_0 + \beta_1 + \epsilon_i, \quad i = 1 \ldots n \tag{1}$$

Since β_0, β_1 and X_i are constant values,

$$E(Y_I) = E(\beta_0 + \beta_1 + \epsilon_i) \tag{2}$$
$$= \beta_0 + \beta_1 + E(\epsilon_i) \tag{3}$$
$$= \beta_0 + \beta_1, \quad E(\epsilon_i) = 0 \tag{4}$$

The student could use either equation (1) or (2), but its incorrect to have $E(Y_i)$ and have real ϵ_i; he should use expected epsilon.

1.7

No, it is not possible to predict the exact probability that Y will fall between 195 and 205. Plugging in the given values, $Y = 200 + \epsilon_i$. So, in order for Y to fall between 195 and 205, ϵ_i must be between 5 and -5. Over the whole model, ϵ_i should conform to a normal distribution, but it is impossible to predict exactly what epsilon will be for a specific X value.

1.8

The new Y value should fit into the normal distribution, but it won't necessarily be 108. 108 is the value for a specific single data point, but not every Y_i value at $X = 45$ will be 108. If the new Y value fits into the normal distribution, the linear model shouldn't change and thus $E(Y)$ would stay the same.