Calibrating a snow gauge

This is a *calibration* or an *inverse regression* problem. With 'gain' as the response Y and 'density' as the predictor X, we obtain the fitted response $\hat{Y}$ based on a linear (or linearized) regression model, given the predictor X.

In a typical calibration setting, Y and X are 2 measurements of the same characteristic (e.g., snow density) obtained via 2 different approaches, with X being 'more accurate' than Y. Given an accurate measurement Y_0 (which can be its true mean at X_0 or an average of several measurements to remove measurement bias), what can we say about the true value of the characteristic? We can thus use the fitted regression model of Y on X as a "calibration curve" to predict X_0.

Building the regression model of *Y* on *X*

From the project brief, we have $P(\textit{gamma ray reaches detector}) \approx e^{bx}$, with density x, so that if M is the constant emission rate at each measurement point (i.e., the constant number of gamma rays the source emits during the time a measurement is taken), we have $Y \propto M \times P(\textit{gamma ray reaches detector}) \approx Me^{bx}$. This suggests the following linear regression model of $\log Y$ on X:

$$\log Y = a + bX + error,$$

with $a = \log M$, with the usual normality assumption on the errors.

To fit this model, we can use the first 3 measurements on Y in Table 1. That is, for each X value, we have 3 corresponding Y values. These values can be the first or the last 3 measurements or they can be 3 randomly selected measurements from the 10 Y values. Also, note that we can use more than 3 measurements, as long as we have enough measurements left on Y to average to obtain a relatively accurate Y_0 value.

Calibrating *X*₀

Given Y_0 as the average of the remaining Y values at a given X value, we can use the fitted regression model $\widehat{\log Y} = \hat{a} + \hat{b}X$ to obtain $\hat{X}_0$, for each X value.

We also need to provide respective $100(1-\alpha)\%$ "fiducial" lower and upper limits $\hat{X}_{0L}$ and $\hat{X}_{0U}$ for $\hat{X}_0$, using the $100(1-\alpha)\%$ confidence interval for Y_0. The method for doing this is outlined on pp. 83-86 of Draper and Smith (1998, *Applied Regression Analysis*, 3rd ed.). Note that you will have to modify the formulas a bit, since the discussion assumes Y_0 is the true mean and not the mean of q=7 Y values (or fewer depending on how many measurements you used in fitting the model); see the note on the bottom of p. 85.

In the end, you will have the predicted values $\hat{X}_0$ at all the 9 X values considered in the study.