Accelerated Testing Case Study

Optical media, such as DVDs and CDs, wear out over time depending on use, temperature and humidity. The rate at which they wear out is known by technology specialists to follow what is called the reduced Eyring equation.

$$Y_i = A\exp(BH_i + C/(T_i + 273)) \tag{1}$$

where T_i is temperature and H_i is humidity and Y_i is the number of minutes the optical media lasted before wearing out, as determined by showing a certain number of defects.

Because it would take a long time to determine how long discs are when they fail at normal conditions, they are subjected to extreme conditions and then extrapolated using Equation (1) to normal conditions. Specifically, they are typically tested at specific combinations of temperature (in ° Celsius) and humidity (in % relative humidity) These temperature/humidity combinations are (65°, 85%), (70°, 75%), (85°, 70%), (85°, 85%) The model is then extrapolated to normal conditions of 25° C and 50% humidity.

Each temperature/humidity combination is tested 20 times and then the fitted model is applied to the standard conditions to predict how long optical media will last. In total, 80 data points are used fit the model, 20 runs of each of 4 different temperature/humidity combinations.

Write the answers to these problems in a single file called eyring.*ext* with the accompanying plot described below.

1. Write a function called `eyring_gen` that will generate accelerated testing data from the reduced Eyring equation. the function should take arguments for the parameters A, B, and C a variance component, which should be a Gaussian error added in on the log scale, i.e. $log(Y_i) \sim N(log(A) + BH_i + C/T_i, \sigma)$. The function should return 80 data points at the specific temperature/humidity combinations mentioned above.

2. Use the function written in Problem 1 to simulate data using $A = \exp(-13.68), B = -0.0422$, and $C = 8,485$ and a variance of 0.0225.. Plot the data in a file called "data.pdf". The figure should display minutes to failure, Y, as a function of humidity H and temperature T. Side by side plots is one easy way to do this. Make sure everything is well labeled and looks professional.

3. It is not difficult to a fit a log linear model to this data using $log(Y_i) \sim N(\log(A) + BH_i + C/T_i, \sigma)$ with standard regression equations. In this case, $\boldsymbol{\beta} = (\log(A), B, C)$, where $\log(\mathbf{Y}) = \mathbf{Z} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, where $\mathbf{X}$ has 3 columns, a column of 1's, the second column contains humidity values, $(H_1, ..., H_n)$ and the third column contains inverse temperatures $(1/(T_1 + 273), ..., 1/(T_n + 273))$. Then $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Z}$. Write a function that is given a data set as the input and outputs the fitted values for A, B, and C.

4. Perform a simulation study where data is simulated repeatedly using $A = \exp(-13.68), B = -0.0422$, and $C = 8,485$ and a variance of 0.0225. Fit the model using the procedure created in problem 3. Compare the actual parameters versus the predicted values and assess any bias.

5. Instead of using a log linear model, fit a model that minimizes the squared errors of the predictions for untransformed data,

$$\sum_{i=1}^{n} \left(Y_i - A\exp(BH_i + C/(T_i + 273))\right)^2 .$$

You can use built-in R functions to optimize to find predicted values of A, B, and C. Perform a simulation study to assess the bias in this case.

6. In comments in your program, state why the methods in problems 3 and 5 are not the same and are not guaranteed to provide the same results.

7. As mentioned in the introductory material, a key element of these models is to be able to correctly predict the expected lifetime under normal conditions. Using the values $A = \exp(-13.68), B = -0.0422$, and $C = 8,485$, the expected lifetime under normal conditions of 25°C and 50% relative humidity is $Y_{normal} = 322,273$ minutes. Use a simulation study to determine how well the procedures above will predict this value. In other words, for both the methods, determine estimates for A, B, and C and use those to provide an estimate for Y_{normal}. Evaluate bias as well as variance for the two methods. Is one clearly better than the other?

8. Write yet another function that will input a data set and will provide a bootstrapped confidence interval for Y_{normal} using the minimizing procedure in problem 5.