

Homework 6

1. Let ϵ_{ij} , $i = 1, 2, \dots, I$, $j = 1, 2, \dots, J$ be independent $N(0, \sigma^2)$. Prove that $\sum_i \sum_j (\bar{\epsilon}_i - \bar{\epsilon}_{..})^2$ and $\sum_i \sum_j (\bar{\epsilon}_j - \bar{\epsilon}_{..})^2$ are statistically independent.
2. The “jointed line” (also called “broken stick”) regression model is as follows. (See section 7.2.1 of Faraway.) Suppose that for $x \leq \xi$, $E(Y|X = x) = \alpha_1 + \beta_1 x$, that for $x > \xi$, $E(Y|X = x) = \alpha_2 + \beta_2 x$, and that the regression function is continuous at ξ .
 - (a) Show that the set of all such functions is a linear space — any linear combination of them is such a function.
 - (b) There are a number of basis sets for this space. Show that the following functions are a basis: $f_1(x) = 1$, $f_2(x) = x$, $f_3(x) = (x - \xi)I\{x \geq \xi\}$. Show how the function could be expressed as a linear model with respect to this basis
3. Old Faithful geyser in Yellowstone National Park, Wyoming, derives its name from the regularity of its eruptions. The file `oldfaithful.csv` contains measurements on eight successive days of the durations of the eruptions (in minutes) and the subsequent time interval before the next eruption. The park posts predicted eruption times for visitors. How well can the time until the next eruption be predicted by the duration of the current one?
 - (a) Plot time intervals versus duration of the previous eruption. Regress intervals on duration and plot the residuals. Comment briefly on the fit.
 - (b) Observe that there may be two regimes, corresponding to short and long eruptions, so that perhaps the regression should be segmented as a “broken stick” regression. Choose a breakpoint by eye, fit a segmented regression, and plot the result. Compare to the result of fitting a single regression.
 - (c) Use an F test to compare the segmented regression to the single regression.
 - (d) The analysis above ignores possible temporal trends. Plot the residuals versus time and comment.
 - (e) How could the breakpoint ξ be estimated other than by eye?