

# Homework 4

*Due: 31 March 2016*

## Problem 4.1

The current project ends in another week and its final error model will be a somewhat ad-hoc one where we make up, and perhaps tweak and tune, probabilities for various kinds of keyboard errors.

Imagine for a moment that the project would continue, and that the error model would involve empirically-observed probabilities of some sort. What would that look like? What kinds of data would you want to gather, and how would you compile it into probability distributions (and use them)? Your answer should include both the counted trait and (symbolically) the probability distribution it would represent. Ponder how you might use generic data (gathered how?) that you could preload into the system, and also data specific to the user, that the system would use for ongoing training. If one or the other would not be useful, say why.