As I have explained in the driver program, the assumption for this algorithm is that letters with higher frequency occurrence has the larger probability to win. This algorithm is trying the most accurate way of winning the game. During the process, we need to update the subset all the way, so the tradeoff is that this algorithm runs a little slow. My final outcome for the words successfully being guessed are 3620 words. The total number of words are 4507. So the winning probability could be calculated as 3520/4507=0.803195