Homework: Week 3

1. From Monday, Sep 15

   (a) (This is the exercise on p 26 of your notes) Suppose we have two returns, Q_1 and Q_2 (with $Q_1 > Q_2$). Suppose we take P_1 and P_2 to be from the softmax algorithm,

   $$P_1 = \frac{\exp(Q_1/\tau)}{\exp(Q_1/\tau) + \exp(Q_2/\tau)}$$

   $$P_2 = \frac{\exp(Q_2/\tau)}{\exp(Q_1/\tau) + \exp(Q_2/\tau)}$$

   Take the limit as $\tau \to 0$ and $\tau \to \infty$. (We'll do one together in class, for homework, do them both.

   (b) (This is from the top of p. 28 of your notes) Consider the two methods in class of decreasing some value from a to b in N steps:

   $$f_1(t) = a + \frac{b-a}{N}t$$

   $$f_2(t) = a\left(\frac{b}{a}\right)^{t/N}$$

   If $a = 10$ and $b = 1$, and $N = 15$ get a plot of the two functions. Think about what your domain ought to be, and explain why one of these might work better than the other.

2. From Tuesday, Sep 16

   (a) (These are from the bottom of p. 28 of your notes, with better wording) Suppose we have three estimated returns, Q_1, Q_2, Q_3 and Q_1 is the best of the three. Show that, given associated values of P_1, P_2, P_3, the updated values will still all sum to 1.

   (b) (From p. 29) Suppose that for some fixed machine a, the return $Q_t(a)$ is never the maximum. Show that, by using the update rule, the corresponding probability $P_t(a)$ goes to zero as $t \to \infty$. HINT: Show that

   $$P_t(a) = (1 - \beta)^t P_0(a)$$

   where $P_0(a)$ is the initial probability.

   (c) Make (all) the following changes to `BinStrings.m` and save the resulting file as `BinStringsHW.m`:

      i. Use the same crossover point for all children, rather than randomly selecting it each time.

ii. Use softmax probabilities in place of the probabilities used (set $\tau = 1$).

iii. Reduce the maximum iteration down to about 30.

iv. Cut the initial population in half.

Report on your results. In the plot of the output, does the maximum ever *decrease*? Why do you think that is, and is there a way to guarantee that the maximum is always increasing? You don't need to answer this yet, but think about the *stopping criteria*.