# Problem 1: Optimization and Probability

1. The sequence is finite, so we differentiate to find the minimum value of θ:

If some of the *w* terms are negative in the series, this can create errors with optimization by creating local extrema, which will make the convergent solution of the optimization problem at least partially dependent on initial conditions.

1. We want to compare

Suppose that there is some x’ for which f(x) is maximized. Then, we have

Since the ranges of the summation are equivalent, we then have

However, because both x and the summation range are constant between f and g, it follows that

And so therefore

1. <REVISIT THIS> For a six-sided fair dice, the probability of rolling any given number is 1/6. Under the specified conditions, the expected number of points accumulated by the time a 1 is rolled is:
2. To optimize L(p) at some p=pmax, we suppose that dL/dp = 0 for some value p = pmax. Taking the logarithm of this function, we see that the optimizing condition becomes

Therefore, the value of p which maximizes L(p) must also maximize L’(p). Therefore, we take the natural logarithm of both sides. We have

Differentiating both sides to find the optimizing value of p yields

This value of p represents the optimal “fairness” of the coin to provide the maximum total probability of obtaining the event sequence {H,H,T,H,T,T,H} in a series of 7 tosses. I.e. a coin which has a 4/7 chance of landing heads will maximize the probability of obtaining said sequence.

1. <REVISIT THIS> Differentiating, we have

The second term reduces to 0 for all components of the vector w that are not wk and 2wk otherwise, so