# Statistical Rethinking: Chapter 6 - Overfitting, Regularization, and Information Criteria

## Chris Hayduk

### January 27, 2019

## 1 Easy

**Problem 6E1.**
State the three motivating criteria that define information entropy.

The three motivating criteria that define information entropy are:

1. The measure of uncertainty should be continuous. If it were not, then an arbitrarily small change in any of the probabilities would result in a massive change in uncertainty.

2. The measure of uncertainty should increase as the number of possible events increases. For example, suppose there are two cities that need weather forecasts. In the first city, it rains on half of the days in the year and is sunny on the others. In the second, it rains, shines, and hails, each on 1 out of every 3 days in the year. We'd like our measure of uncertainty to be larger in the second city, where there is one more kind of event to predict.

3. The measure of uncertainty should be additive. What this means is that if we first measure the uncertainty about rain or shine (2 possible events) and then the uncertainty about hot or cold (2 different possible events), the uncertainty over the four combinations of these events - rain/hot, rain/cold, shine/hot, shine/cold - should be the sum of the separate uncertainties.

**Problem 6E2.**
Suppose a coin is weighted such that, when it is tossed and lands on a table, it comes up heads 70% of the time. What is the entropy of this coin?

$$H(p) = -Elog(p_i) = -\sum_{i=1}^{n} p_i log(p_i) = -(0.7 * log(0.7) + 0.3 * log(0.3)) \approx 0.610864$$

**Problem 6E3.**
Suppose a four-sided die is loaded such that, when tossed onto a table, it shows "1" 20%, "2" 25%, "3" 25%, and "4" 30% of the time. What is the entropy of this die?

$$H(p) = -Elog(p_i) = -\sum_{i=1}^{n} p_i log(p_i)$$
$$= -(0.2 * log(0.2) + 0.25 * log(0.25) + 0.25 * log(0.25) + 0.3 * log(0.3)) \approx 1.37623$$

**Problem 6E4.**
Suppose another four-sided die is loaded such that it never shows "4". The other three sides show equally often. What is the entropy of this die?

$$H(p) = -Elog(p_i) = -\sum_{i=1}^{n} p_i log(p_i)$$
$$= -(0.3333 * log(0.3333) + 0.3333 * log(0.3333) + 0.3333 * log(0.3333)) \approx 1.09498$$

# 2 Medium

**Problem 6M1.**
Write down and compare the definitions of AIC, DIC, and WAIC. Which of these criteria is most general? Which assumptions are required to transform a more general criterion into a less general one?