

# Complexity and Diversity in Low Probability States

Johannes Castner

February 20, 2015

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Models and Belief Systems.</b>	<b>2</b>
<b>3</b>	<b>Priors</b>	<b>2</b>

## Abstract

## 1 Introduction

## 2 Models and Belief Systems.

## 3 Priors

Suppose that at any period,  $t$ , it is known to a decision maker that the system relevant to her decision can be in one of  $K$  potential states. Further, suppose that at period  $t$ ,  $K - r_t$  of those have been experienced at least once, so that  $r_t$  of them have never before been experienced. What prior is a person to optimally place on a never before experienced stimulus? If  $r_t = 0$ , the optimal prior seems trivial: all states have been experienced at least once and it seems that the optimal model should use simple induction. Let  $S = s_1, \dots, s_K$  denote the set of possible states and  $F_t = f_{1,t-1}, \dots, f_{K,t-1}$  the associated observed frequencies of these states before time  $t$ . Then if  $r_t = 0$ , the optimal prior probabilities of the  $K$  states at time  $t$  seem to be  $\{\pi_{i,t}^*\} = \left\{\frac{f_{i,t-1}}{\sum_i f_{i,t-1}}\right\} = \left\{\frac{f_{i,t-1}}{t-1}\right\}$ . However, induction is not optimal if one wants to avoid expressing complete certainty about the impossibility of not yet experienced events.

The fix is to start in period 0, before any sensory experience begins. If an individual is told that there are, say eight possible events that could happen, what likelihood should she attribute to each of the eight events? In the *absence* of information it seems hopeless to place any probabilities on the eight events. But if one wants to be maximally open minded, there is a cure: maximum entropy. In other words, in period 0, set  $\pi_{i,0}^* = \frac{1}{8}$ , for  $i = 1, \dots, 8$ . More generally, for  $K$  states at time  $t$ , the optimal prior probabilities should then be

$$\pi_{i,t}^* = \frac{f_{i,t-1} + \frac{1}{K}}{t}. \quad (1)$$

Surely, this quantity converges to the right limit (there is no bias) and it does so at the highest rate possible, given the information inherent in the data sequence up to period  $t$ . When  $t$  tends to infinity the probability of a state occurring that has up to now not been observed tends asymptotically to zero ( $\lim_{t \rightarrow \infty} \frac{1}{K * t} = 0$ ). In the beginning we are told that swans can be either black or white. We start with an open mind and attribute equal

probability to both, black and white swans. The longer we fail to make a sighting of a black swan, while continuously sighting white ones, the less likely we deem the creature's existence, but we never rule it out completely (only asymptotically).