CHAPTER 6. THE DUAL TO THE MAXIMUM LIKELIHOOD ESTIMATOR

KULLBACK-LEIBLER INFORMATION (ENTROPY)

Before turning to the dual of the maximum likelihood estimator we define the Kullback-Leibler information, and prove a few of its simple properties. The goal of this detour is to provide a natural probabilistic interpretation for this dual as the minimum entropy expectation parameter.

6.1 Definitions

Suppose F, G are two probability distributions with densities f, g relative to some dominating σ-finite measure ν. The *Kullback-Leibler information* of G at F is

$$(1) \qquad K(F, G) = E_F(\ln(f(x)/g(x)))$$

with the convention that $\infty \cdot 0 = 0$, $0/0 = 1$, and $y/0 = \infty$ for $y > 0$. K is also referred to as the *entropy* of G at F.

It can easily be verified that K(F, G) is independent of the choice of dominating measure ν. The existence of K will be established in Lemma 6.2 where it is shown that $0 \le K \le \infty$.

In exponential families it is convenient to write

$$(2) \qquad K(\theta_0, \theta_1) = K(P_{\theta_0}, P_{\theta_1}), \qquad \theta_0, \theta_1 \in N \quad .$$

For $S \subset N$ let

$$(3) \qquad K(S, \theta_1) = \inf\{K(\theta_0, \theta_1): \theta_0 \in S\} \quad ,$$

etc.