CHAPTER 3. THE LIKELIHOOD PRINCIPLE AND GENERALIZATIONS

3.1 INTRODUCTION

The LP deals with situations in which X has a density $f_\theta(x)$ (with respect to some measure ν) for all $\theta \in \Theta$. Of crucial importance is the *likelihood function for* θ *given* x, given by

$$\ell_x(\theta) = f_\theta(x), \tag{3.1.1}$$

i.e., the density evaluated at the observed value X = x and considered as a function of θ. Often we will call $\ell_x(\theta)$ the *likelihood function for* θ or simply the *likelihood function*. The LP, which follows, is stated in a form suitable for easy initial understanding; certain implicit qualifications are discussed at the end of the section.

THE LIKELIHOOD PRINCIPLE. All the information about θ *obtainable from an experiment is contained in the likelihood function for* θ *given* x. *Two likelihood functions for* θ *(from the same or different experiments) contain the same information about* θ *if they are proportional to one another.*

It has been known since Fisher (1925, 1934) that the "random" likelihood function $\ell_X(\theta)$ is a minimal sufficient statistic for θ, and hence contains all information about θ from a classical viewpoint. The LP goes considerably farther, however, maintaining that only $\ell_x(\theta)$ for the actual observation X = x is relevant.