The Problem of Fiducial Volume in The Light of Information Theory and Model Selection

XU Benda

March 6, 2014

1 Introduction

x, y are parameters defined on a unit square U $[0,1] \times [0,1]$. A kind of event f obeys a probability density function (thereafter PDF) $f(x)$, which is known up to a parameter μ. Another kind of event g distributes as $g(y)$, which is not known *a priori*.

A sample of N events are taken in a certain period of time in a subset S of U. N obeys a poisson distribution:

$$N(S) \sim Pois(\iint_S \mathrm{d}x\mathrm{d}y[\lambda f(x) + \sigma g(y)])$$

λ and σ are defined as the *intensity* of f and g, with units of S^{-1}.

2 Binning

The sample N(S) can be binned into $\{N(S_i)\}$, such that $N(S) = \sum_i N(S_i)$ and $\{S_i\}$ is a partition of S.

3 Question

Given a total sample N(U), what is the best knowledge of λ (the intensity of f) and μ (the shape of f)?

4 Reduced Stituation and Possible Strategies

4.1 $g(y) \equiv 0$

Only $f(x)$ exists, λ is best estimated by N and μ by maximum likelihood estimator(thereafter MLE):

$$\mathcal{L}(\mu|x_1, x_2, \ldots, x_N) = \prod_i f(x_i|\mu)$$

4.2 Ignore g

Even if g(y) is non-zero, it could be integrated out and treated as a constant in the MLE:

$$\mathcal{L}(\lambda, \mu, \sigma|x_1, x_2, \ldots, x_N) = \prod_i [\lambda f(x_i|\mu) + \sigma], \quad \sigma := \int \sigma g(y)\mathrm{d}y$$

4.3 Select a Subset of y

If $g(y) \equiv 0$ in a subset of y, U can be partitioned into $S_1 : g(S_1) \equiv 0$ and $S_2 : g(S_2) \neq 0$. The problem is can be reduced into the above two situations.

Generally, if we insist to use MLE, the partition of U can influence the outcome of λ and μ significantly.

In some physics experiment, y is the radius of some detector. The process of *fiducial volume selection* is to select out a subset R of U where g(y) is small enough for further analysis, usually MLE. This process is mostly carried out manually and *ad-hoc*. If g(y) is monotonic, R can be searched by *figure-of-merit*, i.e. apply MLE for each possible R until the inferred uncertainty of λ and μ is minimal.

5 Open Questions

Is MLE the best estimator in all the cases? if so, is there a routine to deduce the best partition of U (i.e. binning of N)?

If not, what is the best estimator and how to construct it? Or does there exist a best estimator after all? How to deduce an upper limit of the information of λ and μ to be inferenced from N? Can partitioned MLE asymptotically approximate the upper limit?

What if g and f both depends on x and y, be it known or unknown?