Re: MLE Related Proof

Source: http://sci.tech−archive.net/Archive/sci.stat.edu/2006−10/msg00075.html

- *From*: "Brenneman" <brennemt@xxxxxxxxxx>
- *Date*: 23 Oct 2006 08:42:14 −0700

Hi Julian,

No not too late at all: this is something I want to understand, but I am willing to believe that it's actually true (;>)) and move on at the same time. In the end, I just want to get a good understanding of why it works.

You mentioned the regularity conditions previously as maybe being the answer, but they are extremely general, and I can't see any such soln to this problem using them. They are
#1) The pdfs are distinct (i.e. q ne q' imples f(x;q) ne f(x;q') for all x)
#2) The true parameter value lies in the interior of the set of all possible param values
#3) The pdfs all have the same common support.

In one book I came across (Wasan, "Point Estimation"), the Lindeberg−Levy thm is invoked to get the limiting distribution of the log of the likelihood ratios. I don't know if this is very useful to understanding what is going though, since I get the impression that the Lindeberg Levy thm is a way of proving the CLT.

Thank you once again,

Matt

J W wrote:

> Not sure if this is finding you too late, but...
>
>> I am confused by what the random variables _are_. Are C&H talking about the random variable
>> log [f(x1,...,xn;q)/f(x1,...,xn;q0)] ?
>
> It might be easier to see if you write f(X1,...Xn) to denote the fact that f represents the (joint) distribution of the random variables X1,

... Xn.

Using this notation, you can rewrite
log [f(X_1, ..., X_n;q) / f(X_1,...,X_n;q0)
=
sum { log [f(X1;q) / f(X1;q0)] + ... + log [f(Xn;q) / f(X1;q0)] }

since the Xi's are independent. And each term in the sum is iid. Nice!

> How can they know that this
> random variable has a finite mean (since its distribution is not obvious)?

Indeed, it is not in general true that log [f(Xi;q) / f(Xi;q0)] for an arbitrary pdf f (imagine that f is defined by the parameter theta via P({X = theta} = 1); then f(X;q) / f(Xi;q0) is always 0 if the Xi come from f with parameter theta = q0, and log(0) is −infinity. My guess is that the regularity conditions which you briefly mentioned at the top of your question either state (or imply) that the expectation of each term is finite.

The more general result is that the sum converges to the Kullback−Leibler distance (http://en.wikipedia.org/wiki/Kullback−Leibler_divergence) between probability measures defined by f(X;q) and f(X;q0); and indeed, the KL distance (without further assumptions) can be +infinity.

Hope this helps.

−Julian

.