The research for this work was partially sponsored by the National Science Foundation under Grant No. GU-2059 and the United States Air Force under Contract No. AFOSR-68-1415.

MAXIMUM LIKELIHOOD ESTIMATION OF A UNIMODAL DENSITY, II

by

Edward J. Wegman

Department of Statistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 647

NOVEMBER 1969
1. Introduction. Several authors, Grenander [3], Robertson [6], and Rao [4], have described the MLE for a unimodal density when the mode was known as well as some of the estimate's properties. The MLE for a unimodal density when the mode is unknown was described in [7]. Strong consistency was also established in [7]. We wish to describe some additional properties in this paper.

2. Asymptotic Distribution. Let \( \hat{\theta}_n \) be the maximum likelihood estimate with unknown mode and \( \hat{\theta}^*_n \) the maximum likelihood estimate with known mode. In defining \( \hat{\theta}_n \), \( \epsilon > 0 \) was a predetermined number. Let \( y_1 < y_2 < \ldots < y_n \) be the ordered observations sampled according to the density \( f \) and let \( A_1 = \{y_1, y_2\}, A_2 = \{y_2, y_3\}, \ldots, A_{\ell(n)} = \{y_{\ell(n)}, L_n\}, \)

\( A_{\ell(n)+1} = \{L_n, R_n\}, A_{\ell(n)+2} = \{R_n, y_{r(n)}\}, \ldots, A_k = \{y_{n-1}, y_n\} \). Here \( R_n - L_n = \epsilon \); the sequences \( \{L_n\} \) and \( \{R_n\} \) converge to \( L \) and \( R \) respectively; and \( y_{\ell(n)} \) and \( y_{r(n)} \) are respectively the largest observation smaller than \( L_n \) and the smallest observation larger than \( R_n \). \( L_n \) and \( R_n \) are determined by the maximum likelihood procedure and at least one of \( L_n \) or \( R_n \) is an observation for each \( n \). If \( L([L_n, R_n]) \) is the \( \sigma \)-lattice

The research for this work was partially sponsored by the National Science Foundation under Grant No. GU-2059 and the United States Air Force under Contract No. AFOSR-68-1415.
of intervals containing \([L_n, R_n]\), the maximum likelihood estimate, \(\hat{f}_n\), is given by the conditional expectation, \(E(\hat{g}_n | L([L_n, R_n]))\), where

\[
\hat{g}_n = \sum_{i=1}^{k} n_i \cdot [n\lambda(A_i)]^{-1} \cdot I_{A_i}.
\]

Here \(n_i\) is the number of observations in \(A_i\), \(\lambda\) is Lebesgue measure and \(I_{A_i}\) is the indicator of \(A_i\).

In a similar manner, let \(A_i^* = [y_1, y_2], \ldots, A_q(n) = [y_q(n), M), A_{q(n)+1} = [M, y_{q(n)} + L], A_{q(n)+2} = (y_{q(n)} + 1, y_{q(n)} + 2], \ldots, A_n = (y_{n-1}, y_n]\). Here \(M\) is the known mode and \(y_q(n)\) is the largest observation smaller than \(M\). Notice with probability one, \(M \neq y_j\) for each \(j\). If \(L(M)\) is the \(\sigma\)-lattice of intervals containing \(M\), the maximum likelihood estimate, \(f_n^*\), is given by \(E(g_n^* | L(M))\) where

\[
g_n^* = \sum_{i=1}^{n} n_i^* \cdot [n\lambda(A_i^*)]^{-1} \cdot I_{A_i^*}.
\]

Of course, \(n_i^*\) is the number of observations in \(A_i^*\). In [7], it is shown that \(M \in (L, R)\), hence \(\hat{g}_n\) and \(g_n^*\) agree except possibly on \([y_{L(n)}, y_{R(n)}]\). A similar situation was the case in Lemma 5.4 in [7].

If we require only that some neighborhood of \(L\), say \(N_L\), is a set of points of increase of \(f\) and similarly some neighborhood of \(R\), say \(N_R\), is a set of points of decrease of \(f\), we may use the arguments of Lemma 5.4 in [7] to obtain

**Lemma 2.1.** Let \(n > 0\) be an arbitrary number such that \(L-n\) and \(R+n\) are elements of \(N_L\) and \(N_R\) respectively. Then with probability one, for sufficiently large \(n\), \(\hat{f}_n\) and \(f_n^*\) agree on \((L-n, R+n)^c\).
Hence, for any \( x \notin [L,R] \), for sufficiently large \( n \), \( \hat{f}_n(x) = f_n^*(x) \). An immediate theorem follows

**Theorem 2.1:** For \( x \notin [L,R] \), \( \hat{f}_n(x) \) has the same asymptotic distribution as \( f_n^*(x) \).

Rao [4] through some very clever, but rather tedious arguments develops the asymptotic distribution of \( f_n^*(x) \). Arguments similar to these could be applied to \( \hat{f}_n(x) \), but are avoided by use of Lemma 2.1. Rao assumes a non-zero derivative of the density, \( f \), at each point \( x \) where the asymptotic distribution is to be found.

3. **A Characterization of \( \hat{f}_n \).** Reid (see [1] and [2]) gave a geometrical interpretation of a conditional expectation with respect to a \( \sigma \)-lattice, \( L \), when \( L \) consists of intervals with the right (or left) endpoint fixed. If the \( \sigma \)-lattice is \( L(M) \), the conditional expectation may be characterized by applying Reid's method individually to the right and to the left of \( M \). To find \( E(h|L(M)) \), the conditional expectation of some function \( h \) with respect to \( L(M) \), determine

\[
H(x) = \int_{(-\infty,x]} h d\lambda.
\]

To the left of \( M \), \( E(h|L(M)) \) is given by the slope of the greatest convex minorant of \( H \) and to the right of \( M \), by the slope of the least concave majorant of \( H \).

Let us assume that \( h \) has bounded support, \( \{x: h(x) \neq 0\} \). Let \( L \) and \( R \) be fixed with \( R-L = \epsilon \). We want a geometrical interpretation of the conditional expectation of \( h \) with respect to \( L([L,R]) \). Robertson
[5] gives a representation of conditional expectations. This representation holds on finite measure spaces, hence the requirement that \( h \) has bounded support. Let \( E(h|L)(x) = y \) and \( P_y = \{x: E(h|L)(x) > y\} \). Let \( H = \{L^* \in L: \lambda(L^*-P_y) > 0\} \). Then
\[
y = \sup_{L^* \in H} \left[\lambda(L^*-P_y)\right]^{-1} \cdot \int_{L^*-P_y} h \, d\lambda.
\]
Letting \( L = L([L,R]) \) and \( x \in [L,R] \), it is clear that \( P_y \) is empty, so that
\[
y = \sup_{L^* \in H} \left(\lambda(L^*)\right)^{-1} \cdot \int_{L^*} h \, d\lambda.
\]
In fact, this supremum is a maximum and \( H \equiv L([L,R]) \). If \( L^* \) is the maximizing interval, for all \( x \in L^* \), \( E(h|L([L,R]))(x) = \lambda(L^*)^{-1} \cdot \int_{L^*} h \, d\lambda \). Let \( a = \inf L^* \) and \( b = \sup L^* \). As in the case of the conditional expectation with respect to \( L(M) \), it is not difficult to see we may apply Reid's method individually to the left of \( a \) and to the right of \( b \). Thus we have,

**Theorem 3.1:** The conditional expectation of a function, \( h \), with bounded support, with respect to a \( \sigma \)-lattice, \( L([L,R]) \), is given by the following procedure.

Find the interval \([a,b]\) containing \([L,R]\) such that \((H(b) - H(a))/(b-a)\) is maximized. On \([a,b]\), the conditional expectation is given by \((H(b) - H(a))/(b-a)\). To the left of \( a \), it is the slope of the greatest convex minorant of \( H \) and to the right of \( b \), it is the slope of the least concave majorant of \( H \).

If \( h = \hat{g}_n \), the theorem applies since the support of \( \hat{g}_n \) is \([y_1,y_n]\). Let \( \hat{g}_n(x) = \int_{(-\infty,x]} \hat{g}_n \, d\lambda \).
Corollary 3.1. If \( h = \frac{1}{n} \) in Theorem 3.1, \( \frac{1}{n} \) may be replaced by \( \frac{1}{F_n} \), the empirical distribution function.

The proof is straightforward and is left to the reader. It is interesting to note that Theorem 3.1 implies Theorem 3.1 of [7] if the condition of \( f \) being continuous is exchanged for \( f \) having bounded support. The author is indebted to the referee of [7] for pointing out this fact.

References.


Maximum Likelihood Estimation of a Unimodal Density, II

by

Edward J. Wegman

Abstract. This paper is a sequel to the earlier paper, "Maximum Likelihood Estimation of a Unimodal Density Function." The MLE of a unimodal density with unknown mode is shown to agree, for sufficiently large $n$ and on certain regions, with the MLE of a unimodal density with known mode. The asymptotic distributions of the MLE's then agree. Also a geometrical interpretation of the MLE of a unimodal density with unknown mode is given.

Keywords: Maximum Likelihood Estimation
Unimodal Density
Asymptotic Distribution
Geometric Characterization
Mode

Subject Classification Numbers:
62.15
62.70
60.05