A REMARK ON THE r-th MEAN DIFFERENTIABILITY

GEORGE D. STAMATELOS

Department of Mathematics
University of Patras - Greece

(Received January 11, 1983 and in revised form May 9, 1983)

ABSTRACT. This paper is concerned with the r-th mean differentiability. In the mathematical developments regarding the asymptotic expansion and the asymptotic distribution of the likelihood function, there arises the question whether the assumptions made on the model imply differentiability in the r'-th mean of the underlying random functions, for integer values r'<r. The present paper provides an answer to this question and also gives the explicit form of the derivatives in the r'-th mean involved.

KEY WORDS AND PHRASES. Stochastic r-recess, derivative of r-th mean, probability measure.

1980 AMS SUBJECT CLASSIFICATION CODE. 60G99

1. INTRODUCTION AND SUMMARY.

For n>0, integer, let X_0, X_1, ..., X_n be the first n+1 r.v.'s from a stochastic process, defined on the probability space (X, A, P_0), whose probability law depends on a k-dimensional parameter \theta \in \Theta, an open subset of \mathbb{R}^k, k>1 and satisfies certain regularity conditions. Then, for \theta, \theta^{*} \in \Theta, let

\[ q_j(\theta; \theta^{*}) = q_j(X_j; \theta, \theta^{*}) = \frac{dP_{\theta}}{dP_{\theta^{*}}}, \quad n \geq 1, \]

be specified versions of the Radon-Nikodym derivatives involved, where X_j \sim (X_0, X_1, ..., X_n) and P_0, \theta is the restriction of P_0 to the \sigma-field \mathcal{A}_n=\sigma(X_0, X_1, ..., X_n) induced by the r-v.'s X_0, X_1, ..., X_n. Also, for any two parameter points \theta and \theta^{*}, let q_j(X_j | X_{j-1}; \theta, \theta^{*}) be the quotient of the probability density functions of the random vectors \bar{X}_j=(X_0, X_1, ..., X_j) and \bar{X}_{j-1}. Then, as is well known, for large sample statistical inference, based on the likelihood function, the underlying conditions are set primarily on the quantities, q_j(X_j | X_{j-1}; \theta, \theta^{*}), j=1, 2, ..., n. Typically these conditions include pointwise differentiability with respect to the "moving" parameter \theta^{*} of up to third order. It is also known, however, that these conditions fail to be satisfied, for
instance, in such a simple and interesting case as that of the double exponential distribution.

In relatively recent asymptotic statistical work, the above mentioned type of conditions have been replaced by conditions referring to the differentiability in the r-th mean of the quantities

$$\varphi_j(\theta, \theta^*; r) = \left[ q_j (X_j | X_{j-1} ; \theta, \theta^*) \right]^{1/r}, \quad j \geq 1,$$  

for some integer r \geq 1. (1.1)

These conditions do not suffer from the kind of inadequacies mentioned above and, in addition, are of probabilistic rather than of analytical nature. (See, for example, LeCam [1,2], Johnson and Roussas [3-5], Roussas [6], Lind and Roussas [7-8] Akritas [9] and, in particular, Akritas and Roussas [10].)

2. ASSUMPTIONS AND MAIN RESULT.

In this section, the assumptions are formulated under which the main result of this note holds true. To this end, let \( \Theta \) be an open subset of \( \mathbb{R}^k \), \( k \geq 1 \), and for each \( \theta \in \Theta \), let \( X_0, X_1, \ldots \) be r.v.'s defined on the probability space \((X, A, P_\theta)\) and taking values in \((S, S)\); here \( S \) is a Borel subset of a Euclidean space and \( S \) is the \( \sigma \)-field of Borel subsets of \( S \). These r.v.'s come from a certain class of stochastic processes which satisfy suitable conditions to be explicitly mentioned below. Let \( A_n \) be the \( \sigma \)-field induced by the first \( n+1 \) r.v.'s, \( A_n = \sigma(X_0, X_1, \ldots, X_n) \), and let \( P_{n, \theta} \) be the restriction of \( P_\theta \) to \( A_n \); that is,

$$P_{n, \theta} = P_\theta | A_n, \quad A_n = \sigma(X_0, X_1, \ldots, X_n).$$  

It will be assumed in the following that, for each \( \varepsilon, \theta^* \in \Theta \), \( P_{n, \theta} \approx P_{n, \theta^*} \) for all \( n \geq 0 \). Thus, the quantities \( \varphi_j(\theta, \theta^*; r), j \geq 1 \), are well defined by (1.1). Then we have

Assumptions

(A1) For each \( n \geq 0 \), the probability measures \( \{P_{n, \theta}; \theta \in \Theta\} \), defined by (2.1), are mutually absolutely continuous.

For \( \theta, \theta^* \) in \( \Theta \), define \( \varphi_j(\theta, \theta^*; r) \) by (1.1)

Then

(A2) For each \( \theta \in \Theta \), the random functions \( \varphi_j(\theta, \cdot; r) \) are differentiable in \( [0, \infty) \) in the r-th mean at \( \theta \), uniformly in \( j \geq 1 \), for some integer \( r \geq 1 \).

Let \( \tilde{\varphi}_j(\theta; r), j \geq 1 \), be the r-th mean derivatives involved. Then

(A3) For each \( \theta \in \Theta \) and each \( h \in \mathbb{R}^k \),

$$E_{\theta} |h^\top \tilde{\varphi}_j(\theta; r)|^r \leq M_r(\theta, h) (\infty), \quad j \geq 1.$$  

We may now formulate the main result of this paper, namely,

Theorem 2.1. Let the number \( r_k \) be defined by

$$r_k = \frac{r}{r_k}, \quad k = r-1, \ldots, \frac{2r}{3}, \quad \frac{r}{2}.$$  

(2.2)
and let the random functions $\varphi_j(\theta, \cdots; r_k)$, $j \geq 1$, be defined by (1.1) with $r$ being replaced by $r_k$. Then, for each $\theta \in \Theta$ and under assumptions (A1)-(A3), these random functions are differentiable in the $r_k$-th mean $[P_{\theta}]$ at $\theta$, uniformly in $j \geq 1$. The $r_k$-th mean derivative $\hat{\varphi}_j(\theta; r_k)$ is given by

$$\hat{\varphi}_j(\theta; r_k) = (r-k)\hat{\varphi}_j(\theta; r)$$

and

$$E_{\theta}|h \hat{\varphi}_j(\theta; r_k)|^{r_k} \leq M_{r_k}(\theta, h) (<<\infty), \quad j \geq 1.$$  

3. PROOF OF MAIN RESULT.

In the course of the proof of Theorem 2.1, the following auxiliary result will be needed which is formulated here as a lemma. Namely,

Lemma 3.1 For $n$, $j \geq 1$ integers, let $X_j$ and $X_{nj}$ be r.v.'s defined on the probability space $(\Omega, F, P)$ and suppose that, for some $r > 0$,

$$X_{nj} \xrightarrow{(r)} X_j, \text{ uniformly in } j \geq 1.$$ 

Then

$$E|X_{nj}|^r \rightarrow E|X_j|^r, \text{ uniformly in } j \geq 1.$$ 

Proof. For the two cases $0 < r \leq 1$ and $r > 1$, use the $c_r$-inequality and the Minkowski inequality, respectively (see, for example, Loève [11] pages 1955-1956), in order to obtain

$$\frac{1}{r} |E|X_{nj}|^r - E|X_j|^r| \leq E|X_{nj} - X_j|^r + E|X_{nj} - X_j|^r.$$ 

Since the right hand side above tends to 0, uniformly in $j \geq 1$, so does the left hand side. The proof is completed.

We may now proceed with the proof of the main result, namely,

Proof of Theorem 2.1. To show that, uniformly on bounded sets of $h$ and as $(0<)\lambda \rightarrow 0$,

$$E_{\theta}|\frac{1}{\lambda} \varphi_j(\theta, \theta + \lambda h; r_k) - 1| - h \hat{\varphi}_j(\theta; r_k)|^{r_k} \rightarrow C, \text{ uniformly in } j \geq 1,$$

and that relation (2.4) holds.

Suppose for a moment that (2.3) holds true. Then the fact that $r_k \leq r$ implies that

$$E_{\theta}|h \hat{\varphi}_j(\theta; r_k)|^{r_k} \leq (r-k)^r E_{\theta} |h \hat{\varphi}_j(\theta; r)|^r.$$ 

Thus, (2.4) is satisfied by means of assumptions (A3).

In the remaining part of the proof of this theorem, all convergences will be taken as above, that is, for $(0<)\lambda \rightarrow 0$ and uniformly on boun-
ded sets of \( h \). Relation (3.1) holds true for \( r_k = r \) by assumption (A2). Assume it to hold true for some \( r_k = 1, 2, \ldots, r - 2 \) and establish it for \( r_k - 1 = r_k - \delta \) where \( \delta \) is defined by the relation

\[
\delta = \frac{(r_k - 1)^2}{k}.
\]

To this end, consider the relation

\[
\varphi_j(\theta, \theta + \lambda h; r_k - 1) = \varphi_j(\theta, \theta + \lambda h; r_k - \delta) = \varphi_j(\theta, \theta + \lambda h; r_k) \varphi_j(0, \theta + \lambda h; r_k - \delta)
\]

(5.2)

which follows from (1.1). Then by means of (3.2) and (2.3), one has the following identity

\[
\frac{1}{\lambda} \left[ \varphi_j(\theta, \theta + \lambda h; r_k - 1) - h \hat{\varphi}_j(\theta; r_k - 1) \right] = \frac{1}{\lambda} \left[ \varphi_j(\theta, \theta + \lambda h; r_k - \delta) - h \hat{\varphi}_j(\theta; r_k - \delta) \right] = \left\{ \frac{1}{\lambda} \left[ \varphi_j(\theta, \theta + \lambda h; r_k) - h \hat{\varphi}_j(\theta; r_k) \right] \right\} + \left\{ \frac{1}{\lambda} \left[ \varphi_j(\theta, \theta + \lambda h; r_k - \delta) - h \hat{\varphi}_j(\theta; r_k - \delta) \right] \right\} + \left\{ \frac{1}{\lambda} \left[ \varphi_j(\theta, \theta + \lambda h; r_k - 1) - \varphi_j(0, \theta + \lambda h; r_k - \delta) \right] \right\}.
\]

(3.5)

Hence the \( c_r \)-inequality gives

\[
E_\theta \left| \frac{1}{\lambda} \left[ \varphi_j(\theta, \theta + \lambda h; r_k - 1) - h \hat{\varphi}_j(\theta; r_k - 1) \right] \right|^{r_k - 1} \leq c^2 \frac{r_k - 1}{r_k} E_\theta \left| \frac{1}{\lambda} \left[ \varphi_j(\theta, \theta + \lambda h; r_k) - h \hat{\varphi}_j(\theta; r_k) \right] \right|^{r_k - 1} + c^2 \frac{r_k - 1}{r_k} E_\theta \left| \frac{1}{\lambda} \left[ \varphi_j(\theta, \theta + \lambda h; r_k - \delta) - h \hat{\varphi}_j(\theta; r_k - \delta) \right] \right|^{r_k - 1} + c^2 \frac{r_k - 1}{r_k} E_\theta \left| \frac{1}{\lambda} \left[ \varphi_j(\theta, \theta + \lambda h; r_k - 1) - \varphi_j(0, \theta + \lambda h; r_k - \delta) \right] \right|^{r_k - 1} \right|^{r_k - 1} = I_1(\lambda, j) + I_2(\lambda, j) + I_3(\lambda, j),
\]

(5.4)

where the quantities \( I_i(\lambda, j) \) stand for the \( i \)-th term, \( i = 1, 2, 3 \) on the right hand side of relation (3.4). Since for all values of \( k \) for which \( r_k \) is integer \( r_k - 1 < r_k < r \),

\[
I_1(\lambda, j) \leq c^2 \frac{(r_k - 1)^2}{r_k} E_\theta \left| \frac{1}{\lambda} \left[ \varphi_j(\theta, \theta + \lambda h; r_k) - h \hat{\varphi}_j(\theta; r_k) \right] \right|^{r_k - 1} \right|^{r_k - 1} \right| \leq 0,
\]

(5.5)

uniformly in \( j > 1 \), by induction hypothesis.
Remark on the \( r \)-th mean differentiability

Next,
\[
I_2(\lambda, j) \leq c \frac{2^{\delta(r_k^{-1})/r}}{r_k} \left| \frac{1}{\lambda} \left[ \varphi_j(\theta, \theta + \lambda h; \frac{r_k}{6}) - 1 \right] - h \varphi_j(\theta; r_k)^{r/6} \right| + 0, \tag{3.6}
\]
uniformly in \( j \geq 1 \), by hypothesis (A2).

Finally to the term \( I_3(\lambda, j) \) apply the H"older inequality with
\[
s = \frac{r}{k} \quad \text{and} \quad t = \frac{r}{r-k}
\]
in order to get
\[
I_3(\lambda, j) \leq c \frac{k/r}{r_k} \left| \frac{1}{\lambda} \left[ \varphi_j(\theta, \theta + \lambda h; r_k) - 1 \right] \right| \left( \frac{r_k}{r-k} \right) \frac{r}{r-k} \tag{3.7}
\]
By the induction hypothesis and Lemma 3.1,
\[
E_{\theta} \left| \frac{1}{\lambda} \left[ \varphi_j(\theta, \theta + \lambda h; r_k) - 1 \right] \right| \frac{r_k}{r-k} \tag{3.8}
\]
uniformly in \( j \geq 1 \).

Also, by assumptions (A2), (A3) and Lemma 3.1,
\[
E_{\theta} \left| \varphi_j(\theta, \theta + \lambda h; \frac{r}{6}) - 1 \right| \frac{r}{6} \tag{3.9}
\]
Relations (3.7)-(3.9) imply that \( I_3(\lambda, j) \to 0 \), uniformly in \( j \geq 1 \).

Hence, from (3.5), (3.6) and (3.9), relation (3.4) gives
\[
E_{\theta} \left| \frac{1}{\lambda} \left[ \varphi_j(\theta, \theta + \lambda h; r_k - 1) - 1 \right] - h \varphi_j(\theta; r_k - 1) \right| \frac{r_k - 1}{r} \to 0,
\]
uniformly in \( j \geq 1 \), which completes the proof of the theorem.

Acknowledgement: The author wishes to express his sincere thanks to referee for his useful comments.

References


Submit your manuscripts at http://www.hindawi.com