The Annals of Mathematical Statistics
1970, Vol. 41, No. 3, 851–864

ASYMPTOTIC EXPANSIONS ASSOCIATED WITH POSTERIOR DISTRIBUTIONS[1]

BY RICHARD A. JOHNSON

University of Wisconsin

1. Introduction and summary. In this paper, an extension of the investigation of Johnson (1967b) is made by giving a larger class of posterior distributions which possess asymptotic expansions having a normal distribution as a leading term. Asymptotic expansions for the related normalizing transformation and percentiles are also presented.

Before asymptotic expansions were treated rigorously, LaPlace (1847) gave an expansion for certain posterior distributions. The method used in this paper is a variation of his technique. Bernstein (1934), page 406, and von Mises (1964), chapter VIII, Section C, also treat special cases of these expansions.

The conditions imposed are sufficient to make the maximum likelihood estimate strongly consistent and asymptotically normal. They also include higher order derivative assumptions on the log of the likelihood. As shown by Schwartz (1966), the posterior distribution may behave well even when the maximum likelihood estimate does not. However, we have not attempted to find the weakest assumptions under which the posterior distribution has an expansion. For general conditions under which the posterior distribution converges in variation to a normal distribution with probability one see LeCam (1953) and (1958) for the independent case and Kallianpur and Borwanker (1968) for Markov processes.

In Section 2, we show that with probability one, the centered and scaled posterior distribution possesses an asymptotic expansion in powers of $n^{-\frac{1}{2}}$ having the standard normal as a leading term. The number of terms in the expansion obtained is two less than the number of continuous derivatives of the log likelihood. All terms beyond the first consist of a polynomial multiplied by the standard normal density. The coefficients of the polynomial depend on the prior density ρ and the likelihood. The moments of the posterior distribution are shown to possess an expansion in Section 3. The following two sections present the normalizing transformation and percentile expansions. These last three expansions also apply for the case considered by Johnson (1967b) as does the information on the form of the terms in the expansion of the posterior distribution. To simplify the already heavy notation, these results are first proved for independent identically distributed random variables. The extension of all these results to the case of certain stationary ergodic Markov processes is immediate; Section 6 presents the necessary modifications.

Throughout this paper, Φ and φ will denote the standard normal cdf and pdf respectively. Also, n will be assumed to range over the positive integers; thus in

Received March 19, 1968; revised December 19, 1968.

[1] This research was supported in part by the National Science Foundation under Grant GP-6519.

some cases, the order of the error term in the expansion may be kept for smaller n if the bounding constant is modified.

2. Expansion of the posterior distribution. Let $X_1, X_2, \cdots$ be a sequence of real-valued random variables independently distributed as $f(x, \theta_0)$ where $f(x, \theta)$ is a family of densities with respect to a dominating measure μ. It is assumed that θ_0 was chosen according to a law $\rho(\cdot)$ which has a density with respect to Lebesgue measure.

We shall require certain assumptions to hold throughout Sections 2–5.

ASSUMPTION 1. Θ is an open interval of the real line. Let $\bar{\Theta}$ denote a closed interval containing Θ.

ASSUMPTION 2. For any $\theta \in \Theta$ and $t \in \bar{\Theta}$, provided $t \neq \theta$,

$$\int_{-\infty}^{\infty} |f(x, \theta) - f(x, t)| \, d\mu > 0.$$

ASSUMPTION 3. $f(x, \theta)$ is jointly measurable in (x, θ).

ASSUMPTION 4. For each $x, f(x, \theta)$ admits partial derivatives of the first and second order with respect to θ and these are continuous in θ for $\theta \in \bar{\Theta}$.

ASSUMPTION 5. The measures $\prod_{i=1}^{n} f(x_i, \theta)$ are mutually absolutely continuous for each $n = 1, 2, \cdots$. Therefore, a null set will have probability zero for all θ. Assumption 4 and Assumption 5 assure that $\log f(x, \theta)$ is well defined and continous in θ.

ASSUMPTION 6. If $\lim_{i \to \infty} |\theta_i| = \infty$, then $\lim_{i \to \infty} f(x, \theta_i) = 0$ for all x except for perhaps a null set not depending on the sequence $\{\theta_i\}$.

ASSUMPTION 7. For all $\theta \in \bar{\Theta}$,

$$E_\theta |\log f(X, \theta)| < \infty$$

$$0 < I(\theta) \equiv -E_\theta\left[\frac{\partial^2}{\partial \theta^2} \log f(X, \theta)\right].$$

ASSUMPTION 8. For each $\theta_0 \in \Theta$, there exist functions $G_1(x)$ and $G_2(x)$ satisfying

$$\left|\frac{\partial}{\partial \theta} \log f(x, \theta)\right| \leqq G_1(x) \qquad \left|\frac{\partial^2}{\partial \theta^2} \log f(x, \theta)\right| \leqq G_2(x)$$

for θ in a neighborhood of θ_0 and also $E_{\theta_0}[G_1(X)] < \infty$ and $E_{\theta_0}[G_2(X)] < \infty$. The functions G_1 and G_2 may depend on θ_0.

ASSUMPTION 9. Let

$$f(x, \theta, \rho) = \sup_{|\theta - \theta'| \leqq \rho} f(x, \theta') \qquad \rho > 0, \quad \text{and}$$

$$Q(x, r) = \sup_{|\theta| > r} f(x, \theta) \qquad r > 0.$$

For every $\theta \in \bar{\Theta}$ and $\rho, r > 0, f(x, \theta, \rho)$ and $Q(x, r)$ are measurable functions of x. Moreover, for sufficiently small ρ and sufficiently large r,

$$E_{\theta_0}[\log f(x, \theta, \rho)]^+ < \infty$$
$$E_{\theta_0}[\log Q(x, r)]^+ < \infty \qquad \text{for each } \theta_0 \in \Theta.$$

These conditions are just one set of the many variations that imply that the maximum likelihood estimate $\hat{\theta}$ is strongly consistent and that $n^{\frac{1}{2}}(\hat{\theta}-\theta_0)$ is asymptotically normal with variance $1/I(\theta_0)$. The consistency conditions are essentially those of Wald while the other conditions are similar to those employed by Wolfowitz (1965) and Weiss–Wolfowitz (1966).

Denoting the maximum likelihood estimate of θ by $\hat{\theta}$, we study the behavior of the posterior distribution of the centered and scaled variable ϕ where

$$\phi = (\theta - \hat{\theta})b(\hat{\theta}) \tag{2.1}$$

and

$$b(\hat{\theta}) = \left[-\frac{1}{n} \sum_{i=1}^{n} \frac{\partial^2}{\partial \theta^2} \log f(x_i, \theta) \big|_{\theta=\hat{\theta}} \right]^{\frac{1}{2}} \tag{2.2}$$

when the prior distribution has pdf $\rho(\theta)$ with respect to Lebesgue measure.

The notation $b(\hat{\theta})$ suppresses the fact that b is the sum of n terms which depend on the observed sequence $\mathbf{x} = (x_1, x_2, \cdots, x_n, \cdots)$.

Denote the posterior cdf of $n^{\frac{1}{2}} \phi$ by F_n. The observations $x_1, x_2, \cdots, x_n$ enter F_n both from the posterior distribution of θ and the centering and scaling constants. The following theorem establishes the existence of an asymptotic expansion for F_n.

THEOREM 2.1. *Let $\theta_0 \in \Theta$ be fixed. In addition to Assumptions* 1 *through* 9, *let K be an integer and assume that, for each x,* $\log f(x, \theta)$ *had* $K+3$ *continuous partial derivatives with respect to θ. Let there exist functions $G_k(x)$ with $E_{\theta_0}[G_k(X)] < \infty$ and $|(\partial^k/\partial\theta^k)\log f(x, \theta)| \leq G_k(x)$ for θ in a neighborhood of θ_0 $k = 3, 4, \cdots, K+3$.*

If $\rho(\theta_0) > 0$ and $\rho(\cdot)$ has $K+1$ continuous derivatives in a neighborhood of θ_0, there exist functions $\{\gamma_j(\xi, \mathbf{x})\}$ and constants D_1 and N_x such that

$$\left|F_n(\xi) - \Phi(\xi) - \textstyle\sum_{j=1}^{K} \gamma_j(\xi, \mathbf{x}) n^{-j/2}\right| \leq D_1 n^{-\frac{1}{2}(K+1)} \qquad \text{for all} \quad n > N_x$$

on an almost sure set $(\pi f(x, \theta_0))$. Here D_1 depends on K and N_x depends on K and the observed sequence $\mathbf{x}$.

If $\rho(\cdot)$ is positive and sufficiently smooth and the derivatives of $\log f$ are dominated, an expansion will exist for every $\theta_0 \in \Theta$.

Information is available on the form of the γ_j.

PROPOSITION 2.1. *Under the Assumptions of Theorem* 2.1, *each $\gamma_j(\xi, \mathbf{x})$ is a polynomial in ξ having coefficients bounded in x multiplied by the standard normal density.*

REMARK 1. The proof below is also valid for the expansion appearing in Theorem 1.1 of Johnson (1967b). It extends Johnson's results since only the first few terms were shown to be polynomials multiplied by $\varphi(\xi)$. Also ρ need only have $K+1$ derivatives for an expansion of $K+1$ terms.

REMARK 2. Specializing these results to the exponential family $C(\phi)\exp(\phi R(x))$, we see that Theorem 2.1 allows us to work with the usual parameter, for instance p in the binomial instead of $\log[p/(1-p)]$, as was necessary under the formulation in Johnson (1967b).

REMARK 3. Since the expansion is uniform in ξ, it may be used to find the rate at which the posterior distribution of $\theta-\hat{\theta}$ approaches the degenerate distribution. In particular, $P[-\varepsilon \leqq \theta-\hat{\theta} \leqq \varepsilon] = F_n(\varepsilon n^{-\frac{1}{2}}b^{-1})-F_n(-\varepsilon n^{-\frac{1}{2}}b^{-1})$.

A proof will be constructed using the lemmas below and will be patterned after the development in Johnson (1967a, 1967b).

2.1 *Preliminaries*. The posterior pdf of ϕ, when viewed as a function of ϕ, is proportional to

$$\rho(\hat{\theta}+\phi b^{-1})\prod_{i=1}^n[f(x_i, \hat{\theta}+\phi b^{-1})/f(x_i, \hat{\theta})] \tag{2.3}$$

where $b = b(\hat{\theta})$ is given by (2.2). Since θ_0 must be an interior point, in the remainder of this section, it will be assumed that selected θ intervals about θ_0 lie entirely within Θ. Further, without loss of generality, the intervals will be taken small enough so that $|(\partial^k/\partial\theta^k)\log f(x, \theta)| \leqq G_k(x)$ holds for $k = 1, 2, \cdots, K+3$. Before establishing certain properties of the function (2.3), we state a version of the uniform strong law which will be employed repeatedly. See LeCam (1953) or Rubin (1956) for the proof.

THEOREM 2.2. *Let C be a compact set and let $u(x, t)$ be a real-valued function measurable in x for each $t \in C$ and continuous in t for each x. Let $H(x)$ satisfy*

$$|u(x, t)| \leqq H(x) \qquad \int H(x)\,dF(x) < \infty.$$

If the X_i are independent F,

$$P[\lim_{n\to\infty}\sup_{t\in C}|\textstyle\sum_{i=1}^n u(X_i, t)n^{-1}-\int u(x, t)\,dF(x)| = 0] = 1.$$

Let

$$a_{2n}(\theta) = \tfrac{1}{2}n^{-1}\sum_{i=1}^{n}\frac{\partial^2}{\partial\theta^2}\log f(x, \theta). \tag{2.4}$$

The next result is well known, but the proof is included for completeness.

LEMMA 2.1. *There exist a null set $S_1{}^c$, $\delta_1 > 0$ and for each $\mathbf{x}\in S_1$ an $N_{1\mathbf{x}}$ such that*

$$|I(\theta_0)+2a_{2n}(\theta)| \leqq I(\theta_0)/2 \quad \textit{all} \quad |\theta-\theta_0| \leqq \delta_1 \quad \textit{when} \quad n > N_{1\mathbf{x}}.$$

PROOF. Since Θ is open, Assumption 8 assures the existence of a $\delta_1 > 0$ such that $|E_{\theta_0}[\partial^2/\partial\theta^2)\log f(X, \theta)]+I(\theta_0)| \leqq I(\theta_0)/3$ for $|\theta-\theta_0| \leqq \delta_1$. An application of Theorem 2.2 to $a_{2n}(\theta)$ gives the desired result.

Noting that $b(\hat{\theta}) = [-2a_{2n}(\hat{\theta})]^{\frac{1}{2}}$, the last result together with the strong consistency of $\hat{\theta}$ provides an $N_{2x}(> N_{1x})$ for each $\mathbf{x} \in S_2$ such that $|\hat{\theta} - \theta_0| \leqq \frac{1}{2}\delta_1$

$$(2.5) \qquad 0 < [I(\theta_0)/2]^{\frac{1}{2}} \leqq b(\hat{\theta}) \leqq [3I(\theta_0)/2]^{\frac{1}{2}} < \infty \qquad \text{for} \quad n > N_{2x}$$

where S_2^c is a null set containing $\{\hat{\theta} \nrightarrow \theta_0\}$.

2.2. *Basic lemmas.* We now turn our attention to the product term in (2.3).

LEMMA 2.2. *There exists a* $\delta_2 (0 < \delta_2 < 1)$ *with*

$$n^{-1} \log \{\prod_{i=1}^n [f(x_i, \hat{\theta} + \phi b^{-1})/f(x_i, \theta)]\} \leqq -\phi^2/6 \quad \textit{for} \quad |\phi| \leqq \delta_2$$

when $n > N_{2x}$ *each* $\mathbf{x} \in S_1 \cap S_2$.

PROOF. For each $\mathbf{x} \in S_1 \cap S_2$, choose δ_2 so that $|\hat{\theta} + \phi b^{-1} - \theta_0| \leqq \delta_1$ when $|\phi| \leqq \delta_2$ and $n > N_{2x}$. Here δ_1 appears in the statement of Lemma 2.1. An application of the extended mean value theorem together with the bound (2.5) yields the desired result.

The next approximation shows that the posterior mass outside of a ϕ-interval may be neglected. A similar argument is given in LeCam (1953), Theorem 5b.

LEMMA 2.3. *There exists an* $\varepsilon > 0$, *a null set* S_3^c *with* $S_3 \subset S_1 \cap S_2$ *and for each* $\mathbf{x} \in S_3$, *an* N_{3x} *such that*

$$\log \{\prod_{i=1}^n [f(x_i, \hat{\theta} + \phi b^{-1})/f(x_i, \hat{\theta})]\} \leqq -n\varepsilon \quad \textit{for} \quad |\phi| \geqq \delta_2$$

whenever $n > N_{3x}$. *This* δ_2 *appears in the statement of Lemma* 2.2.

PROOF. Let $\mathbf{x} \in S_2$ be fixed. According to (2.5), $|\phi b^{-1}| \geqq \delta_2 [3I(\theta_0)/2]^{-\frac{1}{2}}$ for $|\phi| \geqq \delta_2$ whenever $n > N_{2x}$. Let $2\varepsilon_1 = \delta_2 [3I(\theta_0)/2]^{-\frac{1}{2}}$ and make n larger if necessary so that $|\hat{\theta} - \theta_0| \leqq \varepsilon_1$. Then

$$(2.6) \qquad \begin{aligned} &\prod_{i=1}^n \{f(x_i, \hat{\theta} + \phi b^{-1})/f(x_i, \hat{\theta})\} \\ &\qquad \leqq \sup_{|\theta - \theta_0| \geqq \varepsilon_1} [\prod_{i=1}^n f(x_i, \theta)]/[\prod_{i=1}^n f(x_i, \theta_0)]. \end{aligned}$$

According to the argument given by Wald (1949), page 599–600, the right-hand side of (2.6) is bounded by an expression of the form

$$(2.7) \qquad \{[\sum_{j=1}^J \prod_{i=1}^n f(x_i, \theta_j, \rho_{\theta,j}) + \prod_{i=1}^n Q(x_i, r_0)]/\prod_{i=1}^n f(x_i, \theta_0)\}$$

where $\max_j E_{\theta_0} \log [f(X, \theta_j, \rho_{\theta,j}]/f(X, \theta_0)]$ and $E_{\theta_0} \log [Q(X, r_0)/f(X, \theta_0)]$ are negative. By Assumption 9, an application of the strong law of large numbers to the log of each term in (2.7) establishes the existence of an $\varepsilon > 0$ and a set S_3 included in $S_1 \cap S_2$ such that

$$n^{-1} \log [\sup_{|\theta - \theta_0| \geqq \varepsilon_1} \{\prod_{i=1}^n f(x_i, \theta)\}/\{\prod_{i=1}^n f(x_i, \theta_0)\}] \leqq -\varepsilon \qquad n \geqq N_{3x}$$

for some N_{3x} depending on $\mathbf{x} \in S_3$.

Let us introduce some further notation useful in the derivation. Set

$$(2.8) \qquad a_{kn}(\theta) = n^{-1} \sum_{i=1}^n \partial^k/\partial\theta^k \log f(x_i, \theta)/k! \qquad \text{for} \quad k = 3, 4, \cdots$$

where the dependence on $\mathbf{x}$ is suppressed. In this notation, the log of the likelihood ratio under sufficient regularity conditions may be expressed as

(2.9) $\log \prod_{i=1}^{n} [f(x_i, \theta)/f(x_i, \hat{\theta})] = n \sum_{i=2}^{k} a_{in}(\hat{\theta})(\theta-\hat{\theta})^i + n a_{k+1,n}(\theta_*)(\theta-\hat{\theta})^{k+1}$

where θ_* is an intermediate value. Denote the first $k+1$ terms of the Taylor expansion of $\rho(\cdot)$ about $\hat{\theta}$ by $\rho_k(\hat{\theta})$. That is

(2.10) $$\rho_k(\hat{\theta}) = \rho(\hat{\theta}) + \rho^{(1)}(\hat{\theta})(\theta-\hat{\theta}) + \cdots + \rho^{(k)}(\hat{\theta})(\theta-\hat{\theta})^k/k!$$

where $\rho^{(j)}$ denotes the jth derivative of ρ.

In terms of the transformed parameter ϕ, we establish the following lemma.

LEMMA 2.4. *Let the integer K be given. For each x, let $\log f(x, \theta)$ have $K+3$ continuous partial derivatives in θ. Let $\theta_0 \in \Theta$ be fixed and assume that there exist functions $G_k(x)$ satisfying $|\partial^k/\partial\theta^k \log f(x, \theta)| \leqq G_k(x)$, for θ in a neighborhood of θ_0, and $E_{\theta_0}[G_k(X)] < \infty$ each $k = 2, 3, \cdots, K+3$. Also assume that $\rho^{(K+1)}(\cdot)$ is continuous in a neighborhood of θ_0. Then there exists a constant M_1, a null set $S_K{}^c$ and for each $\mathbf{x} \in S_K$, an $N_{4x}(> N_{3x})$ such that*

(2.11) $$\int_{-\delta_2}^{\delta_2} \Big|\exp[n \textstyle\sum_{k=2}^{K+3} a_{kn}(\hat{\theta})(\phi/b)^k]\rho_K(\hat{\theta}+\phi b^{-1}) - \prod_{i=1}^{n} [f(x_i, \hat{\theta}+\phi b^{-1})/f(x_i, \hat{\theta})]\rho(\hat{\theta}+\phi b^{-1})\Big| \, d\phi \leqq M_1 n^{-\frac{1}{2}(K+2)} \quad \text{for} \quad n > N_{4x}.$$

Here a_{kn} is defined by (2.8) and ρ_K by (2.10) and δ_2 appears in the statement of Lemma 2.2.

PROOF. Together with the bound (2.5) for $b(\hat{\theta})$ on the interval $|\hat{\theta}-\theta_0| \leqq \frac{1}{2}\delta_1$, an application of Theorem 2.2 to each a_{kn}, $k = 2, \cdots, K+3$, establishes the existence of a null set $S_K{}^c$, a constant M_2 and for each $\mathbf{x} \in S_K$, an N_{4x} such that $|a_{kn}(\hat{\theta})| \leqq M_2$, $k = 2, \cdots, K+3$, for $n > N_{4x}$ each $\mathbf{x} \in S_K$.

Add and subtract $\rho_K(\hat{\theta}+\phi b^{-1}) \prod_{i=1}^{n} [f(x_i, \hat{\theta}+\phi b^{-1})/f(x_i, \hat{\theta})]$ to the integrand in (2.11). It follows from Lemma 2.2 and the inequality $|\log\{\prod_{i=1}^{n} [f(x_i, \hat{\theta}+\phi b^{-1})/f(x_i, \hat{\theta})]\} - n \sum_{k=2}^{K+3} a_{kn}(\hat{\theta})(\phi/b)^k| \leqq n|\phi|^{K+3}M_2'$ for $|\phi| \leqq \delta_2$ some M_2' that the integrand is bounded by

$$e^{-n\phi^2/6}\{|\rho^{(K+1)}(\theta_*)(\phi/b)^{K+1}/(K+1)!| + |\rho_K(\hat{\theta}+\phi b^{-1})|n|\phi|^{K+3}M_2''\}$$

where θ_* is an intermediate value and M_2'' some constant. For each $x \in S_K$, n is selected greater than N_{4x}. Then the continuity assumptions imply that the integrand is bounded by $M_1 \exp(-n\phi^2/6)\,[|\phi|^{K+1} + n|\phi|^{K+3}]$ for $|\phi| \leqq \delta_2$. The result follows upon making the change of variable $u = n^{\frac{1}{2}}\phi$.

2.3. *Proof of asymptotic nature of the expansion.* Define a function $\psi_{Kn}(z)$ for each K and n by

(2.12) $$\psi_{Kn}(z) = \sum_{k=3}^{K+3} b^{-k} a_{kn}(\hat{\theta}) z^{k-3}$$

where the dependence on $\mathbf{x}$ is suppressed. Recalling that $b^2 = -2a_{2n}(\hat{\theta})$, we write

(2.13) $$\rho_K(\hat{\theta}+\phi b^{-1}) \exp[n \textstyle\sum_{k=2}^{K+3} a_{kn}(\hat{\theta})(\phi/b)^k] = e^{-n\phi^2/2}[\rho_K(\hat{\theta}+\phi b^{-1})\, e^{n\phi^3}\psi_{Kn}{}^{(\phi)}]$$

and note that for fixed n and sequence $\mathbf{x}$, the second factor on the right-hand side is a particular evaluation of the function

$$P_K(w, z, \mathbf{x}) = \rho_K(\hat{\theta}+zb^{-1})e^{w\psi_{Kn}(z)} \qquad |w| \leqq 2, \qquad |z| \leqq 2\delta_2. \tag{2.14}$$

For each $\mathbf{x} \in S_K$ and $n > N_{4x}$, we have

$$P_K(w, z, \mathbf{x}) = \sum_{l,m} c_{lm}(\mathbf{x})w^l z^m \tag{2.15}$$

where the series converges absolutely and uniformly in the region $\{|w| \leqq 3/2, |z| \leqq 3\delta_2/2\}$ (see Markushevich 101–105). The coefficients $c_{lm}(\mathbf{x})$ are given by

$$l!m!\, c_{lm}(\mathbf{x}) = \frac{\partial^{l+m}}{\partial^l w\, \partial^m z} P_K(w, z, \mathbf{x})\bigg|_{z=0, w=0} \qquad l, m = 0, 1, \cdots \tag{2.16}$$

and the usual approximation shows that

$$|c_{lm}(\mathbf{x})| \leqq M_{lm} < \infty \qquad l, m = 0, 1, \cdots \tag{2.17}$$

where M_{lm} does not depend on n for $n > N_{4x}$. These estimates enable us to find constants A_1 and A_2 such that

$$|P_K(w, z, \mathbf{x}) - \textstyle\sum_{l+m \leqq K} c_{lm}(\mathbf{x})w^l z^m| \leqq A_1 |w|^{K+1} + A_2 |z|^{K+1} \tag{2.18}$$

for $|z| \leqq \delta_2$ and $|w| \leqq 1$ when $\mathbf{x} \in S_K$ and $n > N_{4x}$.

Denote the truncated series $\sum_{l+m \leqq K} c_{lm}(\mathbf{x})w^l z^m$ by $P_K{}^t(w, z, \mathbf{x})$. We are now ready to prove the main theorem.

PROOF OF THEOREM 2.1. Let K be an arbitrary but fixed integer, $\mathbf{x} \in S_K$ and $n > N_{4x}$ where the latter quantities appear in the statement of Lemma 2.4. A simple approximation together with the lemmas above show that

$$\begin{aligned}|\textstyle\int_{-\infty}^{\infty} \rho(\hat{\theta}+\phi b^{-1}) \prod_{i=1}^n [f(x_i, \hat{\theta}+\phi b^{-1})/f(x_i, \hat{\theta})]\, d\phi \\ -\textstyle\int_{-\infty}^{\infty} e^{-n\phi^2/2} P_K{}^t(n\phi^3, \phi, \mathbf{x})\, d\phi| \leqq B_1 n^{-\frac{1}{2}(K+2)}\end{aligned} \qquad \text{and} \tag{2.19}$$

$$\begin{aligned}|\textstyle\int_{-\infty}^{\xi n^{-1/2}} \rho(\hat{\theta}+\phi b^{-1}) \prod_{i=1}^n [f(x_i, \hat{\theta}+\phi b^{-1})/f(x_i, \hat{\theta})]\, d\phi \\ -\textstyle\int_{-\infty}^{\xi n^{-1/2}} e^{-n\phi^2/2} P_K{}^t(n\phi^3, \phi, \mathbf{x})\, d\phi| \leqq B_2 n^{-\frac{1}{2}(K+2)}\end{aligned} \tag{2.20}$$

for some B_1 and B_2. The last expression is uniform in ξ.

Integrating the approximation and collecting terms, we obtain an N_{5x} such tha the expansions (2.21) and (2.22) hold for all ξ and all $n > N_{5x}$.

$$\begin{aligned}|\textstyle\int_{-\infty}^{\infty} \rho(\hat{\theta}+\phi b^{-1}) \prod_{i=1}^n [f(x_i, \hat{\theta}+\phi b^{-1})/f(x_i, \hat{\theta})]\, d\phi \\ -\textstyle\sum_{j=0}^K \beta_j(\mathbf{x}) n^{-\frac{1}{2}(j+1)}| \leqq B_3 n^{-\frac{1}{2}(K+2)}\end{aligned} \qquad \text{and} \tag{2.21}$$

$$\begin{aligned}|\textstyle\int_{-\infty}^{\xi n^{-\frac{1}{2}}} \rho(\hat{\theta}+\phi b^{-1}) \prod_{i=1}^n [f(x_i, \hat{\theta}+\phi b^{-1})/f(x_i, \hat{\theta})]\, d\phi \\ -\textstyle\sum_{j=0}^K \alpha_j(\xi, \mathbf{x}) n^{-\frac{1}{2}(j+1)}| \leqq B_4 n^{-\frac{1}{2}(K+2)}\end{aligned} \qquad \text{all } \xi. \tag{2.22}$$

A change of variables gives

$$\alpha_j(\xi, \mathbf{x}) = \textstyle\sum_{s=0}^j c_{s,j-s} \int_{-\infty}^{\xi} y^{2s+j} e^{-y^2/2}\, dy \qquad \text{each } j = 0, 1, \cdots, K \tag{2.23}$$

and $\beta_j(\mathbf{x})$ corresponds to $\alpha_j(\infty, \mathbf{x})$ and thus is zero when j is odd. Bound (2.17) implies that each $|\alpha_j(\xi, \mathbf{x})|$ is bounded for all ξ and consequently $|\beta_j(\mathbf{x})|$ is also bounded. Since $\beta_0(\mathbf{x}) = (2\pi)^{\frac{1}{2}}\rho(\hat{\theta})$, which is bounded away from zero, the quotient series corresponding to $F_n(\xi)$ has leading term $\Phi(\xi)$. The remaining coefficients $\{\gamma_j(\xi, \mathbf{x})\}$ satisfy

$$\alpha_j(\xi,\mathbf{x}) = \beta_0(\mathbf{x})\gamma_j(\xi,\mathbf{x}) + \textstyle\sum_{s=1}^{j-1}\gamma_{j-s}(\xi,\mathbf{x})\beta_s(\mathbf{x}) + \beta_j(\mathbf{x})\Phi(\xi) \quad \text{for} \quad j = 1, 2, \cdots, K. \tag{2.24}$$

An induction argument using the lower bound on β_0 and upper bounds on the other β's and α's gives $|\gamma_j(\xi, \mathbf{x})| \leqq M_K < \infty$ all $j \leqq K$ uniformly in ξ for sufficiently large n. Consideration of the quotient series establishes the existence of a D_1 and an N_x for each $\mathbf{x}$ belonging to an almost sure set where the inequality in the theorem holds. This concludes the proof since the construction is valid for every integer K.

It should be noticed that the terms of the expansion may be obtained by formal division.

PROOF OF PROPOSITION 2.1. In (2.25), it is shown that γ_1 is of the form asserted. From (2.24), it is sufficient to show that $\alpha_j - \beta_j\Phi$ does not involve Φ. If j is odd, this result follows from (2.23). When j is even, it suffices to show

$$\textstyle\int_{-\infty}^{\xi} y^{2r+j}e^{-y^2/2}\,dy - \Phi(\xi)2^{r+(j+1)/2}\Gamma(r+j/2+\frac{1}{2})$$

does not involve Φ. Integration by parts establishes the desired result.

2.4. *Calculation of terms.* Again as in the previous papers (Johnson (1967a, 1967b)), we have the same relationships between the γ's and c_{lm}'s except it must be remembered that both are functions of the observed sequence $\mathbf{x}$. Specifically, the leading term is Φ and

$$\begin{aligned} \gamma_1(\xi, \mathbf{x}) &= -\varphi(\xi)c_{00}^{-1}[c_{10}(\xi^2+2)+c_{01}] \\ \gamma_2(\xi, \mathbf{x}) &= -\varphi(\xi)c_{00}^{-1}[c_{20}\xi^5+(5c_{20}+c_{11})\xi^3+(15c_{20}+3c_{11}+c_{02})\xi] \end{aligned} \tag{2.25}$$

where $\varphi(\xi)$ denotes the first derivative of Φ. Another term is given in the first reference above. The c_{lm} given by (2.16) may also be expressed directly in terms of ρ and the likelihood together with their derivatives. In particular

$$\begin{aligned} &c_{00} = \rho(\hat{\theta}); \quad c_{01} = b^{-1}\rho^{(1)}(\hat{\theta}); \quad c_{02} = b^{-2}\rho^{(2)}(\hat{\theta}) \\ &c_{10} = b^{-3}a_{3n}(\hat{\theta})\rho(\hat{\theta}); \quad c_{11} = b^{-4}a_{4n}(\hat{\theta})\rho(\hat{\theta}) + b^{-4}a_{3n}(\hat{\theta})\rho^{(1)}(\hat{\theta}) \\ &c_{20} = 2^{-1}b^{-6}a_{3n}^2(\hat{\theta})\rho(\hat{\theta}) \end{aligned} \tag{2.26}$$

where b is defined by (2.2) and a_{kn} by (2.9) and $\hat{\theta}$ is the maximum likelihood estimate.

The prior density enters the expansion in the term of order $n^{-\frac{1}{2}}$ as $\rho^{(1)}(\hat{\theta})/\rho(\hat{\theta})$ and the term of order n^{-1} as $\rho^{(2)}(\hat{\theta})/\rho(\hat{\theta})$ and if $c_{11} \neq 0$ as $\rho^{(1)}(\hat{\theta})/\rho(\hat{\theta})$.

3. Expansion of the moments. We also study the moments of the posterior distribution of ϕ. The moments of $n^{\frac{1}{2}}\phi$ are obtained in the obvious manner. For simplicity, only integer moments are considered. Denote the rth moment of the posterior distribution by $E_p(\phi^r)$.

THEOREM 3.1. *Under the assumptions of Theorem* 2.1, *if* $\int|\phi|^r\rho(\phi)\,d\phi < \infty$ *where* $r < K$, *there exist functions* $\{\lambda_{r,j}(\mathbf{x})\}$, *a constant* D_2 *and for all* $\mathbf{x}$ *outside a null set, an* $N_x^{\,2}$ *such that*

$$(3.1)\qquad |E_p[\phi^r]-\textstyle\sum_{j=r}^{K}\lambda_{rj}(\mathbf{x})n^{-j/2}| \leqq D_2\, n^{-\frac{1}{2}(K+1)} \qquad n > N_x^{\,2}.$$

The odd terms are zero.

PROOF. The expansion is the quotient of two expansions of the form (2.21) where the numerator has ρ replaced by $\phi^r\rho$. The denominator has a non-zero leading term and the first few terms of the numerator series are zero. Also both series have odd terms equal to zero. Finding the reciprocal series for the denominator and multiplying, we obtain the desired result.

The terms may be obtained by formal division. In particular, the first non-zero term is

$$(3.2)\qquad \lambda_{rr}(\mathbf{x}) = 2^{r/2}\Gamma(\tfrac{1}{2}(r+1))\Gamma^{-1}(\tfrac{1}{2})$$

if r is even and

$$(3.3)\quad \lambda_{r,r+1}(\mathbf{x}) = \Gamma^{-1}(\tfrac{1}{2})2^{\frac{1}{2}(r+1)}\{2(r+1)a_3(\hat{\theta})\Gamma(\tfrac{1}{2}(r+4))+\Gamma(\tfrac{1}{2}(r+2))\rho^{(1)}(\hat{\theta})/\rho(\hat{\theta})\}$$

when r is odd.

The moments of θ about a fixed point may be obtained from those of ϕ using relation (2.1). Two special cases which hold for sufficiently large n depending on $\mathbf{x}$ belonging to a set of probability one,

$$(3.4)\qquad |E_p(\theta)-\hat{\theta}-b^{-1}(6a_3(\hat{\theta})+\rho'(\hat{\theta})/\rho(\hat{\theta}))n^{-1}| \leqq D_2'n^{-2} \qquad \text{and}$$

$$(3.5)\qquad |E_p(\theta-\hat{\theta})^2-b^{-2}n^{-1}| \leqq D_2''n^{-2},$$

may be compared with Gnedenko (1962), page 413, equations (5′), who further specializes to normal populations.

4. Normalizing transformation. Let $\eta_n(\xi) = \Phi^{-1}(F_n(\xi))$ be the normalizing transformation. The notation does not reflect the dependence on $\mathbf{x}$. For fixed $\mathbf{x}$, n and ξ, η is defined by

$$(4.1)\qquad \Phi(\eta) = F_n(\xi).$$

The following theorem shows that η_n has an asymptotic expansion.

THEOREM 4.1. *Under the assumptions of Theorem* 2.1, *there exist functions* $\{\omega_j(\xi) = \omega_j(\xi, \mathbf{x})\}$, *a constant* D_3 *and for all* $\mathbf{x}$ *outside a null set, an* $N_x^{\,3}$ *such that* (4.2) *holds.*

$$(4.2)\qquad |\eta_n(\xi)-\xi-\textstyle\sum_{j=1}^{K}\omega_j(\xi)n^{-j/2}| \leqq D_3\, n^{-\frac{1}{2}(K+1)} \quad \textit{for} \quad n > N_x^{\,3}.$$

Here D_3 *is independent of* ξ *for* ξ *in a finite interval. The* $\omega_j(\xi)$ *are polynomials in* ξ *having coefficients which are bounded for sufficiently large* n.

PROOF. Let $\mathbf{x} \in S_K$ where S_K is specified in Lemma 2.4. Denote by $\Phi^{-1}(u)$ the branch of the analytic function which is the inverse of Φ and is real valued for $0 < u < 1$. For any finite ξ-interval, $F_n(\xi)$ is monotone and bounded away from 0 and 1 for all sufficiently large n. Substitution of the expansion for F_n into the series for Φ^{-1} establishes the existence of an asymptotic expansion satisfying (4.2) with D_2 independent of ξ over finite ξ-intervals. The ω_j may be obtained by formal substitution. Alternatively, they may be obtained by inserting $\eta = \sum_{j=0}^{K} \omega_j n^{-j/2}$ into

$$\Phi(\eta) \sim \Phi(\xi) + \sum_{j=1}^{K} \gamma_j(\xi, \mathbf{x}) n^{-j/2}$$

and identifying like powers of $n^{-\frac{1}{2}}$. Taking the limit as $n \to \infty$, we see that $\omega_0 = \xi$. The other coefficients satisfy

$$\Phi^{(1)}(\xi)\omega_j(\xi) = \gamma_j(\xi, \mathbf{x}) + L(\omega_0, \cdots, \omega_{j-1}, \xi), \qquad j \geqq 1 \tag{4.3}$$

where L is a polynomial in the ω's without constant terms. Each coefficient of the polynomial is a homogeneous of the first degree in the derivatives $\Phi^{(l)}$ $1 \leqq l \leqq j$. Canceling $\Phi^{(1)}$, each ω_j is seen to be polynomial. The equations (4.3) show that the c_{lm} enter only in the numerator, except for c_{00}, so that the coefficients of the polynomial ω_j are bounded for sufficiently large n. Although the argument is Wasow's, his equation (3.8) seems incorrect and the last portion of his proof should be modified as above.

The terms $\omega_j(\xi)$ of the expansion are expressible directly in terms of the $\gamma_j(\xi, \mathbf{x})$ and by (2.26), they may be written in terms of the $c_{lm}(\mathbf{x})$. For the second and third terms, we have

$$\omega_1 = \gamma_1/\varphi \quad \text{and} \quad \omega_2 = (\gamma_2/\varphi) + \xi(\gamma_1{}^2/2\varphi^2). \tag{4.4}$$

In the next section, this expansion is inverted so as to express η in terms of ξ.

5. Expansion of the percentiles. Here we consider the solution of the equation $\Phi(\eta) = F_n(\xi_n(\eta))$.

THEOREM 5.1. *Under the assumptions of Theorem 2.1, there exist functions $\{\tau_j(\eta) = \tau_j(\eta, \mathbf{x})\}$, a constant D_4 independent of η for η in a finite interval, and for each $\mathbf{x}$ outside a null set, an $N_x{}^4$ such that*

$$\left|\xi_n(\eta) - \eta - \sum_{j=1}^{K} \tau_j(\eta) n^{-j/2}\right| \leqq D_4 n^{-\frac{1}{2}(K+1)} \qquad n > N_x{}^4 \tag{5.1}$$

where the $\tau_j(\eta)$ are polynomials whose coefficients depend on the $c_{lm}(\mathbf{x})$ and are bounded for sufficiently large n.

PROOF. Let $\mathbf{x} \in S_K$ where S_K is specified in Lemma 2.4. F_n has a positive derivative over any finite ξ-interval when n is sufficiently large. Therefore $\xi_n(\eta)$ is unique and $\xi_n(\eta) \to \eta$.

Formally set,

$$\xi = \sum_{j=0}^{K} \tau_j(\eta) n^{-j/2} + \tau_{K+1}^* n^{-\frac{1}{2}(K+1)}. \tag{5.2}$$

Following Wasow, insert this expression into

$$(5.3) \qquad \eta = \sum_{j=0}^{K} \omega_j(\xi) n^{-j/2} + \omega_{K+1}^* n^{-\frac{1}{2}(K+1)}$$

and equate the coefficients of $n^{-j/2}, j = 1, \cdots, K+1$, to zero. The $\tau_j(\xi)$ turn out to be polynomials with bounded coefficients since the same is true of the first term η and the ω_j's.

The condition that the c_{lm} are bounded implies that the equation for determining τ_{K+1}^* has a solution which remains bounded as $n \to \infty$. With this value, the right-hand side of (5.2) is a solution of (5.3) and hence of (4.1). Since the right-hand side of (5.2) is real, it must equal $\xi_n(\eta)$. This proves the theorem.

The coefficients may be obtained formally. We have

$$(5.4) \qquad \tau_1 = c_{00}^{-1}(c_{10}\eta^2 + 2c_{10} + c_{01}) \quad \text{and}$$

$$(5.5) \qquad \begin{aligned} \tau_2 = {} & (5(c_{20}c_{00}^{-1} + c_{11}c_{00}^{-1} - c_{01}c_{10}c_{00}^{-2})\eta^3 \\ & + (2c_{10}^2 c_{00}^{-2} - c_{01}^2 c_{00}^{-2}/2 + 15c_{20}c_{00}^{-1} + 3c_{11}c_{00}^{-1} + c_{02}c_{00}^{-1})\eta \end{aligned}$$

where the c_{lm} depend on **x** and are given by (2.26).

As is usual for such expansions, $\eta + \sum_{j=1}^{K} \tau_j(\eta) n^{-j/2}$ is an approximate αth percentile for F_n when η is the αth percentile of Φ. This follows directly if F_n is expanded about $\xi_n(\eta)$ since $F_n^{(1)}$ is bounded in a fixed neighborhood of η for all sufficiently large n.

THEOREM 5.2. *Under the assumptions of Theorem* 2.1, *there exist a constant* D_6 *independent of* η *for* η *in a finite interval and for each* **x** *outside a null set, an* N_x^6 *such that*

$$\left|F_n\left(\eta + \sum_{j=1}^{K} \tau_j(\eta) n^{-j/2}\right) - \alpha\right| \leqq D_6 n^{-\frac{1}{2}(K+1)} \qquad n > N_x^6.$$

6. Extension of Markov processes. The results in the previous sections extend immediately to include strictly stationary Markov processes. We shall only state sufficient conditions for the existence of a $K+1$ term expansion. Kallianpur and Borwanker (1968) prove convergence in variation and our conditions essentially include theirs.

For each $\theta \in \Theta$, let $\{X_n, n \geqq 0\}$ be a strictly stationary ergodic Markov process defined on $(\mathscr{X}, \mathscr{A}) = \prod_{i=1}(R, \mathscr{B})$ into $(R, \mathscr{B})$ where $(R, \mathscr{B})$ is the Borel real line and the probability measure is P_θ. Each P_θ is induced by a stationary initial distribution $p_\theta(\cdot)$ and the stationary transition probabilities $p_\theta(\cdots)$ defined on $\mathscr{B}$ and $\mathscr{B} \times R$ respectively. Let $P_{n,\theta}$ denote the restriction of P_θ to the σ-field generated by $\{X_0, \cdots, X_n\}$. It will be assumed that the probability measures $\{P_{n,\theta}, \theta \in \Theta\}$ are mutually absolutely continuous. In order to use the transition densities as derivatives, we make the additional but unnecessary assumption that both $p_\theta(\cdot)$ and $p_\theta(y, \cdot)$ are absolutely continuous with respect to a measure μ having densities $f(z, \theta)$ and $f(y, z, \theta)$ for $\theta \in \bar{\Theta}$. The notation $E_\theta[\cdot]$ denotes expectation with respect to $P_{1,\theta}$. The original assumptions together with the additional smoothness conditions become

ASSUMPTION 1′. Θ is an open interval of the real line.

ASSUMPTION 2′. For any $\theta \in \Theta$ and $t \in \bar{\Theta}$, provided $t \neq \theta$

$$\int_{-\infty}^{\infty} |f(y, z, \theta) - f(y, z, t)| \, d\mu > 0 \qquad \text{for all} \quad y.$$

ASSUMPTION 3′. $f(y, z, \theta)$ and $f(z, \theta)$ are jointly measurable in their arguments.

ASSUMPTION 4′. For each (y, z), both $f(z, \theta)$ and $f(y, z, \theta)$ admit partial derivatives of the first and second order with respect to θ and these are continuous for all $\theta \in \bar{\Theta}$.

ASSUMPTION 5′. $\{P_{n,\theta}, \theta \in \bar{\Theta}\}$ for each $n \geqq 0$ are mutually absolutely continuous.

ASSUMPTION 6′. If $\lim_{i\to\infty} |\theta_i| = \infty$, then $\lim_{i\to\infty} f(y, z, \theta_i) = 0$ for all (y, z) except perhaps for a P_{1,θ_0}-null set which does not depend on the sequence $\{\theta_i\}$.

ASSUMPTION 7′. For all $\theta \in \bar{\Theta}$, $E_\theta |\log f(X_0, X_1, \theta)| < \infty$ and $0 < I(\theta) \equiv -E_\theta[\partial^2/\partial\theta^2 \log f(X_0, X_1, \theta)]$.

ASSUMPTION 8′. For each (y, z), there exist $K+3$ partial derivatives of $\log f(y, z, \theta)$ and $\log f(z, \theta)$, continuous for all $\theta \in \bar{\Theta}$. For each $\theta_0 \in \Theta$, there exist functions $G_j, j = 1, 2, \cdots, K+3$ satisfying $|\partial^j\theta/\partial\theta^j \log f(y, z, \theta)| \leqq G_j(y, z)$, for θ in a neighborhood of θ_0, and $E_{\theta_0}[G_j] < \infty$ for $j = 1, 2, \cdots, K+3$.

ASSUMPTION 9′. Let

$$f(y, z, \theta, \rho) = \sup_{|\theta - \theta'| \leqq \rho} f(y, z, \theta') \qquad \rho > 0, \quad \text{and}$$

$$Q(x, r) = \sup_{|\theta| > r} f(y, z, \theta) \qquad r > 0.$$

For every $\theta \in \bar{\Theta}$ and $\rho, r > 0$, $f(y, z, \theta, \rho)$ and $Q(y, z, r)$ are measurable functions of (y, z). Moreover for sufficiently small ρ and sufficiently large r,

$$E_{\theta_0}[\log f(y, z, \theta, \rho)]^+ < \infty \qquad \text{and}$$

$$E_{\theta_0}[\log Q(y, z, r)]^+ < \infty \qquad \text{for} \quad \theta_0 \in \Theta.$$

ASSUMPTION 10′. The prior density $\rho(\cdot)$ has $K+1$ continuous derivatives in a neighborhood of θ_0 and $\rho(\theta_0) > 0$. Here

$$\phi = (\theta - \hat{\theta}) b(\hat{\theta}) \qquad \text{where} \tag{6.1}$$

$$b(\hat{\theta}) = \left\{ -n^{-1} \left[\frac{\partial^2}{\partial\theta^2} \log f(x_0, \theta) + \sum_{i=1}^{n} \frac{\partial^2}{\partial\theta^2} \log f(x_i, \theta) \right]_{\theta = \hat{\theta}} \right\}^{\frac{1}{2}}. \tag{6.2}$$

The a_{kn} are now defined in terms of $n^{-1} \log [f(x_0, \theta) \prod_{i=1}^{n} f(x_{i-1}, x_i, \theta)] / [f(x_0, \hat{\theta}) \prod_{i=1}^{n} f(x_{i-1}, x_i, \hat{\theta})]$ instead of (2.9).

Let $F_n(\cdot)$ denote the posterior cdf of ϕ.

THEOREM 6.1. *The assumptions* 1′ *through* 10′ *imply the conclusion of Theorem* 2.1 *where the null set is taken* w.r.t. $f(x_0, \theta_0) \prod_1^\infty f(x_i, x_{i-1}, \theta_0)$.

Further, the results of Proposition 2.1, Theorem 3.1, Theorem 4.1, Theorem 5.1

and Theorem 5.2 hold under Assumptions 1′ through 10′. The expressions for the terms in the expansions remain the same.

The proofs of Theorem 4.1, Theorem 5.1 and Theorem 5.2 do not require any changes. In the other cases, the initial distribution is handled separately with Assumption 8′. The proof of Theorem 2.1 is modified by employing the uniform strong law for stationary ergodic processes instead of Theorem 2.2. See Kallianpur and Borwanker (1968), Lemma 2.1, for a statement of this result. However the reasoning is the same as previously and there is no need to give new expressions for the terms of any expansion.

This result may be compared with Kallianpur and Borwanker (1968) who show that, almost surely P_{θ_0}, the posterior distribution converges in variation to a normal distribution. This implies that the percentiles also converge. However they do not consider expansions of these quantities.

Another interesting comparison may be made with the work of Welch and Peers (1963). Dealing with the likelihood theory of Barnard, they consider weighted likelihoods which are mathematically identical to posterior distributions. Our results support their formal manipulations.

The author is presently investigating the multivariate extensions.

Acknowledgment. The author wishes to thank Professors J. Borwanker and G. Roussas for helpful discussion relating to Section 6.

REFERENCES

[1] BERNSTEIN, S. N. (1934). *Theory of Probability* (Russian). Moscow.

[2] GNEDENKO, B. (1962). *Theory of Probability*. Chelsea, New York.

[3] JOHNSON, R. A. (1967a). Asymptotic expansions associated with the nth power of a density. *Ann. Math. Statist.* **38** 1266–1272.

[4] JOHNSON, R. A. (1967b). An asymptotic expansion for posterior distributions. *Ann. Math. Statist.* **38** 1899–1907.

[5] KALLIANPUR, G. and BORWANKER, J. (1968). The Bernstein–von Mises Theorem for Markov processes. Technical Report No. 96 *Univ. of Minnesota*.

[6] LAPLACE, P. S. DE (1847). *Oeuvres*, **7**. Imprimerie Royale, Paris.

[7] LECAM, L. (1953). On some asymptotic properties of maximum likelihood estimates and related Bayes estimates. *Univ. California Publ. Statist.* **1** 277–330.

[8] LECAM, L. (1958). Les propriétés asymptotiques des solutions des Bayes. *Publ. Inst. Statist. Univ. Paris* **7** 17–35.

[9] MARKUSHEVICH, A. (1965). *Theory of Functions of a Complex Variable* **2**. Prentice-Hall. Englewood Cliffs.

[10] RUBIN, H. (1956). Uniform convergence of random functions with applications to statistics. *Ann. Math. Statist.* **27** 200–203.

[11] SCHWARTZ, L. (1965). On Bayes procedures. *Z. Wahrscheinlichkeits theorie und Verw. Gebiete* **4** 10–26.

[12] VON MISES, R. (1964). *Mathematical Theory of Probability and Statistics*, ed. Hilda Geiringer. Academic Press, New York.

[13] WALD, A. (1949). Note on the consistency of the maximum likelihood estimate. *Ann. Math. Statist.* **20** 595–601.

[14] WASOW, W. (1956). On the asymptotic transformation of certain distributions into the normal distribution. *Proc. Symposia Appl. Math.* **6** 251–259. McGraw-Hill, New York.

[15] WEISS, L. and WOLFOWITZ, J. (1966). Generalized maximum likelihood estimators. *Theor. Probability Appl.* **11** 58–81.

[16] WELCH, B. L. and PEERS, H. W. (1963). On formulae for confidence points based on integrals of weighted likelihoods. *J. Roy. Statist. Soc., Ser. B.* **25** 318–329.

[17] WOLFOWITZ, J. (1965). Asymptotic efficiency of the maximum likelihood estimator. *Theor. Probability Appl.* **10** 247–260.