AN ASYMPTOTIC EXPANSION FOR POSTERIOR DISTRIBUTIONS

By R. A. Johnson[1]

University of Wisconsin

0. Introduction and summary. Let ϕ be a real valued parameter for the exponential family having densities of the form

$$p_\phi(x) = C(\phi) \exp [\phi R(x)] \tag{0.1}$$

with respect to a σ-finite measure μ over a Euclidean sample space.

Now assume that the parameter ϕ has a prior density $\rho(\phi)$. The posterior density of ϕ, given $(X_1, X_2, \cdots, X_n) = (x_1, x_2, \cdots, x_n)$, is proportional to

$$[C(\phi)e^{\phi r}]^n \rho(\phi) \quad \text{where} \quad r = \textstyle\sum_{i=1}^n R(x_i)/n. \tag{0.2}$$

The expression (0.2) is proportional to a density function and hence defines a random variable ϕ whose density depends on r. But since $r = \sum_{i=1}^n R(x_i)/n$, the distribution of X when $\phi = \phi_0$ generates a sequence $(x_1, x_2, \cdots)$ and a sequence $(R(x_1), \frac{1}{2}[R(x_1) + R(x_2)], \cdots)$ and ultimately an infinite sequence of posterior densities of ϕ. It is the asymptotic form of this sequence with which we shall be concerned.

Neglecting for the moment the stochastic aspect of r, we see that if r is held fixed, perhaps at the expected value of $R(x)$, the density proportional to (0.2) is exactly of the form considered in Johnson (1966, 1967). Accordingly, after suitably centering and scaling, we obtain an asymptotic expansion having the standard normal cdf as the leading term. Closely related to this approach is the work by Bernstein and von Mises. Their results are for the Bernoulli situation, and both use the usual parameter p rather than the $\phi = \log [p/(1 - p)]$ which results if the density is cast into the form (0.2). von Mises actually holds r fixed as he passes to the limit. Their results, which give only the limiting normal term, are reproduced in Bernstein (1934), page 406, and von Mises (1964), Chapter VIII, Section C. von Mises also gives the multinomial generalization. A more recent work following the same line of attack is given in Gnedenko (1962), Section 65.

LeCam, in two basic papers (1953) and (1958), takes into account the stochastic nature of $(x_1, x_2, \cdots)$, and his Theorem 7 (1953) and Lemma 5 (1958) show that under very general conditions, the scaled posterior distribution converges to the normal distribution for almost all sequences $(x_1, x_2, \cdots)$ with respect to the infinite product measure generated by (0.1). His conditions include a more general likelihood than ours and the case where the parameter is

Received 2 March 1967; revised 18 May 1967.

[1] This work is a revision of part of the author's Ph.D. thesis which was written at the University of Minnesota and was partially supported by the National Science Foundation under Grant No. GP 3816. The Office of Naval Research under contract NONR-1202 provided partial support for the revision.

multidimensional. See LeCam (1953), page 278, for a discussion concerning the historical background on the problem of convergence of the scaled posterior distribution.

The main theorem of this paper is given in Section 1 and the following two sections give the details needed to modify the approach of Johnson (1966, 1967) so that it works in the present situation. This theorem shows that when the observations are taken from the population having density (0.1) with $\phi = \phi_0$, not only does the centered and scaled posterior distribution converge to the normal but there exists an asymptotic expansion in powers of $n^{-\frac{1}{2}}$. Section 4 gives the first two correction terms of the expansion together with examples.

Throughout this work, we will use the following notational conventions. Φ and φ are the standard normal cdf and pdf respectively. $F_n(\cdot, r)$ is the cdf of $n^{\frac{1}{2}}\theta$ where θ is defined below by Equation (1.1).

1. A limit theorem for posterior distributions. Let $X_1, X_2, \cdots$ be a sequence of independent, identically distributed random variables with density function p_ϕ given by (0.1). Lehmann (1959), Section 2.7, has shown that p_ϕ is a probability density function for all ϕ belonging to some interval I. Nature chooses a value ϕ_0 for ϕ according to law $\rho(\phi)$, and we assume ϕ_0 is interior to I. This will always be the case when I is open and in any case, will have prior probability one.

In the following, we shall find it convenient to use the parameterization

$$K(\phi) = -\log C(\phi). \tag{1.1}$$

The density (0.1) may now be written as $\exp\{\phi R(x) - K(\phi)\}$ and the ith cumulant of $R(x)$ is the ith derivative of K.

We proceed to study the asymptotic form of the posterior distribution of ϕ after standardizing ϕ by centering at the maximum likelihood estimate $\hat{\phi}(r)$ of ϕ and scaling according to a function of $\hat{\phi}$. Thus we introduce

$$\theta = [\phi - \hat{\phi}(r)]b(r) \tag{1.2}$$

where

$$b(r) = \{K''(\hat{\phi})\}^{\frac{1}{2}}, \hat{\phi} = \hat{\phi}(r). \tag{1.3}$$

Note that $b^2(r)$ may be interpreted as the maximum likelihood estimate of the variance of R or as Fisher's information evaluated at $\hat{\phi}(r)$. The choice of the centering and scaling quantities in (1.2) is discussed below in Section 3 after the proof of Theorem 1.1.

Observe that $r = \sum R(x_i)/n$ enters F_n, the posterior cdf of $n^{\frac{1}{2}}\theta$, both from the posterior distribution of ϕ and through the standardizing quantities $\hat{\phi}$ and b. To avoid a trivial case, we assume that $R(x) \neq$ constant a.s. μ.

We now state the main theorem of this paper.

THEOREM 1.1. *If $\rho(\cdot)$ is analytic in some neighborhood of ϕ_0 and $\rho(\phi_0) > 0$, then there exist functions $\{\gamma_j(\xi, r)\}_{j=1}^{\infty}$ and for each integer N, there exist constants*

A and N_x such that

$$|F_n(\xi, r) - \Phi(\xi) - \textstyle\sum_{j=1}^{N} \gamma_j(\xi, r)n^{-j/2}| \leqq An^{-(N+1)/2} \qquad \text{for all } n > N_x$$

on an almost sure set where the measure is generated by the infinite product $\Pi p_{\phi_0}(x_k)$. *Here A depends on N, and N_x depends on N and the particular sequence* $x = (x_1, x_2, \cdots)$.

2. Preliminaries. We first review some properties of the generalized density (0.1). See Kullback (1959), pages 45–47, for similar statements and further references. Define r_0 by

$$r_0 = E_{\phi_0}R. \tag{2.1}$$

LEMMA 2.1. [*Lehmann* (1959)]. *$C(\cdot)$ is analytic in a neighborhood of the complex variable $\phi + i\lambda$ whenever ϕ is interior to the natural parameter space I and the derivatives of $1/C(\phi)$ may be obtained by differentiation under the integral sign.*

LEMMA 2.2. *There exists a $d_1 > 0$ such that for fixed r with $|r - r_0| \leqq d_1$,* $\exp\{\phi r - K(\phi)\}$ *has a unique maximum at $\hat{\phi}(r)$ where $\hat{\phi}(r)$ satisfies*

$$K'(\hat{\phi}) = r$$

or

$$E_{\hat{\phi}}R = r \tag{2.2}$$

and $\exp\{\phi r - K(\phi)\}$ *strictly decreases as ϕ moves away from $\hat{\phi}$ for all ϕ in I.*

LEMMA 2.3. *$\hat{\phi}(r)$ is a continuous function of r satisfying $\hat{\phi}(r_0) = \phi_0$ and $b(r)$, defined by (1.3), is a positive continuous function of r.*

Define a function $f(\theta, r)$ by

$$\begin{aligned} f(\theta, r) &= \exp\{K(\hat{\phi}) - K(\hat{\phi} + \theta/b) + \theta r/b\}, && \text{if } \hat{\phi} + \theta/b \in I \\ &= 0, && \text{otherwise} \end{aligned} \tag{2.3}$$

where $b = b(r)$ and $\hat{\phi} = \hat{\phi}(r)$.

From the definition of θ given in (1.2), we see that the posterior density of θ, given $\sum R(x_i)/n = r$, is proportional to

$$\rho(\hat{\phi} + \theta/b)f^n(\theta, r). \tag{2.4}$$

Recall that θ is related to ϕ by (1.2).

It is easily seen that for fixed r, $f(0, r) = 1$, $f'(0, r) = 0$ and $f''(0, r) = -1$ where prime denotes differentiation with respect to θ. Since f can be extended to be analytic in θ for some neighborhood of zero, we could employ Laplace's approximation directly and obtain an asymptotic expansion of the posterior distribution of $n^{\frac{1}{2}}\theta$. See Johnson (1966, 1967) for details of the method which follows the development of de Bruijn (1961), Chapter 4. In this paper, we establish bounds similar to those of Johnson except that they hold uniformly for r in a neighborhood of r_0. The previous argument will then provide a proof of Theorem 1.1.

3. Expansion of the posterior distribution. We now establish some lemmas which ultimately lead to the proof of Theorem 1.1. In particular, the integrand (2.4) is approximated on the θ interval $(-n^{-\frac{1}{3}}, n^{-\frac{1}{3}})$ and it is shown that the remaining contribution may be neglected.

LEMMA 3.1. *Let $\rho(\cdot)$ be analytic in some neighborhood of ϕ_0 with $\rho(\phi_0) > 0$ and let $f(\theta, r)$ be defined by (2.3). Then there exist a $\delta_2 > 0$, a sequence of functions $\{c_{lm}(r)\}_{l,m=0}^{\infty}$ and for each integer N, constants A_1 and A_2 depending on N such that*

$$|\rho(\hat{\phi} + \theta/b)f^n(\theta, r) - e^{-n\theta^2/2} \textstyle\sum_{l+m\leqq N} c_{lm}(r)(\theta^3 n)^l \theta^m|$$
$$\leqq e^{-n\theta^2/2}[A_1|\theta|^{N+1} + A_2|\theta^3 n|^{N+1}] \qquad \text{all } |\theta| \leqq \delta_2 \text{ and } |n\theta^3| \leqq 1$$

for $r \varepsilon N_0$ where N_0 is defined below by (3.3). The constants A_1 and A_2 do not depend on r for $r \varepsilon N_0$.

PROOF. Consider $f(\theta, r)$ as being defined for θ complex and r real. Now $f(0, r_0) = 1$ so by the continuity of $\hat{\phi}$ and b expressed in Lemma 2.3, there exist a $\delta_0 > 0$ and $d_0 > 0$ such that the conditions $|\theta| \leqq \delta_0$ and $|r - r_0| \leqq d_0$ imply that $|f(\theta, r) - 1| \leqq \frac{1}{2}$ and that $\rho(\hat{\phi} + \theta/b)$ is different from zero and analytic in θ for fixed r. Make δ_0 and d_0 smaller if necessary so that Lemmas 2.1 and 2.2 are satisfied. Taking the principal branch for the log, define a function $h(\theta, r)$ by

$$h(\theta, r) = \log f(\theta, r). \tag{3.1}$$

It may aid the reader to note that $h(\theta, r)$ satisfies $h(\theta, r) = -I(\hat{\phi}, \hat{\phi} + \theta/b)$ where $I(\cdot, \cdot)$ is the Kullback-Liebler information for the exponential family. From the preceding discussion, there exists an M such that

$$|h(\theta, r)| \leqq M \quad \text{for} \quad |\theta| \leqq \delta_0 \quad \text{and} \quad |r - r_0| \leqq d_0. \tag{3.2}$$

Define N_0 by

$$N_0 = \{r: |r - r_0| \leqq d_0\}. \tag{3.3}$$

For fixed $r \varepsilon N_0$, $h(\theta, r)$ is analytic for $|\theta| \leqq \delta_0$ and its derivatives are given by the Cauchy formula

$$h^{(s)}(0, r)/s! = (2\pi i)^{-1} \textstyle\int_\Gamma h(t, r)/t^{s+1}\, dt \qquad \text{for } s = 1, 2, \cdots \tag{3.4}$$

where $\Gamma = \{t: |t| = 2\delta_0/3\}$. The Taylor expansion becomes

$$nh(\theta, r) = -n\theta^2/2 + (n\theta^3) \textstyle\sum_{s=3}^{\infty} a_s(r)\theta^{s-3} \tag{3.5}$$

for $|\theta| \leqq \delta_0$ and $|r - r_0| \leqq d_0$ when the values of $f(0, r)$, $f'(0, r)$ and $f''(0, r)$ are introduced. The coefficients $a_s(r)$ are given by the right hand side of (3.4).

Let $w = nx^3$ and

$$\psi(\theta, r) = \textstyle\sum_{s=3}^{\infty} a_s(r)\theta^{s-3} \tag{3.6}$$

so that

$$\rho(\hat{\phi} + \theta/b)f^n(\theta, r) = e^{-n\theta^2/2}\rho(\hat{\phi} + \theta/b)e^{w\psi(\theta, r)} \tag{3.7}$$

for $|w| \leqq 1$ and $|\theta| \leqq \delta_0$. Applying the Cauchy inequality to (3.4) with the bound

(3.2), it follows that there exists an M_1 such that $|\psi(\theta, r)| \leqq M_1$ for $|\theta| \leqq \delta_0/2$ and $r \varepsilon N_0$. Bounding ρ separately, we obtain a bound for the second factor on the right in (3.7). In particular, define a function $P(w, z, r)$ by

$$(3.8) \quad P(w, z, r) = \rho(\hat{\phi} + z/b)e^{w\psi(z,r)} \qquad \text{for } |z| \leqq \delta_0, |w| \leqq 1 \quad \text{and} \quad r \varepsilon N_0.$$

$P(w, z, r)$ is analytic for $|z| \leqq \delta_0$ and $|w| \leqq 1$ for each fixed $r \varepsilon N_0$. Also

$$(3.9) \quad |P(w, z, r)| \leqq M_2 \quad \text{for} \quad |z| \leqq \delta_0/2, \quad |w| \leqq 1, \quad \text{all} \quad r \varepsilon N_0.$$

Now for each fixed $r \varepsilon N_0$, we have the expansion

$$(3.10) \qquad P(w, z, r) = \textstyle\sum_{l,m} c_{lm}(r)w^l z^m$$

where the coefficients are given by a two variable Cauchy integral formula (see Fuks (1963), pages 39–40, or Markushevich (1965), pages 101–105). The usual estimates, using the bound (3.9), show that for every integer N, there exist constants A_1 and A_2 such that

$$|\textstyle\sum_{l+m>N} c_{lm}(r)w^l z^m| \leqq A_1|w|^{N+1} + A_2|z|^{N+1} \quad \text{for} \quad |w| \leqq 1, |z| \leqq \delta_0/4 \text{ all } r \varepsilon N_0.$$

The lemma follows with $\delta_2 = \delta_0/4$.

LEMMA 3.2. *Let $f(\theta, r)$ be defined by* (2.3). *Then there exists a $\delta_3 > 0$ depending on N_0 such that*

$$\log f(\theta, r) \leqq -\theta^2/4 \qquad \textit{for all real } \theta \textit{ with } |\theta| \leqq \delta_3 \textit{ and } r \varepsilon N_0.$$

PROOF. Apply the bound on $\psi(\theta, r)$ to the expansion (3.5).

The last two lemmas remain true if δ_3 and δ_2 are replaced by δ where

$$(3.11) \qquad \delta = \min(\delta_3, \delta_2).$$

LEMMA 3.3. *Let $f(\theta, r)$ be defined by* (2.3) *and δ by* (3.11). *There exists an ϵ $(0 < \epsilon < 1)$ such that*

$$f(\theta, r) < \epsilon \qquad \textit{for all real } \theta \textit{ with } |\theta| \geqq \delta \textit{ and } r \varepsilon N_0.$$

PROOF. For each fixed $r \varepsilon N_0$, Lemma 2.2 gives the bound $\max [f(-\delta, r), f(\delta, r)]$ and Lemma 3.2 enables us to bound this last quantity.

The first part of the proof of Theorem 1.1 parallels Johnson (1966, 1967) and we only sketch the details. Denote by $P_N(w, z, r)$ the truncated series $\sum_{l+m\leqq N} c_{lm}(r)w^l z^m$ of $P(w, z, r)$. The right hand side of (3.7) is approximated by $\exp(-n\theta^2/2)P_N(n\theta^3, \theta, r)$.

PROOF OF THEOREM 1.1. Let N be arbitrary but fixed. Let δ be given by (3.11). Bounding $b(r)$ by M and applying Lemma 3.3, we find that

$$\{\textstyle\int_{-\infty}^{-\delta} + \int_{\delta}^{\infty}\}\rho(\hat{\phi} + \theta/b)f^n(\theta, r)\, d\theta \leqq M\epsilon^n \qquad \text{all } n \quad \text{and} \quad r \varepsilon N_0.$$

From Lemma 3.2 and the fact that $\rho(\hat{\phi} + \theta/b)$ is bounded, we assert the existence of an M_1 such that

$$\{\textstyle\int_{-\delta}^{-n^{-1/3}} + \int_{n^{-1/3}}^{\delta}\}\rho(\hat{\phi} + \theta/b)f^n(\theta, r)\, d\theta \leqq M_1 \exp(-n^{\frac{1}{3}}/4) \quad (n > \delta^{-3})$$

for $r \varepsilon N_0$. Since each $c_{lm}(r)$ is bounded uniformly for $r \varepsilon N_0$,

$$\{\int_{-\infty}^{-n^{-1/3}} + \int_{n^{-1/3}}^{\infty}\}|P_N|e^{-n\theta^2/2}\,d\theta \leqq M_2 \exp(-n^{\frac{1}{3}}/4) \quad (n > 1).$$

Combining the above with the approximation in Lemma 3.1, we obtain the estimates

$$(3.12) \qquad |\int_{-\infty}^{\infty} \rho(\hat{\phi} + \theta/b)f^n(\theta, r)\,d\theta - \int_{-\infty}^{\infty} e^{-n\theta^2/2}P_N(n\theta^3, \theta, r)\,d\theta| \leqq B_1 n^{-(N+1)/2} \qquad \text{all } n > N_{B_1}$$

and

$$(3.13) \qquad \int_{-\infty}^{\xi n^{-1/2}} \rho(\hat{\phi} + \theta/b)f^n(\theta, r)\,d\theta - \int_{-\infty}^{\xi n^{-1/2}} e^{-n\theta^2/2}P_N(n\theta^3, \theta, r)\,d\theta| \leqq B_2 n^{-(N+1)/2} \qquad \text{all } n > N_{B_2}$$

some B_1, B_2, N_{B_1}, and N_{B_2} for all $r \varepsilon N_0$. This last expression is uniform in ξ. Integrating the approximation and collecting terms, we obtain the two expansions

$$\int_{-\infty}^{\infty} \rho(\hat{\phi} + \theta/b)f^n(\theta, r)\,d\theta \sim \sum \beta_j(r)n^{-(j+1)/2}$$

and

$$\int_{-\infty}^{\xi n^{-1/2}} \rho(\hat{\phi} + \theta/b)f^n(\theta, r)\,d\theta \sim \sum \alpha_j(\xi, r)n^{-(j+1)/2}$$

where

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

and β_j corresponds to $\alpha_j(\infty, r)$. Clearly, $\alpha_j(\xi, r)$ is bounded uniformly in ξ and r for $r \varepsilon N_0$ and $\beta_j(r)$ is bounded for all $r \varepsilon N_0$. Now $\beta_0(r) = (2\pi)^{\frac{1}{2}}\rho(\hat{\phi}(r))$ which is lower bounded above zero for all $r \varepsilon N_0$. The quotient series then has $\Phi(\xi)$ as a leading term and the remaining coefficients $\{\gamma_j(\xi, r)\}$ satisfy $\alpha_j(\xi, r) = \beta_0(r)\gamma_j(\xi, r) + \sum_{s=1}^{j-1} \gamma_{j-s}(\xi, r)\beta_s(r) + \beta_j(r)\Phi(\xi)$ for each $j = 1, 2, \cdots$ so that it follows by induction that each $\gamma_j(\xi, r)$ is bounded uniformly in ξ and r for $r \varepsilon N_0$. Dividing the two expansions, we conclude that there exist an A independent of $r \varepsilon N_0$ and N_A depending on A such that

$$|F_n(\xi, r) - \Phi(\xi) - \sum_{j=1}^{N} \gamma_j(\xi, r)n^{-j/2}| \leqq An^{-(N+1)/2} \qquad (n > N_A)$$

where each $\gamma_j(\xi, r)$ is bounded uniformly in ξ and r for $r \varepsilon N_0$. The $\gamma_j(\xi, r)$ are obtained by formal division.

Now consider the stochastic aspect of the problem. By the Strong Law of Large Numbers, $\bar{R} = \sum_{i=1}^{n} R(x_i)/n \to r_0$ almost surely on the product space having measure P_{ϕ_0} induced by $\Pi p_{\phi_0}(x_i)$ where p_{ϕ_0} is given by (0.1) with $\phi = \phi_0$. For every $x = (x_1, x_2, \cdots)$ belonging to an almost sure set, there exists an N_x such that $\bar{R} \varepsilon N_0$ if $n > N_x$. Repeating the above argument for each x, we complete the proof.

The conclusion of Theorem 1.1 states that almost surely, the observed sequence x provides a sequence of values of r for which the asymptotic expansion of $F_n(\xi, r)$ is valid. That is, the extra terms may be used as correction terms when n is sufficiently large.

It was decided to center at the maximum likelihood estimate $\hat{\phi}$ after encountering difficulties with the expansion of $\log f(\theta, r)$ while attempting to center at the true value ϕ_0 in which case convergence to the normal distribution may not be true. As to where to evaluate the scaling constants, note that LeCam (1953) used ϕ_0 and in LeCam (1958), $\hat{\phi}$ was used to show convergence to the normal distribution for quite general likelihoods. The result of LeCam, specialized to the density (0.1), is that the posterior distribution converges in variation almost surely. Johnson (1966) shows that the present method leads to the same conclusion for the density in (0.1).

4. Calculation of terms and examples. The results of Johnson (1966, 1967) may be used directly, giving γ_1 and γ_2 in terms of c_{lm}, although it must be remembered that the c_{lm} are functions of r. In particular,

$$\gamma_1(\xi, r) = -\varphi(\xi)c_{00}^{-1}(r)[c_{10}(r)(\xi^2 + 2) + c_{01}(r)] \tag{4.1}$$

and

$$\gamma_2(\xi, r) = -\varphi(\xi)c_{00}^{-1}(r)[c_{20}(r)\xi^5 + (5c_{20}(r) + c_{11}(r))\xi^3 + (15c_{20}(r) + 3c_{11}(r) + c_{02}(r))\xi]. \tag{4.2}$$

The following expressions for the c_{lm} enable us to express $\gamma_1(\xi, r)$ and $\gamma_2(\xi, r)$ in terms of $\rho(\cdot)$ and $C(\cdot)$ and their derivatives. Here K is related to C by (1.1).

$$\begin{aligned} c_{00}(r) &= \rho \\ c_{01}(r) &= \rho'/b(r), \qquad c_{10}(r) = -K'''(\hat{\phi})/6b^3(r) \\ c_{02}(r) &= \rho''/2b^2(r), \qquad c_{20}(r) = \rho\{K'''(\hat{\phi})/b^3(r)\}^2/72 \\ c_{11}(r) &= -1/6(\rho'/b(r))(K'''(\hat{\phi})/b^3(r)) - \rho/24(K''''(\hat{\phi})/b^4(r)) \end{aligned} \tag{4.3}$$

where $\hat{\phi}(r)$ is the solution of (2.2) and $b(r)$ is given by (1.3). Recall that the derivatives of K are related to the cumulants of $R(x)$. In fact, $K'''(\hat{\phi})/b^3(r)$ and $K''''(\hat{\phi})/b^4(r)$ are the maximum likelihood estimates of skewness and kurtosis respectively. Note also that the derivatives of ρ are accompanied by an appropriate number of $b(r)$ to give a dimensionless ratio.

The manner in which the prior density enters the asymptotic expansion of Theorem 1.1 is now apparent. In the term of order $n^{-\frac{1}{2}}$, it enters only as $\rho'(\hat{\phi})/\rho(\hat{\phi})$ and in the term of order n^{-1}, it appears as $\rho''(\hat{\phi})/\rho(\hat{\phi})$ and as $\rho'(\hat{\phi})/\rho(\hat{\phi})$ when $c_{11}(r) \neq 0$.

We now consider a few examples where $p_\phi(x)$ is given by (0.1) and the prior density $\rho(\cdot)$ satisfies the assumptions of Theorem 1.1. The first is the Bernoulli case where $\phi = \log[p/(1 - p)]$ and $K(\phi) = \log(1 + e^\phi)$. The transformed variable θ is equal to $[\phi - \log(r/(1 - r))][r(1 - r)]^{\frac{1}{2}}$.

Upon calculating the derivatives of $K(\cdot)$, we find that

$$\gamma_1(\xi, r) = -\varphi(\xi)[(2r - 1)(\xi^2 + 2)/6 + \rho'(r)/\rho(r)][r(1 - r)]^{-\frac{1}{2}}$$

and

$$\gamma_2(\xi, r) = -\varphi(\xi)[r(1 - r)]^{-1}\{(2r - 1)^2\xi^5/72 - (r^2 - r + 1)\xi^3/36 - (r^2 - r + 1)\xi/12 + (2r - 1)\rho'(r)\xi^3/6\rho(r) + [(2r - 1)\rho'(r) + \rho''(r)]/2\rho(r)\}.$$

The above result is not for the usual parameter. The fact that the posterior distribution of $n^{\frac{1}{2}}(p - r)/[r(1 - r)]^{\frac{1}{2}}$ does converge to the standard normal distribution with probability one follows easily since p is obtained as a smooth transformation of ϕ. This result may be compared with von Mises (1964), pages 345–347, and Bernstein (1934), page 406, who consider $n^{\frac{1}{2}}(p - r)/[r(1 - r)]^{\frac{1}{2}}$ and take limits while ignoring the stochastic aspect of r. Both give only the limiting normal distribution.

As a second example, we consider the normal distribution with known variance σ_0^2 and mean m. Here $p_\phi(x)$ has $\phi = m/\sigma_0^2$, $K(\phi) = \sigma_0^2\phi^2/2$ and $R(x) = x$. The transformed variable θ equals $(\phi - \bar{x}/\sigma_0^2)\sigma_0$ which can be written as $(m - \bar{x})/\sigma_0$. In this case, $\gamma_1 = -\varphi(\xi)\rho'(\bar{x})/\sigma_0\rho(\bar{x})$ and $\gamma_2 = -\varphi(\xi)\rho''(\bar{x})/\sigma_0^2\rho(\bar{x})$. Theorem 1.1 gives an expansion which may be compared with Gnedenko (1962), page 414, second equation. Gnedenko shows that the limiting distribution is normal.

5. Acknowledgment. I wish to express my gratitude to Professor R. J. Buehler for his guidance and advice throughout the preparation of this paper.

REFERENCES

BERNSTEIN, S. (1934). *Theory of Probability* (Russian). Moscow.

BUEHLER, R. J. (1965). The limit of the n-th power of a density. *Ann. Math. Statist.* **36** 1878–1882.

FUKS, B. (1963). *Analytic Functions of Several Complex Variables.* Translations of Mathematical Monographs 8. Amer. Math. Soc., Providence.

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

JOHNSON, R. A. (1966). Asymptotic expansions associated with the nth power of a density. Technical Report No. 80 (Ph.D. Thesis). Univ. of Minnesota.

JOHNSON, R. A. (1967). Asymptotic expansions associated with the nth power of a density. *Ann. Math. Statist.* GTTI, England.

KULLBACK, S. (1959). *Information Theory and Statistics.* Wiley, New York.

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

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

LEHMANN, E. L. (1959). *Testing Statistical Hypotheses.* Wiley, New York.

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

VON MISES, R. (1964). *Mathematical Theory of Probability and Statistics.* (Edited by Hilda Geiringer.) Academic Press, New York.