CONTRIBUTIONS TO THE STUDY OF BAYES ESTIMATES, THE MAXIMUM LIKELIHOOD ESTIMATE AND RAO'S TEST

S.N. JOSHI

RESTRICTED COLLECTION

Thesis submitted to the Indian Statistical Institute
in partial fulfilment of the requirements
for the award of the degree of
Doctor of Philosophy
CALCUTTA
1982

ACKNOWLEDGEMENTS

I thank the Indian Statistical Institute for providing me with all facilities to carry out my research work.

I am extremely grateful to my supervisor Professor J.K. Ghosh for inspiring me to work in asymptotic theory and helping me along at every stage of my research. I also thank him for permitting me to include our joint work in this thesis.

It is a pleasure to record my indebtedness to Drs. Tapas Chandra and Bimal K. Sinha for many helpful discussions and for allowing me to include our joint work here.

I thank every member of the Division of Theoretical Statistics and Mathematics for providing the excellent academic atmosphere.

Finally I thank Mr. Dilip Kumar Bardhan for excellent typing and Mr. Muktalal Bag for cyclostyling of this thesis.

Calcutta,
June 1982

S.N. Joshi

C O N T E N T

*****

INTRODUCTORY CHAPTER

This thesis consists of two parts. In part I we have investigated problems concerning Bayes estimates, especially expansion of the integrated risk of the Bayes estimate (also referred to as the Bayes risk or the integrated Bayes risk), approximation of the Bayes estimate and expansion of the posterior distribution. In part II we have introduced a new optimum property for estimates and have concluded that the maximum likelihood estimate (m.l.e.) enjoys this property; in this part we have also investigated what is known as Rao's conjecture which says that the test based on the score function is " locally " more powerful than the likelihood ratio test and the Wald's test. Our conclusion is that Rao's conjecture is true when the sizes of the above tests are small.

Consider a sequence $X_1, X_2, \ldots$ of independent and identically distributed (i.i.d.) random variables (r.v.'s). X_1 having distribution function (d.f.) $F(x, \theta)$, parametrized by $\theta \in \Theta$, Θ an open subset of $\mathbb{R}$. Let $f(x, \theta)$ be the density of $F(x, \theta)$ w.r.t. some sigma finite measure. The problem of expansion of the posterior distribution function <u>for a fixed value θ_0 of the parameter</u> was considered by Johnson ([1967], [1970]). He proved under certain regularity conditions that with probability one under θ_0 the suitably centered and scaled posterior distribution possesses an asymptotic expansion in powers of $n^{-1/2}$ (n being the sample size) with standard normal as the leading term. The remainder after $(K+1)$ terms is $n^{-(K+1)/2} R_{nK}$ and R_{nK} is bounded by some constant M, $0 < M < \infty$.

Let $(\underline{H})$ be an interval and $(\underline{H})_1$ be a bounded open subinterval whose closure $(\overline{\underline{H}})_1$ is also contained in $(\underline{H})$. Let ρ, the prior density of the parameter θ w.r.t. the Lebesgue measure, be such that $\rho(\theta) > 0$ on $(\overline{\underline{H}})_1$. Choose a constant $r > 0$. Then proceeding as in Johnson [1970], under certain regularity conditions which are stronger than Johnson's and depend on r, one can get the following uniform version of the above mentioned result of Johnson,

$$P_{\theta_o}(|R_{nK}| < M) = 1 - o(n^{-r}) \quad \text{uniformly in } \theta_o \in (\underline{H})_1 \text{ for}$$

some $0 < M < \infty$.

One can also obtain an expansion for the posterior risk under square error loss function using Johnson's other results.

Similar problems were also considered by Gusev ([1975], [1976]). Basically he was interested in getting asymptotic expansions for the Bayes and certain other estimates, which hold good with large P_{θ_o} probability (vide Gusev [1975]); and then getting asymptotic expansions for quantities like $E_{\theta_o}(t_n - \theta_o)^m$ where t_n is, for example, the Bayes estimate (Gusev [1976]). For example one of his results gives us

$$\text{(I.1)} \qquad E_{\theta_o}[E(\theta|\underline{x}_n) - \theta_o]^2 = \sum_{i=2}^{k} h_i \, n^{-i/2} + o(n^{-k/2}).$$

These studies under fixed θ_o are important as they provide the tools for judging the performance of the Bayes estimates as regards consistency, asymptotic normality, risk (under P_{θ_0}) w.r.t. various types of loss functions and robustness.

However, from a Bayesian point of view it is more natural as well as more useful to have versions of abovementioned results of Johnson and Gusev which hold with large P_ρ probability (here P_ρ denotes the marginal probability for X's i.e. for a set A in the sample space of X's $P_\rho(A) = \int P_\theta(A)\ \rho(\theta)d\theta$, ρ being the prior density of θ). For example, the version

(I.2) $$P_\rho(|R_{nK}|n^{-(K+1)/2} < n^{-\lambda}) = 1 - o(n^{-r})$$

of Johnson's result on approximation of the posterior distribution function is useful in deriving posterior confidence sets for the parameter θ; moreover, expansions for the posterior risk and Bayes estimates, which hold good on a large P_ρ probability sets, are useful in getting asymptotically optimal stopping times in the problem of Bayesian sequential estimation.

An expansion for the integrated Bayes risk, namely, $\int E_{\theta_0}[(E(\theta|\underline{x}_n) - \theta_0)^2]\ \rho(\theta_0)d\theta_0 \equiv R_\rho$ of the form

(I.3) $$R_\rho = a_1 n^{-1} + a_2 n^{-2} + o(n^{-2})$$

where a_1 and a_2 are constants, can be used to study robustness of Bayes solutions as well as various other purposes at the planning stage of an experiment. For example it can be used to find the optimum (fixed) sample size, n, if in addition to cost of wrong decision there is a cost c per unit of observation; one has simply to minimize (nc + the R.H.S. of (I.3)).

Apart from these Bayesian considerations, results of the above type can be used as tools in the Neyman Pearsonian framework as well; for example, an approximate Bayes estimate (w.r.t. a sequence of priors) can be used to investigate second order properties (at a fixed θ_0) like second order efficiency of estimates (vide Section 2 of Chapter three).

We prove results of the type (I.2) in Chapter one and (I.3) in Chapter two. One can not get results of these type as easy consequences of results of Johnson or Gusev. The main source of difficulty is the remainder term which depends on the fixed parameter point θ_0, for example the $o(n^{-k/2})$ term of (I.1); Gusev was unable to trace the dependence on θ_0 of this term. These remainder terms blow up as θ_0 approaches the end points a_0 and b_0 (here $[a_0, b_0] \subseteq \bar{H}$ is such that $\rho(\theta) > 0$ on (a_0, b_0) and $\rho(\theta) = 0$ on $(a_0, b_0)^c$; throughout this thesis we will be considering priors of this type and a_0 and b_0 will have this meaning) and hence their treatment becomes very delicate. That neglect of this fact can lead to a wrong result, can be seen from our counterexample (see Section 4 of Chapter two) to Corollary 2 of Alvo [1977]; in this connection see also the consequent changes needed in the statement and proof of his Theorem 4 (vide Chapter three Section 3). Our technique of handling remainder terms of the form $n^{-(K+1)/2} R_{nK}$ involves, mainly, bounding the terms R_{nK} by a suitable positive power of n on a large P_ρ probability set. To do this we have to assume that ρ has smooth contact near the end points a_o and b_o but this alone does not suffice.

To get a flavour of our assumptions and results, consider the following immediate consequence of our Theorem 1.3. Let the common density $f(x, \theta)$ be $c(\theta) \exp[\theta T(x) + \psi(x)]$ and let $[a_o, b_o]$ be a subset of the interior of the natural parameter space. Let ρ be inifinitely differentiable on $[a_o, b_o]$ and let $\rho^{(i)}(\theta)\rho^{-1+v}(\theta) \longrightarrow 0$ as $\theta \longrightarrow a_o$ or b_o for all $v > 0$ and $i = 1,2,\ldots$ where $\rho^{(i)}$ is the ith derivative of ρ. Further assume that ρ is monotone near a_o and b_o then for any $K \geq 0$, $\varepsilon > 0$ and $r > 0$

$$P_\rho(n^{-(K+1)/2}|R_{nK}| < n^{-(K+1-\varepsilon)/2}) = 1 - o(n^{-r}).$$

Results of this type are proved in Chapter one (see Theorems 1.2 and 1.3)

In Chapter two, results regarding approximate Bayes estimates, expansion of the integrated Bayes risk (vide Theorem 2.1) and approximation of the posterior risk (vide Theorem 2.2) are given. For example Theorem 2.1 says that the integrated Bayes risk R_ρ (w.r.t. square error loss function) has an expansion of the form (I.3). This theorem also gives us a simple estimate whose risk differs from R_ρ by $o(n^{-2})$. In this connection it is interesting to note that (vide Example 2 of Chapter two) even in the case when we have i.i.d.

observations from normal distribution with mean θ and variance one, the integrated Bayes risk w.r.t. the square error loss function does not have expansion of the form (I.3), if the prior is uniform on [0,1].

In Chapter three Section 2 we construct a sequence of priors which converges to a degenerate distribution at θ_0 and use the above mentioned simple approximate Bayes estimate to prove the second order efficiency of the m.l.e. (vide Theorem 3.1).

In his Theorem 4 Alvo [1977] obtained a asymptotically optimal stopping rule for the problem of Bayesian sequential estimation but as be used his Corollary 2 to prove this theorem the proof is not valid. In Section 3 of Chapter three we have given a proof of this theorem. Our proof is based on the proof given by Alvo but it uses our results of Chapter two regarding approximations to the posterior risk and Bayes estimates.

Here it may be noted that after most of the above mentioned work was done we were informed by Professor A. Novikov about an annoucement (without proofs) of similar results on expansion of the integrated Bayes risk in case of location parameter families by Burnasev (see Burnasev [1979]). Burnasev's conditions on prior are more general. In Section 6 of Chapter one and Section 5 of Chapter two we have assumed conditions on prior which are similar to those of Burnasev and shown how we can get our results with slight modifications of our earlier arguments. Results of the first three Chapter are based on Ghosh, Sinha and Joshi [1981].

In Chapter two expansion for integrated Bayes risk is obtained under regularity conditions which include continuous defferentiability of $f(x,\theta)$ (and some of it's derivatives) but it is well known (see e.g. Strasser [1978]) that under very general conditions (such as LAN), limit of the integrated Bayes risk for bounded loss functions can be obtained and so the natural question to ask is whether under such weaker conditions is it possible to get an expansion for the integrated Bayes risk. As a first attempt in this direction, in Chapter four, we have investigated the double exponential family of distributions which satisfy LAN condition but for which our conditions of Chapter two do not hold. Our conclusion (see Theorem 4.1) is that in this case Bayes risk w.r.t. square error loss has an expansion but now the term after the n^{-1} term is not of order n^{-2} (as was the case earlier) but is of order $n^{-3/2}$. The method of proof of Theorem 4.1 is different from that of Theorem 2.1. Here we expand the log likelihood ratio around θ_o instead of around the m.l.e., as was the case earlier. Results of Chapter four are based on Joshi [1982].

In the case of i.i.d. set up under certain regularity conditions, the test based on the score function (as proposed by Rao), the likelihood ratio test (as proposed by Neyman and Pearson) and the Wald's test have the same Pitman efficiencies for testing a simple hypothesis $H_o : \theta = \theta_o$ vs $H_1 : \theta \neq \theta_o$ and so Rao [1965] raised the question (see the last paragraph of his Section 6c.2) of higher order discrimination and conjectured that the test proposed by him is likely to be " locally " more powerful than the remaining two (in the second edition, this conjecture was omitted).

In Chapter five of part two of this thesis we have investigated this conjecture ; our conclusion is that Rao's conjecture is true when sizes of these three tests are small (vide Theorem 5.1). Here it may be noted that Peers [1971] attempted to settle this conjecture. As pointed out by Chandra in his thesis (see Section 5 of Chapter three of Chandra [1980]) Peer's approach was not proper ; there Chandra also gave arguments showing that these three test have finite mutual deficiencies. For the sake of completeness, in the introduction of Chapter five we give details of his investigations. Results of this chapter are based on Chandra and Joshi [1982].

In the final chapter we introduce a new optimum property for estimates and show that the m.l.e. possesses it. Under certain regularity conditions on $f(x,\theta)$ we prove that (vide Theorem 6.1) with probability tending to one under a fixed θ_o, the m.l.e. $\hat{\theta}_n$ lies in the $100(1-\alpha)\%$ $(0<\alpha<1)$ confidence set V_n determined by the locally most powerful unbaised tests (LMPU tests) of hypothesis $H_o : \theta = \theta_o$ vs $H_1 : \theta \neq \theta_o$. A sort of converse is also proved, that is, if a consistent estimate T_n is such that $P_{\theta_o}(T_n \varepsilon V_n) \longrightarrow 1$ then $n^{1/2}(T_n - \hat{\theta}_n) \longrightarrow 0$ in P_{θ_o} probability; also if T_n is such that $n^{1/2}(T_n - \hat{\theta}_n) \longrightarrow 0$ in P_{θ_o} probability then $P_{\theta_o}(T_n \varepsilon V_n) \longrightarrow 1$. This chapter is based on Ghosh, Sinha and Joshi [1980].

PART I

CHAPTER ONE

EXPANSION FOR POSTERIOR PROBABILITY

1.1 INTRODUCTION : Asymptotic expansion of the posterior distribution has been investigated by Johnson ([1970], [1967]). Considering a one parameter family of distribution in the independent and identically distributed (i.i.d.) set up, he proves under certain regularity conditions, that with probability one under a fixed value of the parameter, the suitably centered and scaled posterior distribution possesses an asymptotic expansion in powers of $n^{-1/2}$, with the standard normal as the leading term. (Johnson expands only the posterior distribution function but his expansion is valid for the posterior probability of any Borel set.) The proof involves, among other things, repeated use of a version of the uniform strong law. Employing yet another version of the uniform strong law (for a precise statement see Lemma 1.1) and proceeding analogously as in Johnson [1970], the following uniform variant of the main result on asymptotic expansion of the posterior distribution can be obtained.

Let the parameter space $(\underline{H})$ be a possibly unbounded interval. Consider a bounded open subinterval $(\underline{H})_1$ whose closure $\overline{(\underline{H})}_1$ is properly contained in $(\underline{H})$. Fix a prior ρ such that $\rho(\theta) > 0$ on $\overline{(\underline{H})}_1$. Choose a constant $r > 0$. Then under certain regularity conditions which are stronger than Johnson's and depend on r, $\exists\ n_0$ and $M (0 < M < \infty)$ such that $\forall\ n > n_0$

(1.1) $$P_{\theta_0}\Big[|F_n(\xi) - \Phi(\xi) - \sum_{j=1}^{K} \gamma_j(\xi, \underline{x}_n) n^{-j/2}| \le M.n^{-\frac{(K+1)}{2}} \text{ uniformly in } \xi\Big]$$

$$= 1 - O(n^{-r}), \text{ uniformly in } \theta_0 \in (\overline{\underline{H}})_1$$

where $\underline{x}_n = (x_1, \ldots, x_n)$, $F_n(\xi) = P[\sqrt{n}(\theta - \hat{\theta}_n)b \le \xi | \underline{x}_n]$,

$\Phi(\xi) = \frac{1}{\sqrt{2}} \int_{-\infty}^{\xi} \exp(-x^2/2)dx$, $\hat{\theta}_n$ is the maximum likelihood estimate (m.l.e), $(-b^2)$ is the second derivative of the loglikehood at $\hat{\theta}_n$ and quantities γ_j's are as defined in Theorem 1.1. (This is essentially our Corollary 1.1).Here K is such that $\rho(\theta)$ is $(K+1)$ times continuously differentiable. It may be noted that if $r > 1$, by the Borel Cantelli Lemma, (1.1) implies Johnson's result. Conversely, Johnson's result implies (1.1) if $N_{\underline{x}}$ (as defined in Johnson [1970]) has finite r^{th} order moment. In the special case of exponential families (1.1) holds for all $r > 0$ provided $\overline{(\underline{H})}_1$ is contained in the interior of the natural parameter space. This result or rather a slight variant of it has been observed by Alvo [1977]; see in this connection his Corollaries 1 and 2.

While results of Johnson or the one given above in (1.1) are important, a Bayesian would be more interested in versions which hold with a large probability under P_ρ, rather than merely P_{θ_0} (where P_ρ denotes the marginal probability measure of X under the prior ρ). Arguments needed for (1.1) also lead to

$$(1.2) \quad P_\rho \left[|F_n(\xi) - \Phi(\xi) - \sum_{j=1}^{K} \gamma_j(\xi, \underline{x}_n) . n^{-j/2}| \leq M . n^{-\frac{(K+1)}{2}} . |C_n|^{-1} \lambda_n \text{ uniformly in } \xi \right]$$

$$= 1 - O(n^{-r})$$

provided certain regularity conditions hold and ρ is positive on a bounded interval (a_o, b_o), $\rho = 0$ on $(a_o, b_o)^c$ and ρ is $(K+1)$ times continuously differentiable on $[a_o, b_o]$. (All priors in this thesis will satisfy such conditions unless stated otherwise and (a_o, b_o) will have the meaning assigned here) (vide Theorem 1.1). Since $|C_n|^{-1} \lambda_n$ is an unbounded random quantity, (1.1) ceases to be true if the error term in the expansion is replaced by constant $n^{-\frac{(K+1)}{2}}$. (See in this connection Example 1.1 of Section 5 of this chapter).

If ρ behaves like a polynomial of degree K at both end points of the support we shall say ρ is of type D_K ; in this case we can estimate (vide Proposition 1.1) the P_ρ-probability that $\hat{\theta}_n$ lies in neighbourhoods of the two end points whose lengths tend to zero at a certain rate. This enables us to prove (vide Theorem 1.2) that with a large P_ρ-probability the error $M . n^{-\frac{(K+1)}{2}} . |C_n|^{-1} \lambda_n$ as well as the terms $\gamma_j \, n^{-j/2}$ tend to zero as $n \longrightarrow \infty$. For a sub-class D_∞ of infinitely differentiable priors the same end is achieved in Theorem 1.3 via Proposition 1.2. To get a flavour of these results and see the significant role played by the rate of decay near the end-point of the support, consider the following immediate consequence of Theorems 1.1 and 1.3. Let the common density $f(x,\theta)$ be $C(\theta)\exp(\theta.T(x)+\psi(x))$ and let $[a_o, b_o]$ be a subset of the interior of the natural parameter space. Then for any $K \geq 0$, $\varepsilon > 0$ and $r > 0$

$$P_\rho \left[|F_n(\xi) - \Phi(\xi) - \sum_{j=1}^{K} \gamma_j(\xi\ \underline{x}_n).n^{-j/2}| \leq M.n^{-(\frac{K+1}{2}-\varepsilon)} \text{ uniformly in } \xi \right]$$

$$= 1 - o(n^{-r}) \text{ if } \rho \text{ is of type } D_\infty$$

$$= 1 - O(n^{-\frac{(S+1)}{S}(\frac{S-K-2}{2}+\varepsilon)}) - O(n^{-\frac{(S+1)}{K+1}.\varepsilon}) \text{ if } \rho \text{ is of type}$$

D_S, $S > K+2$.

Our regularity assumptions on the density are collected in Section 3. (Note that assumption AV is taken from Pfanzagl [1973a].) The preliminary lemmas needed for Section 2 are presented in Section 4. In Section 5 we give an example which shows that the unboundedness of $|C_n|^{-1}\lambda_n$ can not be ignored with impunity. In Section 6 we prove a version of Theorem 1.2 under Burnasev type of conditions. Results of this chapter are based on the Sections 2 and 4 of Ghosh, Sinha and Joshi [1981].

1.2 MAIN RESULTS.

THEOREM 1.1 : Let the assumptions AI to AV hold for some $r_1 \geq 3$ and $r_2 \geq 2$. Let ρ be K+1 times continously differentiable on $[a_o, b_o]$. Then $\exists$ $M (0 < M < \infty)$ such that

$$(1.2a)\quad P_\rho\Big[|F_n(\xi)-\Phi(\xi)-\sum_{j=1}^{K}\gamma_j(\xi,\underline{x}_n).n^{-j/2}|\le M.n^{-\frac{(K+1)}{2}}|C_n|^{-1}\lambda_n \text{ uniformly in } \xi\Big]$$

$$= 1-O(n^{-r_1/2})-o(n^{-r_2/2}) \quad \text{where}$$

$C_n=\sum_{j=0}^{K}\beta_j(\underline{x}_n)\,n^{-j/2}$, $\lambda_n=1+d_n$, $d_n=\sum_{j+j'\ge K+1} n^{-(j+j'-K-1)/2}|\beta_j\,\gamma_{j'}|$

and β_j's and γ_j's are as in Remark 1.4. (See Remark 1.1 for comments on β_j's and γ_j's; also note that (1.4) and (1.7a) imply that with P_ρ-probability tending to one $|C_n|^{-1}\lambda_n$ is bounded).

Hence forward M will be used as a generic constant independent of ξ.

<u>PROOF</u> : Note that if $|A-a|\le R_1$, $|B-b|\le R_2$ and $|\frac{A}{B}|\le C$ then

$$(1.2b)\qquad |\frac{A}{B}-\frac{a}{b}|\le |b|^{-1}(R_1+C.R_2)$$

The posterior is a ratio whose numerater, say A, and denominator, say B, are approximated in (1.17) and (1.18); also R_n occuring in (1.17) and (1.18) is bounded by $M.n^{-(K+1)/2}$, vide Remark 1.3. Hence using (1.2b) with $C=1$ and $b=C_n$, (1.17), (1.18) and the fact

$$(\sum_{j=0}^{K}\alpha_j\,n^{-(j+1)/2})(\sum_{j=0}^{K}\alpha_j\,n^{-(j+1)/2})^{-1}=\sum_{j=0}^{K}\gamma_j\,n^{-j/2}$$

$+\,(\sum_{j+j'\ge K+1}\beta_j\gamma_{j'}\,n^{-(j+j')/2})\,C_n^{-1}$ we have uniformly in $\theta\in[a_o,b_o]$

$$(1.3)\quad P_\theta\Big[|F_n(\xi)-\Phi(\xi)-\sum_{j=1}^{K}\gamma_j(\xi,\underline{x}_n).n^{-j/2}|\le M.n^{-\frac{(K+1)}{2}}|C_n|^{-1}\lambda_n \text{ uniformly in } \xi\Big]$$

$$= 1-O(n^{-r_1/2})-o(n^{-r_2/2}) \quad \text{which gives us (1.2a)} \;\Box$$

In view of Lemmas 1.2, 1.5 and 1.6, under appropriate assumptions we have uniformly in $\theta\in[a_o,b_o]$

(1.4) $P_\theta(A_n) = 1 - O(n^{-r_1/2}) - o(n^{-r_2/2})$ where

$A_n = [-M < a_{k,n}(\hat{\theta}_n) < M,\ k=1,\ldots,K+2,\ 0 < \delta_1 < b^{-1} < M]$, $a_{k,n}$'s are as defined in Lemma 1.5.

REMARK 1.1 : Note that $\beta_0(\underline{x}_n) = \sqrt{2\pi}\,\rho(\hat{\theta}_n)$,

(1.5) $$\beta_j(\underline{x}_n) = \sum_{i=1}^{j} a_{n,j,i}\ \rho^{(i)}(\hat{\theta}_n) \qquad j \geq 1 \quad \text{and}$$

(1.6) $$\gamma_j(\xi,\underline{x}_n) = \sum_{\underline{i}\,\varepsilon\, I_j} k_{n,j,\underline{i}} \prod_{k\,\varepsilon\,\underline{i}} \rho^{(k)}(\hat{\theta}_n)\,\rho^{-1}(\hat{\theta}_n)$$

$$I_j = (\underline{i} = \{i_1,\ldots,i_\alpha\} : \sum_{r=1}^{\alpha} i_r \leq j, \quad \alpha = 1,\ldots,j)$$

where $a_{n,j,i}$'s $(1 \leq i \leq j,\ 1 \leq j \leq K)$ and $k_{n,j,\underline{i}}$'s $(\underline{i}\ \varepsilon\ I_j,\ 1 \leq j \leq K)$ are bounded (uniformly in ξ, in the later case) on A_n. By (1.5) we get

(1.7) $$C_n = \rho(\hat{\theta}_n)(\sqrt{2\pi} + \sum_{j=1}^{K} C_{n,j}\ \rho^{(j)}(\hat{\theta}_n)\rho^{-1}(\hat{\theta}_n).n^{-j/2}) \quad \text{and}$$

$$|C_n|^{-1}d_n = |\sqrt{2\pi} + \Sigma\, C_{nj}\ \rho^{(j)}(\hat{\theta}_n)\ \rho^{-1}(\hat{\theta}_n)\ n^{-j/2}|^{-1}$$

$$\cdot\Big|\sum_{j+j'\geq K+1} n^{-(j+j'-K-1)/2}\ \rho^{-1}(\hat{\theta}_n)\ \beta_j\ \gamma_{j'}\Big|$$

where $C_{n,j}$ $(0 \leq j \leq K)$ are bounded on A_n.

The quantities $|C_n|^{-1}\ell_n$ and γ_j's $(1 \leq j \leq K)$ are unbounded ; but it is clear that they are bounded on $A_n \cap (\rho(\hat{\theta}_n) > \varepsilon)$ for any $\varepsilon > 0$. Now if $\theta\ \varepsilon\ (\underline{H})_1$, $\overline{(\underline{H})}_1 \subset (a_0, b_0)$, then by Lemma 1.2, uniformly in $\theta\ \varepsilon\ (\underline{H})_1$

(1.7a) $P_\theta(\rho(\hat{\theta}_n) > \varepsilon) = 1 - o(n^{-r_2/2})$ for some $\varepsilon > 0$.

This along with (1.4) and (1.3) gives the following result.

COROLLARY 1.1 : Let the assumptions AI to AV hold for some $r_1 \geq 3$ and $r_2 \geq 2$. Let ρ be K+1 times continuously differentiable on $[a_o, b_o]$. Let $(\underline{H})_1$ be such that $\overline{(\underline{H})_1} \subset (a_o, b_o)$ then uniformly in $\theta \in (\underline{H})_1$

$$P_\theta \left[|F_n(\xi) - \Phi(\xi) - \sum_{j=1}^{K} \gamma_j(\xi, \underline{x}_n)\, n^{-j/2}| < M.n^{-\frac{(K+1)}{2}} \quad \text{uniformly in } \xi \right]$$

$$= 1 - O(n^{-r_1/2}) - o(n^{-r_2/2}).$$

As explained in the introduction we now study the behaviour of the error $n^{-(K+1)/2}.|C_n|^{-1}\lambda_n$ and terms in the expansion under P_ρ. We begin with a few definitions.

DEFINITION 1.1 : Prior ρ is said to be of type D_K, $2 \leq K < \infty$, if

(i) $\rho(\theta) > 0$ on a bounded interval (a_o, b_o) and
$\rho(\theta) = 0$ on $(a_o, b_o)^c$,

(ii) it has (K-1) continuous derivatives on $[a_o, b_o]$,

(iii) $\exists$ $c_i > 0$ and $c_i' > 0 (0 \leq i \leq K-2)$ such that for $0 \leq i \leq K-2$,

$$\rho^{(i)}(\theta) = (\theta - a_o)^{K-i}(c_i + o(1)) \quad \text{for } \theta \text{ near } a_o$$

$$= (b_o - \theta)^{K-i}(c_i' + o(1)) \quad \text{for } \theta \text{ near } b_o$$

where $\rho^{(i)}(\theta) = \frac{d^i}{d\theta^i}\rho(\theta)$ and $\rho^{(0)}(\theta) = \rho(\theta)$.

If the degrees of the polynomials at the two end points do not coincide, only minor modifications are needed in Proposition 1.1 and hence in the Theorem 1.1.

DEFINITION 1.2 : Prior ρ is said to be of type D_∞ if

(i) same as (i) of Definition 1.1,

(ii) it is infinitely differentiable on $[a_o, b_o]$,

(iii) $\rho^{v-1}(\theta).\rho^{(i)}(\theta) \longrightarrow 0$ as $\theta \longrightarrow a_o$ or b_o for all $v > 0$ and $i = 1,2,\ldots.$

(iv) ρ is monotone near a_o and b_o.

An example of this kind of prior is

$$\rho(\theta) = \text{constant}.\exp\left[\frac{-1}{\theta(1-\theta)}\right] \quad 0<\theta<1$$
$$= 0 \qquad \text{elsewhere.}$$

PROPOSITION 1.1 : Let the assumptions AI to AIII and AV hold for some $r_2 \geq 2$. Let ρ be of type D_S. Then for any C, C' $(0 < C, C' < \infty)$ and $0 < 2m \leq S$

$$(1.8) \qquad P_\rho\left[\hat{\theta}_n - a_o < C.n^{-\frac{m}{S}}\right] = O(n^{-\frac{(S+1)}{S}.m}) + o(n^{-r_2/2})$$
$$= P_\rho\left[b_o - \hat{\theta}_n > C.n^{-\frac{m}{S}}\right]$$

and

$$(1.9) \qquad P_\rho\left[\rho(\hat{\theta}_n) > C'.n^{-m}\right] = 1 - O(n^{-\frac{(S+1)}{S}.m}) - o(n^{-r_2/2})$$

PROOF : For simplicity let $a_o = 0$ and $b_o = 1$.

Note that for any $C' \exists\, C\ (0 < C', C < \infty)$ such that

$$(1.10) \qquad (\hat{\theta}_n > C.n^{-m/S}) \cap (\hat{\theta}_n < 1 - C.n^{-m/S}) \implies (\rho(\hat{\theta}_n) > C'.n^{-m})$$

Hence it suffices to prove (1.8).

$$P_\rho[\hat{\theta}_n < C.n^{-m/S}] = E_\rho P_\theta[\sqrt{n}(\hat{\theta}_n - \theta)\beta(\theta) < \sqrt{n}(C.n^{-m/S} - \theta)\beta(\theta)]$$

(see Lemma 1.3 for definition of $\beta(\theta)$)

$$= E_\rho P_\theta[\sqrt{n}(\hat{\theta}_n - \theta)\beta(\theta) < -u.\beta(\theta)],\ u = \sqrt{n}(\theta - C.n^{-m/S})$$

$$\leq \int_{u<0} \rho(\theta)d\theta + \int_{0<u<\alpha\sqrt{\log n}} P_\theta[\sqrt{n}(\hat{\theta}_n - \theta)\beta(\theta) < -u\beta(\theta)]\,\rho(\theta)d\theta$$

$$+ \int_{u>\alpha\sqrt{\log n}} P_\theta[\sqrt{n}(\hat{\theta}_n - \theta)\beta(\theta) < -u\beta(\theta)]\,\rho(\theta)d\theta$$

$$= I_1 + I_2 + I_3 .$$

Note that I_1 is $O(n^{-(S+1)m/S})$ as ρ is of type D_S and I_3 is $o(n^{-r_2/2})$ in view of Lemma 1.3.

In view of Lemma 1.4

$$I_2 = \int_{0<u<\alpha\sqrt{\log n}} [(1 - \Phi(u\beta(\theta))) + \frac{v(u,\theta)}{\sqrt{n}}] (\frac{u}{\sqrt{n}} + \frac{C}{n^{m/S}})^S \frac{du}{\sqrt{n}}$$

for sufficientely large n ;

where $|v(u,\theta)| < K_1\ \forall u,\ \forall \theta \in [0,1]$, for some K_1 and for sufficiently large n.

$$= \int_{0 < u < \alpha \sqrt{\log n}} [1 - \Phi(u\beta(\theta))] \left(\frac{u}{\sqrt{n}} + \frac{C}{n^{m/S}}\right)^S \frac{du}{\sqrt{n}} + o(n^{-m-\frac{1}{2}})$$

$$= n^{-m-\frac{1}{2}} \int \left[1 - \Phi\left(u\beta\left(\frac{u}{\sqrt{n}} + \frac{C}{n^{m/S}}\right)\right)\right] I_{(0,\alpha\sqrt{\log n})} (C + un^{-1/2+m/S})^S du$$

$$+ o(n^{-m-\frac{1}{2}})$$

Let $\beta' = \inf_{\theta \in [0,1]} \beta(\theta)$, then $0 < \beta' < \infty$. Now if $2m < S$, then for sufficiently large n the integrand is dominated by $C_1^S [1 - \Phi(u\beta')] I_{(0,\infty)}$ (for some $0 < C_1 < \infty$) which is integrable, showing $I_2 = o(n^{-m-1/2})$ by dominated convergence theorem. Similar arguments hold for $2m = S$. Other part of (1.8) can be proved analogously □

<u>PROPOSITION 1.2</u> : Let the assumptions AI to AIII and AV hold for some $r_2 \geq 2$. Let ρ be of type D_∞. Then for any C $(0 < C < \infty)$ we have

(1.11) $$P_\rho [\rho(\hat{\theta}_n) > C.n^{-m}] = 1 - O(n^{-m}) - o(n^{-r_2/2}).$$

<u>PROOF</u> : For simplicity let $a_o = 0$ and $b_o = 1$. Note that the inverse function ρ^{-1} can be defined in a neighbourhood of 0 and 1. Let ρ_1^{-1} and ρ_2^{-1} be the inverses of ρ near 0 and 1 respectively.

$$P_\rho \left[\rho(\hat{\theta}_n) < C.n^{-m}\right]$$

$$= E_\rho P_\theta \left[\rho(\hat{\theta}_n) < C.n^{-m}\right]$$

$$= E_\rho P_\theta \left[\rho_1^{-1}(C.n^{-m}) > \hat{\theta}_n\right] + E_\rho P_\theta \left[\rho_2^{-1}(C.n^{-m}) < \hat{\theta}_n\right]$$

As in the Proposition 1.1 and using the same notations as there we need to prove

$$(1.12) \qquad I_2 = \int_{0<u<\alpha\sqrt{\log n}} \left[1-\Phi(u\beta(\theta))\right] \rho\left(\frac{u}{\sqrt{n}} + \rho_1^{-1}(C.n^{-m})\right)\frac{du}{\sqrt{n}} = o(n^{-m})$$

Now, $\rho\left(\frac{u}{\sqrt{n}} + \rho_1^{-1}(C.n^{-m})\right) = C.n^{-m} + \frac{u}{\sqrt{n}}\,\rho^{(1)}\left(\rho_1^{-1}(C.n^{-m})\right) + \dots$

$$\dots + \left(\frac{u}{\sqrt{n}}\right)^{[2m]+1} \cdot \frac{1}{([2m]+1)!}\,\rho^{([2m]+1)}(\theta_n^*)$$

where $[x]$ denotes the integral part of x and θ_n^* is between $\frac{u}{\sqrt{n}} + \rho_1^{-1}(C.n^{-m})$ and $\rho_1^{-1}(C.n^{-m})$. Let $\rho_1^{-1}(C.n^{-m}) = \theta_n$ i.e. $\rho(\theta_n) = C.n^{-m}$ (which implies that θ_n is near zero). Since ρ is of type D_∞, $n^{m-\frac{r}{2}}\,\rho^{(r)}\left(\rho_1^{-1}(C.n^{-m})\right) = \rho^{\frac{r}{2m}-1}(\theta_n)\rho^{(r)}(\theta_n)C^{1-\frac{r}{2m}} \to 0.$ Thus $n^{-m}\rho\left(\frac{u}{\sqrt{n}} + \rho_1^{-1}(C.n^{-m})\right)$ can be expressed as a polynomial of degree $([2m]+1)$ in u with bounded coefficients for sufficiently large n. Now using similar arguments as in the proof of (1.9) proof of (1.12) is completed. □

<u>THEOREM 1.2</u> : Let assumptions AI to AV hold for some $r_1 \geq 3$ and $r_2 \geq 2$. Let ρ be of type D_{K+2} then for $0 < \ell < \frac{K+1}{2}$ $\exists$ $\varepsilon > 0$

such that

$$(1.13)\quad P_\rho\Big[\,|F_n(\xi)-\Phi(\xi)-\sum_{j=1}^{K}\gamma_j(\xi,\underline{x}_n).n^{-j/2}|\le M.n^{-\lambda}$$

$$|\gamma_j(\xi,\underline{x}_n).n^{-j/2}|\le M.n^{-\varepsilon.j}\quad 1\le j\le K \quad \text{uniformly in } \xi\Big]$$

$$= 1-O(n^{-\frac{K+3}{K+2}(\frac{K+1}{2}-\lambda)})-O(n^{-r_1/2})-o(n^{-r_2/2}).$$

<u>PROOF</u> : For simplicity let $a_o = 0$ and $b_o = 1$. Let

$$B_n = \Big[\hat{\theta}_n > C.n^{-\frac{(K+1)}{2(K+2)}+\frac{\lambda}{K+2}}\Big]\bigcap\Big[1-\hat{\theta}_n > C.n^{-\frac{(K+1)}{2(K+2)}+\frac{\lambda}{K+2}}\Big]$$

hence using (1.8) for some C'

$$(1.14)\qquad B_n \Longrightarrow \rho(\hat{\theta}_n) > C'.n^{-\frac{(K+1)}{2}+\lambda}.$$

Now note that on B_n

$$\text{Max}\,(\hat{\theta}_n^{-1}n^{-1/2},(1-\hat{\theta}_n)^{-1}n^{-1/2}) < C.n^{-\varepsilon'} \quad \text{where} \quad \varepsilon' = \Big[\frac{1}{2}-\frac{1}{K+2}(\frac{K+1}{2}-\lambda)\Big]$$

Hence on B_n

$$\Big|\frac{\rho^{(i)}(\hat{\theta}_n)}{\rho(\hat{\theta}_n)}\Big| \le M(\hat{\theta}_n)^{-i} \le M.n^{(\frac{1}{2}-\varepsilon)i}$$

So in view of (1.14) , Remark 1.1 and (1.7)

$$B_n\bigcap A_n \Longrightarrow (|C_n|^{-1}\le M.n^{-\lambda+\frac{K+1}{2}},\ |C_n|^{-1}d_n \le Mn^{-\lambda+\frac{K+1}{2}})$$

$$\bigcap\Big[|\gamma_j(\xi,\underline{x}_n)n^{-j/2}|\le M.n^{-\varepsilon j}\quad 1\le j\le K$$

$$\text{uniformly in } \xi\Big]$$

for some $\varepsilon > 0$. This along with (1.4), (1.9) and Theorem 1.1 gives us the result □

REMARK 1.1 : Suppose AI to AV hold with K replaced by $K'(K' > K)$ for some $r_1 \geq 3$ and $r_2 \geq 2$ and ρ is of type $D_{K'+2}$. Then one would expect better probability bounds in (1.12). This indeed is the case. This can be achieved by first getting a $(K'+1)$ term expansion for $F_n(\xi)$ with $M.n^{-\lambda}(0 < \lambda < \frac{K+1}{2})$ as the error term and then bounding $\sum_{j=K+1}^{K'} |\gamma_j(\xi, \underline{x}_n) n^{-j/2}|$ by $M.n^{-\lambda}$. The probability bound for the statement of (1.12) new turns out to be

$$1 - O(n^{-\frac{(K'+3)}{K'+2}(\frac{K'}{2}-\lambda)}) - O(n^{-(K'+3)(\frac{1}{2}-\frac{\lambda}{K+1})})$$
$$- O(n^{-r_1/2}) - o(n^{-r_2/2}),$$

which clearly becomes better with larger K'.

Below we investigate the case when ρ is of type D_∞.

THEOREM 1.3 : Let the assumptions AI to AV hold for some $r_1 \geq 3$ and $r_2 \geq 2$ and ρ be of type D_∞ then for any $0 < \lambda < \frac{1+1}{2} \leq \frac{K+1}{2}$ and $\varepsilon > 0$

$$P_\rho \Big[|F_n(\xi) - \Phi(\xi) - \sum_{i=1}^{j} \gamma_i(\xi, \underline{x}_n) n^{-i/2}| \leq M.n^{-\lambda},$$

$$|\gamma_i(\xi, \underline{x}_n) n^{-i/2}| \leq M.n^{-\frac{i}{2}+\varepsilon} \quad 1 \leq i \leq j \quad \text{uniformly in } \xi \Big]$$

$$= 1 - O(n^{-(K+1)/2+\lambda}) - O(n^{-r_1/2}) - o(n^{-r_2/2}).$$

PROOF : By (iii) of Definition 1.2 we have for any $v > 0$

$$|\rho^{(i)}(\hat{\theta}_n)\rho^{-1}(\hat{\theta}_n)| = |\rho^{(i)}(\hat{\theta}_n)\rho^{v-1}(\hat{\theta}_n)\rho^{-v}(\hat{\theta}_n)| \leq M.\rho^{-v}(\hat{\theta}_n)$$

Hence using (1.4) and Remark 1.1 for any $\varepsilon > 0$ and $m > 0$

$$(\rho(\hat{\theta}_n) > n^{-m}) \bigcap A_n \Longrightarrow (|C_n|^{-1} \leq M.n^{-m}, \sum_{i=j+1}^{K} |\gamma_i(\xi,\underline{x}_n)n^{-i/2}| \leq M.n^{-\lambda},$$

$$|C_n|^{-1}d_n \leq M.n^{-m} \quad |\gamma_i(\xi,\underline{x}_n)n^{-i/2}| \leq M.n^{-i/2+\varepsilon} \quad 1 \leq i \leq j)$$

This along with (1.4), (1.11) (with $m = \frac{K+1}{2} - \lambda$) and (1.1) gives the result.

1.3 ASSUMPTIONS : Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables having common distribution function $F(x,\theta)$, parametrized by $\theta \varepsilon \underline{(H)}$, $\underline{(H)}$ being an open interval of R. Let $f(x,\theta)$ be the density of $F(x,\theta)$ w.r.t a σ-finite measure μ. Let the parameter θ have a prior distribution which has density $\rho(\cdot)$ w.r.t. Lebesgue measure. Let $[a_o, b_o] \subseteq \underline{(H)}$ be such that $\rho(\theta) > 0$ on (a_o, b_o) and $\rho(\theta) = 0$ on $(a_o, b_o)^c$. Let $c < a_o$ and $d > b_o$ be such that $[c,d] \subseteq \underline{(H)}$. We need to make following assumptions.

AI : $f(x,\theta)$ is measurable in x for each $\theta \varepsilon [c,d]$

AII : $\int |f(x,\theta) - f(x,\theta')|d\mu(x) > 0 \;\forall\; \theta,\theta' \varepsilon [c,d], \theta \neq \theta'$

AIII : For each x, $f(x,\theta)$ admits partial derivatives w.r.t. θ of order upto two which are continuous in $[c,d]$.

AIV : 1) For each x, $f(x,\theta)$ admits partial derivatives of order upto $K+3$ $(K \geq 0)$ which are continuous in $[c,d]$.

2) For every $\theta \in [a_o, b_o]$ there exists a neighbourhood (nhbd) U_θ such that

$$\sup_{\theta' \in U_\theta} E_{\theta'} \left| \frac{d^i}{d\theta^i} \log f(X,\theta') \right|^{r_1} < \infty \quad 1 \leq i \leq K+2$$

3) For every $\theta \in [a_o, b_o] \exists$ a nhbd U_θ and measurable functions $H_\theta(x)$ and $A_\theta(x)$ such that

a) $\left| \frac{d^{K+3}}{d\theta^{K+3}} \log f(x,\theta') \right| \leq H_\theta(x) \; \forall \theta' \in U_\theta$ and $\forall x$,

b) $\left| \frac{d^{K+3}}{d\theta^{K+3}} \log f(x,\theta') - \frac{d^{K+3}}{d\theta^{K+3}} \log f(x,\theta'') \right| \leq |\theta' - \theta''| A_\theta(x)$

$\forall \theta', \theta'' \in U_\theta$ and $\forall x$

c) $\sup_{\theta' \in U_\theta} E_{\theta'} (H_{\theta'}(X)^{r_1}) < \infty$

d) $\sup_{\theta' \in U_\theta} E_{\theta'} (A_{\theta'}(X)) < \infty$

AV : 1) For every $\theta \in (c,d) \exists$ a nhbd U_θ such that

$$\sup_{\theta' \in U_\theta} E_{\theta'} (|\log f(X,\theta)|^{r_2+1}) < \infty$$

2) $E_\theta(\frac{d}{d\theta} \log f(X,\theta)) = 0 \; \forall \theta \in (c,d)$

3) For every $\theta \varepsilon [c,d]$, $\theta' \varepsilon (c,d)$ $\exists$ nhbds V_θ and $W_{\theta'}$ such that for all nhbds $V \subseteq V_\theta$ of θ,

$$\sup_{\alpha \varepsilon W_{\theta'}} E_\alpha\left(\left|\sup_{\sigma \varepsilon V} \log f(X,\sigma)\right|^{r_2+1}\right) < \infty$$

4) For every $\theta \varepsilon (c,d)$

a) $I(\theta) = E_\theta\left(-\frac{d^2}{d\theta^2} \log f(X,\theta)\right) > 0$

b) $I_1(\theta) = E_\theta\left(\frac{d}{d\theta} \log f(X,\theta)\right)^2 > 0$

5) $I(\theta)$ and $I_1(\theta)$ are continuous on (c,d).

6) For every $\theta \varepsilon (c,d)$ $\exists$ a nhbd U_θ and a measurable function $m(x,\theta)$ such that

a) $\left|\frac{d^2}{d\theta^2} \log f(x,\theta') - \frac{d^2}{d\theta^2} \log f(x,\theta'')\right| \leq |\theta' - \theta''| m(x,\theta)$

$\forall\ \theta', \theta'' \varepsilon U_\theta\ \forall x$,

b) $\sup_{\theta' \varepsilon U_\theta} E_{\theta'}(m(X,\theta)^{r_2+1}) < \infty$

c) $\sup_{\theta' \varepsilon U_\theta} E_{\theta'}\left(\left|\frac{d^2}{d\theta^2} \log f(X,\theta')\right|^{r_2+1}\right) < \infty$.

7) For every $\theta \varepsilon (c,d)$ $\exists$ a nhbd U_θ such that

$$\sup_{\theta' \varepsilon U_\theta} E_{\theta'}\left(\left|\frac{d}{d\theta} \log f(X,\theta')\right|^{r_2+2}\right) < \infty.$$

Note that if AIV holds then on $[a_o, b_o]$ $I(\theta)$ equals $I_1(\theta)$ and is continuous (vide Gusev [1975]).

1.4 SOME LEMMAS : In this section we collect all the lemmas needed for the proofs of the results in the Section 2. Lemmas 1.7 and 1.8 are versions of Lemmas 2.2 and 2.4 of Johnson [1970] respectively ; also (1.17) and (1.18) are versions of his (2.21) and (2.22) respectively. For the sake of completeness we sketch their proofs. Following is a version of the Lemma 0 of Ghosh, Sinha and Wieand [1980].

LEMMA 1.1 : Let C be a compact interval and let $U(x,t)$ be a real valued function measurable in x for each $t \varepsilon C$ and continuous in t for each x. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. r.v.'s having a common d.f. F_θ, $\theta \varepsilon \underline{H}$ and let $H(x)$ and $A(x)$ be measurable functions such that

a) $|U(x,t)| \leq H(x) \; \forall \; t \varepsilon C, \; \forall \; x,$

b) $|U(x,t) - U(x,t')| \leq |t - t'| A(x) \; \forall \; t,t' \varepsilon C \; \forall \; x,$

c) $\sup\limits_{\theta \varepsilon \underline{H}} E_\theta(H(X)^r) < \infty, \; r \geq 3$

d) $\sup\limits_{\theta \varepsilon \underline{H}} E_\theta(A(X)) < \infty$

Then for given $\varepsilon > 0 \; \exists \; K(\varepsilon)(0 < K(\varepsilon) < \infty)$ and n_o such that

$$P_{F_\theta}\left[\sup_{t \varepsilon C} \left|n^{-1} \sum_{i=1}^{n} U(x_i,t) - E_{F_\theta}(U(X,t))\right| < \varepsilon\right]$$

$$\geq 1 - K(\varepsilon) n^{-r/2} \; \forall \; \theta \varepsilon \underline{H} \; \forall \; n \geq n_o .$$

LEMMA 1.2 : Let assumptions AI to AIII and AV hold for some $r_2 \geq 2$.

Then for $\varepsilon > 0$ and $\delta > 0$, uniformly on compact subsets of (c,d), we have

$$(1.15) \qquad P_\theta\Big[|\hat{\theta}_n - \theta| \le \varepsilon,\ \sum_{i=1}^{n} \frac{d}{d\theta} \log f(x_i, \hat{\theta}_n) = 0 \text{ and}$$

$$\frac{1}{n}\sum_{i=1}^{n} (\log f(x_i, \hat{\theta}_n + \emptyset) - \log f(x_i, \hat{\theta}_n)) < -\varepsilon\ \forall |\emptyset| > \delta\Big]$$

$$= 1 - o(n^{-r_2/2})$$

where the m.l.e. $\hat{\theta}_n$ is such that

$$\sup_{\theta \in [c,d]} \prod_{i=1}^{n} f(x_i, \theta) = \prod_{i=1}^{n} f(x_i, \hat{\theta}_n).$$

PROOF : (1.15) with r_2 replaced by 2 can be obtained from Lemma 4 of Michel and Pfanzagl [1971]. Now (1.15) can be obtained as suggested in proof of Lemma 3 of Pfanzagl [1973a]. □

Following is Lemma 3 of Pfanzagl [1973a].

LEMMA 1.3 : Let assumptions AI to AIII and AV hold for some $r_2 \ge 2$. Then $\exists$ constant $c_1 (0 < c_1 < \infty)$ such that uniformly on compact subsets of (c,d) :

$$P_\theta\left[\frac{|\hat{\theta}_n - \theta|}{\beta(\theta)} \le c_n\right] = 1 - o(n^{-r_2/2})$$

where $\beta(\theta) = I^{1/2}(\theta)\, I_1^{-1}(\theta)$ and $c_n = c_1 n^{-1/2}(\log n)^{1/2}$.

Following is the Theorem of Pfanzagl [1973b].

<u>LEMMA 1.4</u> : Let assumptions AI to AIII and AV hold for some $r_2 \geq 2$. Then uniformly in $\theta \varepsilon [a,b]$ and uniformly over all measurable convex sets E.

$$|P_\theta [n^{1/2}\beta^{-1}(\theta)(\hat{\theta}_n - \theta) \varepsilon E] - \Phi(E)| = 0(n^{-1/2})$$

where Φ denotes probability measure corresponding to a standard normal variable as well as its d.f..

<u>LEMMA 1.5</u> : Let assumptions AI to AV hold for some $r_1 \geq 3$ and $r_2 \geq 2$. Then $\exists$ M $(0 < M < \infty)$ such that uniformly in $\theta \varepsilon [a_o , b_o]$:

$$P_\theta [-M < a_{k,n}(\hat{\theta}_n) < M] = 1 - o(n^{r_2/2}) - 0(n^{-r_1/2}) \text{ for } k = 1,\ldots,K+3$$

where $a_{k,n}(\theta) = \frac{1}{k!n} \sum_{i=1}^{n} \frac{d^k}{d\theta^k} \log f(x_i , \theta)$

<u>PROOF</u> : Applying Lemma 1.1 to $a_{k,n}$ $(k = 1,\ldots,K+3)$, for every $\delta > 0$ we get M $(0 < M < \infty)$ such that uniformly in $\theta \varepsilon [a_o , b_o]$ we have

$$P_\theta [-M < a_{kn}(\theta') < M \text{ for } \forall |\theta - \theta'| < \delta$$
$$\text{and } k = 1,\ldots,K+3] = 1 - 0(n^{-r_1/2}).$$

Combinging this with Lemma 1.2 we get the result □

<u>LEMMA 1.6</u> : Let assumptions AI to AIV hold with $r_1 \geq 3$. Then for $\varepsilon > 0 \ \exists \ \delta > 0$ such that uniformly in $\theta \varepsilon [a_o , b_o]$

$$P_\theta [E_\theta(-2a_{2,n}(\theta)) - \varepsilon \leq -2a_{2,n}(\theta') \leq E_\theta(-2a_{2,n}(\theta)) + \varepsilon$$
$$\forall |\theta - \theta'| < \delta] = 1 - 0(n^{-r_1/2}).$$

PROOF : Applying Lemma 1.1 to $-2a_{2,n}(\theta)$ we get $\delta_1 > 0$ such that uniformly in $\theta \in [a_o, b_o]$ we have

$$P_\theta \left[E_\theta(-2a_{2,n}(\theta')) - \frac{\varepsilon}{2} \leq -2a_{2,n}(\theta') \leq E_\theta(-2a_{2,n}(\theta')) + \frac{\varepsilon}{2} \ \forall \ |\theta - \theta'| < \delta_1 \right]$$

$$= 1 - O(n^{-r_1/2}) .$$

Now $E_\theta(-2a_{2,n}(\theta')) = E_\theta(-2a_{2,n}(\theta)) + (\theta' - \theta)E_\theta(-a_{3n}(\theta)) \frac{3!}{2} + \ldots$

$$\ldots + \frac{(K+2)!}{(K-1)!} (\theta' - \theta)^{K-1} E_\theta(-a_{K+2,n}(\theta)) + \frac{(K+3)!}{K!} (\theta' - \theta)^K E_\theta(-a_{K+3,n}(\theta_1))$$

where θ_1 lies between θ and θ' hence in view of AIV $\exists \ \delta_2 > 0$ such that

$$|E_\theta(-2a_{2,n}(\theta')) - E_\theta(-2a_{2,n}(\theta))| < \frac{\varepsilon}{2} \ \forall \ |\theta - \theta'| < \delta_2 \ \forall \ \theta \in [a_o, b_o] \ \Box$$

LEMMA 1.7 : Let assumptions AI to AV hold for some $r_1 \geq 3$ and $r_2 \geq 2$. Then $\exists \ \delta_2 > 0$ such that uniformly in $\theta \in [a_o, b_o]$:

$$P_\theta \left[n^{-1} \sum_{i=1}^{n} (\log f(x_i, \hat{\theta}_n + \phi b^{-1}) - \log f(x_i, \hat{\theta}_n)) \leq -\frac{\phi^2}{6} \ \forall \ |\phi| \leq \delta_2 \right]$$

$$= 1 - o(n^{-r_2/2}) - O(n^{-r_1/2}) .$$

Here $b^2(\theta) = (-2a_{2,n}(\theta))$ and $b^2 = b^2(\hat{\theta}_n)$.

PROOF : Note that by using a two term Taylor's expansion and Lemma 1.2 we have for $\varepsilon > 0$

$$P_\theta \Big[n^{-1} \Sigma (\log f(x_i , \hat{\theta}_n + \phi b^{-1}) - \log f(x_i , \hat{\theta}_n))$$

$$= - \frac{\phi^2}{2} (a_{2n}(\hat{\theta}_n + \beta\phi b^{-1})/a_{2n}(\hat{\theta}_n)) \text{ where } 0 \leq \beta \leq 1 \text{ and } |\hat{\theta}_n - \theta| \leq \varepsilon \Big]$$

$$= 1 - o(n^{-r_2/2}) \quad \text{uniformly in} \quad \theta \in [a_o, b_o].$$

Now using Lemma 1.2 and the continuity of $I(\theta)$ the proof is completed □

REMARK 1.3 : Note that (1.16), (1.17) and (1.18) below are similar to the (2.11), (2.21) and (2.22) of Johnson [1970] respectively with slightly different remainder R_n. We need the remainder in this form mainly to display clearly the contribution of ρ to the remainder. In what follows

$$R_n = \sum_{r=1}^{r_o} n^{-(K+r)/2} |\rho^{(K+r)}(\hat{\theta}_n)| + n^{-(K+1)/2} \sum_{r=0}^{K} |\rho^{(r)}(\hat{\theta}_n)|$$

$$+ n^{-(K+r_o+1)/2} \qquad \text{with} \quad \rho^{(0)}(\theta) = \rho(\theta), \; \rho^{(k)}(\theta) = \frac{d^k}{d\theta^k} \rho(\theta).$$

LEMMA 1.8 : Let assumptions AI to AV hold with some $r_1 \geq 3$ and $r_2 \geq 2$. Let ρ be $K+r_o+1$ times continuously differentiable on $[a_o, b_o]$ for some $r_o \geq 0$. Then $\exists\, M \; (0 < M < \infty)$ such that for sufficiently small $\delta > 0$ and uniformly in $\theta \in [a_o, b_o]$ we have

(1.16) $$P_\theta \Big[\int_{-\delta}^{\delta} \Big| \exp\Big[n \sum_{k=2}^{K+3} a_{k,n}(\hat{\theta}_n)(\phi b^{-1})^k\Big] \rho_K(\hat{\theta}_n + \phi b^{-1})$$

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

$$\leq M . R_n n^{-1/2} \Big]$$

$$= 1 - o(n^{-r_2/2}) - o(n^{-r_1/2}).$$

Here $\rho_K(\theta) = \rho(\hat{\theta}_n) + \sum_{r=1}^{K} \frac{(\theta - \hat{\theta}_n)^r}{r!} \rho^{(r)}(\hat{\theta}_n)$.

PROOF : Add and substract

$$\rho_{K+r_o}(\hat{\theta}_n + \phi b^{-1}) \prod_{i=1}^{n} [f(x_i, \hat{\theta}_n + \phi b^{-1})/f(x_i, \hat{\theta}_n)]$$

to the integrand of (1.16). It follows from Lemmas 1.5 and 1.7 and the inequality

$$|\log\left\{\prod_{i=1}^{n} [f(x_i, \hat{\theta}_n + \phi b^{-1})/f(x_i, \hat{\theta}_n)]\right\} - n \sum_{k=2}^{K+3} a_{kn}(\hat{\theta}_n)(\phi/b)^k| \leq M.n|\phi|^{K+3}$$

that there exists a constant M_2 such that for sufficiently small $\delta > 0$ and uniformly in $\theta \in [a_o, b_o]$ we have

$$P_\theta\left\{\int_{-\delta}^{\delta} |\exp\left[n \sum_{k=1}^{K+3} a_{k,n}(\hat{\theta}_n)(\phi b^{-1})^k\right] \rho_K(\hat{\theta}_n + \phi b^{-1}) - \prod_{i=1}^{n} [f(x_i, \hat{\theta}_n + \phi b^{-1})/f(x_i, \hat{\theta}_n)] \rho(\hat{\theta}_n + \phi b^{-1})| d\phi\right.$$

$$\leq \int_{-\delta}^{\delta} e^{-n\phi^2/2} \left[\sum_{r=1}^{r_o} \frac{|\phi b^{-1}|^{K+r}}{(K+r)!} |\rho^{(K+r)}(\hat{\theta}_n)| + M_2 \frac{|\phi b^{-1}|^{K+r_o+1}}{(K+r_o+1)!} |\rho^{(K+r_o+1)}(\theta_n^*)|\right] d\phi \left.\right\}$$

$$= 1 - O(n^{-r_1/2}) - o(n^{-r_2/2}) \quad .$$

where θ_n^* is between $\hat{\theta}_n$ and $\hat{\theta}_n + \phi b^{-1}$. Now making a change of variable $u = n^{1/2}\phi$ and using the continuity of derivatives of ρ we get (1.16) □

LEMMA .9 : Let assumptions AI to AV hold with $r_1 \geq 3$ and $r_2 \geq 2$ and let ρ be $K+r_o+1$ times continuously differentiable on $[a_o, b_o]$ for some $r_o \geq 0$. Then $\exists$ M $(0 < M < \infty)$ such that uniformly in $\theta \varepsilon [a_o, b_o]$ we have

(1.17) $$P_\theta \Big[\Big| \int_{-\infty}^{\infty} \rho(\hat{\theta}_n + \phi b^{-1}) \prod_{i=1}^{n} [f(x_i, \hat{\theta}_n + \phi b^{-1})/f(x_i, \hat{\theta}_n)] d\phi$$

$$- \sum_{j=0}^{K} \beta_j(\underline{x}_n) n^{-(j+1)/2} \Big| \leq M.n^{-1/2} R_n \Big]$$

$$= 1 - O(n^{-r_1/2}) - o(n^{-r_2/2})$$

and

(1.18) $$P_\theta \Big[\Big| \int_{-\infty}^{\xi . n^{-1/2}} \rho(\hat{\theta}_n + \phi b^{-1}) \prod_{i=1}^{n} [f(x_i, \hat{\theta}_n + \phi b^{-1})/f(x_i ; \hat{\theta}_n)]$$

$$- \sum_{j+0}^{K} \omega_j(\xi, \underline{x}_n) n^{-(j+1)/2} \Big| \leq M.n^{-1/2} R_n \text{ uniformly in } \xi \Big]$$

$$= 1 - O(n^{-r_1/2}) - o(n^{-r_2/2}).$$

PROOF : Let $|a_{kn}(\hat{\theta}_n)| \leq M$ $k = 3, \ldots, K+3$, $0 < \delta < b < M$ and $\phi_{K,n}(\phi) = n \sum_{k=3}^{K+3} a_{kn}(\hat{\theta}_n) \phi^{k-3} b^{-k}$ then

$$\rho_K(\hat{\theta}_n + \phi b^{-1}) e^{\psi_{K,n}(\phi)} = (\rho(\hat{\theta}_n) + \phi b^{-1} \rho^{(1)}(\hat{\theta}_n) + \dots + (\phi/b)^K \rho^{(K)}(\hat{\theta}_n))$$

$$\left\{ 1 + \psi_{K,n}(\phi) + \dots + \frac{(\psi_{K,n}(\phi))^K}{K!} + \frac{(\psi_{K,n}(\phi))^{K+1}}{(K+1)!} e^{\psi^*} \right\}$$ where ψ^* is

between 0 and $\psi_{K,n}(\phi)$. Now we collect all the terms in the above expression which after substituting $\phi = n^{-1/2}$ are of order $n^{-(K+1)/2}$ or smaller ; denoting the sum of the rest of the terms by $P_K(\phi, \underline{x}_n)$ we have by using Lemma 1.5

$$P_\theta \Big[|\rho_K(\hat{\theta}_n + \phi b^{-1}) e^{\psi_{K,n}(\phi)} - P_K(\phi, \underline{x}_n)|$$

$$\leq M \sum_{r=0}^{K} |\rho^{(r)}(\hat{\theta}_n)| |\phi|^r (|\phi|^{K+1} + |n\phi^3|^{K+1} + \sum_{K-r+1 \leq i+j \leq K} |\phi|^i |n\phi^3|^j)$$

$$\text{for } |\phi| < n^{-1/3} \Big] = 1 - O(n^{-r_1/2}) - o(n^{-r_2/2}).$$

Note that $P_K(\phi, \underline{x}_n) = \sum_{0 \leq \ell + m \leq K} c_{\ell,m}(\underline{x}_n) \phi^\ell (n\phi^3)^m$ where

$$\ell!\, m!\, c_{\ell,m}(\underline{x}_n) = \frac{\partial^{\ell+m}}{\partial^\ell w\, \partial^m z} \rho_K(\hat{\theta}_n + z b^{-1}) e^{w \psi_{K,n}(z)} \Big|_{\substack{z=0 \\ w=0}} \quad \ell, m = 0, 1, \dots$$

Now using Lemmas 1.5 to 1.8 it is easy to see that (vide (2.19) and (2.20) of Johnson [1970])

$$(1.19) \quad P_\theta \Big[\Big| \int_{-\infty}^{\infty} \rho(\hat{\theta}_n + \phi b^{-1}) \prod_{i=1}^{n} [f(x_i, \hat{\theta}_n + \phi b^{-1}) / f(x_i, \hat{\theta}_n)] d\phi$$

$$- \int_{-\infty}^{\infty} e^{-n\phi^2/2} P_K(\phi, \underline{x}_n) d\phi \Big| \leq M . n^{-1/2} R_n \Big]$$

$$= 1 - O(n^{-r_1/2}) - o(n^{-r_2/2})$$

(1.20) $$= P_\theta \Big[\Big| \int_{-\infty}^{\xi n^{-1/2}} \rho(\hat{\theta}_n + \phi b^{-1}) \prod_{i=1}^{n} [f(x_i, \hat{\theta}_n + \phi b^{-1})/f(x_i, \hat{\theta}_n)] d\phi$$

$$- \int_{-\infty}^{\xi n^{-1/2}} e^{-n\phi^2/2} P_K(\phi, \underline{x}_n) d\phi \Big| \le M.n^{-1/2} R_n$$

$$\text{uniformly in } \xi \Big]$$

Now integrating the approximation and collecting the terms we get (1.17) and (1.18) □

REMARK 1.4 : A change of variables in (1.20) gives

$$\alpha_j(\xi, \underline{x}_n) = \sum_{s=0}^{j} c_{s,j-s} \int_{-\infty}^{\xi} y^{2s+j} e^{-y^2/2} dy, \quad j = 0,1,\ldots,K$$

and $\beta_j(\underline{x}_n)$ corresponds to $\alpha_j(\infty, \underline{x}_n)$ and thus is zero when j is odd. Note that if we write (vide (2.24) to (2.26) of Johnson [1970])

$$\frac{\sum_{j=0} \alpha_j(\xi, \underline{x}_n) n^{-j/2}}{\sum_{j=0} \beta_j(\underline{x}_n) n^{-j/2}} = \sum_{j=0} \gamma_j(\xi, \underline{x}_n) n^{-j/2} \quad \text{then} \quad \gamma_0(\xi, \underline{x}_n) = \Phi(\xi),$$

$$\gamma_1(\xi, \underline{x}_n) = -\phi(\xi) c_{00}^{-1} [c_{10}(\xi^2 + 2) + c_{01}],$$

$$\gamma_2(\xi, \underline{x}_n) = -\phi(\xi) c_{00}^{-1} [c_{20}\xi^5 + (5c_{20} + c_{11})\xi^3 + (15c_{20} + 3c_{11} + c_{02})\xi],$$

$$c_{00} = \rho(\hat{\theta}_n), \; c_{01} = b^{-1}\rho^{(1)}(\hat{\theta}_n), \; c_{02} = b^{-2}\rho^{(2)}(\hat{\theta}_n),$$

$c_{10} = b^{-3}a_{3n}(\hat{\theta}_n)\rho(\hat{\theta}_n)$, $c_{11} = b^{-4}a_{4n}(\hat{\theta}_n)\rho(\hat{\theta}_n) + b^{-4}a_{3n}(\hat{\theta}_n)\rho^{(1)}(\hat{\theta}_n)$

and $c_{20} = 2^{-1}b^{-6}a_{3n}^2(\hat{\theta}_n)\rho(\hat{\theta}_n)$. The coefficients $\{\gamma_j(\xi, \underline{x}_n)\}$ satisfy

$$\alpha_j(\xi, \underline{x}_n) = \beta_0(\underline{x}_n)\gamma_j(\xi, \underline{x}_n) + \sum_{s=1}^{j-1} \gamma_{j-s}(\xi, \underline{x}_n)\beta_s(\underline{x}_n) + \beta_j(\underline{x}_n)\Phi(\xi)$$

for $j = 1, 2, \ldots, K$; and they do not involve $\Phi(\xi)$ (this can be easily seen from the fact that

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

does not involve $\Phi(\xi)$).

1.5 COUNTER EXAMPLE :

In this section we give an example to show that

$$P_\rho\left[|F_n(\xi) - \Phi(\xi)| \leq M.n^{-1/2} \text{ uniformly in } \xi\right]$$

cannot be made of order $1 - o(n^{-r})$ for all $r > 0$ even in the case where $X_1, X_2, \ldots, X_n$ are i.i.d. $N(\theta, 1)$ (An example on similar lines can be constructed to show that (1.2)' does not hold for all $r > 0$ in general, where (1.2)' denotes (1.2) with error term $M.n^{-(K+1)/2}.|C_n|^{-1}\ell_n$ replaced by $M.n^{-(K+1)/2}$).

EXAMPLE 1.1 : Consider $X_1, X_2, \ldots, X_n$ i.i.d. $N(\theta, 1)$ variables, $-\infty < \theta < \infty$ and a prior

$$\rho(\theta) = K.\theta^5(1-\theta)^5 \quad \text{for } 0 \leq \theta \leq 1$$
$$= 0 \quad \text{elsewhere}$$

Now, $0 = P[\theta < 0|\underline{x}_n] = P[\sqrt{n}(\theta - \bar{x}_n) < -\sqrt{n}.\bar{x}_n|\underline{x}_n]$

where $\hat{\theta}_n = n^{-1} \sum_{i=1}^{n} x_i = \bar{x}_n$.

Hence $F_n(\xi) = 0$ for $\xi = -\sqrt{n}\,\bar{x}_n$ and to get a counterexample it is enough to show that for some $0 < \varepsilon < \frac{1}{2}$

(1.21) $\quad P_\rho [\bar{\emptyset}(\xi) > n^{-1/2+\varepsilon}]$ cannot be made of order n^{-r} for all $r > 0$.

Note that if $c < (2-4\varepsilon)^{1/2}$ and n is sufficiently large then

$$P_\rho [\bar{\emptyset}(\xi) > n^{-1/2+\varepsilon}] \geq P_\rho [-\sqrt{n}\,\bar{x}_n > -c\sqrt{\log n}]$$

$$= \int P_\theta [\sqrt{n}(\bar{x}_n - \theta) < c\sqrt{\log n} - \sqrt{n}.\theta]\, \rho(\theta)d\theta$$

$$\geq \frac{1}{2} \int_{0 < \theta < dn^{-1/2}\sqrt{\log n}} \rho(\theta)d\theta \qquad \text{if } (c-d) > 0$$

$$= K(\log n)^3 n^{-3}$$

which proves (1.21).

1.6 <u>RESULTS UNDER BURNASEV TYPE OF CONDITIONS</u> : In this section we consider priors which satisfy Burnasev type of conditions (vide Sec. 5 Ch.2) and show how a version of Theorem 1.2 can be obtained.

Let assumptions AI to AV hold for some $r_1 \geq 3$ and $r_2 \geq 2$. Let the prior ρ be three times continuously differentiable on $[a_0, b_0]$ and let (i) below hold for some $\delta' > 0$ and $m_i > 0$, $i = 1, 2$

(i) $\int_{D^c(\varepsilon)} \eta_\theta^{m_1}(\varepsilon)\, \rho(\theta)d\theta = O(\varepsilon^{-m_2})$ as $\varepsilon \longrightarrow 0$

where $\eta_\theta(\varepsilon) = \left\{ (|\rho^{(1)}(\theta)| + |\rho^{(2)}(\theta)|)^2 \rho^{-2}(\theta) + \sup_{|z|<\varepsilon} \frac{|\rho^{(3)}(\theta+z)|}{\rho(\theta)} \right\}$

and $D(\varepsilon) = \left\{ \theta : |\rho^{(1)}(\theta)|\rho^{-1}(\theta) > \varepsilon^{-1+\delta'}\right.$ or $|\rho^{(2)}(\theta)|\rho^{-1}(\theta) > \varepsilon^{-2+\delta'}$

or $\left.\sup_{|z|<\varepsilon} |\rho^{(3)}(\theta+z)|\rho^{-1}(\theta) > \varepsilon^{-3+\delta'} \right\}$

Then

(1.22) $P_\rho \left[|F_n(\xi) - \sum_{j=1}^{2} \gamma_j(\xi, \underline{x}_n)\, n^{-j/2}| \leq M\, n^{-(3/2-\delta)} \right.$ $\left.\text{uniformly in } \xi \right]$

$$= 1 - O(n^{-r_1}) - o(n^{-r_2}) - O(n^{-(m_1\delta - m_2/2)}) - \int_{D(2c_n)} \rho(\theta)d\theta$$

c_n as in Lemma 1.3.

<u>PROOF</u> : Let $D_n = D(2c_n) \bigcup (\rho(\theta) < n^{-r_1})$ and

$A_{n\theta} = A_n \bigcap [|\hat{\theta}_n - \theta| < c_n]$, A_n of (1.4).

Note that by expanding $\rho(\hat{\theta}_n + \phi b^{-1})$ only upto 4 terms Lemmas 1.8 and 1.9 hold for $K = 2$ with R_n replaced by R_{1n},

$$R_{1n} = n^{-3/2} \left[e^{-n\varepsilon_1} + \sum_{i=0}^{2} |\rho^{(i)}(\hat{\theta}_n)| + \sup_{|z|<c_n} |\rho^{(3)}(\hat{\theta}_n + z)| \right]$$

for some $\varepsilon_1 > 0$. Now using Lemma 1.3 and expanding $\rho^{(i)}(\hat{\theta}_n)$ around θ upto (4-i) terms (i = 0 to 2) this R_{1n} can be replaced by $n^{-3/2} R_{2n}$, $R_{2n} = \rho(\theta) \left[e^{-n\varepsilon_1}\rho^{-1}(\theta) + \sum_{i=0}^{2} |\rho^{(i)}(\theta)|\rho^{-1}(\theta) + y_{n\theta} \right]$

where $y_{n\theta} = \sup_{|z| < 2c_n} |\rho^{(3)}(\theta + z)| \rho^{-1}(\theta)$.

Thus (1.3) holds for $K = 2$ with $|C_n|^{-1} \ell_n$ replaced by $|C_n|^{-1}(R_{2n} + d_n)$ (vide (1.2b)) where now $d_n = |\beta_2|(|\gamma_1| + n^{-1/2}|\gamma_2|)$

Note that on $A_{n\theta}$ and $\theta \in D_n^c$ by expanding $\rho^{(i)}(\hat{\theta}_n)$ around θ upto (4-i) terms ($i = 0$ to 2) $|C_n|^{-1} \leq M \rho^{-1}(\theta)$ and $|C_n|^{-1} d_n \leq M \eta_{n\theta}$ where $\eta_{n\theta} = \eta_\theta (2c_n)$.

Also on $A_{n\theta}$, $\theta \in D_n^c$ $|C_n|^{-1} R_{2n} < M \eta_{n\theta}$.

Thus (1.3) holds for $K = 2$ with $|C_n|^{-1} \ell_n$ replaced by $\eta_{n\theta}$.

Now using (i), (1.22) follows $\square$

CHAPTER TWO

EXPANSIONS FOR POSTERIOR RISK AND INTEGRATED BAYES RISK

2.1 <u>INTRODUCTION</u> : One of the main applications of the techniques developed in Chapter one is in expanding the integrated Bayes risk and getting an approximate Bayes estimate (vide Definition 2.1). For estimation with squared error loss the integrated Bayes risk R_ρ can be expanded (vide Theorem 2.1 (a)) in the form

$$R_\rho = a_1 . n^{-1} + a_2 . n^{-2} + o(n^{-2})$$

where $a_1 = \int (I(\theta))^{-1} \rho(\theta) d\theta$ and a_2 is as defined in Proposition 2.1. It may be noted that the problem of expanding $E_\theta [(E(\theta|\underline{x}_n)-\theta)^2]$ was considered by Gusev [1976] where e.g. results of type (see his Theorem 1)

$$E_\theta \left[(\sqrt{n}\, (E(\theta|\underline{x}_n) - \theta))^2 \right] = I^{-1}(\theta) + \sum_{i=1}^{k} h_i n^{-i} + o(n)^{-k}$$

were obtained. One can not get expantion for integrated Bayes risk using above expansion as the dependence on θ of the remainder term is not clear from his work. Using above mentioned techniques of ours we are (in a sense) able to trace this dependence (vide Remark 1.3). (In our case one of the quantities that needs attention is $|C_n|^{-1}$ of Theorem 1.1).

Main results of this chapter are given in Section 2. Our method of proof for getting above expansion for R_ρ consists of first getting a approximation, B_n, for the Bayes estimate which holds good on a large P_ρ probability set $\triangle_n$ (vide (2.5)). Contribution to R_ρ from $\triangle_n^c$ being small now it is enough to get a expansion for the quantity $E_\rho[(B_n - \theta)^2 . I_{\triangle_n}]$ which is done in Proposition 2.1. In Theorem 2.1 we also prove that the natural truncation of B_n is a approximate Bayes estimate. Here it may be noted that (vide (2.1a)) B_n depends on the whole sample viz. $\underline{x}_n$ where as B_n' depends on the sample only through the m.l.e. $\hat{\theta}_n$. In Theorem 2.1 we prove that the natural truncation of B_n' is also an approximate Bayes estimate ; one of the important application of this is in proving the second order efficiency of the m.l.e. (vide Section 2 of Chapter three). In Section 2 we also obtain expansion upto $o(n^{-2})$ for the posterior risk under square error loss (viz. $E[(E(\theta|\underline{x}_n) - \theta)^2|\underline{x}_n])$ which holds good on a large P_ρ probability set (vide Theorem 2.2). Section 3 gives results needed for the proofs of Section 2. In Section 4 we give a counterexample to show that Corollary 2 of Alvo [1977] is false ; Alvo ignored the dependence on $|C_n|^{-1}$ of a remainder term and got his Corollary 2. In this section it is also shown that even for a normal distribution with unknown mean, there is no expansion for the integrated Bayes risk of the form $R_\rho = a_1.n^{-1} + a_2.n^{-2} + o(n^{-2})$ where a_1 and a_2 are constants, when the prior is uniform on [0,1]. In Section 5 we assume conditions on prior which are similar to those of Burnasev [1979] and show how the above mentioned expansion for R_ρ can be obtained by minor modifications of arguments of Section 2 of Chapter one and of Section 3. This chapter is based on Sections 5,6 and 9 of Ghosh, Sinha and Joshi [1981].

2.2 MAIN RESULTS : In this chapter we assume all the notations of Chapter one and the m.l.e. $\hat{\theta}_n$ is as defined in Lemma 1.2. Let

$$\lambda_{1,2,n}(\underline{x}_n) = b^{-2}(6\, a_{3n}(\hat{\theta}_n)b^{-2} + \frac{\rho^{(1)}(\hat{\theta}_n)}{\rho(\hat{\theta}_n)})$$

Following two statistics are going to play an important roll in this chapter. Let

(2.1a) $$B_n = \hat{\theta}_n + \lambda_{1,2,n}(\underline{x}_n)n^{-1} \quad \text{and}$$

$$B_n' = \hat{\theta}_n + (\tfrac{1}{2}L_{001}(\hat{\theta}_n)\, I^{-1}(\hat{\theta}_n) + \frac{d}{d\theta}\log \rho(\hat{\theta}_n))I^{-1}(\hat{\theta}_n)$$

where $L_{001} = E_\theta [\frac{d^3}{d\theta^3} \log f(X, \theta)]$.

Let B_n^* and B_n'' be the natural truncations of B_n and B_n' respectively that is

$$\begin{aligned} B_n^* &= B_n \quad \text{if} \quad B_n \varepsilon (a_o, b_o) \\ &= a_o \quad \text{if} \quad B_n \leq a_o \\ &= b_o \quad \text{if} \quad B_n \geq b_o \end{aligned}$$

and similarly for B_n'' ; here (a_o, b_o) has the same meaning as in the introduction of Chapter one.

DEFINITION 2.1 : An estimate $T_n(\underline{x}_n)$ is called approximate Bayes estimate of order β w.r.t. loss function L and prior ρ if

$$|E_\rho [L(T_n, \theta)] - \operatorname*{Inf}_{T_n'} E_\rho [L(T_n', \theta)]| = o(n^{-\beta})$$

the above infinum being taken over all estimates depending only on $\underline{x}_n$.

THEOREM 2.1 : Let assumptions AI to AV of Section 3 of Chapter one hold for some $K \geq 2$, $r_1 > 4$ and $r_2 \geq 4$. Let ρ be of type D_S, $11 < S \leq \infty$. Then

(a) The integrated Bayes risk, R_ρ, w.r.t. squared error loss function has following expansion

$$R_\rho = a_1 n^{-1} + a_2 n^{-2} + o(n^{-2}),$$

a_1 and a_2 as defined in Proposition 2.1,

(b) B_n^* and B_n'' are approximate Bayes estimates of order 2 w.r.t. squared error loss function.

PROOF : (a) As B_n^* is natural truncation of B_n we have (with $\triangle_n$ of (2.5))

(2.1b) $$E_\rho [(B_n - \theta)^2 I_{\triangle_n}] \geq E_\rho [(B_n^* - \theta)^2 I_{\triangle_n}]$$
$$= E_\rho [(B_n^* - \theta)^2] + o(n^{-2})$$

The last equality is due the fact that B_n^* is bounded and $P_\rho(\triangle_n^c) = o(n^{-2})$ (vide (2.5)). Now

(2.1c) $$E_\rho [E(\theta|\underline{x}_n) - \theta]^2$$
$$\geq E_\rho [\{E(\theta|\underline{x}_n) - B_n + B_n - \theta\}^2 I_{\triangle_n}]$$

$$= E_\rho [(B_n - \theta)^2 I_{\triangle_n}] + E_\rho [\{E(\theta|\underline{x}_n) - B_n\}^2 I_{\triangle_n}]$$

$$+ 2 E_\rho [(E(\theta|\underline{x}_n) - \theta)(B_n - \theta) I_{\triangle_n}]$$

$$= I_1 + I_2 + I_3, \text{ say.}$$

Note that

$$I_1 = a_1 n^{-1} + a_2 n^{-2} + o(n^{-2}) \quad \text{(vide (2.6))}$$

$$I_2 = o(n^{-2}) \quad \text{(vide (2.5)) and}$$

$$I_3 \leq 2 \{E_\rho [(B_n - \theta)^2 I_{\triangle_n}]\}^{1/2} \{E_\rho [(E(\theta|\underline{x}_n) - B_n)^2 I_{\triangle_n}]\}^{1/2}$$

$$= o(n^{-2}) \quad \text{(vide (2.5) and (2.6)).}$$

Hence in view of (2.1b)

$$E_\rho [(B_n^* - \theta)^2] \leq a_1 n^{-1} + a_2 n^{-2} + o(n^{-2}) \leq E_\rho [\{E(\theta|\underline{x}_n) - \theta\}^2] + o(n^{-2})$$

Thus proof of part (a) is completed. Note that the last expression also shows that B_n^* is approximate Bayes estimate of order 2. To prove the same about B_n'' note that (vide 2.18)

$$E_\rho [(B_n' - \theta)^2 I_{\triangle_n}] = a_1 n^{-1} + a_2 n^{-2} + o(n^{-2})$$

and as B_n'' is natural truncation of B_n' we have

$$E_\rho [(B_n'' - \theta)^2 I_{\triangle_n}] \leq E_\rho [(B_n' - \theta)^2 I_{\triangle_n}]$$

Thus in view of (2.5) and boundedness of B_n'' we have

$$E_\rho [(B_n'' - \theta)^2] = E_\rho [(B_n'' - \theta)^2 I_{\triangle_n}] + o(n^{-2}) \leq a_1 n^{-1} + a_2 n^{-2} + o(n^{-2})$$

This in view of part (a) completes the proof of part (b) □

In the following theorem we get an approximation to the Posterior risk w.r.t. squared error loss function, $E\left[(E(\theta|\underline{x}_n)-\theta)^2|\underline{x}_n\right]$, which holds good on a large P_ρ probability set.

<u>THEOREM 2.2</u> : Let the assumptions AI to AV hold with some $K \geq 4$, $r_1 > 4$, $r_2 \geq 4$ and ρ be of type D_S $(20 < S \leq \infty)$ then for some $\varepsilon > 0$, $\delta > 0$ we have

$$P_\rho\left[\left|E\left[\left\{E(\theta|\underline{x}_n)-\theta\right\}^2|\underline{x}_n\right]-n^{-1}P_1(\underline{x}_n)-n^{-2}P_2(\underline{x}_n)\right|\right.$$

$$\left.\leq M.n^{-(2+\varepsilon)},\ |n^{-i}P_i(\underline{x}_n)|\leq M.n^{-\delta},\ i=1,2\right]$$

$$= 1 - o(n^{-2})$$

where $P_1(\underline{x}_n) = b^{-2}$,

$$P_2(\underline{x}_n) = b^{-2}\ \lambda_{2,4}(\underline{x}_n) - \lambda^2_{1,2,n}(\underline{x}_n),$$

$\lambda_{2,4}(\underline{x}_n)$ as in Lemma 2.2 and $\lambda_{1,2,n}(\underline{x}_n)$ as in Lemma 2.1.

<u>PROOF</u> : Combining Lemmas 2.1 and 2.2 the proofs follows easily □

<u>REMARK 2.1</u> : If assumptions AI to AV hold with some $K \geq 2$, $r_1 > 3$, $r_2 > 2$ and ρ is of type D_S $(5 \leq S \leq \infty)$ then it is easy to see that $\hat{\theta}_n$ is approximate Bayes estimate of order 1 w.r.t. the squared error loss function and using similar arguments as in the case of Theorem 2.1 and Proposition 2.1, we have

$$R_\rho = a_1 n^{-1} + o(n^{-1}).$$

2.3 LEMMAS AND PROPOSITIONS : In this section we prove some results which are needed for the proofs of the results of Section 2.2. The following lemma gives us approximation to the Bayes estimate $E(\theta|\underline{x}_n)$ whereas Lemma 2.2 gives us an approximation needed for the proof of Theorem 2.2. Method of the proof of these lemmas is similar to that of Theorem 1.2. First we get approximations to the numerator and denominator (of the quantity for which the approximation is saught) with precise remainder terms which in turn gives the approximation to the ratio with suitable remainder term. This remainder now is bounded by using Proposition 1.1.

Let $A_n = \{-M < a_{3,n}(\theta_n) < M,\ 0 < \delta < b^{-1} < M\}$.

Then $\exists$ $\delta > 0$ and M such that under appropriate assumptions (vide Lemmas 1.2, 1.4 and 1.6)

(2.1) $$P_\rho(A_n) = 1 - o(n^{-2}).$$

LEMMA 2.1 : Let assumptions AI to AV hold with some $K \geq 2$, $r_1 > 4$ and $r_2 \geq 4$. Let ρ be of type D_{5+r_o} for some r_o, $6 < r_o \leq \infty$. Then $\exists$ $\varepsilon > 0$ and $\delta > 0$ such that

(2.2) $$P_\rho \left[|E(\theta|\underline{x}_n) - \hat{\theta}_n - \lambda_{1,2,n}(\underline{x}_n)n^{-1}| \leq M\, n^{-(3/2+\varepsilon)},\ |n^{-1}\rho^{(1)}(\hat{\theta}_n)/\rho(\hat{\theta}_n)| \leq M\, n^{-\delta} \right] = 1 - o(n^{-2})$$

where $\lambda_{1,2,n}(\underline{x}_n) = (6\, a_{3,n}(\hat{\theta}_n)\, b^{-4} + \rho^{(1)}(\hat{\theta}_n)\, b^{-2}/\rho(\hat{\theta}_n))$.

Expression for $\lambda_{1,2,n}$ which one gets from Johnson [1970] is not correct. It should be as given above (here it may be noted that our $\lambda_{1,2,n}$ is by definition b^{-1} times Johnson's $\lambda_{1,2}$ and expression for his $\lambda_{1,2}$ as given by him is incorrect). In this connection note that Johnson expands to less terms than are allowed by the assumptions for his Theorem 3.1 ; in the following we are in effect expanding the Bayes estimate upto $n^{-3/2}$ and using the fact that the coefficient of this term is zero.

PROOF : Following the proof of (1.17) (vide Theorem 3.1 of Johnson [1970]) with $\rho(\hat{\theta}_n + \phi\, b^{-1})$ replaced by $\phi\,\rho\,(\hat{\theta}_n + \phi\, b^{-1})$ we get

(2.3a) $$P_\rho\Big[\Big| \int_{-\infty}^{\infty} \phi\rho(\hat{\theta}_n + \phi b^{-1}) \prod_{i=1}^{n} [f(x_i, \hat{\theta}_n + \phi b^{-1})/f(x_i, \hat{\theta}_n)]\, d\phi - (b^{-1}\rho^{(1)}(\hat{\theta}_n) + b^{-3}a_{3,n}(\hat{\theta}_n)\rho(\hat{\theta}_n))n^{-3/2}\sqrt{2\pi}\Big| \le M\, n^{-1/2}R_{1n}\Big] = 1 - o(n^{-2})$$

and from (1.17) we have

(2.3b) $$P_\rho\Big[\Big| \int_{-\infty}^{\infty} \rho(\hat{\theta}_n + \phi\, b^{-1}) \prod_{i=1}^{n} [f(x_i, \hat{\theta}_n + \phi\, b^{-1}) / f(x_i, \hat{\theta}_n)]\, d\phi - \sum_{j=0}^{2} \beta_j(\underline{x}_n)\, n^{-(j+1)/2} \Big| \le M\, R_{1n}\Big] = 1 - o(n^{-2})$$

where $R_{1n} = n^{-(4+r_0)/2} + n^{-2}\sum_{r=0}^{2} |\rho^{(r)}(\hat{\theta}_n)| + \sum_{r=1}^{r_0} n^{-(3+r)/2} |\rho^{(2+r)}(\hat{\theta}_n)|$

(In (2.3a) we have an extra $n^{-1/2}$ in the remainder because we are able to expand upto n^{-2} ; coefficient of this term is zero.)

Now using (2.3b), (2.3a), (1.2b) and the fact

(2.3c) $[(b^{-1}\rho^{(1)}(\hat{\theta}_n) + b^{-3}a_{3,n}(\hat{\theta}_n)\rho(\hat{\theta}_n))n^{-3/2}\sqrt{2\pi}](\sum_{j=o}^{2}\beta_j(\underline{x}_n)\, n^{-(j+1)/2})^{-1}$

$= \lambda_{1,2,n}(\underline{x}_n)\, b\, n^{-1} - \lambda_{1,2,n}(\underline{x}_n)\, b\, \beta_2(\underline{x}_n)\, n^{-2}\, C_n^{-1}$

where $C_n = \rho(\hat{\theta}_n) + n^{-1}\beta_2(\underline{x}_n)(\sqrt{2\pi})^{-1}$ (vide Remark 1.4) we have

$$P_\rho[|E(\theta - \hat{\theta}_n)|\underline{x}_n)| \leq n^{-1/2} M\delta_n] = 1 - o(n^{-2})$$

where $\delta_n = n^{-1/2}|\lambda_{1,2,n}(\underline{x}_n)| + |C_n|^{-1} n\, R_{1n}$

$+ |\lambda_{1,2,n}(\underline{x}_n)||\beta_2(\underline{x}_n)||C_n|^{-1} b\, n^{-3/2}$.

Now again using (2.3b), (2.3a) and (1.2b) but this time (1.2b) with $C = n^{-1/2} M\delta_n b(\hat{\theta}_n)$ we get

(2.3) $P_\rho[|E(\theta|\underline{x}_n) - \hat{\theta}_n - \lambda_{1,2,n}(\underline{x}_n)\, n^{-1}| \leq M|C_n|^{-1} R_n]$

$= 1 - o(n^{-2})$

where $R_n = n^{-2}|\lambda_{1,2,n}(\underline{x}_n)||\beta_2(\underline{x}_n)| + R_{1n}[1 + n^{-1/2}|\lambda_{1,2,n}(\underline{x}_n)|$

$+ n^{-3/2}|\lambda_{1,2,n}(\underline{x}_n)||\beta_2(\underline{x}_n)||C_n|^{-1} + n\, R_{1n}|C_n|^{-1}]$.

Now we proceed to bound $|C_n|^{-1} R_n$. For simplicity let $a_o = 0$ and $b_o = 1$. Consider the case $r_o < \infty$. First note that by using the form of $\beta_2(\underline{x}_n)$ and the fact

(2.4a) $|\rho^{(i)}(\theta)(\rho(\theta))^{-1}| \leq M(\theta^{-i}, (1-\theta)^{-i})$

we have for some $\varepsilon_1 > 0$

$$(\hat{\theta}_n \geq M.n^{-(1/6-\varepsilon_1)}, \hat{\theta}_n \leq 1 - M.n^{-(1/6-\varepsilon_1)}) \cap A_n$$

$$\Rightarrow (|C_n|^{-1} = (\rho(\hat{\theta}_n))^{-1}(1 + d_n)^{-1} \text{ where } d_n = o(1)) .$$

Now consider a term of $|C_n|^{-1} R_n$, say,

$\eta_n \equiv n^{-3/2}|\lambda_{1,2,n}(\underline{x}_n)||\beta_2(\underline{x}_n)||C_n|^{-1} R_{1n}|C_n|^{-1}$; by using the forms of $\lambda_{1,2,n}(\underline{x}_n)$, $\beta_2(\underline{x}_n)$ and R_{1n} and (2.4a) and (1.10) we have for some $\varepsilon_1 > 0$

$$(\hat{\theta}_n \geq M.n^{-(1/6-\varepsilon_1)}, \hat{\theta}_n \leq 1 - M.n^{-(1/6-\varepsilon_1)}) \cap A_n$$

$$\Rightarrow \Big[\eta_n \leq M.n^{-3/2}\, n^{-2}(\hat{\theta}_n^{-6} + (1 - \hat{\theta}_n)^{-6} + n^{-(2+r_0)}(\rho(\hat{\theta}_n))^{-1}) ,$$

$$(\rho(\hat{\theta}_n))^{-1} \leq M.n^{(1/6-\varepsilon_1)(r_0+5)} \quad \text{and}$$

$$\hat{\theta}_n^{-6} + (1 - \hat{\theta}_n)^{-6} < M.n^{1-6\varepsilon_1} \Big]$$

Handling other terms of $|C_n|^{-1} R_n$ in similar way we have for every $\varepsilon_1 > 0$ $\exists$ $\varepsilon > 0$ and $\delta > 0$ such that

$$(2.4) \qquad (\hat{\theta}_n \geq M.n^{-(1/6-\varepsilon_1)}, \hat{\theta}_n \leq 1 - M.n^{-(1/6-\varepsilon_1)}) \cap A_n$$

$$\Rightarrow \Big[|C_n|^{-1} R_n \leq M.n^{-3/2-\varepsilon}), |n^{-1}\lambda_{1,2,n}(\underline{x}_n)| \leq M.n^{-\delta}\Big]$$

For the case $r_0 < \infty$, proof of (2.2) is completed by using (2.3), (2.4), (1.8) and (2.1). The case $r_0 = \infty$ follows easily □

LEMMA 2.2 : Let assumptions AI to AV hold with some $K \geq 4$, $r_1 > 4$, $r_2 \geq 4$ and ρ be of type D_S $(20 < S \leq \infty)$ then $\exists$ $\varepsilon > 0$, $\delta > 0$ such that

$$P_\rho \Big[|E[b^2(\theta - \hat{\theta}_n)^2 | \underline{x}_n] - n^{-1}\lambda_{2,2}(\underline{x}_n) - n^2\lambda_{2,4}(\underline{x}_n)| \leq M.n^{-(2+\varepsilon)}, |n^{-i}\lambda_{2,2i}(\underline{x}_n)| \leq M.n^{-\delta} \quad i = 1,2 \Big]$$

$= 1 - o(n^{-2})$ where

$\lambda_{2,2}(\underline{x}_n) = 1$ and

$$\lambda_{2,4}(\underline{x}_n) = 12b^{-4}a_{4n}(\hat{\theta}_n) + 45b^{-6}a_{3n}^2(\hat{\theta}_n) + 12b^{-4}a_{3n}(\hat{\theta}_n)\rho^{(1)}(\hat{\theta}_n)/\rho(\hat{\theta}_n)$$

Proof of this lemma is similar to that of Lemma 2.1 and hence is omitted.

Note that $\triangle_n$ (see (2.5) below) has lagge P_ρ probability and that on $\triangle_n$ B_n is a good approximation for the Bayes estimate $E(\theta|\underline{x}_n)$; so roughly speaking, contribution to the Bayes-risk comes (mainly) from $E_\rho[(B_n - \theta)^2 I_{\triangle_n}]$. Proposition 5.1 below gives expansion for this quantity. To prove this proposition we first throw away the part of the parameter space which is near to the two end points a_o and b_o (and on which the contribution is small)(vide (2.11)). For θ in the remaining part we expand $B_n(\hat{\theta}_n)$ around θ upto sufficiently many terms and take term by term expectation (under P_θ). The remainder of this expansion has small contribution (vide (2.12)) ; this is due to the fact that θ (and hence $\hat{\theta}_n$) is away from a_o and b_o and hence the troublesome (i.e. exploding) part of the expansion (viz. darivatives of log ρ) can be bounded by suitable quantities (vide (2.14)).

Let $\Delta_n = \left[|E(\theta|\underline{x}_n) - B_n| \leq M.n^{-(3/2+\varepsilon)},\ |n^{-1}\lambda_{1,2,n}(\underline{x}_n)| \leq M.n^{-\delta} \right] \cap A_n$

(2.5) where $B_n = \hat{\theta}_n + n^{-1}\lambda_{1,2,n}(\underline{x}_n)$ and $\varepsilon > 0$ and $\delta > 0$ as in Lemma 2.1. Note that B_n is bounded on Δ_n and

$$P_\rho(\Delta_n) = 1 - o(n^{-2}).$$

<u>PROPOSITION 5.1</u> : Let assumptions AI to AV hold with some $K \geq 2$, $r_1 > 4$ and $r_2 \geq 4$. Let ρ be of type D_S, $11 < S \leq \infty$. Then

(2.6) $$E_\rho [(B_n - \theta)^2 I_{\Delta_n}] = a_1 n^{-1} + a_2 n^{-2} + o(n^{-2})$$

where $a_1 = \int I^{-1}(\theta)\ \rho(\theta)d\theta$,

$$a_2 = \int [H_{2,2}(\theta) - \frac{1}{4} I^{-4}(\theta)L^2_{001}(\theta) - 4\, I^{-3}(\theta)\frac{d}{d\theta} \log \rho(\theta)\frac{d}{d\theta} I(\theta)$$
$$- I^{-2}(\theta)(\frac{d}{d\theta} \log \rho(\theta))^2 + I^{-3}(\theta)\frac{d}{d\theta} L_{001}(\theta)$$
$$- 3\, I^{-4}(\theta)\ L_{001}(\theta)\frac{d}{d\theta} I(\theta)]\ \rho(\theta)d\theta,$$

$L_{ijk\ell}(\theta) = E_\theta [h_1^i(\theta)\ h_2^j(\theta)\ h_3^k(\theta)\ h_4^\ell(\theta)]$ where

$$h_i(\theta) = \frac{d^i}{d\theta^i} \log f(x,\theta),$$

$L_{ijk}(\theta) = L_{ijko}(\theta)$, $L_{ij}(\theta) = L_{ijo}(\theta)$, $L_i(\theta) = L_{io}(\theta)$ and

$$H_{2,2}(\theta) = I^{-4}(\theta)(L_{301}(\theta) + \frac{15}{4} L^2_{001}(\theta) + 12L_{11}(\theta)$$
$$+ 12L_{001}(\theta) + 6L^2_{11}(\theta))$$
$$+ I^{-3}(\theta)(2L_{21}(\theta) + L_{0001}(\theta) + 3L_{101}(\theta) + L_{02}(\theta))$$

(vide (3.3) of Gusev [1976]).

We also use following notations

$$\lambda_{i,n}(\theta) = (i!)\, a_{i,n}(\theta) \quad \text{and} \quad I_n(\theta) = -2a_{2,n}(\theta).$$

PROOF : Note that under the assumptions of the Proposition, by Gusev [1976] we have uniformly in $\theta \in [a_o , b_o]$

(2.7) $$E_\theta(\hat{\theta}_n - \theta) = \left[(L_{001}(\theta) + 2L_{11}(\theta))/2\, I^2(\theta)\right] n^{-1} + o(n^{-2})$$

$$E_\theta(\hat{\theta}_n - \theta)^2 = I^{-1}(\theta)\, n^{-1} + H_{2,2}(\theta)\, n^{-2} + o(n^{-2})$$

$$E_\theta(\hat{\theta}_n - \theta)^i = o(n^{-2}) \quad i = 3,4 \quad \text{and}$$

(2.7a) $$L_{101}(\theta) + L_{0001}(\theta) = \frac{d}{d\theta} L_{001}(\theta), -L_{11}(\theta) - L_{001}(\theta) = \frac{d}{d\theta} I(\theta).$$

Let $I_{n,\theta} = [\, |\hat{\theta}_n - \theta| \le c_n , \sum_{i=1}^{n} \frac{d}{d\theta} \log f(x_i , \hat{\theta}_n) = 0, |\lambda_{2,n}(\hat{\theta}_n)$

$+ I(\hat{\theta}_n)| \le c_n,$

$|\lambda_{3,n}(\hat{\theta}_n) - L_{001}(\hat{\theta}_n)| \le c_n , 0 < \delta < I_n(\theta^*_n) < M$ and $M < \lambda_{i,n}(\theta^*_n) < M$

for $i = 3,4,5$ and θ^*_n between $\hat{\theta}_n$ and $\theta\,]$ with c_n of Lemma 1.3. Note that by applying Lemma 1 of Pfanzagl [1973a] to $\lambda_{2,n}(\theta)$ and

$\lambda_{3,n}(\theta)$ and our Lemma 1.1 to $\lambda_{i,n}$ $(i = 3,4,5)$ in view of our Lemma 1.3 we have uniformly in $\theta \varepsilon [a_o , b_o]$

(2.8) $$P_\theta(I_{n,\theta}) = 1 - o(n^{-2})$$

Now,

(2.9) $$E_\theta [(B_n - \theta)^2 I_{\Delta_n}] = E_\theta [(B_n - \theta)^2 (I_{\Delta_n \cap I_{n,\theta}} + I_{\Delta_n \cap I^c_{n,\theta}})]$$

For simplicity, let $a_o = 0$ and $b_o = 1$. First we consider the case $16 < S < \infty$. Let $a_n = 2c_n$ and $b_n = 1 - 2c_n$. Note that

(2.10) $$\int_0^{a_n} E_\theta [(B_n - \theta)^2 I_{\Delta_n}] \rho(\theta) d\theta = o(n^{-2})$$

$$= \int_{b_n}^{1} E_\theta [(B_n - \theta)^2 I_{\Delta_n}] \rho(\theta) d\theta$$

and $$\int_0^1 E_\theta [(B_n - \theta)^2 I_{\Delta_n \cap I^c_{n,\theta}}] \rho(\theta) d\theta = o(n^{-2}).$$

Using (2.9) and (2.10) we get

(2.11) $$E_\rho [(B_n - \theta)^2 I_{\Delta_n}] = \int_{a_n}^{b_n} E_\theta [(B_n - \theta)^2 I_{\Delta_n \cap I_{n,\theta}}] \cdot \rho(\theta) d\theta + o(n^{-2})$$

Expanding $\lambda_{1,2,n}(\hat{\theta}_n)$ around θ and denoting the first and second derivatives of $\lambda_{1,2,n}$ by $\lambda^{(1)}_{1,2,n}$ and $\lambda^{(2)}_{1,2,n}$ respectively we have

$$(B_n - \theta) = (\hat{\theta}_n - \theta) + \lambda_{1,2,n}(\theta) n^{-1} + (\hat{\theta}_n - \theta) n^{-1} \lambda^{(1)}_{1,2,n}(\theta)$$
$$+ n^{-1} (\hat{\theta}_n - \theta)^2 / 2 \, \lambda^{(2)}_{1,2,n}(\xi)$$

$$= (\hat{\theta}_n - \theta) + n^{-1} I_n^{-1}(\theta) [\lambda_{3,n}(\theta) I_n^{-1}(\theta)/2 + \frac{d}{d\theta} \log \rho(\theta)]$$

$$+ (\hat{\theta}_n - \theta) n^{-1} [I_n^{-3}(\theta) \lambda_{3,n}^2(\theta) + I_n^{-2}(\theta) \lambda_{4,n}(\theta)/2$$

$$+ I_n^{-1}(\theta) \frac{d^2}{d\theta^2} \log \rho(\theta) + I_n^{-2}(\theta) \lambda_{3,n}(\theta) \frac{d}{d\theta} \log \rho(\theta)]$$

$$+ n^{-1} (\hat{\theta}_n - \theta)^2 \lambda_{1,2,n}^{(2)}(\xi)/2 \quad \text{where } \xi \text{ is between } \hat{\theta}_n \text{ and } \theta .$$

Now we use

$$(I_n(\theta))^{-r} = (I(\theta))^{-r} - r(I_n(\theta) - I(\theta))(I(\theta))^{-(r+1)}$$

$$+ r(r+1)(I_n(\theta) - I(\theta))^2 (I_n^*)^{-(r+2)}/2$$

where I_n^* is between $I_n(\theta)$ and $I(\theta)$ and

$$(\hat{\theta}_n - \theta) = I^{-1}(\theta) \lambda_{1,n}(\theta) - (\hat{\theta}_n - \theta) I^{-1}(\theta)(I_n(\theta) - I(\theta))$$

$$+ (\hat{\theta}_n - \theta)^2 \lambda_{3,n}(\theta_n^*) I^{-1}(\theta)/2 \quad \text{where } \theta_n^* \text{ is between}$$

$\hat{\theta}_n$ and θ and write

$$(B_n - \theta)^2 = E_{1,n} + E_{2,n} \quad \text{where}$$

$$E_{1,n} = (\hat{\theta}_n - \theta)^2 + n^{-2} I^{-2}(\theta) \left[\frac{d}{d\theta} \log \rho(\theta) + \frac{I^{-1}(\theta)}{2} L_{001}(\theta) \right]^2$$

$$+ n^{-1} I^{-2}(\theta) \Big[I^{-1}(\theta) \lambda_{1,n}(\theta) [\lambda_{3,n}(\theta) - L_{001}(\theta)]$$

$$+ (\hat{\theta}_n - \theta)^2 L_{0001}(\theta) \Big] + 2 n^{-1} I^{-3}(\theta) L_{001}(\theta) \Big[(\hat{\theta}_n - \theta)^2 L_{001}(\theta)$$

$$- I^{-1}(\theta) \lambda_{1,n}(\theta) [I_n(\theta) - I(\theta)] \Big]$$

$$+ 2 n^{-1} I^{-1}(\theta) (\hat{\theta}_n - \theta)^2 \frac{d^2}{d\theta^2} \log \rho(\theta)$$

$$+ 2 n^{-1} I^{-2}(\theta)\frac{d}{d\theta} \log \rho(\theta) \Big[(\hat{\theta}_n - \theta)^2 L_{001}(\theta) - I^{-1}(\theta) \lambda_{1,n}(\theta) [I_n(\theta) - I(\theta)]\Big]$$

$$+ 2 n^{-1} I^{-1}(\theta)(\hat{\theta}_n - \theta) [I^{-1}(\theta) L_{001}(\theta)/2 + \frac{d}{d\theta} \log \rho(\theta)].$$

We now prove that

$$(2.12) \qquad \int_{a_n}^{b_n} E_\theta [E_{2,n} I_{\triangle_n \cap I_{n,\theta}}] \rho(\theta) d\theta = o(n^{-2})$$

Choose a term of $E_{2,n}$, say, $2 n^{-2}\Big[\rho^{(1)}(\theta)(I(\theta))^{-1/2}/\rho(\theta)\Big]$. $(\hat{\theta}_n - \theta)^2 \lambda^{(2)}_{1,2,n}(\xi)/2$. We prove that

$$(2.13) \qquad n^{-2} \int_{a_n}^{b_n} \frac{\rho^{(1)}(\theta)}{\rho(\theta)} I^{-1}(\theta) E_\theta \Big[(\hat{\theta}_n - \theta)^2 \lambda^{(2)}_{1,2,n}(\xi) I_{\triangle_n \cap I_{n,\theta}}\Big] \rho(\theta) d\theta$$

$$= o(n^{-2}) .$$

Note that for sufficiently small $\varepsilon > 0$

$$(2.14) \qquad n^{-2} \int_{a_n}^{\varepsilon} E_\theta \Big[(\hat{\theta}_n - \theta)^2 \frac{\rho^{(1)}(\theta)}{\rho(\theta)} \Big(\frac{\rho^{(1)}(\xi)}{\rho(\xi)} \Big)^3 I_{\triangle_n \cap I_{n,\theta}}\Big] \rho(\theta) d\theta$$

$$\leq M.n^{-3} \log n \int_{2c_n}^{\varepsilon} \theta^{-1} (\theta - c_n)^{-3} \theta^S d\theta$$

$$= M.n^{-3} \log n \int_{c_n}^{\varepsilon - c_n} (\theta + c_n)^{S-1} \theta^{-3} d\theta$$

$$= o(n^{-2}) \text{ as } S > 11 .$$

Now it is not difficult to see that (2.14) and similar observations give us (2.13). Similarly handling other terms of $E_{2,n}$ one gets (2.12). Using (2.11) and (2.12) we get

(2.15) $$E_\rho[(B_n-\theta)^2 I_{\Delta_n}] = \int_{a_n}^{b_n} E_\theta[E_{1,n} I_{\Delta_n \cap I_{n,\theta}}]\rho(\theta)d\theta + o(n^{-2}).$$

We have

(2.16) $$\int_{a_n}^{b_n} E_\theta[E_{1,n} I_{(\Delta_n \cap I_{n,\theta})^c}]\rho(\theta)d\theta$$

$$\leq \int_{a_n}^{b_n} [E_\theta(E_{1,n}^2)\rho(\theta)d\theta]^{1/2}[P_\theta(\Delta_n^c \cup I_{n,\theta}^c)\rho(\theta)]^{1/2} d\theta$$

$$\leq \left[\int_{a_n}^{b_n} E_\theta(E_{1,n}^2)\rho(\theta)d\theta\right]^{1/2}\left[\int_{a_n}^{b_n} P_\theta(\Delta_n^c \cup I_{n,\theta}^c)\rho(\theta)d\theta\right]^{1/2}$$

$= o(n^{-2})$ using (2.5), (2.8), (2.7) and (iii) of Definition 1.1.

Moreover, integrating term by term and utilizing the polynomial decay of $\rho^{(1)}(\theta)$, $\rho^{(2)}(\theta)$ and $(\rho^{(1)}(\theta))^2/\rho(\theta)$ near the two end points, in view of (2.7) we have

(2.17) $$\int_0^{a_n} E_\theta(E_{1,n})\rho(\theta)d\theta = o(n^{-2}) = \int_{b_n}^{1} E_\theta(E_{1,n})\rho(\theta)d\theta .$$

Combining (2.15), (2.16) and (2.17) we get

$$E_\rho\left[(\hat{\theta}_n-\theta)^2 I_{\triangle_n}\right] = \int_0^1 E_\theta(E_{1,n})\cdot\rho(\theta)d\theta + o(n^{-2}) .$$

This in view of (2.17) proves the Proposition 2.1 for $11 < S < \infty$.

When ρ is of type D_∞, let $a_n = \rho_1^{-1}(n^{-3})$ and $b_n = \rho_2^{-1}(n^{-3})$ (see Proposition 1.2 for the definitions of ρ_1^{-1} and ρ_2^{-1}). Now to get (2.6), similar arguments as above can be given, only difference being demonstration of conclusion of (2.17). Note that for sufficiently small $\varepsilon > 0$ and for sufficiently large n

$$n^{-2}\int_{a_n}^{\varepsilon} E_\theta\left[(\hat{\theta}_n-\theta)^2\frac{\rho^{(1)}(\theta)}{\rho(\theta)}\left(\frac{\rho^{(1)}(\xi)}{\rho(\xi)}\right)^3 I_{\triangle_n}\cap I_{n,\theta}\right]\rho(\theta)d\theta$$

$$\leq M.n^{-3}\int_{a_n}^{\varepsilon}\rho^{-v_1}(\theta)\,\rho^{-v_2}(\theta-c_n)\rho(\theta)d\theta \quad \text{where} \quad v_1 > 0 \quad \text{and} \quad v_2 > 0$$

can be made arbitrarily small by choosing $\varepsilon > 0$.

Now

$$\rho^{-v_2}(\theta-c_n) = \left[\rho(\theta) + \sum_{i=1}^{6}\frac{(-c_n)^i}{i!}\rho^{(i)}(\theta) + \frac{(-c_n)^7}{7!}\rho^{(7)}(\theta_n)\right]^{-v_2}$$

where θ_n is between θ and $\theta - c_n$

$$= \rho^{-v_2}(\theta)\left[1 + \sum_{i=1}^{6}\frac{(-c_n)^i\rho^{(i)}(\theta)}{i!\,\rho(\theta)} + \frac{(-c_n)^7\rho^{(7)}(\theta_n)}{7!\;\rho(\theta)}\right]^{-v_2}$$

$$\leq M\,\rho^{-v_2}(\theta) \quad \text{for} \quad \theta = \theta_n .$$

Conclusion of (2.17) is now proved in view of monotone nature of ρ near the end points along with other observations. This completes the proof of the proposition □

Let $B_n' = \hat{\theta}_n + I^{-1}(\hat{\theta}_n)(\frac{1}{2}I^{-1}(\hat{\theta}_n)\, L_{oo1}(\hat{\theta}_n) + \frac{d}{d\theta}\log \rho(\hat{\theta}_n))\, n^{-1}$.

PROPOSITION 2.2 : Under the assumptions of Proposition 2.1

(2.18) $$E_\rho [(B_n' - \theta)^2\, I_{\triangle_n}] = a_1 n^{-1} + a_2 n^{-2} + o(n^{-2}).$$

PROOF : Let $C(\theta) = I^{-1}(\theta)(\frac{1}{2}I^{-1}(\theta)\, L_{oo1}(\theta) + \frac{d}{d\theta}\log \rho(\theta))$.

$$(B_n' - \theta) = (\hat{\theta}_n - \theta) + n^{-1} C(\theta) + n^{-1} (\hat{\theta}_n - \theta)\, C^{(1)}(\theta)$$

$$+ n^{-1}(\hat{\theta}_n - \theta)^2/2\, C^{(2)}(\theta_n^*) \quad \text{where} \quad C^{(i)}(\theta) = \frac{d^i}{d\theta^i} C(\theta)$$

and θ_n^* is between $\hat{\theta}_n$ and θ.

As in the case of B_n, we write

$(B_n' - \theta)^2 = E_{1,n}' + E_{2,n}'$ where

$$E_{1,n}' = (\hat{\theta}_n - \theta)^2 + n^{-2} I^{-2}(\theta)\, [I^{-1}(\theta)\, L_{oo1}(\theta)/2 + \frac{d}{d\theta}\log \rho(\theta)]^2$$

$$+ n^{-1} I^{-2}(\theta)\, [I^{-1}(\theta)\, \ell_{1,n}(\theta)(\ell_{3,n}(\theta) - L_{oo1}(\theta)) + (\hat{\theta}_n - \theta)^2 L_{ooo1}(\theta)]$$

$$- 2 n^{-1} I^{-3}(\theta)\, L_{oo1}(\theta)(\hat{\theta}_n - \theta)^2 \frac{d}{d\theta} I(\theta)$$

$$+ 2 n^{-1} I^{-1}(\theta)(\hat{\theta}_n - \theta)\, [\frac{I^{-1}(\theta)}{2} L_{oo1}(\theta) + \frac{d}{d\theta}\log \rho(\theta)]$$

$$+ 2 n^{-1} I^{-1}(\theta)(\hat{\theta}_n - \theta)^2 \frac{d^2}{d\theta^2}\log \rho(\theta)$$

$$- 2 n^{-1} I^{-2}(\theta)(\hat{\theta}_n - \theta)^2 \frac{d}{d\theta}\log \rho(\theta)\, \frac{d}{d\theta} I(\theta).$$

Now proceeding as in the case of Proposition 2.1 the proof is completed □

2.4 COUNTEREXAMPLES : The Example 1.1 of Chapter one also suggests that Corollary 2 of Alvo [1977] is not true. Consider the same example. To get a counterexample to the corollary it is enough to show that for some $0 < \varepsilon_1 < 1$

(2.19) $P_\rho [|E(\theta|\underline{x}_n) - \hat{\theta}_n| < M.n^{-1+\varepsilon_1}]$ cannot be made of order n^{-1}, implying that even the first moment of Alvo's $N_{\underline{x}}$ is infinity.

Now as in the case of our Lemma 2.1 we have

$$(2.20) \quad P_\rho [|E(\theta|\underline{x}_n) - \hat{\theta}_n - \lambda_{1,2,n}(\underline{x}_n)\, n^{-1}| \leq M.n^{-(1+\varepsilon)}]$$

$$\geq 1 - O(n^{-(1-\varepsilon)6/5}) \quad \text{for} \quad 0 < \varepsilon < 1 .$$

Here $\lambda_{1,2,n}(\underline{x}_n) = \rho^{(1)}(\hat{\theta}_n)/\rho(\hat{\theta}_n)$.

Also by modifying the proof of Proposition 2.1 it is not difficult to get that for any $0 < \varepsilon_1 < 1/2$

$$P_\rho [|\lambda_{1,2,n}(\underline{x}_n)\, n^{-1}| > M.n^{-1+\varepsilon_1}] \geq O(n^{-6.\varepsilon_1}).$$

This along with (2.20) gives

$$P_\rho [|E(\theta|\underline{x}_n) - \hat{\theta}_n| > M.n^{-1+\varepsilon_1}] \geq O(n^{-6\,\varepsilon_1}) - O(n^{-(1-\varepsilon)6/5})$$

Now choosing ε_1 sufficiently small we get (2.19).

Now we give an example which demonstrates (A) the Bayes risk need not have expansion of the form $a_1\, n^{-1} + a_2\, n^{-2} + o(n^{-2})$ where a_1 and a_2 are constants and (B) a Bayes estimate, in general, cannot be improved by an estimate of the form $\hat{\theta}_n + d(\hat{\theta}_n)/n$ where

$d(\theta)$ is smooth (i.e. twice continuously differentiable) function of θ, in the sense

(2.21) $E_\theta(E(\theta|\underline{x}_n) - \theta)^2 \geq E_\theta(\hat{\theta}_n + d(\hat{\theta}_n)/n - \theta)^2 + o(n^{-2})$

uniformly in θ belonging to an interval say $(0,1)$.

EXAMPLE 2 : Let $X_1, X_2, \ldots, X_n$ be iid $N(\theta,1)$ and the prior $\rho(\theta)$ be uniform over $(0,1)$. It is easy to see that for any $K > 0$ if we take $r_K = [(K/2+1)^2 - 1]/2$ and $d_K(\theta) = -\theta^{r_K}(K+2)/2$ then $\exists\ n_o(K)$ such that Bayes risk of $\hat{\theta}_n + d_K(\hat{\theta}_n)n^{-1}$

(2.22) $R(\hat{\theta}_n + d_K(\hat{\theta}_n)n^{-1}) = n^{-1} - K\,n^{-2} + R_n$

where $|R_n| < n^{-2}\ \forall\ n \geq n_o(K)$.

We shall prove that (for another proof of (2.23) see Remark 2.2)

(2.23) $R(E(\theta|\underline{x}_n)) = n^{-1} + o(n^{-1})$

Since K in (2.22) is arbitrary, (2.22) and (2.23) imply that $R(E(\theta|\underline{x}_n)) = n^{-1} - a_n$ where $a_n > 0$ and $a_n n^2 \longrightarrow \infty$. This demonstrates (A). Also (B) is demonstrated for if there exists a twice continuously differentiable function $d(\theta)$ such that (2.21) holds then $n^{-1} + d_1 n^{-2} + o(n^{-2}) \leq n^{-1} - a_n$ where

$$R(\hat{\theta}_n + d(\hat{\theta}_n)n^{-1}) = n^{-1} + d_1 n^{-2} + o(n^{-2})$$ and d_1 is a constant.

To prove (2.23), note that for $\theta \in (n^{-1/2}\log n,\ 1 - n^{-1/2}\log n) \equiv S_n$

$$P_\theta(\bar{x}_n \in (0,1)) = 1 - o(n^{-3})$$

i.e. $P_\theta(\rho(\bar{x}_n) = 1) = 1 - o(n^{-3})$ and hence we can get

$$P_\theta\left[|E(\theta|\underline{x}_n) - \bar{x}_n| \leq n^{-3}\right] = 1 - o(n^{-3}).$$

Thus we have for $\theta \varepsilon S_n$

$$E_\theta(E(\theta|\underline{x}_n) - \theta)^2 = E_\theta(\bar{x}_n - \theta)^2 + o(n^{-2}).$$

Now it is easy to see that (2.23) follows if we prove

$$(2.24) \qquad E_\theta\left[h^2(\underline{x}_n)\, I_{\left[\sqrt{n}|\bar{x}_n - \theta| < c(\log n)^{1/2}\right]}\right] < M \quad \forall\, \theta \varepsilon S_n^o$$

where $E(\theta|\underline{x}_n) = \bar{x}_n + h(\bar{x}_n)/\sqrt{n}$ and c is choosen suitably.

First assume $\theta \varepsilon (0, n^{-1/2}\log n)$ and note that

$$h(\underline{x}_n) = \left[\int_{-\sqrt{n}\,\bar{x}_n}^{-\sqrt{n}\,\bar{x}_n + \sqrt{n}} z\, e^{-z^2/2} dz\right]\left[\int_{-\sqrt{n}\,\bar{x}_n}^{-\sqrt{n}\,\bar{x}_n + \sqrt{n}} e^{-z^2/2} dz\right]^{-1}$$

$$< e^{-n\bar{x}_n^2/2}\left[\Phi(-\sqrt{n}\,\bar{x}_n + \sqrt{n}) - \Phi(-\sqrt{n}\,\bar{x}_n)\right]^{-1}$$

$$< M(-\sqrt{n}\,\bar{x}_n) \quad \text{if} \quad 0 < c' < -\sqrt{n}\,\bar{x}_n \quad \text{for some} \quad c'.$$

Since $\theta \varepsilon (0, n^{-1/2}\log n)$, $\exists\, t_o > 0$ such that

$$|h(\bar{x}_n)|\, I_{\left[\sqrt{n}|\bar{x}_n - \theta| < c(\log n)^{1/2}\right]}$$

$$< \sqrt{t_o} \quad \text{on} \quad (\sqrt{n}\,\bar{x}_n > 0) \cup (0 < -\sqrt{n}\,\bar{x}_n < c').$$

Note that

$$h(\bar{x}_n)\, I_{\left[\sqrt{n}|\bar{x}_n - \theta| < c(\log n)^{1/2}\right]} \geq 0 \quad \text{for sufficiently large} \quad n.$$

Hence $E_\theta\, h^2(\bar{x}_n)\, I_{[\sqrt{n}|\bar{x}_n - \theta| < c(\log n)^{1/2}]}$

$$< t_o + \int_{t_o}^{\infty} P_\theta[\sqrt{n}(\bar{x}_n - \theta) < - t^{1/2}/M - \sqrt{n}\,\theta]\,dt$$

$$\leq M \quad \text{for} \quad \theta \in (0, n^{-1/2} \log n)$$

and similarly (2.24) can be proved for $\theta \in (1 - n^{-1/2} \log n, 1)$.

2.5 <u>RELATION WITH THE WORK OF BURNASEV</u> : Let the prior $\rho \in C^3(R)$ be such that $\rho(\theta) > 0$ on (a_o, b_o) and $\rho(\theta) = 0$ on $(a_o, b_o)^c$. Let for some $\delta > 0\ \exists\ \delta_1 > 0$ such that

i) $\int |\frac{d}{d\theta} \log \rho(\theta)|^{2+\delta_1} \rho(\theta)\,d\theta < \infty$

ii) $\int |\frac{d^2}{d\theta^2} \log \rho(\theta)|^{1+\delta_1} \rho(\theta)\,d\theta < \infty$

iii) $\int_{D(\varepsilon)} \rho(\theta)\,d\theta = O(\varepsilon^{2+\delta_1})$ as $\varepsilon \longrightarrow o$

where

$$D(\varepsilon) = \left\{\theta : \sup_{|z| < \varepsilon} \frac{|\rho^{(3)}(\theta+z)|}{\rho(\theta)} > \varepsilon^{-(3-\delta)}\right\}$$

(Burnasev [1979] has conditions (i) and (ii), and (iii) with $D(\varepsilon) = \left\{\theta : \sup_{|z| < \varepsilon} |\frac{d^3}{d\theta^3} \log \rho(\theta + z)| > \varepsilon^{-(3-\delta)}\right\}$ along with condition

$$\int_{D^c(\varepsilon)} \rho(\theta) \sup_{|z| < \varepsilon} |\frac{d^3}{d\theta^3} \log \rho(\theta + z)|\,d\theta = O(\varepsilon^{-(1-\delta)})).$$

Below we sketch a proof of Theorem 2.1 when ρ satisfies (i) to (iii) above and assumptions AI to AV hold with some $r_1 > 4$ and $r_2 > 6$. Note that if $|A - a| \leq n^{-1/2} R_{1n}$, $|B - b_1| \leq R_{1n}$ and $|\frac{A}{B}| \leq C$ then using (1.2b) repeatedly it is easy to get

$$(2.25) \qquad |\frac{A}{B} - \frac{a}{b_1}| \leq (C + n^{-1/2})(|b_1|^{-1} R_{1n})^K + (|\frac{a}{b_1}| + n^{-1/2}) \sum_{r=1}^{K-1} (|b_1|^{-1} R_{1n})^r = R''_{nK} \text{ (say) for } K = 1, 2, \ldots .$$

Let with $I_{n,\theta}$ of Section 3,

$$A_{n\theta} = \{\underline{x}_n : |E(\theta|\underline{x}_n) - \hat{\theta}_n - n^{-1} \lambda_{1,2,n}(\underline{x}_n)| < M R'_{nK}\} \cap I_{n\theta}$$

where R'_{nK} is R''_{nK} of (2.25) with

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

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

$b_1 = n^{-1/2} C_n$, a is approximation for A, as given in (2.3a) and $R_{1n} = M c_n^4 \{ \sum_{i=0}^{2} |\rho^{(i)}(\hat{\theta}_n)| + \sup_{|z| < c_n} |\rho^{(3)}(\hat{\theta}_n + z)| + e^{-n\varepsilon_1} \}$, where C_n is C_n of Theorem 1.1 for $K = 2$ and c_n as in Lemma 1.3 and $\varepsilon_1 > 0$.

Hence using (2.8) and arguments similar to the proofs of (2.3a) and (1.17) we get (here we expand $\rho(\hat{\theta}_n + \phi b^{-1})$ only upto 4 terms)

$$(2.26) \qquad P_\theta(A_{n\theta}) = 1 - o(n^{-2}) \text{ uniformly in } \theta \in [a_0, b_0] .$$

In view of (i) to (iii) $\exists$ $\delta_o > 0$ and $\delta_2 > 0$ such that

(2.27) $\quad P_\rho(D_n) = o(n^{-1-\delta_2})$ where

$$D_n = \left\{\theta : \frac{|\rho^{(1)}(\theta)|}{\rho(\theta)} > n^{1/2-\delta_o} \text{ or } \frac{|\rho^{(2)}(\theta)|}{\rho(\theta)} > n^{1-\delta_o} \text{ or } y_{n\theta} = \sup_{|z|<2c_n} \frac{|\rho^{(3)}(\theta+z)|}{\rho(\theta)} > (2c_n)^{-3+\delta_o}\right\} \cup (\rho(\theta) < n^{-2})$$

Now we prove that

(2.28) $\quad E(R'^2_{nK}\, I_{A_{n\theta}}\, I_{D_n^c}) = o(n^{-2})$

for sufficiently large K depending on δ_o. Towards this end we first bound $|a\, b_1^{-1}|$ by a suitable function of θ.

Note that (vide (2.3c))

$$|a\, b_1^{-1}| \le M(|\lambda_{1,2,n}(\underline{x}_n)|b\, n^{-1} + |\lambda_{1,2,n}(\underline{x}_n) b\, \beta_2(\underline{x}_n)|n^{-2}|C_n|^{-1})$$

Now on $A_{n\theta}$, $\theta \varepsilon D_n^c$ using forms of $\lambda_{1,2,n}$, β_2 and C_n

$$|\lambda_{1,2,n}(\underline{x}_n)| \le M(\sum_{i=1}^{2} |\rho^{(i)}(\theta)|(\rho(\theta))^{-1} c_n^{i-1} + c_n^2\, y_{n\theta})$$

and

$$|\lambda_{1,2,n}(\underline{x}_n)||\beta_2(\underline{x}_n)||C_n|^{-1}$$

$$\le M(\sum_{i=1}^{2} |\rho^{(i)}(\theta)|(\rho(\theta))^{-1} c_n^{i-1} + c_n^2\, y_{n\theta}) .$$

$$(1 + \sum_{i=1}^{2} |\rho^{(i)}(\theta)|(\rho(\theta))^{-1} c_n^{i} + c_n^3\, y_{n\theta})$$

Also on $A_{n\theta}$, $\theta \varepsilon D_n^c$ we have

$$|b_1|^{-1} R_{1n} \leq M n^{1/2} c_n^4 \Big(\sum_{i=o}^{2} |\rho^{(i)}(\theta)(\rho(\theta))^{-1}| + y_{n\theta}$$
$$+ e^{-n\varepsilon_1}(\rho(\theta))^{-1}\Big)$$

Thus on $A_{n\theta}$, $\theta \varepsilon D_n^c$ all terms of R'_{nK} are of the form

$$\Big(\prod_{i=1}^{2} |\rho^{(i)}(\theta)(\rho(\theta))^{-1}|^{\ell_i}\Big) y_{n\theta}^m c_n^k$$

for some ℓ_1, ℓ_2, m and k all nonnegative.

Now we take a term of R'_{nK}, say, $\eta_n \equiv C(|b_1|^{-1} R_{1n})^K$ and show that

(2.28a) $$E(\eta_n^2 I_{A_{n\theta}} I_{D_n^c}) = o(n^{-2}).$$

Note that on $A_{n\theta}$, $\theta \varepsilon D_n^c$

$$\eta_n^2 \leq M n^K c_n^{8K}\Big(1 + \sum_{i=1}^{2} |\rho^{(i)}(\theta)(\rho(\theta))^{-1}| + y_{n\theta}\Big)^{2K}$$

Consider a term, say, $\eta_{1n} = M n^K c_n^{8K} y_{n\theta}^{2K}$

Note that

$$E(\eta_{1n} I_{A_{n\theta}} I_{D_n^c}) \leq M n^K c_n^{8K} (2c_n)^{-6K + 2K\delta_o}$$

$$= o(n^{-2}) \quad \text{for } K \text{ sufficiently large}$$

depending on δ_o.

Other terms of η_n^2 can be handled in a similar way by using (i), (ii) and the fact $\int y_{n\theta}\, \rho(\theta)d\theta < M$. Thus we get (2.28a). Other terms of R'^{2}_{nK} can be handled using similar arguments as needed to prove (2.28a). Thus the proof of (2.28) is completed.

CLAIM 1 : (2.6) holds with $I_{\triangle_n}$ replaced by $I_{A_{n\theta}}\, I_{D_n^c}$.

PROOF : Note that the statements

$$(2.29) \qquad E_\rho\left[(B_n-\theta)^2 I_{A_{n\theta}}\, I_{D_n^c}\right] = E_\rho\left[E_{1n}\, I_{A_{n\theta}}\, I_{D_n^c}\right] + o(n^{-2})$$

$$\text{and} \qquad E_\rho\left[E_{1n}\, I_{A_{n\theta}^c}\, I_{D_n^c}\right] = o(n^{-2})$$

are analogous to (2.15) and (2.16) respectively and can be proved similarly if we bound quantities like $\rho^{-i}(\theta)\rho^{(i)}(\theta)$ and $y_{n\theta}$ on D_n^c by using the definition of D_n^c, and use the integrability conditions (i) and (ii) and the fact $\int y_{n\theta}\, \rho(\theta)d\theta < M$. Also in view of (i), (ii) and (2.27) it is easy to see that

$$E(E_{1n}\, I_{D_n}\, I_{A_{n\theta}}) = o(n^{-2}) \quad \square$$

CLAIM 2 : $\displaystyle\int_{D_n} E_\theta(E(\theta|\underline{x}_n)-\theta)^2\, \rho(\theta)d\theta = o(n^{-2})$.

PROOF : Let
$$\tilde\theta_n = E(\theta|\underline{x}_n) \quad \text{if} \quad |\hat\theta_n - E(\theta|\underline{x}_n)| < 3c_n$$
$$= \hat\theta_n \qquad \text{otherwise}.$$

Let $W_n = \left\{\underline{x}_n : |\hat\theta_n - E(\theta|\underline{x}_n)| > 3c_n\right\}$, $G_{n\theta} = \left\{(\theta,\underline{x}_n) : |\hat\theta_n - \theta| < c_n\right\}$, and let W_n also denote $(a_o, b_o)\times W_n$.

Clearly $P_\rho(G^c_{n\theta}) = o(n^{-2} c_n^2)$ (vide Lemma 1.3 with $r_2 > 6$).

On $G_{n\theta} \cap W_n$, $c_n^2 < (E(\theta|\underline{x}_n) - \theta)^2 - (\hat{\theta}_n - \theta)^2$

hence $0 \geq R(\rho, E(\theta|\underline{x}_n)) - R(\rho, \hat{\theta}_n) \geq c_n^2 P_\rho(G_{n\theta} \cap W_n) - K P_\rho(G^c_{n\theta})$

and so we have $P_\rho(G_{n\theta} \cap W_n) = o(n^{-2})$.

Now $E_\rho(E(\theta|\underline{x}_n) - \theta)^2 I_{D_n} = J_1 + J_2 + J_3 + J_4$ where

$$J_1 = E_\rho(E(\theta|\underline{x}_n) - \theta)^2 I_{D_n \cap W_n^c \cap G_{n\theta}}$$

$$\leq c_n^2 P_\rho(D_n) = o(n^{-2}) \quad \text{(vide (2.27))},$$

$$J_2 = E_\rho(E(\theta|\underline{x}_n) - \theta)^2 I_{D_n \cap W_n^c \cap G^c_{n\theta}}$$

$$\leq K P_\rho(G^c_{n\theta}) = o(n^{-2}),$$

$$J_3 = E_\rho(E(\theta|\underline{x}_n) - \theta)^2 I_{D_n \cap W_n \cap G_{n\theta}}$$

$$\leq K P_\rho(G_{n\theta} \cap W_n) = o(n^{-2}),$$

$$J_4 = E_\rho(E(\theta|\underline{x}_n) - \theta)^2 I_{D_n \cap W_n \cap G^c_{n\theta}}$$

$$\leq K P_\rho(G^c_{n\theta}) = o(n^{-2}). \quad \square$$

Now using Claim 2 and (2.26) we have

$$E_\rho [\sqrt{n}(E(\theta|\underline{x}_n) - \theta)]^2 = I_1 + I_2 + I_3 + o(n^{-1})$$

where I_i's are defined as in (2.1c) with I_{Δ_n} replaced by

$n\, I_{A_{n\theta}}\, I_{D_n^c}$. Using (2.28) we get $I_2 = o(n^{-1})$. Also it is easy to see that $I_3 = o(n^{-1})$, hence $E_\rho[\sqrt{n}(E(\theta|\underline{x}_n) - \theta)]^2 = I_1 + o(n^{-1}) = a_1 + a_2 n^{-1} + o(n^{-1})$.

REMARK 2.2 : If ρ is continuously differentiable with $\rho(\theta) > 0$ on (a_o, b_o) and $\rho(\theta) = 0$ on $(a_o, b_o)^c$ and if for some $\delta > 0$

$$\int_{D_1^c(\varepsilon)} \sup_{|z| < \varepsilon} \frac{|\rho^{(1)}(\theta+z)|^2}{\rho(\theta)} d\theta = O(\varepsilon^{-(1-\delta)}) \quad \text{where}$$

$$D_1(\varepsilon) = \left\{ \theta : \sup_{|z|<\varepsilon} \frac{|\rho^{(1)}(\theta+z)|}{\rho(\theta)} > \varepsilon^{-(1-\delta)} \right\}$$

then we have $R_\rho = a_1 n^{-1} + o(n^{-1})$. Also it is clear that when ρ is uniform over (a_o, b_o) the above expansion for R_ρ holds. Also note that if $X_1, X_2, \ldots, X_n$ are iid $N(\theta, 1)$ and ρ is uniform over $(-\frac{h}{\sqrt{n}}, \frac{h}{\sqrt{n}})$ then Bayes estimate is given by

$$\frac{z}{\sqrt{n}} + \frac{1}{\sqrt{n}} \int_{-h-z}^{h-z} t\, e^{-t^2/2}\, dt \Big/ \int_{-h-z}^{h-z} e^{-t^2/2}\, dt$$

where z is $N(\sqrt{n}\,\theta, 1)$ and so coefficient of n^{-1} term in the Bayes risk is not one. This is due to the fact that some of the error terms neglected earlier now assume magnitude of order n^{-1}. This would be true for certain other smooth priors also if they are supported on $(-\frac{h}{\sqrt{n}}, \frac{h}{\sqrt{n}})$.

REMARK 2.3 : Priors of type D_S $(S \geq 4)$ or D_∞ satisfy both Burnasev's as well as our conditions (i) to (iii).

CHAPTER THREE

TWO APPLICATIONS

3.1 INTRODUCTION : In this chapter we give two applications of the results proved in Chapter two.

The first one consists of using the expansion of the integrated Bayes risk to prove Second Order Efficiency (SOE) of the m.l.e. whereas the expansions for posterior risk and for Bayes estimates are applied to a problem of sequential Bayes estimation; this is our second application. In Section 2, roughly speaking, we show how given an efficient estimate T_n we can get an estimate θ_n^* depending only on $\hat{\theta}_n$ such that θ_n^* is better upto $o(n^{-2})$ than T_n at any $\theta_o \in [a_o, b_o]$ (for a precise definition of SOE see Definition 3.1). Our proof of SOE of the m.l.e. is based on the heuristic arguments of Section 4 of Ghosh and Subramanyam [1974]. The basic idea is to get an approximate Bayes estimate for a sequence of priors which converge to a point θ_o at a certain rate and then utilize (4.18) of Ghosh and Subramanyam [1974]. This is done in Theorem 3.1. We have come to know (through a personal communication) that Bickel and Goetze have recently obtained a very general result of this type.

It may be observed that a similar result for quite general loss functions was proved by Ghosh, Sinha and Wieand [1981] but there, instead of obtaining an expansion of the concerned risk, expected loss is computed by integrating with respect to the relevant

Edgeworth expansions. This explains how our result differs from theirs. In the matter of proof also there is an important difference. There the estimation problem is reduced to a testing problem and then the asymptotic Bayes solutions of the testing problem are studied. This turns out to be simpler than the study of the Bayes estimates.

We should also mention in this connection Pfanzagl and Wefelmeyer [1978] and Akahira and Takeuchi [1976] where results similar to that of Ghosh, Sinha and Wieand [1981] are obtained independently. In these papers existence of stochastic expansions for the estimates permit rather direct calculations as in Section 2 of Ghosh and Subramanyam [1974] or Ghosh, Sinha and Subramanyam [1979]. It may be noted that by strengthening some of the assumptions about T_n and using arguments of Theorem 3.1 it is easy to prove the other version of SOE viz. the m.l.e. corrected for its bias is better than T_n corrected for its biase (vide Ghosh and Subramanyam [1974]).

We have defined the m.l.e. as in Lemma 1.2. It may be noted that by defining the m.l.e. in the usual way one can prove SOE of the m.l.e. using our arguments.

Our second application is to Bayesian sequential estimation. As pointed out in Section 4 of Chapter two, Corollary 2 of Alvo [1977] is false and hence his proof of Theorem 4 is not correct. The conclusion of his Theorem 4 happens to be essentially true if the prior is of type D_S $(S \geq 30)$ or D_∞. In Section 3 we adopt his notation and give a proof of a version of his Theorem 4 (vide Theorem 3.2).

Alvo [1977] considered the problem of estimating sequentially the parameter of one parameter exponential family when the loss is square error and the prior has compact interval as a support. In his Theorem 2 he obtained a asymptotic lower bound (as, the cost of single observation, $c \longrightarrow o$) for the Bayes risk of a stopping time. It was important, in general, to develop techniques by which one can get asymptotic expansion for the Bayes risk of a stopping time and in particular to get a stopping time which is asymptotically optimal. He proposes a simple stopping rule and his Theorem 4 asserts that Bayes risk of this stopping time (denoted by N in Section 3) has an expansion, the leading term of which is of order $c^{1/2}$ and agrees with the leading term of the above mentioned lower bound. The difference from the lower bound is $O(c)$. For the sake of completeness we describe the Alvo's setup in the beginning of the Section 3. This chapter is based on Sections 7 and 8 of Ghosh, Sinha and Joshi [1981].

3.2 SECOND ORDER EFFICIENCY : We begin with the definition of SOE. Let the parameter space $\textcircled{H}$ be an open subset of $\mathbb{R}$ and let $\mathcal{C}$ be a class of estimates.

DEFINITION 3.1 : An estimate $T_n \in \mathcal{C}$ is said to be SOE in $\mathcal{C}$ w.r.t. square error loss function and a compact subset $\underline{\underline{M}}$ of $\textcircled{H}$ if given any other estimate $T_n' \in \mathcal{C}$ there exists a function h such that uniformly in $\theta_o \in \underline{\underline{M}}$ we have

$$E_{\theta_o}[(T_n + h(T_n)n^{-1} - \theta_o)^2] \leq E_{\theta_o}[(T_n' - \theta_o)^2] + o(n^{-2}) .$$

Note that $h(T_n)$ depends on the sample only through T_n.

Now let assumptions AI to AV hold with some $K \geq 3$, $r_1 > 4$ and $r_2 > 4$. Define the m.l.e. $\hat{\theta}_n$ as in Lemma 1.2. Note that under above assumptions we have (vide (2.7))

$E_{\theta_o}(\hat{\theta}_n - \theta_o) = b_o(\theta_o)\, n^{-1} + o(n^{-1})$ uniformly in $\theta_o \in [a_o, b_o]$

where $b_o(\theta) = (L_{oo1}(\theta) + 2\, L_{11}(\theta))(2\, I^2(\theta))^{-1}$.

We also assume that $b_o(\theta)$ is twice continuously differentiable on $[c,d]$ and $H_{22}(\theta)$ (vide Proposition 2.1) is continuously differentiable on $[a_o, b_o]$.

Let ϕ be the class of the estimates T_n's satisfying (t_1) to (t_3) below

(t_1) $E_\theta(T_n - \theta)^2 = n^{-1}\, I^{-1}(\theta) + n^{-2}\, g(\theta) + o(n^{-2})$ uniformly in $\theta \in [a_o, b_o]$ where $g(\theta)$ is continuously differentiable in $\theta \in [a_o, b_o]$,

(t_2) $E_\theta(T_n - \theta) = n^{-1}\, b(\theta) + o(n^{-1-\varepsilon})$ for some $\varepsilon > 0$, uniformly in $\theta \in [a_o, b_o]$ and $b(\theta)$ is twice continuously differentiable in $\theta \in [c, d]$,

(t_3) $\sup_{\theta \in [a_o,\, b_o]} E_\theta(T_n - \theta)^4 < \infty$.

In Theorem 3.1 below we prove SOE of $\hat{\theta}_n$ w.r.t. square error loss function, class ϕ (as defined above) and $[a_o, b_o]$. As pointed out in the introduction, under appropriate assumptions (which include compactifiability of the parameter space) and defining the m.l.e. with $[c,d]$ of Lemma 1.2 replaced by closure, in $[-\infty, \infty]$, of $(\underline{\bar{H}})$ we can prove SOE of the m.l.e. for any compact subset of $(\underline{\bar{H}})$, by making minor modifications in the proof given below.

<u>THEOREM 3.1</u> : For any $T_n \in \phi$ and $\theta_o \in [a_o, b_o]$

$$E_{\theta_o}(\theta_n^* - \theta_o)^2 \leq E_{\theta_o}(T_n - \theta_o)^2 + o(n^{-2})$$

where $\theta_n^* = \hat{\theta}_n - (b_o(\hat{\theta}_n) - b(\hat{\theta}_n))n^{-1}$ and b is as in (t_2).

<u>PROOF</u> : Step 1 : We fix a $\theta_o \in [a_o, b_o]$; we prove that Proposition 2.2 and Theorem 2.1 hold with ρ replaced by ρ_n, ρ_n as defined below.

We will be using the fact that

(3.1a) $E_{\rho_n}[(B_n' - \theta)^2 I_{\triangle_n}]$ differs from the Bayes risk by $o(n^{-2})$

where B_n' and $\triangle_n$ are respectively B_n' and $\triangle_n$ of Chapter two with ρ replaced by ρ_n.

Let

$$\rho_n(\theta) = K_n \exp[-(\theta - \theta_o + (\log n)^{-1/4})^{-1}(\theta_o - \theta + (\log n)^{-1/4})^{-1}]$$

$$\text{for } \theta \in V_n \equiv (\theta_o - (\log n)^{-1/4}, \theta_o + (\log n)^{-1/4})$$

$$= 0 \qquad \text{elsewhere}.$$

Note that

$$1 \geq \int_{\theta_o-(\log n)^{-1/2}}^{\theta_o+(\log n)^{-1/2}} \rho_n(\theta)d\theta \geq K_n \, 2(\log n)^{-1/2} \, e^{-c(\log n)^{1/2}}$$

for some $c > 0$; this provides an upper bound for K_n.

In what follows let ε denote arbitrary positive number. It is easy to check that

(3.1) $\quad |\rho_n^{(i)}(\theta)| \leq n^{\varepsilon} \;\forall\, \theta \in V_n \,,\; \forall\, n \geq n_o(\varepsilon, i),\; i = 0,1,\ldots;$

$\rho_n(\theta)$ is of type D_∞ for each $n \geq 1$ and

(3.2) $\quad |\rho_n^{(i)}(\theta)|\,\rho_n^{v-1}(\theta) \leq K_n^v \, M(k,v) \;\forall\, \theta \in V_n \,,\; \forall\, n \geq 1,\; i = 1,\ldots,k$

and where $0 < M(k,v) < \infty$ depends only on v and k.

Now we show that Theorem 2.1 holds with ρ replaced by ρ_n. Towards this end first note that Proposition 1.2 holds that is

(3.3) $\quad P_{\rho_n}(\rho_n(\hat{\theta}_n) > n^{-m}) = 1 - O(n^{-m}) - o(n^{-(r_2-\varepsilon)/2})$

and that Lemmas 1.8 and 1.9 hold with R_n (vide Remark 1.3) having term $n^{-(K+r_o+1-\varepsilon)/2}$ instead of $n^{-(K+r_o+1)/2}$ (vide (3.1)).

Now using (3.1), (3.2), (3.3) and the upper bound n^{ε} for K_n, it is easy to see that Propositions 2.1 and 2.2 hold and hence Theorem 2.1 holds.

Let $\quad d_n(\theta) = I^{-1}(\theta)(I^{-1}(\theta)\; L_{co1}(\theta)/2 + \frac{d}{d\theta}\log \rho_n(\theta))$ then

$$B_n' = \hat{\theta}_n + d_n(\hat{\theta}_n)n^{-1}$$

Below are some more observations which we need in the sequel.

Let $c_{1n} < \theta_0$ and $c_{2n} > \theta_0$ be such that

(3.4a) $\rho_n(c_{1n}) = n^{-3} = \rho_n(c_{2n})$

then it is not defficult to see that with c_n of Lemma 1.3 we have

(3.4) $(|\hat{\theta}_n - \theta| < c_n) \Longrightarrow \rho_n(\hat{\theta}_n) > M\, n^{-3}$ for $\theta \in (c_{1n}, c_{2n})$ and

(3.5) $P_\theta(\triangle_n) \geq 1 - M\, n^{-(2+\varepsilon)}$ for $\theta \in (c_{1n}, c_{2n})$

(where $\triangle_n$ is $\triangle_n$ of (2.5) with ρ replaced by ρ_n).

STEP 2 : Now we prove that (4.18) of Ghosh and Subramanyam [1974] holds that is

(3.6) $$\int E_\theta(T_n - \theta)^2\, \rho_n(\theta)d\theta \geq \int E_\theta(\theta_n^* - \theta)^2 \rho_n(\theta)d\theta + o(n^{-2}).$$

Let $J_{n\theta} = (|\hat{\theta}_n - \theta| < c_n)$, note that (vide Lemma 1.3)

(3.7) $P_\theta(J_{n\theta}) = 1 - o(n^{-(2+\varepsilon)})$ uniformly in $\theta \in [a_0, b_0]$

Now on $\triangle_n \cap J_{n\theta}$ and $\theta \in (c_{1n}, c_{2n})$ in view of (3.2) and (3.4) we have with

$$f(\theta) = (b_0(\theta) - b(\theta)) \quad \text{and} \quad g_n(\theta) = (b_0(\theta) - b(\theta) + d_n(\theta))$$

and primes denoting the derivatives for g_n, f_n, b_0 and b

$$(B_n' - \theta)^2 = (\theta_n^* - \theta)^2 + n^{-2}g_n^2(\theta) + 2n^{-1}(\hat{\theta}_n - \theta)g_n(\theta) - 2n^{-2}g_n(\theta)f(\theta)$$
$$+ 2n^{-1}(\hat{\theta}_n - \theta)^2 g_n'(\theta) + o(n^{-(2+\varepsilon)}) \quad \text{and}$$

$$(T_n - \theta)^2 = (T_n' - \theta)^2 - n^{-2} g_n^2(\theta) - 2n^{-1} g_n(\theta)(T_n - \theta)$$
$$- 2n^{-1} g_n'(\theta)(T_n - \theta)(\hat{\theta}_n - \theta) - n^{-1}(\hat{\theta}_n - \theta)^2 (T_n - \theta) g_n''(\theta_{1n})$$
$$+ o(n^{-(2+\varepsilon)})$$

uniformly in $\theta \varepsilon (c_{1n}, c_{2n})$ where

$T_n' = T_n + n^{-1} g_n(\hat{\theta}_n)$ and θ_{1n} is between $\hat{\theta}_n$ and θ.

Hence using (t_1) to (t_3), (3.2), (3.5), (3.7) and (2.7) along with the fact that $\hat{\theta}_n \varepsilon [c,d]$ we get uniformly in $\theta \varepsilon (c_{1n}, c_{2n})$

(3.8) $$E_\theta(\theta_n^* - \theta)^2 - E_\theta [(B_n' - \theta)^2 I_{\Delta_n \cap J_{n\theta}}]$$
$$= n^{-2}[2g_n(\theta) f(\theta) - 2g_n'(\theta) I^{-1}(\theta) - 2g_n(\theta) b_o(\theta) - g_n^2(\theta)]$$
$$+ M\, n^{-(2+\varepsilon)}$$ and

(3.9) $$E_\theta [(T_n - \theta)^2 I_{\Delta_n \cap J_{n\theta}}] - E_\theta [(T_n' - \theta)^2 I_{\Delta_n \cap J_{n\theta}}]$$
$$= 2n^{-1} g_n'(\theta) E_\theta (T_n - \hat{\theta}_n)^2 - n^{-2} [g_n^2(\theta) + 2g_n(\theta) b(\theta)$$
$$+ 2g_n'(\theta)\, I^{-1}(\theta)] + M\, n^{-(2+\varepsilon)}.$$

Let $T_n'^*$ be the natural truncation of T_n'; using the boundedness of B_n' on Δ_n (vide (2.5)), (3.5), (3.7), (3.8) and (3.9) we have uniformly in $\theta \varepsilon (c_{1n}, c_{2n})$

$$E_\theta(T_n - \theta)^2 - E_\theta(\theta_n^* - \theta)^2$$
$$\geq E_\theta(T_n'^* - \theta)^2 - E_\theta [(B_n' - \theta)^2 I_{\Delta_n}] + 2n^{-1} E_\theta (T_n - \hat{\theta}_n)^2 + M\, n^{-(2+\varepsilon)}.$$

Now using the fact that θ_n^* and $T_n'^*$ are bounded and that B_n' is bounded on $\triangle_n$ we have, in view of (3.1a)

$$\int E_\theta(T_n - \theta)^2 \rho_n(\theta)d\theta - \int E_\theta(\theta_n^* - \theta)^2 \rho_n(\theta)d\theta$$

$$\geq 2n^{-1} \int_{c_{1n}}^{c_{2n}} E_\theta(T_n - \hat{\theta}_n)^2 g_n'(\theta) \rho_n(\theta)d\theta + o(n^{-2})$$

Now note that

$$|2n^{-1} \int_{c_{1n}}^{c_{2n}} E_\theta(T_n - \hat{\theta}_n)^2 g_n'(\theta) \rho_n(\theta)d\theta |$$

$$\leq 2n^{-1+\varepsilon} E_{\rho_n}(T_n - \hat{\theta}_n)^2 \qquad \text{(vide (3.2))}$$

$$\leq 2n^{-1+\varepsilon} \left[E_{\rho_n}^{1/2} [T_n - E(\theta|\underline{x}_n)]^2 + E_{\rho_n}^{1/2} [\hat{\theta}_n - E(\theta|\underline{x}_n)]^2 \right]^2$$

$$\leq 2n^{-2-\varepsilon} \qquad \text{as}$$

$$E_{\rho_n} [T_n - E(\theta|\underline{x}_n)]^2 = E_{\rho_n}(T_n - \theta)^2 - E_{\rho_n} [E(\theta|\underline{x}_n) - \theta]^2$$

$$= o(n^{-2}) \qquad \text{(vide } (t_1) \text{ and (a) of Theorem 2.1)}$$

and similarly $E_{\rho_n} [\hat{\theta}_n - E(\theta|\underline{x}_n)]^2 = o(n^{-2})$.

Thus the proof of (3.6) is completed.

<u>STEP 3</u> : Note that coefficient $g_1(\theta)$ of n^{-2} in the expansion (using (2.7)) of $E_\theta(\theta_n^* - \theta)^2$ is given by

$$g_1(\theta) = H_{2,2}(\theta) + f^2(\theta) - 2f(\theta)b_o(\theta) - 2f'(\theta)b_o(\theta) .$$

Using (t_1), (2.7) and (3.6) we have

$$n^{-2}\int (g(\theta) - g_1(\theta))\, \rho_n(\theta)d\theta + o(n^{-2}) \geq 0 \quad \text{hence}$$

$$n^{-2}\int [\, g(\theta_o) - g_1(\theta_o) + (\theta - \theta_o)(g'(\theta_1) - g_1'(\theta))]\, \rho_n(\theta)d\theta + o(n^{-2}) \geq 0$$

where θ_1 is between θ and θ_o.

Now using the continuity of g' and g_1' and the shrinking nature of V_n we have

$$n^{-2}(g(\theta_o) - g_1(\theta_o)) + o(n^{-2}) \geq 0 .$$

The proof of the theorem is now completed by using (t_1) and the expansion for $E_{\theta_o}(\theta_n^* - \theta_o)^2$ □

3.3 BAYESIAN SEQUENTIAL ESTIMATION : Let $X_1, X_2, \ldots, X_n$, be a sequence of iid r.v.; X_1 having p.d.f. $f(x, \theta)$ w.r.t. a signa finite measure, given by

$$f(x, \theta) = \exp(\theta x - K(\theta))$$

$\theta \in D = (a_2, a_1)$ where a_2, a_1 could be $-\infty$ and $+\infty$ respectively. Let θ have a prior density ψ w.r.t. Lebesgue measure. For any function w of θ let $\hat{w}_n$ and $\bar{w}_n$ denote respectively the m.l.e. and the Bayes estimate of w based on $\underline{x}_n$. Consider the problem of estimating the parameter θ sequentially when cost of single observation is c and the loss is square error. The object is to obtain an asymptotic expansion (as $c \rightarrow o$) for $\rho(\psi, N)$, the

Bayes risk, of a stopping time N given by

N = least integer $n \geq 1$ such that $n \geq c^{-1/2}\, \bar{\sigma}_n$

where $\sigma^{-2}(\theta)$ is the Fisher information number.

Following is a version of Theorem 4 of Alvo [1977]

THEOREM 3.2 : Let prior ψ be of type D_∞ or D_S $(S \geq 30)$ with $[a_o, b_o] \subset D$ then we have

$$\rho(\psi, N) = 2c^{1/2}\, E\, \sigma(\theta) + c\, E\, \ell(\theta) + o(c) \tag{3.10}$$

where

$$\ell(\theta) = (\lambda_{24}(\theta) - \lambda_{12}^2(\theta) - 2\sigma'(\theta)\, \lambda_{12}(\theta) - \sigma''(\theta)\sigma(\theta)),$$

primes denoting the derivatives and as in Johnson [1970] $\lambda_{ij}(\hat{\theta}_n)$ is the coefficient of $n^{-j/2}$ in the expansion of $E\,[((\theta - \hat{\theta})\sigma^{-1}(\hat{\theta}_n))^i \,|\, \underline{x}_n]$.

NOTE : Our proof of this theorem is based on the proof given by Alvo for his Theorem 4 and hence below we give only a sketch of it.

PROOF : Let ψ be of type D_∞; proof when ψ is of type D_S $(S \geq 30)$ is similar. Note that as in the case of our Lemma 2.1 we can get a set A_n such that for some $\beta > 3$, $P_\psi(A_n) \geq 1 - M.n^{-\beta}$ $\forall\, n \geq 1$ and such that on A_n we have

(a) $|E(\theta|\underline{x}_n) - \hat{\theta}_n - \hat{\sigma}\, \hat{\lambda}_{1,2}\, n^{-1} - \hat{\sigma}\, \hat{\lambda}_{1,4}\, n^{-2}\,| < R_n\, n^{-3}$,

(b) $|(\bar{\sigma}_n)^2 - (\hat{\sigma}_n)^2| < R_n\, n^{-1}$,

(c) $|(\bar{\sigma}_{n-1})^2 - (\hat{\sigma}_{n-1})^2| < R_n\, n^{-1}$,

(d) $|\bar{\lambda}_n - \hat{\lambda}_n| < R_n\, n^{-1}$,

(e) $|E\,[(\theta - \hat{\theta}_n)^2 | \underline{x}_n] - (\hat{\sigma}_n)^2 (n^{-1} + \hat{\lambda}_{2,4}\, n^{-2})| < R_n\, n^{-3}$,

(f) $|\hat{\lambda}_n| < R_n$,

(g) $|\bar{x}_{n-1}| < R_n$ and $|x_n| < R_n$,

where R_n is the generic notation for a random term with the property that $R_n\, n^{-\delta} < M$ for some $0 < \delta < 1/2\ \forall\ n \geq 1$. Note that (b) and (d) can be obtained by replacing $\phi^r \rho$ by $\sigma(\theta)\rho(\theta)$ and $\lambda(\theta)\rho(\theta)$ respectively in Theorem 3.1 of Johnson [1970] and then proceeding as in the case of our Lemma 2.1). Let $\rho_o(\psi_N)$ be the posterior risk of N ; we now estimate $E(\rho_o(\psi_N)\, I_{A_N})$.

Following Alvo and using (a), (b), (d) and (e) we have on A_n

(3.11) $$\rho_o(\psi_n) = 2c^{1/2}\,\bar{\sigma}_n + \bar{\lambda}_n c + n^{-1}(\bar{\sigma}_n - n\, c^{1/2})^2 + \bar{\lambda}_n\,[(\bar{\sigma}_n)^2 - n^2 c]\, n^{-2} + R_n\, n^{-3}$$

Now on $(N = n) \bigcap A_n$

$$0 \geq (\bar{\sigma}_n)^2 - n^2 c > (\bar{\sigma}_n)^2 - (\bar{\sigma}_{n-1})^2 - (2n-1)c$$
$$= \sigma^2(\hat{\theta}_n) - \sigma^2(\hat{\theta}_{n-1}) - (2n-1)c + R_n\, n^{-1}$$

by using (b) and (c)

$$= \frac{1}{n}(x_n - \bar{x}_{n-1})D'(\bar{x}_n^*)\ 2\sigma(\theta_n^*)\sigma'(\theta_n^*)$$

$$- (2n-1)c + R_n\ n^{-1}$$

where $\bar{x}_n^*$ is between $\bar{x}_n$ and $\bar{x}_{n-1}$, θ_n^* is between $\hat{\theta}_n$ and $\hat{\theta}_{n-1}$ and D is the map which takes $\bar{x}_n$ to $\hat{\theta}_n$. Also on $(N = n)$ we have

$$\delta.c^{-1/2} \leq n \leq 1 + [K\ c^{-1/2}] \quad \text{where} \quad 0 < \delta \leq \bar{\sigma}_n \leq K < \infty \quad \text{and}$$

$[y]$ denotes integer part of y (existence of such δ and K is ensured by the compact support of ϕ).

Hence on $(N = n) \cap A_n$ in view of (d), (f) and (g)

$$|\bar{\lambda}_n [(\bar{\sigma}_n)^2 - n^2c]\ n^{-2}| < M.c^{1+\varepsilon} \quad \text{for some} \quad \varepsilon > 0.$$

Using similar arguments, on $(N = n) \cap A_n$ we have

$$n^{-1}(\bar{\sigma}_n - n.c^{1/2})^2 < M.c^{1+\varepsilon} \quad \text{for some} \quad \varepsilon > 0.$$

Thus

(3.12) $$\left| \sum_{n \geq 1} E\left[\left\{ n^{-1} [\bar{\sigma}_n - n\ c^{1/2}]^2 + \bar{\lambda}_n [(\bar{\sigma}_n)^2 - n^2 c] n^{-2} + R_n\ n^{-3} \right\} I_{(N = n) \cap A_n} \right] \right|$$

$$\leq M\ c^{1+\varepsilon} \sum_{n \leq 1} P(N = n)$$

$$= M\ c^{1+\varepsilon} \quad \text{for some} \quad \varepsilon > 0$$

Now

$$E[\bar{\sigma}_n I_{(N = n)}\ I_{(N = n) \cap A_n^c}]$$

$$\leq E^{1/2}[(\bar{\sigma}_n)^2 I_{(N = n)}]\ P^{1/2}((N = n) \cap A_n^c)$$

$$\leq \left\{ 1 + E\left[(\bar{\sigma}_n)^2 \, I_{(N=n)} \right] \right\} P^{1/2}((N=n) \cap A_n^c)$$

hence

$$\sum_{n \geq 1} E\left[2\, \bar{\sigma}_n \, c^{1/2} \, I_{(N=n) \cap A_n^c} \right]$$

$$\leq 2\, c^{1/2} \left\{ M\, c^{\beta/2}\, E\sigma^2 + M\, c^{\beta/2} \left(\left[K\, c^{1/2} \right] + 1 \right) \right\}$$

(3.13) as $E^2(\sigma \,|\, \underline{x}_n) = (\bar{\sigma}_n)^2 \leq E(\sigma^2 \,|\, \underline{x}_n) = (\overline{\sigma^2})_n$

which is a martingale,

$$\leq M\, c^{1+\varepsilon} \quad \text{for some} \quad \varepsilon > 0 .$$

Using similar arguments.

(3.14) $$\sum_{n \geq 1} E\left\{ \bar{\ell}_n \cdot c\, I_{(N=n) \cap A_n^c} \right\}$$

$$\leq M\, c^{1+\varepsilon} \quad \text{for some} \quad \varepsilon > 0 .$$

Since $\bar{\sigma}_n$ and $\bar{\ell}_n$ are martingales and N is finite, $E(\bar{\sigma}_N) = E\sigma(\theta)$ and $E(\bar{\ell}_N) = E\,\ell(\theta)$.

Hence by (3.13) and (3.14) we have

$$E\left[\left\{ c^{1/2}\, \bar{\sigma}_N + c\, \bar{\ell}_N \right\} I_{A_N} \right]$$

$= c^{1/2}\, E\sigma(\theta) + c\, E\,\ell(\theta) + o(c)$ and so by (3.11) and (3.12)

$$E\left[\rho_o(\phi_N)\, I_{A_N} \right] = 2\, c^{1/2}\, E\sigma(\theta) + c\, E\,\ell(\theta) + o(c)$$

The proof of (3.10) is completed by noticing that

$$\rho(\phi, N) = E\left[\rho_o(\phi_N) \right]$$

$$= \sum_{n \geq 1} E\left\{ \rho_o(\phi_n)\, I_{(N=n) \cap A_n} \right\} + o(c) \qquad \Box$$

CHAPTER FOUR

EXPANSION OF BAYES RISK IN THE CASE OF DOUBLE EXPONENTIAL FAMILY

4.1 INTRODUCTION : In Chapter two we obtained expansion upto $o(n^{-2})$ for the Bayes risk under certain regularity conditions which include differentiability of log likelihood function (and some of it's darivatives). Under vary general conditions (such as LAN condition) the limit of the Bayes risk (w.r.t. bounded lossfunction)has been obtained. (see Strasser [1978] Proposition 2) So a natural question to ask is whether, under such general conditions as LAN, it is possible to get an expansion for the Bayes risk.

In this Chapter we consider a family of distributions (viz. double exponential with location parameter) which satisfy LAN condition but for which conditions of Chapter two do not hold and show that the Bayes risk w.r.t. square error loss function has expansion but now the term after the n^{-1} term is not of order n^{-2} (as was the case in Chapter two) but is of order $n^{-3/2}$. This indicates that under suitable strengthening of the LAN condition, it may be possible to get an expansion of the Bayes risk. In the following paragraph we sketch the method of proof which we followed in our special case and which is likely to succeed in the general case too.

A sequence of family of distribution is said to satisfy LAN condition if the log likelihood function $\Lambda(\theta_o, \theta_o + \delta n^{-1/2}) = \log$

$L(\theta_o + \delta n^{-1/2}|x_1,\ldots,x_n) - \log L(\theta_o|x_1,\ldots,x_n)$ can be approximated in the following way

$$|\Lambda(\theta_o, \theta_o + \delta n^{-1/2}) - \delta \sum_{i=1}^{n} h(\theta_o, x_i) - \delta^2 A(\theta_o)/2| \xrightarrow{P_{\theta_o}} 0$$

where $h(\theta_o, x_1)$ is normalized r.v. and $A(\theta_o) = n \operatorname{Var}_{\theta_o}(h(\theta_o, X_1))$. Now if we have following type of approximation

$$P_{\theta_o}\left\{|\Lambda(\theta_o, \theta_o + \delta n^{-1/2}) - \delta\Sigma h(\theta_o, x_i) - \delta^2 A(\theta_o)/2 - n^{-\alpha} V_n(\theta_o)| < n^{-\beta}\right\} \geq 1 - 0(n^{-\gamma})$$

for some suitable α, β and γ all positive then it is likely that by plugging in both, the numerator and the denominator of $B_n(\theta_o) = E[n^{1/2}(\theta - \theta_o)|x_1,\ldots,x_n]$ (w.r.t. some suitable prior for θ), the above approximation for the likelihood function one can get an approximation for $B_n(\theta_o)$ upto some suitable order (vide (4.9)). This in turn (vide (4.11)) gives the desired expansion of the Bayes risk in the present case; to extend it to the general LAN case may require non-trivial modifications. It may be observed that one of the main differences between the present investigation and Chapter two is that here we expand the likelihood around θ_o rather than the m.l.e. $\hat{\theta}_n$. Our main result (vide Theorem 4.1) is presented in Section 2. It may be interesting to note that the Bayes risk does not depend on the prior upto $0(n^{-3/2})$. In Section 3 some important lemmas are given; Section 4 gives some auxiliary lemmas. This chapter is based on Joshi [1982].

4.2 MAIN RESULT : Let $X_1, X_2, \ldots, X_n$ be iid r.v.'s with p.d.f. $f(x,\theta) = \frac{1}{2}\exp(-|x-\theta|)$, $-\infty < x < \infty$, $\theta \in \mathbb{R}$. Let ρ be the prior density of θ w.r.t. Lebesgue measure. Let for some $\eta_j > 0$ ($j = 1$ to 4) ρ satisfy (i) to (iv) below.

(i) $\rho(\theta) > 0$ for $\theta \in (a_o, b_o)$ and $\rho(\theta) = 0$ on $(a_o, b_o)^c$ for some $-\infty < a_o < b_o < \infty$,

(ii) $\int \rho(\theta)\,[(\rho)^{(1)}(\theta)]^{2+\eta_1}\,d\theta < \infty$

where $(\rho)^{(i)}(\theta) = \frac{d^i}{d\theta^i}\log\rho(\theta)$,

(iii) $\int_{D_1(\varepsilon)} \rho(\theta)d\theta = O(\varepsilon^{1+\eta_2})$ as $\varepsilon \longrightarrow 0$

where $D_1(\varepsilon) = [\sup_{|z|<\varepsilon} |(\rho)^{(2)}(\theta+z)| > \varepsilon^{-1+\eta_3}]$,

(iv) $\int_{(a_o, a_o+\varepsilon)\bigcup(b_o-\varepsilon,\, b)} \rho(\theta)d\theta = O(\varepsilon^{1+\eta_4})$ as $\varepsilon \longrightarrow 0$.

Let

(4.1) $$D_n^c = \Big[\sup_{|z|<n^{-1/2}\log^2 n} |(\rho)^{(2)}(\theta+z)| < n^{1/2-\eta},\ |(\rho)^{(1)}(\theta)| < n^{1/4-\eta}\Big]$$

$$\bigcap(\rho(\theta) > n^{-1})\bigcap(a_o + n^{-1/2}\log^3 n,\ b_o - n^{-1/2}\log^3 n)$$

then using (i) to (iv) we have for some $\eta > 0$ and $\eta' > 0$

(4.2) $$\int_{D_n} \rho(\theta)d\theta = O(n^{-1/2-\eta'}).$$

In this chapter, hence forward η and η' will be used as generic notations for positive constants.

Fix a $\theta_o \varepsilon D_n^c$; all the probability statements in the rest of the chapter hold good uniformly in $\theta_o \varepsilon D_n^c$.

(4.3) Note that $\text{Min}\,(\log^2 n, \sqrt{n}(b_o - \theta_o), \sqrt{n}(\theta_o - a_o)) = \log^2 n$.

<u>THEOREM</u> 4.1 : Bayes risk, $R(\rho)$, w.r.t. squared error loss function has following expansion

$$R(\rho) = n^{-1} + b_o n^{-3/2} + o(n^{-3/2})$$

where b_o is given by (4.12).

<u>PROOF</u> : Step 1 : We first approximate the likelihood ratio. Note that

(4.4) $$\exp[-\Lambda(\theta_o,\theta_o + \delta n^{-1/2})] = \exp(\sum_{i=1}^{n} |x_i-\theta_o| - \sum_{i=1}^{n} |x_i-\theta_o - \delta n^{-1/2}|)$$

$$= \exp[-(\delta+t)^2/2 + t^2/2 + n^{-1/4}(V_{1n}(\delta) + V_{2n}(\delta))] \text{ where}$$

$$V_{1n}(\delta) = n^{1/4} I(\delta > 0)[\Sigma 2(x_i-\theta_o-\delta n^{-1/2}) I(0 < x_i-\theta_o < \delta n^{-1/2}) + \delta^2/2],$$

$$V_{2n}(\delta) = n^{1/4} I(\delta < 0)[\Sigma 2(\theta_o+\delta n^{-1/2}-x_i) I(\delta n^{-1/2} < x_i-\theta_o < 0) + \delta^2/2],$$

$$t = n^{-1/2}\, 2(\Sigma I(x_i - \theta_o \leq 0) - n/2)$$

and <u>$I(A)$ represents indicator function of set A.</u>

<u>Step 2</u> : We now approximate $B_n(\theta_o)$. Let

$$B_n(\theta_o) = E(\sqrt{n}(\theta - \theta_o) | \underline{x}_n), \ \underline{x}_n = (x_1,\ldots,x_n) \text{ and}$$

$$N(i) = \int \delta^i I(-\sqrt{n}(\theta_o - a_o), \sqrt{n}(b_o - \theta_o))\phi(\delta + t) \times$$

(4.5) $$\exp\left[n^{-1/4}(V_{1n}(\delta) + V_{2n}(\delta)) + \log \rho(\theta_o + \delta n^{-1/2}) - \log \rho(\theta_o)\right] d\delta, \quad i = 0,1$$

where $\phi(x) = (2\pi)^{-1/2} \exp[-x^2/2]$; also $\Phi(x) = \int_{-\infty}^{x} \phi(x)dx$.

Then using (4.4)

(4.6) $$B_n(\theta_o) = N(1)/N(0).$$

Let $I_n(\delta) = I(|\delta| < \log^2 n)$ and

$$R_{1n}^{(i)} = \int I(|\delta| \geq \log^2 n)\, \delta^i \exp\left[n^{-1/4}(V_{1n}(\delta) + V_{2n}(\delta)) + \log \rho(\theta_o + \delta n^{-1/2}) - \log \rho(\theta_o)\right] \phi(t + \delta) d\delta.$$

We have, in view of (4.3),

$$N(i) = \int I_n(\delta)\delta^i \exp\left[n^{-1/4}(V_{1n}(\delta) + V_{2n}(\delta)) + \log \rho(\theta_o + \delta n^{-1/2}) - \log \rho(\theta_o)\right] \phi(t + \delta) d\delta + R_{1n}^{(i)}$$

$$= \int I_n(\delta)\delta^i \phi(t+\delta)\Big\{1 + n^{-1/4}(V_{1n}(\delta) + V_{2n}(\delta)) + n^{-1/2}(V_{1n}(\delta) + V_{2n}(\delta))^2/2 + n^{-3/4}(V_{1n}(\delta) + V_{2n}(\delta))^3\, e^{x_{1n}}/3!\Big\}\Big\{1 + n^{-1/2}\delta(\rho)^{(1)}(\theta_o) + n^{-1}((\rho)^{(1)}(\theta_o))^2\delta^2 e^{x_{2n}}/2\Big\}\Big\{1 + n^{-1}\delta^2(\rho)^{(2)}(\theta_1)e^{x_{3n}}/2\Big\} d\delta + R_{1n}^{(i)}$$

where x_{1n} is between 0 and $n^{-1/4}(V_{1n}(\delta) + V_{2n}(\delta))$, x_{2n} is between 0 and $n^{-1/2}(\rho)^{(1)}(\theta_o)$, x_{3n} is between 0 and

$n^{-1}\delta^2(\rho)^{(2)}(\theta_1)/2$ and θ_1 is between θ_o and $\theta_o + \delta n^{-1/2}$. Hence

(4.7) $N(i) = a_{on}^{(i)} + n^{-1/4}\, a_{1n}^{(i)} + n^{-1/2}\, a_{2n}^{(i)} + R_{1n}^{(i)} + R_{2n}^{(i)}$ where

$$a_{on}^{(i)} = \int I_n(\delta)\delta^i \phi(t+\delta)d\delta,$$

$$a_{1n}^{(i)} = \int I_n(\delta)\delta^i \phi(t+\delta)(V_{1n}(\delta) + V_{2n}(\delta))d\delta,$$

$$a_{2n}^{(i)} = \int I_n(\delta)\phi(t+\delta)\delta^i \left[(V_{1n}(\delta)+V_{2n}(\delta))^2/2 + \delta(\rho)^{(1)}(\theta_o)\right]d\delta \text{ and}$$

$$R_{1n}^{(i)} + R_{2n}^{(i)} = N(i) - (a_{on}^{(i)} + a_{1n}^{(i)} + a_{2n}^{(i)}).$$

Note that on A_n (vide (4.23)) we have by Lemmas 4.2 to 4.4 and (4.1)

(4.8) $a_{on}^{(o)} = 1 + o(n^{-1/2-\eta}),\ a_{on}^{(1)} = -t + o(n^{-1/2-\eta})$

and $|R_{jn}^{(i)}| = o(n^{-1/2-\eta})$ for $i = 0,1,\ j = 1,2$.

Absorbing $o(n^{-1/2-\eta})$ terms of $a_{on}^{(i)}$ in $R_{2n}^{(i)}$ and writting $c_{1n} = a_{1n}^{(1)} + ta_{1n}^{(o)}$;

$$c_{2n} = a_{2n}^{(1)} - a_{1n}^{(o)}\, a_{1n}^{(1)} - ta_{1n}^{(o)2} + ta_{2n}^{(o)};$$

$$R_{3n} = \Big\{R_{1n}^{(1)} + R_{2n}^{(1)} - (R_{1n}^{(o)} + R_{2n}^{(o)})(c_{on} + n^{-1/4}c_{1n} + n^{-1/2}c_{2n}) - n^{-3/4}(a_{1n}^{(o)}c_{2n} + a_{2n}^{(o)}c_{1n}) - n^{-1}a_{2n}^{(o)}c_{2n}\Big\}\cdot \Big\{1 + n^{-1/4}a_{1n}^{(o)} + n^{-1/2}a_{2n}^{(o)} + R_{1n}^{(o)} + R_{2n}^{(o)}\Big\}^{-1}$$

and $B_n'(\theta_o) = -t + c_{1n}\, n^{-1/4} + c_{2n}\, n^{-1/2}$ in view of (4.6), (4.7) and (4.8) we have on A_n

$$(4.9)\qquad B_n(\theta_o) = B_n'(\theta_o) + R_{3n},\ |B_n'(\theta_o)| < (\log n)^{\eta'},$$

$$|R_{3n}| = o(n^{-1/2-\eta}) \quad \text{and also}$$

$$(4.10)\qquad B_n'^2(\theta_o) = t^2 - 2tc_{1n}\, n^{-1/4} + (c_{1n}^2 - 2tc_{2n})n^{-1/2} + R_{4n},$$

$$|R_{4n}| = o(n^{-1/2-\eta}).$$

<u>Step 3</u> : We can now approximate the Bayes risk. Note that by using arguments similar to those given to prove claim 2 (Section 5) of Chapter two (here we use Lemma 4.1 in place of Lemma 1.3 and (4.2) in place of (2.27)) we have

$$\int I(D_n)E_{\theta_o}(B_n^2(\theta_o))\rho(\theta_o)d\theta_o = o(n^{-1/2-\eta})$$

Hence in view of (4.9)

$$(4.11)\qquad R(\rho) = n^{-1}\int E_{\theta_o}[B_n^2(\theta_o)]\,\rho(\theta_o)d\theta_o$$

$$= n^{-1}\int I(D_n^c)E_{\theta_o}[B_n'^2(\theta_o)I(A_n)]\,\rho(\theta_o)d\theta_o + o(n^{-3/2-\eta})$$

(vide (4.23))

$$= n^{-1}\int I(D_n^c)E_{\theta_o}\Big\{[t^2 + n^{-1/2}(-2tc_{2n} + c_{1n}^2 - 2tn^{1/4}c_{1n})]\cdot I(A_n)\Big\}\,\rho(\theta_o)d\theta_o + o(n^{-3/2-\eta})$$

$$= n^{-1} + n^{-3/2}\int I(D_n^c)E_{\theta_o}\Big\{(D_{1n} + D_{2n} + D_{3n})I(A_n)\Big\}\,\rho(\theta_o)d\theta_o + o(n^{-3/2-\eta})$$

where $D_{1n} = -2tc_{2n}$, $D_{2n} = c_{1n}^2$ and $D_{3n} = -2tn^{1/4}c_{1n}$.

Thus (vide Lemma 4.5)

$$R(\rho) = n^{-1} + b_o\, n^{-3/2} + o(n^{-3/2}),$$

where by Lemma 4.5

(4.12) $b_o = 4f_2 - 4f_3 - 4f_4 - 2f_5 - 2f_6$; and by (4.26)

$$= 24\int \bar{\Phi}(w)\phi^2(w)dw + 4\int (1-\bar{\Phi}(w))^2\, \bar{\Phi}(w)dw - \frac{6}{\sqrt{\pi}}$$

$$\geq 4\int (1-\bar{\Phi}(w))^2\, \bar{\Phi}(w)dw \quad \text{as} \int \bar{\Phi}(w)\phi^2(w)dw \geq \frac{1}{4\sqrt{\pi}}$$

This completes the proof of the theorem □

4.3 <u>SOME LEMMAS</u> : The following Lemma is a consequence of Lemma 3.1 of Reiss [1976].

<u>LEMMA 4.1</u> : Let $\bar{\theta}_n$ be the median of sample of size n from $f(x,\theta_o)$. For each $k > 0$ ther exists a constant $c > 0$ such that

$$P_{\theta_o}(n^{1/2}|\bar{\theta}_n - \theta_o| < c\, \log^{1/2} n) = 1 - 0(n^{-k}) .$$

Following lemma is a well know result (vide e.g. Serfling [1980] page 95 Lemma A) (below t is as in (4.4)).

<u>LEMMA 4.2</u> : $P_{\theta_o}(|t| < ((s+1)\log n)^{1/2}) \geq 1-0(n^{-s})$ for $s > 0$.

<u>LEMMA 4.3</u> : For each positive λ and ℓ

(4.13) $$P_{\theta_o}\Big\{\int I(\log^2 n,\ n^{1/2}(b_o - \theta_o))\delta\phi(t+\delta)\ \exp[n^{-1/4}V_{1n}(\delta)]d\delta < n^{-\ell},$$

$$\int I(-n^{1/2}(\theta_o - a_o), -\log^2 n)\delta\phi(t+\delta)\exp[n^{-1/4}V_{2n}(\delta)]\, d\delta < n^{-\ell},$$

$$|t| < ((\lambda+4)\log n)^{1/2}\Big\} \geq 1 - 0(n^{-\lambda}).$$

PROOF : Step 1 : For each $\varepsilon_1 > 0$ and $\log^2 n < \delta < n^{1/2}(b_0-\theta_0)\log^{-1} n$

(4.14) $$P_{\theta_0}\left[n^{-1/4} V_{1n}(\delta) + (\varepsilon_1 - 1)\delta^2/2 > 0 \right] \leq \exp\left[-\delta^2(1-\varepsilon_1 + 0(\log^{-1} n))/4 \right]$$

To get (4.14) note that LHS of (4.14) $\leq \exp[\varepsilon_1\delta^2/4]\, E^n \exp[X']$ where $X' = (X - \delta n^{-1/2}) I(0 < X < \delta n^{-1/2})$ and r.v. X has density $\exp[-|x|]/2 \quad -\infty < x < \infty$.

Hence LHS of (4.14) $\leq \exp[\varepsilon_1\delta^2/4 + n\log(1 - \delta^2 n^{-1}/4 + 0(\delta^3 n^{-3/2}))]$

Step 2 : For $n^{1/2}(b_0-\theta_0)\log^{-1} n \leq \delta \leq n^{1/2}(b_0-\theta_0)$ and $\varepsilon_2 > 0$

(4.15) $$\text{LHS of (4.14)} \leq \exp\left[\varepsilon_2\delta^2/4 - n(b_0-\theta_0)^2(\log^{-2} n)(1 + 0(\log^{-1} n))/4 \right].$$

To get (4.15) note that

$$\text{LHS of (4.14)} \leq \exp\left[\varepsilon_2\delta^2/4 + n\log(1 + e^{-\delta n^{-1/2}} + \delta n^{-1/2} e^{-\delta n^{-1/2}}) - n\log 2 \right]$$

$$\leq \exp\Big[\varepsilon_2\delta^2/4 + n\log\Big\{ 1 + \exp[-(b_0-\theta_0)\log^{-1} n]$$
$$+ (b_0-\theta_0)(\log^{-1} n)\exp[-(b_0-\theta_0)\log^{-1} n] \Big\} - n\log 2 \Big]$$

and now (4.15) follows easily.

Step 3 : For $\log^2 n < \delta < n^{1/2}(b_0 - \theta_0)$

(4.16) $$P_{\theta_0}\left[\inf_{(\delta - n^{-1}) \leq \delta' \leq \delta} n^{-1/4}\left[V_{1n}(\delta) - V_{1n}(\delta') \right] < -\log n \right]$$
$$\leq \exp[-\log^2 n].$$

To get (4.16) note that

$$(4.17) \qquad \inf_{\delta - n^{-1} \le \delta' \le \delta} n^{-1/4} [V_{1n}(\delta) - V_{1n}(\delta')]$$

$$= \inf_{\delta - n^{-1} \le \delta' \le \delta} [2\Sigma (x_i - \theta_o - \delta n^{-1/2}) I (\delta' n^{-1/2} < x_i - \theta_o \le \delta n^{-1/2})$$

$$-2(\delta - \delta') n^{-1/2} I (0 < x_i - \theta_o \le \delta' n^{-1/2}) - (\delta - \delta')(\delta + \delta')/2]$$

$$\ge 2\Sigma (x_i - \theta_o - \delta n^{-1/2}) I((\delta - n^{-1}) n^{-1/2} \le x_i - \theta_o \le \delta n^{-1/2})$$

$$- 2n^{-1/2} - n^{1/2}(b_o - \theta_o) n^{-1} .$$

Now (4.16) can be obtained by using exponential probability inequality for the first term of the last expression.

<u>Step 4</u> : Note that

$$(4.18) \qquad P_{\theta_o} \left[\int I(\log^2 n, n^{1/2}(b_o - \theta_o)) \delta \phi(t+\delta) \exp[n^{-1/4} V_{1n}(\delta)] d\delta \right.$$

$$\left. > n^{-\lambda}, \quad |t| \quad ((\lambda + 4) \log n)^{1/2} \right]$$

$$\le \sum_{i \varepsilon J_n} P_i \quad \text{where}$$

$$J_n = \left\{ [\log^2 n] + j/n : j = 0, 1, \ldots, n([n^{1/2}(b_o - \theta_o)] + 1 - [\log^2 n]) \right\},$$

$[g]$ = integer part of g and

$$P_i = P_{\theta_o} \left[\int I(i - n^{-1} < \delta < i) \delta \phi(t+\delta) \exp[n^{-1/4} V_{1n}(\delta)] d\delta > n^{-3/2-\lambda}, \right.$$

$$\left. |t| < ((\lambda + 4) \log n)^{1/2} \right].$$

Now choosing ε_1 of step 1 to be 1/2, ε_2 of step 2 to be $n^{-1/4}$ in view of step 3 and Lemma 4.2 it is easy to see that

$P_i \leq O(n^{-\lambda-3/2})$ uniformly in $i \in J_n$.

(4.19) Hence LHS of (4.18) $\leq O(n^{-\lambda})$.

A statement analogues to (4.19) for the part containing $V_{2n}(\delta)$ can be proved in a similar way; this completes the proof of the Lemma □

LEMMA 4.4 : For $\lambda > 0$ we have

(4.20) $$P_{\theta_o}\Big[\sup_{0 \leq |\delta| \leq \log^2 n} |V_{in}(\delta)| < 2\log^3 n\Big] \geq 1 - O(n^{-\lambda}) \text{ for } i = 1,2.$$

PROOF : Step 1 : For $0 \leq \delta \leq \log^2 n$

(4.21) $$P_{\theta_o}\big[|V_{1n}(\delta)| < \log^3 n\big] \geq 1 - \exp\big[-\log^2 n\,(1 + o(1))/2\big]$$

To prove (4.21) first note that

$$P_{\theta_o}\big[V_{1n}(\delta) > \log^3 n\big] \leq P_{\theta_o}\big[2\Sigma(x_i-\theta_o-\delta n^{-1/2}) I(0 < x_i-\theta_o \leq \delta n^{-1/2}) + \delta^2/2 > n^{-1/4}\log^3 n\big]$$

$\leq \exp\big[\delta^2 h/4 - n^{-1/4}(\log^3 n)\, h/2\big]\, E^n \exp\big[X'h\big]$, X' as defined in step 1 of Lemma 4.3

$\leq \exp\big[-(\log^2 n)(1 + o(1))/2\big]$ by choosing $h = n^{1/4}\log^{-1} n$.

Remaining part of (4.21) can be proved in a similar way.

Step 2 : For $0 \leq \delta \leq \log^2 n$

(4.22) $$P_{\theta_o}\Big[\sup_{\delta - n^{-1} \leq \delta' \leq \delta} |V_{1n}(\delta) - V_{1n}(\delta')| > \log^2 n\Big] \leq \exp\big[-\log^2 n/2\big].$$

To prove (4.22) first note that for $n^{-1} \leq \delta \leq \log^2 n$

$$\sup_{\delta-n^{-1}\le\delta'\le\delta} |V_{1n}(\delta)-V_{1n}(\delta')|$$

$$= \sup_{\delta-n^{-1}\le\delta'\le\delta} |2\Sigma(x_i-\theta_o-\delta n^{-1/2})I(\delta' n^{-1/2}\le x_i-\theta_o\le\delta n^{-1/2})$$

$$-(\delta-\delta')(\delta+\delta')/2-2(\delta-\delta')n^{-1/2}\Sigma I(0<x_i-\theta_o\le\delta' n^{-1/2})|n^{1/4}$$

$$\le n^{1/4}[n^{-1/2}\log^2 n+2\Sigma(\theta_o+\delta n^{-1/2}-x_i)I((\delta-n^{-1})n^{-1/2}$$

$$<x_i-\theta_o\le\delta n^{-1/2})].$$

Hence LHS of (4.22)

$\le \exp[n^{-1/2}(\log^2 n)h/2-n^{-1/4}(\log^2 n)h/2]\, E^n\exp[X''h]$ where

$X''=(\delta n^{-1/2}-X)I((\delta-n^{-1})n^{-\frac{1}{2}}\le X\le\delta n^{-1/2})$ and X as in Step 1 of Lemma 4.3.

$\le \exp[-\log^2 n/2]$ by choosing $h=n^{1/4}\log n$.

For $0\le\delta\le n^{-1}$ (4.22) can be proved in a similar way.

Step 3 : Proof of (4.20) for $i=1$ is completed by combining steps 1 and 2 ; for $i=2$ (4.20) can be proved analogously, completing the proof of the lemma □

Let A_n denote the common part of events of (4.13) and (4.20) then

(4.23) $\quad P_{\theta_o}(A_n)\ge 1-O(n^{-\lambda})$ for ℓ and λ both positive.

4.4 AUXILIARY LEMMAS : Let $P(t)$ be generic notation for a polynomial in t of finite degree and bounded coefficients. Let with t of (4.4)

$t_o \equiv \sum_{i=1}^{n} I(x_i - \theta_o \leq 0) = n/2 + n^{1/2} t/2$ and $F_n(t)$ be d.f. of t.

LEMMA 4.5 :

(4.24) $$E_{\theta_o}(D_{1n} I(A_n)) = -4(f_5 + f_4)/3 - 4f_3 + 0(n^{-\eta})$$

$$E_{\theta_o}(D_{2n} I(A_n)) = 4f_2 + 0(n^{-\eta}) \quad \text{and}$$

$$E_{\theta_o}(D_{3n} I(A_n)) = -2(4f_4 + f_5)/3 - 2f_6 + 0(n^{-\eta})$$

where $$f_1 = \int\int_t^{\infty} [\phi(w) - w(1 - \Phi(w))]^2 t^2 \phi(t) dw dt$$

$$f_2 = \int\int_t^{\infty} (1 - \Phi(w))^2 \phi(t) dw dt$$

(4.25) $$f_3 = \int\int_t^{\infty} 2(1 - \Phi(w)) [w(1 - \Phi(w)) - \phi(w)] t\phi(t) dw dt$$

$$f_4 = \int\int_t^{\infty} (w-t)^3 t^2 \phi(w)\phi(t) dw dt$$

$$f_5 = \int\int_t^{\infty} t(w-t)^4 \phi(w)\phi(t) dw dt$$

$$f_6 = \int\int_t^{\infty} t^3 (w-t)^2 \phi(w)\phi(t) dw dt .$$

Note that

(4.26) $$f_2 = \int (1 - \Phi(w))^2 \Phi(w) dw ,$$

$$f_3 = \frac{3}{\sqrt{\pi}} - 6 \int \Phi(w) \phi^2(w) dw ,$$

$$f_4 = 37/8\sqrt{\lambda}, \quad f_5 = -85/8\sqrt{\lambda} \quad \text{and} \quad f_6 = -13/8\sqrt{\lambda}.$$

PROOF : For $i = 0$ and 1 and $j = 1$ and 2 let

$$W_j^{(i)} = \int I_n(\delta)\delta^i V_{jn}(\delta)\phi(t+\delta)d\delta \quad \text{and}$$

$$W_{2+j}^{(i)} = \int I_n(\delta)\delta^i (V_{jn}(\delta))^2 \phi(t+\delta)d\delta$$

then on A_n

$$a_{1n}^{(i)} = W_1^{(i)} + W_2^{(i)} \quad \text{and}$$

$$a_{2n}^{(i)} = (W_3^{(i)} + W_4^{(i)})/2 + (\rho)^{(1)}(\theta_o)\left[1 + (i-1)(t+1) + it^{2i}\right] + o(n^{-\eta})$$

and hence

$$D_{1n} = 2t\left[(W_3^{(1)} + W_4^{(1)})/2 + (\rho)^{(1)}(\theta_o) - (W_1^{(1)}W_1^{(o)} + W_2^{(1)}W_2^{(o)} + W_1^{(1)}W_2^{(o)} + W_1^{(o)}W_2^{(1)})\right] + 2t^2\left[(W_1^{(o)})^2 + (W_2^{(o)})^2 + 2W_1^{(o)}W_2^{(o)} - (W_3^{(o)} + W_4^{(o)})/2\right]$$

$$D_{2n} = \left[(W_1^{(1)})^2 + (W_2^{(1)})^2 + 2W_1^{(1)}W_2^{(1)}\right] - 2t\left[W_1^{(1)}W_1^{(o)} + W_2^{(1)}W_2^{(o)} + W_1^{(o)}W_2^{(1)} + W_1^{(1)}W_2^{(o)}\right] + t^2\left[(W_1^{(o)})^2 + (W_2^{(o)})^2 + 2W_1^{(o)}W_2^{(o)}\right] \quad \text{and}$$

$$D_{3n} = n^{1/4}\left[2t(W_1^{(1)} + W_2^{(1)}) - 2t^2(W_1^{(o)} + W_2^{(o)})\right]$$

Note that for $i = 1$ to 3 $|D_{in}| \leq n^2 P(t)$ and hence $E_{\theta_o} D_{in}\, I(A_n^c) = o(n^{-\eta})$ by choosing λ (vide (4.23)) sufficiently large. Thus we have

$$E_{\theta_o}(D_{in}\, I(A_n)) = E_{\theta_o} E(D_{in}|t) + o(n^{-\eta}).$$

Now using Lemma 4.8 we get (4.24) □

<u>LEMMA</u> 4.6 : Let $z_1,\ldots,z_{(n-t_o)}$ be i.i.d. with p.d.f. e^{-z}, $z > 0$ and $y_1,\ldots,y_{t_o}$ be i.i.d. with p.d.f. e^{y}, $y < 0$ and let z_i's and y_i's be independent of each other and of $X_1,\ldots,X_n$ then conditional joint distribution of $\sum_{i=1}^{n} g_1(x_i-\theta_o)I(x_i-\theta_o \leq 0)$ and $\sum_{i=1}^{n} g_2(x_i-\theta_o)I(x_i-\theta_o > 0)$ given t_o is same as the joint distribution of $\sum_{i=1}^{t_o} g_1(y_i)$ and $\sum_{i=1}^{n-t_o} g_2(z_i)$ where $g_1 : R^- \longrightarrow R$ and $g_2 : R^+ \longrightarrow R$ are measurable functions.

<u>LEMMA</u> 4.7 : For $0 \leq \delta \leq \log^2 n$ we have

(4.27) $$E(z_1-\delta n^{-1/2})I(z_1 \leq \delta n^{-1/2}) = -\delta^2(2n)^{-1} + \delta^3(6n^{3/2})^{-1} + o(n^{-2}\delta^4)$$

(4.28) $$E(z_1-\delta n^{-1/2})^2\, I(z_1 \leq \delta n^{-1/2}) = \delta^3(3n^{3/2})^{-1} + o(n^{-2}\delta^4)$$

and for $-\log^2 n \leq \delta \leq 0$

(4.29) $$E(-y_1+\delta n^{-1/2})I(y_1 \geq \delta n^{-1/2}) = -\delta^2(2n)^{-1} - \delta^3(6n^{3/2})^{-1} + o(n^{-2}\delta^4)$$ and

(4.30) $$E(-y_1+\delta n^{-1/2})^2\, I(y_1 \geq \delta n^{-1/2}) = -\delta^3(3n^{3/2})^{-1} + o(n^{-2}\delta^4)$$

<u>LEMMA 4.8</u> : For $i = 1,2$

(a) $$E_{\theta_o} E\left[(W_i^{(1)})^2 | t\right] = 2\,(f_1 + f_2 + f_3) + O(n^{-\eta}),$$

$$E_{\theta_o} E\left[t^2 (W_i^{(o)})^2 | t\right] = 2f_1 + O(n^{-\eta}),$$

$$E_{\theta_o} E\left[t W_i^{(o)} W_i^{(1)} | t\right] = -(2f_1 + f_3) + O(n^{-\eta}),$$

$$n^{1/4} E_{\theta_o} E\left[t W_i^{(1)} | t\right] = \tfrac{1}{2} f_4 + \tfrac{1}{6} f_5 + O(n^{-\eta}),\quad E\left[W_i^{(j)} | t\right] = n^{-1/4+\eta} P(t)$$

for $j = 0,1,$

and

$$n^{1/4} E_{\theta_o} E\left[t^2 W_i^{(o)} | t\right] = \tfrac{1}{6} f_4 + \tfrac{1}{2} f_6 + O(n^{-\eta}).$$

(b) $$E_{\theta_o} E\left[t W_{2+i}^{(1)} | t\right] = \tfrac{2}{3} f_5 + O(n^{-\eta}) \quad \text{and}$$

$$E_{\theta_o} E\left[t^2 W_{2+i}^{(o)} | t\right] = \tfrac{2}{3} f_4 + O(n^{-\eta}).$$

<u>PROOF</u> : Let $u_{in} = 2\int I_n(\delta)\,\delta n^{1/4}(z_i - \delta n^{-1/2})\, I(z_i < \delta n^{-1/2})\,\phi(t+\delta)\,d\delta$

then by (4.27)

(4.31) $$E(u_{1n}) = n^{-3/4}\int I_n(\delta) I(\delta > 0)(-\delta^2 + \delta^3 n^{-5/4}/3)\,\phi(t+\delta) + O(n^{-7/4+\eta})$$

hence $(n - t_o) E(u_{1n}) = -\frac{n}{2}^{1/4}\int I_n(\delta) I(\delta > 0)\,\delta^3 \phi(t+\delta) + n^{-1/4+\eta} P(t).$

Now using Lemma 4.6 we have

$$E_{\theta_o} E\left[(W_1^{(1)})^2 \mid t\right] = E_{\theta_o} E\left\{ \sum_{i=1}^{n-t_o} (u_{in} - E(u_{1n})) + n^{-1/4+\eta} P(t) \right\}^2$$

$$= E_{\theta_o}\left\{ (n-t_o) E(u_{1n}^2) + n^{-1/2+\eta} P(t) \right\} \text{ in view of (4.31)}$$

$$= E_{\theta_o}\left(\frac{n}{2} E(u_{1n}^2) + n^{-1/4+\eta} P(t)\right) \text{ as } E(u_{1n}^2) \leq n^{-1/4+\eta} P(z_1 \leq n^{-1/2} \log^2 n)$$

$$= O(n^{-3/4+\eta})$$

$$= 2\int\int_0^{\log^2 n} \exp\left[-wn^{-1/2}\right]\left\{ \int_w^{\log^2 n} \delta(w-\delta)\phi(t+\delta)\,d\delta \right\}^2 dw\,dF_n(t) + O(n^{-\eta})$$

$$= 2\int I(|t| < 4\log^{1/2} n) \int_0^{\log^2 n} \exp\left[- wn^{-1/2}\right]\left\{ \int_w^{\log^2 n} \delta(w-\delta)\phi(t+\delta)\,d\delta \right\}^2 dw\,dF_n(t)$$

$$+ O(n^{-\eta})$$

$$= 2\int I(|t| < 4\log^{1/2} n) \int_0^{\log^2 n} \exp\left[- wn^{-1/2}\right]\left\{ \int_{w+t}^{\infty} (w(\delta-t) - (\delta-t)^2)\phi(\delta)\,d\delta \right\}^2$$

$$dw\,dF_n(t) + O(n^{-\eta})$$

$$= 2\int I(|t| < 4\log^{1/2} n) \int_t^{\log^2 n+t} \exp\left[- (w-t)n^{-1/2}\right]\left\{ t\phi(w) - (1-\Phi(w))(1+tw) \right\}^2$$

$$dw\,dF_n(t) + O(n^{-\eta})$$

$$= 2\int I(|t| < 4\log^{1/2} n) \int_t^{\infty} \left[t\phi(w) - (1-\Phi(w))(1+tw)\right]^2 dw\,dF_n(t) + O(n^{-\eta})$$

$$= 2\int\int_t^{\infty} \left[t\phi(w) - (1-\Phi(w))(1+tw)\right]^2 \phi(t)\,dw\,dt + O(n^{-\eta})$$

as $\int_t^{\infty} \left[t\phi(w) - (1-\Phi(w))(1+tw)\right]^2 dw \leq P(t)$ and

$$\int P(t)\,dF_n(t) = \int P(t)\phi(t)\,dt + O(n^{-\eta}) .$$

Thus we have $E_{\theta_o} E\left[(W_1^{(1)})^2 | t\right] = 2(f_1 + f_2 + f_3) + o(n^{-\eta})$.

Other assertions of (a) can be proved in a similar way.

Now we prove $E_{\theta_o} E\left[tW_3^{(1)} | t\right] = \frac{2}{3} f_5 + o(n^{-\eta})$; other assertions of (b) can be proved in a similar way.

Note that using Lemma 4.6 and (4.27) we have

$$E_{\theta_o} E\left[tW_3^{(1)} | t\right]$$

$$= E_{\theta_o} tE\Big\{ \int I_n(\delta)\delta n^{1/2}\Big[\sum_{i=1}^{n-t_o} \Big\{2(z_i - \delta n^{-1/2}) I(z_i < \delta n^{-1/2})$$

$$- 2E(z_1 - \delta n^{-1/2}) I(z_1 < \delta n^{-1/2})\Big\}$$

$$+ n^{-1/2+\eta} P(t)\Big]^2 \phi(t+\delta)d\delta \Big\}$$

$$= E_{\theta_o} t\Big\{\int I_n(\delta) I(\delta > 0)\delta n^{1/2}\Big[4(n-t_o)((n^{3/2})^{-1}\delta^3/3 + o(n^{-2}\delta^4))\Big]\phi(t+\delta)d\delta$$

$$+ n^{-1/2+\eta} P(t)\Big\} \text{ by using (4.27) and (4.28)}$$

$$= \frac{2}{3} E_{\theta_o} t \int I(0, \log^2 n)\, \delta^4 \phi(t+\delta)d\delta + o(n^{-\eta})$$

$$= \frac{2}{3} E_{\theta_o} I(|t| < 4\log^{1/2} n) \int I(t,\infty)(w-t)^4 \phi(w)dw + o(n^{-\eta})$$

$$= \frac{2}{3} \int t \int_t^\infty (w-t)^4 \phi(w)\phi(t)dwdt + o(n^{-\eta}) \quad \Box$$

PART II

CHAPTER FIVE

COMPARISON OF THE LIKELIHOOD RATIO, RAO'S AND WALD'S TESTS AND A CONJECTURE OF C.R. RAO

5.1 INTRODUCTION : Let $\{X_n\}_{n \geq 1}$ be a sequence of iid random variables with a common density $f(x,\theta)$ w.r.t. some sigma finite measure μ, $\theta \in \textcircled{H}$, $\textcircled{H}$ being an open subset of R^p. Consider the problem of testing a simple hypothesis $H_o : \theta = \theta_o = 0$ vs. $H_1 : \theta \neq \theta_o$ and three tests based on the statistics proposed by Neyman and Pearson, Rao and Wald (see Rao [1965], pages 347-352). We denote them by λ_n^1, λ_n^2 and λ_n^3 respectively. Under certain regularity conditions, these tests have same Pitman efficiency and so Rao [1965] raised the question (see the last paragraph of Section 6e.2) of higher order discrimination between these statistics and conjectured that the test proposed by him is likely to be " locally " more powerful than the others (in the second edition, this conjecture has omitted).

Peers [1971] attempted to settle the above conjecture. For contiguous alternatives of the form $\theta_n = \delta\, n^{-1/2}$, he expanded formally $P_{\theta_n}(\lambda_n^i > \chi^2_{p,\alpha})$ up to $o(n^{-1/2})$ where $\chi^2_{p,\alpha}$ is the 100α per cent point of the χ^2-distribution with p degrees of freedom and noted that these expansions do not support Rao's conjecture.

Under the assumptions of Section 3 of Chandra and Ghosh [1980], the validity of these expansions can be established. It then follows from Hodges and Lehmann [1970] that the mutual deficiencies (as computed by Peers) of these tests are $\pm\infty$, both values being possible. At first this seems to contradict Pfanzagl's well-known result for one-sided tests that "the first-order efficiency implies second-order efficiency" ; see, for example Pfanzagl ([1973a], [1980]). The following arguments will make it clear **that** Peers has really compared different cut-off points of the same statistic rather than different tests.

Take $p = 1$ and let $\hat{\theta}$ be the maximum likelihood estimator. It follows from Chandra and Ghosh ([1979], [1980]) that λ_n^1 can be written as

$$(n^{1/2}\hat{\theta})^2 (1 + n^{-1/2} T_1)^2 + o(n^{-1/2})$$

for a suitable statistic T_1 on a set A_n' with $P_{\theta_n}(A_n') = 1 + o(n^{-1/2})$, In the rest of this paragraph, we consider only subsets of A_n'. Then the region $\lambda_n^1 > \chi^2_{1,\alpha}$ reduces to the two-sided region,

(5.1) $$\left\{ n^{1/2}\hat{\theta}(1 + n^{-1/2} T_1) > z \text{ or } < -z \right\}$$

where

(5.2) $$z = (\chi^2_{1,\alpha})^{1/2}.$$

Using calculations similar to those of Ghosh, Sinha and Subramanyam [1979] (for details, see Chandra [1980]), the last critical region can be transformed in the form

$$(5.3) \qquad \left\{ n^{1/2}\,\hat{\theta} > t^1_{n,1} \text{ or } < t^1_{n,2} \right\}$$

where $t^1_{n,1}$ and $t^1_{n,2}$ are determined up to $o(n^{-1/2})$ such that the region (5.3) has the same size (and hence the same power because of uniformity of Edgeworth expansion of the m.l.e. over compact subsets of $(\underline{H})$) as that of the region (5.1). Similar results hold for Rao's and Wald's tests. Thus each of these tests can be viewed as a two-sided test based on the m.l.e. ; moreover, if we choose the cut-off points $t^i_{n,1}$ and $t^i_{n,2}$ such that the right-tailed tests have the same size as well as the left-tailed tests have the same size, then each of these (two-sided versions of) tests has the same power function, at least up to $o(n^{-1/2})$. In the special case of the one-parameter exponential family, this equivalence holds up to $o(n^{-1/2})$ for any $j \geq 1$ and hence the relative deficiencies of these tests are zero.

To really compare the tests, we adopt the following procedure. First we get statistics W^i_n $i = 1,2,3$ and a set A_n (see Section 2) such that

$$(5.4) \qquad P_{\theta_n}(A_n) = 1 + o(n^{-1}), \text{ uniformly over compact subsets of } \delta,$$

and on A_n,

$$(5.5) \qquad \lambda^i_n = (W^i_n)^2 + o(n^{-1}) \qquad i = 1,2,3 \;;$$

the existence of W^i_n follows as before from Chandra and Ghosh ([1979], [1980]). Here we have choosen W^i_n as functions of the first three derivatives at θ_0 of the loglikelihood --- details are given in Section 2. Let $\phi(t) = (2\pi)^{-1/2} \exp(-t^2/2)$, $\int_z^\infty \phi(t)dl = \frac{\alpha}{2}$ and

$$(5.6) \qquad \begin{aligned} P_{\theta_0}(W_n^1 > z) &= \frac{\alpha}{2} + (n^{-1/2} d_1 + n^{-1} d_2)\phi(z) + o(n^{-1}) \\ P_{\theta_0}(W_n^1 < -z) &= \frac{\alpha}{2} + (n^{-1/2} d_1' + n^{-1} d_2')\phi(z) + o(n^{-1}) \end{aligned}$$

and determine $t_{n,j}^i$ $(i = 2,3,\ j = 1,2)$ up to $o(n^{-1})$ such that

$$(5.7) \qquad \begin{aligned} P_{\theta_0}(W_n^i > t_{n,1}^i) &= P_{\theta_0}(W_n^1 > z) + o(n^{-1}) \\ P_{\theta_0}(W_n^i < t_{n,2}^i) &= P_{\theta_0}(W_n^1 < -z) + o(n^{-1}) \end{aligned}$$

The problem is to compare the asymptotic performances of the following tests (critical regions) :

$$(5.8) \qquad \begin{aligned} &\{W_n^i > t_{n,1}^i \text{ or } < t_{n,2}^i\} \quad i = 1,2,3 \\ &t_{n,1}^1 = z, \quad t_{n,2}^1 = -z \end{aligned}$$

Using the formal delta method (see Bhattacharya and Ghosh [1978] and Remark 3 of Chandra and Ghosh [1980]), one gets the asymptotic expansions of $P_{\theta_0}(W_n^i > t_{n,1}^i)$ and hence those of the powers of the tests (5.8). As argued in the previous paragraph, these expansions will agree upto $o(n^{-1})$, leading to <u>finite</u> deficiencies ; the last fact was noted by Chandra ([1980], Section 5, Chapter 3).

Coming back to the conjecture of Rao, we write the powers, $P^i \equiv P^i_{n,\delta,\alpha}$, of (5.8) as follows :

(5.9) $$P^i = P_0 + n^{-1/2} P^i_1 + n^{-1} P^i_2 + o(n^{-1}) \qquad i = 1,2,3$$

and compare P^i_2 as δ tends to zero as P^i_1 is same for $i = 1,2,3$ (vide Remark 5.2). Let $P^i_2(0)$, $P^i_2(\delta^2)$ be respectively the constant term and the coefficient of δ^2 in P^i_2 ; the coefficient of δ in P^i_2 is zero, since these tests are unbiased up to $o(n^{-1})$. Clearly

(5.10) $$P^1_2(0) = P^2_2(0) = P^3_2(0) = d_2$$

and so one can as well compare $P^i_2(\delta^2)$. In case of Cauchy distribution with θ as the location parameter, it can be shown that

(5.11) $$P^2_2(\delta^2) > P^1_2(\delta^2) \text{ iff } z > 2^{1/2}$$
$$P^2_2(\delta^2) > P^3_2(\delta^2) \text{ iff } z > 1 .$$

<u>We, therefore, say that Rao's conjecture is true if</u>

$$P^2_2(\delta^2) > \underline{\max}\, (P^1_2(\delta^2),\ P^3_2(\delta^2))$$

<u>for all sufficiently small</u> α. Since $P^i_2(\delta^2)$ is usually a polynomial (in z) of degree 3, another equivalent condition is

$$\lim_{\alpha \to 0} \lim_{\delta \to 0} \lim_{n \to \infty} n\delta^{-2} z^{-3} \left\{ P^2_{n,\delta,\alpha} - \max\, (P^1_{n,\delta,\alpha},\ P^3_{n,\delta,\alpha}) \right\} > 0.$$

Note that $P^i_{n,\delta,\alpha}$ depend on α through z (cf. (5.2)).

It is shown here that in the case $p=1$, Rao's conjecture is true. More precisely, we have shown that (see, in this connection, Equations (1) and (2) of Pfanzagl [1975])

(5.12) $$\lim_n n(P^2 - P^1) = \left\{\frac{1}{4}\delta^2 I \gamma_{\theta_0}^2 (z^3 - 2z) + o(\delta^3)\right\} \phi(z - \delta I^{1/2})$$

and

(5.13) $$\lim_n n(P^2 - P^3) = \left\{\delta^2 I \gamma_{\theta_0}^2 (z^3 - z) + o(\delta^3)\right\} \phi(z - \delta I^{1/2})$$

and therefore (5.11) holds (note that $z > 2^{1/2}$ iff $\alpha < 0.254$ (approx.) and $z > 1$ iff $\alpha < 0.317$ (approx.)). Here I is the Fisher information at θ_0, γ_{θ_0} is the statistical curvature at θ_0 introduced by Efron [1975] and is given by

$$\gamma_{\theta_0} = |N_{\theta_0}|^{1/2} I^{-3/2},$$

$|N_{\theta_0}|$ being the determinant of the dispersion matrix (under θ_0) of the first two derivatives at θ_0 of the loglikelihood. Finally, the deficiencies of the LR and Wald's tests with respect to Rao's test are given by

(5.14) $$D(LR, Rao) = \frac{1}{4}\gamma_{\theta_0}^2 (z^2 - 2) + o(\delta)$$

(5.15) $$D(Wald, Rao) = \gamma_{\theta_0}^2 (z^2 - 1) + o(\delta),$$

implying that

$$D(Wald, Rao) = 4D(LR, Rao) + \gamma_{\theta_0}^2 \quad \text{approximately.}$$

In Section 4, we have stated a general result on the Edgeworth expansions under θ_n of statistics of the form W_n^i (cf.(5.5)). As these expansions depend on the "approximate cumulants" under θ_n of W_n^i, it is convenient to express the derivatives at θ_o of the loglikelihood in terms of the (normalised) derivatives at θ_n --- this leads to the consideration of the random variables $W_n^{'}$ of the form (5.19). We do not need the exact values of W_n^i but they can be obtained, if desired, by substituting $\delta = 0$ in (5.19) and Appendix 1 (in particular, in (5.19) v_i and y_i are to be treated as arbitrary while u_i and x_i are suitable functions of v_i and y_i). Section 2 gives the main results i.e. Rao's conjecture and mutual dificiencies of the tests (5.8). Our assumptions and notations are given in the Section 3. The three appendices give the expression for various constants appearing in the text of the Chapter.

This chapter is based on Chandra and Joshi [1982].

5.2 MAIN RESULTS :

THEOREM 5.1 : Under assumptions of Section 3 Rao's conjecture is true.

PROOF : First we show that (5.4) and (5.5) hold. Note that (see Section 3 for notations)

$$\lambda_n^1 = 2\,(\sum_i \log f(x_i\,,\theta_o) - \sum_i \log f(x_i\,,\hat{\theta})),$$

$$\lambda_n^2 = (nI)^{-1}(\sum_i \frac{d}{d\theta} \log f(x_i\,,\theta_o))^2 \quad \text{and}$$

$$\lambda_n^3 = n\,(\hat{\theta} - \theta_o)^2\, I(\hat{\theta}).$$

Let

(5.15a) $$A_n = \left\{ \|n^{1/2} U_n\| < 3 \log n ,\ |V_n - E_{\theta_o}(V_n)| < 3n^{-1/2} \log n \right\}$$

where U_n is the vector whose components are

$$\hat{\theta},\ n^{-1} \sum_i \left[\frac{d^j}{d\theta^j} \log f(x_i, \theta_o) - E_{\theta_o}\left(\frac{d^j}{d\theta^j} \log f(X, \theta_o)\right)\right] \quad j = 2,3,4;$$

and $V_n = n^{-1} \sum \sup_{|\theta - \theta_o| < \varepsilon} \left| \frac{d^5}{d\theta^5} \log f(x_i, \theta) \right|$ for some $\varepsilon > 0$.

Note that under A (i) (5.4) holds; now we show that approximation of (5.5) holds on A_n.

First consider λ_n^1; any upper suffix i for a quantity will denote the corresponding quantity for λ_n^i e.g. W_n^i is W_n corresponding to λ_n^i $(i = 1,2,3)$. To get W_n, we proceed as in Section 4 of Chandra and Ghosh [1979] :

$$\lambda_n^1 = n\,\hat{\theta}^2 \Big\{ I_n - n^{-1/2}\left(\Delta_2 - \tfrac{1}{3} n^{1/2} \hat{\theta}\, \ell_3^* + \theta_n'\, \ell_3^*\right)$$
$$- n^{-1}\left(\theta_n' \Delta_3 + \tfrac{1}{2} \theta_n'^2\, \ell_4^* - \tfrac{1}{3} n^{1/2} \hat{\theta}\, \Delta_3\right.$$
$$\left. - \tfrac{1}{3} n^{1/2} \hat{\theta}\, \theta_n'\, \ell_4^* + \tfrac{1}{12} n\, \hat{\theta}^2\, \ell_4^*\right) + o(n^{-1}).$$

So by taking W_n^1 as below, (5.5) holds for λ_n^1;

$$W_n^1 = (nI_n)^{1/2}\,\hat\theta - \tfrac{1}{2}(nI_n)^{-1/2}\Big\{ n^{1/2}\hat\theta\Delta_2 + n^{1/2}\hat\theta\,\theta_n'\,\lambda_3^* - \tfrac{1}{3} n\,\hat\theta^2\lambda_3^*\Big\}$$

$$- \tfrac{1}{8} n^{-1} I_n^{-3/2}\Big\{ 4n^{1/2}\hat\theta\,\theta_n' I_n\Delta_3 + 2n^{1/2}\hat\theta\,\theta_n'^2 I_n\lambda_4^* - \tfrac{4}{3} n\,\hat\theta^2 I_n\Delta_3$$

$$- \tfrac{4}{3} n\,\hat\theta^2\theta_n' I_n\lambda_4^* + \tfrac{1}{3} n^{3/2}\hat\theta^3 I_n\lambda_4^*$$

$$+ n^{1/2}\hat\theta\Delta_2^2 + n^{1/2}\hat\theta\,\theta_n'^2\lambda_3^{*2} + \tfrac{1}{9} n^{3/2}\hat\theta^3\lambda_3^{*2}$$

$$+ 2n^{1/2}\hat\theta\,\theta_n'\Delta_2\lambda_3^* - \tfrac{2}{3} n\,\hat\theta^2\Delta_2\lambda_3^* - \tfrac{2}{3} n\,\hat\theta^2\theta_n'\lambda_3^{*2}\Big\} + o(n^{-1}).$$

Now using the following expansion for θ_n', $\theta_n' = \sqrt{n}(\hat\theta - \theta_n)$

$$\theta_n' = I_n^{-1}\Delta_1 + n^{-1/2}\Big\{ I_n^{-2}\Delta_1\Delta_2 + \tfrac{1}{2}\lambda_3^* I_n^{-3}\Delta_1^2\Big\}$$

$$+ n^{-1}\Big\{ I_n^{-3}\Delta_1\Delta_2^2 + \tfrac{3}{2} I_n^{-4}\Delta_1^2\Delta_2\lambda_3^* + \tfrac{1}{2} I_n^{-5}\Delta_1^3\lambda_3^{*2}$$

$$+ \tfrac{1}{6} I_n^{-4}\Delta_1^3\lambda_4^* + \tfrac{1}{2} I_n^{-3}\Delta_1^2\Delta_3\Big\} + o(n^{-1})$$

we can express W_n^1 in the form (5.19) with values of u, v etc. as in Appendix 1.

Clearly by taking

$$W_n^2 = (nI)^{-1/2}\sum_i \frac{d}{d\theta}\log f(x_i, \theta_o) \quad \text{and}$$

$$W_n^3 = (nI(\hat\theta))^{1/2}(\hat\theta - \theta_o)$$

(5.5) holds. Now expanding W_n^2 and W_n^3 on A_n upto $o(n^{-1})$ and using above expansion for θ_n' they can be put in the form

(5.19) (see Appendix 1 for values of u, v etc.). Now note that in view of (5.31) the term free from δ and coefficient of δ in P_2^i (vide (5.9)) is same for $i = 1,2,3$. Thus to settle Rao's conjecture one needs to obtain coefficient of the highest power of z in $I^{1/2} z Q + R$ (vide (5.26) and (E) of Lemma 5.2). Thus we have to obtain (vide (C) and (D) of Lemma 5.2), ψ^i, the coefficient of z^2 in Q^i and the coefficient of z^3 in R. In view of (5.28), (5.30) and Appendix 3 we have

$$\psi^1 = \frac{1}{24} I^{-3/2}(L_4 - 3L_{o2}) + \frac{1}{72} I^{-5/2}(9L_{11}^2 - 4L_3^2) + \frac{1}{6} I^{-1} L_3 d_1$$

$$\psi^2 = \frac{1}{24} I^{-3/2}(L_4 - 3I^2) - \frac{1}{18} I^{-5/2} L_3^2 + \frac{1}{6} I^{-1} L_3 d_1$$

$$\psi^3 = \frac{1}{24} I^{-3/2}(9I^2 - 12L_{o2} + L_4) - \frac{1}{18} I^{-5/2}(L_3^2 - 9L_{11}^2) + \frac{1}{6} I^{-1} L_3 d_1$$

and coefficient of z^3 in R^i (vide (5.26)) is $\frac{1}{72} I^{-2} L_3^2$. Hence using (5.3) it follows that the coefficients of z^3 in $\lim n(P^2 - P^i)$ $i = 1,3$ are as in (5.12) and (5.13). This completes the proof of the theorem $\square$

<u>REMARK 5.1</u> : Consider the Cauchy distribution with θ as the location parameter. Here $K_{j1}^i = 0$ for $i,j = 1,2,3$ and hence the powers of the tests (5.8) can be obtained very easily from (5.23). Using the integral

$$\int_{-\infty}^{\infty} \frac{x^{2n}}{(1+x^2)^m} = \frac{\Gamma(n+\frac{1}{2})\,\Gamma(m-n-\frac{1}{2})}{\Gamma m} \quad \text{if } m-n > \frac{1}{2}, \text{ one gets}$$

$I = \frac{1}{2}$, $L_4 = \frac{3}{8}$, $L_{101} = -\frac{3}{4}$, $L_{o2} = \frac{7}{8}$ and $L_{21} = -\frac{1}{8}$.

The powers upto $o(n^{-1})$ are given by

$$P^1 = A + n^{-1}\phi(z - \delta/\sqrt{2})\left\{(3/4)z - (3\sqrt{2}/8)z^2\delta + (6/8)z\delta^2 + 0(\delta^3)\right\}$$

$$P^2 = A + n^{-1}\phi(z - \delta/\sqrt{2})\left\{(3/4)z - (3\sqrt{2}/8)z^2\delta + (1/8)z\delta^2 + (5/16)z^3\delta^2 + 0(\delta^3)\right\}$$

$$P^3 = A + n^{-1}\phi(z - \delta/\sqrt{2})\left\{(3/4)z - (3\sqrt{2}/8)z^2\delta + (11/8)z\delta^2 - (15/16)z^3\delta^2 + 0(\delta^3)\right\}$$

$$A = \int_{z-\delta/\sqrt{2}}^{\infty} \phi(t)dt + \int_{z+\delta/\sqrt{2}}^{\infty} \phi(t)dt .$$

Thus the relation (5.11) holds.

THEOREM 5.2 : Under assumptions of Section 3 (5.14) and (5.15) hold.

PROOF : Note that the deficiency of LR test w.r.t to Rao's test is given by

$$D(\text{LR}, \text{Rao}) = \frac{2(P_2^2 - P_2^1)}{\delta(\phi(z - \delta I^{1/2}) - \phi(z + \delta I^{1/2}))I^{1/2}}$$

$$= \frac{(P_2^2 - P_2^1)}{\delta^2 I\, \phi(z - \delta I^{1/2}) + 0(\delta^3)}$$

and expression for D (Wald, Rao) is similar.

Hence to prove (5.14) and (5.15) it sufficies to prove that (5.12) and (5.13) hold.

Now to get (5.12) and (5.13) in view of (5.31) one needs the term free from z in Q^i $(i = 1,2,3)$ and the coefficient, Γ^i, of

z in R^i $(i = 1,2,3)$ in addition to ψ^i and coefficient of z^3 in R^i $(i = 1,2,3)$. Towards this end note that by using Lemma 5.3 and Appendix 3 we have the term free from z in Q^i is

$-\frac{1}{24}(L_4 - 3I^2)\, I^{-3/2} + \frac{1}{24} L_3^2\, I^{-5/2}$ for $i = 1,2,3$ and

$$\Gamma^1 = I^{-1}\left(\frac{1}{4} L_{21} + \frac{1}{4} L_{02} + \frac{1}{24} L_4 - \frac{1}{8} I^2\right) - I^{-2}\left(\frac{1}{4} L_{11}^2 + \frac{1}{4} L_{11} L_3 + \frac{1}{24} L_3^2\right)$$

$$\Gamma^2 = I^{-1}\left(\frac{1}{4} L_{21} + \frac{1}{24} L_4 + \frac{1}{8} I^2\right) - I^{-2}\left(\frac{1}{4} L_{11} L_3 + \frac{1}{24} L_3^2\right) \quad \text{and}$$

$$\Gamma^3 = I^{-1}\left(\frac{1}{4} L_{21} + \frac{1}{2} L_{02} + \frac{1}{24} L_4 - \frac{3}{8} I^2\right) - I^{-2}\left(\frac{1}{2} L_{11}^2 + \frac{1}{4} L_{11} L_3 + \frac{1}{24} L_3^2\right).$$

Now using (5.31) it is easy to see that (5.12) and (5.13) hold. This completes the proof of the theorem □

<u>REMARK 5.2</u> : A straightforward computation using Appendix 3 shows that

$$\frac{1}{2} K_{21}^i(\delta) - \frac{1}{3} I^{1/2} K_{31}^i(0) = \frac{1}{6} I^{-1} L_3$$

$$K_{11}^i(\delta^2) - \frac{1}{2} I^{1/2} K_{21}^i(\delta) + \frac{1}{6} I K_{31}^i(0) = \frac{1}{6} I^{1/2}(3L_{11} + L_3).$$

Thus in the view of (5.23) P_1^i (vide (5.9)) is same for $i = 1,2,3$. Also by using (5.23) for W_n^1 and the facts $k_{11}(\delta) = 0 = k_{21}(0) = k_{31}^1(0)$ we have the coefficient of δ in P_1^1 equal to

$$\phi(z - \delta I^{1/2})\, z\, (k_{21}^1(\delta) + 2I^{1/2} k_{11}^1(0)) = 0 \quad \text{(vide Appendix 3)}$$

Thus LR test of (5.8) is unbaised upto $o(n^{-1/2})$; as $P_1^i = P_1^1$ for $i = 2,3$ the same is true for other tests of (5.8).

5.3 NOTATIONS AND ASSUMPTIONS : First we give some notations.

Let $L_{i,j,k,\ell}(\theta) = E_\theta [(h_1(\theta))^i (h_2(\theta))^j (h_3(\theta))^k (h_4(\theta))^\ell]$

where $h_i(\theta) = \frac{d^i}{d\theta^i} \log f(x, \theta) \quad i = 1 \text{ to } 4$;

$L_{i,j,k}(\theta) = L_{i,j,k,o}(\theta)$, $L_{i,j}(\theta) = L_{i,j,o}(\theta)$, $L_i(\theta) = L_{i,o}(\theta)$

$L_{i,j,k,\ell} = L_{i,j,k,\ell}(\theta_o)$ etc.

When there is no possibility of confusion the commas in the sufix will be dropped e.g. $L_{1,1}(\theta)$ will be written as $L_{11}(\theta)$.

We also use the following notations.

$$\theta_n = \theta_o + \delta\, n^{-1/2}$$

$$\ell_i^* = E_{\theta_n}\left(\frac{d^i}{d\theta^i} \log f(x, \theta_n)\right), \; I_n = -\ell_2^*,$$

$$\ell_{ijk} = E_{\theta_n} [(h^{(1)})^i \, (h^{(2)})^j \, (h^{(3)})^k] \quad \text{where}$$

$$h^{(i)} = \frac{d^i}{d\theta^i} \log f(x, \theta_n) - \ell_i^* \quad \text{for } i = 1,2,3 ;$$

$$\ell_{ij} = \ell_{ijo} \quad \text{and} \quad \ell_i = \ell_{ioo} ;$$

$$\Delta_i = n^{-1/2} \left\{ \sum_{j=1}^{n} \frac{d^i}{d\theta^i} \log f(x_j, \theta_n) - n\, \ell_i^* \right\} \quad i = 1,2,3.$$

Let k_{ij} denote the coefficient of $n^{-j/2}$ in the ith

" approximate cumulant " of W_n (vide (5.18) and (5.19)), $i = 1$ to 4, $j = 0$ to 2 and k_{ij} being free from n; $k_{ij}(0)$, $k_{ij}(\delta)$ and $k_{ij}(\delta^2)$ denote respectively term free from δ, coefficient of δ and coefficient of δ^2 in k_{ij}.

Our m.l.e. $\hat{\theta}$ is a solution of the likelihood equation. For the proofs of Sections 2 and 4 we need to assume following facts :

A(i) : Let A_n be defined as in (5.15a) ;

$$P_{\theta_n}\left[A_n \bigcap \left(\sum_{i=1}^{n} \frac{d}{d\theta} \log f(x_i , \hat{\theta}) = 0 \right) \right] = 1 - o(n^{-1})$$

uniformly over compact subsets of δ.

A(ii) : Let W_n' and W_n be as in (5.18) and (5.19) ; Edgeworth expansions upto $o(n^{-1})$ under θ_n for W_n' and W_n, obtained by the formal delta method are valid with remainder term $o(n^{-1})$ uniformly over compact subsets of δ.

A(iii): Let U_{θ_o} be a nhbd of θ_o :

(a) $L_{8,0,4}(\theta)$, $L_{8,4}(\theta)$, $L_{4,8}(\theta)$, $L_{0001}(\theta)$ and $L_{12,0}(\theta)$ are all bounded on U_{θ_o}.

(b) $\int \frac{d^i}{d\theta^i} f(x , \theta) \; d\mu(x) = 0$ for $\forall\, \theta \in U_{\theta_o}$ $i = 1$ to 4.

(c) $L_{02}(\theta)$, $L_{101}(\theta)$, $L_{21}(\theta)$, $L_4(\theta)$ and $L_{0001}(\theta)$ are all continuous on U_{θ_o}.

(d) $L_{11}(\theta)$, $L_{001}(\theta)$ and $L_3(\theta)$ are continuously defferentiable on U_{θ_o}.

(e) $L_2(\theta) = I(\theta)$ is twice continuously differentiable on U_{θ_o}

Note that A(i) and A(ii) hold if for example (A_1) to (A_6) of Bhattacharya and Ghosh [1978] hold with $s = 4$ (see also Remark 3 of Chandra and Ghosh [1980]) and A(iii) holds if for example (a) of A(iii) holds and conditions (3) and (4) of Gusev [1976] hold with $k = 5$. Also note that if we define the m.l.e $\hat{\theta}$ as in our Lemma 1.2 then under appropriate conditions A(i) holds.

In Sections 2 and 4 we will be using following consequences of A(iii) without any further reference to A(iii) (primes denote derivatives)

$$L_{001}(\theta) + 3L_{11}(\theta) + L_3(\theta) = 0 \;;$$

$$L_{0001}(\theta) + 4L_{101}(\theta) + 3L_{02}(\theta) + 6L_{21}(\theta) + L_4(\theta) = 0 \;;$$

$$L_2'(\theta) = 2L_{11}(\theta) + L_3(\theta) \;; \quad L_{11}'(\theta) = L_{21}(\theta) + L_{101}(\theta) + L_{02}(\theta) \quad \text{etc.}$$

5.4 <u>SOME LEMMAS</u> : Let A(i) to A(iii) holds. Consider an asymptotically normally distributed statistic W_n and a set A_n such that (5.4) holds and on A_n

(5.18) $$W_n = W_n' + o(n^{-1}) \quad \text{where}$$

(5.19) $$W_n' = \delta I_n^{1/2} + \Delta_1 I_n^{-1/2} + n^{-1/2}\left\{u_1\delta\Delta_2 + u_2\delta^2 + u_3\delta\Delta_1 + v_1\Delta_1\Delta_2 + v_2\Delta_1^2\right\}$$

$$+ n^{-1}\{x_1\delta\Delta_1\Delta_2 + x_2\delta\Delta_1^2 + x_3\delta\Delta_1\Delta_3 + x_4\delta^2\Delta_3 + x_5\delta^2\Delta_1 + x_6\delta^3 + x_7\delta\Delta_2^2 + x_8\delta^2\Delta_2 + y_1\Delta_1\Delta_2^2 + y_2\Delta_1^2\Delta_2 + y_3\Delta_1^3 + y_4\Delta_1^2\Delta_3\} + o(n^{-1})$$

and the coefficients u_i, v_i, x_i and y_i are nonrandom. Note that Edgeworth expansions for W_n' and W_n obtained by the formal delta method are valid. Let k_{ni} be the ith " approximate cummulant " of W_n' (under θ_n) and let

(5.20) $$k_{ni} = k_{io} + n^{-1/2} k_{i1} + n^{-1} k_{i2} + o(n^{-1}) \quad i = 1 \text{ to } 4$$

where k_{ij}'s are free from n. Then using delta method we have

(5.21) $$P_{\theta_n}(W_n - \delta I^{1/2} \leq x) = \int_{-\infty}^{x} g_n(y)\phi(y)dy + o(n^{-1})$$

uniformly in $x \in R$ and over compact subsets of δ where

$$g_n(x) = 1 + n^{-1/2}\left\{xk_{11} + \frac{x^2-1}{2} k_{21} + \frac{x^3-3x}{6} k_{31}\right\} + n^{-1}\left\{xk_{12} + \frac{x^2-1}{2}[k_{22} + (k_{11})^2] + \frac{x^3-3x}{6}(k_{32} + 3k_{21}k_{11}) + \frac{x^4-6x^2+3}{24}[k_{42} + 3(k_{21})^2 + 4k_{11}k_{31}] + \frac{x^5-10x^3+15x}{12} k_{21}k_{31} + \frac{x^6-14x^4+42x^2-15}{72}(k_{31})^2\right\}.$$

It is easy to verify that

(5.22) $$k_{10} = \delta I^{1/2},\ k_{20} = 1,\ k_{30} = k_{40} = k_{41} = 0.$$

Appendix 2 gives the expressions for remaining k_{ni}'s.

Clearly for any a, b and c we have

$$(5.23)\quad P_{\theta_n}(W_n - \delta I^{1/2} > a + n^{-1/2}b + n^{-1}c)$$

$$= \int_a^\infty \phi(x)dx + n^{-1/2}\phi(a)\left\{ -b+k_{11}+\frac{a}{2}k_{21} + \frac{a^2-1}{6}k_{31}\right\}$$

$$+ n^{-1}\phi(a)\left\{\frac{1}{2}ab^2 - c - abk_{11} + \frac{b-a^2b}{2}k_{21}\right.$$

$$+ \frac{3ab-a^3b}{6}k_{31} + k_{12} + \frac{1}{2}a\left[k_{22} + (k_{11})^2\right]$$

$$+ \frac{a^2-1}{6}(k_{32} + 3k_{21}k_{11}) + \frac{a^3-3a}{24}\left[k_{42} + 3(k_{21})^2 + 4k_{11}k_{31}\right]$$

$$\left. + \frac{a^4-6a^2+3}{12}k_{21}k_{31} + \frac{a^5-10a^3+15a}{72}(k_{31})^2\right\} + o(n^{-1}),$$

uniformly over compact subsets of δ.

In what follows let $k_{ij}(0)$ denote term free from δ in k_{ij} and $k_{ij}(\delta)$ and $k_{ij}(\delta^2)$ denote respectively the coefficients of δ and δ^2 in k_{ij}.

Note that W_n^1 (vide (5.5) and proof of Theorem 5.1) is of the form W_n above and hence (5.2) holds for W_n^1.

Let d_i and d_i' $(i = 1,2)$ be defined by (5.6). Let b_i and c_i $(i = 1,2)$ be such that

$$(5.24)\quad P_{\theta_0}(W_n \geq z + n^{-1/2}b_1 + n^{-1}c_1) = \frac{\alpha}{2} + \phi(z)(n^{-1/2}d_1 + n^{-1}d_2) + o(n^{-1})$$

$$(5.25)\quad P_{\theta_0}(W_n \leq -z + n^{-1/2}b_2 + n^{-1}c_2) = \frac{\alpha}{2} + \phi(z)(n^{-1/2}d_1' + n^{-1}d_2') + o(n^{-1})$$

where $1 - \Phi(z) = \frac{\alpha}{2}$,

LEMMA 5.1 :

(A) $b_1 = k_{11}(0) + \frac{1}{6}(z^2 - 1)\, k_{31}(0) - d_1$,

$$d_1 = -\frac{1}{6} I^{-3/2} L_3, \quad d_1' = -d_1$$

i.e. b_1 is an even polynomial of degree 2 in z.

Also $b_2 = b_1$

(B) c_1 is a odd polynomial of degree 5 in z.

Also $c_2 = -c_1$.

PROOF : Using (5.23) and then equating coefficients of $n^{-1/2}$ on LHS and RHS of (5.24) we have

$$b_1 = k_{11}(0) + \frac{z}{2} k_{21}(0) + \frac{(z^2-1)}{6} k_{31}(0) - d_1$$

$$= k_{11}(0) + \frac{(z^2-1)}{6} k_{31}(0) - d_1 \quad \text{as} \quad k_{21}(0) = 0 \quad \text{(vide Appendix 2).}$$

Similarly

$$b_2 = k_{11}(0) + \frac{(z^2-1)}{6} k_{31}(0) + d_1' .$$

Now using (5.23) (as applied to W_n^1) and (5.6) we have

$$d_1 = k_{11}^1(0) + \frac{1}{6}(z^2-1)\, k_{31}^1(0) = -\frac{1}{6} I^{-3/2} L_3$$

(vide Appendix 3) and $d_1 = -d_1'$. Thus $b_1 = b_2$. This completes the proof of part (A); part (B) follows using similar arguments (here we use the fact $d_2 = -d_2'$) □

Let the coefficient of n^{-1} in $P_{\theta_0}(W_n \geq z + n^{-1/2} b_1 + n^{-1} c_1)$ be written in the form

(5.26) $\phi(z-\delta I^{1/2})\left\{d_2+\delta Q+\delta^2 R+O(\delta^3)\right\}$.

LEMMA 5.2 : (C) Q is an even polynomial of degree 2 in z ; in fact

(5.27) $$Q=-\frac{1}{2}I^{1/2}(b_1)^2+I^{1/2}b_1k_{11}(0)+\frac{1}{2}b_1(1-z^2)[k_{21}(\delta)-I^{1/2}k_{31}(0)]+k_{12}(\delta)-\frac{1}{2}I^{1/2}[k_{22}(0)+(k_{11}(0))^2]+(z^2-1)[\frac{1}{6}k_{32}(\delta)+\frac{1}{2}k_{21}(\delta)k_{11}(0)-\frac{1}{8}I^{1/2}k_{42}(0)-\frac{1}{2}I^{1/2}k_{11}(0)k_{31}(0)]+(z^4-6z^2+3)[\frac{1}{12}k_{21}(\delta)k_{31}(0)-\frac{5}{72}I^{1/2}(k_{31}(0))^2].$$

Morever, the coefficient, ϕ, of z^2 in Q is

(5.28) $$\phi=\frac{1}{6}k_{32}(\delta)-\frac{1}{8}I^{1/2}k_{42}(0)+\frac{1}{18}I^{1/2}(k_{31}(0))^2-\frac{1}{9}I^{-1}L_3\,k_{31}(0)+\frac{1}{6}I^{-1}d_1L_3.$$

(D) R is a odd polynomial of degree 3 in z ; in fact

(5.29) $$R=zb_1[-k_{11}(\delta^2)+I^{1/2}k_{21}(\delta)-\frac{1}{2}I\,k_{31}(0)]+z[\frac{1}{2}k_{22}(\delta^2)+k_{11}(0)k_{11}(\delta^2)-\frac{1}{3}I^{1/2}k_{32}(\delta)-I^{1/2}k_{21}(\delta)k_{11}(0)+\frac{1}{8}Ik_{42}(0)+\frac{1}{2}Ik_{11}(0)k_{31}(0)]+(z^3-3z)[\frac{1}{8}(k_{21}(\delta))^2+\frac{1}{6}k_{11}(\delta^2)k_{31}(0)-\frac{1}{3}I^{1/2}k_{21}(\delta)k_{31}(0)+\frac{10}{72}I(k_{31}(0))^2]$$

and the coefficient of z^3 in R is

(5.30) $$\frac{1}{2}\left[\frac{1}{2}k_{21}(\delta) - \frac{1}{3}I^{1/2}k_{31}(0)\right]^2$$

(E) the coefficient of n^{-1} in

(5.31) $$P_{\theta_n}(W_n \geq z + n^{-1/2}b_1 + n^{-1}c_1) + P_{\theta_n}(W_n \leq -z + n^{-1/2}b_2 + n^{-1}c_2)$$

is

$$2\phi(z - \delta I^{1/2})\,[d_2 - \delta I^{1/2}zd_2 + \delta^2(I^{1/2}zQ + R) + O(\delta^3)].$$

PROOF : (C) follows easily in view of (5.23), (A) of Lemma 5.1 and the fact that $k_{11}(\delta)$, $k_{21}(0)$, $k_{32}(0)$, $k_{12}(0)$, $k_{31}(\delta)$, $k_{22}(\delta)$ and $k_{31}(\delta^2)$ are all zero (vide Appendix 2); part (D) follows easily in view of (5.23) and the fact that $k_{31}(\delta)$, $k_{31}(\delta^2)$, $k_{21}(0)$, $k_{21}(\delta^2)$, $k_{32}(0)$ and $k_{32}(\delta^2)$ are all zero. Part (E) is straightforward $\square$.

LEMMA 5.3 : (F) The term free from z in Q is

$$k_{12}(\delta) - \frac{1}{2}I^{1/2}\,[k_{22}(0) + (k_{11}(0))^2] - \frac{1}{6}k_{32}(\delta)$$
$$- \frac{1}{2}k_{21}(\delta)k_{11}(0) + \frac{1}{8}I^{1/2}k_{42}(0) + \frac{1}{2}I^{1/2}k_{11}(0)k_{31}(0)$$
$$+ \frac{1}{4}k_{21}(\delta)k_{31}(0) - \frac{5}{24}I^{1/2}(k_{31}(0))^2$$

and (G) the coefficient, Γ, of z in R is

$$\Gamma = \frac{1}{2}k_{22}(\delta^2) - \frac{1}{3}I^{1/2}k_{32}(\delta) + \frac{1}{8}I\,k_{42}(0)$$
$$- L_{11}k_{11}(0)I^{-1/2} + \frac{3}{2}L_{11}d_1\,I^{-1/2} - \frac{1}{24}L_3^2\,I^{-2}.$$

PROOF : Straightforward.

Appendix 1. Constants appearing in (5.19)

Here primes stand for differentiation.

IR Test :

$$u_1 = -\frac{1}{2} I_n^{-1/2}, \quad u_2 = \frac{1}{6} \lambda_3^* I_n^{-1/2}, \quad u_3 = -\frac{1}{6} \lambda_3^* I_n^{-3/2}$$

$$v_1 = \frac{1}{2} I_n^{-3/2}, \quad v_2 = \frac{1}{6} \lambda_3^* I_n^{-5/2}, \quad y_1 = \frac{3}{8} I_n^{-5/2}$$

$$y_2 = \frac{5}{12} \lambda_3^* I_n^{-7/2}, \quad y_3 = \frac{1}{9} \lambda_3^{*2} I_n^{-9/2} + \frac{1}{24} \lambda_4^* I_n^{-7/2}$$

$$y_4 = \frac{1}{6} I_n^{-5/2}, \quad x_1 = -\frac{1}{4} \lambda_3^* I_n^{-5/2},$$

$$x_2 = -\frac{1}{12} \lambda_3^{*2} I_n^{-7/2} - \frac{1}{24} \lambda_4^* I_n^{-5/2}, \quad x_3 = -\frac{1}{6} I_n^{-3/2}$$

$$x_4 = \frac{1}{6} I_n^{-1/2}, \quad x_5 = \frac{1}{24} \lambda_4^* I_n^{-3/2} + \frac{1}{24} \lambda_3^{*2} I_n^{-5/2}$$

$$x_6 = -\frac{1}{24} \lambda_4^* I_n^{-1/2} - \frac{1}{72} \lambda_3^{*2} I_n^{-3/2}, \quad x_7 = -\frac{1}{8} I_n^{-3/2}$$

$$x_8 = \frac{1}{12} \lambda_3^* I_n^{-3/2}$$

Rao's Test :

$$u_1 = -I_n^{-1/2}, \quad u_2 = \frac{1}{2}(\lambda_3^* + I_n')I_n^{-1/2}, \quad u_3 = \frac{1}{2} I_n' I_n^{-3/2}$$

$$x_4 = \frac{1}{2} I_n^{-1/2}, \quad x_5 = -\frac{1}{4} I_n'' I_n^{-3/2} + \frac{3}{8} I_n'^2 I_n^{-5/2}$$

$$x_6 = -\frac{1}{6} \lambda_4^* I_n^{-1/2} + \frac{3}{8} I_n'^2 I_n^{-3/2} + \frac{1}{4} \lambda_3^* I_n' I_n^{-3/2} - \frac{1}{4} I_n'' I_n^{-1/2}$$

$$x_8 = -\frac{1}{2} I_n' I_n^{-3/2}$$

he rest are all zero.

ald's Test :

$$u_3 = \frac{1}{2} I_n' I_n^{-3/2}, \quad v_1 = I_n^{-3/2}, \quad v_2 = \frac{1}{2}(\lambda_3^* + I_n') I_n^{-5/2}$$

$$y_1 = I_n^{-5/2}, \quad y_2 = (\frac{3}{2}\lambda_3^* + I_n') I_n^{-7/2}$$

$$y_3 = (\frac{1}{6}\lambda_4^* + \frac{1}{4} I_n'') I_n^{-7/2} + (\frac{1}{2}\lambda_3^{*2} - \frac{1}{8} I_n'^2 + \frac{1}{2}\lambda_3^* I_n') I_n^{-9/2}$$

$$y_4 = \frac{1}{2} I_n^{-5/2}, \quad x_1 = \frac{1}{2} I_n' I_n^{-5/2}$$

$$x_2 = \frac{1}{4} I_n'' I_n^{-5/2} + (\frac{1}{4}\lambda_3^* I_n' - \frac{1}{8} I_n'^2) I_n^{-7/2}$$

e rest are all zero.

Appendix 2. Cumulants under θ_n of $W_n' - \delta I^{1/2}$

$$K_{n,1} = \delta(I_n^{1/2} - I^{1/2}) + n^{-1/2}(u_2\delta^2 + v_1\lambda_{11} + v_2\lambda_2)$$
$$+ n^{-1}\delta(x_1\lambda_{11} + x_2\lambda_2 + x_3\lambda_{101} + x_6\delta^2 + x_7\lambda_{02}) + o(n^{-1})$$

$$K_{n,2} = 1 + 2n^{-1/2}\,\delta I_n^{-1/2}(u_1\lambda_{11} + u_3\lambda_2)$$
$$+ n^{-1}\Big[2I_n^{-1/2}(v_1\lambda_{21} + v_2\lambda_3 + x_4\delta^2\lambda_{101} + x_5\delta^2\lambda_2$$
$$+ x_8\delta^2\lambda_{11} + y_1(\lambda_2\lambda_{02} + 2\lambda_{11}^2) + 3y_2\lambda_2\lambda_{11} + 3y_3\lambda_2^2$$
$$+ 3y_4\lambda_2\lambda_{101}) + u_1^2\delta^2\lambda_{02} + u_3^2\delta^2\lambda_2 + v_1^2(\lambda_2\lambda_{02} + \lambda_{11}^2)$$
$$+ 2v_2^2\lambda_2^2 + 2u_1u_3\delta^2\lambda_{11} + 4v_1v_2\lambda_{11}\lambda_2\Big] + o(n^{-1})$$

$$K_{n,3} = n^{-1/2}(\lambda_3 I_n^{-3/2} + 6v_1\lambda_{11} + 6v_2\lambda_2)$$
$$+ 3\delta n^{-1}\Big[I_n^{-1}(u_1\lambda_{21} + u_3\lambda_3 + 2x_1\lambda_{11}\lambda_2 + 2x_2\lambda_2^2 + 2x_3\lambda_2\lambda_{101}$$
$$+ 2x_7\lambda_{11}^2) + I_n^{-1/2}(2u_1v_1(\lambda_2\lambda_{02} + \lambda_{11}^2) + 4u_1v_2\lambda_2\lambda_{11}$$
$$+ 4u_3v_2\lambda_2^2 + 4u_3v_1\lambda_2\lambda_{11})\Big] + o(n^{-1})$$

$$K_{n,4} = n^{-1}\Big[(\lambda_4 - 3\lambda_2^2)I_n^{-2} + 12v_1(\lambda_{21} + \lambda_{11}\lambda_3\lambda_2^{-1})\, I_n^{-1/2}$$
$$+ 24v_2\lambda_3 I_n^{-1/2} + 12v_1^2(3\lambda_{11}^2 + \lambda_{02}\lambda_2) + 48v_2^2\lambda_2^2 + 96v_1v_2\lambda_{11}\lambda_2$$
$$+ 24(y_1\lambda_{11}^2 + y_2\lambda_2\lambda_{11} + y_3\lambda_2^2 + y_4\lambda_2\lambda_{101})\, I_n^{-1/2}\Big] + o(n^{-1})$$

To get these expressions, the following moments about θ_n are needed (all equalities are correct up to $o(n^{-1})$ unless otherwise stated) :

$$E\Delta_1^2 = \lambda_2 \qquad E\Delta_1\Delta_2 = \lambda_{11}$$

$$E\Delta_1^2\Delta_2 = n^{-1/2}\lambda_{21} \qquad E\Delta_1^3 = n^{-1/2}\lambda_3$$

$$E\Delta_1^3\Delta_2 = 3\lambda_2\lambda_{11} + O(n^{-1})$$

$$E\Delta_1^4 = 3\lambda_2^2 + n^{-1}(\lambda_4 - 3\lambda_2^2)$$

$$E\Delta_1^2\Delta_2^2 = \lambda_2\lambda_{02} + 2\lambda_{11}^2 + O(n^{-1})$$

$$E\Delta_1^3\Delta_2^2 = n^{-1/2}(3\lambda_2\lambda_{12} + 3\lambda_{21}\lambda_{11} + \lambda_3\lambda_{02})$$

$$E\Delta_1^4\Delta_2 = n^{-1/2}(4\lambda_3\lambda_{11} + 6\lambda_2\lambda_{21})$$

$$E\Delta_1^4\Delta_2^2 = 3\lambda_2^2\lambda_{02} + 12\lambda_2\lambda_{11}^2 + O(n^{-1}) \qquad E\Delta_1^5 = 10n^{-1/2}\lambda_3\lambda_2$$

$$E\Delta_1^5\Delta_2 = 15\lambda_2^2\lambda_{11} + O(n^{-1}) \qquad E\Delta_1^6 = 15\lambda_2^3 + O(n^{-1})$$

Appendix 3 . Below are the values of some of the " approximate cummulants " of W_n^i

IR Test :

$$k_{11}^1(0) = -\frac{1}{6} I^{-3/2} L_3 , \quad k_{11}^1(\delta^2) = \frac{1}{6} (3L_{11} + 2L_3) I^{-1/2} ,$$

$$k_{12}^1(\delta) = -\frac{1}{8} I^{-3/2}(2L_{21} + L_4 - I^2) + \frac{1}{12} I^{-5/2} L_3 (3L_{11} + 2L_3)$$

$$k_{21}^1(\delta) = \frac{1}{3} I^{-1} L_3 ,$$

$$k_{22}^1(\delta^2) = \frac{1}{4} I^{-1}(2L_{21} + L_4 - I^2) - \frac{1}{9} I^{-2} L_3(3L_{11} + 2L_3),$$

$$k_{22}^1(0) = \frac{1}{4}(L_{02} - 2L_{21} - L_4)I^{-2} + \frac{1}{36}(14L_3^2 + 18L_{11}L_3 - 9L_{11}^2)I^{-3},$$

$$k_{31}^1(0) = 0 = k_{42}^1(0)$$

$$k_{32}^1(\delta) = \frac{1}{4} I^{-3/2}(L_4 - 3L_{02}) + \frac{1}{12} I^{-5/2}(9L_{11}^2 - 4L_3^2)$$

Rao's Test :

$$k_{11}^2(0) = 0, \quad k_{11}^2(\delta^2) = \frac{1}{2} (L_{11} + L_3) I^{-1/2} ,$$

$$k_{12}^2(\delta) = 0, \quad k_{21}^2(\delta) = I^{-1} L_3 , \quad k_{22}^2(\delta^2) = \frac{1}{2}(L_{21} + L_4 - 2I^2)I^{-1},$$

$$k_{22}^2(0) = 0, \quad k_{31}^2(0) = I^{-3/2} L_3, \quad k_{32}^2(\delta) = I^{-3/2}(L_4 - 3I^2),$$

$$k_{42}^2(0) = I^{-2}(L_4 - 3I^2) .$$

Wald's Test :

$$k_{11}^3(0) = \frac{1}{2} I^{-3/2} L_{11}, \quad k_{11}^3(\delta^2) = \frac{1}{2}(2L_{11} + L_3)I^{-1/2} ,$$

$$k^3_{12}(\delta) = \frac{1}{4}(7L_{21} + 4L_{101} + 4L_{02} + L_4)\, I^{-3/2}$$

$$- \frac{1}{8}\,(2L_{11} + L_3)(10L_{11} + 3L_3)\, I^{-5/2}$$

$$k^3_{21}(\delta) = I^{-1}(2L_{11} + L_3),$$

$$k^3_{22}(\delta^2) = (2L_{02} + 2L_{101} + 5L_{21} + L_4)I^{-1} - \frac{3}{4}\,(2L_{11} + L_3)^2 I^{-2}$$

$$k^3_{22}(0) = \frac{1}{2}\,(L_4 + 7L_{21} + 4L_{101} + 6L_{02} - 2I^2)I^{-2}$$

$$- \frac{1}{4}\,(3L_3^2 + 16L_{11}L_3 + 22L_{11}^2)I^{-3},$$

$$k^3_{31}(0) = I^{-3/2}(L_3 + 3L_{11}),$$

$$k^3_{32}(\delta) = I^{-3/2}(\frac{27}{2}\,L_{21} + 6L_{101} + 6L_{02} + \frac{5}{2}\,L_4) - \frac{9}{4}\,I^{-5/2}(2L_{11} + L_3)^2$$

$$k^3_{42}(0) = I^{-2}(L_4 - 3I^2 + 18L_{21} + 12L_{02} + 2L_4 + 8L_{101})$$

$$- 3\,I^{-3}(2L_{11} + L_3)^2.$$

CHAPTER SIX

A PROPERTY OF MAXIMUM LIKELIHOOD ESTIMATE

6.1 INTRODUCTION : Optimality or efficiency of the m.l.e. in a class of estimates is usually defined via comparision of its asymptotic variance with that of rival estimates. This approach originates in the work of Fisher and has been developed by LeCam, Rao, Bahadur and others; in recent works the variance is replaced by the probability of complement of symmetric intervals around the true value and other loss functions (vide Weiss and Wolfowitz [1974]). In a closely related approach Wilks had studied (see e.g. Wilks [1961] pp 374) the asymptotic lengths of confidence intervals arising from different methods of estimation. In this chapter we define a different kind of optimal property using unbaised confidence sets and show that the m.l.e. enjoys this property.

Let $X_1, X_2, \ldots$ be a sequence of iid r.v.'s with common d.f. $F(x,\theta)$, $\theta \in \textcircled{H}$; $\textcircled{H}$ an open subset of $\mathbb{R}$. Let $f(x,\theta)$ be the density of $F(x,\theta)$ w.r.t. some dominating measure μ. Under certain regularity conditions on $f(x,\theta)$ we prove that with probability tending to one, the m.l.e. $\hat{\theta}_n$ lies in the $100(1-\alpha)$ % $(0<\alpha<1)$ confidence set V_n determined by the locally most powerful unbaised tests (LMPU tests) of $H_o : \theta = \theta_o$ vs $H_1 : \theta \neq \theta_o$. A sort of converse is also proved that is, if a consistent estimate T_n is such that $P_{\theta_o}(T_n \in V_n) \longrightarrow 1$ then $n^{1/2}(T_n - \hat{\theta}_n) \longrightarrow 0$ in P_{θ_o}; also if T_n is such that $n^{1/2}(T_n - \hat{\theta}_n) \longrightarrow 0$ in P_{θ_o} then $P_{\theta_o}(T_n \in V_n) \longrightarrow 1$. (vide Theorem 6.1). Our regularity conditions are given in Section 3. Section 4 gives some preliminary Lemmas. This chapter is based on Ghosh, Sinha and Joshi [1980].

6.2 MAIN RESULT : Assumptions I to VI guarantee (see Lehmann, [1959] p. 83) the existence of a IMPU test of $H_o : \theta = \theta_o$ vs. $H_1 : \theta \neq \theta_o$ with critical function

$$\phi_{\theta_o} = \begin{cases} 1 & \text{if } W_{n\theta_o} + Z^2_{n\theta_o} > K_{1n\theta_o} + K_{2n\theta_o} Z_{n\theta_o} \\ 0 & \text{if } \quad " \quad < \quad " \\ \text{arbitrary} & \text{if } \quad " \quad = \quad " \end{cases}$$

where

$$Z_{n\theta_o} = n^{-1/2} I^{-1/2}(\theta_o) \sum_1^n \frac{d}{d\theta} \log f(x_i , \theta_o) ,$$

$$W_{n\theta_o} = n^{-1} I^{-1} (\theta_o) \sum_1^n \frac{d^2}{d\theta^2} \log f(x_i , \theta_o) ,$$

and

$K_{1n\theta_o}$ and $K_{2n\theta_o}$ are such that

$$E_{\theta_o} (\phi_{\theta_o}) = \alpha \quad \text{and} \quad E_{\theta_o} (\phi_{\theta_o} Z_{n\theta_o}) = 0 .$$

Let V_n be the randomized confidence set arising from this family of tests i.e. it consists of all θ accepted by the test ϕ_θ. The set V_n will depend on the randomising device in addition to $X_1 , X_2, \ldots, X_n$ but will contain

$$w_n = \left\{ \theta : W_{n\theta} + Z^2_{n\theta} < K_{1n\theta} + K_{2n\theta} Z_{n\theta} \right\} .$$

Similarly

$$V_n \subset \left\{ \theta : W_{n\theta} + Z^2_{n\theta} \leq K_{1n\theta} + K_{2n\theta} Z_{n\theta} \right\} = w'_n \quad \text{(say)} .$$

Now we state our result. Let $\hat{\theta}_n$ denote the m.l.e.

THEOREM 6.1 : Under assumptions I to VI

(a) For every $\theta_o \varepsilon (\underline{H})$ and for every $0 < \alpha < 1$,

$P_{\theta_o} (\hat{\theta}_n \varepsilon V_n) \longrightarrow 1$ as $n \longrightarrow \infty$.

(b) Let T_n be any other estimate of θ; then for $\theta_o \varepsilon (\underline{H})$,

$P_{\theta_o} (T_n \varepsilon V_n) \longrightarrow 1$ for every $0 < \alpha < 1$ if $\sqrt{n}(\hat{\theta}_n - T_n) \xrightarrow{P_{\theta_o}} 0$.

(c) Let T_n be any consistent estimate of θ such that for $\theta_o \varepsilon (\underline{H})$,

$P_{\theta_o} (T_n \varepsilon V_n) \longrightarrow 1$ for every $0 < \alpha < 1$. Then $\sqrt{n}(\hat{\theta}_n - T_n) \xrightarrow{P_{\theta_o}} 0$.

<u>PROOF</u> :

(6.1a) Fix a $\theta_o \varepsilon (\underline{H})$ and a bounded open set Ω containing θ_o, $\Omega \varepsilon (\underline{H})$ such that assumptions III and IV hold on the closure of Ω.

Choose $\delta > 0$ such that $(|\theta_o - \theta| < \delta) \subset \Omega$. By assumption VI for any $\eta > 0 \; \exists \; n_o$ such that

(6.1) $$P_{\theta_o} (|\hat{\theta}_n - \theta_o| < \delta, Z_{n,\hat{\theta}_n} = 0) \geq 1 - \eta , n \geq n_o .$$

Choose $\varepsilon > 0$ such that $\xi^2_{\alpha/2} > 2\varepsilon$ (see Lemma 6.4 for the definition of ξ_α) and use Lemma 6.4 to get n_o such that

(6.2) $$K_{1n\theta} > \xi^2_{\alpha/2} - 1 - \varepsilon > - 1 + \varepsilon , n \geq n_o , \forall \theta \varepsilon \Omega .$$

Using Lemma 1.1 we get (as applied to $W_{n\theta}$) η such that

(6.3) $$P_{\theta_o}\left[\sup_{\theta \varepsilon \underline{\Omega}} W_{n\theta} \leq -1+\varepsilon\right] \geq 1-\eta, \forall n \geq n_o .$$

Note that on $\{Z_{n,\hat{\theta}_n} = 0\}$

(6.4) $$\hat{\theta}_n \varepsilon w_n \text{ iff } W_{n\hat{\theta}_n} < K_{1n\hat{\theta}_n} .$$

Combining (6.1), (6.2), (6.3) and (6.4) we get part (a) of the theorem.

Under hypothesis of part (b) or (c) we have

(6.5) $$P_{\theta_o}[T_n \varepsilon \underline{\Omega}] \longrightarrow 1 .$$

This along with Lemmas 6.1 and 6.2 implies

(6.6) $$K_{2nT_n} \xrightarrow{P_{\theta_o}} 0 , \quad K_{1nT_n} \xrightarrow{P_{\theta_o}} \xi^2_{\alpha/2} - 1 \quad \text{and} \quad W_{nT_n} \xrightarrow{P_{\theta_o}} 1 ,$$

so that

(6.7) $$K_{1nT_n} - W_{nT_n} + \frac{1}{4} K^2_{2nT_n} \xrightarrow{P_{\theta_o}} \xi^2_{\alpha/2} > 0, \; 0 < \alpha < 1 .$$

Now expanding Z_{nT_n} around $\hat{\theta}_n$ on the set $\{Z_{n\hat{\theta}_n} = 0\}$, we have

$$(T_n \varepsilon w_n) = \left\{W_{nT_n} + Z^2_{nT_n} < K_{1nT_n} + K_{2nT_n} Z_{nT_n}\right\}$$

$$= \Big[\Big\{\sqrt{n}(T_n - \hat{\theta}_n) \cdot \frac{1}{n} \cdot \Sigma \frac{d^2}{d\theta^2} \log f(x_i , \theta^*_n)\Big\}^2 I^{-1} (T_n)$$

$$- K_{2nT_n}\Big\{\sqrt{n}(T_n - \hat{\theta}_n) \cdot \frac{1}{n} \cdot \Sigma \frac{d^2}{d\theta^2} \log f(x_i , \theta^*_n)\Big\} I^{-1/2}(T_n)$$

$$+ W_{nT_n} - K_{1nT_n} < 0\Big]$$ where θ^*_n is between $\hat{\theta}_n$ and T_n

$$= \Big\{ [\sqrt{n}(T_n - \hat{\theta}_n) - \tfrac{1}{2} K_{2nT_n} W^{-1}_{n\theta^*_n} I^{-1}(\theta^*_n) I^{1/2}(T_n)]^2$$

$$< W^{-2}_{n\theta^*_n} I^{-2}(\theta^*_n) I(T_n) \Big[K_{1nT_n} - W_{nT_n} + \frac{K^2_{2nT_n}}{4}\Big]\Big\}.$$

Observe that by (6.5) and Lemma 1.1,

(6.8) $$W_{n\theta^*_n} I^{-2}(\theta^*_n) I(T_n)$$

is bounded away from zero and infinity with probability tending to 1.

Let

(6.9) $$\Sigma_n = \Big\{ W^{-1}_{n\theta^*_n} I^{-1}(\theta^*_n) I^{1/2}(T_n) \Big[\frac{K_{2nT_n}}{2} - (K_{1nT_n} - W_{nT_n} + \frac{K^2_{2nT_n}}{4})^{1/2}\Big]$$

$$< \sqrt{n}(T_n - \hat{\theta}_n) < W^{-1}_{n\theta^*_n} I^{-1}(\theta^*_n) I^{1/2}(T_n) \Big[\frac{K_{2nT_n}}{2} + (K_{1nT_n} - W_{nT_n}$$

$$+ \frac{K^2_{2nT_n}}{4})^{1/2}\Big]\Big\} \text{ and } K_{1nT_n} - W_{nT_n} + \frac{K^2_{2nT_n}}{4} > 0 .$$

et Σ'_n denote the set with strict inequalities replaced by " $\leq$ "
n (6.9).

If $P_{\theta_0}(T_n \varepsilon w'_n) \longrightarrow 1$ for each $0 < \alpha < 1$ then, it is
lear from (6.7) that $P_{\theta_0}[\Sigma'_n] \longrightarrow 1$. This in view (6.6), (6.7),
5.8) and the fact $\xi_{\alpha/2} \longrightarrow 0$ as $\alpha \longrightarrow 1$ gives us part (c) of
ie theorem.

Finally, if $\sqrt{n}(\hat{\theta}_n - T_n) \xrightarrow{P_{\theta_0}} 0$, clearly $P_{\theta_0}(\Sigma_n) \longrightarrow 1$ for
ery $0 < \alpha < 1$ and hence $P_{\theta_0}(T_n \varepsilon w_n) \longrightarrow 1$ for every $0 < \alpha < 1$
ich gives part (b) of the theorem □

since consistency of T_n was used only to derive (6.5), we have the following

<u>COROLLARY 6.1</u> : Suppose $P_{\theta_o}(T_n \varepsilon C_{\theta_o}) \longrightarrow 1$ and $P_{\theta_o}(T_n \varepsilon V_n) \longrightarrow 1$, for all $\theta_o \varepsilon (\underline{H})$. Then $\sqrt{n}(T_n - \hat{\theta}_n) \xrightarrow{P_{\theta_o}} 0$, $\forall\, \theta_o \varepsilon (\underline{H})$.

<u>REMARK 6.1</u>: If instead of the randomized confidence set V_n one of the nonrandomized confidence sets w_n or w_n' be used, the resultant size of the test will be $\alpha_n(\theta)$ which will eventually be α as $n \longrightarrow \infty$ for every $\theta \varepsilon (\underline{H})$ (vide proof of Lemma 6.2.) The theorem remains true if V_n is replaced by w_n or w_n' throughout. This is so because the proof of (a) and (b) uses $(\hat{\theta}_n \varepsilon w_n)$ and the proof of (c) uses $(\hat{\theta}_n \varepsilon w_n')$.

6.3 <u>ASSUMPTIONS</u> :

Assumption I : For each $x, f(x,\theta)$ is twice continuously differentiable in $\theta \varepsilon (\underline{H})$.

Assumption II : Let

$$I(\theta) = E_\theta\left[-\frac{d^2}{d\theta^2}\log f(x\,,\theta)\right];$$

then $0 < I(\theta) < \infty$ for $\theta \varepsilon (\underline{H})$, and $I(\theta)$ is continuous in $\theta \varepsilon (\underline{H})$.

Assumption III : For every $\theta_o \varepsilon (\underline{H})$, $\exists$ a nhbd C_{θ_o} of θ_o such that

$$\sup_{\theta \varepsilon C_{\theta_o}} E_\theta\left|\frac{d}{d\theta}\log f(X\,,\theta)\right|^3 < \infty .$$

Assumption IV : For every $\theta_o \varepsilon (\underline{H})$, $\exists$ a nhbd C_{θ_o} of θ_o such that

$$\left|\frac{d^2}{d\theta^2} \log f(x,\theta)\right| \leq H(x), \ \forall \ \theta \ \varepsilon \ C_{\theta_o} ,$$

$$\left|\frac{d^2}{d\theta^2} \log f(x,\theta) - \frac{d^2}{d\theta^2} \log f(x,\theta')\right| \leq |\theta - \theta'| A(x)$$

for $\forall \ \theta,\theta' \ \varepsilon \ C_{\theta_o}$; and for some $\delta > 0$

$$\sup_{\theta \varepsilon C_{\theta_o}} E_\theta H^{2+\delta}(X) < \infty , \quad \sup_{\theta \varepsilon C_{\theta_o}} E_\theta A(X) < \infty .$$

Assumption V : If ϕ_n is any test function based on n observations then $E_\theta \ \phi_n$ is twice continuously differentiable in $\theta \ \varepsilon \ (\underline{H})$; moreover for every $\theta \ \varepsilon \ (\underline{H})$ and $n \geq 1$

$$\frac{d}{d\theta} E_\theta \ \phi_n(\underline{X}_n) = \int \phi_n(\underline{x}_n) \frac{d}{d\theta} \prod_{i=1}^{n} f(x_i , \theta) d\mu(\underline{x}_n) \quad \text{and}$$

$$\frac{d^2}{d\theta^2} E_\theta \ \phi_n(\underline{X}_n) = \int \phi_n(\underline{x}_n) \frac{d^2}{d\theta^2} \prod_{i=1}^{n} f(x_i,\theta) d\mu(\underline{x}_n)$$

<u>Assumption VI</u> : The m.l.e $\hat{\theta}_n$ of θ exists and for every $\theta_o \ \varepsilon \ (\underline{H})$ and $\varepsilon > 0$

$$P_{\theta_o} \left[|\theta_n - \theta_o| < \varepsilon , \frac{d}{d\theta} \log \prod_{i=1}^{n} f(x_i , \hat{\theta}_n) = 0 \right] \longrightarrow 1$$

as $n \longrightarrow \infty$.

<u>REMARK 6.2</u> : VI holds if conditions of Wald [1949] or Bahadur ([1971], p. 34) hold.

6.4 <u>SOME LEMMAS</u> : We quote a version of Theorem 3 of Michel [1976] which will be needed in the sequel.

<u>LEMMA 6.1</u> : Let $X_1, X_2, \ldots,$ be a sequence of i.i.d. r.v's having a common d.f. F_θ, $\theta \varepsilon$ $(\underline{H})$ such that $E_\theta(X_1) = 0$, $E_\theta(X_1^2) = 1$. If for some $\delta > 0$

$$\sup_{\theta \varepsilon (\underline{H})} E_\theta |X_1|^{2+\delta} < \infty$$

then there exists a constant M such that for $n \geq 1$, $\forall \theta \varepsilon (\underline{H})$ and for all $t \varepsilon R$,

$$|F_{n\theta}(t) - \Phi(t)| \leq M n^{-\delta^*} [1 + |t|^{2+\delta}]^{-1}$$

where $F_{n\theta}$ is the d.f. of $n^{-1/2} \sum_1^n X_i$ under F_θ and $\delta^* = \frac{1}{2} \min(\delta, 1)$.

<u>NOTE</u> : The assumptions II and III enable us to apply Lemma 6.1 to the d.f. of $Z_{n\theta}$ and assumptions II and IV enable us to apply Lemma 1.1 to $W_{n\theta}$.

Let $R_{n\theta} = \{\underline{X}_n : Z_{n\theta}^2 + W_{n\theta} > K_{1n\theta} + K_{2n\theta} Z_{n\theta}\}$,

$R'_{n\theta} = \{\underline{X}_n : Z_{n\theta}^2 + W_{n\theta} = K_{1n\theta} + K_{2n\theta} Z_{n\theta}\}$,

$\widetilde{R}_{n\theta} = \{\underline{X}_n : Z_{n\theta}^2 \geq K_{1n\theta} + K_{2n\theta} Z_{n\theta} + 1\}$

and $A \triangle B = (A^c \cap B) \cup (A \cap B^c)$ for any two sets A and B.

We fix a $\theta_0 \varepsilon (\underline{H})$ and let $\underline{(\ )}$ be as in (6.1a).

<u>LEMMA 6.2</u> : Uniformly in $\theta \varepsilon \underline{(\ )}$

(6.10) $$P_\theta(\widetilde{R}_{n\theta}) \longrightarrow \alpha \quad \text{and}$$

(6.11) $$E_\theta(I_{\widetilde{R}_{n\theta}} Z_{n\theta}) \longrightarrow 0.$$

PROOF : Note

$$|E_\theta [I_{\hat{R}_{n\theta}} Z_{n\theta}] - E_\theta [\phi_\theta Z_{n\theta}]| \leq P_\theta^{1/2}(R_{n\theta} \Delta \hat{R}_{n\theta}) + P_\theta^{1/2}(R'_{n\theta})$$

and $|P_\theta(\hat{R}_{n\theta}) - E_\theta \phi_\theta| \leq P_\theta(R_{n\theta} \Delta \hat{R}_{n\theta}) + P_\theta(R'_{n\theta})$.

Hence (6.10) and (6.11) are proved if we prove uniformly in

$$\theta \in \Omega,\ P_\theta(R_{n\theta} \Delta \hat{R}_{n\theta}) \longrightarrow 0 \quad \text{and} \quad P_\theta(R'_{n\theta}) \longrightarrow 0 .$$

In view of assumptions II and IV it is clear that for every $\varepsilon > 0$

(6.12) $\quad P_\theta(|W_{n\theta} + 1| \leq \varepsilon) \longrightarrow 1$ uniformly in $\theta \in \Omega$.

Let $$A_{n,\varepsilon,\theta} = \Big\{ \underline{X}_n : \frac{K_{2n\theta}^2}{4} + K_{1n\theta} + 1 - \varepsilon \leq (Z_{n\theta} - \frac{K_{2n\theta}}{2})^2 \leq \frac{K_{2n\theta}^2}{4} + K_{1n\theta} + 1 + \varepsilon \Big\}$$ then

(6.13) $$A_{n,\varepsilon,\theta} \supset (R_{n\theta} \Delta \hat{R}_{n\theta}) \cap (|W_{n\theta} + 1| \leq \varepsilon).$$

Also, $$A_{n,\varepsilon,\theta} \supset (|W_{n\theta} + 1| \leq \varepsilon) \cap R'_{n\theta}.$$

On the other hand

(6.14) $$A_{n,\varepsilon,\theta} \subset \Big\{ \underline{X}_n : (Z_{n\theta} - \frac{K_{2n\theta}}{2})^2 \leq 2\varepsilon \Big\} \text{ if } \frac{K_{2n\theta}^2}{4} + K_{1n\theta} + 1 \leq \varepsilon$$

and $$A_{n,\varepsilon,\theta} = \Big\{ \underline{X}_n : (x - \varepsilon)^{1/2} + \frac{K_{2n\theta}}{2} \leq Z_{n\theta} \leq \frac{K_{2n\theta}}{2} + (x + \varepsilon)^{1/2} \Big\}$$

$$\cup \Big\{ \underline{X}_n : \frac{K_{2n\theta}}{2} - (x+\varepsilon)^{1/2} \leq Z_{n\theta} \leq \frac{K_{2n\theta}}{2} - (x - \varepsilon)^{1/2} \Big\}$$

if $x = \frac{K_{2n\theta}^2}{4} + K_{1n\theta} + 1 \geq \varepsilon$.

Now using the Berry-Essen theorem for $Z_{n\theta}$ along with assumption III we get for any $\varepsilon' > 0 \ \exists \ \varepsilon > 0$ such that

$P_\theta(A_{n,\varepsilon,\theta}) < \varepsilon'$ uniformly in $\theta \in \Omega$; this along with (6.12) and (6.13) completes the proof of the lemma □

Since $\alpha < 1$, from Lemma 6.1 and (6.10) it is clear that $\exists \ n_o$ such that

$$\frac{K_{2n\theta}^2}{4} + K_{1n\theta} + 1 \geq 0 \ \forall \ n \geq n_o , \forall \ \theta \in \Omega .$$

Let for $n \geq n_o$

$$C_{1n\theta} = \frac{K_{2n\theta}}{2} + \left(\frac{1}{4} K_{2n\theta}^2 + K_{1n\theta} + 1 \right)^{1/2} \quad \text{and}$$

$$C_{2n\theta} = \frac{1}{2} K_{2n\theta} - \left(\frac{1}{4} K_{2n\theta}^2 + K_{1n\theta} + 1 \right)^{1/2}.$$

LEMMA 6.3 : Let Z be normal with zero mean and unit variance. Then uniformly in $\theta \in \Omega$

(6.15) $$E\left[I_{(C_{2n\theta} \leq Z \leq C_{1n\theta})}\right] \longrightarrow 1 - \alpha \quad \text{and}$$

(6.16) $$E\left[I_{(C_{2n\theta} \leq Z \leq C_{1n\theta})} Z\right] \longrightarrow 0 .$$

PROOF : Note for $n \geq n_o$, $\hat{R}_{n\theta} = (Z_{n\theta} \leq C_{2n\theta}) \bigcup (Z_{n\theta} \geq C_{1n\theta})$ and hence $P_\theta(C_{2n\theta} \leq Z_{n\theta} \leq C_{1n\theta}) \longrightarrow 1 - \alpha$ uniformly in $\theta \in \Omega$ by (6.10). The Berry-Essen theorem along with assumption III completes the proof of (6.15).

Let $F_{n\theta}$ be the d.f. of $Z_{n\theta}$. For a d.f. $F(z)$ we have

$$\int_{C_{2n\theta}}^{C_{1n\theta}} z\, dF(z) = C_{1n\theta}F(C_{1n\theta}) - C_{2n\theta}F(C_{2n\theta}) - \int_{C_{2n\theta}}^{C_{1n\theta}} F(z)dz .$$

Hence we have, with $\Phi(z) = P(Z \le z)$,

$$(6.17) \quad \left| \int_{C_{2n\theta}}^{C_{1n\theta}} z\, dF_{n\theta}(z) - \int_{C_{2n\theta}}^{C_{1n\theta}} z\, d\,\Phi(z) \right|$$

$$\le |C_{1n\theta}|\,|F_{n\theta}(C_{1n\theta}) - \Phi(C_{1n\theta})| + |C_{2n\theta}|\,|F_{n\theta}(C_{2n\theta}) - \Phi(C_{2n\theta})|$$

$$+ \int_{C_{2n\theta}}^{C_{1n\theta}} |F_{n\theta}(z) - \Phi(z)|dz .$$

Lemma 6.1 applied to $F_{n\theta}$ imples,

$$\text{R.H.S. of } (6.17) \le bn^{-1/2}\Big\{ |C_{1n\theta}|(1 + |C_{1n\theta}|^3)^{-1} + |C_{2n\theta}|\cdot$$

$$(1 + |C_{2n\theta}|^3)^{-1} + \int_{C_{2n\theta}}^{C_{1n\theta}} (1 + |t|^3)^{-1} dt \Big\} \text{ for some } b > 0 .$$

This in view of (6.11) complete the proof of the lemma □

<u>LEMMA 6.4</u> : $K_{2n\theta} \longrightarrow 0$ and $K_{1n\theta} \longrightarrow \xi^2_{\alpha/2} - 1$ both uniformly in $\theta \in \Omega$, where $\Phi(\xi_\alpha) = 1 - \alpha$.

<u>PROOF</u> : Note that $0 < \alpha < 1$ and (6.15) imply existence of n_o, $0 < M < \infty$ and $\delta > 0$ such that

$$(6.18) \quad C_{1n\theta} - C_{2n\theta} > \delta, \ \forall\, n \ge n_o, \ \forall\, \theta \in \Omega$$

(6.19) $\quad \min(|C_{1n\theta}|, |C_{2n\theta}|) < M, \ \forall\, n \geq n_o, \ \forall\, \theta \in \Omega.$

Hence $\exists\ 0 < M' < \infty$ such that

(6.20) $\quad \max(e^{-C^2_{1n\theta}/2}, e^{-C^2_{2n\theta}/2}) > M', \ \forall\, n \geq n_o, \ \forall\, \theta \in \Omega.$

Using (6.16) we get

(6.21) $$\left| \int_{C_{2n\theta}}^{C_{1n\theta}} z . e^{-z^2/2} dz \right| = e^{-C^2_{1n\theta}/2} \left| 1 - e^{(C_{1n\theta} + C_{2n\theta})(C_{1n\theta} - C_{2n\theta})/2} \right|$$

$$= e^{-C^2_{2n\theta}/2} \left| 1 - e^{-(C_{1n\theta} + C_{2n\theta})(C_{1n\theta} - C_{2n\theta})/2} \right|$$

$$\longrightarrow 0 \text{ uniformly in } \theta \in \Omega.$$

(6.21) along with (6.20) implies

(6.22) $\quad K_{2n\theta} = (C_{1n\theta} + C_{2n\theta}) \longrightarrow 0$ uniformly in $\theta \in \Omega$.

To prove the second part, note that if for every $n \ \exists\ M > n$ and $\theta' \in \Omega$ such that $C_{2n\theta'}$ and $C_{1n\theta'}$ are on the same side of zero then we get a contradiction to (6.16) using (6.18) and (6.19). Hence $\exists\ n_o$ such that $C_{2n\theta} \leq 0 \leq C_{1n\theta}, \ \forall\, n \geq n_o, \ \forall\, \theta \in \Omega$.

Now, $$2\int_0^{C_{1n\theta}} d\,\Phi(z) = \left(\int_0^{C_{1n\theta}} d\,\Phi(z) + \int_0^{-C_{2n\theta}} d\,\Phi(z) \right)$$

$$+ \left(\int_0^{C_{1n\theta}} d\,\Phi(z) - \int_0^{-C_{2n\theta}} d\,\Phi(z) \right) \longrightarrow 1 - \alpha \text{ uniformly in } \theta \in \Omega$$

because of (6.15) and (6.22); hence in view of (6.22) we have

$$K_{1n\theta} \longrightarrow \xi^2_{\alpha/2} - 1 \text{ uniformly in } \theta \in \Omega \quad \square$$

REFERENCES

Akahira, M. and Takeuchi, K. [1976]. On the second order asymptotic efficiencies of estimators. Proceedings of the Third Japan-USSR Symposium on Probability Theory (G. Naruyama and J.V. Prokhorov, Eds.), 604-638. Lecture Notes in Mathematics 550, Springer-Verlag, Berlin.

Alvo, M. [1977]. Bayesian sequential estimation. Ann. Statist., 5, 955-968.

Bahadur, R.R. [1971]. Some Limit Theorems in Statistics. SIAM, Philadelphia.

Bhattacharya, R.N. and Ghosh, J.K. [1978]. On the validity of the formal Edgeworth expansions. Ann. Statist., 6, 434-451.

Burnasev, M.V. [1979]. Asymptotic expansions of the integral risk of statistical estimators of location parameter in a scheme of independent observations. Soviet Math. Dokl. Vol.20, No.4, 788-791. (English translation).

Burnasev, M.V. [1981]. Investigation of second order properties of statistical estimators in the scheme of independent observations. (in Russian) Izvestiya Akademii Nauk USSR Ser. Matemat. 45, 509-539.

Chandra, T.K. [1980]. Asymptotic Expansions and Deficiency. Doctoral Thesis, Indian Statistical Institute, Calcutta.

Chandra, T.K. and Ghosh, J.K [1979]. Valid asymptotic expansions for the likelihood ratio statistic and other perturbed chi-square variables, Sankhyā, Ser. A, 41, 22-47.

________ [1980]. Valid asymptotic expansions for the likelihood ratio and other statistics under contiguous alternatives, Sankhya, Ser. A, 42, 174-184.

Chandra, T. and Joshi, S.N. [1982]. Comparison of the likelihood ratio, Rao's and Walds tests and a conjecture of C.R. Rao. (Submitted to Sankhyā)

Efron, B. [1975]. Defining the curvature of a statistical problem (with applications to second order efficiency), Ann. Statist., 6, 1189-1242.

Ghosh, J.K., Sinha, B.K. and Joshi, S.N. [1980]. A property of maximum likelihood estimator. Sankhyā, Series B, 42, 143-152.

Ghosh, J.K., Sinha, B.K. and Joshi, S.N. [1981]. Expansion for posterior probability and integrated Bayes risk. (To appear in Proc of Purdue Symp. on Decision Theory, 1981).

Ghosh, J.K., Sinha, B.K. and Subramanyam, K. [1979]. Edgeworth expansions for Fisher-consistent estimators and second order efficiency. Calcutta Statist. Assoc. Bull., 28, 1-18.

Ghosh, J.K., Sinha, B.K. and Wieand, H.S. [1981]. Second order efficiency of the M.L.E. w.r.t. any bounded bowl-shaped loss function. Ann. Statist., 8, 506-521.

Ghosh, J.K. and Subramanyam, K. [1974]. Second order efficiency of maximum likelihood estimators. Sankhyā, Series A, 36, 325-358.

Gusev, S.I. [1975]. Asymptotic expansions associated with some statistical estimators in the smooth case. 1. Expansions of random variables. Theory Prob. Applications, 20, 470-498. (English translation)

———— [1976]. Asymptotic expansions associated with some statistical estimators in the smooth case II. Expansions of moments and distributions. Theory Prob. Applications, 21, 1, 14-32. (English translation)

Hodges, J.L. (jr.) and Lehmann, E.L. [1970]. Deficiency, Ann. Math. Statist., 41, 783-801.

Johason, R.A. [1967]. An asymptotic expansion for posterior distributions. Ann. Math. Statist., 38, 1899-1906.

———— [1970]. Asymptotic expansions associated with posterior distributions. Ann. Math. Statist. 41, 851-864.

Joshi, S.N. [1982]. Expansion of Bayes risk in the case of double exponential family. (Submitted to Sankhyā).

Lawley, D.N. [1956]. A general method for approximating to the distribution of the likelihood ratio criteria, Biometrika, 43, 295-308.

Lehmann, E.L. [1959]. Testing Statistical Hypotheses. Wiley, New York.

Michel, R. [1976]. Nonuniform central limit bounds with application to probabilities of deviations. Ann. Prob., 4, 102-106.

Michel, R. and Pfanzagl, J. [1971]. The accuracy of the normal approximation for minimum contrast estimates. Z. Wahr. and Berw. Gebiete, 18, 73-84.

Peers, H.W. [1971]. Likelihood ratio and associated test criteria, Biometrika, 58, 577-587.

Pfanzagl, J. [1973a]. Asymptotic expansions related to minimum contrast estimators. Ann. Statist., 1, 993-1026.

________ [1973b]. The accuracy of the normal approximation for estimates of vector parameters. Z. Wahr. and Verw. Gebiets, 25, 171-198.

________ [1975]. Discussions on Efron (1975) Ann. Statist., 6.

________ [1980]. Asymptotic expansions in parametric decision theory, Developments in Statistics, Edited by P.R. Krishnaiah, 3, 1-97.

Pfanzagl, J. and Wefelmeyer, W. [1978]. A third order optimum property of the maximum likelihood estimator. J. Multivariate Analysis, 8, 1-29.

Rao, C.R. [1962]. Efficient estimates and optimum inference procedures in large samples (with discussion). J. Roy. Statist. Sec. B, 24, 46-72.

________ [1965]. Linear Statistical Inference and Its Applications. Wiley, New York.

Reiss, R.D. [1976]. Asymptotic expansions for sample quantiles. Ann. Probability, 4. 249-258.

Serfling, R.J. [1980]. Approximation theorems of mathematical statistics. John Wiley.

Strasser, H.L. [1978]. Global asymptotic properties of risk functions in estimation. Z. Wahr. and Verw. Gebiete, 45, 35-48.

Wald, A. [1949]. Note on the consistency of the maximum likelihood estimate. Ann. Math. Statist., 20, 595-601.

Weiss, L. and Wolfowitz, J. [1974]. Maximum probability estimators and related topics. Lecture notes in mathematics, Springer-Verlog.

Wilks, S.S. [1961]. Mathematical statistics. Wiley, New York.