Available online at www.sciencedirect.com
SCIENCE DIRECT®
Statistics & Probability Letters 74 (2005) 212–220
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

A note on Moore's conjecture

Richard Lockhart[a,*], Federico O'Reilly[b]

[a]Department of Statistics, Simon Fraser University, Burnaby, BC, Canada V5A 1S6
[b]IIMAS-UNAM, México, D.F. México

Received 23 September 2004; received in revised form 10 December 2004; accepted 1 April 2005
Available online 27 June 2005

Abstract

We establish the conjecture of Moore [1973. A note on Srinivasan's goodness-of-fit test. Biometrika 60, 209–211] that the usual plug-in estimate of a distribution function and the Rao–Blackwell estimate of the distribution function are asymptotically equivalent for a wide class of exponential family distributions.

MSC: primary 62G20; 62G30

Keywords: Empirical distribution function; Goodness-of-fit; Local central limit theorem; Rao–Blackwell

1. Introduction

Let $X_1, \ldots, X_n$ be independent and identically distributed according to a distribution G which under a null hypothesis, H_0, is known to belong to the parametric family $\{F(\cdot;\theta); \theta \in \Theta\}$. Under H_0 let T_n be a minimal sufficient statistic for θ and let $\hat{\theta}_n$ be the maximum likelihood estimate (mle) of θ. By the plug-in estimate of the unknown cumulative distribution function (cdf) $F(\cdot;\theta)$ we mean $\hat{F}_n = F(\cdot;\hat{\theta}_n)$. The Rao–Blackwell estimate is $\tilde{F}_n$ given by

$$\tilde{F}_n(x) = \mathrm{P}(X_1 \leqslant x | T_n).$$

*Corresponding author. Fax: +1 604 291 4368.
E-mail address: lockhart@sfu.ca (R. Lockhart).

0167-7152/$ - see front matter
doi:10.1016/j.spl.2005.04.050

Lilliefors (1967, 1969) proposed Kolmogorov–Smirnov tests for the null hypothesis of an exponential distribution with unknown scale parameter and for the null hypothesis of a normal distribution with unknown mean and variance. The test statistics proposed were

$$\hat{D}_n = \sqrt{n} \sup_x |F_n(x) - \hat{F}_n(x)|,$$

where F_n is the usual empirical distribution function, that is,

$$F_n(x) = n^{-1} \sum_{i=1}^{n} 1(X_i \leqslant x).$$

Srinivasan (1970) proposed, for the normal and exponential null hypotheses analogous Kolmogorov–Smirnov test statistics based on $\tilde{F}_n$:

$$\tilde{D}_n = \sqrt{n} \sup_x |F_n(x) - \tilde{F}_n(x)|.$$

Both Srinivasan and Lilliefors studied their tests by simulation. Kac et al. (1955) derived large sample distribution theory for $\hat{D}_n$ in the case of tests for the normal distribution. Sukhatme (1972) extended this to general regular families by showing that the process

$$\sqrt{n}\{F_n(x) - \hat{F}_n(x)\}$$

converges weakly, under H_0, to a mean zero Gaussian process whose covariance depends on the model being tested. Moore (1973) showed the process

$$\sqrt{n}\{F_n(x) - \tilde{F}_n(x)\}$$

converges weakly to the same limit in the exponential, Normal(μ, σ^2) and Uniform $[0, \theta]$ families, by establishing that for these families

$$n^{1-\delta} \sup_x |\hat{F}_n(x) - \tilde{F}_n(x)| \to 0 \tag{1}$$

in probability for any $\delta > 0$ fixed. The weak convergence result is a consequence of this in the special case $\delta = 1/2$. For a detailed practical discussion of tests of this type see Stephens (1986).

We refer to (1) as Moore's conjecture. In addition to the cases established by Moore (who studied explicit forms for $\tilde{F}_n$ in the normal, exponential and uniform cases), the conjecture (1) has been shown to hold for the inverse Gaussian family by O'Reilly and Rueda (1992) when $\delta = 1/2$.

In Section 2 we use a uniform version of the local central limit theorem to establish (1) for exponential families where the complete sufficient statistic has a density relative to Lebesgue measure. We give a corresponding result for exponential families supported on a lattice and end the section with an example showing the conjecture is not much more general than the cases covered by our theorem. In particular the result does not hold for the $N(\theta, \theta^2)$ curved exponential family. The paper finishes with proofs and lemmas.

2. Main results

2.1. Absolutely continuous distributions

We suppose our interest is to test the hypothesis that G belongs to a natural exponential family with density, relative to Lebesgue measure, of the form

$$f(x;\theta) \equiv c(x)\exp\{\theta^t T(x) - K(\theta)\}$$

with natural parameter space $\Theta \subset R^k$. The statistic

$$T_n \equiv T_n(X_1,\ldots,X_n) = \sum_{i=1}^{n} T(X_i)$$

is complete and sufficient. The notation $P_{\hat{\theta}_n}(A)$ describes the plug-in estimate of $P_\theta(A)$, that is, the function $\theta \mapsto P_\theta(A)$ evaluated at $\theta = \hat{\theta}_n$.

Theorem 1. *Suppose the true value θ_0 of θ is in the interior of Θ. Assume that there is an integer r and a neighbourhood $\mathcal{N}$ of θ_0 such that T_r has a bounded density relative to Lebesgue measure for each $\theta \in \mathcal{N}$. Then for each fixed integer m and each $\delta > 0$ we have*

$$n^{1-\delta} \sup_H |P_{\hat{\theta}}\{(X_1,\ldots,X_m) \in H\} - P\{(X_1,\ldots,X_m) \in H | T_n\}| \to 0$$

almost surely as $n \to \infty$. The supremum is over all Borel sets H in R^m.

Remark 1. Moore's conjecture is the special case $m = 1$ with the supremum taken over the smaller class of H of the form $(-\infty, x]$. The condition that T_r have a bounded density for some r is mild; it is equivalent to integrability of some power of the characteristic function. (That is, if $\eta(u) = E_\theta\{\exp(iuT_1)\}$ and there is a $0<\gamma<\infty$ such that $\int_{-\infty}^{\infty} |\eta(u)|^\gamma \, du < \infty$ then T_r has a bounded density for all $r > \gamma$.) If T_r has a bounded density for some r_0 it has a bounded density for all larger r (see Bhattacharya and Ranga Rao (1976, Section 19) for a discussion). For the normal case, for instance, we have $T_n = (\sum X_i, \sum X_i^2)$ which has a bounded density if $n \geqslant 3$. (For $\mu = 0$ and $\sigma = 1$ for instance the density is a constant multiple of $f(u,v) = (v - u^2/n)^{(n-3)/2} e^{-v/2} 1(v > u^2/n)$ which is bounded for $n \geqslant 3$.)

Remark 2. The supremum over H is the total variation distance between the measures $P_{\hat{\theta}}\{(X_1,\ldots,X_m) \in \cdot\}$ and $P\{(X_1,\ldots,X_m) \in \cdot \,|T_n\}$.

Remark 3. Through the rest of the paper all convergences of random quantities to 0 are almost sure. It is well known that $\hat{\theta}_n \to \theta_0$ almost surely.

Remark 4. Moore (1973) notes his conjecture holds in the normal and exponential cases even when H_0 is false. It will be seen, by examining our proof, that it is not necessary for H_0 to be true. It is, however, necessary that T_n/n converge in large samples to some value of θ for which the conditions of the theorem apply. In the exponential case for instance we need $\sum X_i/n$ to converge to a positive limit; we cannot condition on a negative value of $\bar{X}$ to compute a Rao–Blackwell estimate. Typically, of course, exponential family models would not be used when the statistic T_n/n takes values outside the range of the mean parameter of the model.

Remark 5. Remark 4 has implications for the consistency of goodness-of-fit tests such as that based on $\tilde{D}_n$. Suppose the true cdf of the X_i is G and that G is not in the closure of H_0 (that is, there is no sequence θ_n such that $F(x;\theta_n) \to G(x)$ for all x). Then the Kolmogorov–Smirnov statistic $\hat{D}_n$ converges to ∞ as $n \to \infty$ and the test is consistent against G. A similar conclusion holds for $\tilde{D}_n$ provided G satisfies the conditions in Remark 4.

2.2. The discrete case

Now suppose $X_1, \ldots, X_n$ are discrete with

$$f(x;\theta) \equiv P_\theta(X_1 = x) = c(x)\exp\{\theta^t T(x) - K(\theta)\}$$

for x in some countable set $\mathscr{X}$. We assume that as x ranges over $\mathscr{X}$ the function $T(x)$ takes values in a k dimensional lattice, namely, a set of the form $\{(a + \ell_1 h_1 + \cdots + \ell_k h_k); \ell_i \in \{0, \pm 1, \ldots\}, i = 1, \ldots, k\}$ for some k dimensional vectors $a, h_1, \ldots, h_k$. We can then use a different local central limit theorem to obtain an analogue of Theorem 1. For simplicity we assume in the proofs that each component $T_i(x)$ of $T(x)$ is actually integer valued and that the lattice size of the distribution of $T_i(X)$ is 1. This amounts to saying that the greatest common divisor of

$$\{j - \ell : P\{T_i(X_1) = j\} > 0, P\{T_i(X_1) = \ell\} > 0\}$$

is 1. Notice that the support of the distribution of T does not depend on θ.

Theorem 2. *Suppose θ_0 is in the interior of Θ. Assume that* $\mathrm{Var}_{\theta_0}\{T(X_1)\}$ *is nonsingular. Then for each fixed integer m and each $\delta > 0$ we have*

$$n^{1-\delta} \sup_H |P_{\hat{\theta}}\{(X_1, \ldots, X_m) \in H\} - P\{(X_1, \ldots, X_m) \in H | T_n\}| \to 0$$

almost surely as $n \to \infty$. The supremum is over all subsets H in $\mathscr{X}^m$.

2.3. A counterexample

In both the discrete and continuous cases covered by our theorem the minimal sufficient statistic has the same dimension as the parameter space. When this is not the case (1) will generally not hold for $\delta = 1/2$ (or any smaller δ) as the following example shows. Suppose $X_1, \ldots, X_n$ are an iid sample from the $N(\theta, \theta^2)$ distribution where the unknown parameter θ belongs to $\Theta = (-\infty, \infty)\backslash\{0\}$. It is easily seen that the statistic $T_n \equiv (\sum X_i, \sum X_i^2)$ is minimal sufficient for this model. Since this statistic is the canonical sufficient statistic for the larger $N(\mu, \sigma^2)$ model the Rao–Blackwell estimate $\tilde{F}_n$ of the underlying cdf of the X_i is identical in the two models.

For the $N(\mu, \sigma^2)$ model the plug-in estimator of $F(x, \theta)$ is

$$\hat{F}_{\text{full}}(x) = \Phi\{(x - \bar{X})/\hat{\sigma}\},$$

where $\bar{X}$ is the usual sample mean, $\hat{\sigma}^2 = \sum X_i^2/n - \bar{X}^2$ and Φ is the standard normal cdf. As observed in the Introduction it is well known that

$$\sqrt{n}[F_n\{\mu + \sigma\Phi^{-1}(\cdot)\} - \hat{F}_{\text{full}}\{\mu + \sigma\Phi^{-1}(\cdot)\}]$$

converges weakly in $D[0,1]$ to a mean 0 Gaussian process with covariance

$$\rho_{\text{full}}(s,t) = \min(s,t) - st - J_1(s)J_1(t) - J_2(s)J_2(t)/2,$$

where $J_1(s) = \phi\{\Phi^{-1}(s)\}$, $J_2(s) = \Phi^{-1}(s)J_1(s)$ and ϕ is the standard normal density. By Moore's original result the same conclusion holds for the process

$$\tilde{W}_n \equiv \sqrt{n}[F_n\{\mu + \sigma\Phi^{-1}(\cdot)\} - \tilde{F}_n\{\mu + \sigma\Phi^{-1}(\cdot)\}].$$

For the $N(\theta, \theta^2)$ model $\tilde{F}_n$ is unchanged so the weak limit of $\tilde{W}_n$ is unchanged. The mle of θ, however, is now a root of the equation $\sum X_i(X_i - \theta) = n\theta^2$. There are two roots, one positive and one negative; the mle is the one of these roots which maximizes the likelihood. It is easily seen that this root is consistent and that Sukhatme's (1972) theory applies to show that

$$\sqrt{n}[F_n\{\mu + \sigma\Phi^{-1}(\cdot)\} - \hat{F}_{\text{rest}}\{\mu + \sigma\Phi^{-1}(\cdot)\}]$$

converges weakly to a mean 0 Gaussian process with covariance function

$$\rho_{\text{rest}}(s,t) = \min(s,t) - st - \{J_1(s) + J_2(s)\}\{J_1(t) + J_2(t)\}/3.$$

Since the restricted and full covariance functions are different we cannot have

$$n^{1/2}\{\tilde{F}_n(x) - \hat{F}_{\text{rest}}(x)\} \to 0$$

uniformly in x and so (1) does not hold for $\delta \leqslant 1/2$.

The same sort of argument may be expected to apply in any curved exponential model with parameter space of dimension say p embedded in a natural exponential family of dimension $k > p$ (provided the curved family is not flat so the minimal sufficient statistic has dimension higher than p).

3. Proofs

Proof of Theorem 1. We do the notationally simpler case $m = 1$ but the extension to general m is easy. By shrinking $\mathcal{N}$ if necessary we may assume that the closure of $\mathcal{N}$ lies in the interior of Θ and that the conditions on the existence of a density hold for all θ in the closure of $\mathcal{N}$. Let $f_n(t;\theta)$ denote the density of T_n with respect to Lebesgue measure; this density exists and is bounded over t for all $\theta \in \mathcal{N}$ and all $n \geqslant r$. For $n \geqslant r+1$ the pair (X_n, T_{n-1}) has joint density (because X_n is independent of T_{n-1})

$$f_{X_n,T_{n-1}}(x,t) \equiv f(x;\theta)f_{n-1}(t;\theta).$$

Since $T_n = X_n + T_{n-1}$ we see that (X_n, T_n) has joint density

$$f_{X_n,T_n}(x,t) \equiv f(x;\theta)f_{n-1}\{t - T(x);\theta\}.$$

Now we observe that $P(X_1 \in H|T_n) = P(X_n \in H|T_n)$ so we evaluate

$$P(X_n \in H|T_n = t) = \int_H f_{X_n|T_n}(x|t)\,\mathrm{d}x = \frac{\int_H f(x;\theta)f_{n-1}\{t - T(x);\theta\}\,\mathrm{d}x}{f_n(t;\theta)}.$$

The right-hand side of this formula does not depend on θ. Since the mle $\hat{\theta}_n$ converges almost surely to θ_0 it must lie in $\mathcal{N}$ for all large n. From now on we work on the event $\hat{\theta}_n \in \mathcal{N}$ and write

$$P(X_n \in H | T_n = t) = \frac{\int_H f(x; \hat{\theta}_n) f_{n-1}\{t - T(x); \hat{\theta}_n\}\,\mathrm{d}x}{f_n(t; \hat{\theta}_n)}$$

or

$$P(X_1 \in H | T_n) = \frac{\int_H f(x, \hat{\theta}_n) f_{n-1}\{T_n - T(x); \hat{\theta}_n\}\,\mathrm{d}x}{f_n(T_n; \hat{\theta}_n)}.$$

On the other hand

$$P_{\hat{\theta}_n}(X_1 \in H) = \int_H f(x; \hat{\theta}_n)\,\mathrm{d}x.$$

Comparison of these two formulas shows that for $\hat{\theta}_n \in \mathcal{N}$

$$\sup_H |P_{\hat{\theta}_n}(X_1 \in H) - P(X_n \in H|T_n)| \leqslant \frac{\int_R f(x; \hat{\theta}_n) |f_{n-1}\{T_n - T(x); \hat{\theta}_n\} - f_n(T_n; \hat{\theta}_n)|\,\mathrm{d}x}{f_n(T_n; \hat{\theta}_n)}.$$

We will show that $n^{1-\delta}$ times the right-hand side of this inequality tends to 0. Our proof uses a uniform version of the local central limit theorem following Bhattacharya and Ranga Rao (1976) but using the uniform version of their results outlined by Yuan and Clarke (2004). Our lemma below contains a number of well-known facts about exponential families which we use in the sequel and in the proof of the local limit conclusion.

Lemma 1. *Under the conditions of Theorem 1 the random vector T_n has:*

(1) *moment generating function* $E_\theta[\exp\{\phi^t T(X)\}] = \exp\{K(\phi + \theta) - K(\theta)\}$;
(2) *mean vector* $n\mu(\theta) \equiv nK'(\theta)$;
(3) *covariance matrix* $nV(\theta) \equiv nK''(\theta)$ *which is nonsingular for* $\theta \in \mathcal{N}$;
(4) *finite moments of all orders which depend continuously on* θ;

Moreover, for $n \geqslant r$ the quantity $\{T_n - n\mu(\theta)\}/\sqrt{n}$ has a density $q_n(\cdot;\theta)$. There is a function, $\psi(u, \theta) = \sum_1^k a_i(\theta)u_i + \sum_{ij\ell} b_{ij\ell}(\theta)u_i u_j u_\ell$ such that

$$\varepsilon_n \equiv n^{1-\delta} \sup_{\theta \in \mathcal{N}} \sup_u |q_n(u; \theta) - \phi\{u, V(\theta)\}\{1 + \psi(u, \theta)/\sqrt{n}\}| \to 0. \tag{2}$$

Here $\phi(u, V)$ is the multivariate normal density with mean 0 and covariance matrix V. Finally the functions a_i and $b_{ij\ell}$ depend continuously on θ.

We will not prove this lemma in detail. The conclusions in the enumerated list are well known properties of exponential families. Non-singularity of $V(\theta)$ follows from the existence of a density for T_r. Assertion (2) is a consequence of a uniform version of Theorem 19.2 in Bhattacharya and Ranga Rao (1976). The proof of our minor generalization is essentially that outlined by Yuan and Clarke (2004). To get the conclusion with $n^{1-\delta}$ it is necessary to use the Edgeworth expansion up to order 4 given by Bhattacharya and Ranga Rao. Finally the quantities a_i and $b_{ij\ell}$ are functions of moments of order 3 and so are continuous by the earlier assertions in the theorem.

The density of T_n and the density q_n are related by

$$f_n(t;\theta) = n^{-k/2} q_n[n^{-1/2}\{t - n\mu(\theta)\};\theta].$$

Moreover T_n and $\hat{\theta}_n$ are related by $T_n = n\mu(\hat{\theta}_n)$. Thus by Lemma 1

$$n^{k/2} f_n(T_n, \hat{\theta}_n) \to \phi\{0, V(\theta_0)\} = (2\pi)^{-k/2} \det\{V(\theta_0)\}^{-1/2}. \tag{3}$$

We must therefore show that

$$n^{(k+1)/2-\delta} \int_R f(x;\hat{\theta}_n) |f_{n-1}\{T_n - T(x);\hat{\theta}_n\} - f_n(T_n;\hat{\theta}_n)| \,\mathrm{d}x \to 0.$$

Introduce the shorthand notation

$$A_n(x) = \frac{\{T_n - T(x) - (n-1)\mu(\hat{\theta}_n)\}}{\sqrt{n-1}} = \frac{\{\mu(\hat{\theta}_n) - T(x)\}}{\sqrt{n-1}}.$$

Written in terms of q_n we then have

$$f_{n-1}\{T_n - T(x);\hat{\theta}_n\} = (n-1)^{-k/2} q_{n-1}\{A_n(x);\hat{\theta}_n\}.$$

Thus our problem reduces to showing that

$$n^{1-\delta} \int_R f(x,\hat{\theta}_n) \left| \left(\frac{n}{n-1}\right)^{k/2} q_{n-1}\{A_n(x);\hat{\theta}_n\} - q_n(0;\hat{\theta}_n) \right| \mathrm{d}x \to 0. \tag{4}$$

Lemma 1 guarantees $\sup_{u\in R^k} \sup_{\theta\in\mathcal{N}} \sup_{n\geqslant r} q_n(u,\theta) < \infty$. Thus

$$n^{1-\delta} \int_R f(x,\hat{\theta}_n) \left| \left\{ \left(\frac{n}{n-1}\right)^{k/2} - 1 \right\} q_{n-1}\{A_n(x);\hat{\theta}_n\} \right| \mathrm{d}x \to 0. \tag{5}$$

Put $H_n(x) = 1 + \psi\{A_n(x);\hat{\theta}_n\}/\sqrt{n}$. In view of Lemma 1 we have

$$n^{1-\delta} \int_R f(x,\hat{\theta}_n) |q_{n-1}\{A_n(x);\hat{\theta}_n\} - \phi\{A_n(x);\hat{\theta}_n\} H_n(x)| \,\mathrm{d}x \leqslant \varepsilon_n \to 0. \tag{6}$$

Similarly

$$n^{1-\delta} \int_R f(x,\hat{\theta}_n) |q_n\{0; V(\hat{\theta}_n)\} - \phi\{0, V(\hat{\theta}_n)\}| \,\mathrm{d}x \leqslant \varepsilon_n \to 0. \tag{7}$$

From (4)–(7) and the triangle inequality we need only show

$$n^{1-\delta} \int_R f(x,\hat{\theta}_n) |\phi\{A_n(x);\hat{\theta}_n\} H_n(x) - \phi\{0, V(\hat{\theta}_n)\}| \,\mathrm{d}x \to 0. \tag{8}$$

Split the domain of integration into two pieces. Fix α with $0 < \alpha < \delta/2$. Put $I_1 = \{x : |\mu(\hat{\theta}_n) - T(x)| \leqslant n^\alpha\}$ and $I_2 = \{x : |\mu(\hat{\theta}_n) - T(x)| > n^\alpha\}$. Over I_1 use Taylor expansion of ϕ and ψ near 0 and over I_2 Markov's inequality.

For $x \in I_1$ we have $||A_n(x)|| \leqslant n^{\alpha}/\sqrt{n-1} \leqslant 2n^{\alpha-1/2}$. The smallest eigenvalue of $V(\theta)$ is bounded away from 0 over $\mathcal{N}$ and so

$$A_n(x)^t\{V(\hat{\theta}_n)\}^{-1}A_n(x) \leqslant C_1 n^{2\alpha-1}$$

for some constant C_1 and all $x \in I_1$. Since $|1-\exp(-x)| \leqslant x$ for all $x \geqslant 0$ and $\inf_{\theta\in\mathcal{N}} \det\{V(\theta)\} > 0$ there is a constant C_2 such that for all $x \in I_1$

$$|\phi\{A_n(x); V(\hat{\theta}_n)\} - \phi\{0, V(\hat{\theta}_n)\}| \leqslant C_2 n^{2\alpha-1}. \tag{9}$$

Finally the polynomial structure of ψ shows there is a constant C_3 such that

$$|\psi\{A_n(x); V(\hat{\theta}_n)\}| \leqslant C_3 n^{\alpha-1/2} \tag{10}$$

for all $x \in I_1$. Now combine (9) and (10) to see that

$$\begin{aligned} &n^{1-\delta}\int_{I_1} f(x,\hat{\theta}_n)|\phi\{A_n(x);\hat{\theta}_n\}H_n(x) - \phi\{0, V(\hat{\theta}_n)\}|\,\mathrm{d}x \\ &\quad \leqslant C_2 n^{1-\delta+(2\alpha-1)} + C_3 n^{1-\delta+(\alpha-1/2)-1/2} = C_2 n^{2\alpha-\delta} + C_3 n^{\alpha-\delta} \to 0. \end{aligned} \tag{11}$$

Since the statistic T has finite moments of all orders and all these moments depend continuously on θ there is, for each s, a constant D_s such that

$$\sup_{\theta\in\mathcal{N}} P_\theta\{|T(X) - \mu(\theta)| > n^{\alpha}\} \leqslant \frac{D_s}{n^{s\alpha}}.$$

Thus on the event $\hat{\theta}_n \in \mathcal{N}$ we have

$$\int_{I_2} f(x,\hat{\theta}_n)\,\mathrm{d}x \leqslant \frac{D_s}{n^{s\alpha}}. \tag{12}$$

Take $s = 1/\alpha$. Combine (12) and (11) to get (8), finishing the proof. □

Proof of Theorem 2. Again take $m = 1$. Let $f_n(t;\theta) = P_\theta(T_n = t)$. Then

$$P(X_n \in H | T_n = t) = \frac{\sum_{x\in H} P_\theta\{X_n = x; T_{n-1} = t - T(x)\}}{P_\theta\{T_n = t\}}.$$

We use the independence of T_{n-m} and $(X_{n-m+1},\ldots,X_n)$ and the fact that this conditional probability does not depend on θ to write

$$P\{X_n \in H | T_n\} = \frac{\sum_{x\in H} f(x;\hat{\theta}_n) f_{n-1}\{T_n - T(x);\hat{\theta}_n\}}{f_n(T_n;\hat{\theta}_n)}.$$

As in the continuous case this gives the bound

$$\begin{aligned} &\sup_H |P_{\hat{\theta}}\{X_1 \in H\} - P\{X_1 \in H | T_n\}| \\ &\quad \leqslant \frac{\sum_x f(x;\hat{\theta}_n)|f_{n-1}\{T_{n-1} - T(x);\hat{\theta}_n\} - f_n(T_n;\hat{\theta}_n)|}{f_n(T_n;\hat{\theta}_n)}. \end{aligned}$$

Finish the proof as for Theorem 1 using the following local central limit theorem. □

Lemma 2. *Under the conditions of Theorem* 2 *the vector* T_n *has mean* $n\mu(\theta) \equiv nK'(\theta)$ *and covariance* $nK''(\theta)$. *There is a function,* $\psi(u, \theta)$ *of the form* $\psi(u, \theta) = \sum_1^k a_i(\theta)u_i + \sum_{ij\ell} b_{ij\ell}(\theta)u_i u_j u_\ell$ *such that*

$$n^{1-\delta} \sup_{\theta \in \mathcal{N}} \sup_{u \in R^k} |n^{1/2} P_\theta(T_n = t) - \phi\{u, V(\theta)\}\{1 + \psi(u, \theta)/\sqrt{n}\}| \to 0,$$

where $u = n^{-1/2}\{t - n\mu(\theta)\}$. *The functions* a_i *and* $b_{ij\ell}$ *are continuous.*

Acknowledgements

The authors acknowledge grant support from the Natural Sciences and Engineering Research Council of Canada.

References

Bhattacharya, R.N., Ranga Rao, R., 1976. Normal Approximations and Asymptotic Expansions. R.B. Krieger, Malabar, FL.

Kac, M., Kiefer, J., Wolfowitz, J., 1955. On tests of normality and other tests of goodness of fit based on distance methods. Ann. Math. Statist. 26, 189–211.

Lilliefors, H.W., 1967. On the Kolmogorov–Smirnov test for normality with mean and variance unknown. J. Amer. Statist. Assoc. 62, 399–402.

Lilliefors, H.W., 1969. On the Kolmogorov–Smirnov test for the exponential distribution with mean unknown. J. Amer. Statist. Assoc. 64, 387–389.

Moore, D.S., 1973. A note on Srinivasan's goodness-of-fit test. Biometrika 60, 209–211.

O'Reilly, F.J., Rueda, R., 1992. Goodness of fit for the inverse Gaussian distribution. Canad. J. Statist. 20, 387–397.

Srinivasan, R., 1970. An approach to testing the goodness of fit of incompletely specified distributions. Biometrika 57, 605–611.

Stephens, M.A., 1986. Tests based on EDF statistics. In: D'Agostino, R.B., Stephens, M.A. (Eds.), Goodness-of-fit Techniques. Marcel Dekker, New York, pp. 97–193 (Chapter 4).

Sukhatme, S., 1972. Fredholm determinant of a positive definite kernel of a special type and its application. Ann. Math. Statist. 43, 1914–1926.

Yuan, A., Clarke, B., 2004. Asymptotic normality of the posterior given a statistic. Canad. J. Statist. 32, 119–137.