Extremes of normed empirical moment generating function processes

Michael Stewart[*]**and John Robinson**[†]
University of Sydney

March 16, 2004

Abstract

We present a method for deriving the limiting distribution of the maximum of a normed empirical moment generating function process indexed by one parameter. We first extend slightly the results of Csörgő *et al.* (1986b) to provide the rate of convergence for a Gaussian approximation to a non-Donsker empirical process. In cases we consider, the maximum tends to infinity in probability, but when appropriately scaled has a limiting Gumbel extreme value distribution.

Keywords: Empirical processes, extreme value distribution, limiting distribution, empirical moment generating function, extremes of Gaussian processes.

AMS 2000 subject classifications: 62E20, 60G70.

1 Introduction

Suppose $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables with common distribution F and quantile function given by $F^{-1}(u) = \inf\{\, x \mid F(x) \geq u \,\}$, and write $Fg = \int g\, dF$ if this exists. Let $\mathscr{G}$ be a class of functions satisfying $Fg^2 < \infty$ for all $g \in \mathscr{G}$. The empirical process indexed by $\mathscr{G}$ is $S_n(g) = n^{-1/2}\sum_{i=1}^{n}[g(X_i) - Fg]$. We are interested in the limiting behaviour of the quantity $M_n(\mathscr{G}) = \sup_{g\in\mathscr{G}} S_n(g)$ as $n \to \infty$. If $\mathscr{G}$ is a Donsker class, then

[*]email: `michaels@maths.usyd.edu.au`
[†]email: `johnr@maths.usyd.edu.au`

the empirical process $\{ S_n(g) \mid g \in \mathscr{G} \}$ converges weakly in the space of bounded functions on $\mathscr{G}$ to a tight, mean-zero Gaussian process $\{ S_0(g) \mid g \in \mathscr{G} \}$ with covariance function $Fgh - FgFh$, $g, h \in \mathscr{G}$. In this case, $M_n(\mathscr{G})$ converges in distribution to $M_0(\mathscr{G}) = \sup_{g\in\mathscr{G}} S_0(g)$.

In various statistical applications a statistic of the form $M_n(\mathscr{G})$ occurs but $\mathscr{G}$ is not necessarily Donsker. What usually happens is that the class is restricted to some $\mathscr{G}_0 \subset \mathscr{G}$ which is Donsker, so that the restricted maximum $M_n(\mathscr{G}_0)$ converges in distribution to $M_0(\mathscr{G}_0)$ (see Dacunha-Castelle & Gassiat (1999) for an example of this).

One particular example is the normed empirical moment generating function (emgf) process corresponding to the one parameter class of functions $\mathscr{G} = \{ g_\theta \mid \theta \in \Theta/2 \}$ where $\Theta/2 = \{ \theta/2 \mid \theta \in \Theta \}$, $\Theta = \left\{ \theta \in \mathbb{R} \;\middle|\; \int e^{x\theta}\, dF(x) < \infty \right\}$, $g_\theta(x) = \exp\{x\theta - K(2\theta)/2\}$ and $K(\theta) = \log \int e^{x\theta}\, dF(x)$. Write $S_n(\theta) = S_n(g_\theta)$ and for any $I \subset \Theta/2$ write $M_n(I) = \sup_{\theta\in I} S_n(\theta)$. Several authors have noted that the unrestricted maximum $M_n(\Theta/2)$ tends to infinity in probability, mainly in connection with finite mixture models; see for instance Bickel & Chernoff (1993) and Hartigan (1985) for a similar problem. However if I is a closed interval of finite length, then results in Csörgő (1982) imply that $M_n(I)$ converges weakly to $M_0(I) = \sup_{\theta\in I} S_0(g_\theta)$. The same conclusion can be reached using Theorems A and 2.1 from Csörgő *et al.* (1986a). Theorem A implies that

$$\sup_{\theta\in\Theta/2} \left| \int_0^1 g_\theta \circ F^{-1}(u)\, d\alpha_n(u) - \int_{1/n}^{1-1/n} g_\theta \circ F^{-1}(u)\, dB_n(u) \right| \xrightarrow{P} 0 \;, \qquad (1)$$

where $\alpha_n(u) = n^{-1/2}[G_n(u) - u]$ is the empirical process based on the empirical cumulative distribution function $G_n(u) = n^{-1}\sum_{i=1}^n 1\{U_i \le u\}$ of a $U(0,1)$ random sample $U_1, U_2, \ldots, U_n$ and $\{ B_n(u) \mid 0 \le u \le 1 \}$ is a sequence of Brownian Bridge processes, all defined on the same probability space. Under a further metric entropy condition (similar to conditions in chapter 2.5 of van der Vaart & Wellner (1996)), Theorem 2.1 of Csörgő *et al.* (1986a) shows that the limits of integration in the second integral in (1) can be replaced by 0 and 1. This is shown to hold if the supremum in (1) is taken only over a closed interval of finite length $I \subset \Theta/2$.

A different approach is taken in Bickel & Chernoff (1993); they show that when F is the standard normal distribution, writing $M_n = M_n(\Theta/2)$, for a certain sequence $a_n = O(\log n)^{1/2}$,

$$P\left\{a_n\left[M_n - a_n\right] + \log 2\pi \leq x\right\} \rightarrow \exp\left\{-e^{-x}\right\} \ . \tag{2}$$

We prove (2) for a general F, replacing M_n with $M_n(\Theta_n)$ where $\{\Theta_n\}$ is any sequence of subsets satisfying $\cup_n \Theta_n = \Theta/2$ and a certain technical condition (see remarks following Theorem 3.1 below). In many cases we can take $\Theta_n \equiv \Theta/2$. The sequence $a_n \rightarrow \infty$ then depends on F and the sequence of indexing sets $\{\Theta_n\}$ chosen.

There are at least two applications of the work in this paper. Firstly, our Theorem 4.2 gives the limiting behaviour of test statistics for tests of homogeneity in two component exponential family mixture models, thus generalising work in Bickel & Chernoff (1993) and Liu *et al.* (2003). Secondly, Theorem 3.1 provides an improved approximation to the empirical moment generating function, which is relevant to various tests of goodness of fit and other diagnostic procedures; see for instance Epps *et al.* (1982) and Ghosh (1996).

In section 2 we present some preliminary results on boundary crossings by Gaussian processes. In section 3 we establish an approximation of the form (1). In section 4 we apply results from section 2 to derive our version of (2). In section 5 we give sufficient conditions and in section 6 examples are considered.

Proofs appear in section 7.

2 Some Gaussian preliminaries

Our main result in section 4 below relies on some results on boundary crossings by Gaussian processes, based on results of Hüsler (1990, 1995), who considers a single standardised Gaussian process $\{\, X(t) \mid t \in [0,T] \,\}$, by which is meant $EX(t) = 0$ and $EX^2(t) = 1$ for all t, with $(0,\infty)$-valued boundary functions $\{\, b_T(t) \mid t \in [0,T] \,\}$ and computes $\lim_{T\to\infty} P\{X(t) \le b_T(t),\, t \in [0,T]\}$.

We generalise this to a sequence of such processes $\{\, X_n(t) \mid t \in T_n \,\}$ whose correlation functions may change with n, where the intervals $\{T_n\}$ are possibly of infinite length and boundary functions $\{\, b_n(t) \mid t \in T_n \,\}$ are possibly infinite-valued. We are particularly interested in the case where for each n the boundary is essentially flat in a central region and increases suddenly outside this region, and where the minimum boundary height and the lengths of the T_n's tend to infinity with n. Our results are weaker than Hüsler's though, in the sense that we demand more smoothness from the correlation functions than he does. "Note that for the rest of this section, all Gaussian processes are standardised".

Condition A. *The processes $\{\, X_n(t) \mid t \in T_n \,\}$ are almost surely continous, the correlation functions $r_n(s,t) = EX_n(s)X_n(t)$, have continous partial derivatives up to order 2, satisfying*

$$r_n^{(1,1)}(t,t) = -r_n^{(0,2)}(t,t) = -r_n^{(2,0)}(t,t) \equiv 1 \ , \tag{3}$$

for all $t \in T_n$ and each n, where $r_n^{(i,j)}(s,t) = \partial^{i+j} r_n(u,v)/\partial u^i \partial v^j|_{u=s,v=t}$.

That (3) holds is not as restrictive as it might initially seem. It is possible to transform the scale of a standardised process with suitably smooth correlation function to obtain a new process whose correlation function satisfies (3), as we do in section 4. See also Hüsler (1995).

The next condition divides T_n into a region where the boundary functions are essentially flat and another where they increase suddenly. Henceforth, for any interval I, $\lambda(I)$ denotes its Lebesgue measure.

Condition B. *There is a sequence of intervals* $\{T_{n1} \subset T_n\}$ *such that as* $n \to \infty$,

1. *writing* $\ell_{n1} = \inf\{\, b_n(t) \mid t \in T_{n1} \,\}$ *and* $m_{n1} = \sup\{\, b_n(t) \mid t \in T_{n1} \,\}$, $\ell_{n1} \to \infty$, $m_{n1}/\ell_{n1} \to 1$ *and* $\limsup \lambda(T_{n1}) \exp\{-\ell_{n1}^2/2\} < \infty$;

2. $T_n \backslash T_{n1}$ *can be partitioned into a finite number of intervals* $\cup_{j=2}^{k} T_{nj}$, *where for each* $j = 2, 3, \ldots, k$, $b_n(\cdot)$ *is bounded below on* T_{nj} *by a continuous monotone function* $B_{nj}(\cdot)$ *which satisfies* $\inf\{\, B_{nj}(t) \mid t \in T_{nj} \,\} \to \infty$ *and* $\int_{T_{nj}} \exp\{-B_{nj}^2(t)/2\}\, dt \to 0$.

The next condition also applies to a sequence of intervals $\{T_{n1}\}$ and has two parts. The first is a stronger version of condition A which controls local dependence uniformly. It ensures that over T_{n1} the processes behave like locally stationary processes (see Hüsler (1990, 1995), Berman (1974)). The second is a decay condition that controls long-range dependence within T_{n1}.

Condition C. *There is a sequence of intervals* $\{T_{n1} \subset T_n\}$, *such that as* $n \to \infty$,

1. *for any* $h_n \to 0$,

$$\sup_{\{\, t \mid t, t+h_n \in T_{n1} \,\}} h_n^{-2} |r_n(t, t+h_n) - 1 + h_n^2/2| \to 0 \tag{4}$$

2. *for any* $u_n \to \infty$ *such that the set* $S(u_n) = \{\, (s,t) \mid s, t \in T_{n1},\ |s-t| \geq u_n \,\}$

remains non-empty,

$$\sup \{ r_n(s,t) \mid (s,t) \in S(u_n) \} \log u_n \to 0 \ . \tag{5}$$

Theorem 2.1. *Suppose we have a sequence of standardised Gaussian processes* $\{ X_n(t) \mid t \in T_n \}$ *with correlation functions* $\{ r_n(s,t) \mid s,t \in T_n \}$ *for intervals* $\{T_n\}$ *so that condition A holds. Suppose that there also exists a sequence of intervals* $\{T_{n1} \subset T_n\}$ *and a sequence of* $(0,\infty]$*-valued boundary functions* $\{ b_n(t) \mid t \in T_n \}$*, so that conditions B and C both hold. Then with* $\ell_{n1} = \inf \{ b_n(t) \mid t \in T_{n1} \}$ *and* $J_n = \lambda(T_{n1}) \exp\{-\ell_{n1}^2/2\} /(2\pi)$*,*

$$P\{X_n(t) \le b_n(t),\, t \in T_{n1}\} - \exp\{-J_n\} \to 0$$

and

$$P\{X_n(t) \le b_n(t),\, t \in T_n \setminus T_{n1}\} \to 1$$

as $n \to \infty$*. Hence*

$$P\{X_n(t) \le b_n(t),\, t \in T_n\} - \exp\{-J_n\} \to 0 \ .$$

The proof of Theorem 2.1 appears in section 7.

3 Establishing the Gaussian approximation

For a fixed sequence $\{\Theta_n\}$ of subsets of $\Theta/2$, define the envelope function $\breve{g}_n(x) = \sup_{\theta \in \Theta_n} g_\theta(x)$. For any g write $F_n g = \int_{1/n}^{1-1/n} g \circ F^{-1}(u)\, du$ whenever Fg exists (note that $F_n g \to Fg$ for each g as $n \to \infty$).

Theorem 3.1. *Suppose that, writing $\delta_n = n^{-1/2} \log n$, we have*

$$\sup_{F(x) \leq \delta_n} \breve{g}_n(x) F(x)^{1/2} = O(c_n) \tag{6}$$

and

$$\sup_{F(x) \geq 1-\delta_n} \breve{g}_n(x) [1 - F(x)]^{1/2} = O(c_n) \tag{7}$$

for some rate $c_n \to 0$ as $n \to \infty$. Then it is possible to construct a sequence of processes equal in distribution to $\{ S_n(\theta) \mid \theta \in \Theta_n \}$ and a sequence of Gaussian processes $\{ Z_n(\theta) \mid \theta \in \Theta_n \}$ with zero mean and covariance functions

$$C_n(\eta, \theta) = F_n g_\eta g_\theta - F_n g_\eta F_n g_\theta \tag{8}$$

on a common probability space such that

$$\sup_{\theta \in \Theta_n} |S_n(\theta) - Z_n(\theta)| = O_p(c_n) \ . \tag{9}$$

Remarks So long as Θ_n is an interval, the envelope functions are given by $\breve{g}_n(x) = g(x; \tilde{\theta}(x))$, where $g(\,\cdot\,; \theta) = g_\theta$ and $\tilde{\theta}(x) = (\inf \Theta_n) \vee \hat{\theta}(x)/2 \wedge (\sup \Theta_n)$, where $\hat{\theta}(x)$ satisfies $K'(\theta) = x$ (derivatives of all orders exist for θ in the interior of Θ, see Barndorff-Nielsen (1978, 1980)). In most cases we can take the sequence $\Theta_n \equiv \Theta/2$ to be maximal, in which case $\breve{g}_n \equiv g(\,\cdot\,; \hat{\theta}(\cdot)/2)$, but in some examples where the corresponding envelope functions $\breve{g}_n$ do not satisfy both (6) and (7), it is possible to choose a sequence $\{\Theta_n\}$ such that $\Theta_n \subset \Theta_{n+1}$ for each n, $\cup_{n=1}^{\infty} \Theta_n = \Theta/2$ and (6) and (7) are both satisfied. For instance, it can be shown that when F is the standard normal distribution and $\Theta_n \equiv \Theta/2 = \mathbb{R}$, (6) and (7) both hold with $c_n = (\log n)^{-1/4}$. However, if F is the gamma distribution on $(0, \infty)$ with shape parameter α (and unit scale), and $\Theta_n \equiv \Theta/2 = (-\infty, 1/2)$, then (7) holds with $c_n = (\log n)^{-1/2}$, but as $x \to 0$, $\check{g}_n(x)F(x)^{1/2} \to [(\alpha/e)^\alpha/\Gamma(\alpha+1)]^{1/2}$, not even tending to zero, so (6) does not hold. However if $\Theta_n = (-\theta_n, 1/2)$ for $\theta_n = \left[n^{1/2}(\log n)^{-3/2}\right]^{1/\alpha}$, then (6) also holds with $c_n = (\log n)^{-1/2}$ (see Stewart (2002) for details).

In the case that each $\Theta_n \subset [0, \infty)$ (respectively $(-\infty, 0]$) and each g_θ is thus non-decreasing (respectively non-increasing), only condition (7) (respectively (6)) is required. In the case $\Theta_n \equiv \Theta_0$ is a fixed set of non-negative real numbers, it is of interest to compare the condition (7) with a sufficient condition that the class of non-decreasing functions $\mathscr{G} = \{ g_\theta \mid \theta \in \Theta_0 \}$ is Donsker, given in van der Vaart (1996). Write $\check{g}(x) = \sup_{\theta \in \Theta_0} g_\theta(x)$ for the envelope function. If $\check{g}(x)[1 - F(x)]^{1/2} \to 0$, or equivalently

$$\check{g}(x)[1 - F(x)]^{-1/2} f(x) = o\left(f(x)/[1 - F(x)]\right) \tag{10}$$

as $F(x) \to 1$, then (7), and hence (9), holds for some $c_n \to 0$. If

$$\int_{-\infty}^{\infty} \check{g}(x)[1 - F(x)]^{-1/2} f(x)\, dx < \infty \tag{11}$$

then by Corollary 3.1 in van der Vaart (1996) $\mathscr{G}$ is a Donsker class. That the condition (11) can be much stronger than (10) can be illustrated by considering what (10) reduces to for various F's. For instance, when $F(x) = 1 - \exp\{-(x \vee 0)\}$, the function required to be finitely integrable by (11) is only required to tend to zero as $x \to \infty$ by (10). When F is the standard normal distribution, the function required to be finitely integrable in (11) is only required to be $o(x)$ as $x \to \infty$ by (10), so can even tend to infinity!

Theorem 3.2 below is a slight generalisation of Theorem 3.2 of Csörgő *et al.* (1986b) and is used to prove Theorem 3.1.

Theorem 3.2. *Let $\{\mathscr{L}_n\}$ be a sequence of classes of functions such that each $\ell \in \mathscr{L}_n$ can be written as $\ell = \ell_1 - \ell_2$ where for $i = 1, 2$, ℓ_i is a left-continous non-decreasing function defined on (0,1). Define for any positive sequence $\delta_n \to 0$,*

$$N_n(\delta_n, \mathscr{L}_n) = \sup_{\substack{0 \le u \le \delta_n \\ \ell \in \mathscr{L}_n}} \{|\ell_1(u)| + |\ell_2(u)| + |\ell_1(1-u)| + |\ell_2(1-u)|\} \, u^{1/2} \ . \quad (12)$$

If, for some rate $c_n \to 0$ as $n \to \infty$, we have

$$N_n\left(n^{-1/2} \log n, \mathscr{L}_n\right) = O(c_n) \ , \quad (13)$$

then there exists a probability space $(\Omega, \mathscr{A}, P)$ with a sequence of $U(0,1)$ empirical processes $\{\alpha_n(\cdot)\}$ and a sequence of standard Brownian Bridges $\{B_n(\cdot)\}$ such that

$$\sup_{\ell \in \mathscr{L}_n} \left| \int_0^1 \ell(u) \, d\alpha_n(u) - \int_{1/n}^{1-1/n} \ell(u) \, dB_n(u) \right| = O_p\left(c_n\right) \ . \quad (14)$$

Proof. This is proved exactly as Theorem 3.2 of Csörgő *et al.* (1986b), where the analogue of the left-hand side of (14) is bounded above by the sum of a finite number of terms of the form

$$N_n\left(n^{-1/2} \log n, \mathscr{L}_n\right) \, O_p\left(1\right) \ .$$

This, together with (13), implies (14). □

Proof of Theorem 3.1 In the language of Theorem 3.2, for each n, each element ℓ of the monotone, left-continuous class of functions defined by $\mathscr{L}_n = \left\{ \ell_\theta = g_\theta \circ F_0^{-1} \mid \theta \in \Theta_n \right\}$ satisfies $|\ell_1| + |\ell_2| = |\ell|$. Conditions (6) and (7) then imply that $N_n(\delta_n, \mathscr{L}_n) = O(c_n)$ (see (12)). Equation (9) then follows from

Theorem 3.2.

4 Main Result

In this section we derive the limiting distribution of the sequence of suprema of the processes $\{ Z_n(\theta) \mid \theta \in \Theta_n \}$ appearing in (9). We introduce some definitions and present Theorem 4.1, the proof of which utilises Theorem 2.1. However the conditions of Theorem 4.1, while easy to formulate, are quite difficult to verify directly in examples. In section 5 we present conditions that imply Theorem 4.1 and are also easier to verify in examples.

We firstly introduce the modified process $\left\{ \tilde{Z}_n(\theta) = Z_n(\theta) + \tilde{X} F_n g_\theta \mid \theta \in \Theta_n \right\}$, where $\tilde{X} \sim \mathscr{N}(0,1)$ independently of Z_n, which is mean-zero Gaussian and has covariance function $E\tilde{Z}_n(\eta)\tilde{Z}_n(\theta) = \tilde{C}_n(\eta,\theta) = F_n g_\eta g_\theta$. This is the generalisation of a trick used in Bickel & Chernoff (1993); the new covariance function is nicer, but the added term has a negligible effect on the maximum. Write $v_n(\theta) = \mathrm{Var}\tilde{Z}_n(\theta) = \tilde{C}_n(\theta,\theta)$. Next define the standardised processes $\left\{ Y_n(\theta) = v_n^{-1/2}(\theta)\tilde{Z}_n(\theta) \mid \theta \in \Theta_n \right\}$, which have correlation functions $\rho_n(\eta,\theta) = [v_n(\eta)v_n(\theta)]^{-1/2}\, \tilde{C}_n(\eta,\theta)$.

Writing partial derivatives as in the statement of condition A, since ρ_n is a correlation function, $\gamma_n(\theta) = \rho_n^{(1,1)}(\theta,\theta) = -\rho_n^{(2,0)}(\theta,\theta) = -\rho_n^{(0,2)}(\theta,\theta)$. Define the transformation $\theta \mapsto \tau_n(\theta) = \int_0^\theta \gamma_n^{1/2}(\eta)\, d\eta$, and write $t \mapsto \theta_n(t)$ for its inverse. Define also $T_n = \{ \tau_n(\theta) \mid \theta \in \Theta_n \}$. Then the standardised Gaussian processes given by $\{ X_n(t) = Y_n(\theta_n(t)) \mid t \in T_n \}$ have correlation functions

$$r_n(s,t) = \rho_n(\theta_n(s), \theta_n(t)) \tag{15}$$

which satisfy condition A in section 2, that is $r_n^{(1,1)}(t,t) = -r_n^{(2,0)}(t,t) = -r_n^{(0,2)}(t,t) \equiv 1$ for each n and t.

Theorem 4.1. *Suppose that there exists a sequence of intervals $\{T_{n1} \subset T_n\}$, each containing 0, such that the correlation functions* (15) *satisfy condition C and if for each real x we define $a_n = [2\log\lambda(T_{n1})]^{1/2}$ and $u_n(x) = a_n + (x - \log 2\pi)/a_n$ then the boundary functions $\left\{ b_{n,x}(t) = v_n^{-1/2}[\theta_n(t)]u_n(x) \mid t \in T_{n1} \right\}$ satisfy condition B. Then for each real x,*

$$P\left\{a_n\left[\sup_{\theta\in\Theta_n} Z_n(\theta) - a_n\right] + \log 2\pi \le x\right\} \to \exp\{-e^{-x}\} \ . \tag{16}$$

See section 7 for the proof.

Combining this with Theorem 3.1 gives our main result:

Theorem 4.2. *Suppose that Theorem 3.1 holds for some rate $c_n \to 0$ and that Theorem 4.1 holds for some sequence of intervals $\{T_{n1}\}$. Suppose that $c_n = o(a_n^{-1})$, where a_n is as in the statement of Theorem 4.1. Then* (2) *holds for the same a_n, with $M_n = M_n(\Theta_n) = \sup_{\theta\in\Theta_n} S_n(\theta)$.*

Proof. If $c_n = o(a_n^{-1})$ then by (9), $\sup_{\theta\in\Theta_n} S_n(\theta) = \sup_{\theta\in\Theta_n} Z_n(\theta) + o_p\,(a_n)^{-1}$ and so

$$a_n\left[\sup_{\theta\in\Theta_n} S_n(\theta) - a_n\right] = a_n\left[\sup_{\theta\in\Theta_n} Z_n(\theta) - a_n\right] + o_p\,(1) \ ,$$

and so the final statement of Theorem 4.1 implies (2) with M_n replaced by $M_n(\Theta_n)$. □

5 Sufficient Conditions for Theorem 4.1

We identify in this section an exponential family of distributions in terms of which easier sufficient conditions may be formulated.

Define $q_n^- = F^{-1}(1/n)$ and $q_n^+ = F^{-1}(1 - 1/n)$ as the lower and upper $1/n$ quantiles of F. Write $A_n = \{\, x \mid q_n^- \le x \le q_n^+ \,\}$. Note that we can write

$\tilde{C}_n(\theta,\eta) = F_n g_\theta g_\eta = \rho(\theta,\eta)\int_{A_n} dF_{\theta+\eta}$, where F_θ is the distribution satisfying $dF_\theta(x) = e^{\theta x - K(\theta)}dF(x)$ and $\rho(\theta,\eta) = \exp\{K(\theta+\eta) - [K(2\theta)+K(2\eta)]/2\}$ is the pointwise limit of both $\tilde{C}_n(\theta,\eta)$ and $\rho_n(\theta,\eta)$ as $n \to \infty$. Identifying the full, linear exponential family $\mathscr{F} = \{\, F_\theta \mid \theta \in \Theta \,\}$ provides nice interpretations of various quantities. For example, $v_n(\theta) = \int_{A_n} dF_{2\theta}$ is the mass put on the set A_n by $F_{2\theta}$.

The first two derivatives $K'(\theta)$ and $K''(\theta)$ give the mean and variance of F_θ. The function $x \mapsto \hat{\theta}(x)$ solving $K'(\theta) = x$ (already remarked upon below Theorem 3.2) gives the maximum likelihood estimate for the parameter θ within $\mathscr{F}$ based on a single observation x. The mapping $\theta \mapsto \gamma_n(\theta)$ defined in the previous subsection can be rewritten as the conditional variance of X, given $X \in A_n$, when unconditionally $X \sim F_{2\theta}$. Hence the pointwise limit as $n \to \infty$ of $\gamma_n(\theta)$ is $K''(2\theta)$, the unconditional variance. Furthermore, the pointwise limit of $\tau_n(\theta)$ defined in the previous subsection is $\tau_0(\theta) = \int_0^\theta K''(2\eta)^{1/2}\,d\eta = \phi(2\theta)/2$, where $\phi(\cdot)$ is the variance-stablising transformation of $\mathscr{F}$.

Condition D. *There exists a monotone sequence of intervals* $\{\, \Theta_{n1} \subset \Theta_n \mid n \in \mathbb{N} \,\}$ *with* $\cup_n \Theta_{n1} = \cup_n \Theta_n$ *such that*

$$\sup_{\substack{a_i \ge 0,\, \sum_i a_i \le 4 \\ \theta_i,\eta_i \in \Theta_{n1}}} \left| \prod_{i=1}^4 \int_{A_n} x_i^{a_i}\, dF_{\theta_i+\eta_i}(x_i) - \prod_{i=1}^4 \int_{\mathbb{R}} x_i^{a_i}\, dF_{\theta_i+\eta_i}(x_i) \right| \to 0 \ .$$

The pointwise limit of $r_n(s,t)$ is given by $r_0(s,t) = \rho(\theta_0(s),\theta_o(t))$, where $t \mapsto \theta_0(t)$ is the inverse of the transformation $\theta \mapsto \tau_0(\theta) = \phi(2\theta)/2$. In section 7 we show that condition D implies that $r_n(s,t) = r_0(s,t)[1+o(1)]$ and $r_n^{(2,0)}(s,t) = r_0^{(2,0)}(s,t)[1+o(1)]$, with both $o(1)$ terms uniform over $s,t \in T_{n1} = \{\, \tau_n(\theta) \mid \theta \in \Theta_{n1} \,\}$. If $r_0(\cdot,\cdot)$ satisfies condition E below, the functions $r_n(\cdot,\cdot)$ can be shown to satisfy condition C in section 2 for this choice of T_{n1}. Write $T_0 = \cup_n T_{n1}$.

Condition E. $\lambda(T_0) = \infty$*, and the following two conditions are satisfied:* $\sup_{t \in T_0} h^{-2} \left| r_0(t, t+h) + 1 - h^2/2 \right| \to 0$ *as* $h \to 0$ *and* $\sup \{ r_0(s,t) \mid s, t \in T_0,\ |s-t| \geq u \} \log u \to 0$ *as* $u \to \infty$.

The conditions D and E imply condition C. We still need more to imply B. Define $\theta_n^- = \hat{\theta}(q_n^-)/2$ and similarly θ_n^+.

Condition F. *There exist sequences* $\theta_n^- < \theta_{n1}^- \leq 0 \leq \theta_{n1}^+ < \theta_n^+$, $\theta_{n2}^- \leq \theta_{n1}^-$ *and* $\theta_{n1}^+ \leq \theta_{n2}^+$ *taking values in* $\Theta/2$ *such that, defining* $\Theta_{n1} = \Theta_n \cap [\theta_{n1}^-, \theta_{n1}^+]$, $\Theta_{n2}^- = \Theta_n \cap (\theta_{n2}^-, \theta_{n1}^-)$, $\Theta_{n2}^+ = \Theta_n \cap (\theta_{n1}^+, \theta_{n2}^+)$, $v_{n2}^- = \inf_{\theta \in \Theta_{n2}^-} v_n(\theta)$*, similarly* v_{n2}^+, $\Theta_{n3}^- = \Theta_n \cap (-\infty, \theta_{n2}^-]$, $\Theta_{n3}^+ = \Theta_n \cap [\theta_{n2}^+, \infty)$, $L_{n1} = \lambda(\tau_0(\Theta_{n1}))$ $L_{n2}^- = \lambda(\tau_0(\Theta_{n2}^-))$*, similarly* L_{n2}^+*, we have that*

1. *either* $\Theta_{n2}^- = \emptyset$ *or* $(v_{n2}^-)^{-1/2} L_{n2}^- / L_{n1} \to 0$ *and*

2. *either* $\Theta_{n3}^- = \emptyset$ *or* $v_{n2}^- \to 0$ *and*

$$\sup_{\theta \in \Theta_{n3}^-} v_n(\theta) K''(2\theta)^{1/2} / |q_n^- - K'(2\theta)| = O\left([v_{n2}^-]^{1/2} L_{n1} \log L_{n1} \right) \tag{17}$$

and both statments also hold if $-$ *is replaced with* $+$ *(here we interpret the "open interval"* (x,x) *as the empty set).*

In subsection 6 we verify these conditions for certain examples of interest.

Theorem 5.1. *Suppose there exist sequences* $\theta_{n2}^- \leq \theta_{n1}^- < \theta_{n1}^+ \leq \theta_{n2}^+$ *taking values in* $\Theta/2$ *satisfying condition F and such that for* $\Theta_{n1} = [\theta_{n1}^-, \theta_{n1}^+]$*, conditions D and E hold. Then Theorem 4.1 holds taking* $T_{n1} = \{ \tau_n(\theta) \mid \theta \in \Theta_{n1} \}$.

6 Examples

In this subsection we show that conditions D, E and F hold for certain examples of interest.

6.1 Normal

Suppose F is the standard normal distribution. Then the first product appearing in condition D can be written as

$$\prod_{i=1}^{4}\left[\int_{\mathbb{R}} x^{a_i}\, dF_{\theta_i+\eta_i}(x) - \int_{\mathbb{R}\setminus A_n} x^{a_i}\, dF_{\theta_i+\eta_i}(x)\right]$$

which when expanded gives a linear combination of sixteen terms of the form $\prod_{i=1}^{4}\int_{S_i} x^{a_i} dF_{\theta_i+\eta_i}(x)$, where each S_i is either $\mathbb{R}$ or $\mathbb{R}\setminus A_n$; the term corresponding to all S_i's equalling $\mathbb{R}$ is the second product in condition D. The remaining 15 terms constitute the remainder which we wish to show tends to zero, uniformly over θ_i, η_i in some Θ_{n1}. In this example each of these terms can be written as

$$\prod_{i=1}^{4}\left\{\sum_{j=0}^{a_i}\binom{a_i}{j}(\theta_i+\eta_i)^{a_i-j}\int_{I_i} y^j f(y)\, dy\right\} \tag{18}$$

where f is the standard normal density, and at most 3 of the I_i's are $\mathbb{R}$, the others being $\mathbb{R}\setminus(-q_n-\theta_i-\eta_i, q_n-\theta_i-\eta_i)$. Suppose we take

$$\Theta_{n1} = \{\theta\colon |\theta| \le (q_n-\Delta_n)/2\}\ , \tag{19}$$

for some $\Delta_n \to \infty$ more slowly than $q_n = O(\log n)^{1/2}$. When $I_i = \mathbb{R}$, each integral is $O(1)$, so the biggest term in the sum in (18) is

$$\sup_{\theta,\eta\in\Theta_{n1}} |\theta+\eta|^{a_i}\, O(1) = O(q_n^{a_i})\ .$$

Otherwise, the integral is bounded above by $2\int_{\Delta_n}^{\infty} y^j f(y)\, dy = O\left(\Delta_n^{j-1}\exp\{-\Delta_n^2/2\}\right)$ and the biggest term in the sum is $O(q_n^{a_i}\exp\{-\Delta_n^2/2\}$. The biggest each product of the form (18) can be is when exactly 3 of the I_i's are $\mathbb{R}$ and only 1 is not. In this case the product is $O\left(q_n^{\sum_i a_i}\exp\{-\Delta_n^2/2\}\right) = O\left(q_n^4\exp\{-\Delta_n^2/2\}\right)$. Any Δ_n such that this last quantity goes to zero will suffice in the definition (19) above, e.g. $\Delta_n = [(4+\varepsilon)\log\log n]^{1/2}$ or perhaps $2(\log\log n + \frac{1}{4}\log\log\log n)^{1/2}$. In any case we must have $\Delta_n = O(\log\log n)^{1/2}$.

The verification of condition E is easier; the limiting correlation function is $r_0(s,t) = \exp\{-(s-t)^2/2\}$, and since this is the correlation function of a stationary process, verifies the local stationarity condition (4). Also, since $e^{-u^2/2}\log u \to 0$ as $u \to \infty$, (5) is also satisfied.

As for condition F, $v_n(\theta) = F(q_n - 2\theta) - F(-q_n - 2\theta)$. Define $\theta_{n1}^+ = (q_n - \Delta_n)/2$ and for some $\alpha_n \to \infty$, define $\theta_{n2}^+ = (q_n + \alpha_n)/2$. Also define $\theta_{nj}^- = -\theta_{nj}^+$ for $j = 1, 2$. Then $v_{n2}^+ = \sup_{\theta_{n1}^+ \le \theta \le \theta_{n2}^+} v_n(\theta) = v_n(\theta_{n2}^+) + O(n^{-1}) = O\left(F(-\alpha_n)\right) = O(\exp\{-\alpha_n^2/2\}/\alpha_n)$. Using the notation of condition F, $\tau_0(\theta) \equiv \theta$ so $L_{n1} = q_n - \Delta_n = O(\log n)^{1/2}$ and $L_{n2}^+ = [\Delta_n + \alpha_n]/2$. The first statement of condition F is satisfied then if

$$\exp\{\alpha_n^2/4\}\,\alpha_n^{1/2}[\Delta_n + \alpha_n]/q_n \to 0\ .$$

Take $\alpha_n = (\log\log n - 4\log\log\log n)^{1/2}$. Then with $q_n = O(\log n)^{1/2}$, $\Delta_n = O(\log\log n)^{1/2}$, the previous display is

$$O(1)\left[(\log n)^{1/4}(\log\log n)^{-1}\right](\log\log n)^{1/2}\left[\log\log n/\log n\right]^{1/2} = O(\log n)^{-1/4} \to 0$$

so the first statement of condition F is satisfied. For the second statement, the supremum on the left hand side of the '+' version of (17) is easily seen to occur at θ_{n2}^+, so the left hand side is $O(v_{n2}^+/\alpha_n)$. Dividing both sides then by v_{n2}^+ gives that the left hand side tends to zero whereas the right-hand side tends to infinity, and so (17) is satisfied.

6.2 Poisson

Now suppose that F is the Poisson distribution with mean μ. Note then that F_θ is the Poisson distribution with mean μe^θ.

Define $q_n = F^{-1}(1-1/n)$, $x_n = q_n + 1$ and θ_{nj}, $j = 1, 2$ via

$$\begin{aligned}
\mu e^{2\theta_{n1}} &= x_n - (8 x_n \log x_n)^{1/2} \\
\mu e^{2\theta_{n2}} &= x_n + (x_n \log x_n)^{1/2} .
\end{aligned}$$

Exploiting the form of the Poisson density, it is sufficient to show the condition D holds with x^a replaced by $1\{x \geq a_i\} x!/(x-a_i)!$. The first product there becomes

$$\begin{aligned}
&\prod_{i=1}^{4} \sum_{x=a_i}^{q_n} \frac{\exp\{-\mu e^{\theta_i+\eta_i}\} (\mu e^{\theta_i+\eta_i})^x}{(x-a_i)!} \\
&= \prod_{i=1}^{4} \left\{ (\mu e^{\theta_i+\eta_i})^{a_i} \sum_{x-a_i=0}^{q_n-a_i} \frac{\exp\{-\mu e^{\theta_i+\eta_i}\} (\mu e^{\theta_i+\eta_i})^{x-a_i}}{(x-a_i)!} \right\} \\
&= \prod_{i=1}^{4} \left\{ (\mu e^{\theta_i+\eta_i})^{a_i} \left[1 - \sum_{y=q_n-a_i+1}^{\infty} \frac{\exp\{-\mu e^{\theta_i+\eta_i}\} (\mu e^{\theta_i+\eta_i})^{y}}{y!} \right] \right\} ,
\end{aligned}$$

which is $\prod_{i=1}^4 (\mu e^{\theta_i+\eta_i})^{a_i}$, the second product appearing in condition D, plus 15 other terms, the biggest of which is bounded above by

$$\begin{aligned}
&(\mu e^{2\theta_{n1}})^4 \sum_{y=x_n-4}^{\infty} \frac{\exp\{-\mu e^{2\theta_{n1}}\} (\mu e^{2\theta_{n1}})^y}{y!} \\
&\sim (\mu e^{2\theta_{n1}})^4 \frac{\exp\{-\mu e^{2\theta_{n1}}\} (\mu e^{2\theta_{n1}})^{x_n-4}}{(x_n-4)!} \\
&\sim \frac{\exp\{-\mu e^{2\theta_{n1}}\} (\mu e^{2\theta_{n1}})^{x_n}}{\sqrt{2\pi(x_n-4)}} \left(\frac{e}{x_n-4} \right)^{x_n-4}
\end{aligned}$$

using Stirling's approximation. Upon inserting the defintion of θ_{n1} given above, this becomes $(2\pi)^{-1/2}$ times

$$\begin{aligned}
&\exp\Big\{(8x_n \log x_n)^{1/2} - x_n\Big\} \Big[x_n - (8x_n \log x_n)^{1/2}\Big]^{x_n} (e/x_n)^{x_n} (x_n - 4)^{3.5} \\
&= \exp\Big\{(8x_n \log x_n)^{1/2}\Big\} \exp\left\{x_n \log\left[1 - \left(\frac{8\log x_n}{x_n}\right)^{1/2}\right]\right\} (x_n - 4)^{3.5} \\
&\sim \exp\Big\{(8x_n \log x_n)^{1/2}\Big\} \exp\Big\{-(8x_n \log x_n)^{1/2} - 4\log x_n\Big\} (x_n - 4)^{3.5} \\
&\sim x_n^{-1/2} \to 0
\end{aligned}$$

as $n \to \infty$, establishing condition D.

Condition E is easily shown to be satisfied, since the limiting correlation function $r_0(s,t)$ is again $\exp\big\{(s-t)^2/2\big\}$, and so the same argument applies as in the previous example.

Next, note that in this case $\tau(\theta) = \mu^{1/2}\left[e^{\theta} - 1\right]$, so $L_{n1} = O\left(x_n^{1/2}\right)$. Also

$$\begin{aligned}
L_{n2} &= O\left(e^{\theta_{n2}} - e^{\theta_{n1}}\right) \\
&= O(1)\left[x_n + (x_n \log x_n)^{1/2}\right]^{1/2} - \left[x_n - (8x_n \log x_n)^{1/2}\right]^{1/2} \\
&= x_n^{1/2}\left\{\left[1 + \left(\frac{\log x_n}{x_n}\right)^{1/2}\right]^{1/2} - \left[1 - \left(\frac{8\log x_n}{x_n}\right)^{1/2}\right]^{1/2}\right\} \\
&= O(\log x_n)^{1/2} .
\end{aligned}$$

Using the well-known relation between tail probabilities for the Poisson and Gamma distributions,

$$\begin{aligned}
v_{n2} &= P\left(\mathrm{Pois}(\mu e^{2\theta_{n2}}) < x_n\right) = P\left(\mathrm{Gam}(x_n, 1) > \mu e^{2\theta_{n2}}\right) \qquad (20) \\
&\sim \frac{\left(\mu e^{2\theta_{n2}}\right)^{x_n} \exp\left\{-\mu e^{2\theta_{n2}}\right\}}{\Gamma(x_n)(\mu e^{2\theta_{n2}} - x_n)} ,
\end{aligned}$$

using an integration-by-parts approximation for the upper tail of the Gamma distribution (see also relation (101) in Stewart (2002)). Using Stirlings approximation for the Gamma function, and inserting the definition of θ_{n2}, we obtain an expression which is asymptotic to a constant times

$$\begin{aligned}
&\frac{\left[x_n + (x_n \log x_n)^{1/2}\right]^{x_n} \exp\left\{-(x_n \log x_n)^{1/2}\right\}}{(x_n-1)^{x_n-1/2}(x_n \log x_n)^{1/2}} \\
&\quad \sim \quad (\log x_n)^{-1/2} \exp\left\{-(x_n \log x_n)^{1/2} + x_n \log\left[1 + \left(\frac{\log x_n}{x_n}\right)^{1/2}\right]\right\} \\
&\quad \sim \quad (\log x_n)^{-1/2} \exp\left\{-\frac{1}{2}\log x_n\right\} = (x_n \log x_n)^{-1/2} \ .
\end{aligned}$$

Thus $v_{n2}L_{n2}/L_{n1} = O\left(x_n^{-1/4}[\log x_n]^{3/4}\right) \to 0$, verifying the first statement of condition F.

Finally, note that $K'(2\theta) = K''(2\theta) = \mu e^{2\theta}$, and so using a similar approximation to that used in (20) above, we have

$$\begin{aligned}
\sup_{\theta \geq \theta_{n2}} \frac{v_n(\theta)K''(2\theta)}{K'(2\theta) - q_n} \quad &\sim \quad \sup_{\theta \geq \theta_{n2}} \frac{\left(\mu e^{2\theta}\right)^{x_n} \exp\left\{-\mu e^{2\theta}\right\}}{\Gamma(x_n)(\mu e^{2\theta} - x_n)} \frac{(\mu e^{2\theta})^{1/2}}{(\mu e^{2\theta} - x_n + 1)} \\
&\leq \quad \sup_{\theta \geq \theta_{n2}} \frac{\left(\mu e^{2\theta}\right)^{x_n + 1/2} \exp\left\{-\mu e^{2\theta}\right\}}{\Gamma(x_n)(\mu e^{2\theta_{n2}} - x_n)^2} \ .
\end{aligned}$$

The numerator is maximised over all θ at $\mu e^{2\theta} = x_n + 1/2$, and this, together with Stirlings approximation again, gives that the left-hand side of equation (17) is at most asymptotic to a constant times

$$\frac{\left(x_n + \frac{1}{2}\right)^{x_n + \frac{1}{2}} e^{-x_n - \frac{1}{2}} e^{x_n - 1}}{(x_n - 1)^{x_n - \frac{1}{2}} \ x_n \log x_n} = O(\log x_n)^{-1} \ .$$

Note that $v_{n2}^{1/2} L_{n1} \log L_{n1}$ has exact rate $x_n^{1/4}(\log x_n)^{3/4} \to \infty$, and so (17), and thus condition F, is easily satisfied.

6.3 Gamma

Now suppose F is the standard Gamma distribution on $(0,\infty)$ with shape parameter α. We take $\Theta_n = (-\theta_n, 1/2)$, where $\theta_n = [n^{1/2}(\log n)^{-3/2}]^{1/\alpha}$, as remarked on below Theorem 3.1. We claim that for $j=1,2$, with $\theta_{nj}^- = -\theta_n$, $\theta_{nj}^+ = [1-\alpha\Delta_n^{3-2j}/c_n]/2$ and $\Delta_n = (5/\alpha)\log\log n$, conditions D, E and F are satisfied. We use the same approach as in subsection 6.1. For convenience write $b_n = q_n^-$ and $c_n = q_n^+$. Again we have a linear combination of 15 terms of the form

$$\prod_{i=1}^{4}\int_{S_i} x^{a_i}\, dF_{\theta_i+\eta_i}(x) \ , \tag{21}$$

where this time S_i is either $(0,\infty)$ or $(0,b_n)\cup(c_n,\infty)$. Standard tail approximations for the gamma distribution (see, for example, Stewart (2002)) give that (note $dF_\theta(x) = x^{\alpha-1}(1-\theta)^\alpha e^{-x(1-\theta)}/\Gamma(\alpha)\,dx$)

$$\begin{aligned}
&\int_I x^a dF_{\theta+\eta}(x) \\
&= \begin{cases} O(1-\theta-\eta)^{-a} & \text{if } I=(0,\infty)\ , \\ O(1)b_n^a\{b_n(1-\theta-\eta)\}^\alpha & \text{if } I=(0,b_n) \ \text{ and} \\ O(1)c_n^a\{c_n(1-\theta-\eta)\}^{\alpha-1}\exp\{-c_n(1-\theta-\eta)\} & \text{if } I=(c_n,\infty)\ . \end{cases}
\end{aligned}$$

Now for $\theta,\eta\in\Theta_{n1}$, $(1-\theta-\eta)\geq 1-2\theta_{n1}^+ = c_n^{-1}\alpha\Delta_n$. So the above three bounds respectively reduce to

$$\begin{aligned}
&O(1)c_n^a(\log\log n)^{-a} && \\
O(1)b_n^a n^{-1/2}(\log n)^{-3/2} &= o(1)c_n^a n^{-1/2}(\log n)^{-3/2} \\
O(1)c_n^a\Delta_n^a\exp\{-\alpha\Delta_n\} &= O(1)c_n^a(\log\log n)^{\alpha-1}(\log n)^{-5}\ .
\end{aligned} \tag{22}$$

The largest a product of the form (21) can be is where three of the S_i's are $(0,\infty)$ and one is not. In this case, the product can be written as

$$O(1)c_n^4(\log\log n)^{\alpha-1}(\log n)^{-5} = O(1)(\log\log n)^{\alpha-1}(\log n)^{-1} = o(1) \qquad (23)$$

since $c_n = O(\log n)$ for any $\alpha > 0$. This proves that condition D is satisfied.

The limiting correlation function is $r_0(s,t) = \cosh(|s-t|/\alpha^{1/2})^{-\alpha}$, which again is the correlation function of a stationary process. Also, $r_0(s,t) = O(e^{-\alpha^{-1/2}|s-t|})$ as $|s-t| \to \infty$, and so Condition E is satisfied.

Note that, using the notation of condition F that $\Theta_{nj}^- = \emptyset$ for $j = 2, 3$, so we only need show the '+'-versions of the two statements appearing there. The transformation $\tau_0(\cdot)$ in this case is given by $\theta \mapsto -(\alpha^{1/2}/2)\log(1-2\theta)$. So

$$\begin{aligned} L_{n1} &= \tau_0(\theta_{n1}^+) - \tau_0(-\theta_n) = O(\log n) \ ; \\ L_{n2}^+ &= \tau_0(\theta_{n2}^+) - \tau_0(\theta_{n1}^+) = O(\log\log\log n) \ . \end{aligned}$$

Also, since $v_{n2}^+ = \int_{b_n}^{c_n} dF_{2\theta_{n2}^+}$, again, using standard lower-tail approximations for the gamma distribution, $v_{n2}^+ \sim \{c_n(1-2\theta_{n2}^+)\}^\alpha/\Gamma(\alpha+1)$ which has exact rate $(\log\log n)^{-\alpha}$. Since $(\log\log n)^{\alpha/2}\log\log\log n/\log n \to 0$ for any $\alpha > 0$, the first statement of condition F holds. Furthermore, the right hand side of (17) becomes $O(1)(\log n)(\log\log n)^{1-\alpha/2}$, which tends to infinity for any α. The left-hand side becomes a constant times

$$\sup_{\theta\in\Theta_{n3}^+} \frac{\int_{b_n}^{c_n} dF_{2\theta}}{1 - c_n(1-2\theta)} = O(1)$$

since $\sup_{\theta\in\Theta_{n3}} c_n(1-2\theta) \le c_n(1-2\theta_{n2}^+) \to 0$. Hence condition F is satisfied.

7 Proofs

7.1 Proof of Theorem 2.1

Lemma 7.1. *If h_n and ℓ_n satisfy $\ell_n \to \infty$, $h_n \to 0$ and $h_n \ell_n \to \infty$ and equation 4 in condition C holds, then defining $R_n^*(s)$ via*

$$P\left\{\sup(X_n(t), s \le t \le s + h_n) > \ell_n\right\} = \frac{h_n \exp\left\{-\ell_n^2/2\right\}}{2\pi} \left[1 + R_n^*(s)\right] \ ,$$

as $n \to \infty$ we have

$$\sup_{s, s+h_n \in T_{n1}} |R_n^*(s)| \to 0 \ . \tag{24}$$

Proof. Define the correlation functions

$$r_n^{\sup}(h) = \sup_{t, t+h \in T_{n1}} r_n(t, t+h) \quad \text{and} \quad r_n^{\inf}(h) = \inf_{t, t+h \in T_{n1}} r_n(t, t+h) \ .$$

Then by 4 we have, as $h \to 0$,

$$r_n^{\sup}(h) = 1 - h^2/2 + o(h^2) \ ,$$

and the same holds for $r_n^{\inf}(h)$. If we now define two sequences of stationary, standardised Gaussian processes $\{ X_n^{\sup}(t) \mid t \in T_{n1} \}$ and $\left\{ X_n^{\inf}(t) \mid t \in T_{n1} \right\}$ with respective correlation functions

$$E X_n^{\sup}(t) X_n^{\sup}(t+h) \equiv r_n^{\sup}(h) \quad \text{and} \quad E X_n^{\inf}(t) X_n^{\inf}(t+h) \equiv r_n^{\inf}(h)$$

then Theorem 2.1 of Hüsler (1990) and stationarity mean that (24) holds but with X_n replaced by both $X_n^{\sup}$ and $X_n^{\inf}$ in the definition of $R_n^*(s)$. But by two applications of Slepian's Lemma (see for instance Leadbetter *et al.* (1983), Theorem 7.4.2), the maximum of X_n is stochastically between the maxima of $X_n^{\sup}$ and $X_n^{\inf}$. This proves (24). □

Lemma 7.2. *Suppose that for a sequence of intervals* $\{ T_{n1} \mid n \in \mathbb{N} \}$*, the standardised Gaussian processes* $\{ X_n(t) \mid t \in T_{n1} \}$*, satisfy condition C. Then, for any level* $\ell_n \to \infty$*, if* $J_n = \lambda(T_{n1}) \exp\{-\ell_n^2/2\} / (2\pi)$ *satisfies* $\limsup J_n < \infty$ *then*

$$P\left\{ \sup_{0 \le t \le t_{n1}} X_n(t) \le \ell_n \right\} - \exp\{-J_n\} \to 0 \ . \tag{25}$$

Proof. The method of proof is almost identical to the proof of Theorem 4.1 in Hüsler (1990), taking $\alpha = 2$ throughout. The only notable difference is that our Lemma 7.1 takes the place of Hüsler's Theorem 2.2. Otherwise, the method of proof is identical; our equation (4) in condition C takes the place of Hüsler's conditions (5) and (6), equation (5) takes the place of Hüsler's condition (10), and since the boundary is constant Hüsler's smoothness conditions (8) and (9) are trivially satisfied. □

Lemma 7.3. *Suppose that condition C and the first statement of condition B hold for some sequence of intervals* $\{ T_{n1} \mid n \in \mathbb{N} \}$*. Then for any other sequence of intervals* $\{ T_{n0} \mid n \in \mathbb{N} \}$ *satisfying* $T_{n0} \subset T_{n1}$ *for each* n*, defining* $\ell_{n0} = \inf \{ b_n(t) \mid t \in T_{n0} \}$ *and* $J_{n0} = \lambda(T_{n0}) \exp\{-\ell_{n0}^2/2\} / (2\pi)$*, we have*

$$P\{X_n(t) \le b_n(t),\, t \in T_{n0}\} - \exp\{-J_{n0}\} \to 0 \ . \tag{26}$$

Proof. Write $m_{n0} = \sup \{ b_n(t) \mid t \in T_{n0} \}$. Then using the notation introduced in the statement of condition B, we have $\ell_{n1} \le \ell_{n0} \le m_{n0} \le m_{n1}$, and since

$m_{n1}/\ell_{n1} \to 0$ we have both

$$\ell_{n0}/\ell_{n1} \to 1 \tag{27}$$

and

$$m_{n0}/\ell_{n0} \to 1 \ . \tag{28}$$

Since $\lambda(T_{n0}) \le \lambda(T_{n1})$, (27) gives

$$\begin{aligned} \limsup J_{n0} &\le \limsup \lambda(T_{n1}) \exp\{-\ell_{n0}^2/2\} \\ &= \limsup \lambda(T_{n1}) \exp\{-\ell_{n1}^2/2\} [1+o(1)] < \infty \end{aligned} \tag{29}$$

from condition B, and since

$$\begin{aligned} P\left\{\sup_{t\in T_{n1}} X_n(t) \le \ell_{n0}\right\} &\le P\{X_n(t) \le b_n(t),\, t \in T_{n0}\} \\ &\le P\left\{\sup_{t\in T_{n1}} X_n(t) \le m_{n0}\right\} \ , \end{aligned}$$

(28) means that (25) holds with J_n replaced by J_{n0} and ℓ_n replaced by both ℓ_{n0} and m_{n0}, by condition C, (29) and Lemma 7.2. This proves (26). □

Lemma 7.4. *Suppose condition A and the sequence of intervals* $\{\, T_{n1} \mid n \in \mathbb{N} \,\}$ *satisfies the second statement of condition B. Then*

$$P\{X_n(t) \le b_n(t),\, t \in T_n \setminus T_{n1}\} \to 1 \ .$$

Proof. If the second statement of condition B holds then for each $j = 2, 3, \ldots, k$,

$$P\{X_n(t) \le b_n(t),\, t \in T_{nj}\} \ge P\{X_n(t) \le B_{nj}(t),\, t \in T_{nj}\} \ .$$

We intend to show that for each j the probability on the right-hand side tends to 1.

For each n, define $\ell_{nj} = \inf\{\, B_{nj}(t) \mid t \in T_{nj} \,\}$ and let $\{I_{njk}\}$ form a partition of T_{nj} of intervals of width $h_{nj} = (2\pi)^{1/2}/\ell_{nj}$. Condition A and equation 13.3.2 of Cramér & Leadbetter (1967) give us that for any $u > 0$,

$$EN_{u,n,j,k} = h_{nj} \exp\{-u^2/2\}/(2\pi) \;, \tag{30}$$

where $N_{u,n,j,k}$ is the number of upcrossings of the level u by the process $\{\, X_n(t) \mid t \in I_{njk} \,\}$. Assume without loss of generality that $\inf I_{nj} \in I_{njk}$. Since the processes are a.s. continous, we have

$$P\left\{\max_{t \in I_{njk}} X_n(t) > u\right\} \leq EN_{u,n,j,k} + P(X_n(\inf I_{nj}) > u) \;, \tag{31}$$

(see for instance Piterbarg (1996), Lemma F.1, page 28). Write now $\ell_{njk} = \inf_{t \in I_{njk}} B_{nj}(t)$.

$$\begin{aligned} P\left\{\sup_{t \in T_{nj}} [X_n(t) - B_{nj}(t)] > 0\right\} &\leq \sum_k P\left\{\sup_{t \in I_{njk}} [X_n(t) - B_{nj}(t)] > 0\right\} \\ &\leq \sum_k P\left\{\sup_{t \in I_{njk}} X_n(t) > \ell_{njk}\right\} \\ &\leq \sum_k \exp\{-\ell_{njk}^2/2\}\left[h_{nj}(2\pi)^{-1} + \ell_{njk}^{-1}(2\pi)^{-1/2}\right] \\ &\leq \pi^{-1} h_{nj} \sum_k \exp\{-\ell_{njk}^2/2\} \;, \end{aligned} \tag{32}$$

using (30), (31), the usual upper bound for the standard normal upper tail and the fact that each $\ell_{njk}^{-1} \leq \ell_{nj}^{-1} = (2\pi)^{-1/2} h_{nj}$. The last expression is π^{-1} times the upper Riemann sum based on intervals of width h_{nj} which approximates the integral $\int_{T_{nj}} \exp\{-B_{nj}^2(t)\}\, dt \to 0$ by condition B. Since the integrand is continuous and monotone over T_{nj}, the upper and lower Riemann sums approximating this integral only differ by at most two terms, corresponding to the endpoints of each interval, since the upper bound of the integrand over one interval is the lower bound for an adjacent interval. Thus the most the upper and lower Riemann sums could differ by is $2h_{nj}\exp\{-\ell_{nj}^2/2\} = O\left(\ell_{nj}^{-1}\exp\{-\ell_{nj}^2/2\}\right) \to 0$ since $\ell_{nj} \to \infty$ by condition B. Since any lower Riemann sum is at most the integral it approximates, this completes the proof. $\square$

Theorem 2.1 is proved by using Lemma 7.3 with $T_{n0} = T_{n1}$ and Lemma 7.4.

7.2 Proof of Theorem 4.1

Note firstly that since $0 \in T_{n1}$ for all n,

$$1 \geq \sup_{t \in T_{n1}} v_n[\theta_n(t)] \geq v_n(0) = 1 - 2/n \to 1 \ ,$$

and so writing $l_{n1}(x) = \inf_{t \in T_{n1}} b_{n,x}(t)$ we have

$$J_n(x) = \lambda(T_{n1})\exp\left\{-\ell_{n1}^2(x)\right\}/(2\pi) \sim \lambda(T_{n1})\exp\left\{-u_n^2(x)/2\right\}/(2\pi) \to e^{-x} \ .$$

This, together with the fact that

$$\left\{a_n\left[\sup_{\theta \in \Theta_n} \tilde{Z}_n(\theta) - a_n\right] + \log 2\pi \leq x\right\} = \{X_n(t) \leq b_{n,x}(t),\, t \in T_n\} \ ,$$

shows that the statement (16), with Z_n replaced by $\tilde{Z}_n$, follows from Theorem 2.1.

Next note that for any sequence of intervals $\{\, T_{n0} \mid n \in \mathbb{N} \,\}$ satisfying $T_{n0} \subset$

T_{n1}, with $a_{n0} = [2\log\lambda(T_{n0})]^{1/2}$ Lemma 7.3 implies that,

$$P\left\{a_{n0}\left[\sup_{t\in T_{n0}}\tilde{Z}_n(\theta_n(t)) - a_{n0}\right] + \log 2\pi \le x\right\} \to \exp\{-e^{-x}\} \ ,$$

and so $\sup_{t\in T_{n0}}\tilde{Z}_n(\theta_n(t)) = O_p(a_{n0})$. In particular, taking $T_{n0} \equiv T_{n1}$, we have by Theorem 2.1 that

$$\sup_{\theta\in\Theta_n}\tilde{Z}_n(\theta) = \sup_{t\in T_{n1}}\tilde{Z}_n(\theta_n(t)) = O_p(a_n) = O_p(\log\lambda(T_{n1}))^{1/2} \ ,$$

whereas if we define $T_{n0} = \left\{ t\in T_{n1} \,\middle|\, |t| \le \exp\left\{[\log\lambda(T_{n1})]^{1/2}\right\} \right\}$, we have

$$\sup_{t\in T_{n0}}\tilde{Z}_n(\theta_n(t)) = O_p(a_{n0}) = O_p(\log(T_{n0}))^{1/2} = O_p(\log\lambda(T_{n1}))^{1/4} \ .$$

So with probability tending to 1, $\sup_{\theta\in\Theta_n}\tilde{Z}_n(\theta) = \sup_{t\in T_n\setminus T_{n0}}\tilde{Z}_n(\theta_n(t))$.

Note that we can write $F_n g_\theta = \rho_n(0,\theta)$ and $\rho_n(0,\theta_n(t)) = r_n(0,t)$. Furthermore, $\sup_{t\in T\setminus T_{n0}} r_n(0,t) = o(\log\lambda(T_{n0}))^{-1}$ by equation (5) in condition C. So

$$\begin{aligned}
\sup_{t\in T_n\setminus T_{n0}}\tilde{Z}_n(\theta_n(t)) &= \sup_{t\in T_n\setminus T_{n0}}\left[Z_n(\theta_n(t)) + O_p(1)\, r_n(0,t)\right] \\
&= \sup_{t\in T_n\setminus T_{n0}} Z_n(\theta_n(t)) + o_p(\log\lambda(T_{n1}))^{-1/2} \ .
\end{aligned}$$

Finally, note that $\sup_{t\in T_{n0}} Z_n(\theta_n(t)) = \sup_{t\in T_{n0}}\tilde{Z}_n(\theta) + O_p(1) = O_p(\log\lambda(T_{n1}))^{1/4}$. Combining these gives that with probability tending to 1, $\sup_{\theta\in\Theta_n} Z_n(\theta) = \sup_{\theta\in\Theta_n}\tilde{Z}_n(\theta) + o_p(a_n)^{-1}$, and so

$$a_n\left[\sup_{\theta\in\Theta_n} Z_n(\theta) - a_n\right] = a_n\left[\sup_{\theta\in\Theta_n}\tilde{Z}_n(\theta) - a_n\right] + o_p(1)$$

and so (16) holds with $\tilde{Z}_n$ replaced by Z_n.

7.3 Proof of Theorem 5.1

Firstly note that by condition F each Θ_{n1}, and thus each T_{n1}, contains 0.

Conditions D and E imply C

Differentiation yields that

$$\begin{aligned} r_n^{(0,2)}(t,s) &= \rho_n^{(1,1)}\theta_n(s),\theta_n(s))^{-1}\rho_n^{(0,2)}\theta_n(t),\theta_n(s)) - \\ &\quad \rho_n^{(1,1)}\theta_n(s),\theta_n(s))^{-2}\rho_n^{(0,1)}\theta_n(s),\theta_n(s))\rho_n^{(1,2)}\theta_n(s),\theta_n(s)) \end{aligned} \tag{33}$$

Let $E_{nj}(\theta) = \int_{A_n} x^j \, dF_\theta(x) \Big/ \int_{A_n} dF_\theta$, denote the j-th conditional moment of $X \sim F_\theta$ given $X \in A_n$. Then $\rho_n^{(0,1)}(\theta,\eta) = \rho_n(\theta,\eta)\,[E_{n1}(\theta+\eta) - E_{n1}(2\eta)]$. Note also that the derivatives with respect to θ satisfy $E'_{nj}(\theta) = E_{n(j+1)}(\theta) - E_{nj}(\theta)E_{n1}(\theta)$ for all $j \geq 1$. Repeated differentiation yields that $\rho_n^{(i,j)}(\theta,\eta)$ equals $\rho(\theta,\eta)$ times a linear combination of terms of the form $\prod_{k=1}^{i+j} E_{na_k}^{b_k}(\theta_k)$, where $\sum_k a_k b_k \leq (i+j)$ and each θ_k equals either $\theta+\eta$, 2θ or 2η.

So the first term on the right hand side of (33) is $\rho(\theta_n(t),\theta_n(s))$ times a ratio of such linear combinations, with $i+j=2$; the numerator is evaluated at values $\theta_n(s)+\theta_n(t)$, $2\theta_n(t)$ or $2\theta_n(s)$ but the denominator is only evaluated at $2\theta_n(s)$. The second term is simply a ratio of such linear combinations with $i+j=4$ evaluated only at $2\theta_n(s)$.

Suppose condition D holds for some $\{\,\Theta_{n1} \mid n \in \mathbb{N}\,\}$. Then $r_n^{(0,2)}(t,s) = r_0^{(0,2)}(t,s)[1+o(1)]$, with $o(1)$ uniform over $t,s \in T_{n1} = \{\,\tau_n(\theta) \mid \theta \in \Theta_{n1}\,\}$. A fortiori, we have also $r_n(t,s) = r_0(t,s)[1+o(1)]$ with the same uniformity properties. Thus for any $h_n \to 0$ we have

$$\begin{aligned}&\sup_{t,t+h_n\in T_{n1}}\left|r_n^{(0,2)}(t,t+h_n)/r_0^{(0,2)}(t,t+h_n)-1\right|\\&\quad\leq\sup_{t,s\in T_{n1}}\left|r_n^{(0,2)}(t,t+s)/r_0^{(0,2)}(t,t+s)-1\right|\to 0\ ,\end{aligned}$$

and the first statement of condition E implies that for any $h_n \to 0$,

$$r_n^{(0,2)}(t,t+h_n) = r_0^{(0,2)}(t,t+h_n)[1+o(1)] = -1+o(1)$$

with $o(1)$ uniform over t such that t and $t+h_n$ are in T_{n1} defined above. Finally, since $r_n^{(0,1)}(t,t) \equiv 0$, a Taylor series expansion with mean-value remainder gives that $r_n(t,t+h_n) = 1 - r_n^{(0,2)}(t,t+h_n^*)$, for h_n^* between 0 and h_n. Equation (34) then implies that equation (4) in condition C holds. In the same way, with appropriately uniform $o(1)$, $r_n(t,s) = r_0(t,s)[1+o(1)]$ and so the second statement of condition E then implies that equation (5) holds.

Conditions D and F imply B

Two consequences of D are that

$$v_n(\theta) = 1 + o(1) \tag{34}$$

and

$$\gamma_n(\theta) = K''(2\theta)[1+o(1)]\ , \tag{35}$$

with $o(1)$ uniform over $\theta \in \Theta_{n1}$. The first of these implies that for each x, the boundary functions defined in Theorem 4.1, with $T_{n1} = \{\, \tau_n(\theta) \mid \theta \in \Theta_{n1}\,\}$

satisfy the first statement of condition B. Note also that (35) implies that

$$\begin{aligned}
\lambda(T_{n1}) &= \tau_n(\theta_{n1}^+) - \tau_n(\theta_{n1}^-) = \int_{\theta_{n1}^-}^{\theta_{n1}^+} \gamma_n^{1/2}(\theta)\,d\theta = \int_{\theta_{n1}^-}^{\theta_{n1}^+} K''(2\theta)^{1/2}\,d\theta[1+o(1)] \\
&= \left[\tau_0(\theta_{n1}^+) - \tau_0(\theta_{n1}^-)\right][1+o(1)]\ . \qquad (36)
\end{aligned}$$

We prove the second statement of condition B in two parts. The minimum boundary height tends to infinity by construction, so we only need to verify the second statement. We partition $T_n = \{\,\tau_n(\theta)\mid\theta\in\Theta_n\,\}$ into T_{n1}, $T_{n2}^+ = T_n\cap\{\,\tau_n(\theta)\,\big|\,\theta_{n1}^+<\theta<\theta_{n2}^+\,\}$, $T_{n3}^+ = T_n\cap\{\,\tau_n(\theta)\,\big|\,\theta\in\Theta_n,\,\theta\ge\theta_{n2}^+\,\}$, $T_{n2}^- = T_n\cap\{\,\tau_n(\theta)\,\big|\,\theta_{n2}^-<\theta<\theta_{n1}^-\,\}$, and $T_{n3}^- = T_n\cap\{\,\tau_n(\theta)\,\big|\,\theta\in\Theta_n,\,\theta\le\theta_{n2}^-\,\}$.

The argument for T_{nj}^+ also applies to T_{nj}^-, for $j=2,3$. We assume for simplicity that T_{n2}^+ and T_{n3}^+ are both non-empty. Write $v_{n2}^+ = \inf_{\theta\in\Theta_{n2}^+} v_n(\theta)$. We henceforth drop the $+$ superscript. We use the inequality $\mathrm{Var}X \ge E\mathrm{Var}(X|A) \ge \mathrm{Var}(X|A)P(A)$, which when $A=\{X\in A_n\}$ and unconditionally $X\sim F_{2\theta}$ reduces to

$$\gamma_n(\theta) \le K''(2\theta)/v_n(\theta)\ . \qquad (37)$$

Using this gives

$$\begin{aligned}
\lambda(T_{n2}) = \tau_n(\theta_{n2}) - \tau_n(\theta_{n1}) = \int_{\theta_{n1}}^{\theta_{n2}} \gamma_n^{1/2}(\theta)\,d\theta &\le v_{n2}^{-1/2}\int_{\theta_{n1}}^{\theta_{n2}} K''(2\theta)^{1/2}\,d\theta \\
&= v_{n2}^{-1/2}\left[\tau_0(\theta_{n2}) - \tau_0(\theta_{n1})\right]\ .
\end{aligned}$$

On T_{n2}, $b_n(\cdot)$ is bounded below by the constant (thus montone and continuous) function $B_{n2}\equiv u_n(x)$. Since $\exp\{-u_n^2(x)/2\} = \lambda(T_{n1})^{-1}$, the previous display, together with (36) give that $\int_{T_{n2}}\exp\{-B_{n2}^2(t)/2\}\,dt\to 0$.

For $\theta>\theta_n^+$, the mean of $F_{2\theta}$ given by $K'(2\theta)$ is strictly greater than q_n^+.

However, the derivative

$$
\begin{aligned}
v_n'(\theta) &= \int_{A_n} 2[x - K'(2\theta)]\, dF_{2\theta} \\
&= 2v_n(\theta)\left\{\left[\int_{A_n} dF_{2\theta}\right]^{-1}\int_{A_n} x\, dF_{2\theta} - K'(2\theta)\right\} \\
&\leq 2v_n(\theta)[q_n^+ - K'(2\theta)]\ , \qquad (38)
\end{aligned}
$$

since the first term inside the curly brackets is the conditional mean of X given $X \in A_n$, which is at most q_n^+. This is strictly negative for any $\theta > \theta_n^+$. Thus

$$
v_n'(\theta) < 0 \quad \text{for} \quad \theta \geq \theta_{n2}\ . \qquad (39)
$$

Consequently, the continuous function $t \mapsto b_n(t)$ is strictly increasing over T_{n3}. Hence

$$
\begin{aligned}
\int_{T_{n3}} \exp\{-b_n^2(t)/2\}\ dt &\leq \int_{\theta_{n2}}^{\theta_n(\infty)} \gamma_n^{1/2}(\theta)\exp\{-u_n^2(x)v_n^{-1}(\theta)/2\}\ d\theta \\
&\leq \int_{\theta_{n2}}^{\theta_n(\infty)} v_n^{-1/2}(\theta)K''(2\theta)\exp\{-u_n^2(x)v_n^{-1}(\theta)/2\}\ d\theta
\end{aligned}
$$

using (37). Inequality (39) implies that the function $\theta \mapsto 1/v_n(\theta)$ is strictly increasing. Make a change of variable in the last integral to $y = 1/v_n(\theta)$. If we write $y \mapsto w_n(y)$ as the inverse of this transformation, then $d\theta = w_n'(y)dy$, where $w_n'(y) = -v_n^2(w_n(y))/v_n'(w_n(y))$. This integral is at most

$$
\begin{aligned}
&\int_{1/v_n(\theta_{n2})}^{\infty} K''(2w_n(y))^{1/2}y^{1/2}e^{-u_n^2(x)y/2}w_n'(y)\, dy \\
&\quad\leq\ 2u_n^{-3}(x)\int_{u_n^2(x)/[2v_n(\theta_{n2})]}^{\infty} z^{1/2}e^{-z}\, dz \sup_{y\geq 1/v_n(\theta_{n2})} K''(2w_n(y))w_n'(y) \qquad (40)
\end{aligned}
$$

An integration by parts approach (see Barndorff-Nielsen & Cox (1989) or the appendix of Stewart (2002)) gives that as $c \to \infty$, $\int_c^\infty z^{1/2}e^{-z}\, dz = O(c^{1/2}e^{-c})$.

Implementing this enables us to rewrite the integral in (40) as

$$O(1)u_n(x)v_{n2}^{-1/2}\exp\left\{-u_n^2(x)v_{n2}^{-1}/2\right\} \ .$$

Also the supremum in (40) can be written as

$$\sup_{\theta\geq\theta_{n2}} K''(2\theta)^{1/2}v_n^2(\theta)/[-v_n'(\theta)] \leq \sup_{\theta\geq\theta_{n2}} K''(2\theta)^{1/2}v_n(\theta)/\left\{2[K'(2\theta)-q_n^+]\right\}$$

Combining permits us to rewrite (40) as

$$O(1)u_n^{-2}(x)v_{n2}^{-1/2}\exp\left\{-u_n^2(x)v_{n2}^{-1}/2\right\}\sup_{\theta\geq\theta_{n2}} K''(2\theta)^{1/2}v_n(\theta)/\left\{2[K'(2\theta)-q_n^+]\right\} \quad (41)$$

Since $1\geq v_{n2}\to 0$ by condition F, $0\leq v_{n2}^{-1}-1\to\infty$ and so

$$\exp\left\{-u_n^2(x)v_{n2}^{-1}/2\right\} = \exp\left\{-u_n^2(x)/2\right\}\exp\left\{-u_n^2(x)[v_{n2}^{-1}-1]/2\right\} = o(\lambda(\tau_0(\Theta_{n1})))^{-1} \ ,$$

since $\exp\left\{-u_n^2(x)/2\right\} = O(\lambda(T_{n1}))^{-1} = O(L_{n1})^{-1}$, where $L_{n1}=\lambda(\tau_0(\Theta_{n1}))$. Finally, the supremum in (41) is $O(v_{n2}^{1/2}L_{n1}\log L_{n1})$ by condition F. This completes the proof.

References

Barndorff-Nielsen, O. 1978. *Information and Exponential Families in Statistical Theory.* Wiley series in probability and mathematical statistics. John Wiley and Sons.

Barndorff-Nielsen, O. 1980. Exponential Families. *Memoirs, Department of Theoretical Statistics, Institute of Mathematics, University of Aarhus.*

Barndorff-Nielsen, O. E., & Cox, D. R. 1989. *Asymptotic Techniques for Use in Statistics.* Monographs on Statistics and Applied Probability. Chapman and Hall.

Berman, S. 1974. Sojourns and extremes of Gaussian processes. *Annals of Probability*, **2**, 999–1026.

Bickel, P., & Chernoff, H. 1993. Asymptotic distribution of the likelihood ratio statistic in a prototypical non regular problem. *Pages 83–96 of:* Ghosh, J. K., Mitra, S. K., Parthasararthy, K. R., & Prakasa Rao, B. L. S. (eds), *Statistics and Probability: A Raghu Raj Bahadur Festschrift*. Wiley Eastern Limited.

Cramér, H., & Leadbetter, M. R. 1967. *Stationary and Related Stochastic Processes*. John Wiley and Sons.

Csörgő, M., Csörgő, S., Horváth, L., & Mason, D.M. 1986a. Sup-norm convergence of the empirical process indexed by functions and applications. *Probability and Mathematical Statistics*, **7**, 13–26.

Csörgő, M., Csörgő, S., Horváth, L., & Mason, D. M. 1986b. Weighted empirical and quantile processes. *Annals of Probability*, **14**(1), 31–85.

Csörgő, S. 1982. The empirical moment generating function. *Pages 139–150 of:* Gnedenko, B.V., Puri, M.L., & Vincze, I. (eds), *Nonparametric Statistical Inference*, vol. 1. Amsterdam: North-Holland, for János Bolyai Mathematical Society.

Dacunha-Castelle, D., & Gassiat, E. 1999. Testing the order of a model using locally conic parametrization: population mixtures and stationary ARMA processes. *Annals of Statistics*, **27**, 1178–1209.

Epps, T.W., Singleton, L.B., & Pulley, L.B. 1982. A test of separate families of distributions based on the empirical moment generating function. *Biometrika*, **69**(2), 391–399.

Ghosh, Sucharita. 1996. A new graphical tool to detect non-normality. *J. Roy. Statist. Soc. Ser. B*, **58**(4), 691–702.

Hartigan, J. A. 1985. A failure of likelihood ratio asymptotics for normal mixtures. *In: Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer*.

Hüsler, J. 1990. Extreme values and high boundary crossings of locally stationary Gaussian processes. *Annals of Probability*, **18**(3), 1141–1158.

Hüsler, J. 1995. A note on extreme values of locally stationary Gaussian processes. *Journal of Statistical Planning and Inference*, 203–213.

Leadbetter, M.R., Lindgren, G., & Rootzén, H. 1983. *Extremes and Related Properties of Random Sequences and Processes*. Springer-Verlag.

Liu, X., Pasarica, C., & Shao, Y. 2003. Testing homogeneity in gamma mixture models. *Scaninavian Journal of Statistics*, **30**, 227–240.

Piterbarg, V. I. 1996. *Asymptotic Methods in the Theory of Gaussian Processes and Fields*. American Mathematical Society.

Stewart, M.I. 2002. *Asymptotic methods for tests of homogeneity in finite mixture models*. Ph.D. thesis, University of Sydney.

van der Vaart, A. 1996. New Donsker Classes. *Annals of Probability*, **24**(4), 2128–2140.

van der Vaart, A. W., & Wellner, J. A. 1996. *Weak Convergence and Empirical Processes*. Springer Series in Statistics. Springer.