Approximations and limit theorems for log-likelihood ratio processes

Alexander Gushchin, Esko Valkeila

Le Mans, March 16, 2009

Setting of the problem

Consider a sequence

$$\mathbb{E}^n = (\Omega^n, \mathcal{F}^n, \mathbb{F}^n = (\mathcal{F}^n_t)_{t\in\mathbb{R}_+}, (P^{n,\vartheta})_{\vartheta\in\mathbb{R}^k})$$

of filtered statistical models. Denote $P^n := P^{n,0}$. If T is an $\mathbb{F}^n$-stopping time then $P^n_T := P^n\Big|_{\mathcal{F}^n_T}$, $P^{n,\vartheta}_T := P^{n,\vartheta}\Big|_{\mathcal{F}^n_T}$.

For simplicity we shall assume that for all n and ϑ

$$P^{n,\vartheta}_0 = P^n_0.$$

Let $Z^{n,\vartheta}$ be the density process of $P^{n,\vartheta}$ with respect to P^n, i.e. a càdlàg $\mathbb{F}^n$-adapted process $(Z_t^{n,\vartheta})_{t\in\mathbb{R}_+}$ with values in $\mathbb{R}_+$ such that, for any $\mathbb{F}^n$-stopping time T, $Z_T^{n,\vartheta}$ is the density of the absolutely continuous part of $P_T^{n,\vartheta}$ with respect to P_T^n. Then $Z^{n,\vartheta}$ is a P^n-supermartingale and

$$P^n(\sup_t Z_t^{n,\vartheta} \geq a) \leq a^{-1} \qquad \text{for any} \quad a > 0 \quad \text{and} \quad \vartheta \in \Theta.$$

Let T_n be an $\mathbb{F}^n$-stopping time for every n.

Denote by $\mathcal{W} = \mathcal{W}(\{T_n\})$ the set of all sequences $\{w^n\}$ with the following properties: $w^n = (w^{n,1}, \ldots, w^{n,k})$, $w_0^n = 0$, is a locally square-integrable martingale on $(\Omega^n, \mathcal{F}^n, \mathbb{F}^n, P^n)$ with values in $\mathbb{R}^k$ for each n, the quadratic characteristic $\langle w^n, w^n \rangle_{T_n}$ is bounded in P^n-probability, i.e.

$$\text{the sequence} \quad \left(\sum_{i=1}^{k} \langle w^{n,i}, w^{n,i} \rangle_{T_n} \Big| P^n \right) \quad \text{is } \mathbb{R}\text{-tight,}$$

and the Lindeberg-type condition on jumps of w^n holds, i.e.

$$\|x\|^2 \mathbf{1}_{\{\|x\|>\varepsilon\}} \star \nu_{T_n}^{w^n} \xrightarrow{P^n} 0, \quad n \to \infty, \quad \text{for all } \varepsilon > 0.$$

Our **objective** is to find conditions that imply the existence of a sequence $\{w^n\} \in \mathcal{W}$ such that (at least)

$$\sup_{s \le T_n} \left| \log Z_s^{n,\vartheta_n} - \left(\vartheta_n^\top w_s^n - \frac{1}{2} \vartheta_n^\top \left\langle w^n, w^n \right\rangle_s \vartheta_n \right) \right| \xrightarrow{P^n} 0, \quad n \to \infty,$$

for each bounded sequence $\{\vartheta_n\}$ in $\mathbb{R}^k$.

For brevity, if such a sequence exists, we say that there is a quadratic approximation and call $\{w^n\}$ a central sequence.

Example: i.i.d. observations with density $p_\vartheta(x)$, $x \in \mathbb{R}$, $\vartheta \in \Theta$, where Θ is an open subset of $\mathbb{R}$.

Take $\Omega^n = \mathbb{R}^n$, $\mathcal{F}^n = \mathcal{B}(\mathbb{R}^n)$. Let $\vartheta_0 \in \Theta$ and $\varphi_n = \varphi_n(\vartheta_0)$ be a sequence with $\varphi_n \to 0$. If $\vartheta_0 + \varphi_n \vartheta \in \Theta$, define $P^{n,\vartheta}$ as the measure with the density

$$\prod_{i=1}^{n} p_{\vartheta_0+\varphi_n\vartheta}(x_i).$$

Otherwise, $P^{n,\vartheta}$ is arbitrary. There are two ways to introduce the filtration $\mathbb{F}^n$ and the stopping times T_n:

(1) $\mathcal{F}^n_t = \mathcal{G}^n_{[t]\wedge n}$, $T_n = n$, (2) $\mathcal{F}^n_t = \mathcal{G}^n_{[nt]\wedge n}$, $T_n = 1$,

where $\mathcal{G}^n_m$ be the σ-field in $\mathbb{R}^n$ generated by the first m coordinates.

If $\vartheta \rightsquigarrow p_\vartheta^{1/2}(x)$ is differentiable in quadratic mean (DQM) at ϑ_0, $g(x)$ is the (appropriately defined) score function, $J = \int g^2(x) p_{\vartheta_0}(x)\, dx$ is the Fisher information, $\varphi_n = n^{-1/2}$,

$$S_n = \varphi_n \sum_{i=1}^{n} g(x_i) \tag{1}$$

then

$$\log Z_{T_n}^{n,\vartheta_n} - \left(\vartheta_n S_n - \frac{1}{2} \vartheta_n^2 J \right) \xrightarrow{P^n} 0, \quad n \to \infty, \tag{2}$$

for each bounded sequence $\{\vartheta_n\}$,

$$\text{Law}(S_n | P^n) \to \mathcal{N}(0, J), \tag{3}$$

and we have (uniform) local asymptotic normality (if $J > 0$).

Let

$$w_t^n = \varphi_n \sum_{i=1}^{[nt]\wedge n} g(x_i)$$

(we choose the second way to define $\mathbb{F}^n$ and T_n). Then w^n is a P^n-square-integrable martingale, $\langle w^n, w^n\rangle_t = \left(\frac{[nt]}{n} \wedge 1\right) J$, the Lindeberg-type condition for jumps is the usual Lindeberg condition and is trivially satisfied, and $\{w^n\}$ is a central sequence in our sense.

If DQM is not satisfied, it may happen that (1), (2), and (3) are still valid with $\varphi_n = n^{-1/2}L(n)$, where $L(n)$ is a slowly varying sequence with $L(n) \to 0$. A typical case is when the score function $g(x)$ considered as a r.v. on $(\mathbb{R}, \mathcal{B}(\mathbb{R}), p_{\vartheta_0}(x)\,dx)$ is not square-integrable but belongs to the domain of attraction of the normal law.

Moreover, it can happen that there does not exist a function g such that (2) and (3) hold with

$$S_n = a_n \sum_{i=1}^{n} g(x_i).$$

However, at the same time, (2) and (3) are valid with

$$S_n = \sum_{i=1}^{n} g_n(x_i)$$

An example of such a case was given in Pfanzagl (2002):

$$p_\vartheta(x) = \frac{1}{4}\exp(-|x-\vartheta|^{1/2}), \qquad \varphi_n = (n \log n)^{-1/2}.$$

What **follows** from the existence of a quadratic approximation:

What **follows** from the existence of a quadratic approximation:

- for every bounded sequence ϑ_n

$$(P^n_{T_n}) \lhd (P^{n,\vartheta_n}_{T_n});$$

What **follows** from the existence of a quadratic approximation:

• for every bounded sequence ϑ_n

$$(P^n_{T_n}) \lhd (P^{n,\vartheta_n}_{T_n});$$

• If $\langle w^n, w^n \rangle_{T_n}$ converges in P^n-probability to a deterministic matrix J, then

$$\text{Law}(w^n_{T_n} | P^n) \to \mathcal{N}(0, J),$$

$$(P^n_{T_n}) \lhd \rhd (P^{n,\vartheta_n}_{T_n})$$

for every bounded sequence ϑ_n, and we have asymptotic normality (if J is nonsingular);

• If $T_n \equiv t$, $\langle w^n, w^n \rangle_s$ converges in P^n-probability to a deterministic matrix J_s for every $s \in [0, t]$, and $s \rightsquigarrow J_s$ is continuous, then, additionally,

$$\text{Law}(w^n | P^n) \to \text{Law}(w) \quad \text{in } \mathbb{D}([0,t], \mathbb{R}^k)$$

and

$$\text{Law}(\log Z^{n,\vartheta_n} | P^n) \to \text{Law}(\vartheta^\top w - 1/2 \vartheta^\top J \vartheta^\top) \quad \text{in } \mathbb{D}([0,t], \mathbb{R}^k),$$

for every bounded sequence $\vartheta_n \to \vartheta$, where $w = (w_s)_{s \le t}$ is a k-dimensional Gaussian martingale with covariance matrix J_s;

• If $T_n \equiv t$, $\langle w^n, w^n \rangle_s$ converges in P^n-probability to a deterministic matrix J_s for every $s \in [0, t]$, and $s \rightsquigarrow J_s$ is continuous, then, additionally,

$$\text{Law}(w^n | P^n) \to \text{Law}(w) \quad \text{in } \mathbb{D}([0,t], \mathbb{R}^k)$$

and

$$\text{Law}(\log Z^{n,\vartheta_n} | P^n) \to \text{Law}(\vartheta^\top w - 1/2\vartheta^\top J \vartheta^\top) \quad \text{in } \mathbb{D}([0,t], \mathbb{R}^k),$$

for every bounded sequence $\vartheta_n \to \vartheta$, where $w = (w_s)_{s \le t}$ is a k-dimensional Gaussian martingale with covariance matrix J_s;

• Some other limit theorems can be deduced in a similar way (e.g., asymptotic mixed normality);

•

$$\sup_{s \le T_n} |\Delta \log Z_s^{n,\vartheta_n}| \xrightarrow{P^n} 0, \quad n \to \infty,$$

for every bounded sequence ϑ_n. This condition, called ANJ (asymptotic negligibilility of jumps) below, is necessary for the existence of a quadratic approximation in our sense and indicates a range of applicability of our methods.

Basic ingredients

Put $Y^{n,\vartheta} := \sqrt{Z^{n,\vartheta}}$ and define the stochastic logarithm $y^{n,\vartheta} := \mathcal{L}og(Y^{n,\vartheta})$ on $\Gamma^{n,\vartheta} = \{Z_{-}^{n,\vartheta} > 0\}$ by

$$y_t^{n,\vartheta} = \int_0^t \frac{dY_s^{n,\vartheta}}{Y_{s-}^{n,\vartheta}}.$$

Basic ingredients

Put $Y^{n,\vartheta} := \sqrt{Z^{n,\vartheta}}$ and define the stochastic logarithm $y^{n,\vartheta} := \mathcal{L}og(Y^{n,\vartheta})$ on $\Gamma^{n,\vartheta} = \{Z_-^{n,\vartheta} > 0\}$ by

$$y_t^{n,\vartheta} = \int_0^t \frac{dY_s^{n,\vartheta}}{Y_{s-}^{n,\vartheta}}.$$

The process $y^{n,\vartheta}$ is a P^n-local supermartingale on $\Gamma^{n,\vartheta}$ and has the Doob–Meyer decomposition

$$y^{n,\vartheta} = m^{n,\vartheta} - h^{n,\vartheta},$$

where $m^{n,\vartheta}$ is a P^n-local martingale on $\Gamma^{n,\vartheta}$ and $h^{n,\vartheta}$ is a predictable increasing process which is the Hellinger process of order $1/2$ for P^n and $P^{n,\vartheta}$.

Define also $\overline{h}^{n,\vartheta,\eta}$ as the predictable increasing process in the Doob–Meyer decomposition of $\mathcal{L}og(Y^{n,\vartheta}Y^{n,\eta})$. In general, $\overline{h}^{n,\vartheta,\eta}$ is not the Hellinger process of order 1/2 for $P^{n,\vartheta}$ and $P^{n,\eta}$ but it is if $\eta = 0$:

$$\overline{h}^{n,\vartheta,0} = h^{n,\vartheta},$$

or if $P^{n,\vartheta} \ll P^n$ and $P^{n,\eta} \ll P^n$.

Define also $\overline{h}^{n,\vartheta,\eta}$ as the predictable increasing process in the Doob–Meyer decomposition of $\mathcal{L}og(Y^{n,\vartheta}Y^{n,\eta})$. In general, $\overline{h}^{n,\vartheta,\eta}$ is not the Hellinger process of order 1/2 for $P^{n,\vartheta}$ and $P^{n,\eta}$ but it is if $\eta = 0$:

$$\overline{h}^{n,\vartheta,0} = h^{n,\vartheta},$$

or if $P^{n,\vartheta} \ll P^n$ and $P^{n,\eta} \ll P^n$.

Denote $\iota^{n,\vartheta} := \overline{h}^{n,\vartheta,\vartheta}$. $\iota^{n,\vartheta}$ is the Hellinger process of order 0 since it is the predictable increasing process in the Doob–Meyer decomposition of $\mathcal{L}og(Z^{n,\vartheta})$. Of course, $\iota^{n,\vartheta} \equiv 0$ if $P^{n,\vartheta} \ll P^n$.

Lemma 1 *The process* $m^{n,\vartheta}$ *is a* P^n*-locally square-integrable martingale on* $\Gamma^{n,\vartheta}$*. On* $\Gamma^{n,\vartheta} \cap \Gamma^{n,\eta}$ *we have*

$$\overline{h}^{n,\vartheta,\eta} = h^{n,\vartheta} + h^{n,\eta} - \langle m^{n,\vartheta}, m^{n,\eta} \rangle - [h^{n,\vartheta}, h^{n,\eta}].$$

In particular, on $\Gamma^{n,\vartheta}$,

$$\langle m^{n,\vartheta}, m^{n,\vartheta} \rangle = 2h^{n,\vartheta} - [h^{n,\vartheta}, h^{n,\vartheta}] - \iota^{n,\vartheta}.$$

Lemma 1 *The process* $m^{n,\vartheta}$ *is a* P^n*-locally square-integrable martingale on* $\Gamma^{n,\vartheta}$*. On* $\Gamma^{n,\vartheta} \cap \Gamma^{n,\eta}$ *we have*

$$\overline{h}^{n,\vartheta,\eta} = h^{n,\vartheta} + h^{n,\eta} - \langle m^{n,\vartheta}, m^{n,\eta} \rangle - [h^{n,\vartheta}, h^{n,\eta}].$$

In particular, on $\Gamma^{n,\vartheta}$,

$$\langle m^{n,\vartheta}, m^{n,\vartheta} \rangle = 2h^{n,\vartheta} - [h^{n,\vartheta}, h^{n,\vartheta}] - \iota^{n,\vartheta}.$$

Example: Let $(X_i, \mathcal{X}_i, (Q_i^\vartheta)_{\vartheta \in \mathbb{R}^k})$ be statistical models, $i = 1, 2, \ldots$. Put

$$\Omega = \prod X_i, \quad \mathcal{F} = \bigotimes \mathcal{X}_i, \quad P^\vartheta = \prod Q_i^\vartheta, \quad P = P^0.$$

$\mathcal{F}_t$ is generated by the projections on the first $[t]$ coordinates. Let $\alpha_i(\vartheta)$ be the density of the absolutely continuous part of Q_i^ϑ with respect to Q_i^0. Then

$$y_t^{\vartheta} = \sum_{m=1}^{[t]} \Big(\sqrt{\alpha_i(\vartheta)} - 1\Big),$$

$$h_t^{\vartheta} = \sum_{m=1}^{[t]} E\Big(1 - \sqrt{\alpha_i(\vartheta)}\Big) = \sum_{m=1}^{[t]} \rho^2(Q_i^0, Q_i^{\vartheta}),$$

$$m_t^{\vartheta} = \sum_{m=1}^{[t]} \Big(\sqrt{\alpha_i(\vartheta)} - E\sqrt{\alpha_i(\vartheta)}\Big), \qquad \iota_t^{\vartheta} = \sum_{m=1}^{[t]} \Big(1 - E\alpha_i(\vartheta)\Big),$$

$$h_t^{\vartheta,\eta} = \sum_{m=1}^{[t]} E\Big(1 - \sqrt{\alpha_i(\vartheta)\alpha_i(\eta)}\Big).$$

Assumptions: For all bounded sequences $\{\vartheta_n\}$ and $\{\eta_n\}$

(O) $\iota_{T_n}^{n,\vartheta_n} \xrightarrow{P^n} 0, \quad n \to \infty.$

Assumptions: For all bounded sequences $\{\vartheta_n\}$ and $\{\eta_n\}$

(O) $\iota_{T_n}^{n,\vartheta_n} \xrightarrow{P^n} 0, \quad n \to \infty.$

(H) There is a sequence $\{H^n\}$ of symmetric positive-definite matrices such that that

$$\text{the sequence} \quad (\operatorname{tr} H^n | P^n) \quad \text{is} \quad \mathbb{R}\text{-tight.}$$

and

$$\overline{h}_{T_n}^{n,\vartheta_n,\eta_n} - (\vartheta_n - \eta_n)^\top H^n (\vartheta_n - \eta_n) \xrightarrow{P^n} 0, \quad n \to \infty.$$

Assumptions: For all bounded sequences $\{\vartheta_n\}$ and $\{\eta_n\}$

(O) $\iota_{T_n}^{n,\vartheta_n} \xrightarrow{P^n} 0, \quad n \to \infty.$

(H) There is a sequence $\{H^n\}$ of symmetric positive-definite matrices such that that

$$\text{the sequence} \quad (\operatorname{tr} H^n | P^n) \quad \text{is} \quad \mathbb{R}\text{-tight.}$$

and

$$\overline{h}_{T_n}^{n,\vartheta_n,\eta_n} - (\vartheta_n - \eta_n)^\top H^n (\vartheta_n - \eta_n) \xrightarrow{P^n} 0, \quad n \to \infty.$$

(L)

$$x^2 \mathbf{1}_{\{|x|>\varepsilon\}} \star \nu_{T_n}^{m^{n,\vartheta_n}} \xrightarrow{P^n} 0, \quad n \to \infty, \quad \text{for all } \varepsilon > 0.$$

The basic tool is

Theorem 1 (Gushchin and Valkeila (2003)) *If assumptions* (O), (H), *and* (L) *are satisfied for a fixed sequence* $\{\vartheta_n\}$ (*and* $\eta_n \equiv 0$) *then*

$$(P^n_{T_n}) \lhd (P^{n,\vartheta_n}_{T_n})$$

and

$$\sup_{s \le T_n} \left| \log Z^{n,\vartheta_n}_s - \left(2m^{n,\vartheta_n}_s - 2\left\langle m^{n,\vartheta_n}, m^{n,\vartheta_n} \right\rangle_s \right) \right| \xrightarrow{P^n} 0, \quad n \to \infty.$$

This theorem suggests that if

(W) there exists a sequence $\{w^n\} \in \mathcal{W}$ such that

$$\left\langle m^{n,\vartheta_n} - \frac{1}{2}\vartheta_n^\top w^n, m^{n,\vartheta_n} - \frac{1}{2}\vartheta_n^\top w^n \right\rangle_{T_n} \xrightarrow{P^n} 0, \quad n \to \infty.$$

for every bounded sequence $\{\vartheta_n\}$

then there is a quadratic approximation.

Main result:

Theorem 2 *Under assumption* O, *the pair of assumptions* (H) *and* (L) *is equivalent to* (W) *and implies the quadratic approximation with the central sequence* w^n *from* (W).

Main result:

Theorem 2 *Under assumption* O, *the pair of assumptions* (H) *and* (L) *is equivalent to* (W) *and implies the quadratic approximation with the central sequence* w^n *from* (W)*.*

The main ideas go back to LeCam. In the discrete-time case similar results were obtained in

Greenwood, P. E. and Shiryayev, A. N. (1985) *Contiguity and the statistical invariance principle*. Stochastics Monographs, 1. Gordon & Breach Science Publishers, New York.

Fabian, V. and Hannan, J. (1987) Local asymptotic behavior of densities. *Statist. Decisions*, **5**, 105–138

For continuous time related results (with certain restrictions on the model) were obtained in

Luschgy, H. (1992) Local asymptotic mixed normality for semimartingale experiments. *Probab. Theory Related Fields*, **92**, 151–176.

Luschgy, H. (1994) Asymptotic inference for semimartingale models with singular parameter points. *J. Statist. Plann. Inference*, **39**, 155–186.

Luschgy, H. (1995) Local asymptotic quadraticity of stochastic process models based on stopping times. *Stochastic Process. Appl.*, **57**, 305–317.

Some further references:

Mémin, J. (1985) Théorèmes limite fonctionnels pour les processus de vraisemblance (cadre asymptotiquement non gaussien). *Publications de l'Institut de Recherche Mathématique de Rennes*.

Jacod, J. and Shiryaev, A. N. (2003) *Limit Theorems for Stochastic Processes*. Springer, Berlin.

Vostrikova, L. (1988) On the weak convergence of likelihood ratio processes of general statistical parametric models. *Stochastics*, **23**, 277–298.

Jacod, J. (1989) Convergence of filtered statistical models and Hellinger processes. *Stochastic Process. Appl.*, **32**, 47–68.

Coquet, F. and Jacod, J. (1990) Convergence des surmartingales. Application aux vraisemblances partielles. *Séminaire de Probabilités, XXIV, 1988/89*, 282–299, Lecture Notes in Math., **1426**, Springer, Berlin.

Kramkov, D. O. (1993) Convergence of filtered experiments to the experiment generated by a semimartingale. In Niemi, H. (ed.) et al., *Proceedings of the Third Finnish-Soviet symposium on probability theory and mathematical statistics. Turku, Finland, August 13-16, 1991*, Utrecht: VSP. Front. Pure Appl. Probab. **1**, 145-164.

Necessity of conditions

Theorem 3 *Let*

$$(P^{n,\vartheta_n}_{T_n}) \lhd \rhd (P^n_{T_n})$$

for each bounded sequence $\{\vartheta_n\}$ *and condition* ANJ *is satisfied. Assume also that there exist càdlàg processes* X^n *with values in* $\mathbb{R}^k$ *and predictable processes* $B^{n,\vartheta}$ *with finite variation such that*

$$\text{the sequence} \quad \left(\mathrm{Var}\,(B^{n,\vartheta})_{T_n}\middle|P^n\right) \quad \text{is } \mathbb{R}\text{-tight},$$

and

$$\sup_{s\le T_n}\left|\log Z^{n,\vartheta_n}_s - \left(\vartheta_n^\top X^n_s - B^{n,\vartheta_n}_s\right)\right| \xrightarrow{P^n} 0, \quad n\to\infty,$$

for every bounded sequence $\{\vartheta_n\}$. *Then assumptions* (O), (H), (L), *and* (W) *are satisfied.*

Theorem 4 *Let* $T \in \mathbb{R}_+$ *and* $T_n \equiv T$. *Assume that, for every* $t \in S$, *where* S *is a dense subset of* $[0, T]$ *containing* 0 *and* T, *there are random vectors* V_t^n *and deterministic positive definite symmetric matrices* K_t *such that*

$$\log Z_t^{n,\vartheta_n} - \left(\vartheta_n^\top V_t^n - \frac{1}{2}\vartheta_n^\top K_t \vartheta_n\right) \xrightarrow{P^n} 0, \quad n \to \infty,$$

and

$$\mathrm{Law}(V_t^n | P^n) \Rightarrow \mathcal{N}(0, K_t), \quad n \to \infty.$$

Assume also that $[0, T] \ni t \rightsquigarrow K_t$ *is a continuous increasing function. Then conditions* (O), (H), (L) *and* (W) *are satisfied.*

Introduce the following assumption:

(D) For all bounded sequences $\{\vartheta_n\}$ and $\{\eta_n\}$ there is a deterministic sequence $d_n = d_n(\{\vartheta_n\}, \{\eta_n\})$ such that

$$h_{T_n}^{n,\vartheta_n,\eta_n} - d_n \xrightarrow{P^n} 0, \quad n \to \infty.$$

Of course, (D) is satisfied if all the processes $h^{n,\vartheta,\eta}$ admit deterministic versions and T_n are deterministic, in particular, if our experiments correspond to independent observations on a non-random time interval.

Theorem 5 *Assume that there are random vectors* V^n *and a deterministic positive definite symmetric matrix* K^n *such that*

$$\log Z_{T_n}^{n,\vartheta_n} - \left(\vartheta_n^\top V^n - \frac{1}{2}\vartheta_n^\top K^n \vartheta_n\right) \xrightarrow{P^n} 0, \quad n \to \infty,$$

and

$$\mathrm{Law}(V^n | P^n) \Rightarrow \mathcal{N}(0, K), \quad n \to \infty.$$

Assume also that conditions ANJ *and* (D) *hold. Then assumptions* (O), (H), (L), *and* (W) *are satisfied.*

Pfanzagl, J. (2002) On distinguished LAN-representations. *Math. Methods Statist.* **11**, 477–488.

Gushchin, A. A. and Valkeila, E. (2003) Approximations and limit theorems for likelihood ratio processes in the binary case. *Statist. Decisions*, **21**, 219–260.