DERIVATIVE IN THE MEAN OF A DENSITY AND STATISTICAL APPLICATIONS

J. JUREČKOVÁ
Charles University

X. MILHAUD
Université Paul Sabatier

In the parametric model $\mathcal{P} = \{P_\theta : \theta \in \boldsymbol{\Theta}\}$ with p-dimensional parameter θ, such that P_θ are mutually absolutely continuous with densities differentiable in the mean, we prove an identity transforming the derivative in the mean of the likelihood ratio in $\mathcal{P}$ into the derivative in the mean of a general statistic. This identity makes possible, among others, to approximate the expectation (or other moments) of a statistic, locally in a neighborhood of θ_0 and non-asymptotically under a finite number of observations. This in turn provides the local power of a general test of the simple hypothesis $\mathbf{H}_0$: $\theta = \theta_0$. Using these results, we show that the classical χ^2-test is the locally most powerful invariant test for the hypothesis of balanced multinomial trials.

1. Introduction

Consider the parametric model $(\mathcal{X}, \mathcal{B}, \mathcal{P})$, where $\mathcal{P} = \{P_\theta : \theta \in \boldsymbol{\Theta}\}$ is a family of probability distributions on $\mathcal{B}$ with $\boldsymbol{\Theta}$ being an open subset of $\mathbb{R}^p$. Assume that the members of the family $\mathcal{P}$ are mutually absolutely continuous with square roots of densities differentiable in quadratic mean. More precisely, we impose the following conditions on $\mathcal{P}$:

A.1 The probabilities P_θ, $\theta \in \boldsymbol{\Theta}$ are mutually absolutely continuous. Denote $f_\theta(\cdot) = dP_\theta / dP_{\theta_0}$ the density of P_θ with respect to P_{θ_0}, $\theta \in \boldsymbol{\Theta}$, with a fixed $\theta_0 \in \boldsymbol{\Theta}$.

A.2 The mapping $\theta^* \mapsto f_{\theta^*}/f_\theta$ is P_θ-differentiable in the mean for all $\theta, \theta^* \in \boldsymbol{\Theta}$ with derivative $\mathbf{l}(\cdot, \theta) = \big(\ell_1(\cdot,\theta), \ldots, \ell_p(\cdot,\theta)\big)'$ such that $\mathbb{E}_\theta \|\mathbf{l}(\cdot,\theta)\|^2 < \infty$ if there exists a random vector $\mathbf{l}(\cdot,\theta)$ satisfying

$$\mathbb{E}_\theta \left\{ \sum_{k=1}^{p} \left| \frac{1}{h_k} \left(\frac{f_{\theta+\mathbf{h}}}{f_\theta} - 1 \right) - \ell_k(\cdot,\theta) \right| \right\} \to 0 \quad \text{as } \mathbf{h} \to \mathbf{0}. \tag{1.1}$$

Then, obviously

$$\mathbb{E}_\theta\big(\mathbf{l}(\cdot,\theta)\big) = \lim_{\mathbf{h}\to 0} \mathbb{E}_\theta \left\{ \frac{1}{|\mathbf{h}|} \left(\frac{f_{\theta+h}}{f_\theta} - 1 \right) \right\} \equiv 0. \tag{1.2}$$

Keywords and phrases: differentiability in the mean; differentiability in quadratic mean; local power of the test; locally optimal rank test; multinomial distribution; Rao score test; score function; χ^2 test.

The derivative in the mean $\mathbf{l}(\cdot,\theta)$ (L_1-derivative) is a generalization of the logarithmic derivative of the density, $\dot{L}_\theta(\cdot) = \nabla \log f(\cdot,\theta)$ (*the score function*); compare (1.2) and the fact that the L_1-derivative of the product measure is equal to the sum of L_1-derivatives of the marginal distributions, and hence to a sum of independent random variables. The L_1-derivative and the score function play a basic role in the statistical inference; they are basic tools in the maximum likelihood estimation, in the study of score tests, and in the locally or asymptotically optimal rank tests, among others. The standardized vector $\mathbf{l}(\cdot,\theta)$ plays a similar role in the general model as the vector of observations $\mathbf{X}$ in the normal model, as has been demonstrated in the literature (e.g., Bondesson (1974), Jurečková and Milhaud (1994, 1999), Jurečková (1999), and others).

We shall first show that the differentiability of the likelihood ratio in the mean follows from the stronger differentiability of the square root of the same in the quadratic mean, which in turn leads to the LAN (local asymptotic normality) of the model; for a detailed explanation of this concept see, among others, Hájek (1970, 1972), LeCam and Yang (1990) and Pollard, Yang and Torgersen (1997), where other references are cited.

The mapping $\theta^* \mapsto \sqrt{f_{\theta^*}/f_\theta}$ is P_θ-differentiable in the quadratic mean, $\theta,\theta^* \in \mathbf{\Theta}$, if there exists a random vector $\mathbf{l}(\cdot,\theta) = \big(\ell_1(\cdot,\theta),\ldots,\ell_p(\cdot,\theta)\big)'$ such that $\mathbb{E}_\theta\|\mathbf{l}(\cdot,\theta)\|^2 < \infty$

$$\mathbb{E}_\theta\Bigg(\sum_{k=1}^{p}\Bigg[\frac{1}{h_k}\Bigg(\sqrt{\frac{f_{\theta+\mathbf{h}}}{f_\theta}}-1\Bigg)-\frac{1}{2}\ell_k(\cdot,\theta)\Bigg]^2\Bigg) \to 0 \quad \text{as } \mathbf{h}\to\mathbf{0}. \tag{1.3}$$

The following lemma shows that the differentiability of $\sqrt{f_{\theta^*}/f_\theta}$ in quadratic mean implies the differentiability of f_{θ^*}/f_θ in the mean and that both L_1- and L_2-derivatives are linearly related:

Lemma 1.1. *Let $\mathcal{P} = \{P_\theta : \theta \in \mathbf{\Theta}\}$ be a parametric model satisfying condition* A.1. *Then the differentiability in quadratic mean* (1.3) *implies the differentiability in the mean* (1.1).

Proof. We have

$$\begin{aligned}
&\mathbb{E}_\theta\Bigg\{\sum_{k=1}^{p}\Bigg|\frac{1}{h_k}\Bigg(\frac{f_{\theta+\mathbf{h}}}{f_\theta}-1\Bigg)-\ell_k(\cdot,\theta)\Bigg|\Bigg\}\\
&=\sum_{k=1}^{p}\mathbb{E}_\theta\Bigg\{\Bigg|\Bigg[\frac{1}{h_k}\Bigg(\sqrt{\frac{f_{\theta+\mathbf{h}}}{f_\theta}}-1\Bigg)-\frac{1}{2}\ell_k(\cdot,\theta)\Bigg]\Bigg(1+\sqrt{\frac{f_{\theta+\mathbf{h}}}{f_\theta}}\Bigg)\\
&\qquad\qquad\qquad\qquad+\frac{1}{2}\ell_k(\cdot,\theta)\Bigg(\sqrt{\frac{f_{\theta+\mathbf{h}}}{f_\theta}}-1\Bigg)\Bigg|\Bigg\}
\end{aligned}$$

$$\le \left\{\mathbb{E}_\theta\Big(\sqrt{\frac{f_{\theta+\mathbf{h}}}{f_\theta}}+1\Big)^2\right\}^{1/2}\sum_{k=1}^{p}\left\{\mathbb{E}_\theta\Big[\frac{1}{h_k}\Big(\sqrt{\frac{f_{\theta+\mathbf{h}}}{f_\theta}}-1\Big)-\frac{1}{2}\ell_k(\cdot,\theta)\Big]^2\right\}^{1/2}$$

$$+\frac{1}{2}\left\{\mathbb{E}_\theta\Big(\sqrt{\frac{f_{\theta+\mathbf{h}}}{f_\theta}}-1\Big)^2\right\}^{1/2}\sum_{k=1}^{p}\left\{\mathbb{E}_\theta\Big(\ell_k(\cdot,\theta)\Big)^2\right\}^{1/2}$$

$$\le \sqrt{p}\left\{\mathbb{E}_\theta\Big(\sqrt{\frac{f_{\theta+\mathbf{h}}}{f_\theta}}+1\Big)^2\right\}^{1/2}\left\{\mathbb{E}_0\Big(\sum_{k=1}^{p}\Big[\frac{1}{h_k}\Big(\sqrt{\frac{f_{\theta+\mathbf{h}}}{f_\theta}}-1\Big)-\frac{1}{2}\ell_k(\cdot,\theta)\Big]^2\Big)\right\}^{1/2}$$

$$+\frac{\sqrt{p}}{2}\left\{\mathbb{E}_\theta\|\mathbf{l}(\cdot,\theta)\|^2\right\}^{1/2}\left\{\mathbb{E}\Big(\sqrt{\frac{f_{\theta+\mathbf{h}}}{f_\theta}}-1\Big)^2\right\}^{1/2}.$$

From A.2 and from the fact that

$$\mathbb{E}_\theta\Big(1+\sqrt{\frac{f_{\theta+\mathbf{h}}}{f_\theta}}\Big)^2\le 5 \quad\text{and}\quad \mathbb{E}_\theta\Big(\sqrt{\frac{f_{\theta+\mathbf{h}}}{f_\theta}}-1\Big)^2\to 0 \text{ as } \mathbf{h}\to\mathbf{0}$$

we conclude that the above inequality implies (1.1). □

Jurečková (1999) expressed the logarithmic derivative of the density of a general statistic S in the model with a single location parameter by means of conditional expectation of the score function of the sample, given $S = s$. Here we shall extend this result and show that, under conditions A.1–A.2, the likelihood ratio of an arbitrary statistic S is differentiable in the mean and that its L_1-derivative can be explicitly expressed as the conditional expectation of $\mathbf{l}(\cdot,\theta)$ given S. This applies to the score function as a special case and makes possible to derive the local power of a general test of the hypothesis $\mathbf{H}$: $\theta=\theta_0$ in a neighborhood of θ_0. This, in turn, permits a more profound study of the structure of locally optimal tests. Several tests are considered as examples; it is shown that the classical χ^2-test is the locally optimal invariant test of the hypothesis of balanced multinomial trial.

The L_1-derivative of S is studied in Section 2; the results of Section 2 are then used in Section 3 to derive local approximations of $\mathbb{E}_\theta S$. The local power of the test is derived in Section 4, along with some illustrative examples. The classical χ^2-test and its local optimality are considered in Section 5.

2. L_1-derivative of a statistic

Let S be an arbitrary statistic, i.e. a $\mathcal{B}$-measurable mapping $S\colon(\mathcal{X},\mathcal{B})\mapsto(\mathcal{S},\mathcal{A})$. Denote by $P_\theta^S=P_\theta(S^{-1})$ the probability distribution of S under $\theta\in\mathbf{\Theta}$ and $Q=P_{\theta_0}^S$ for a fixed $\theta_0\in\mathbf{\Theta}$. We shall first show that, under conditions A.1–A.2, the probabilities P_θ^S are also mutually absolutely continuous, and characterize the likelihood ratio of S as the conditional expectation of the likelihood ration of X, given S.

Lemma 2.1. *Under the conditions* A.1–A.2, *the probabilities* P_θ^S *are mutually absolutely continuous and*

$$\frac{g_{\theta^*}(S)}{g_\theta(S)} = \mathbb{E}_\theta\left[\frac{f_{\theta^*}(X)}{f_\theta(X)} \,\Big|\, S\right] \quad a.s.\ [Q] \tag{2.1}$$

where $g_\theta = dP_\theta^S/dQ$.

Proof. By condition A.1, we can write for any $A \in \mathcal{S}$ and $\forall\theta, \theta' \in \mathbf{\Theta}$,

$$P_\theta^S(A) = P_\theta\big(S^{-1}(A)\big) = 0 \iff P_{\theta'}\big(S^{-1}(A)\big) = P_{\theta'}^S(A) = 0,$$

hence P_θ^S are mutually absolutely continuous.

For any $\mathcal{A}$-measurable bounded function $k\colon \mathcal{S} \mapsto \mathbb{R}$ and for $\theta, \theta^* \in \mathbf{\Theta}$, we can write

$$\begin{aligned}\int_{\mathcal{S}} k(s)P_{\theta^*}^S(ds) &= \int_{\mathcal{S}} k(s)g_{\theta^*}(s)Q(ds)\\ &= \int_{\mathcal{S}} k(s)\frac{g_{\theta^*}(s)}{g_\theta(s)}g_\theta(s)Q(ds) = \int_{\mathcal{S}} k(s)\frac{g_{\theta^*}(s)}{g_{\theta_0}(s)}P_\theta^S(ds)\end{aligned}$$

and, on the other hand,

$$\begin{aligned}\int_{\mathcal{S}} k(s)P_{\theta^*}^S(ds) &= \int_{\mathcal{X}} k\big(S(x)\big)f_{\theta^*}(x)P_{\theta_0}(dx)\\ &= \int_{\mathcal{X}} k\big(S(x)\big)\frac{f_{\theta^*}(x)}{f_\theta(x)}f_\theta(x)P_{\theta_0}(dx)\\ &= \int_{\mathcal{X}} k\big(S(x)\big)\mathbb{E}_\theta\left[\frac{f_{\theta^*}(\cdot)}{f_\theta(\cdot)} \,\Big|\, S(x)\right]P_\theta(dx)\\ &= \int_{\mathcal{S}} k(s)\mathbb{E}_\theta\left[\frac{f_{\theta^*}(\cdot)}{f_\theta(\cdot)} \,\Big|\, S(\cdot) = s\right]P_\theta^S(ds);\end{aligned}$$

hence

$$\frac{g_{\theta^*}}{g_\theta} = \mathbb{E}_\theta\left[\frac{f_{\theta^*}}{f_\theta} \,\Big|\, S = s\right] \text{ a.s. } [Q]. \qquad \square$$

As our first main result, we shall show that, under the conditions A.1–A.2, the likelihood ratio of S is differentiable in the mean and its L_1 derivative can be explicitly expressed as the conditional expectation of $\mathbf{l}(\cdot, \theta)$ given S.

Theorem 2.1. *Let* $S\colon (\mathcal{X}, \mathcal{B}) \mapsto (\mathcal{S}, \mathcal{A})$ *be a statistic and let* $P_\theta^S = P_\theta(S^{-1})$ *be the probability distribution of* S *under* $\theta \in \mathbf{\Theta}$. *Denote by* $Q = P_{\theta_0}(S^{-1})$ *the distribution under a fixed* $\theta_0 \in \mathbf{\Theta}$ *and* $g_\theta = dP_\theta^S/dQ$, $\theta \in \mathbf{\Theta}$. *Then, under the conditions* A.1–A.2,

(i) *the mapping* $\theta^* \mapsto g_{\theta^*}/g_\theta$ *is differentiable in the mean with respect to* P_θ^S, $\theta, \theta^* \in \boldsymbol{\Theta}$, *and its* L_1*-derivative is equal to* $\mathbb{E}_\theta(\mathbf{l}(\cdot,\theta) \mid S = s)$; *i.e.,*

$$(2.2)\quad \mathbb{E}_\theta\left\{\sum_{k=1}^{p}\left|\frac{1}{h_k}\left(\frac{g_{\theta+\mathbf{h}}}{g_\theta}-1\right)-\mathbb{E}_\theta(\ell_k(\cdot,\theta) \mid S=s)\right|\right\} \to 0 \quad \textit{as } \mathbf{h}\to\mathbf{0}.$$

(ii) *The* L_1*-derivative of* $\theta^* \mapsto g_{\theta^*}/g_\theta$ *belongs to* $L^2(P_\theta^S)$, *i.e.,*

$$\mathbb{E}_\theta\big[(\mathbb{E}_\theta(\mathbf{l}(\cdot,\theta) \mid S))'\mathbb{E}_\theta(\mathbf{l}(\cdot,\theta) \mid S)\big] \le \mathbb{E}_\theta\big(\mathbf{l}'(\cdot,\theta)\mathbf{l}(\cdot,\theta)\big) < \infty.$$

Proof. (i) Denote

$$\Delta(\mathbf{h}) = \int_{\mathcal{S}}\sum_{k=1}^{p}\left|\frac{1}{h_k}\left(\frac{g_{\theta+\mathbf{h}}(s)}{g_\theta(s)}-1\right)-\mathbb{E}_\theta\left(\ell_k(\cdot,\theta)\;\middle|\; S(\cdot)=s\right)\right|P_\theta^S(ds).$$

Then

$$\begin{aligned}\Delta(\mathbf{h}) &= \int_{\mathcal{S}}\sum_{k=1}^{p}\left|\frac{1}{h_k}\left[\mathbb{E}_\theta\left(\frac{f_{\theta+\mathbf{h}}(\cdot)}{f_\theta(\cdot)}\;\middle|\; S(\cdot)=s\right)-1\right]\right.\\ &\qquad\qquad\qquad\qquad \left. -\,\mathbb{E}_\theta(\ell_k(\cdot,\theta) \mid S(\cdot)=s)\right|P_\theta^S(ds)\\ &= \int_{\mathcal{X}}\sum_{k=1}^{p}\left|\frac{1}{h_k}\left[\mathbb{E}_\theta\left(\frac{f_{\theta+\mathbf{h}}(\cdot)}{f_\theta(\cdot)}\;\middle|\; S(X)\right)-1\right]-\mathbb{E}_\theta\big(\ell_k(\cdot,\theta) \mid S(X)\big)\right|P_\theta(dx).\end{aligned}$$

Further, by the Jenssen inequality,

$$0 \le \Delta(\mathbf{h}) \le \int_{\mathcal{X}}\sum_{k=1}^{p}\left|\frac{1}{h_k}\left(\frac{f_{\theta+\mathbf{h}}(x)}{f_\theta(x)}-1\right)-\ell_k(x,\theta)\right|P_\theta(dx)$$

and the right member tends to 0 by A.2.

(ii) By (1.2),

$$\mathbb{E}_\theta\big(\mathbb{E}_\theta(\mathbf{l}(\cdot,\theta) \mid S)\big) = \mathbb{E}_\theta\big(\mathbf{l}(\cdot,\theta)\big) = \mathbf{0}$$

and by the Jenssen inequality

$$\begin{aligned}&\int_{\mathcal{S}}\mathbb{E}_\theta\Big((\mathbf{l}(\cdot,\theta))'\;\Big|\; S(\cdot)=s\Big)\mathbb{E}_\theta\big(\mathbf{l}(\cdot,\theta) \mid S(\cdot)=s\big)P_\theta^S(ds)\\ &\qquad = \int_{\mathcal{X}}\mathbb{E}_\theta\Big((\mathbf{l}(\cdot,\theta))'\;\Big|\; S(X)\Big)\mathbb{E}_\theta\big(\mathbf{l}(\cdot,\theta) \mid S(X)\big)P_\theta(dx)\\ &\qquad \le \int_{\mathcal{X}}(\mathbf{l}(x,\theta))'\mathbf{l}(x,\theta)P_\theta(dx) < \infty.\end{aligned}$$

□

Consider an application of Theorem 2.1 when $\mathbf{X} = (X_1, \ldots, X_n)$ is a sample from a distribution $F(\cdot, \theta)$ with density $f(\cdot, \theta)$, $\theta \in \boldsymbol{\Theta} \subset \mathbb{R}^p$ with respect to a σ-finite measure μ, such that the Fisher's information matrix $\mathbf{J}_f(\theta)$ is positively definite. Denote by

$$L(\mathbf{x}, \boldsymbol{\theta}) = \sum_{i=1}^{n} \log f(x_i, \boldsymbol{\theta}) \tag{2.3}$$

the log-likelihood function of the sample $(x_1, \ldots, x_n)$. The vector function with p components

$$\begin{aligned} l(\mathbf{x}, \boldsymbol{\theta}) &= (\ell_1(\mathbf{x}, \boldsymbol{\theta}), \ldots, \ell_p(\mathbf{x}, \boldsymbol{\theta}))' \\ &= \left(\frac{\partial L(\mathbf{x}, \boldsymbol{\theta})}{\partial \theta_1}, \ldots, \frac{\partial L(\mathbf{x}, \boldsymbol{\theta})}{\partial \theta_p} \right)', \quad \theta \in \boldsymbol{\Theta} \end{aligned} \tag{2.4}$$

is called the (Fisher) *score function* corresponding to $f(x, \boldsymbol{\theta})$. Notice that $\mathbb{E}_{\boldsymbol{\theta}}(l(\mathbf{X}, \boldsymbol{\theta})) = \mathbf{0}$ and that $\mathrm{cov}_{\boldsymbol{\theta}}\{l(\mathbf{X}, \boldsymbol{\theta})\} = \mathbf{J}_f(\boldsymbol{\theta})$ is the Fisher information matrix.

Let $S_n(X_1, \ldots, X_n)$ be a real statistic with distribution function $G(s, \theta)$ and density $g(s, \theta)$ which is assumed being differentiable in θ. Then we have the following corollary to Theorem 2.1:

Corollary 2.1. *Let $X_1, \ldots, X_n$ be i.i.d. observations from a distribution with distribution function $F(x, \theta)$, $\theta \in \boldsymbol{\Theta} \subset \mathbb{R}^p$, and density $f(x, \theta)$, differentiable in the components of θ, and with positively definite Fisher's information matrix $\mathbf{J}_f(\theta)$. Let $S_n(X_1, \ldots, X_n)$ be a statistic; denote $G(s, \theta)$ and $g(s, \theta)$ its distribution function and density, respectively; assume that $g(s, \theta)$ is differentiable in θ. Then*

$$\left(\frac{\partial \log g(s, \theta)}{\partial \theta_1}, \ldots, \frac{\partial \log g(s, \theta)}{\partial \theta_p} \right)' = \mathbb{E}_\theta\{l(\mathbf{X}, \theta) \mid S(X_1, \ldots, X_n) = s\}, \tag{2.5}$$

where $\mathbb{E}_\theta$ denotes the expectation with respect to the density $\prod_{i=1}^n f(x_i, \theta)$.

3. Finite sample expansion of $\mathbb{E}_{\boldsymbol{\theta}} S$

Theorem 2.1 provides a tool to compute local and non-asymptotic approximations of moments of real statistics in a neighborhood of a fixed θ. These approximations, in turn, enable to approximate powers of tests.

Theorem 3.1. (i) *Let $Z\colon (\mathcal{X}, \mathcal{B}) \mapsto (\mathbb{R}, \mathcal{B}^1)$ be a bounded real statistic. Then, under conditions* A.1–A.2,

$$\mathbb{E}_{\theta+\mathbf{h}}(Z) = \mathbb{E}_\theta(Z) + \mathbb{E}_\theta(\mathbf{h}'\mathbf{l}(X, \theta) Z) + \|\mathbf{h}\| \varepsilon(\mathbf{h}, \theta), \quad \theta \in \boldsymbol{\Theta}, \tag{3.1}$$

where $\varepsilon(\mathbf{h}, \theta) \to 0$ as $\mathbf{h} \to 0$.

(ii) *Let* $W = \psi(S)$ *be a bounded real function of statistic* $S\colon (\mathcal{X}, \mathcal{B}) \mapsto (\mathcal{S}, \mathcal{A})$. *Then,*

$$(3.2)\quad \mathbb{E}_{\theta+\mathbf{h}}(W) = \mathbb{E}_\theta(W) + \mathbb{E}_\theta\big(\mathbf{h}'\mathbb{E}_\theta\big(\mathbf{l}(X,\theta) \mid S\big)W\big) + \|\mathbf{h}\|\varepsilon^*(\mathbf{h},\theta), \qquad \theta \in \boldsymbol{\Theta},$$

where $\varepsilon^*(\mathbf{h},\theta) \to 0$ *as* $\mathbf{h} \to \mathbf{0}$.

Proof. (i) Using Lemma 1.1 and considering that Z is bounded, we have

$$\begin{aligned}
|\mathbb{E}_{\theta+\mathbf{h}}(Z) &- \mathbb{E}_\theta(Z) - \mathbb{E}_\theta(\mathbf{h}'\mathbf{l}(X,\theta)Z)| \\
&= \left|\mathbb{E}_\theta\left\{Z\left[\frac{f_{\theta+\mathbf{h}}}{f_\theta} - 1 - \mathbf{h}'\mathbf{l}(X,\theta)\right]\right\}\right| \\
&= \left|\sum_{\substack{k=1\\ h_k\neq 0}}^{p} \mathbb{E}_\theta\left\{h_k Z\left[\frac{1}{h_k}\left(\frac{f_{\theta+\mathbf{h}}}{f_\theta} - 1\right) - \ell_k(X,\theta))\right]\right\}\right| \\
&\leq \|\mathbf{h}\|\varepsilon(\mathbf{h},\theta).
\end{aligned}$$

(ii) The proof is similar as in step (i) where we replace $g_{\theta+\mathbf{h}}/g_\theta$ by $\mathbb{E}_\theta(f_{\theta+\mathbf{h}}/f_\theta \mid S)$. □

4. Local approximation of the power of a test

The above results enable us to derive an approximation of the power of a test, a non-asymptotic one, in a neighborhood of the hypothesis. This, in turn, is a starting point to a more profound study of the structure of locally most powerful tests.

Consider a random sample $\mathbf{X} = (X_1, \ldots, X_n)$ from a distribution function $F(x,\theta)$, $\theta \in \boldsymbol{\Theta} \subset \mathbb{R}^p$, with density $f(x,\theta)$ and with positively definite Fisher's information matrix $\mathbf{J}_f(\theta)$. Let $l(\mathbf{X},\theta)$ be the Fisher score function (2.4).

Consider a test Φ of the simple hypothesis $\mathbf{H}_0$: $\theta = \theta_0$ against the alternative $\mathbf{K}$: $\theta \neq \theta_0$ of size α, i.e., $\mathbb{E}_{\theta_0}[\Phi(\mathbf{X})] = \alpha$. Assume that the test is based on the criterion $S_n = S_n(X_1, \ldots, X_n)$, $S\colon (\mathcal{X}, \mathcal{B}) \mapsto (\mathcal{S}, \mathcal{A})$ in such a way that $\Phi(\mathbf{x}) = 1$ if $\mathbf{x} \in \mathcal{K}_\alpha$ and $\Phi(\mathbf{x}) = 0$ otherwise. Let $g(s,\theta)$ be the density of S with respect to a σ-finite measure μ on $\mathcal{A}$, and assume that $g(\cdot,\theta)$ is twice differentiable under an integral sign with respect to components of θ.

The next theorem gives the local power of Φ for a fixed n:

Theorem 4.1. *Assume that the density* $f(x,\theta)$ *is three times differentiable in the components of* θ. *Then, for* $\theta \neq \theta_0$,

$$\begin{aligned}
(4.1)\quad \mathbb{E}_\theta[\Phi(\mathbf{X})] &= \alpha + (\theta-\theta_0)'\mathbb{E}_{\theta_0}\{l(\mathbf{X},\theta_0)\Phi(\mathbf{X})\} \\
&\quad + \tfrac{1}{2}(\theta-\theta_0)'\mathbb{E}_{\theta_0}\{\Phi(\mathbf{X})\mathbf{A}(\mathbf{X},\theta_0)\}(\theta-\theta_0) + \mathcal{O}(\|\theta-\theta_0\|^3),
\end{aligned}$$

where $\mathbf{A}(\mathbf{x},\theta_0)$ *is the* $p\times p$ *matrix,*

$$\mathbf{A}(\mathbf{x},\theta_0)=\left[\ell_j(\mathbf{x},\theta_0)\ell_k(\mathbf{x},\theta_0)+\frac{\partial\ell_j(\mathbf{x},\theta)}{\partial\theta_k}\bigg|_{\theta=\theta_0}\right]_{j,k=1}^p. \tag{4.2}$$

Proof. Denote by $G(s,\theta)$ the distribution function of S and $\dot g_j(s,\theta)=\partial g(s,\theta)/\partial\theta_j$, $j=1,\dots,p$ and $\ddot g_{jk}(s,\theta)=\partial^2 g(s,\theta)/\partial\theta_j\partial\theta_k$, $j,k=1,\dots,p$. The power of the test at θ can be written in the form

$$\mathbb{E}_\theta[\Phi(\mathbf{X})]=\int_{\mathcal{K}_\alpha} g(s,\theta)\,d\mu \tag{4.3}$$

and, using the Taylor expansion in θ, we obtain

$$\begin{aligned}\mathbb{E}_\theta[\Phi(\mathbf{X})]&=\alpha+\sum_{j=1}^p(\theta_j-\theta_{0j})\int_{\mathcal{K}_\alpha}\frac{\dot g_j(s,\theta_0)}{g(s,\theta_0)}g(s,\theta_0)d\mu\\&+\frac12\sum_{j=1}^p\sum_{k=1}^p(\theta_j-\theta_{0j})(\theta_k-\theta_{0k})\int_{\mathcal{K}_\alpha}\ddot g_{jk}(s,\theta_0)\,ds+\mathcal{O}(\|\theta-\theta_0\|^3).\end{aligned} \tag{4.4}$$

By (2.5), the second term on the right-hand side of (4.4) equals to

$$\begin{aligned}(\theta-\theta_0)'\int_{\mathcal{K}_\alpha}\mathbb{E}_{\theta_0}\{l(\mathbf{X},\theta_0)\mid S=s\}g(s,\theta_0)\,d\mu&\\=(\theta-\theta_0)'\mathbb{E}_{\theta_0}\{l(\mathbf{X},\theta_0)\Phi(\mathbf{X})\}&\end{aligned} \tag{4.5}$$

and similarly we obtain that the third term equals to

$$\frac12(\boldsymbol{\theta}-\boldsymbol{\theta}_0)'\mathbb{E}_{\theta_0}\{\Phi(\mathbf{X})\mathbf{A}(\mathbf{X},\theta_0)\}(\theta-\theta_0). \tag{4.6}$$

(4.4), (4.5) and (4.6) give the desired approximation. □

Example 1. Let f belong to the exponential family, i.e.

$$\prod_{i=1}^n f(x_i,\theta)=\exp\left\{\sum_{j=1}^p\theta_jT_j(\mathbf{x})+A(\theta)+B(\mathbf{x})\right\}. \tag{4.7}$$

Then $\ell_j(\mathbf{x},\theta)=T_j(\mathbf{x})-\mathbb{E}_\theta\,[T_j(\mathbf{X})]$, hence $\frac{\partial\ell_j(\mathbf{x},\theta)}{\partial\theta_k}=0$ and thus

$$\begin{aligned}\mathbb{E}_\theta[\Phi(\mathbf{X})]&=\alpha+(\theta-\theta_0)'\mathbb{E}_{\theta_0}\{\Phi(\mathbf{X})[\mathbf{T}(\mathbf{X})-\mathbb{E}_{\theta_0}\mathbf{T}(\mathbf{X})]\}\\&+\tfrac12(\theta-\theta_0)'\mathbb{E}_{\theta_0}\{\Phi(\mathbf{X})\mathbf{A}_{\theta_0}(\mathbf{X})\}(\theta-\theta_0)+\mathcal{O}(\|\theta-\theta_0\|^3),\end{aligned} \tag{4.8}$$

where $\mathbf{T}(\mathbf{x})=(T_1(\mathbf{x}),\dots,T_p(\mathbf{x}))'$ and

$$\begin{aligned}&\mathbb{E}_{\theta_0}\{\Phi(\mathbf{X})\mathbf{A}_{\theta_0}(\mathbf{X})\}\\&=\left[\operatorname{cov}_{\theta_0}\left\{\left(T_j(\mathbf{X})-\mathbb{E}_{\theta_0}T_j(\mathbf{X})\right)\left(T_k(\mathbf{X})-\mathbb{E}_{\theta_0}T_k(\mathbf{X})\right),\Phi(\mathbf{X})\right\}\right]_{j,k=1}^p.\end{aligned} \tag{4.9}$$

The test $\Phi(\mathbf{x})$ is locally unbiased if

$$\text{(4.10)}\qquad \mathbb{E}_{\theta_0}[\Phi(\mathbf{X})\mathbf{T}(\mathbf{X})] = \mathbb{E}_{\theta_0}[\Phi(\mathbf{X})]\cdot\mathbb{E}_{\theta_0}[\mathbf{T}(\mathbf{X})] = \alpha\mathbb{E}_{\theta_0}[\mathbf{T}(\mathbf{X})],$$

hence the second term on the right-hand side of (4.8) vanishes. Consequently, the local power of the locally unbiased test is the larger, the more $\Phi(\mathbf{X})$ is correlated with the matrix

$$\left[\big(T_j(\mathbf{X}) - \mathbb{E}_{\theta_0}T_j(\mathbf{X})\big)\big(T_k(\mathbf{X}) - \mathbb{E}_{\theta_0}T_k(\mathbf{X})\big)\right]_{j,k=1}^{p}.$$

As a special case, the test with the critical region

$$\{\mathbf{X} : T_1(\mathbf{X}) - \mathbb{E}_{\theta_0}T_1(\mathbf{X}) \le C_1 \text{ or } T_1(\mathbf{X}) - \mathbb{E}_{\theta_0}T_1(\mathbf{X}) \ge C_2\},$$

where C_1 and C_2 are determined by the conditions $\mathbb{E}_{\theta_0}[\Phi(\mathbf{X})] = \alpha$ and (4.10), is the locally most powerful, locally unbiased test for $\mathbf{H}_0$ against the alternative $\theta = (\theta_1, \theta_{02}, \dots, \theta_{0p})'$ in the exponential family (4.7).

Example 2. Consider the simple regression model, where the vector $\mathbf{X} = (X_1, \dots, X_n)$ has the Lebesgue density $\prod_{i=1}^{n} f(x_i - c_i\theta)$, $\theta \in \mathbb{R}^1$ with given $c_1, \dots, c_n$, and we wish to test the hypothesis $\mathbf{H}$: $\theta = 0$ against $\mathbf{K}$: $\theta > 0$. The pertaining score function is $\ell(\mathbf{x}, \theta) = -\sum_{i=1}^{n} c_i f'(x_i - c_i\theta)/f(x_i - c_i\theta)$, and the test statistic of the score test of $\mathbf{H}$ has the form $\big(-\sum_{i=1}^{n} f'(x_i)/f(x_i)\big)$. Let $\mathbf{S}_n(\mathbf{X}) = \mathbf{R}_n = (R_1, \dots, R_n)$ be the vector of ranks of $X_1, \dots, X_n$. It follows from Theorem 2.1 that the score function of $\mathbf{S}_n(\mathbf{X}) = \mathbf{R}_n$ is

$$\text{(4.11)}\quad \mathbb{E}_\theta\left[-\sum_{i=1}^{n} c_i \frac{f'(x_i - c_i\theta)}{f(x_i - c_i\theta)} \;\middle|\; \mathbf{R}_n = \mathbf{r}\right] = \mathbb{E}_\theta\left[-\sum_{i=1}^{n} c_i \frac{f'(x_{(r_i)} - c_i\theta)}{f(x_{(r_i)} - c_i\theta)}\right],$$

where $(X_{(1)}, \dots, X_{(n)})$ are the order statistics corresponding to $(X_1, \dots, X_n)$ and $\mathbf{r}$ runs over the set $\mathcal{R}_n$ of permutations of $\{1, \dots, n\}$. It is well known that the locally most powerful rank test of $\mathbf{H}$ rejects $\mathbf{H}$ provided

$$\sum_{i=1}^{n} c_i a_n(R_i, f) > C_\alpha$$

where

$$a_n(R_i, f) = \mathbb{E}_0\left(-\frac{f'(X_{(i)})}{f(X_{(i)})}\right), \quad i = 1, \dots, n$$

and $\mathbb{E}_0$ denotes the expectation under $\theta_0 = 0$ (see Hájek and Šidák, 1967). By (4.11), we can also interprete this test as the score test based on ranks.

The finite sample approximation of Theorem 4.1 is applicable also to the Rao score test of the simple hypothesis $\mathbf{H}_0$: $\theta = \theta_0$ against the alternative $\mathbf{K}$: $\theta \ne \theta_0$, with the test criterion

$$\text{(4.12)}\qquad R_0 = (l(\mathbf{X}, \theta_0))'(\mathbf{J}_f(\theta_0))^{-1}(l(\mathbf{X}, \theta_0)).$$

The asymptotic null distribution of R_0 is $\chi^2(p)$ under general conditions. Many authors studied the asymptotic properties of the Rao test, namely its local asymptotic efficiency with respect to the likelihood ratio and the Wald tests; see Mukherjee (1993), where other references are cited. However, these considerations are asymptotic for $n \to \infty$ and are often based on the Edgeworth expansions of the distribution of R_0. It would be of interest to compare the asymptotic and non-asymptotic approximations; this will be an object of a special study.

The test of balanced multinomial trials is studied in a special Section 6.

5. An identity for moments of a statistic

As another application of the expansions in Section 3, we shall derive an identity for a derivative of the νth moment of a statistic, that can be useful in estimation and testing. Let $S = S(X_1, \ldots, X_n)$ be a statistic having a density $g(s, \theta)$ with respect to a σ-finite measure μ; let $\gamma^{(\nu)}(\theta) = \mathbb{E}_\theta\big(S(\mathbf{X})\big)^\nu$ be the ν-th moment of S and assume that it is finite and differentiable θ; $\nu > 0$ is not necessary an integer. Then the identity (2.5) makes possible to derive a useful identity for the moments of S:

Theorem 5.1. *Let $X_1, \ldots, X_n$ be i.i.d. observations with distribution function $F(x, \theta)$, $\theta \in \mathbf{\Theta} \subset \mathbb{R}^p$, and density $f(x, \theta)$, differentiable in the components of θ. Let $S_n(X_1, \ldots, X_n)$ be a statistic, whose νth moment $\gamma^{(\nu)}(\theta)$ is differentiable in components of θ, $\nu > 0$. Then, for all $\theta \in \mathbf{\Theta}$,*

$$\dot{\gamma}^{(\nu)}(\theta) = \mathbb{E}_\theta\big[\big(S(\mathbf{X})\big)^\nu l(\mathbf{X}, \theta)\big], \tag{5.1}$$

where

$$\dot{\gamma}^{(\nu)}(\theta) = (\dot{\gamma}_1^{(\nu)}(\theta), \ldots, \dot{\gamma}_p^{(\nu)}(\theta))', \quad \dot{\gamma}_j^{(\nu)}(\theta) = \frac{\partial \dot{\gamma}^{(\nu)}(\theta)}{\partial \theta_j}, \ j = 1, \ldots, p. \tag{5.2}$$

Proof. We can differentiate $\gamma^{(\nu)}(\theta) = \int s^\nu g(s, \theta)\, d\mu$, obtaining with the aid of (2.5)

$$\begin{aligned} \dot{\gamma}_j^\nu(\theta) &= \int s^\nu \frac{\dot{g}_j(s, \theta)}{g(s, \theta)} g(s, \theta)\, d\mu \\ &= \int s^\nu \mathbb{E}_\theta[\ell_j(\cdot, \theta) \mid S(\cdot) = s] g(s, \theta)\, d\mu \\ &= \mathbb{E}_\theta\big[\big(S(\mathbf{X})\big)^\nu \ell_j(\mathbf{X}, \theta)\big], \quad j = 1, \ldots, p. \end{aligned}$$ □

6. Test of balanced multinomial trials

Let $(\mathbf{X}_1, \ldots, \mathbf{X}_n)'$ be an independent sample from the multinomial $\mathcal{M}(1; p_1, \ldots, p_k)$ distribution, $k \geq 2$, $0 \leq p_j \leq 1$, $j = 1, \ldots, k$, $\sum_{j=1}^k p_j = 1$. More precisely, $\mathbf{X}_i' = (X_{i1}, \ldots, X_{ik})$, $i = 1, \ldots, n$ are i.i.d. random vectors,

$P(\mathbf{X}_i = \mathbf{e}_j) = p_j$, $j = 1, \dots, k$, $i = 1, \dots, n$, where $\mathbf{e}_j \in \mathbb{R}^k$ is the unit vector, $e_{js} = \delta_{js}$, $j, s = 1, \dots, k$.

The problem is to test the balanced multinomial trial

$$\text{(6.1)} \qquad \mathbf{H}_0 \colon p_j = \frac{1}{k},\ j = 1, \dots, k \quad \text{against} \quad \mathbf{H}_1 : p_j \neq \frac{1}{k} \text{ for some } j.$$

This problem is invariant with respect to the group $\mathcal{R}_k$ of permutations of the numbers $\{1, 2, \dots, k\}$. Using Theorem 4.1, we shall show that, for the hypothesis of the balanced multinomial trial, the classical χ^2-test is the locally most powerful invariant test of $\mathbf{H}_0$ against $\mathbf{H}_1$.

Denote $\mathbf{S} = \sum_{i=1}^n \mathbf{X}_i$; then $\mathbf{S} = (S_1, \dots, S_k)'$ follows the multinomial $\mathcal{M}(n; p_1, \dots, p_k)$ distribution. The optimal invariant test should be optimal over the family of probability distributions

$$\text{(6.2)} \qquad \mathcal{M}_{\mathcal{R}_k}(n; p_1, \dots, p_k) = \frac{1}{k!} \sum_{\mathbf{r} \in \mathcal{R}_k} \mathcal{M}(n; p_{r_1}, \dots, p_{r_k}),$$

where $\mathbf{r} = (r_1, \dots, r_k)$ runs over the permutations of $\{1, \dots, k\}$. The likelihood ratio of the sample from the distribution $\mathcal{M}_{\mathcal{R}_k}(n; p_1, \dots, p_k)$ with respect to $\mathcal{M}(1/k, \dots, 1/k)$ equals to

$$\text{(6.3)} \qquad \frac{1}{k!} \sum_{\mathbf{r} \in \mathcal{R}_k} \prod_{j=1}^{k} (kp_{r_j})^{S_j} = \frac{1}{k!} \sum_{\mathbf{r} \in \mathcal{R}_k} \prod_{j=1}^{k} (kp_j)^{S_{r_j}}.$$

Hence, large values of (6.3) would be significant for rejecting $\mathbf{H}_0$ against the alternative $\mathcal{M}_{\mathcal{R}_k}(n; p_1, \dots, p_k)$. We are interested to know whether there exists a test of $\mathbf{H}_0$, locally most powerful against the family of such alternatives.

Given $p_1, \dots, p_k$, consider the family of distributions

$$\text{(6.4)} \qquad \mathcal{M}_{\mathcal{R}_k}\left(n; \frac{1-\lambda}{k} + \lambda p_1, \dots, \frac{1-\lambda}{k} + \lambda p_k\right), \quad 0 \leq \lambda \leq 1.$$

Then the likelihood ratio (6.3) corresponding to (6.4) can be rewritten as

$$\text{(6.5)} \qquad \frac{1}{k!} \sum_{\mathbf{r} \in \mathcal{R}_k} \exp\left\{ \sum_{j=1}^{k} \left[S_{r_j} \log\left(1 + \lambda(kp_j - 1)\right)\right] \right\}.$$

Expanding (6.5) around $\lambda = 0$ up to the second order term, we obtain

$$\begin{aligned}(6.6)\quad & \frac{1}{k!}\sum_{\mathbf{r}\in\mathcal{R}_k} \exp\Big\{\sum_{j=1}^{k}\big[S_{r_j}\log\big(1+\lambda(kp_j-1)\big)\big]\Big\} \\ &= \frac{1}{k!}\sum_{\mathbf{r}\in\mathcal{R}_k}\Big\{1+\lambda\sum_{j=1}^{k} S_{r_j}(kp_j-1) - \frac{\lambda^2}{2}\sum_{j=1}^{k} S_{r_j}(kp_j-1)^2 \\ &\qquad + \frac{\lambda^2}{2}\Big(\sum_{j=1}^{k} S_{r_j}(kp_j-1)\Big)^2 + o(\lambda^2)\Big\}.\end{aligned}$$

For simplicity, put $kp_j - 1 = h_j$ and $S_j - n/k = V_j$, $j = 1, \dots, k$. Then

$$\sum_{j=1}^{k} h_j = 0 \quad \text{and} \quad \sum_{j=1}^{k} V_j = 0 \tag{6.7}$$

hence

$$\sum_{\mathbf{r}\in\mathcal{R}_k}\sum_{j=1}^{k} S_{r_j}(kp_j-1) = 0. \tag{6.8}$$

Moreover,

$$\sum_{\mathbf{r}\in\mathcal{R}_k}\sum_{j=1}^{k} S_{r_j}(kp_j-1)^2 = C\|\mathbf{h}\|^2, \tag{6.9}$$

where $C = C(n,k) > 0$ does not depend on S_j, $j = 1, \dots, k$ and on $\|\mathbf{h}\|^2 = \sum_{j=1}^{k} h_j^2$. It remains to treat the value of

$$\begin{aligned}\sum_{\mathbf{r}\in\mathcal{R}_k}\Big(\sum_{j=1}^{k} S_{r_j}(kp_j-1)\Big)^2 &= \sum_{\mathbf{r}\in\mathcal{R}_k}\Big(\sum_{j=1}^{k} V_{r_j}h_j\Big)^2 \\ &= \sum_{\mathbf{r}\in\mathcal{R}_k}\sum_{j=1}^{k} V_{r_j}^2 h_j^2 + \sum_{\mathbf{r}\in\mathcal{R}_k}\sum_{j\neq s} V_{r_j}V_{r_s}h_jh_s = A + B \quad \text{(say)}.\end{aligned}$$

By (6.7),

$$\begin{aligned} A &= (k-1)!\Big(\sum_{j=1}^{k} V_j^2\Big)\|\mathbf{h}\|^2, \\ B &= -(k-2)!\sum_{j\neq s} h_jh_s\sum_{z=1}^{k} V_z^2 = (k-2)!\sum_{j=1}^{k} V_j^2\|\mathbf{h}\|^2. \end{aligned} \tag{6.10}$$

Combining (6.5)–(6.10) together, we obtain

$$\frac{1}{k!}\sum_{\mathbf{r}\in\mathcal{R}_k}\exp\Big\{\sum_{j=1}^{k}\big[S_j\log\big(1+\lambda(kp_j-1)\big)\big]\Big\}$$
$$=1+\frac{\lambda^2}{2}\frac{1}{k-1}\sum_{j=1}^{k}\Big(S_j-\frac{n}{k}\Big)^2\sum_{j=1}^{k}(kp_j-1)^2+o(\lambda^2).$$

Hence, if $0<\frac{1}{2}\lambda^2\sum_{j=1}^{k}(p_j-1/k)^2\le\varepsilon$, then, by Theorem 4.1, the local power of test $\Phi(\mathbf{S})$ is bounded by

$$\mathbb{E}_\lambda(\Phi(\mathbf{S}))=\alpha+\varepsilon\frac{\lambda^2}{2}\|\mathbf{h}\|\frac{1}{k-1}\mathbb{E}\Big(\Phi(\mathbf{S})\sum_{j=1}^{k}\Big(S_j-\frac{n}{k}\Big)\Big)^2+o(\lambda^2)+o(\varepsilon^2),$$

hence the locally most powerful invariant test of $\mathbf{H}_0$ against $\mathbf{H}_1$ has the form

$$\Phi(S_1,\dots,S_k)=\begin{cases}1 & \text{if } \sum_{j=1}^{k}(S_j-n/k)^2>C_\alpha\\ \gamma & \text{if } \sum_{j=1}^{k}(S_j-n/k)^2=C_\alpha\\ \Phi(S_1,\dots,S_k)=0 & \text{if } \sum_{j=1}^{k}(S_j-n/k)^2<C_\alpha,\end{cases}\tag{6.11}$$

where $C_\alpha>0$ and $\gamma\in[0,1)$ are determined by the condition

$$\mathbb{E}_{\mathbf{H}_0}(\Phi(S_1,\dots,S_k))=\alpha.\tag{6.12}$$

This is the classical χ^2-test.

Acknowledgements. The research of J. Jurečková was supported by the Czech Republic Grant GAČR 201/02/0621 and by the Research Project CEZ: J13/98:113200008 "Mathematical Methods in Stochastics."

REFERENCES

Bondesson, L. (1974). Characterization of probability laws through constant regression. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 30, 93–115.

Hájek, J. (1970). A characterization of limiting distributions of regular estimates. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 14, 323–330.

Hájek, J. (1972). Local asymptotic minimax and admissibility in estimation. *Proc. 6th Berkeley Symposium Math. Statist. Probab.*, pp. 175–194. Univ. California Press, Berkeley, CA.

Hájek, J. and Šidák, Z. (1967). *Theory of Rank Tests.* Academia, Prague.

Jurečková, J. (1999). Equivariant estimators and their asymptotic representations. *Tatra Mt. Math. Publ.* 17, 1–9.

Jurečková, J. and Milhaud, X. (1994). Shrinkage of maximum likelihood estimator of multivariate location. *Asymptotic Statistics* (P. Mandl and M. Hušková, eds.), pp. 303–318. Physica, Heidelberg.

Jurečková, J. and Milhaud, X. (1999). Characterization of distributions in invariant models. *J. Statist. Plann. Inference* 75, 353–361.

LeCam, L. and Yang, G.L. (1990). *Asymptotic Statistics. Some Basic Concepts.* Springer-Verlag, New York.

Lehmann, E. L. (1986). *Testing Statistical Hypotheses*, 2nd ed. Wiley, New York.

Mukherjee, R. (1993). Rao's score test: Recent asymptotic results. *Econometrics* (G.S. Maddala, C.R. Rao and H.D. Vinod, eds.), pp. 363–379. *Handbook of Statistics*, vol. 11. Elsevier, Amsterdam.

Pollard, D., Yang, G. and Torgersen, E. (eds.) (1997). *Festschrift for Lucien LeCam.* Springer-Verlag, New York.

Jana Jurečková
Dept. of Probability and Statistics
Charles University
Sokolovska 83
CZ-186 75 Prague 8
Czech Republic
jurecko@karlin.mff.cuni.cz

Xavier Milhaud
Laboratoire de Statistique et Probabilités
Université Paul Sabatier
118, Route de Narbonne
31062 Toulouse Cedex 4
France
milhaud@cict.fr