Journal of ELECTRICAL ENGINEERING, VOL. 51, NO. 12/s, 2000, 46–50

ON THE COMPLEMENTARITY OF MAXIMUM LIKELIHOOD AND MiniMax ENTROPY

Marian Grendár* — Marián Grendár**

The article 1) extends the notion of the exponential family into a general exponential form, 2) utilizes an analogy between Boltzmann's deduction of equilibrium distribution of ideal gas in an external potential field and a probability density function, and consequently, based on it, 3) investigates a complementary relationship between Maximum Likelihood (ML) and Maximum Entropy (MaxEnt) methods. MiniMaximization of Entropy (MiniMax Ent), for the case of parametric inverse problem is proposed, and demonstrated to be complementary to ML on the general exponential form. The complementary relationship with ML method seems to be a specific property of Shannon's entropy (mini)maximization.

K e y w o r d s: simple and general exponential form/potentials, ML task, ME task, MiniMax Entropy task, complementarity

2000 *Mathematics Subject Classification*: 62A10, 62A99, 62F10

1 INTRODUCTION

Maximization of the likelihood function is a well-established method of finding estimators with desirable asymptotical properties. The method relies upon a sample and an assumed pmf/pdf the sample came from. The choice of the sample generating pmf/pdf is ambiguous. Shannon's entropy maximization should be, in order to bring a nontrivial result, confined by some constraints, and it is where its ambiguity lies.

2 DEFINITIONS AND NOTATION

The notion of exponential family is extended into simple and general exponential forms.

Definition 1. Let X be a random variable with pmf/pdf $f_X(x)$. If $f_X(x)$ can be written in the form of

$$f_X(x|\boldsymbol{\lambda}) = k(\boldsymbol{\lambda})e^{-U(x,\boldsymbol{\lambda})},$$

where $U(x,\boldsymbol{\lambda}) = \boldsymbol{\lambda}'\mathbf{u}(x)$ is a linear combination of functions $\mathbf{u}(x)$ depending on other parameters, and $k(\boldsymbol{\lambda})$ is a normalizing factor, then it has a *simple exponential form.* $u(x)$ is called *simple potential.* If the pmf/pdf can be written in the form of

$$f_X(x|\boldsymbol{\lambda},\boldsymbol{\alpha}) = k(\boldsymbol{\lambda},\boldsymbol{\alpha})e^{-U(x,\boldsymbol{\lambda},\boldsymbol{\alpha})}$$

where $U(x,\boldsymbol{\lambda},\boldsymbol{\alpha}) = \boldsymbol{\lambda}'\mathbf{u}(x,\boldsymbol{\alpha})$ is a linear combination of functions $\mathbf{u}(x,\boldsymbol{\alpha})$ depending on other parameters $\boldsymbol{\alpha}$, and $k(\boldsymbol{\lambda},\boldsymbol{\alpha})$ is a normalizing factor, then it has a *general exponential form.* $u(x,\boldsymbol{\alpha})$ is called the *general potential.* The $U(\cdot)$ function is called the *total potential.*

N o t e . Any class of pmf/pdf which can be written in the exponential form is equivalently characterized by its exponential form pmf/pdf or by its potentials.

Example 1. $\Gamma(\alpha,\beta)$ distribution has a simple exponential form, with total potential $U(x,\boldsymbol{\lambda}) = \lambda_1 x + \lambda_2 \ln x$; $\lambda_1 = \frac{1}{\beta}$ and $\lambda_2 = 1-\alpha$; $u_1(x) = x$ and $u_2(x) = \ln x$ are the potentials. The normalizing factor $k(\lambda_1,\lambda_2) = \frac{1}{\Gamma(1-\lambda_2)\lambda_1^{\lambda_2-1}}$. *Logistic* (μ,β) distribution has a general exponential form with a total potential $U(x,\boldsymbol{\lambda},\boldsymbol{\alpha}) = \lambda_1 u_1(x,\boldsymbol{\alpha}) + \lambda_2 u_2(x,\boldsymbol{\alpha})$, with $\boldsymbol{\lambda} = [\frac{1}{\alpha_2}, 2]$, and the potentials $u_1(\cdot) = \frac{x-\alpha_1}{\alpha_2}$, $u_2(\cdot) = \ln(1+e^{-\frac{x-\alpha_1}{\alpha_2}})$, and $\boldsymbol{\alpha} = [\mu,\beta]$. $k(\alpha_2) = 1/\alpha_2$. Discrete normal distribution $dn(\lambda,\alpha)$, defined over a support by $f_X(x_i|\lambda) = e^{-\lambda(x_i-\alpha)^2}/\sum_i e^{-\lambda(x_i-\alpha)^2}$ has a total potential $U(x,\lambda,\alpha) = \lambda(x-\alpha)^2$. It can be equivalently expressed in a simple form with $U(x,\lambda_1,\lambda_2) = \lambda_1 x + \lambda_2 x^2$, where $\lambda_1 = -2\alpha\lambda$ and $\lambda_2 = \lambda$.

Standard definitions of the moment and sample mean are extended.

Definition 2. V-moment of random variable X, $\mu(V)$, is for any function $V(X,\boldsymbol{\alpha})$ defined as

$$\mu(V) = \mathrm{E}\,V(X,\boldsymbol{\alpha})$$

* Institute of Measurement Science, Slovak Academy of Sciences, Dúbravská cesta 5, 852 19 Bratislava, Slovakia, e-mail: umergren@savba.sk

** Railways of the Slovak Republic, Klemensova 8, 813 61 Bratislava, Slovakia, e-mail: grendar.marian@zsr.sk

It is a pleasure to thank George Judge (UC Berkeley), Gejza Wimmer (MI SAS) and Viktor Witkovský (IMS SAS) for valuable discussions.

ISSN 1335-3632 © 2000 FEI STU

Definition 3. Sample V-moment of random variable X, $m(V)$, is for any function $V(X,\boldsymbol{\alpha})$ defined as

$$m(V)=\sum_{i=1}^{m} r_i V(X_i,\boldsymbol{\alpha})$$

where r_i is the frequency of i-th element of support in the sample.

Definition 4. Let $\mu(V)$, $m(V)$ be V-moment and sample V-moment, respectively. Then the requirement

$$\mu(V)=m(V)$$

will be called *V-moment consistency condition.*

N o t a t i o n . $\boldsymbol{\lambda}$, $\mathbf{u}(\cdot)$, $\boldsymbol{\mu}(\cdot)$ and $\mathbf{m}(\cdot)$ are $[J,1]$ vectors, indexed by j. $\mathbf{x}$, $\mathbf{p}$ and $\mathbf{r}$ are $[m,1]$ vectors, indexed by i, with m finite or infinite. $\boldsymbol{\alpha}$ is $[T,1]$ vector indexed by t.

Since entropy maximization can be reasonably constrained by constraints other than the moment consistency constraints (see for instance [2], or proceedings of MaxEnt conferences), in order to be specific, we will speak about a *ME task*. Also, *ML task* is defined.

Definition 5. *ML task on* $f_X(x|\boldsymbol{\theta})$. Let $X_1,X_2,\dots,X_n$ be a random sample from population $f_X(x|\boldsymbol{\theta})$. The maximum likelihood task is to find the maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ for a given sample.

Definition 6. *ME task on* $\mathbf{u}(\cdot)$. Given a sample and a vector of known potential functions $\mathbf{u}(\cdot)$, the maximum entropy task is to find the most entropic distribution $\mathbf{p}$ consistent with a set of $\mathbf{u}$-moment consistency conditions.

3 ML TASK AND ME TASK

Simple exponential form, simple potential case

Theorem 1. Complementarity of ML and ME tasks, identity of solutions.

Let $X_1,X_2,\dots,X_n$ be a random sample. Then

(i) complementarity of tasks

a) *ML estimator $\hat{\boldsymbol{\lambda}}$ of $\boldsymbol{\lambda}$ on simple exponential form $f_X(x|\boldsymbol{\lambda}) = k(\boldsymbol{\lambda})e^{-\boldsymbol{\lambda}'\mathbf{u}}$ is obtained as a solution of system of J u_j-moment consistency conditions,*

b) *the most entropic distribution $\mathbf{p}$ satisfying the system of J u_j-moment consistency conditions is the simple exponential form pmf/pdf $f_X(x|\hat{\boldsymbol{\lambda}})$.*

(ii) identity of solutions

necessary and sufficient conditions for ML task on simple exponential form pmf or pdf $f_X(x|\boldsymbol{\lambda})=k(\boldsymbol{\lambda})e^{-\boldsymbol{\lambda}'\mathbf{u}(x)}$ and ME task on the simple potentials $\mathbf{u}(x)$ are identical, and they are

$$\mu(u_j)=m(u_j) \qquad j=1,2,\dots,J \tag{1}$$

P r o o f . For the proof see [3].

N o t e . ML task on simple exponential form and ME task on simple potentials are complementary in the sense that where one starts the other one ends, and vice versa. ML starts with a known simple exponential form of pmf/pdf and ends up with ML estimators of the parameters, found out of the potential moment consistency equations. ME, working on the sample, starts with an assumed form of potential functions, forming potential moment consistency constraints. The most entropic distribution resolved is just the exponential form pmf/pdf ML has assumed. And the ME estimators of its parameters are the same as the ML estimators. We say that ML task on simple exponential form pmf/pdf and ME task on simple potentials are *complementary.*

ML and ME tasks are complementary in set-up but identical in solution. Both the tasks end up with the same mathematical problem of solving estimators of $\boldsymbol{\lambda}$ out of the system of potential moment consistency equations (1).

Example 2. Let $X_1,X_2,\dots,X_n$ be a random sample of size n from discrete normal distribution $dn(\lambda_1,\lambda_2)$, taken in the simple exponential form. ML task of estimation leads to solving λ_1,λ_2 out of the system of equations

$$\frac{\sum_{i=1}^{m} x_i e^{-(\lambda_1 x_i+\lambda_2 x_i^2)}}{\sum_{i=1}^{m} e^{-(\lambda_1 x_i+\lambda_2 x_i^2)}}=\sum_{i=1}^{m} r_i x_i$$
$$\frac{\sum_{i=1}^{m} x_i^2 e^{-(\lambda_1 x_i+\lambda_2 x_i^2)}}{\sum_{i=1}^{m} e^{-(\lambda_1 x_i+\lambda_2 x_i^2)}}=\sum_{i=1}^{m} r_i x_i^2$$

which is just the system of x-moment and x^2-moment consistency conditions.

ME task constrained by the system of x-moment, and x^2-moment consistency conditions

$$\begin{aligned}\sum_{i=1}^{m} p_i x_i &= \sum_{i=1}^{m} r_i x_i\\ \sum_{i=1}^{m} p_i x_i^2 &= \sum_{i=1}^{m} r_i x_i^2\end{aligned} \tag{2}$$

finds the most entropic distribution consistent with the constraints to have form (after normalization)

$$p_i=\frac{e^{-(\lambda_1 x_i+\lambda_2 x_i^2)}}{\sum_{i=1}^{m} e^{-(\lambda_1 x_i+\lambda_2 x_i^2)}} \tag{3}$$

where, λ_1,λ_2 should be found out of the system (2), after plugging (3) in.

In passing we mention an identity of ML and modified method of moments (MMM) in the case of an exponential family discovered by [5] and explored further by [1]. The identity holds also for the simple exponential form, making ME complementary to both ML and MMM. Note that MMM starts with moment consistency conditions, where the understanding of moments is enhanced as done here by Definitions 2, 3, 4.

General exponential form, general potential case

Complementarity of the general exponential form ML task and general potential ME task can not be assessed analytically in full extent, for sufficient conditions for maximum of likelihood or entropy function do not allow, in general, for it. We show, analytically, that ML task on the general exponential form and ME task on the general potentials lead to the same FOC's. This could be called 'weak complementarity'.

Theorem. *Let $X_1, X_2, \ldots, X_n$ be a random sample. Then, necessary conditions for*

a) *ML task on general exponential form pmf/pdf* $f_X(x|\boldsymbol{\lambda}, \boldsymbol{\alpha}) = k(\boldsymbol{\lambda}, \boldsymbol{\alpha}) e^{-\boldsymbol{\lambda}' \mathbf{u}(x, \boldsymbol{\alpha})}$

b) *ME task on the general potentials* $\mathbf{u}(x, \boldsymbol{\alpha})$

are identical, and they are

$$\mu(u_j) = m(u_j) \qquad j = 1, 2, \ldots, J\,,$$
$$\boldsymbol{\lambda}'\boldsymbol{\mu}\Big(\frac{\partial \mathbf{u}}{\partial \alpha_t}\Big) = \boldsymbol{\lambda}'\mathbf{m}\Big(\frac{\partial \mathbf{u}}{\partial \alpha_t}\Big) \qquad t = 1, 2, \ldots, T\,.$$

P r o o f . Discrete r.v. case.

1. ML task.

$$\max_{\boldsymbol{\lambda}, \boldsymbol{\alpha}} l(\boldsymbol{\lambda}, \boldsymbol{\alpha}) = \ln(k(\boldsymbol{\lambda}, \boldsymbol{\alpha})) - \sum_{j=1}^{J} \sum_{i=1}^{m} \lambda_j r_i u_j(x_i, \boldsymbol{\alpha})$$

leads to system of $J + T$ first order conditions

$$\begin{aligned} \mu(u_j) &= m(u_j) \qquad j = 1, 2, \ldots, J \\ \boldsymbol{\lambda}'\boldsymbol{\mu}\Big(\frac{\partial \mathbf{u}}{\partial \alpha_t}\Big) &= \boldsymbol{\lambda}'\mathbf{m}\Big(\frac{\partial \mathbf{u}}{\partial \alpha_t}\Big) \qquad t = 1, 2, \ldots, T\,. \end{aligned} \tag{4}$$

2. ME task.

$$\max_{\mathbf{p}(\boldsymbol{\alpha})} H(\mathbf{p}(\boldsymbol{\alpha})) = -\sum_{i=1}^{m} p_i \ln p_i$$

subject to $\mu(u_j) = m(u_j)\,, \quad j = 1, 2, \ldots, J$
which can be accomplished by means of Lagrangean

$$L(\mathbf{p}(\boldsymbol{\alpha})) = -\sum_{i=1}^{m} p_i \ln p_i + \sum_{j=1}^{J} \lambda_j (m(u_j) - \mu(u_j))$$

leading to the system of $m + T$ FOC's

$$\begin{aligned} p_i &= e^{-\boldsymbol{\lambda}' \mathbf{u}(x_i, \boldsymbol{\alpha})} \qquad i = 1, 2, \ldots, m \\ -\sum_{i=1}^{m} \Big(\frac{\partial p_i}{\partial \alpha_t} \ln p_i + \frac{\partial p_i}{\partial \alpha_t}\Big) &+ \sum_{j=1}^{J} \lambda_j \Big(\frac{\partial m(u_j)}{\partial \alpha_t} \\ &- \sum_{i=1}^{m} \{p_i \frac{\partial u_j(x_i, \boldsymbol{\alpha})}{\partial \alpha_t} + \frac{\partial p_i}{\partial \alpha_t} u_j(x_i, \boldsymbol{\alpha})\}\Big) = 0 \quad \forall t\,. \end{aligned} \tag{5}$$

The most entropic distribution after normalization takes general exponential form

$$p_i = \frac{e^{-\boldsymbol{\lambda}' \mathbf{u}(x_i, \boldsymbol{\alpha})}}{\sum_{i=1}^{m} e^{-\boldsymbol{\lambda}' \mathbf{u}(x_i, \boldsymbol{\alpha})}} \qquad i = 1, 2, \ldots, m$$

where 'ME estimators' of λ have to be found out of the system (5).

The T of equations of the system (5) simplifies heavily into

$$\boldsymbol{\lambda}'\boldsymbol{\mu}\Big(\frac{\partial \mathbf{u}}{\partial \alpha_t}\Big) = \boldsymbol{\lambda}'\mathbf{m}\Big(\frac{\partial \mathbf{u}}{\partial \alpha_t}\Big) \qquad t = 1, 2, \ldots, T$$

which are the same as the T equations of FOC's for ML task (4). Thus, the ME and ML tasks indeed lead to the same necessary conditions (4).

Continuous r.v. case.

In analogy to the proof of Theorem 1.

Corollary. *Due to the linearity of $U(x, \boldsymbol{\lambda}, \boldsymbol{\alpha})$ in $\boldsymbol{\lambda}$, the necessary conditions (4) can be rewritten in a compact form*

$$\mu\Big(\frac{\partial U}{\partial \lambda_j}\Big) = m\Big(\frac{\partial U}{\partial \lambda_j}\Big) \qquad j = 1, 2, \ldots, J\,,$$
$$\mu\Big(\frac{\partial U}{\partial \alpha_t}\Big) = m\Big(\frac{\partial U}{\partial \alpha_t}\Big) \qquad t = 1, 2, \ldots T\,.$$

Example 3. Let $X_1, X_2, \ldots, X_n$ be a random sample from discrete normal distribution $dn(\lambda, \alpha)$, taken in the general exponential form, so $u(x, \alpha) = (x - \alpha)^2$.

ML task of estimation leads to solving λ, α out of the system of equations

$$\begin{aligned} \mu(u) &= m(u)\,, \\ \mu\Big(\frac{\partial u}{\partial \alpha_t}\Big) &= m\Big(\frac{\partial u}{\partial \alpha_t}\Big)\,. \end{aligned} \tag{6}$$

ME task constrained by moment consistency condition

$$\sum_{i=1}^{m} p_i (x_i - \alpha)^2 = \sum_{i=1}^{m} r_i (x_i - \alpha)^2$$

leads to the FOC's

$$p_i = e^{-\lambda (x_i - \alpha)^2}$$
$$\mu\Big(\frac{\partial u}{\partial \alpha_t}\Big) = m\Big(\frac{\partial u}{\partial \alpha_t}\Big)$$

where λ, α has to be found out of (6), after normalizing p's.

So, ML and ME tasks lead to the same necessary conditions. Also, note that the ML and ME estimators are the same as in Example 2, where $dn(\cdot)$ was taken in the simple exponential form.

Sufficient conditions do not allow, in general, for assessing the kind of extremum attained in the points chosen out by FOC's.

Theorem 3. *Second derivatives for the ML task are*

$$\frac{\partial^2 l(\boldsymbol{\lambda},\boldsymbol{\alpha})}{\partial\lambda_j^2} = -\operatorname{Var}(U'_{\lambda_j}),$$

$$\frac{\partial^2 l(\boldsymbol{\lambda},\boldsymbol{\alpha})}{\partial\lambda_j\partial\lambda_\iota} = -\operatorname{Cov}(U'_{\lambda_j}U'_{\lambda_\iota}),$$

$$\frac{\partial^2 l(\boldsymbol{\lambda},\boldsymbol{\alpha})}{\partial\alpha^2} = -\big(\operatorname{Var}(U'_\alpha) + m(U''_\alpha) - \mu(U''_\alpha)\big),$$

$$\frac{\partial^2 l(\boldsymbol{\lambda},\boldsymbol{\alpha})}{\partial\alpha_t\partial\alpha_\tau} = -\mu(U'_{\alpha_t})\mu(U'_{\alpha_\tau}) - \sum_{j=1}^{J}\lambda_j\mu(U''_{\lambda_j\alpha_t}U''_{\alpha_t\alpha_\tau}) - m(U''_{\alpha_t\alpha_\tau}) + \mu(U''_{\alpha_t\alpha_\tau}),$$

$$\frac{\partial^2 l}{\partial\lambda_j\partial\alpha_t} = -\lambda_j\operatorname{Cov}(U'_{\lambda_j},U''_{\lambda_j\alpha_t}) - \sum_{k\neq j}^{J}\lambda_k\mu(U'_{\lambda_j}U''_{\lambda_k\alpha_t}) - m(U''_{\lambda_j\alpha_t}) + \mu(U''_{\lambda_j\alpha_t})$$

and for the ME task they are

$$\frac{\partial^2 L(\mathbf{p}(\boldsymbol{\alpha}))}{\partial p_i^2} = -\frac{1}{p_i},$$

$$\frac{\partial^2 L(\mathbf{p}(\boldsymbol{\alpha}))}{\partial\alpha_t^2} = \operatorname{Var}(U'_{\alpha_t}) + m(U''_{\alpha_t}) - \mu(U''_{\alpha_t}),$$

$$\frac{\partial^2 L(\mathbf{p}(\boldsymbol{\alpha}))}{\partial\alpha_t\partial\alpha_\tau} = m(U''_{\alpha_t\alpha_\tau}) - \mu(U''_{\alpha_t\alpha_\tau}) + \operatorname{Cov}(U'_{\alpha_t},U'_{\alpha_\tau}).$$

P r o o f . Differentiating twice the loglikelihood function, and the Lagrange function lead to the stated results.

In the following simple instance of the general potential the sufficient conditions are analytically tractable, showing that at the points chosen by the necessary conditions (4) entropy function attains its *maximum* in $\mathbf{p}(\boldsymbol{\alpha})$, and *minimum* in $\boldsymbol{\alpha}$, hence the chosen distribution has a minimal entropy in the class of the most entropic distributions, consistent with the moment consistency constraints. Likelihood function at the points attains its maximum.

Example 4. Find the sufficient conditions for the Example 3 set-up.

The general total potential is $U(x,\lambda,\alpha) = \lambda(x-\alpha)^2$, so the potential is $u(x,\alpha) = (x-\alpha)^2$. The second derivatives stated in the above Theorem then simplifies into

$$\frac{\partial^2 l(\boldsymbol{\lambda},\alpha)}{\partial\lambda^2} = -\operatorname{Var}(u)$$

$$\frac{\partial^2 l(\boldsymbol{\lambda},\alpha)}{\partial\alpha^2} = -(\lambda^2\operatorname{Var}(u'_\alpha) + \lambda(m(u''_\alpha) - \mu(u''_\alpha)))$$

$$\frac{\partial^2 l(\boldsymbol{\lambda},\alpha)}{\partial\lambda\partial\alpha} = -(\lambda\operatorname{Cov}(u,u'_\alpha) + m(u'_\alpha) - \mu(u'_\alpha))$$

for the ML task, and into

$$\frac{\partial^2 L(\mathbf{p}(\alpha))}{\partial p_i^2} = -\frac{1}{p_i}$$

$$\frac{\partial^2 L(\mathbf{p}(\alpha))}{\partial\alpha^2} = \lambda^2\operatorname{Var}(u'_\alpha) + \lambda(m(u''_\alpha) - \mu(u''_\alpha))$$

for the ME task. Furthermore, in this case

$$m(u''_\alpha) - \mu(u''_\alpha) = 0$$

and also, due to the FOC's (4)

$$m(u'_\alpha) - \mu(u'_\alpha) = 0$$

Thus, the second derivatives for the ML task form a hessian matrix

$$H_{ML} = -\begin{pmatrix} \operatorname{Var}(u) & \lambda\operatorname{Cov}(u,u') \\ \lambda\operatorname{Cov}(u,u') & \lambda^2\operatorname{Var}(u') \end{pmatrix}$$

which is negative definite, assuring in this case, that the global maximum was attained.

ME task second derivatives are

$$\frac{\partial^2 L(\mathbf{p}(\alpha))}{\partial p_i^2} = -\frac{1}{p_i}$$

$$\frac{\partial^2 L(\mathbf{p}(\alpha))}{\partial\alpha^2} = 4\lambda^2\operatorname{Var}(x)$$

showing that entropy attains its maximum in distribution $\mathbf{p}$, and minimum in α, at the same point where likelihood attains its maximum.

This result was also supported by numerical investigations, elucidating the behavior. In the α suggested by FOC's entropy function attains its *minimum*, whilst the maximum is attained for an $\tilde{\alpha}$ degenerating $\mathbf{p}$ into an uniform distribution. No surprise, since the value of parameter α of $u(x,\alpha)$ is free to choose, and attaining the goal of maximal entropy the value is set up such that the uniform distribution is reached.

The above analytically tractable case of the sufficient conditions and several numerical investigations of more complex general potentials lead us to propose a *hypothesis* about complementarity of ML and *MiniMax Entropy* tasks and identity of their solutions, under the general exponential form, general potentials.

For the sake of completeness, the MiniMax Ent task is defined.

Definition. *MiniMax Entropy task.* Given a sample and a vector of known general potentials $\mathbf{u}(x,\boldsymbol{\alpha})$, the MiniMax Entropy task is to find in the class of all most entropic distributions $\mathbf{p}(\boldsymbol{\alpha})$ consistent with the set of $\mathbf{u}$-moment consistency conditions, a pmf/pdf with minimal entropy.

N o t e . If the potentials are simple, MiniMax Ent task reduces into the ME task on simple potentials.

4 SUMMARY

The complementary tasks of likelihood maximization on the general exponential form, and Shannon's entropy minimaximization on the general potentials, are in favor of each other, in a circular way; supporting jointly the exponential form.

To the both, the question about the choice of the potential functions remains open.

Finally, we would like to note that the complementary relationship of (Mini)MaxEnt with ML seems to be a specific property of the Shannon entropy function. In [4] we have shown that the so-called maximum empirical likelihood (MEL) criterion constrained by moment consistency constraints, proposed by [6] in the context of noiseless linear inverse problem, is not complementary with ML on the MEL recovered class of pmf/pdf.

References

[1] DAVIDSON, R. R.—SOLOMON, D. L.: Moment-Type Estimation in the Exponential Family, Communications in Statistics **3** (1974), 1101–1108.

[2] GOLAN, A.—JUDGE, G.—MILLER, D.: Maximum Entropy Econometrics. Robust Estimation With Limited Data, John Wiley & Sons, New York, 1996.

[3] GRENDÁR, M.—GRENDÁR, M.: Maximum Likelihood and Minimax Entropy. Complementarity of Tasks, Identity of Solutions, Tech Rep IMS SAS, TR 2/1999.

[4] GRENDÁR, M.—GRENDÁR, M.: Criterion Choice Problem. An ML-complementarity approach., Technical Report TR 1/2000 of IMS SAS.

[5] HUZURBAZAR, V. S.: On a Property of Distributions Admitting Sufficient Statistics, Biometrika **36** (1949), 71–74.

[6] MITTELHAMMER, R.—JUDGE, G.—MILLER, D.: Econometric Foundations, Cambridge University Press, New York, 2000 (to appear).

Received 14 June 2000

Marian Grendár (Mgr) is with the Institute of Measurement Science of Slovak Academy of Sciences (SAS) and a graduate student of mathematical statistics at the Mathematical Institute of SAS.

Marián Grendár (Ing) is a communication engineer at the RSR.