Math-Net.Ru

Общероссийский математический портал

О. Е. Barndorff-Nielsen, L -nonformation, L -ancillarity, and L -sufficiency, *Теория вероятн. и ее примен.*, 1999, том 44, выпуск 1, 225–229

DOI: <https://doi.org/10.4213/tvp620>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.197.217.227

26 апреля 2023 г., 15:46:38

11. Varadhan S. R. S. Large Deviations and Applications, Philadelphia: SIAM, 1984.
12. Wentzell A. D. Limit Theorems in Large Deviations for Markov Random Processes. Dordrecht: Reidel, 1984.

© 1999 г.

BARNDORFF-NIELSEN O. E.*
L-NONFORMATION, L-ANCILLARITY,
AND L-SUFFICIENCY

Определяется понятие отсутствия информации об интересующем параметре *L*-неинформативности. Из него вытекают понятия *L*-подчиненности и *L*-достаточности (последнее у нее было изучено ранее). В статье *L* означает правдоподобие. В предлагаемой работе обсуждаются свойства этих понятий.

Ключевые слова и фразы: *L*-неинформативность (*L*-nonformation), *L*-достаточность (*L*-sufficiency), *L*-подчиненность (*L*-ancillarity), правдоподобие, параметризация, информация.

1. Introduction. Rémon [6] introduced a concept of generalized sufficiency. He called it *L*-sufficiency, but for reasons of consistency with other notation that term is now being used in a restricted sense (cf. [2], [5]). Here *L* as a prefix is short for likelihood.

Inherent in the concept of *L*-sufficiency is a notion of «no information about a parameter of interest». In the present paper this notion is extracted and termed *L*-nonformation. In addition to *L*-sufficiency, it leads to a concept of generalized ancillarity, to be termed *L*-ancillarity.

L-nonformation is defined and discussed in Section 2 and its relation to *L*-sufficiency is considered in Section 3. Section 4 introduces the concept of *L*-ancillarity and derives some properties of this, particularly in the context of exponential models. The rest of the present section sets out a few preliminaries.

Let $p(x; \omega)$ be the model function of a parametric model M and let (ψ, χ) be a partition of the parameter ω into an interest parameter ψ and a nuisance parameter χ . We assume that ψ and χ are variation independent (i.e., $\Omega = \Psi \times X$, where Ω , Ψ , and X are the domains of variation of ω , ψ , and Ξ , respectively). Reparametrizations $(\psi, \chi) \rightarrow (\varphi, \lambda)$ of M of the form

$$\varphi = \varphi(\psi), \quad \lambda = \lambda(\psi, \chi)$$

are called *interest respecting reparametrizations*.

The maximum likelihood estimate for χ when ψ is considered as known is denoted by $\hat{\chi}_\psi$ and for any combinant $q = q(x; \omega)$ we use the convention that the symbols « $\hat{}$ » and « $\tilde{}$ » placed above q indicate evaluation at the full and partial maximum likelihood estimate, respectively. That is, $\hat{q} = q(x; \hat{\omega})$ and $\tilde{q} = q(x; \psi, \hat{\chi}_\psi)$. The evaluation is taken to be the last operation in the calculation of the quantity concerned, in order to avoid possible ambiguities.

We denote the profile likelihood for the interest parameter ψ by $L_P(\psi)$ or by $\tilde{L}(\psi)$, i. e., $L_P(\psi) = L(\psi, \hat{\chi}_\psi) = \tilde{L}(\psi)$. Furthermore, the normed version of the profile likelihood for ψ will be denoted by $\bar{L}_P(\psi)$, i. e.,

$$\bar{L}_P(\psi) = \frac{\tilde{L}(\psi)}{\tilde{L}(\hat{\psi})}.$$

Let u and v be statistics. For given value of ψ we write $\hat{\chi}_\psi(\tilde{v} \mid \tilde{u})$ for the maximum likelihood estimate of χ based on the conditional distribution of v given u , and we denote

* Institute of Mathematics, Aarhus University, DK-8000 Aarhus C, Denmark.

the profile likelihood for ψ based on the conditional distribution of v given u by $L_P(\psi; v | u)$, i.e.,

$$L_P(\psi; v | u) = L(\psi, \hat{\chi}_\psi(v | u); v | u).$$

We shall call any combinant of the form $c(x)\tilde{L}(\psi; v | u)$ a *version* of the profile likelihood for ψ based on the conditional distribution of v given u , the point here being that we allow arbitrary functions c of x , and not just functions of (u, v) , as factors. Further, we introduce the notation

$$Q(\psi, \psi'; v | u) = \frac{L_P(\psi; v | u)}{L_P(\psi'; v | u)} \quad (1.1)$$

for the relative likelihoods.

It is of some interest to write the likelihood function of the original model as

$$L(\psi, \chi) = \tilde{L}(\psi) \bar{L}(\chi | \psi), \quad (1.2)$$

where

$$\bar{L}(\chi | \psi) = \frac{L(\psi, \chi)}{\tilde{L}(\psi)} \quad (1.3)$$

is the normed likelihood function for χ when ψ is assumed to be known. The observed information matrices calculated respectively from $\tilde{L}(\psi)$ and $\bar{L}(\chi | \psi)$ yield, by inversion, the standard estimates of the asymptotic variance matrices of $\hat{\psi}$ and of $\hat{\chi}$ given $\hat{\psi}$, under the asymptotic normal distribution of $(\hat{\psi}, \hat{\chi})$.

Finally, we shall denote arbitrary functions by *fn*. Thus, if we write, for instance, $Q(\psi, \psi'; v | u) = \text{fn}(\psi, \psi', u)$ this means that $Q(\psi, \psi'; v | u)$ is a function of ψ , ψ' , and u only.

2. L -exhaustivity and L -nonformation. Let w be a function of u and v , i.e., $w = w(u, v)$. If

$$Q(\psi, \psi'; v | u) = \text{fn}(\psi, \psi', u, w) \quad (2.1)$$

then w is said to be *L -exhaustive* for the interest parameter ψ relative to v and given u .

In the special case where $\psi = \omega$ (i.e., there is no nuisance parameter), u is constant (i.e., no conditioning) and $v = x$, the statistic w is *L -exhaustive* if and only if it is sufficient in the classical sense.

Another special case is where the right-hand side of (2.1) does not depend on w , i.e.,

$$Q(\psi, \psi'; v | u) = \text{fn}(\psi, \psi', u). \quad (2.2)$$

We then say that the conditional distribution of v given u is *L -nonformative* about ψ .

Remark 2.1. We interpret *L -nonformation* as meaning that there is no information, in the sense of likelihood, about ψ in the conditional distribution of v given u (provided no other information is available to the statistician than the form of this distribution and the values of u and v).

Remark 2.2. Note that the definition of *L -nonformativity* is independent of which version of the likelihood function $L(\psi, \chi; v | u)$ one considers and is invariant under interest respecting reparametrizations.

Remark 2.3. Trivially, if $L_P(\psi; v | u)$ does not depend on ψ then the conditional distribution of v given u is *L -nonformative* with respect to ψ . This occurs in the definitions of *S -sufficiency* and *S -ancillarity*, cf. below.

We now consider the particular case where it is known that the marginal distribution of u depends on ψ only and where $v = x$.

Theorem 2.1. Suppose the marginal distribution of u depends on ψ only and take $v = x$.

Then $\hat{\chi}_\psi(x | u) = \hat{\chi}_\psi(x)$.

Furthermore, the following four conditions are equivalent:

- (i) the conditional distribution of x given u is *L -nonformative* about ψ ;
- (ii) the ratio $\tilde{L}(\psi)/\tilde{L}(\psi')$ depends on ψ , ψ' , and u only;
- (iii) there is a version of the profile likelihood for ψ which depends on ψ and u only;
- (iv) the normed profile likelihood for ψ depends on ψ and u only.

When this is the case, the maximum likelihood estimate $\hat{\psi}$ of ψ depends on u only.

Note that (i), (ii), and (iv) above can be expressed as:

(i) $Q(\psi, \psi'; v | u) = \text{fn}(\psi, \psi', u)$;

(ii) $\tilde{L}(\psi)/\tilde{L}(\psi') = \text{fn}(\psi, \psi', u)$;

(iv) $\bar{L}_P(\psi) = \text{fn}(\psi, u)$.

P r o o f. We have

$$p(x; \psi, \chi) = p(u; \psi) p(x; \psi, \chi | u) \quad (2.3)$$

and hence $\hat{\chi}_\psi(x | u) = \hat{\chi}_\psi (= \hat{\chi}_\psi(x))$ and

$$\tilde{L}(\psi) = L(\psi; u) L(\psi, \hat{\chi}_\psi; x | u). \quad (2.4)$$

Consequently

$$\frac{\tilde{L}(\psi)}{\tilde{L}(\psi')} = \left\{ \frac{L(\psi; u)}{L(\psi'; u)} \right\} Q(\psi, \psi'; x | u). \quad (2.5)$$

Since the term in curly brackets in (2.5) depends on ψ , ψ' , and u only, the equivalence of (i) and (ii) follows.

Conditions (ii) and (iii) are clearly equivalent, and they imply that $\hat{\psi}$ depends on u only. By the latter property and the fact that on setting $\psi' = \hat{\psi}$ in (2.5) we have

$$\bar{L}_P(\psi) = \left\{ \frac{L(\psi; u)}{L(\hat{\psi}; u)} \right\} Q(\psi, \hat{\psi}; x | u)$$

we find that (i) (and hence (ii) and (iii)) entails (iv).

Finally, it is obvious that (iv) implies (ii).

3. *L*-sufficiency.

D e f i n i t i o n 3.1. A statistic u is said to be *L*-sufficient for the interest parameter ψ if the following two conditions are satisfied:

(i) the marginal distribution of u depends on ψ only,

(ii) the conditional distribution of x given u is *L*-nonformative about ψ .

(The original definition of *L*-sufficiency, given by Rémon in [6] corresponds to employing condition (ii) only.)

R e m a r k 3.1. *S*-sufficiency (as defined in [1]) is a special instance of *L*-sufficiency, corresponding to the case where (i) holds and the conditional distribution given u does not depend on ψ .

R e m a r k 3.2. The concept of *L*-sufficiency also generalizes that of *G*-sufficiency (see [2, Section 1.9]).

R e m a r k 3.3. Theorem 2.1 provides three alternatives to condition (ii) of the definition of *L*-sufficiency (and also shows that when u is *L*-sufficient for ψ then $\hat{\psi}$ depends on u only). Those alternatives are important because they involve only the profile likelihood for ψ and not any prior knowledge about existence of u . Thus they can be used constructively to find a possible *L*-sufficient statistic. Specifically, given an interest parameter ψ one calculates the normed profile likelihood for ψ and inspects this to see if it depends on some function u of x only. If so and if the marginal distribution of u depends on ψ only, then u is *L*-sufficient for ψ .

4. *L*-ancillarity.

D e f i n i t i o n 4.1. A statistic u is said to be *L*-ancillary with respect to the interest parameter ψ if the following two conditions are satisfied:

(i) the conditional distribution of x given u does not depend on χ ,

(ii) the marginal distribution of the statistic u is *L*-nonformative with respect to ψ .

Condition (ii) says that $Q(\psi, \psi'; u) = \text{fn}(\psi, \psi')$.

R e m a r k 4.1. *S*-ancillarity (as defined in [1]) is a special instance of *L*-ancillarity, corresponding to the case where (i) holds and the marginal distribution of u does not depend on ψ .

R e m a r k 4.2. When condition (i) is satisfied we have

$$\hat{\chi}_\psi(x) = \hat{\chi}_\psi(u). \quad (4.1)$$

R e m a r k 4.3. Given the model function $p(x; \psi, \chi)$, the question of whether there exists an L -ancillary statistic for the interest parameter may be approached by calculating $\bar{L}(\chi \backslash \psi)$, defined by formula (1.3), and investigating whether this depends on the data x through some function u only (as will be the case if condition (i) of Definition 4.1 is satisfied, cf. also Remark 4.2). Such a statistic u is then candidate for being L -ancillary with respect to ψ .

We conclude by a discussion of L -ancillarity in the context of exponential models.

Suppose the model is a prime exponential model (as defined in [4])

$$\exp\{\theta \cdot t - k(\theta) - h(t)\} \quad (4.2)$$

and that the maximum likelihood estimator exists with probability 1 and has range equal to $T = \tau(\Theta)$, the set of mean values of the canonical statistic t . Let (t_1, t_2) be a partition of the canonical statistic t and let (τ_1, θ_2) be a corresponding mixed parametrization of the model. By $\tau_2(\theta_2 | t_1)$ we denote the conditional mean of t_2 given t_1 .

Theorem 4.1. *If θ_2 is the parameter of interest, then t_1 is L -ancillary if and only if for some function $v(\theta_2)$ we have that $\tau_2(\theta_2 | \tau_1)$, i.e., the value of $\tau_2(\theta_2 | \cdot)$ at $t_2 = \tau_2$, is of the form*

$$\tau_2(\theta_2 | \tau_1) = \tau_2 + v(\theta_2). \quad (4.3)$$

P r o o f. As is well known the conditional distribution of t_2 depends on θ_2 only, so to prove the statement we need only show that t_1 is L -noninformative about θ_2 if and only if (4.3) holds. Now, the marginal likelihood based on t_1 is of the form

$$L(\theta_1; t_1) = \exp\{\theta_1 \cdot t_1 - k(\theta)\} \mathbf{E}[\exp\{\theta_2 \cdot t_2\} | t_1] \quad (4.4)$$

and t_1 is L -noninformative about θ_2 if and only if substitution of $\tilde{\theta}_1 = \theta_1(t_1, \theta_2) (= \hat{\theta}_{1\theta_2})$ for θ_1 in (4.4) results in an expression, which equals a function of t_1 times a function of θ_2 or, equivalently, if and only if

$$\tilde{\theta}_1 \cdot t_1 - k(\tilde{\theta}_1, \theta_2) + k(\theta_2 | t_1) = f(t_1) + g(\theta_2) \quad (4.5)$$

for some functions f and g and where

$$k(\theta_2 | t_1) = \log \mathbf{E}[\exp\{\theta_2 \cdot t_2\} | t_1] \quad (4.6)$$

is the cumulant function of the conditional exponential model given t_1 . Equation (4.5) is, in turn and via differentiation with respect to θ_2 , equivalent to

$$\tau_2(\theta_2 | t_1) = \tau_2(t_1, \theta_2) + v(\theta_2). \quad (4.7)$$

Under the assumptions we have made concerning the exponential model, relation (4.7) is equivalent to (4.3).

For the following theorem and its proof we rely on terminology and results from [3].

Theorem 4.2. *Suppose the exponential model possesses a τ -parallel foliation of the form*

$$\theta_1 = -\theta_2 h(\tau_1) + k(\tau_1). \quad (4.8)$$

Then t_1 is L -ancillary with respect to θ_2 .

P r o o f. Condition (4.8) is equivalent to the existence of functions H and m such that τ_2 is of the form

$$\tau_2 = H(\tau_1) + m(\theta_2). \quad (4.9)$$

It is moreover known that $t_2 - H(t_1)$ is independent of t_1 and that the distribution of $t_2 - H(t_1)$ depends on θ_2 only. Hence

$$\tau_2(\theta_2 | t_1) = \mathbf{E}\{t_2 | t_1\} = \mathbf{E}\{t_2 - H(t_1) | t_1\} + H(t_1) = \delta(\theta_2) + H(t_1) \quad (4.10)$$

for some function δ . The theorem follows on combining (4.9) and (4.10).

R e m a r k 4.4. It follows also from the properties mentioned in the proof that in addition to t_1 being L -ancillary with respect to θ_2 we also have that $t_2 - H(t_1)$ is L -sufficient with respect to θ_2 .

E x a m p l e 4.1. For a sample of size $n > 1$ from the Gamma distribution

$$\frac{\alpha^\lambda}{\Gamma(\lambda)} x^{\lambda-1} e^{-\alpha x}$$

we find that the arithmetic mean $\bar{x}$ is L -ancillary with respect to λ .

REFERENCES

1. *Barndorff-Nielsen O. E.* Information and Exponential Families in Statistical Theory. Chichester: Wiley, 1978.
2. *Barndorff-Nielsen O. E.* Parametric statistical models and likelihood. Lect. Notes in Statist., v. 50, 1988.
3. *Barndorff-Nielsen O. E., Blæsild P.* Exponential models with affine dual foliations. — Ann. Statist., 1983, v. 11, p. 753–769.
4. *Barndorff-Nielsen O. E., Cox D. R.* Inference and Asymptotics. London: Chapman and Hall, 1994.
5. *Barndorff-Nielsen O. E., Jupp P. E.* Differential geometry, profile likelihood, L -sufficiency and composite transformation models. — Ann. Statist., 1988, v. 16, p. 1009–1043.
6. *Rémon M.* On a concept of partial sufficiency: L -sufficiency. — Internat. Statist. Rev., 1984, v. 52, p. 127–136.

© 1999 г.

MORDECKI E.*

NECESSARY CONDITIONS FOR STABLE
CONVERGENCE OF SEMIMARTINGALES

В статье доказан результат обратной теореме об устойчивой сходимости семимартингалов, полученный в [1]. Как следствие утверждается (при условии некоторого контроля за скачками), что последовательность мартингалов X^n устойчиво сходится к непрерывному мартингалу X с условно независимыми приращениями тогда и только тогда, когда квадратичные отклонения X^n сходятся по вероятности к квадратичному отклонению X для каждого $t \in \mathbf{R}^+$.

Ключевые слова и фразы: семимартингал, устойчивая сходимость, независимые приращения.

1. Introduction and results. The subject of the present paper is the characterization of the behaviour of a sequence of semimartingales when they converge stably to a continuous semimartingale with conditionally independent increments. In view of the results due to [3], and more recently to [1], it remains only to prove some inverse theorems, and this is done, based on the extension of existing results for weak convergence, and a convenient characterization of stable convergence. The results presented here were announced in [4].

For general concepts about semimartingales and their characteristics we refer to ([2, Chapters I, II]).

Let us now define the sequence of semimartingales and the limit.

Let h denote a truncation function, and all the characteristics will be considered with respect to h .

(S) Consider for each $n \in \mathbf{N}$ a d -dimensional semimartingale X^n defined on a stochastic basis $\mathcal{B}^n = (\Omega, \mathcal{F}, \mathbb{F}^n, \mathbf{P})$, with $\mathbb{F}^n = (\mathcal{F}_t^n: t \geq 0)$. Denote by $(B^n, \tilde{C}^n, \nu^n)$ the triplet of modified characteristics of X^n under $\mathbf{P}$, and by $[X^n, X^n]$ the $\mathbf{R}^d \otimes \mathbf{R}^d$ -valued processes whose components are the (co-)quadratic variation $[X^{n,j}, X^{n,k}]$. Set for simplicity $X_0^n = 0$.

As for the limit consider a transition probability $\mathbf{Q}(\omega, d\gamma)$ from $(\Omega, \mathcal{F})$ into the Skorohod space

$$(\mathbf{D}(\mathbf{R}^d), \mathcal{D}(\mathbf{R}^d), (\mathcal{D}_t: t \geq 0)),$$

* Universidad de la República, Montevideo, Uruguay.