Steffen L. Lauritzen

PROJECTIVE STATISTICAL FIELDS

Preprint 1976 No. 6

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

March 1976

Abstract.

A projective statistical field is defined as a projective system of measurable spaces and Markov kernels, equipped with a system of statistical models consistent with the projections.

A statistical population can then be defined as the projective limit of such a system.

An appropriate definition of a sufficient reduction of a projective statistical field is given and shown to coincide with Bahadurs notion of a sufficient and transitive sequence of statistics in the sequential case. A canonical projective statistical field is defined as a field where the parameter space is the projective limit of a sufficient system.

Construction of canonical fields and its relation to foundational questions in statistical inference is touched upon.

Finally examples are given illustrating the impacts of the theory on questions as conditional inference and extension of statistical models.

Key words: Projective systems, repetitive structures, statistical inference, sufficiency, transitivity.

AMS 1970 subject classifications.

Primary 62B05

Secondary 62M99

1. Introduction and Summary.

The present paper is a further development of ideas from Martin-Löf (1970), (1974) and Lauritzen (1972), (1974a), (1974b), (1975).

It can be seen as an attempt to give a mathematical framework in which it is possible to discuss the relation between a statistical model and the reference population it is supposed to describe.

Usually, a statistical field is thought of as a measurable space equipped with a parametrised family of probability measures

$$\{\Omega, (P_\theta \quad \in \Theta)\} \cdot$$

The observation can then be thought of as a sample from a hypothetically infinite population of independent, identically distributed random variables following the probability law $\theta \in \Theta$. This is the population one uses to interpret testprobabilities in the Neyman-Pearson theory. However, in many situations, where we would do statistical analyses this population becomes very hypothetical and is not at all the population one really would like to describe by the statistical model.

Consider for example an observation from a time series

$$x_t, \ t = 1,\dots,T$$

where t really denotes time. It seems awkward to think of this as a sample from a population of independent random vectors with identical distributions. One would rather think of it as a sample from the whole series itself

$$\{x_t, \ t = 0, \pm 1, \pm 2,\dots\} \ .$$

This difficulty is sometimes overcome by assumptions like stationarity and ergodicity, which gives a connection between two populations in the sense that they have same averages. It remains nevertheless that the two populations are different and the probabilistic model gives a description of the wrong one, seen from a statistical point of view.

Now consider a situation with many nuisance parameters

$$X_i \sim P_{\theta,\gamma_i} , \quad \theta \in \Theta , \quad \gamma_i \in \Gamma_i , \quad i = 1,2,\ldots,N .$$

Also here one would like to think of the observation as a sample from a hypothetically infinite population, but of the form

$$X_i \sim P_{\theta,\gamma_i} , \quad \theta \in \Theta , \quad \gamma_i \in \Gamma_i , \quad i = 1,2,\ldots \quad .$$

The usual probability model refers to a population of the form

$$X_{ij} \sim P_{\theta,\gamma_i} , \quad \theta \in \Theta \quad , \gamma_i \in \Gamma_i , \quad \begin{matrix} i = 1,2,\ldots,N \\ j = 1,2,\ldots \end{matrix} \quad .$$

This is again "the wrong" population seen from a statistical point of view. A resulting difficulty is for example the non-consistency of maximum likelihood estimators when the asymptotics are referred to the former population see e.g. Neyman and Scott (1948). This is then usually taken care of by conditioning on statistics that are ancillary to θ and sufficient for γ_i.

This procedure is still under discussion but can be thought of as an adaption of the model to the population of interest.

Many other examples could be given, where there is a "conflict" between the hypothetical population and the relevant one. Some of these are:

1) Sampling from <u>finite</u> populations.

2) Two-way analysis of variance. (Here the two-way scheme should be considered a part of a doubly infinite two-way scheme, since mostly, the combination of blocks and treatments cannot be repeated.

3) Pairwise comparisons.

4) All types of inference in stochastic processes.

To be able to discuss such things in a reasonably rigorous way, one needs a mathematical definition of a population and a sample.

To give such a definition we consider simultaneously a whole family of experiments i.e. a <u>family</u> of measurable spaces

$$(\Omega_i, i \in I) .$$

These spaces have to be related in the sense that some experiments are subexperiments of others, i.e. there is a <u>partial ordering</u> of the spaces

$$\Omega_i \text{ "}\subset\text{" } \Omega_j .$$

When we want to give statistical models for such families the distributions of experiments that are related, also have to be related. We formalise this by specifying the <u>conditional distribution</u> on the space Ω_i given the value on Ω_j, if Ω_i is a subexperiment of Ω_j. Often this distribution is degenerate, but it turns out to be convenient not always to assume so.

These have to satisfy certain consistency conditions given the fact that when we have

$$\Omega_i \text{ "}\subset\text{" } \Omega_j \text{ "}\subset\text{" } \Omega_k$$

we also have

$$\Omega_i \text{ "}\subset\text{" } \Omega_k .$$

Systems of this kind is what is formalised as projective systems of measurable spaces and Markov kernels. Sections 2 and 3 contain the definition and elementary properties.

The "population" now corresponds to the smallest experiment containing all other experiments in the system. This is exactly what is called the projective limit and the theory is described in section 4.

We then construct statistical models for these experiments by specifying "consistent" systems of parametrised probability measures. We use the term "projective statistical fields" for such models and give the definition and examples on section 5.

We would also like to consider sufficient reductions of experiments. But then we have to take care that the reduced experiments again can be thought of as an experiment from a system of the type defined. Therefore we need a definition of sufficiency that is a bit more restrictive that usual. This definition is given in section 6 and is shown to coincide with the notion of a sufficient and transitive sequence in the sequential case.

In section 7 we define a canonical projective statistical field which is to be thought of as a statistical model with an especially simple relation between population, parameter and statistics.

In section 8 it is shown how such canonical fields can be constructed, and how one can compare a given projective statistical field to the corresponding canonical statistical field , thus getting an insight into the relation between population, parameters and model in the given field. This is illustrated by examples.

Throughout the paper we consider the situation of hypergeometric sampling or sampling without replacement from infinite populations to illustrate the theory.

2. Measurable Spaces and Markov Kernels.

In the following, a measurable space Ω is a pair $\Omega = (S,\mathcal{S})$, where S is a set and $\mathcal{S}$ a σ-algebra of subsets of S. We shall sometimes write $\Omega = S$ when it is clear which σ-algebra of subsets one would think of. This is to avoid a complicated notation.

If $\Omega = (S,\mathcal{S})$ and $\Omega' = (S',\mathcal{S}')$ are measurable spaces, a Markov kernel P from Ω to Ω' is a mapping

$$P: \quad S \times \mathcal{S}' \rightarrow [0,1]$$

satisfying

i) $\quad \forall\, s \in S$, $P(s,\cdot)$ is a probability measure on $\mathcal{S}'$

ii) $\quad \forall\, A \in \mathcal{S}'$, $P(\cdot,A)$ is an $\mathcal{S}$-measurable function.

For convenience we shall write

$$P: \Omega \rightarrow \Omega' \quad \text{or} \quad \Omega \overset{P}{\rightarrow} \Omega'.$$

Let us introduce the indicatorfunction of a set A:

$$1_A(x) = \begin{cases} 1 & \text{for } x \in A \\ 0 & \text{otherwise} \end{cases}$$

Via the indicatorfunction, any ordinary measurable mapping from one probability space to another can be thought of as a Markov kernel. Suppose namely that $f: \Omega \rightarrow \Omega'$ is such a mapping. Define the Markov kernel

$$P_f(s,A) = 1_A(f(s)).$$

Such a Markov kernel is said to be induced by the mapping f. When misunderstanding is impossible we shall not distinguish between a mapping and the Markov kernel induced by it.

Example 2.1

Let $S = \{0,1,\ldots,n\}$ and $\mathcal{S}$ consist of all subsets of S. Let $S' = \{0,1,\ldots,N\}$ and $\mathcal{S}'$ be all subsets of S'. Suppose $n < N$. One can define a Markov kernel

$$(S',\mathcal{S}') = \Omega' \overset{P}{\to} \Omega = (S,\mathcal{S}),$$

by

$$P(s',A) = \sum_{s\in A} \frac{\binom{s'}{s}\binom{N-s'}{n-s}}{\binom{N}{n}} .$$

This will in the following be referred to as the hypergeometric kernel.

Let us now suppose that we have two Markov kernels P and Q as follows

$$\Omega \overset{P}{\to} \Omega' \overset{Q}{\to} \Omega'' .$$

We can then compose P and Q and construct the Markov kernel QP as

$$QP(s,A'') = \int_{\Omega'} Q(s',A'')\, P(s,ds').$$

To each measurable space Ω there is an identity on Ω, I_Ω, defined as

$$I_\Omega(s,A) = 1_A(s) .$$

If $P: \Omega \to \Omega'$ is a Markov kernel we have

$$PI_\Omega = I_{\Omega'}P = P .$$

We shall say, that two measurable spaces Ω and Ω' are isomorphic and write

$$\Omega \sim \Omega'$$

if there are Markov kernels M and M'

$$\Omega \overset{M}{\to} \Omega'$$

$$\Omega' \overset{M'}{\to} \Omega$$

Manuskriptark. A4 = A5

such that

$$MM' = I_{\Omega} \quad \text{and} \quad MM' = I_{\Omega'} \quad .$$

It is easy to see that both M and M' must be induced by bimeasurable mappings that are one-to-one and onto.

Example 2.2

Let $S = \{0,1,\dots,i\}$, $S' = \{0,1,\dots,j\}$, $S'' = \{0,1,\dots,k\}$ where $i < j < k$ and let the σ-algebras consist of all subsets. Define $P: \Omega' \to \Omega$ and Q from Ω'' to Ω' as the hypergeometric kernels. We then have

$$PQ(s'',A) = \sum_{s'=0}^{j} \sum_{s\in A} \frac{\binom{s'}{s}\binom{j-s'}{i-s}}{\binom{j}{i}} \frac{\binom{s''}{s'}\binom{k-s''}{j-s'}}{\binom{k}{j}}$$

$$= \sum_{s\in A} \frac{\binom{s''}{s}\binom{k-s''}{i-s}}{\binom{k}{i}} \quad , \qquad (2.1)$$

i.e. the corresponding hypergeometric kernel from Ω'' to Ω.

3. Projective Systems of Measurable Spaces.

Let $(I,<)$ be a partially ordered set, directed to the right, i.e.

$$\forall\, i,j \in I \quad \exists\, k \in I : \; i,j < k \; .$$

Let $(\Omega_i, i \in I)$ be a family of measurable spaces indexed by I. Suppose further that there is given a family of Markov kernels $(P_{ij}, i < j)$

$$\Omega_i \xleftarrow{P_{ij}} \Omega_j \; . \tag{3.1}$$

Such a system is called a projective system of measurable spaces provided

$$P_{ij}\, P_{jk} = P_{ik} \quad \text{whenever} \quad i < j < k \; . \tag{3.2}$$

Examples of such systems are plenty.

Example 3.1

If I contains only one element, a projective system corresponding to I is just a measurable space Ω.

Example 3.2

Another degenerate projective system appears when $\Omega_i = \Omega$ for all $i \in I$ and corresponding $P_{ij} = I_\Omega$.

Example 3.3

Let $I = \mathbb{N}$ with the usual ordering and let Ω_i be the space defined in example 2.2. For $i < j$ we define P_{ij} to be the relevant hypergeometric kernels.

(2.1) shows that this is in fact a projective system.

Example 3.4

Let I consist of all finite subsets of a set T, ordered by inclusion. Define for $D \in I$

$$\Omega_D = \prod_{t \in D} \Omega_t^* ,$$

where $(\Omega_t^*, t \in T)$ is a family of measurable spaces. For $D_1 \subset D_2$ define $P_{D_1 D_2}$ as the Markov kernel induced by usual coordinate projection. Clearly we have

$$P_{D_1 D_2} P_{D_2 D_3} = P_{D_1 D_3} \quad \text{for} \quad D_1 \subset D_2 \subset D_3$$

and the system considered is thus a projective system.

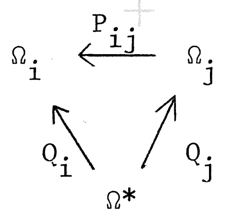
The formulae (3.1) and (3.2) together imply that one can think of a projective system as a "reverse Markov chain". This gives the connection to the work in Lauritzen (1974a).

4. Projective Limits.

We consider a projective system

$$\{(\Omega_i, i \in I), (P_{ij}, i < j)\} \tag{4.1}$$

of measurable spaces. Suppose Ω^* is a measurable space and $(Q_i, i \in I)$ a system of Markov kernels, so that the diagram below is commutative



i.e. satisfying for all $i < j$

$$Q_i = P_{ij}\, Q_j \; . \tag{4.2}$$

A measurable space Ω is said to be the projective limit of the system (4.1), and we write

$$\Omega \sim \varprojlim_{i \in I} \Omega_i$$

if

i) for all $i \in I$ there are Markov kernels $P_i : \Omega \to \Omega_i$ so that
$$P_i = P_{ij}\, P_j \tag{4.3}$$

ii) for all systems $(\Omega^*, (Q_i, i \in I))$ satisfying (4.2) there is one and only one Markov kernel $R: \Omega^* \to \Omega$ such that for all $i \in I$

$$P_i\, R = Q_i \tag{4.4}$$

The projective limit, if it exists, is well defined up to isomorphism. Assume namely that both Ω and Ω' are projective limits of (4.1). Then (4.3) and (4.4) combined imply existence of Markov kernels M and M' so that

$$P'_i M' = P_i \quad \text{and} \quad P_i M = P'_i \ .$$

But then

$$P_i MM' = P_i \quad \text{and} \quad P'_i M'M = P'_i \ .$$

The uniqueness of the mapping R in (4.4) then implies

$$MM' = I_\Omega \quad \text{and} \quad M'M = I_\Omega \ ,$$

and thus that Ω and Ω' are isomorphic.

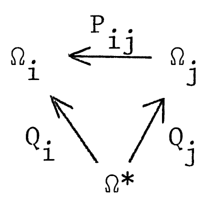
Projective limits do not exist in general and usually their existence is a non-trivial matter. The Kolmogorov consistency theorem for stochastic processes is a special case of an existence theorem for projective limits, see ex.4.3. In the case where the spaces Ω_i are compact topological spaces and the Markov kernels are continuous, the existence has been proved by Scheffer (1971). In the theoretical considerations of the present work we shall plainly assume existence of projective limits.

Example 4.1

In the two degenerate systems of examples 3.1 and 3.2, the projective limit is just Ω itself.

Example 4.2

Consider the projective system defined in example 3.3 (hypergeometric kernels). A consistent system



is equivalent to a family $(Q_i(s,\cdot), s \in S^*)$ of distributions reproducible under hypergeometric sampling, cf. Hald (1960), since the equation

$$Q_i = P_{ij}\, Q_j$$

is equivalent to

$$Q_i(s,A) = \sum_{y\in A} \sum_{x=0}^{j} \frac{\binom{x}{y}\binom{j-x}{i-y}}{\binom{j}{i}}\, Q_j(s,\{x\}).$$

It can be proved that this implies that $Q_i(s,\cdot)$ are mixed binomial distributions with the mixing measure being independent of $i \in \mathbb{N}$

$$Q_i(s,\{x\}) = \int_0^1 \binom{i}{x}\, \theta^x (1-\theta)^{i-x}\, \mu_s(d\theta) \ . \qquad (4.5)$$

Obviously, the measure μ_s is uniquely defined and therefore [0,1] with the usual Borel σ-algebra is the projective limit of this system, since (4.5) can be written

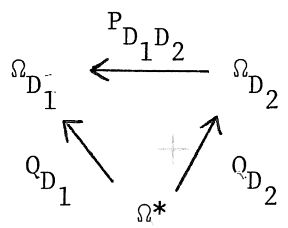
$$Q_i = P_i\, M \ .$$

when P_i is defined as

$$P_i(\theta,\{x\}) = \binom{i}{x}\, \theta^x (1-\theta)^{1-x} \ .$$

<u>Example 4.3</u>

Consider the projective system defined in example 3.4. A system of the form



where the diagram is commutative is exactly a consistent family of finite-dimensional distributions. An obvious candidate for the limit is the infinite product space

$$\prod_{t\in T} \Omega_t^* \overset{?}{\sim} \varprojlim_{D\in I} \Omega_D \ .$$

To say that this is the limit is equivalent to saying that to any consistent family $\{Q_D(s,\cdot), D \subset T \text{ finite}\}$ of finite-dimensional distributions there is a uniquely defined probability measure μ_s on the procuct $\prod_{t \in T} \Omega^*_t$ such that

$$Q_D(s,\cdot) = P_D \, \mu_s \quad .$$

This is exactly Kolmogorovs consistency theorem for stochastic processes which is known to be <u>false</u> without further assumptions on T or the family $(\Omega^*_t, t \in T)$, cf. Sparre Andersen and Jessen (1948).

We shall now informally state an important theorem about construction of projective limits. It says that, roughly, the projective limit can be identified with the limits of the measures

$$P_{ij}(s_j,\cdot), \quad j \to \infty.$$

The theorem is a direct generalisation of the theorem on p.279 in Martin-Löf (1974). We shall name it "Boltzmann's law" as it is analogous to a theorem used to prove Boltzmann's law in statistical mechanics, Khinchin (1949). See also Martin-Löf (1970).

Let there be given a projective system of measurable spaces

$$\{(\Omega_i, i \in I), (P_{ij}, i < j)\} \quad .$$

Define the following system of measures for $i,j \in I$, $s_j \in S_j$, $A_i \in \mathcal{S}_i$

$$\mu_i^{(j,s_j)}(A_i) = \begin{cases} P_{ij}(s_j, A_i) & \text{if} \quad i < j \\ 0 \text{ otherwise} \end{cases} \qquad (4.6)$$

Suppose that we have a suitable topology given on the measures on Ω_i.

A system $(\mu_i, i \in I)$ is said to be an <u>accumulation</u> point of the system (4.6) if there is a <u>cofinal</u> subset $J \subset I$, i.e. a subset satisfying

$$\forall\, i \in I\ \exists j \in J :\ i < j\ ,$$

and points s_j, $j \in J$ so that for all $i \in I$

$$\lim_{j \in J} \mu_i^{(j,s_j)} = \mu_i$$

in the topology mentioned before.

Let B consist of all such accumulation points and let $\mathcal{B}$ be the Borel sets of B induced by the pointwise topology.

Let $\Omega' = (B,\mathcal{B})$ and let the Markov kernels $P_i' : \Omega' \to \Omega_i$ be given as

$$P_i'((\mu_k, k \in I), A_i) = \mu_i(A_i).$$

Under certain regularity conditions one can now prove the following

Theorem 4.1 : (Boltzmann's law) If Ω^* is a measurable space and Q_i a system of Markov kernels from Ω^* to Ω_i such that

$$P_{ij}\, Q_j = Q_i \qquad \forall\, i,j \in I:\ i < j$$

then there is at least one Markov kernel M^* from Ω^* to Ω' so that

$$Q_i = P_i'\, M^* . \qquad \forall\, i \in I \tag{4.7}$$

We shall not here give a proof of the theorem nor state the exact regularity conditions needed. There is no reason to believe the theorem to be true in full generality.

If one assumes the spaces Ω_i to be locally compact topological spaces, the Markov kernels to be continuous with limits at infinity and the index set I to have a cofinal sequence, the theorem can be proved by a direct adaption of the proof of the theorem on p. 279 in Martin-Löf (1974).

The significance of the theorem is that apart from the uniqueness of the Markov kernel in (4.7) the projective limit can be identified with Ω'. A way to find the projective limit will offen be to find Ω' and then prove uniqueness of the integral representation (4.7). In many cases indeed it is so that Ω' in fact is the projective limit itself.

Example 4.4

Consider the projective system of examples 3.3 and 4.2, i.e. the hypergeometric kernels. The measures $\mu_i^{(j,s_j)}$ are here given as

$$\mu_i^{(j,s_j)}(\{x\}) = \begin{cases} \dfrac{\binom{s_j}{x}\binom{j-s_j}{i-x}}{\binom{j}{i}} \\ 0 \quad \text{if} \quad i > j \end{cases}$$

It is well known that if $(j,s_j) \to \infty$ in such a way that

$$s_j / j \to \theta ,$$

then

$$\mu_i^{(j,s_j)}(\{x\}) \to \binom{i}{x} \theta^x (1-\theta)^{1-x} \quad .$$

Comparing this result with the considerations in ex. 4.2, we see that

$$\Omega' \sim \varprojlim_{i \in I} \Omega_i \quad .$$

5. Projective Statistical Fields

In this section we shall consider statistical models for projective systems of measurable spaces or, in short, projective statistical fields.

A proejctive statistical field consists of

i) a projective system of measurable spaces

$$\{(\Omega_i, i \in I), (P_{ij}, i < j)\};$$

ii) a parameter space $\Theta = (T, \mathcal{T})$ that for convenience is considered equipped with a measurable structure.

iii) a parametrisation, which is a system of Markov kernels $(\mu_i, i \in I)$

$$\mu_i : \quad \Theta \to \Omega_i$$

that is consistent with the projections

$$\forall i < j : \quad \mu_i = P_{ij} \mu_j \; .$$

To interpret the definition one should think of an increasing system of experiments

$$\Omega_i \subset \Omega_j$$

each of them giving rise to random variables

$$X_i \in \Omega_i$$

with distributions given by $\mu_i(\theta)$ and depending on the unknown parameter θ.

To say that

$$\mu_i = P_{ij} \mu_j$$

means that the measure

$$P_{ij}(s,\cdot) \quad \text{on} \quad \Omega_i$$

is the <u>conditional distribution of</u> X_i <u>given</u> X_j, since one could write it as

$$\mu_i(\theta,A) = \int P_{ij}(s,A)\ \mu_j(\theta,ds) \ .$$

Frequently one has $I = \mathbb{N}$ and $X_i = (Y_1,\dots,Y_i)$.

The P_{ij}'s would then just be coordinate projections mapping the distribution μ_j of X_j into the <u>marginal</u> distribution of:

$$X_i = (Y_1,\dots,Y_i) \ .$$

The reason for considering more complicated structures is that one would like to reduce data by sufficiency, and as we shall see in the next section, this generates projective statistical fields where P_{ij} are truly random.

If the projective system has a limit

$$\Omega = \varprojlim_{i\in I} \Omega_i$$

we shall call this the <u>reference population</u>.

The fundamental property of the projective limit then implies existence of a unique Markov kernel μ

$$\mu : \Theta \to \Omega$$

such that

$$\mu_i = P_i\ \mu \ .$$

Thus the parametrisation can be identified with the single Markov kernel μ via the projective limit.

In the examples we consider, the projective limit does exist and probably we are not at all interested in dealing with situations where this is not the case.

Example 5.1

If we consider the degenerate system with I having only one element, a projective statistical field is just a usual statistical field.

Example 5.2

Let $I = \mathbb{N}$ and $\Omega_i = \{0,1\}^i$, and let P_{ij} be induced by the coordinate projections.

Let $\Theta = [0,1]$ and define

$$\mu_i(\theta,\{(x_1,\ldots,x_i)\}) = \theta^{\sum_{j=1}^{i} x_j} (1-\theta)^{\sum_{j=1}^{i}(1-x_j)}$$

The Markov kernel

$$\mu: \Theta \to \{0,1\}^{\mathbb{N}} \sim \underset{\overleftarrow{i\in I}}{\lim} \Omega_i$$

assigns to a $\theta \in \Theta$ the probability measure corresponding to an infinite sequence of independent Bernoulli random variables with probability θ of succes.

Example 5.3

Let $I = \{(i,j), i,j \in \mathbb{N}\}$ with the ordering

$$(i_1,j_1) < (i_2,j_2) \Leftrightarrow i_1 < i_2 \wedge j_1 < j_2 .$$

Let $\Omega_{(i,j)} = \{0,1\}^{i\cdot j}$ and let $P_{(i_1,j_1)(i_2,j_2)}$ be induced by coordinate projections.

Let $\Theta = \mathbb{R}^{\mathbb{N}} \times \mathbb{R}^{\mathbb{N}}$ and define

$$\mu_{(i,j)}((\alpha_k,\beta_\ell,k,\ell \in \mathbb{N}),(x_{mn},m \leq i,n \leq j))$$

$$= \prod_{m=1}^{i} \prod_{n=1}^{j} \frac{e^{(\alpha_m+\beta_n)x_{mn}}}{1+e^{\alpha_m+\beta_n}} .$$

This example corresponds to Rasch's models for the item analysis, cf. Rasch (1960).

Example 5.4

Let I consist of all closed subintervals of the positive real axis

$$I = \{[a,b],\ 0 \leq a,b \leq +\infty\},$$

partially ordered by inclusion. Let $\Omega_{[a,b]}$ be $\mathbb{R}^{[a,b]}$ with the cylinder σ-algebra and let $P_{[a,b],[c,d]}$ be induced by coordinate projection.

Let $\Theta = [0,\infty[$ and let $\mu_{[a,b]}(\theta)$ be the measure induced on $\Omega_{[a,b]}$ by the gaussian process ξ with mean

$$E\ \xi(t) = \theta t\ , \quad t \geq 0$$

and covariance

$$E\ \xi(s)\ \xi(t) = \min\ \{s,t\}\ ,$$

i.e. the Wiener proces with unknown drift.

6. Sufficient Reductions of Projective Statistical Fields.

Consider a projective statistical field

$$F = \{(\Omega_i, i \in I), (P_{ij}, i < j), \Theta, (\mu_i, i \in I)\},$$

a family of measurable spaces $(\Omega_i', i \in I)$ and Markov kernels $(R_i, i \in I)$

$$R_i : \quad \Omega_i \to \Omega_i' .$$

We shall say that $(R_i, i \in I)$ is a <u>sufficient system of reductions</u> if there is an "inverse" family of Markov kernels

$$R_i^* : \quad \Omega_i' \to \Omega_i$$

so that the following three conditions are fulfilled

i) $\quad R_i R_i^* = I_{\Omega_i'} \qquad \forall i \in I$

ii) $\quad R_i^* R_i \mu_i = \mu_i \qquad \forall i \in I$

iii) $\quad R_i^* R_i P_{ij} R_j^* = P_{ij} R_j^* \qquad \forall i,j \in I: i < j$

<u>Comments</u>: Condition i) ensures that R_i is induced by a mapping; i) and ii) combined says that R_i^* is the conditional distribution in $\mu_i(\theta)$ given R_i. Since R_i^* is chosen independently of $\theta \in \Theta$, i) and ii) is just the "usual" definition of sufficiency.

The last condition combined with i) says analogously, that R_i^* also is the conditional distribution on Ω_i given <u>both</u> R_i and R_j and thus that the random elements

$$X_i \in \Omega_i \quad \text{and} \quad R_j(X_j) \in \Omega_j'$$

are <u>conditionally independent</u> given $R_i(X_i)$. This is obvious because $P_{ij} R_j^*$ is the conditional distribution on Ω_i given R_j and $P_{ij} R_j^*$ plays the same role in

iii) as does μ_i in ii). The reasoning is eased by the diagram below

$$\begin{array}{ccc} \Omega_i & \xleftarrow{P_{ij}} & \Omega_j \\ R_i \downarrow \uparrow R_i^* & & R_j \downarrow \uparrow R_j^* \\ \Omega_i' & & \Omega_j' \end{array} .$$

<u>Example 6.1</u>

Consider the projective statistical field defined in example 5.2 and let

$$\Omega_i' = \{0,1,\ldots,i\} .$$

The system of reductions $(R_i, i \in I)$

$$R_i : \ \Omega_i \rightarrow \Omega_i'$$

induced by the mappings

$$R_i(x_1,\ldots,x_i) = x_1 + \ldots + x_i$$

is sufficient because it is in the usual sense and because iii) is satisfied. We define namely R_i^* as

$$R_i^*(t,\{(x_1,\ldots,x_i)\}) = \frac{1}{\binom{i}{t}} \cdot 1_{\{x_1+\ldots+x_i=t\}}$$

and have

$$P_{ij}R_j^*(t,\{(x_1,\ldots,x_i)\}) = \sum_{\{(x_{i+1},\ldots,x_j):x_1+..+x_j=t\}} \frac{1}{\binom{j}{t}} = \frac{\binom{j-i}{t-s}}{\binom{j}{t}} \cdot 1_{\{x_1+..+x_i=s\}} .$$

Then

$$R_i \ P_{ij} \ R_j^* = \frac{\binom{i}{s}\binom{j-i}{t-s}}{\binom{j}{t}}$$

and obviously

$$R_i^* \, R_i \, P_{ij} \, R_j^*(t,\{(x_1,\dots,x_i)\}) = 1_{\{x_1+\dots+x_i=s\}} \frac{\binom{j-i}{t-s}}{\binom{j}{t}}$$

$$= P_{ij} \, R_j^*(t,\{(x_1,\dots,x_i)\}) \, .$$

A consequence of the above comments is, that in the case $I = \mathbb{N}$, the definition here given of a sufficient system (apart form technicalities) coincides with that of a <u>sufficient and transitive</u> system as defined by Bahadur (1954). The present definition is therefore a direct generalisation of Bahadur's notion.

We can, whenever we have a sufficient system of reductions define a system of Markov kernels

$$\Omega_j' \xleftarrow{Q_{ij}} \Omega_i' \ , \quad i < j$$

by

$$Q_{ij} = R_i \, P_{ij} \, R_j^* \, .$$

The last condition in the definition then implies for $i < j < k$

$$Q_{ij} \, Q_{jk} = R_i \, P_{ij} \, R_j^* \, R_j \, P_{jk} \, R_k^* = R_i \, P_{ij} \, P_{jk} \, R_k^* = R_i \, P_{ik} \, R_k^* = Q_{ik} \ ,$$

and thus that $\{(\Omega_i, i \in I),(Q_{ij}, i < j)\}$ is a projective system and that the diagram

$$\begin{array}{ccc} \Omega_i & \xleftarrow{P_{ij}} & \Omega_j \\ R_i^* \uparrow & & \uparrow R_j^* \\ \Omega_i' & \xleftarrow{Q_{ij}} & \Omega_j' \end{array}$$

is commutative.

If we define ν_i: $\Theta \rightarrow \Omega_i'$ by

$$\nu_i = R_i \mu_i ,$$

we have that

$$F' = \{(\Omega_i', i \in I), (Q_{ij}, i < j), \Theta, (\nu_i, i \in I)\}$$

is a projective statistical field.

Here Q_{ij} denotes the conditional distribution of the sufficient statistic R_i in the "small" experiment given the statistic R_j in the "large" experiment and will in most cases be truly random.

Introducing the notation $R = (R_i, i \in I)$ we can write

$$F \overset{R}{\rightarrow} F' .$$

We shall now state an important but simple result about composition of sufficient transformations.

<u>Theorem 6.1</u>: <u>If we have projective statistical fields</u>

$$F \overset{R}{\rightarrow} F' \overset{T}{\rightarrow} F''$$

<u>where</u> R <u>and</u> T <u>are sufficient then the reduction</u>

$$F \overset{TR}{\longrightarrow} F''$$

<u>defined as</u>

$$TR_i = T_i R_i , \quad i \in I$$

<u>is also sufficient</u>.

<u>Proof</u>: We assume that we have

A) i) $R_i R_i^* = I_{\Omega_i'}$

ii) $R_i^* R_i \mu_i = \mu_i$

iii) $R_i^* R_i P_{ij} R_j^* = P_{ij} R_j$

and

B) i) $T_i \, T_i^* = I_{\Omega_i''}$

ii) $T_i^* \, T_i \, R_i \, \mu_i = R_i \, \mu_i$

iii) $T_i^* \, T_i \, R_i \, P_{ij} \, R_j^* \, T_j^* = R_i \, P_{ij} \, R_j^* \, T_j^*$

and want to prove the existence of Markov kernels TR_i^*, $i \in I$ so that

C) i) $TR_i \, TR_i^* = I_{\Omega_i''}$

ii) $TR_i^* \, TR_i \, \mu_i = \mu_i$

iii) $TR_i^* \, TR_i \, P_{ij} \, TR_j^* = P_{ij} \, TR_j^*$

The proof is then simple algebraic manipulations. Define

$$TR_i^* = R_i^* \, T_i^* \,.$$

To prove C i) we have

$$T_i \, R_i \, R_i^* \, T_i^* \overset{A\,i)}{=} T_i \, T_i^* \overset{B\,i)}{=} I_{\Omega_i''} \,.$$

And C ii) follows as

$$R_i^* \, T_i^* \, T_i \, R_i \, \mu_i \overset{B\,ii)}{=} R_i^* \, R_i \, \mu_i \overset{A\,ii)}{=} \mu_i \,.$$

Finally, to establish C iii) we get

$$R_i^* \, T_i^* \, T_i \, R_i \, P_{ij} \, R_j^* \, T_j^* \overset{B\,iii)}{=} R_i^* \, R_i \, P_{ij} \, R_j^* \, T_j^* \overset{A\,iii)}{=} P_{ij} \, R_j^* \, T_j^* = P_{ij} \, TR_j^* \,.$$

Remark. The converse, i.e. that C implies A and B is not true in general (It is true that C i) and ii) imply A and B i) and ii)).

A system $(R_i, i \in I)$ is called minimal sufficient if to any sufficient system $(S_i, i \in I)$ there is a family of Markov kernels $(T_i, i \in I)$ such that

$$R_i = T_i \, S_i \quad \forall i \in I \,.$$

It is obvious, that if for each i, R_i is minimal sufficient in the usual sense, and the system $(R_i, i \in I)$ is sufficient in the sense defined here, then $(R_i, i \in I)$ is a minimal sufficient system of reductions.

Example 6.2

Consider the projective statistical field in example 5.4. Let $R_{[a,b]}$ be induced by the mapping

$$R_{[a,b]}(x(t), t \in [a,b]) = (x(a), x(b)) ,$$

$$R_{[a,b]} : \mathbb{R}^{[a,b]} \to \mathbb{R}^2 .$$

The conditional distribution of $\{\xi(t), t \in [a,b]\}$ given $(\xi(a), \xi(b))$ is the so-called Brownian bridge.

The system $\{R_{[a,b]}, a \leq b\}$ is a minimal sufficient system because of the Markov property of the process ξ.

7. Canonical Projective Statistical Fields.

This section gives a definition of a canonical projective statistical field which is to be seen as a generalisation of the notion of an extreme family of Markov chains as defined in Lauritzen (1974a).

In certain projective statistical fields there is a special coherence between the parameter space and the sufficient reductions. Such fields shall be called canonical and a precise definition is following.

Let

$$F = \{(\Omega_i, i \in I), (P_{ij}, i < j), \Theta, (\mu_i, i \in I)\}$$

be a projective statistical field. F is said to be <u>canonical</u> if

i) $\quad \Omega \sim \varprojlim_{i \in I} \Omega_i$ exists

ii) $\quad$ there is a sufficient system of reductions $(R_i, i \in I)$

$$R_i: \; \Omega_i \to \Omega_i'$$

such that the parameter space is the projective limit of the reduced system

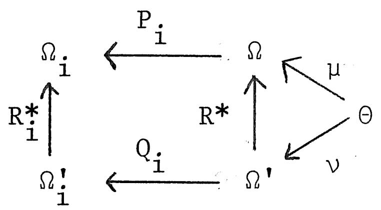
$$\Theta \sim \varprojlim_{i \in I} \Omega_i'$$

If a projective statistical field is canonical, we have the following diagram

$$\begin{array}{ccccc} \Omega_i & \xleftarrow{\;P_{ij}\;} & \Omega_j & \xleftarrow{\;P_i\;} & \Omega \\ R_i \downarrow \uparrow R_i^* & & R_i \downarrow \uparrow R_j^* & & \uparrow \mu \\ \Omega_i' & \xleftarrow{\;Q_{ij}\;} & \Omega_j' & \xleftarrow{\;Q\;} & \Theta \end{array} .$$

A verbal interpretation of the definition is that a projective statistical field is canonical exactly when the parameter is the limit of the sufficient statistic.

Another interpretation can be given from the following diagram for a projective statistical field having a projective limit and sufficient reductions



As mentioned earlier, Ω, the projective limit of the system can be thought of as the population or the maximal possible observation. Because of the existence of $R^*: \Omega' \to \Omega$, Ω' can be considered statistically equivalent to Ω. Condition ii) can be stated as the existence of a Markov kernel M so that

$$M\,\nu = I_\Theta \quad \text{and} \quad \nu\, M = I_{\Omega'}$$

$$M: \Omega' \to \Theta,$$

or, in other words that ν is induced by a bimeasurable mapping that is one-to-one and onto. This connection between parameter and population is a fundamental property of a statistical model.

<u>Example 7.1</u>

In examples 4.2, 5.2 and 6.1 we have considered the projective systems given by

$$\Omega_i = \{0,1\}^i \;, \quad i = 1,2,\ldots$$

$$\Omega_i' = \{0,1,\ldots,i\} \quad i = 1,2,\ldots$$

$$P_{ij}: \Omega_j \to \Omega_i \qquad \text{"coordinate projections"}$$

$$Q_{ij}: \Omega_j' \to \Omega_i' \qquad \text{hypergeometric kernels,}$$

and the system of reductions $R_i: \Omega_i \to \Omega_i'$ given by

$$R_i(x_1,\ldots,x_i) = x_1 + \ldots + x_i \;.$$

The projective limits have been shown to be

$$\varprojlim_{i \in I} \Omega_i \sim \{0,1\}^{\mathbb{N}}$$

and

$$\varprojlim_{i \in I} \Omega_i' \sim [0,1]$$

It follows that the system of binomial distributions considered in examples 5.2 and 6.1 constitutes a canonical projective statistical field.

Genereally speaking, all the "canonical models" introduced in Martin-Löf (1974), are canonical projective statistical fields. This will be clear in the next section.

8. Construction of Canonical Fields.

Suppose that we have given two projective systems of probability spaces

$$\{(\Omega_i, i \in I), (P_{ij}, i < j)\} \quad \text{and}$$

$$\{(\Omega'_i, i \in I), (Q_{ij}, i < j)\} \quad ,$$

and a system of reductions and their "inverses"

$$R_i : \Omega_i \to \Omega'_i \; , \quad R^*_i : \Omega'_i \to \Omega_i \; , \quad i \in I \; ,$$

satisfying

a) $\quad R_i \, R^*_i = I_{\Omega'_i} \qquad \forall\, i \in I$

b) $\quad R^*_i \, Q_{ij} = P_{ij} \, R^*_j \qquad \forall\, i,j \in I : \; i < j$

Suppose that both systems have projective limits

$$\varprojlim_{i \in I} \Omega_i \sim \Omega$$

$$\varprojlim_{i \in I} \Omega'_i \sim \Omega'$$

We can then define a canonical projective statistical field by letting

$$\Theta = \Omega' \sim \varprojlim_{i \in I} \Omega'_i$$

and the parametrisations be given as

$$\mu_i = R^*_i \, Q_i \; .$$

To verify that the field

$$\mathcal{F} = \{(\Omega_i, i \in I), (P_{ij}, i < j), \Theta, (\mu_i, i \in I)$$

is canonical, we just have to prove that the system of reductions $(R_i, i \in I)$ is sufficient.

Condition i) is already supposed to be satisfied (assumption a)). ii) follows since

$$R_i^* R_i \mu_i = R_i^* R_i (R_i^* Q_i) = R_i^* (R_i R_i^*) Q_i = R_i^* Q_i = \mu_i \quad .$$

Finally iii) is true because of

$$R_i^* R_i P_{ij} R_j^* = R_i^* R_i R_i^* Q_{ij} = R_i^* (R_i R_i^*) Q_{ij} = R_i^* Q_{ij} = P_{ij} R_j^* \quad .$$

The construction procedure described here is exactly a formalisation of the procedure given in Martin-Löf (1970) and (1974).

There, the system $(\Omega_i, i \in I)$ is defined, the R_i's are considered to be given statistics and the R_i's are the conditional distributions on Ω_i given R_i with respect to certain canonically determined uniform measures (given by e.g. Riemannian metrics).

The consistency condition on some combinatorial coefficients given in the latter of the above mentioned papers is exactly corresponding to condition b) here.

In various situations it would be interesting to do the following. Suppose a usual statistical field is given

$$\{\Omega, P_\theta, \theta \in \Theta\} \quad . \tag{8.1}$$

One could then embed this statistical field into a relevant projective statistical field. Sometimes the relevant projective statistical field is formed by independent repetitions of the experiment leading to the field (8.1). But in many cases it would be a very different projective field, depending on the way the observations are produced. By doing this, one obtains a mathematically specified reference population.

Suppose this is done and we now have a projective statistical field

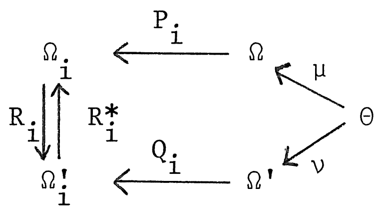
$$F = \{(\Omega_i, i \in I), (P_{ij}, i < j), \Theta, (\mu_i, i \in I)\} .$$

One would then look for a (minimal) sufficient system of reductions $(R_i, i \in I)$. Defining

$$Q_{ij} = R_i \, P_{ij} \, R_j^* \quad ,$$

one could then find the canonical statistical field generated by the system. If this is the same as F, the original field was canonical. If this is not the case, it might be convenient to modify the original statistical field by substituting the canonical field for F. At least it might throw some light on the relation between F and the corresponding canonical field.

If all projective limits exist, a projective statistical field can be non-canonical for three different reasons. We have the diagram



,

and the field is canonical iff ν is induced by a one-to-one and onto mapping. The deficiences can be divided into the following three types

I. <u>The Markov kernel ν is induced by a mapping, but this mapping is not one-to-one</u>.

In this case, the parameter $\theta \in \Theta$ will not be identifiable and it is convenient to introduce a new, identifiable parametrisation via Ω'.

II <u>ν is induced by a mapping that is one-to-one but not onto</u>.

In this case the parameter space has inconvenient restraints and there is not variability enough in the parameter space to take care of the variability in the population. It might be convenient to extend the model by substituting Ω' for Θ as parameter space.

III ν is not induced by a mapping but is truly random.

In this case the parameter will not be estimable even from complete observation of the entire population. This suggests that the parameter is not well-defined as a characteristic of an empirically observable phenomenon. The model contains superfluous randomisation. It might be convenient to remove this randomisation by a conditioning procedure, i.e. by substituting Ω' for Θ.

It should be noted that the "new" statistical field induced on a given field by embedding this into a projective field, considering reduction of this and finally substituting the generated canonical field is very sensitive and depends drastically on into which projective system it is embedded.

This should encourage one to be careful, using the procedure. However, statistical models ought to be very dependent on the reference populations that they are supposed to describe!

In Lauritzen (1975) projective statistical fields corresponding to independent identically distributed random variables, whose distributions are supposed to be member of a (somewhat generalised) exponential family of distributions are discussed. In such situations it is shown that the canonical fields roughly correspond to "full" exponential families.

Below we shall give two examples where the relation between a projective statistical field and the corresponding canonical field are different.

<u>Example 8.1</u>

Here we shall investigate a classical problem first discussed by Neyman and Scott (1948).

Let

$$\Omega_i = \{(x_1,y_1;\ldots;x_i,y_i)\colon (x_j,y_j) \in \mathbb{R}^2,\ j=1,\ldots,i\}\ .$$

and let $P_{ij}\colon \Omega_j \to \Omega_i$ be induced by the coordinate projections. Let

$$\Theta = \mathbb{R}^{\mathbb{N}} \times [0,\infty[\ ,$$

and let the measure $\mu_i(\theta)$ of a parameter

$$\theta = [(\xi_j, j \in \mathbb{N}), \sigma^2]$$

be the product of 2i normal distributions with means

$$E\,X_j = E\,Y_j = \xi_j\ , \quad j = 1,\ldots,i$$

and variances σ^2.

This is the model for determining the accuracy $(1/\sigma^2)$ of a measuring instrument by considering double measurements (X_j,Y_j) of unknown quantities ξ_j, $j=1,2,\ldots$.

We then define

$$\Omega_i' = \mathbb{R}^i \times [0,\infty[\ ,$$

the reductions $R_i\colon \Omega_i \to \Omega_i'$ by

$$R_i((x_1,y_1;\ldots;x_i,y_i)) = (x_1+y_1,\ldots,x_i+y_i;\ \sum_{j=1}^{i} (x_i - y_i)^2)$$

and R_i^* as the conditional distribution of $(X_1,Y_1,\ldots,X_i,Y_i)$ given R_i in the normal distributions mentioned.

Clearly $(R_i, i \in I)$ is a minimal sufficient system of reductions.

Now we let

$$Q_{ij} = R_i \; P_{ij} \; R_j^* \quad .$$

$Q_{ij}\{(s_1,\ldots,s_j,S_j^2)\}$ is the probability measure on $\mathbb{R}^j \times [0,\infty[$ which is the product of a measure degenerate in $(s_1,\ldots,s_j)$ and a beta-distribution with density w.r.t. Lebesgue-measure

$$f_{ij}(x \mid S_j^2) = \frac{1}{S_j^2 \; B(\frac{i}{2},\frac{j-i}{2})} \left(\frac{x}{S_j^2}\right)^{\frac{i}{2}-1} \left(1 - \frac{x}{S_j^2}\right)^{\frac{j-i}{2}-1} \quad .$$

It is an easy consequence of Stirlings formula that

$$\lim_{j\to\infty} f_{ij}(x \mid S_j^2) = \frac{1}{\Gamma(\frac{i}{2})} \left(\frac{x}{2\sigma^2}\right)^{\frac{i}{2}-1} e^{-\frac{x}{2\sigma^2}}$$

when $S_j^2 / j \to 2\sigma^2$.

Combining this fact, Boltzmann's law and the fact that a genuine mixture of gamma densities cannot be a gamma density, one obtains the result, that the projective limit of $(\Omega_i^!, i \in I)$ is

$$\varprojlim_{i\in I} \Omega_i^! \sim \mathbb{R}^{\mathbb{N}} \times [0,\infty[\; = \Omega'$$

where the probability measures $Q_i(\omega)$, $\omega \in \Omega'$, are given as follows.

If $\omega = ((z_j, j \in \mathbb{N}),\sigma^2)$ then $Q_i(\omega)$ is the distribution on $\mathbb{R}^i \times [0,\infty[$ given as the product of a distribution degenerate in $(z_1,\ldots,z_i)$ and a χ^2-distribution with i degrees of freedom and σ^2 as scale parameter.

Defining $\Theta^* = \Omega'$ and $\mu_i^* = R_i^* \; Q_i$, we have that the modified (canonical) projective statistical field

$$\mathcal{F}^* = \{(\Omega_i, i \in I),(P_{ij}, i < j),\Theta^*,(\mu_i^*, i \in I)\}$$

is obtained from the original field

$$F = \{(\Omega_i, i \in I), (P_{ij}, i < j), \Theta, (\mu_i, i \in I)\}$$

by <u>conditioning</u> on the statistics $(x_1,y_1;\dots;x_i,y_i) \to (x_1+y_1,\dots,x_i+y_i)$, i.e., the statistics minimal sufficient for the "nuisance parameters" $(\xi_j, j \in \mathbb{N})$.

It is worth noting that the vague statement "$\xi_j, j \in \mathbb{N}$ are nuisance parameters and σ^2 the parameter of interest" is substituted by a precise specification of the reference population in question. The procedure suggested to remove deficiencies of type III here plays the role of a conditionality principle.

<u>Example 8.2</u> (Autoregressive processes).

Let I consist of bounded intervals on the integer axis ordered by inclusion

$$I = \{(s,s+1,\dots,t) : s < t, s,t \in \mathbb{Z}\}.$$

Let for $[s,t] \in I$

$$\Omega_{[s,t]} = \mathbb{R}^{t-s+1}$$

and let as usual $P_{[s,t][u,v]}$ be induced by coordinate projections.

Let

$$\Theta =]-1,1[\times [0,\infty[$$

and let for $\theta = (\theta_1,\theta_2) \in \Theta$, $\mu_{[s,t]}(\theta)$ be the normal distribution on $\mathbb{R}^{t-s+1}$ with mean zero and inverse covariance given as

$$\Sigma^{-1}_{s,t}(\theta) = \frac{1}{\theta_2(1-\theta_1^2)} \begin{pmatrix} 1 & -\theta_1 & 0 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ -\theta_1 & 1+\theta_1^2 & -\theta_1 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & -\theta_1 & 1+\theta_1^2 & -\theta_1 & 0 & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & -\theta_1 & 1+\theta_1^2 & -\theta_1 & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & 0 & -\theta_1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -\theta_1 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 & -\theta_1 & 1+\theta_1^2 & -\theta_1 \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 & 0 & -\theta_1 & 1 \end{pmatrix}.$$

This corresponds to the autoregressive scheme of order one defined as the unique gaussian stationary process

$$\{X(t), t \in Z\}$$

satisfying the stochastic difference-equation

$$X(t+1) - \theta_1 X(t) = \varepsilon(t),$$

where $\{\varepsilon(t), t \in Z\}$ is supposed to be gaussian white noise with mean zero and variance σ^2. Equivalently the covariance structure is given as

$$E\, X(s)\, X(t) = \theta_2\, \theta_1^{|s-t|}.$$

From the expression defining $\Sigma_{s,t}^{-1}(\theta)$ one clearly has that the likelihood function is proportional to

$$L(x_s, \ldots, x_t; \theta_1, \theta_2)$$

$$\sim \psi(\theta_1, \theta_2) \exp\{-\frac{1}{2\theta_2(1-\theta_1^2} \cdot [x_s^2 + x_t^2 + (1+\theta_1^2) \sum_{i=s+1}^{t-1} x_i^2 - 2\theta_1 \sum_{i=s}^{t-1} x_i\, x_{i+1}]\}\ .$$

Thus the statistic

$$(x_s, \ldots, x_t) \xrightarrow{\tilde{R}_{s,t}} (x_s^2 + x_t^2, \sum_{i=s+1}^{t-1} x_i^2, \sum_{i=s}^{t-1} x_i\, x_{i+1})$$

is minimal sufficient in the usual sense.

$\tilde{R}_{s,t}$ is <u>not</u> a sufficient system of reductions in the sense defined here. One can show that the statistics $R_{s,t}$ defined as

$$(x_s, \ldots, x_t) \xrightarrow{R_{s,t}} (x_s, \sum_{i=s+1}^{t-1} x_i^2, \sum_{i=s}^{t-1} x_i\, x_{i+1}, x_t)$$

constitute a minimal sufficient system of reductions.

It is quite a bit technical, and the details are omitted in the present paper, but possible to show, that the canonical projective statistical field generated by this system of reductions has parameter space

$$\Theta^* = \mathbb{R} \times \mathbb{R} \times]-1,1[\times [0,\infty[$$

and that the canonical distributions of $(X(s),\ldots,X(t))$ with parameter

$$\theta^* = (\alpha,\beta,\theta_1,\theta_2)$$

is given as the multivariate normal distribution with the <u>same</u> covariance as before but with <u>mean</u>

$$E\ X(t) = \alpha\theta_1^{t} + \beta\theta_1^{-t} ,$$

i.e. a model with a <u>trend</u> of form as a hyperbolic sine/cosine superimposed with autoregression errors.

Hence in this example the original model has a deficiency of type II compared to the reference population.

References

Andersen, E. Sparre, B. Jessen (1948). On the introduction of measures in infinite product sets. Danske Videnskabers Selskab. Mat.-Fys.Medd. 25 No.4.

Bahadur, R.R. (1954). Sufficiency and statistical decision functions. Ann.Math.Statist, 25, 423 - 462.

Hald, A. (1960). The compound hypergeometric distribution and a system of single sampling inspection plans based on prior distributions and costs. Technometrics, 2, 275 - 340.

Khinchin, A.I. (1949). Mathematical Foundations of Statistical Mechanics. Dover. New York.

Lauritzen, S.L. (1972). Sufficiency and Time Series Analysis. Preprint No. 11. UCIMS. (Mimeographed report).

Lauritzen, S.L. (1974a). Sufficiency, Prediction and Extreme Models. Scand.J. Statist., 1, 128 - 134.

Lauritzen, S.L. (1974b). On the interrelationships among suffifiency, total sufficiency and some related concepts. Preprint No. 8. UCIMS. (Mimeographed report).

Lauritzen, S.L. (1975). General Exponential Models for Discrete Observations. Scand.J.Statist., 2, 23 - 33.

Martin-Löf, P. (1970). Statistiska Modeller. Institutionen för Matematisk Statistik, Stockholm University.(Mimeographed report.)

Martin-Löf, P. (1974). Repetitive Structures. Proceedings of conference on foundational questions in statistical inference. Aarhus, may 7-12, 1973. Memoirs No.1. 1974. Department of theoretical statistics. Aarhus University.

Neyman, J. and Scott, E. (1948). Consistent estimates based on partially consistent observations. Econometrika, 16, 1 - 32.

Rasch, G. (1960). Probabilistic Models for Some Intelligence and Attainment Tests, Studies in Mathematical Psychology I. Danmarks Pædagogiske Institut.

Scheffer, C.L. (1971). La limite projective des espaces topologiques de Hausdorff compacts dans la catégorie des diffusions continues sousmarkoviennes. C.R.Acad.Sc.Paris A, 272, 1198 - 1201.