Bollettino
Unione Matematica Italiana

Anna Gerardi, Barbara Torti

Symmetric models for lifetimes: the role of exchangeable equivalence relations

Bollettino dell'Unione Matematica Italiana, Serie 8, Vol. **6-B** (2003), n.1, p. 111–123.

Unione Matematica Italiana

<http://www.bdim.eu/item?id=BUMI_2003_8_6B_1_111_0>

Bollettino dell'Unione Matematica Italiana, Unione Matematica Italiana, 2003.

Bollettino U. M. I.
(8) **6**-B (2003), 111-123

Symmetric Models for Lifetimes: the Role of Exchangeable Equivalence Relations.

ANNA GERARDI - BARBARA TORTI

Sunto. – *Si considera un modello relativo ad una popolazione eterogenea suddivisa in un numero finito di classi, in accordo con una relazione di equivalenza scambiabile. Con questa motivazione si studiano le proprietà delle relazioni di equivalenza scambiabili ed in particolare se ne caratterizza la struttura delle classi di equivalenza.*

Summary. – *A model of a heterogeneous population partitioned into a finite number of classes according an exchangeable equivalence relation is studied. With this motivation the properties of exchangeable equivalence relations are investigated and, in particular, the structure of its equivalence classes is characterized.*

1. – Introduction.

Exchangeable Equivalence Relations (EER) have been introduced by Kingman ([5]) in order to study genealogical properties of biological populations consisting of identical individuals.

In several papers (see [5], [6], [7], also for further references) Kingman points out that some properties of a given population, such as the age or genetic mutations derive from the correlation among individuals. To characterize these properties, he introduces the EER's which, at any time, define both correlation and exchangeability. An EER on $\mathbb{N}$ is, roughly speaking, a random partition of $\mathbb{N}$ whose distribution is invariant under permutation (see Def. 2.2 below).

Moreover, when the dimension of the population is infinite, Kingman shows that the distribution of any EER has an explicit representation.

Another point of interest is to study the connection between EER's and exchangeable sequences $\{Z_i\}$ of real-valued r.v.'s. Actually, given $\{Z_i\}$ it is quite obvious that the relation $\mathcal{R}$ defined by

$$(i, j) \in \mathcal{R} \iff Z_i = Z_j \tag{1}$$

is an EER. We say, in this case, that the sequence $\{Z_i\}$ generates $\mathcal{R}$. On the other hand Aldous proves in ([1]) that any EER is generated by an exchangeable sequence that is, given $\mathcal{R}$, there exists a (not unique) sequence of exchangeable r.v.'s such that Eq. (1) holds.

Two are the main results in this paper.

The first one provides a characterization of the equivalence classes of an arbitrary EER in terms of the atoms of the De Finetti measure associated to any sequence generating the EER under consideration.

The second one consists in using this characterization in order to improve a model arising in reliability and survival analysis.

More precisely, let $\mathcal{R}$ be an EER and, taking into account the result in [1], let $\{Z_i\}$ be a sequence generating $\mathcal{R}$. As a consequence of De Finetti theorem, we can define a random measure α such that the r.v.'s Z_i are i.i.d. given α. We will prove that the equivalence classes of $\mathcal{R}$ can be characterized a.s. in terms of the atoms of the purely discrete component α^d of α. Moreover we can obtain the same representation given by Kingman, pointing out that the distribution of $\mathcal{R}$ only depends on the (random) weights of the atoms of α^d. This implies that, while different sequences can generate the same EER, the weights of the atoms of α^d are uniquely determined.

The structure of the reliability model makes of interest to consider, in particular, the case when α is purely atomic. The model we deal with can be described as follows.

Consider a population $\{U_i\}$ of indistinguishable individuals, divided into a finite number of different types according to their surviving capacity. We suppose that the subdivision in types induces a non-observable EER, $\mathcal{R}$, on the population. Let T_i be the r.v.'s representing the lifetime of the i-th item. The r.v.'s T_i are supposed to be independent given the partition induced by $\mathcal{R}$. Moreover we suppose that any T_i has a distribution function G_k, given the event $\{U_i$ is of type k$\}$. Our aim is to compute the joint distribution of the lifetimes of an n-dimensional sample drawn from the population.

Observe that in this model, given $\mathcal{R}$ one can merely say whether two individuals are of the same type, without specifying their type. In this respect our model is more general than the models considered in [8] and in [3]. In [8] the population was supposed to be finite and labelled into two different types. In [3] the population is assumed to consist of an infinite sequence of items partitioned into many different types. In both models an exchangeable sequence Z_i of non-observable r.v.'s is considered, such that $Z_i = k$ if and only if the i-th item is of type k.

Nevertheless, even if our assumptions are more general, the joint distribution of lifetimes for the sample, turns out to have the same expression obtained in [3], [4].

The paper is organized as follows. In Section 2 basic definitions are given and Aldous theorem is recalled. In Section 3, by using De Finetti theorem, we obtain the general representation of an EER distribution. In Section 4 the structure of the equivalence classes is investigated, and finally, in the last Section, the application to reliability model is given. Moreover, in the Appendix, some technical properties of random probability measures are discussed.

2. – Exchangeable equivalence relations and exchangeable sequences.

Let E be the set of equivalence relations on the natural numbers $\mathbb{N}$. Any $\xi \in E$ can be identified with the family of the pairs of individuals in relation according to $\mathcal{R}$, so that E can be regarded as a closed subset of $\{0, 1\}^{\mathbb{N}\times\mathbb{N}}$ endowed with the product topology. With respect to the relative topology E is a Polish space. Let $\mathcal{B}(E)$ be the Borel σ-algebra of E. Let us note that any finite permutation $\pi : \mathbb{N} \to \mathbb{N}$ defines an application (which we again call π, with a little abuse of notations)

$$(2) \qquad \pi : E \to E \qquad \xi \to \pi\xi = \{(i, j) : (\pi^{-1} i, \pi^{-1} j) \in \xi\}$$

DEF. 2.1. – A probability measure P on $(E, \mathcal{B}(E))$ is called exchangeable if it is invariant under the application π, that is $P\pi^{-1} = P$, $\forall\pi$.

DEF. 2.2. – An E-valued random variable $\mathcal{R}$ is an Exchangeable Equivalence Relation (EER) if its distribution is exchangeable, that is, $\forall\pi$, $\mathcal{L}(\mathcal{R}) = \mathcal{L}(\pi\mathcal{R})$.

The following theorem can be found in [1].

THEOREM 2.3. – Let $\mathcal{R}$ be an EER on $\mathbb{N}$. There exists an exchangeable sequence $\{Z_i\}_{i\in\mathbb{N}}$ of non-negative r.v's, such that

$$(3) \qquad (i, j) \in \mathcal{R} \iff Z_i = Z_j$$

PROOF. – Consider, for any $i \in \mathbb{N}$, the random variable

$$M_i = \min\{j \in \mathbb{N} : (i, j) \in \mathcal{R}\}$$

Obviously

$$(i, j) \in \mathcal{R} \iff M_i = M_j.$$

Now, let $\zeta = \{\zeta_i\}$ be a sequence of independent random variables uniformly distributed on the interval $(0, 1)$. Suppose also the sequence ζ independent on $\mathcal{R}$ and put

$$Z_i = \zeta_{M_i}.$$

It is easy to see that $\{Z_i\}_{i\in\mathbb{N}}$ is an exchangeable sequence such that Eq. (3) holds. ∎

REMARK 2.4. – Observe that since $M_1 = 1$, the common marginal law of Z_i is uniform on (0, 1). Moreover, it is easy to see that any other sequence of i.i.d. r.v.'s could be used. The only requirement is that the common marginal law has to be purely continuous to guarantee that

$$Z_i = Z_j \Leftrightarrow M_i = M_j .$$

Different sequences can generate the same EER. In the following example we consider a case of a sequence of purely discrete r.v.'s. and we construct an EER with a procedure known as the «paintbox construction» ([7]). This construction plays a crucial role in the representation theorem.

Let $Z = \{Z_i\}_{i\in\mathbb{N}}$ be a sequence of i.i.d. random variables taking values in $\mathbb{N} \cup \{0\}$ and let $x_k = P(Z_i = k)$, $\forall i \geq 1, k \geq 0$.

Note that $0 \leq x_k \leq 1$ for any k and $\sum x_k = 1$. Define the E-valued random variable $\mathcal{R}_x$ by the rule

$$\mathcal{R}_x = \{(i, j) : i = j \text{ or } Z_i = Z_j \geq 1\}.$$

Let us observe that when $Z_i = Z_j = 0$, $(i, j) \in \mathcal{R}$ iff $i = j$. In other words, $Z_i = 0$ implies that the equivalence class containing i reduces to $\{(i, i)\}$.

PROP. 2.5. – $\mathcal{R}_x$ is an EER and its distribution only depends on the sequence $x = \{x_k, k \geq 0\}$.

PROOF. – It is easy to see that $\mathcal{R}_x$ is an EER. Moreover, $\forall i \neq j$

$$P((i, j) \in \mathcal{R}_x) = \sum_{k \geq 1} x_k^2 \tag{4}$$

Then we see that the distribution of $\mathcal{R}_x$ is an exchangeable probability measure P_x on E characterized by the sequence $x \in \nabla$ where

$$P_x(\cdot) = P(\mathcal{R}_x \in \cdot); \qquad \nabla = \left\{x = (x_0, x_1, \dots) : x_k \geq 0 \text{ and } \sum_{k \geq 1} x_k \leq 1\right\} \qquad ∎$$

In the sequel we consider ∇ as a subspace of $\mathbb{R}^{\mathbb{N}}$ endowed with the usual topology.

3. – The representation theorem.

In this section we will characterize the structure of the law of an EER by using some elementary properties of random measures (see Thm. 3.5. below). Results of this kind are well known (see [1], [7]). However we give a different proof, and let us notice that our proof allows to obtain an useful description of the equivalence classes of $\mathcal{R}$. The probabilistic properties of such classes will be investigated in the next Section.

First we analyze some properties of exchangeable sequences.

Consider an exchangeable sequence of r.v.'s $\{Z_i\}_{i\in\mathbb{N}}$, and let $\mathcal{R}$ be the EER defined by Eq. (3). Without loss of generality (see Thm. 2.3.) we can assume, $\forall i$, $Z_i \geqslant 0$.

De Finetti theorem states that there exists a σ-algebra $\mathcal{E}$ (the exchangeable σ-algebra), such that the r.v.'s Z_i are independent and identically distributed given $\mathcal{E}$. This allows us to define a random variable α_Z taking values on the space $\Pi(\mathbb{R})$ of probability measures on $\mathbb{R}$, endowed with the weak convergence topology, such that, for each $I \in \mathcal{B}(\mathbb{R})$,

$$\alpha_Z(I) = P(Z_i \in I \mid \mathcal{E}) \qquad \forall i \tag{5}$$

The measure α_Z satisfies

$$\alpha_Z([0, +\infty)) = 1, \qquad P(Z_1 \in H_1, \ldots, Z_n \in H \mid \alpha_Z = a) = \prod_{i=1}^{n} a(H_i).$$

Note that the purely atomic component α_Z^d of α_Z is a measure-valued r.v. such that $\alpha_Z^d(\mathbb{R}) \leqslant 1$. Moreover it can be represented as

$$\alpha_Z^d = \sum_{k\geqslant 1} X_k \delta_{\{Y_k\}}$$

where $X = (X_1, X_2, \ldots)$, $Y = (Y_1, Y_2, \ldots)$ are two sequences of real-valued r.v.'s such that

$$X_k \geqslant 0, \qquad \sum_{k\geqslant 1} X_k \leqslant 1, \qquad Y_k \geqslant 0 .$$

(see the appendix for technical details).

PROP. 3.1. – Let $a \in \Pi(\mathbb{R})$ be such that

$$a^d = \sum_{k\geqslant 1} x_k \delta_{\{y_k\}}$$

(a^d is the purely atomic components of a). Then

$$P(Z_i = Z_j \mid \alpha_Z = a) = P((i,j) \in \mathcal{R}_x) \tag{6}$$

where $\mathcal{R}_x$ is defined by Eq. (4).

PROOF. – The conditional independence of the Z_i implies

$$P(Z_i = Z_j \mid \alpha_Z = a) = \int_A a(dx)\, a(dy) = \int_{\mathbb{R}} \int_{A_x} a(dx)\, a(dy)$$

where $A = \{(x, y) \in \mathbb{R}^2 : x = y\}$ and, $\forall x \in \mathbb{R}$, $A_x = \{y \in \mathbb{R} : (x, y) \in A\} = \{x\}$.

On the other hand a discrete measure a^d and a continuous measure a^c such that $a = a^d + a^c$ are uniquely determined and $a^c(A_x) = 0$. Thus

$$P(Z_i = Z_j \mid \alpha_Z = a) = \int_{\mathbb{R}} \int_{A_x} a(dx)\, a^d(dy) = \int_{\mathbb{R}} a^d(\{x\})\, a(dx) = \sum_{k \geq 1} x_k^2$$

and the statement is proved, by Eq. (4). ■

As a straightforward consequence we have

PROP. 3.2. – If $a \in \Pi(R)$ is a purely continuous measure, then $P(Z_i = Z_j \mid \alpha_Z = a) = 0$.

REMARK 3.3. – From Prop. 3.1. we get the following equalities

$$P(Z_i = Z_j \mid \alpha_Z) = P(Z_i = Z_j \mid \alpha_Z^d) = P(Z_i = Z_j \mid X, Y) = P(Z_i = Z_j \mid X).$$

We are now able to give the main result of this section

THEOREM 3.4. – For any exchangeable sequence Z_i of non-negative r.v.'s, there exists a ∇-valued r.v. $\underline{X} = (X_0, X_1, X_2, \ldots)$ such that

$$P(Z_i = Z_j) = \int_{\nabla} P((i, j) \in \mathcal{R}_x)\, \nu_{\underline{X}}(dx)$$

where $\nu_{\underline{X}}$ is the law of $\underline{X}$ on ∇.

PROOF. – The statement is a consequence of Prop. 3.1. In fact the sequence $(X_1, X_2, \ldots)$ of the weights of the atoms of α_Z uniquely defines a ∇-valued r.v. $\underline{X}$ by setting

$$\underline{X} = (X_0, X_1, \ldots) \quad \text{and} \quad X_0 = 1 - \sum_{k \geq 1} X_k = \alpha_Z^c(\mathbb{R}). \qquad ■$$

By using this result, we are able to give our proof of the representation theorem.

THEOREM 3.5. – Let $\mathcal{R}$ be an EER. There exists a random variable $\underline{X}$ on ∇, such that, its distribution $\nu_{\underline{X}}$ is uniquely determined, and

$$P(\mathcal{R} \in \cdot) = \int_{\nabla} P_x(\cdot)\, \nu_{\underline{X}}(dx)$$

where P_x is the measure defined by the paintbox construction.

PROOF. – Let $\{Z_i\}_{i\in\mathbb{N}}$ be an exchangeable sequence of non-negative r.v.'s generating $\mathcal{R}$. For instance one can consider the sequence given in Thm. 2.3. Since

$$P(Z_i = Z_j) = P((i, j) \in \mathcal{R})$$

the statement is an easy consequence of Thm. 3.4. ■

REMARK 3.6. – For any $\mathcal{R}$, the sequence Z generating $\mathcal{R}$ is not uniquely determined. However, for any exchangeable non-negative sequence generating $\mathcal{R}$, are uniquely determined the weights $X_1, X_2, \ldots$ of the purely atomic component of the associated random measure.

4. – Equivalence classes and occupation numbers.

Let $\mathcal{R}$ be an EER and let $Z = (Z_1, Z_2, \ldots)$ be any exchangeable sequence of non-negative r.v.'s which generates $\mathcal{R}$. We remember that Z uniquely defines two sequences $X = (X_1, X_2, \ldots)$ and $Y = (Y_1, Y_2, \ldots)$ such that the random measure associated to Z is

$$\alpha_Z = \sum_{k\geq 1} X_k \delta_{\{Y_k\}} + \alpha_Z^c .$$

We want to discuss the structure of the equivalence classes of $\mathcal{R}$. Consider the a.s. disjoint sets

$$\mathcal{C}_k = \{i \in \mathbb{N} \text{ s.t. } Z_i = Y_k\} \qquad \forall k \geq 1 . \tag{7}$$

Each nonempty $\mathcal{C}_k$ is an equivalence class of $\mathcal{R}$. Let

$$\mathcal{C}_0 = \mathbb{N} - \bigcup_{k\geq 1} \mathcal{C}_k$$

and observe that $\forall i \in \mathcal{C}_0$, the equivalence class of $\mathcal{R}$ containing i reduces to $\{(i, i)\}$ a.s., that is

$$\forall i = j \qquad P(i \in \mathcal{C}_0, (i, j) \in \mathcal{R}) = P(\{Z_i \neq Y_k, \forall k\} \cap \{Z_i = Z_j\}) = 0 .$$

Since $P(i \in \mathcal{C}_0 \mid X) = X_0$, the family of the nonempty sets defined by Eq. (7) coincides with the equivalence classes of $\mathcal{R}$ iff $X_0 = 0$, that is iff α_Z is purely atomic. As we will see, this is the case when $\mathcal{R}$ has a finite number of classes. To study such a situation we need a preliminary result.

Let E_n be the set of equivalence relations on $\{1, 2, \ldots, n\}$ and define the (measurable) map $\varrho_n : E \to E_n$ by

$$\forall \xi \in E \qquad \varrho_n \xi = \big\{(i, j) \in \xi \text{ s.t. } i, j \in \{1, 2, \ldots, n\}\big\} .$$

Consider the r.v. $\varrho_n \mathcal{R}$. It is clear that it is an equivalence relation on

$\{1, 2, \dots, n\}$ whose distribution is invariant under permutation of $\{1, 2, \dots, n\}$. Define, $\forall k \geq 1$

$$\lambda_k^n = \sum_{i=1}^{n} \mathbb{I}_{\{Z_i = Y_k\}}.$$

Call $\{\lambda_k^n\}$ the occupation numbers of $\varrho_n \mathcal{R}$.

PROP. 4.1.

$$\lim_{n \to \infty} \frac{\lambda_k^n}{n} = X_k \tag{8}$$

PROOF. – The sequence $\{\mathbb{I}_{\{Z_i = Y_k\}}\}$ is an exchangeable sequence and its exchangeable σ-algebra coincides with $\mathcal{E}$. Then, the thesis is a consequence of the strong law of large numbers for exchangeable sequences [2]

$$\lim_{n \to \infty} \frac{\lambda_k^n}{n} = E[\mathbb{I}_{\{Z_i = Y_k\}} \mid \mathcal{E}] = \alpha_Z(\{Y_k\}) = X_k. \qquad \blacksquare$$

Previous proposition allows us to look at the r.v.'s $X_1, X_2, \dots$ as frequencies of the elements belonging to the equivalent classes of $\varrho_n \mathcal{R}$. This result was proved by Kingman together with Thm. 3.4. ([7]); our proof is directly related to the approach followed in this paper.

From now on we assume that $\mathcal{R}$ has a finite number of classes: $|\mathcal{R}| \leq d$ a.s.. Setting

$$\mathcal{C}_k^n = \{i \in \{1, \dots, n\} \text{ s.t. } Z_i = Y_k\} \qquad k = 1, \dots, d$$

we have that $\lambda_k^n = \mathrm{Card}(\mathcal{C}_k^n)$.

PROP. 4.2. – Suppose $|\mathcal{R}| \leq d$ a.s.. Then, the random measure defined by Eq. (5) is purely atomic.

PROOF. – We note that Prop. 4.1 implies that $X_k \neq 0$ for all but a finite number of indexes. Then,

$$\sum_{k \geq 1} \frac{\lambda_k^n}{n} = 1 \quad \forall n \geq 1 \quad \text{and} \quad \sum_{k \geq 1} X_k = \lim_{n \to \infty} \sum_{k \geq 1} \frac{\lambda_k^n}{n} = 1. \qquad \blacksquare \tag{9}$$

We want to find the joint distribution of the occupation numbers. To this end we first need the following

LEMMA 4.3. – Let $k_1, \ldots, k_n$ be natural numbers such that $k_i \in \{1, \ldots, d\}$, $\forall i = 1, \ldots, n$.

$$P(1 \in \mathcal{C}^n_{k_1}, \ldots, n \in \mathcal{C}^n_{k_n}) = E[X_{k_1} \cdots X_{k_n}]. \tag{10}$$

PROOF. – For any n,

$$\begin{aligned} P(1 \in \mathcal{C}^n_{k_1}, \ldots, n \in \mathcal{C}^n_{k_n}) &= E\big[P\big(Z_1 = Y_{k_1}, \ldots, Z_n = Y_{k_n} \,|\, X, Y\big)\big] = \\ &= E\Big[\prod_{i=1}^{n} P\big(Z_i = Y_{k_i} \,|\, X, Y\big)\Big]. \qquad \blacksquare \end{aligned} \tag{11}$$

PROP. 4.4. – The joint distribution of the occupation numbers is a mixture of multinomials and it is a multinomial distribution only when the distribution ν_X of X is a degenerate distribution.

PROOF. – Recall that

$$\lambda^n_k = \sum_{i=1}^{n} \mathbb{I}_{\{Z_i = Y_k\}} \qquad \text{and} \qquad \sum_{k=1}^{n} \lambda^n_k = n\,.$$

Let $h_1, \ldots, h_d$ be natural numbers such that $h_1 + \ldots + h_d = n$. By the exchangeability of the Z_i

$$P(\lambda^n_1 = h_1, \ldots, \lambda^n_d = h_d \,|\, XY) = \frac{n!}{h_1! \cdots h_d!} X_1^{h_1} \cdots X_d^{h_d}.$$

Then

$$P(\lambda^n_1 = h_1, \ldots, \lambda^n_d = h_d) = E[P(\lambda^n_1 = h_1, \ldots, \lambda^n_d = h_d \,|\, X, Y)] \tag{12}$$

and finally

$$P(\lambda^n_1 = h_1, \ldots, \lambda^n_d = h_d) = \frac{n!}{h_1! \cdots h_d!} E[X_1^{h_1} \cdots X_d^{h_d}]. \qquad \blacksquare \tag{13}$$

5. – The reliability model.

Consider a countable heterogeneous population $\mathcal{P} = \{U_i\}$ and let $\{T_i\}$ be the family of non-negative r.v.'s denoting the lifetimes of each item $U_i \in \mathcal{P}$. Let also $\{G_k\}_{k=1,\ldots,d}$ be given distribution functions and suppose that the population is divided into d different types and that the lifetime T_i of any individual U_i of type k admits the distribution function G_k.

In this section we suppose that the information about the relationships among individuals are characterized by an EER which selects the individuals of the same type. Under these conditions we want to determine the joint survival function of a sample lifetimes $T_1, \dots, T_n$.

More precisely, define the equivalence relations $\mathcal{R}$ by the rule

$$(i, j) \in \mathcal{R} \Leftrightarrow U_i \text{ and } U_j \text{ are of the same type}$$

and assume that $\mathcal{R}$ is any EER. Let $\{Z_i\}$ be an exchangeable sequence of nonnegative r.v.'s generating $\mathcal{R}$. Since $|\mathcal{R}| \leqslant d$, α_Z is a purely atomic measure

$$\alpha_Z = \sum_{k \geqslant 1} X_k \delta_{\{Y_k\}}.$$

We are then in the frame discussed in Section 4, and the equivalence classes of $\mathcal{R}$ coincide with the nonempty sets of the family

$$\mathcal{C}_k = \{i \in \mathbb{N} \text{ s.t. } Z_i = Y_k\} \qquad k = 1, \dots, d.$$

We make the following assumption $\forall n \in \mathbb{N}$, $j_1, \dots, j_n \in \mathbb{N}$, $k_1, \dots, k_n \in \{1, \dots, d\}$, and for any permutation π on $\{1, \dots, d\}$,

$$P(T_{j_1} \leqslant t_1, \dots, T_{j_n} \leqslant t_n \,|\, j_1 \in \mathcal{C}_{\pi k_1}, \dots, j_n \in \mathcal{C}_{\pi k_n}) = \prod_{i=1}^{n} G_{k_i}(t_i). \tag{14}$$

Note that Eq. (14) implies that $T_1, \dots, T_n$ are conditionally independent given $\mathcal{R}$. We are also assuming that each item belonging to the class $\mathcal{C}_{\pi_k}$ is of type k, and this is an arbitrary choice. The particular choice $\Pi_k = k$ reproduces the model discussed in [3]. On the other hand, the structure of the joint survival function of $T_1, \dots, T_n$ turns out to be independent of this choice.

In fact, taking into account Eq. (10) the joint distribution function of $T_1, \dots, T_n$ is given by

$$P(T_1 \leqslant t_1, \dots, T_n \leqslant t_n) = \sum_{k_1, \dots, k_n}^{1, \dots, d} G_{k_1}(t_1) \cdots G_{k_n}(t_n)\, E[X_{k_1} \cdots X_{k_n}] \tag{15}$$

which coincides with the same equation obtained in [3]. Moreover

PROP. 5.1. – The sequence $\{T_i\}$ is exchangeable.

PROOF. – For any $n > 1$ and for any permutation π on $\{1, \dots, n\}$, Eq. (10) implies that

$$P(T_1 \leqslant t_1, \dots, T_n \leqslant t_n) = P(T_{\pi 1} \leqslant t_1, \dots, T_{\pi n} \leqslant t_n). \qquad \blacksquare$$

The previous Proposition allows us to give a concluding remark. Let us point out that we can define a EER on $\mathcal{P}$ by the rule

$$(i, j) \in \mathcal{R}_T \iff T_i = T_j$$

and $\mathcal{R}_T$ has the properties discussed in Section 4.

Then the following holds.

PROP. 5.2. – Given the distribution functions $\{G_1, \ldots, G_d\}$, and a sequence $\{T_i\}$ of non negative random variables, under the assumptions

(i) At most d different types can be found in $\mathcal{P}$.

(ii) The sequence $\{T_i\}$ is exchangeable,

then, taking into account Eq. (15), the model is uniquely determined by a law on the set

$$\nabla_d = \left\{ x = (x_1, \ldots, x_d) : x_k \geq 0 \text{ and } \sum_{k=1}^{d} x_k = 1 \right\}.$$

Let us finally observe that a different ordering of the classes implies a different assignment of the type of an item belonging to the class $\mathcal{C}_k$. This in turn reduces to a different choice of the law on ∇_d.

Appendix.

Let $\Pi(\mathbb{R}^+)$ be the space of the probability measures on $\mathbb{R}^+$ endowed with the topology of the weak convergence. It is well known that it is a separable and complete metrizable space.

If α is a $\Pi(\mathbb{R}^+)$-valued r.v., there exist two r.v.'s $X = (X_1, X_2, \ldots)$ and $Y = (Y_1, Y_2, \ldots)$ such that

i) $X_k \geq 0$, $\forall k$, and $\sum_{k \geq 1} X_k \leq 1$,

ii) $Y_k \geq 0$, $\forall k$,

iii) the discrete component α^d of α is a r.v. that admits the representation

$$\alpha^d = \sum_{k \geq 1} X_k \delta_{\{Y_k\}}.$$

The proof is a consequence of the following properties

(**1**) $\forall T > 0$, $\forall t \in [0, T]$, let $a(t)$ be a real-valued non-decreasing cadlag function and $\forall \delta > 0$, set

$$\omega_T(\delta) = \sup_{|u - v| < \delta} |a(u) - a(v)|.$$

Then

$$\lim_{\delta \to 0} \omega_T(\delta) = \sup_{t \in [0, T]} (a(t) - a(t-)) .$$

PROOF. – First note that

$$\lim_{\delta \to 0} \omega_T(\delta) \geqslant \sup_{t \in [0, T]} (a(t) - a(t-)) .$$

Suppose that the strictly inequality holds.

Then there exist $\varepsilon > 0$, an infinitesimal sequence $\{\delta_k\}$, and two sequences $\{u_k\}$, $\{v_k\}$ such that $0 < v_k - u_k \leqslant \delta_k$ and

$$a(v_k) - a(u_k) > \sup_{t \in [0, T]} (a(t) - a(t-)) + \varepsilon . \tag{16}$$

Without loss of generality, we can assume $\{u_k\}$, $\{v_k\}$ bounded and s.t.

$$\lim_{k \to \infty} u_k = \lim_{k \to \infty} v_k = u .$$

Then Eq. (16) implies that u is a discontinuity point and that there exist subsequences such that $u_k \to u-$, $v_k \to u+$. Then

$$\lim_{k \to \infty} (a(v_k) - a(u_k)) = a(u) - a(u-) \geqslant \sup_{t \in [0, T]} (a(t) - a(t-)) + \varepsilon$$

which is a contradiction. ■

(2) $\forall t \in [0, +\infty)$, $A_t = \alpha[0, t]$ is a stochastic process with non decreasing cadlag trajectories.

PROOF. – Let $\{f_n\}$ be a sequence of bounded continuous functions on $[0, +\infty)$ s.t. $f_n \uparrow \mathbb{I}_{[0, t]}$. then

$$A_t = \lim_{k \to +\infty} \int_{\mathbb{R}} f_n(x)\, \alpha(dx) . \qquad ■$$

(3) $\forall H > 0$, $\tau = \inf\{t \geqslant 0 \text{ s.t. } A_t - A_{t-} > H\}$ is a random variable.

PROOF. – $\forall t \geqslant 0$,

$$\{\tau > t\} = \{\forall s \in [0, t], A_s - A_{s-} \leqslant H\} = \left\{ \sup_{s \in [0, t]} (A_s - A_{s-} \leqslant H \right\} =$$

$$\left\{ \lim_{\delta \to 0} \omega_t(\delta) \leqslant H \right\} . \qquad ■$$

Finally the sequence Y can be defined recursively as follows

$$Y_1 = \inf\left\{t \geqslant 0 \text{ s.t. } \frac{1}{2} \leqslant A_t - A_{t-} \leqslant 1\right\}$$

and, for $n > 1$, $k = 2^{n-1} + 1, \ldots, 2^n - 1$

$$Y_{2^{n-1}} = \inf\left\{t \geqslant 0 \text{ s.t. } \frac{1}{2^n} \leqslant A_t - A_{t-} \leqslant \frac{1}{2^{n-1}}\right\}$$

$$Y_k = \inf\left\{t \geqslant Y_{k-1} \text{ s.t. } \frac{1}{2^n} \leqslant A_t - A_{t-} \leqslant \frac{1}{2^{n-1}}\right\}$$

while, as far as the sequence X is concerned we set

$$X_k = \alpha(\{Y_k\})\, \mathbb{I}_{\{Y_k < +\infty\}}. \qquad \blacksquare$$

Finally, let us observe that, when α has at most d atoms a.s., $X_k = \neq 0$ for at most d indexes.

REFERENCES

[1] D. J. ALDOUS, *Exchangeability and related topics*, Lect. Notes in Math., Springer (1983).
[2] P. BILLINGSLEY, *Probability and measures*, J. Wiley (1979).
[3] A. GERARDI - F. SPIZZICHINO - B. TORTI, *Exchangeable mixture models for lifetimes: the role of «occupation numbers»*, Stat. Probab. Letters, **49** (2000).
[4] A. GERARDI - F. SPIZZICHINO - B. TORTI, *Filtering equations for the conditional law of residual lifetimes from a heterogeneous population*, J. Appl. Probab., **37**, 3 (2000).
[5] J. F. C. KINGMAN, *The representation of partitions structures*, J. London Math. Soc., **18** (1978).
[6] J. F. C. KINGMAN, *Exchangeability and evolution of the large populations*, in *Exchangeability in Probability and Statistics*, G. Koch and F. Spizzichino eds. (1982).
[7] J. F. C. KINGMAN, *The coalescent*, Stoch. Proc. Appl., **13** (1982).
[8] F. SPIZZICHINO, *A probabilistic model for heterogeneous populations and related burn-in design problems*. In *Reliability and Maintenance of Complex Systems* (S. Ozekici, Ed.), Springer (1996).

Gerardi Anna: Dip. Ingegneria Elettrica, Università de L'Aquila, Italy
Torti Barbara: Dip. di Matematica, Università di Roma «Tor Vergata», Roma, Italy

Pervenuta in Redazione
il 29 marzo 2002