A NOTE ON SOME COMPOUND POISSON DISTRIBUTIONS

CARL PHILIPSON
Stockholm

At the Lundberg Symposium, Stockholm 1968 Jung and Lundberg presented a report on similar problems as those treated in this note, and to the Astin colloquium, Berlin 1968 the present author presented a report with the same title as this note, where some of the results in the first-mentioned report were commented upon. Jung and Lundberg kindly discussed the topic here concerned with the present author some time after the colloquium. On account of this discussion, the present author withdrew his report from the publication in its original form. The following context is a revision and a completion of the author's report to the colloquium.

1. *Basic definitions*

Let τ be a parameter measured on its original, absolute scale, and let $s = \bar{s}(\tau)$ (or $t = \bar{t}(\tau)$) be the same parameter measured on an operational scale with respect to the probability distribution $[\bar{s}(\tau)]^m \exp [-\bar{s}(\tau)] / m!$ (or the corresponding for t). The parameter will often be referred to as "time", which does not imply a restriction of the theory to proper time parameters.

A random function $X(s)$ is said to be distributed in a *cPd i.w.s.* (*compound Poisson distribution in the wide sense*), if the distribution function of $X(s)$ for every fixed parameter point (s, τ) in a finite or infinite domaine of the parametric space as a function of τ can be written in the following general form

$$\sum_{m}^{\infty} \int_{0}^{\infty} e^{-vs} (v\bar{s})^m \, W^{m*}(x, \bar{s}) \, d_v U(v, \tau) \, / \, m!, \; \bar{s} = \bar{s}(\tau). \qquad (1a)$$

where the asterisk power m^*, here and throughout this note, is taken to mean, for $m > 0$, the m times iterated convolution of the distribution function with itself, and, for $m = 0$, unity. $W(x, s)$

$= \int_0^s V(x, u)\, du/s$, $V(x, s)$ being the conditional distribution function of the size of one change in $X(s)$ relative to the hypothesis that the change has occurred at s, here abbreviated to the *change distribution*. $U(v, \tau)$ is a distribution function, called the *structure function*. In the general case $V(x, s)$ and $U(v, \tau)$ may depend on s and τ respectively. If, particularly, these functions are supposed to be independent of the parameter, they will be denoted $V(x)$, $U(v)$ respectively. In the particular case, where $V(x) = \varepsilon(v—c_1)$, c_1 being an arbitrary but fixed constant and $\varepsilon(\xi)$, here, and in the following context, the *unity distribution* equal to zero for negative values, and to unity for non-negative values of ξ, the cPd is said to be *elementary* and, in the opposite case, *non-elementary*. In the elementary case the distribution of $X(s)$ is defined by the integral appearing in (1a) with $x = c_1$, so that $W(x, s) = W(x) = 1$.

In this note, the distribution defined by (1a), in the particular case, where $U(v, \tau) = U(v)$ independently of τ, will of reasons given below, be called a *cPd* :1. If, in addition, $U(v) = \varepsilon(v—\gamma_1)$, γ_1 being an arbitrary but fixed constant, the cPd : 1 reduces to a Poisson probability distribution, in the general case, a *non-elementary*, or, if $V(x) = \varepsilon(x—c_1)$, an *elementary* distribution, defined by the integral with $W(x, s) = W(x) = 1$. If, on the other hand, $U(v, \tau)$, particularly, is in the form $\sum_{m_1=0}^{\infty} \overline{Q}_{m_1}(\tau)\, {}_2U^{m_1*}(v)$, where $Q_{m_1}(\tau)$ is a not specified probability distribution of the variable m_1 assuming only integer values, the distribution defined by (1a) may be called an *aco cPd* : 1 (average of convoluted cPd : 1), (cf. section 6, here below). An aco cPd : 1 can, as proved later in this note, be interpreted as a bunch distribution, as for the elementary case defined by Thyrion ([1]*, p. 68), provided that this definition is extended to include the non-elementary case (which under certain conditions is possible, see section 7 here below), and to allow the number of events within each bunch to depend on a parameter.

*) Numbers within square brackets refer to the list of literature at the end of this note.

If the term *equivalent* is used for two variables, for which the distribution function of one of the variables can be transformed to the distribution function of the other variable, by a simple transformation of the parameter or of the parameter vector, a variable with the *ch.f.* (characteristic function) in the form $\varphi(\eta) = {}_1\hat{\varphi}[{}_2\varphi(\eta)]$, where an angle over the symbol for a ch.f., here and in the following context, is taken to mean the generating function, defined by $\hat{\varphi}(z) = \varphi(-i \log \eta)$, $\varphi(\eta)$ being a ch.f., η a real variable and i the imaginary unit, is said to be *equivalent* to a *generalized variable* with the ch.f. ${}_1\varphi(\eta)$ by a *generalizing variable* with the ch.f. ${}_2\varphi(\eta)$. Consider a (ν—1) times iterated generalization for which the ch.f. can be written in the following general form

$$\varphi_\nu(\eta) = {}_1\hat{\varphi}_1\{{}_2\hat{\varphi}_1[{}_3\hat{\varphi}_1(\ldots {}_{\nu-1}\hat{\varphi}_1[{}_\nu\varphi_1(\eta)]])]\} \tag{1b}$$

A variable with such a ch.f. is by Thyrion [1], said to be distributed in bunches ($\nu = 2$), or in bunches of bunches ($\nu > 2$), provided that ${}_j\varphi_1(\eta)$ are ch.f. of variables assuming only integer values. If assuming the conditions referred to in section 7 are fulfilled, ${}_\nu\varphi_1(\eta)$ may be allowed to be a ch.f. of a continuous variable, and, if the functions ${}_j\varphi(\eta, s_j)$ are substituted for ${}_j\varphi(\eta)$, the random function $X(\mathcal{S}_\nu)$, $\mathcal{S}_\nu = (s_1, s_2 \ldots s_\nu)$ will, in this note, be said to be distributed with a *bunch distribution of order* ν, if it is equivalent with the variable defined by $\varphi_\nu(\eta, \mathcal{S}_\nu)$ in this extended form of (1b). Thyrion has proved, that, for $\nu > 2$, $\varphi_\nu(\eta)$ can also be written in the form ${}_1\varphi_1[\varphi_{\nu-1}(\eta)]$, where $\varphi_{\nu-1}(\eta)$ is in the form of (1b). It might be remarked, that $\varphi_\nu(\eta)$, $\nu > 2$ can also be written ${}_1\psi_{\nu-1}[{}_\nu\varphi_1(\eta)]$ with ${}_1\psi_{\nu-1}(\eta)$ in the form of $\varphi_{\nu-1}(\eta)$ according to (1b).

The ch.f. of an aco cPd : 1 can be transformed in the form ${}_1\hat{\varphi}_1[{}_2\varphi_1(\eta, s_2), s_1]$, s_1, s_2 being parameters on some operational scales (in the non-elementary case under the conditions referred to in section 7). Thus, an aco cPd : 1 can be interpreted as a bunch distribution (cf. section 6 here below). This bunch distribution is of the order 2, if $Q_{m_1}(s_1)$ is not a bunch distribution, in the opposite case the aco cPd : 1 is of an order greater than 2. Thus, an aco cPd : 1 is a particular case both of (1a) and, subject to the conditions mentioned in section 7, of (1b).

The class *cPd* : ν (cPd of the order ν) shall for $\nu = 1,2 \ldots$ be defined, for $\nu \geq 2$, as a bunch distribution of the order ν, where particularly, ${}_j\varphi_1(\eta, s_j)$ defines a cPd : 1 for each value $j = 1, 2 \ldots \nu$, and, for $\nu = 1$, as a distribution in the form of (1a) with $U(v, \tau) = U(v)$ independently of τ; in fact, the ch.f. in the form defining cPd : ν for $\nu \geq 2$, reduces to this form for $\nu = 1$.

If, and only if, the joint probabilities of $X(\mathfrak{s}_\nu)$ on disjoint domaines of the parametric space are well-defined in the sense of stochastic process theory, $X(\mathfrak{s}_\nu)$ constitutes a stochastic process. For such a process, the following terms shall be used *cPp i.w.s.*, *aco cPp* : 1 and *cPp* : ν, if the distribution of $X(\mathfrak{s}_\nu)$ is defined by a cPd i.w.s., aco cPd : 1 and cPd : ν respectively.

2. *Remarks to the terminology*

The terminology in the field treated in this note is rather confusing.

The author's own terminology has previously not accounted for a differentiation between a random function, which for each value of the parameter, eventually restricted to a certain domaine of the parametric space, has an absolute distribution in one of the forms defined in the previous section, and a random function, which, in addition, fulfils the conditions for constituting a stochastic process. This is due to the belief, that in cases, where the existence of well-defined joint probabilities for random functions with given such distributions has not been established, it would later be possible to establish the conditions for the existence of such probabilities, wide enough to cover all actual applications to phenomena, for which a priori process models seem to be rational. Later in this note, such conditions for some of the distributions considered will be dealt with.

As, however, the recent study by Jung and Lundberg (in the report mentioned in the introduction here above), which will be commented upon, later in this context, has given very negative results with respect to the application of process models of this type, it seems necessary to restrict the previous notation cPp to cases, where the conditions for the existence of well-defined joint probabilities either can be postulated, or deducted from other assumptions. In all other cases the term cPd shall be used. — Further, the terms cPp i.n.s. (in the narrow sense) and stationary or non-stationary cPp, previously introduced by the author for a cPp : 1, and a cPp : 2 respectively, have, so far, been used by other authors in quotations. They may, therefore, be replaced by cPd : 1, cPp : 1 and cPd : 2, cPp : 2 respectively without risk for confusion.

The translation of the word "compound", as used here above, as distinct from the word "composed" used by Rényi et al [2], for the *composed Poisson processes*, which are principally different from cPp, into French and German, implies certain difficulties. In French the word "composé" has been used both for "compound" and "composed"; in German a cPp has often been called ein zusammengesetzter Poisson Prozess; to the knowledge of the present author composed Poisson processes have not been treated by German authors. It seems also difficult to find another German translation for "composed" than zusammengesetzt. In French the word "compound" exists, but is very seldom used, e.g. in une compound machine à vapeur. The word "mixed", sometimes used for a cPp, is, however, easily translated into French and German. It remains, however, to find translations of "composed"; for this word it would not be advisable to use "composé" and "zusammengesetzt", which would lead to misunderstanding of earlier works, where these terms have been used for "compound". Also the term, a weighted Poisson process, in German translated into ein gewichteter Poisson Prozess, has — though more seldom — been used for a cPp. A translation of this term into French could, eventually, be un processus de Poisson pondéré. Perhaps the words "mixed" for "compound" and "weighted" for "composed" could be accepted.

For a non-elementary process the confusion seems to be still greater. A non-elementary Poisson process $[U(v) = \varepsilon(v - \gamma_1)]$ has, thus, been called a compound Poisson process (Ge. ein zusammengesetzter Poisson Prozess), which is the same term as that commonly used for a general cPp independently of being elementary or non-elementary. This is apt to lead to serious misunderstandings (in fact, the definitions for a "compound" Poisson process in the first and second edition of Feller's well-known textbook Part I seem not to be consistent). Therefore, the use of "compound" for the designation of a non-elementary Poisson process, must be avoided, which applies also to the German translation. The random function distributed in a cPd is equivalent to the number of changes generalized by the size of a change. Therefore, the term "generalized", in German "verallgemeinert", is often used for "non-elementary", the corresponding word "généralisé" in French, is — as far as is known by the author, — always used for "non-elementary". However, even in certain elementary distributions, e.g. a cPd : 2, the variable is often equivalent to a generalized variable, thus, the term "generalized' is a wider concept than "non-elementary". This last-mentioned term, quoted from Lundberg [3], affords a distinction between the case, where the generalizing variable is the size of a change, from other cases of generalizing variables possible. The use of the terms "elementary" and "non-elementary" is facilitated by the possibility of a direct translation into French and German.

3. *Some remarks on the Poisson process*

In the report by Jung and Lundberg [4], which was referred to in the introduction of this note, the conditions for the constitution of processes by random functions, with given absolute distribution functions in one of the forms defined in section 1, were discussed. In fact, these conditions, as formulated in [4], seem to be very restrictive, and it is, further, said in [4], that, if these conditions are not satisfied, the random functions may not constitute stochastic processes, or, if such processes are constituted, they will not be sufficiently specified, and may include processes with less realistic properties. In this section the results with respect to Poisson-distributed random functions will be commented upon, and, in later sections, other results in the quoted paper will be discussed.

In [4] it is said, that a random function known to be distributed with a Poisson probability distribution (elementary or non-elementary) constitutes a Poisson process, if it is homogeneous in τ, and the intensity, which in this case is constant, may be either an arbitrary, but fixed constant or an estimate from a random experiment. If, however, the random function, $X(\tau)$ say, is heterogeneous in τ, with the intensity $\tilde{\varkappa}w(\tau)$ say, where $\tilde{\varkappa}$ is constant, the random function $X(t)$ obtained by the transformation of τ in $X(\tau)$ according to the relation $\tau = \bar{t}(\tau) = \tilde{\varkappa}\int_0^t w(u)du$, is said to constitute a Poisson proces if and only if $w(\tau)$ is a known, integrable function of τ.

It is said that "As soon as $w(\tau)$ is unknown, or contains a random element, the model of a Poisson process is no longer applicable".

If, on the other hand, the usual basic conditions for a Poisson process, except the homogenity in time, a) homogenity in space, that is any finite number of (t_j, y_i) being the time point and the size of the ith change are independent random variables, and that any t_i will have the probability density e^{-t} for $t > 0$, and b) rarity of multiple events, are satisfied, the random function concerned constitutes a Poisson process (according to Bühlmann [5] it should even not be necessary to include b) in the basic conditions, as according to him, b) is a direct consequence of the homogeneity in space and in transformed time, and of the simple properties of the sample functions in the restricted space; an assertion for which a proof will later be published). It seems, therefore, admissible to use an estimate of the intensity $w(\tau)$, or of the mean number of events for certain periods of absolute time τ, as a function of τ, for the definition of a Poisson process. Should this assertion not be true, many results reached in numerical investigations of the theory of the insurance risk were to be rejected. The mapping of the parametric space, defined by $t = \bar{t}(\tau)$, implies that $\bar{t}$ is a never decreasing function of τ, as, even in random experiments, $\bar{t}'(\tau)$ is essentially positive, with one-to-one correspondence, in both directions, so that the estimating problem is reduced to the estimation of a function with very simple properties. According to the present author's opinion, a denial of the admissibility of using such estimates calls in question the general principles, which are the basis for the application of estimated statistics depending on a parameter, to several problems in a wide field of statistic research. The reader will also be referred to Cramér's well-known survey of the risk theory [6], where the Poisson process was rigoroulsy deducted, after the transformation of the parameter, from the basic assumptions. The author has not found any statement in Cramér's survey to the effect, that the theory developed by him should not be applicable, if the parameter t has to be estimated. In fact, numerical examples on the expansion of the distribution functions (one of these examples is applied in a calculation of the ruin probability) with the application of chosen values of t have been given.

It is a common objection against the estimations of quantities dependent on time, that the applications of the estimates for prediction must be based on very uncertain extrapolations of the estimated trend. It is very easy to construct examples of a sequence of numbers of events for several years, in situations, where it seems a priori quite irrational to use a process model, and, where the sequence may be described as a sequence of results of a Poisson process during the observation period. In such examples the extrapolation to future time of the observed trend seems to be meaningless. But for phenomena of a type for which a process model is likely to be applicable, the state of things is different. It is, then, necessary to apply the commonly used precautions for the estimation of a time trend. Such a precaution is an analysis of the time variation of all circumstances of influence on the statistic to be estimated and extrapolated. As an example, for which such methods are used, is the prediction of risk premiums for the establishment of new rates for non-life insurance, e.g. motor insurance, where the risk premium depends on a great number of circumstances variable in time, and which, in addition, is the result of a more complicated process than the Poisson process. In this connection statistical methods of the types multiple regression analysis, and the so called factor analysis are used. The factor analysis implies a graduation of a function of time and other factors of influence on the risk in a form, which makes it possible to separate the effects of the different arguments on the risk (cf. e.g. Philipson [7]). As an example of this kind, also the development of the market price for a commodity may be mentioned. The estimation of the time trend of the number of events, particularly, in situations, where the increase of the number during disjoint intervals may be considered independent variables, is often much simpler than the prediction of future risk premiums. This is i.a. due to the fact, that the number of factors of influence on the number of events is often materially less than the number of factors influencing the risk premium. Also the ordinary statistical tests, e.g. the χ^2-test, is often simpler for the number of events. Methods for the differentiation between the Poisson model, and some standard models, based on heterogeneity in space, have also been devised. In Swedish Motor Insurance it has been possible to give safe numerical evidence for the necessity of using models of the last-mentioned type. In an investigation of this insurance, for which the observation period related to the calendar years from 1947 to 1954, both years inclusive, the time trend both of the claim frequency, and of the risk premium was estimated for a very fine classification of the risk factors by ihc methods indicated here above, The rating table introduced. in 1955 was based on an extrapolation of these statistics. These were later compared with the actual experience, and the extrapolated risk premiums were found to agree satisfactorily with the results of this experience for three and four years after 1954; for the claim frequency this was valid for a greater number of years.

In cases, where it can be proved with sufficient precision that the number of events in disjoint intervals are mutually independent, and/or that the number is Poisson-distributed, and that a process model is likely to be applicable, by the estimation of the time trend in the expected number of events, the Poisson process must — due to all experience — be applicable, even for a prediction provided that the estimation, and the extrapolation, is made by using appropriate methods. This principle shall in this note be called *proposition G*. An extension of this principle to cases, where the number of changes in disjoint intervals are dependent, and where the absolute probabilities are of any of the types defined in section 1, the Poisson distribution exclusive, seems also to be possible, as judged from the experience of motor insurance referred to in the previous paragraph. In fact, the proposition has been applied in several numerical investigations, e.g. in [3].

4. *The developments with respect to a cPd*:1 *in* [3] *and* [4]

In his deduction of a Polya process by using the limit of a Polya-Eggenberger distribution, Lundberg [3] introduces the notation $p_n(t)\Delta_h(t)$ "for the conditional probability of an event" in the interval $(t, t+\Delta_h(t))$ "when n events have taken place up to t". This must, tacitly, imply a postulate of the existence of such a probability. Further, the expression obtained for $p_n(t)$ is continuous in t (p. 17, l.c.). In the general theory, the starting point is a process, so that the existence of the conditional probabilities is postulated, which applies also to the continuity of $p_n(t)$ with respect to t (3), p. 27, l.c.), and to the equations (29), (30) and (30*) (l.c.), which are deducted on these assumptions. In chapter 4 (l.c.) the condition 4) (p. 27, l.c.) is disregarded, so that $p_n(t)$ is no longer necessarily uniformly bounded for all n. It is, then, proved that the forward and backward differential equations have a unique solution, provided that $p_n(t)$ is a non-negative, continuous function of t for every fixed n (or at least integrable over a finite interval of t), and that this solution satisfies all the fundamental conditions of a generalized Markov process under the additional assumption that $[\text{Max } p_n(u)]^{-1}$ from the result m in the parameter point s, which defines the starting point of the forward differential equation, forms a divergent sequence in n for every finite interval $0 \leq u < T$. Generalized is here taken to mean a process, for which the fundamental conditions are disregarded for intervals starting in (m, s) with zero probability. A cPp : 1, first mentioned in [3], p. 20 and described by the words, that "these processes are characterized by the" (absolute) "probability for n events up to t for each real value $t > 0$ being equal to a Laplace-Stieltjes integral" in the form of the mth term in (1a) with $V(x) = \varepsilon(x—c_1)$, $\bar{U}(v, \tau) = U(v)$ independently of τ. A cPp : 1 is later (p. 76, l.c.) defined by the words "if and only if the intensity function of an elementary process satisfies the recurrence formula", $p_{n+1}(t) = p_n(t) — p'_n(t)/p_n(t)$, "it will define a compound Poisson process", cPp : 1 After the definition of the absolute probability of a cPp : 1 (p. 71, l.c.), Lundberg seeks "an approximate expression of the probability of an event occurring during a short interval $(t, t+\Delta t)$ under the condition that n events occurred during the preceding interval $(0, t)$", interpreting t as a time parameter. This probability is, then, denoted $p_n(t)\,\Delta t$, where $p_n(t)$ is said to express the "*intensity*" The existence of such conditional probabilities is, therefore, tacitly understood in the words quoted (pp. 20, 71, l.c.). The expression obtained for $p_n(t)$ is continuous in t for fixed n, and leads to the recurrence formula quoted here above. Further, the definition (p. 72, l.c.) contains the following sentence "if an elementary random process exists with an absolute probability function" in the form of the Laplace-Stieltjes integral defined here above "the process will be called a compound Poisson process . . ." (cPp : 1), which, thus, contains an explicit statement of the postulate tacitly understood in the quoted parts (pp. 20, 71, l.c.). In Theorem 6

(p. 73, l.c.) it is stated that: "Given a function $P(t)$ with the properties 1) $P(t)$ completely monotonic for $t > 0$, 2) $\lim_{t \to 0} P(t) = 1$, the function $= p_n(t)$", $= P^{(n+1)}(t) \,/\, P^{(n)}(t)$, "constitutes for $n = 0, 1 \ldots$ and for $t \geq 0$ an intensity function of an elementary process in the generalized sense. The process defined by this intensity function is a compound Poisson process . . ." (cPp : 1). It is, then, proved that the solutions of the forward differential equations defined by $p_n(t)$, fulfil all the fundamental conditions of a Markov process, and that the absolute probabilities have the form of the Laplace-Stieltjes integral referred to above, so that the process is a cPp : 1.

The only possible interpretation of Theorem 6 (p. 73, l.c.) is that, if $P(t)$ fulfils the conditions given in the theorem, there exists an intensity function deduced from $P(t)$, which defines a cPp : 1 constituted by a random function $Y(t)$ say. It does not follow, however, that every random function $X(t)$, for which the probability for non-occurrence in the interval $(0, t)$ is given by $P(t)$ satisfying the conditions of the theorem, and, consequently, the probability distribution of the number n of changes in the interval $(0, t)$ is given in the form of the Laplace-Stieltjes integral referred to above in terms of $P(t)$, necessarily constitutes a random process. If and only if the joint probabilities of $X(t)$ are well-defined in the sense of stochastic process theory, $X(t)$ is identical with $Y(t)$, and constitutes a cPp : 1 defined by $p_n(t)$. This is consistent with the previously quoted parts from Lundberg's book and with the last paragraph on p. 84 (l.c.).

In [4], Theorem 6 in [3] has been quoted by saying, that a completely monotone function for $t > 0$, $P(t)$, with the limit for $t \to 0$ equal to unity "may always define a compound Poisson process" (cPp : 1). By the discussion above, $p_n(t)$ deduced from $P(t)$ always defines a cPp : 1, even if a random function, assumed to be distributed with a cPd : 1, deduced from a function with the assumed properties, does not constitute a stochastic process, unless the joint probabilities are well-defined; but, if the last condition is fulfilled, the random function, thus defined, constitutes a cPp : 1. The following words in ([4], p. 8), that every process with absolute probabilities in the form of a cPd : 1, does not necessarily be a cPp : 1, are — in the author's opinion — not consistent with the developments in [3] (cf. p. 72, l.c.), as, if and only if a random function distributed with a cPd : 1 constitutes a random process, this process must be a cPp : 1. The degenerate counter example in [4] for a random function with a given cPd : 1, which does not constitute a cPp : 1, does not — as far as the present author can see — constitute a stochastic process whatsoever, at least not a Markov process. Even if the example were to lead to a stochastic process, the example is so unrealistic, that the conditions of the example are not likely to be satisfied for any phenomenon, for which one chooses to apply process models.

It is in [4], further, asserted, that, if the risk intensity can be written $\tilde{\varkappa}\, w(\tau)$, where $\tilde{\varkappa}$ may be estimated by a random experiment, and $w(\tau)$ may contain a random element, and, if $\bar{N}(\tau)$ is distributed in a cPd : 1, $\bar{N}(\tau)$ is, in general, neither a cPp : 1, nor a well-defined random process, if the transition probabilities are not defined. According to the interpretation given above of Theorem 6, [3], it is always possible to deduct an intensity function $p_n(t)$, which defines a cPp : 1, the random function $N(t)$, obtained by the transformation $t = \bar{t}\,(\tau)$, does, however, constitute the said cPp : 1, if and only if the transition probabilities of $N(t)$ are well-defined, in which case they are the solutions of the differential equations in terms of $p_n(t)$ derived from the given cPd : 1. By this interpretation $N(t)$ cannot constitute a stochastic process, which is *not* a cPp : 1. As far as an extension of these statements to the case, where $\bar{t}(\tau)$ or $\tilde{\varkappa}\; w(\tau)$ are estimated from random experiments, is concerned, the reader is referred to what has been said in the previous section (proposition G). It is, further, referred to Chapter VII of [3] (particularly p. 137), where t has been estimated from the experience, and used in two process models for comparison with the reality. Thus, proposition G seems to have been tacitly understood in this chapter (VII, l.c.).

In [4] it has been remarked, that a general stochastic process is not defined by the absolute probabilities in each parameter point. A birth process is completely defined by the conditional probabilities, or by the intensity function. If, in addition, the recurrence relation for $p_n(t)$ of a cPp : 1 holds ([3], p. 68, [4] p. 7, quoted here above), the process is a cPp : 1. The definition of $p_n(t)$ in Theorem 6 [3] leads to this recurrence relation. Further, in most cases, where it is deemed appropriate to apply stochastic process models, it must be natural to postulate the existence of well-defined joint probabilities. In cases, where it cannot be assumed, that the phenomenon is homogeneous in space, the simplest basic assumption, seems to be the assumption of exchangeability of the sum total of the changes on each interval of a sequence of an arbitrary number of intervals of equal length, without change in the probability function of the process. According to Bühlmann [5], a cPp : 1 can be deducted from such an assumption, which seems to be a natural postulate for phenomena for which a cPd : 1 has been found to hold, and, where it is deemed rational to apply stochastic process models.

For purposes of a further elucidation of the comments in this section, and for obtaining direct proofs of the assertions made, the author will in the next section give two therorems which are based on modern process theory and which contain assertions intended to give a clear picture of the author's interpretation of Theorem 6, [3].

5. *A short review of the theory of translator operators and two theorems on the problems discussed in the previous section*

The following review of the definition and certain properties of translator operators has been drawn out from a book by Dynkin, [8]. The reader is for a full description referred to [8], I ch. 3.

In I 3.1 (l.c.) a *Markov process* is defined. The following items are given: a) a function $\zeta(\omega)$ in some space Ω taking values in $[0, \infty]$ (ζ may also be an arbitrary fixed finite constant or equal to infinity), b) a function $x_t(\omega)$ for $\omega \in \Omega$, $t \in [0, \zeta(\omega)]$, taking values in the state space $(E, \mathbf{B})$, c) a σ-algebra $\mathbf{M}_t$ defined on $\Omega_t = \{\omega : \zeta(\omega) > t\}$, for each $t \geq 0$, and d) a function $P_x(A)$ for each $x \in E$, on some σ-algebra $\mathbf{M}^0$ on Ω, which contains $\mathbf{M}_t$ for all $t \geq 0$. Then, if and only if certain conditions 3.1A — 3.1G are satisfied (if 3.1G is not satisfied, this can be achieved by a suitable elargement of Ω), the quadruple $(x_t, \zeta, \mathbf{M}_t, P_x)$ defines a Markov process ([8], I 3.1). Let $\psi(t)$ be a function taking values on the interval $0 \leq t < \lambda$ in E, and the *shift by the amount* t of $\psi(t)$, c_t, be defined by $c_t\, \psi(u) = \psi(t+u)$ $(0 \leq u < \lambda - t)$. The condition 3.1G requires that the set of trajectories of the process is invariant under all shifts ([8], I 3.1).

Let ν be the mapping of Ω_0 on $\{\psi(t)\}$, the family of $\psi(t)$ for all intervals $(0 \leq t < \lambda)$; this mapping associates each $\omega \in \Omega$ to a trajectory $x_u(\omega)$, $[0 \leq u < \zeta(\omega)]$. Let the *translator operators* $\varkappa_t$ be defined by the following relation

$$\varkappa_t B = \nu^{-1} c_t^{-1} \nu B \text{ for every } B \in \mathbf{N}^*, \tag{2a}$$

where $\mathbf{N}^*$ denotes the minimal system of subsets in Ω_0, that contains all the sets $\{\omega : x_t(\omega) \in \Gamma\}$ $(t \geq 0,\ \Gamma \subseteq E)$, and is closed under the union and the intersection of any number of sets, and under the operation of taking complements. The translator operators fulfil, i.a., the following relations

$$\varkappa_t \{x_h \in \Gamma\} = \{x_{t+h} \in \Gamma\} \text{ for any } t \geq 0,\ h \geq 0,\ \Gamma \subseteq E \tag{2b}$$

([8] I 3.5). If $\varkappa_t$ with these properties exists, 3.1 G is fulfilled.

Let $\overline{\mathbf{M}}_t$ be a completion of $\mathbf{M}_t$ with respect to P_x [by including all sets Γ such that $\Gamma_1 \subseteq \Gamma \subseteq \Gamma_2$, $P_x(\Gamma_1) = P_x(\Gamma_2)$], $\mathbf{N}$ the σ-algebra on Ω_0 generated by the sets $\{x_u \in \Gamma\}$ $(u \geq 0,\ \Gamma \in \mathbf{B})$, and $\overline{\mathbf{N}}$ the intersection of $\mathbf{N}^*$ defined under (2a) with the completion of $\mathbf{N}$, with respect to P_μ corresponding to all initial distributions μ. Then, the following assertion holds

$$P_x(A \varkappa_t B) = \int_A P_{x_t}(B)\ P_x(d\omega) \text{ for } A \in \overline{\mathbf{M}}_t,\ B \in \overline{\mathbf{N}} \tag{2c}$$

([8], I 3.6). The quadruple $(x_t, \zeta, \overline{\mathbf{M}}_t, P_x)$, where the measures P_x have been extended to the σ-algebra $\overline{\mathbf{M}}_0$, defines a Markov process, as, in fact, the existence of $\varkappa_t$ with the properties given in I 3.5 — 3.6 (l.c.) implies, that the conditions 3.1A — 3.1G, referred to in the definition of a Markov process, are satisfied. ([8] I 3.6, Corollary 2).

If for a random function $X(t)$, $\varkappa_t$ with these properties exist and the σ-algebras can be defined in such a way that (2c) holds, $X(t)$ shall in the following context be said to be *admissible*, and, in the opposite case, *non-admissible*.

Theorem 1

Let a random function $N(t)$, which assumes only integer values, be admissible or non-admissible in the sense of the definition just given, and for every fixed value of a continuous real parameter $t \in [0,\zeta]$, where ζ may be finite or infinite, be distributed with a given distribution $P_n(t)$ in the following form, where the structure function $U(v)$ is a distribution function independent of t.

$P_n(t) = \int\limits_0^\infty e^{-vt}(vt)^n \, dU(v)/n!$, and the functions $\pi_n(t)$, $g_{n,\nu}(t, \Delta t)$ be defined by the following relations.

$$\pi_n(t) = -P_0^{(n+1)}(t)/P_0^{(n)}(t), \text{ where } P_0^{(n)}(t) = \partial^n P_0(t)/\partial t^n$$

$$g_{n,\nu}(t, \Delta t) = \begin{cases} 1 - \pi_n(t)\,\Delta t + o(\Delta t) \text{ for } \nu = 0 \\ \pi_n(t)\,\Delta t + o(\Delta t) \text{ for } \nu = 1 \\ o(\Delta t) \text{ for } \nu > 1 \end{cases}$$

and for sufficiently small values of Δt.

Then,

1) $P_0(t)$ is a completely monotone function of t for $t > 0$, equal to 1 for $t = 0$.

2) $\pi_n(t)$ is for fixed n a positive, continuous function of $t < \zeta$ for every fixed n.

3) $P_n(t+\Delta t) = g_{n,0}(t, \Delta t) \; P_n(t) + g_{n-1,1} \; (t, \Delta t) \; P_{n-1}(t) + o(\Delta t)$ for every $t < \zeta$, and for all n. (3)

4) The forward differential equation, obtained by the limit passage of (3) for $\Delta t \to dt$, substituting the conditional probabilities for $P_n(t)$, and the corresponding backward differential equation have a unique solution, which fulfils the fundamental conditions of a generalized Markov process, being a cPp : 1, defined by $\pi_n(t)$ and the termination point ζ.

5) $N(t)$ may or may not constitute a process, as defined in 4).

Proof. By assumption, $P_n(t)$ is defined for any t in the interval $0 \leqslant t < \zeta$, thus, $P_n(t + \Delta t)$ is, for $0 \leqslant t + \Delta t < \zeta$, a well-defined probability.

By the insertion of $t + \Delta t$ for t in $P_n(t)$, and by using an asymptotic expression for the product of the functions $e^{-v\Delta t}$ and $(t + \Delta t)^n/t^n$, to be deducted here below, an easy calculation leads to (3). The asymptotic expression of the said product is obtained by using the MacLaurin expansions for the functions concerned and by the following calculation $\{1 + n[\Delta t + o(\Delta t)]/t\} \times \{1 - v\Delta t + o(\Delta t)\} \simeq 1 - v\Delta t + o(\Delta t) + n[\Delta t + o(\Delta t)]/t$. The assertion 3) is, thus, proved. 1) is a direct consequence of the well-known theorem given by Bernstein and Feller; 2) is a consequence of the definition of $\pi_n(t)$ and $P_n(t)$. For the demonstration of 4) the reader is referred to the proof of Theorem 6 [3], and to the Bernstein-Feller theorem. 5) is a consequence of $N(t)$ being, not necessarily, an admissible random function.

THEOREM 2

Let the random function $N(t)$ be defined as in Theorem 1, and, assume, particularly, that $N(t)$ is admissible, as defined before the Theorem 1.

Then, $N(t)$ constitutes a cPp : 1 defined by the intensity function $p_n(t) = \pi_n(t)$, as defined in Theorem 1, and by the termination point ζ. The forward differential equation of this process is in the form of the limit, when Δt tends to dt, of the difference equation (3), with the substitution of $p_n(t)$ and the conditional probabilities for $\pi_n(t)$ and $P_n(t)$, respectively.

Proof. Let the conditional probability for an increase of ν units in $N(t)$ on the interval $(t, t + \Delta t)$ relative to the hypothesis that n events have occurred on the interval $(0, t)$ be designated by $f_{n,\nu}(t, \Delta t)$. By assumption, these probabilities are well-defined in the sense of stochastic process theory. By using (2c), which, by assumption is applicable, in this case, the following relation is obtained

$$P_n(t + \Delta t) = \sum_{\nu=0}^{n} f_{n-\nu,\nu}(t, \Delta t)\ P_{n-\nu}(t),\ \zeta > t + \Delta t \tag{4}$$

This may also be derived from the following arguments. Let $B_n(t)$ be the set of elements for which $N(t) = n$ and $A_{n-\nu,\nu}(t, \Delta t)$ the intersection of $B_{n-\nu}(t)$ and $B_n(t + \Delta t)$. Then for $\zeta > t + \Delta t$

$$B_n(t + \Delta t) = \begin{cases} A_{00}(t, \Delta t) \\ \sum_{\nu=0}^{n} A_{n-\nu,\nu}(t, \Delta t) \end{cases}$$

The probability for fixed values of n and of $t < \zeta - \Delta t$ of the set in the left membrum of this relation is given by the left membrum of (4), and the probabilities of the sets in the right membra for $n = 0$, $n > 0$ respectively given by the right membrum of (4); thus, (4) holds.

By assumption, (3) of Theorem 1 also holds. For each value of n and $t + \Delta t$ the left membra of (3) and (4) are equal, which, thus, applies also to the right membra, for all values of n, and of $t < \zeta - \Delta t$, so that $f_{n,\nu}(t, \Delta t) = g_{n,\nu}(t, \Delta t)$ for all $\nu = 0, 1 \ldots n$, for all n and for every $t < \zeta - \Delta t$.

Thus, there exist asymptotic expressions for $f_{n,\nu}(t, \Delta t)$ in the form given for $g_{n,\nu}(t, \Delta t)$ in Theorem 1, with the substitution of $p_n(t)$ for $\pi_n(t)$, and the asymptotic expression for $P_n(t + \Delta t)$ is in the form given in (3) with the same substitution. Therefore, the assertions 4) of Theorem 1 hold for $N(t)$, subject to the assumptions of Theorem 2. Thus, $N(t)$ constitutes a cPp : 1, defined by the intensity function $p_n(t) = \pi_n(t)$ and the termination point ζ, with the differential equation defined in the assertion, which, thus, has been proved.

Remark: Consider the distribution function $\sum\limits_{n=0}^{\infty} P_n(t)\ V^{n*}(x)$, where $P_n(t)$ is defined in the Theorem 1. It is easily seen, that the Theorems 1 and 2, can be extended to the case, where $X(t)$, distributed with a given distribution function in this form, is substituted for $N(t)$. The extension of a cPp : 1 with absolute distribution functions, thus defined to the case, where the change distribution is dependent on t, $V(x, t)$ say, is a particular consequence of a theorem given by Jung [9], the absolute distribution function of $X(t)$ is, then, given in the form of (1a), with $U(v, \tau) = U(v)$ independently of τ. Independently of Jung, this result was obtained by the present author [10], later extended to the general case in (1a). In [10] the starting point was the forward differential equation of a cPp : 1 with the change distribution $V(x, t)$, assuming, that there exists a modification of $V(x, t)$, $\tilde{W}(x,t)$ say, such that the distribution function of $X(t)$ is in the form $\sum\limits_{n\ 0}^{\infty} P_n(t)\ \tilde{W}^{n*}(x, t)$ for every fixed t. An easy deduction leads to a differential equation in the form $t\tilde{z}'(t) = z_t - \tilde{z}_t$; $\tilde{z}_t$, z_t being ch.f. corresponding to $\tilde{W}(x,t)$ and $V(x, t)$ respectively; a solution of this equation leads, after conversion, under mild regularity conditions to $\tilde{W}(x, t) = W(x, t)$ as defined under (1a) (cf. Cramér, [6], 6.1). If in Theorem 1, $X(t)$ with the change distribution $V(x, t)$ is substituted for $N(t)$, (3) will be obtained in the form:

$$P_n(t+\Delta t)\ \tilde{W}(x,t+\Delta t) = g_{n,0}(t,\Delta t)\ P_n(t)\ \tilde{W}^{n*}(x,t) + \\ + g_{n-1,1}(t,\Delta t)\ P_{n-1}(t)\ \tilde{W}^{(n-1)*}(x,t) * V(x,t) + o\,(\Delta t),$$

where $\tilde{W}(x, t)$ is defined as here above. This leads, if quantities of smaller order than the order of Δt are neglected, to the differential equation for $\tilde{z}_t$ just given, which, thus, holds asymptotically for $X(t)$; this gives the following corollary to Theorems 1 and 2.

Corollary. If the random function $X(t)$ is distributed with a non-elementary cPd : 1, defined by $P_n(t)$, $V(x, t)$ and ζ, there exists a random function $Y(t)$, which constitutes a cPp : 1 defined by $\pi_n(t)$, $V(x, t)$ and ζ, with the absolute distribution functions given by $\sum\limits_{n=0}^{\infty} P_n(t)\ W^{n*}(x, t)$, where $W(x, t)$ is defined as under (1a). If and only if $X(t)$ is admissible, as defined before Theorem 1, it constitutes such a process.

6. *About the amalgamation of independent random processes*

In [4] the following statements with respect to the amalgamation of independent elementary cPp : 1 were made. For greater clarity the term cPp : 1 will be added to the term compound Poisson process, as used in [4]. Also the statements in [4] with respect to a random function of the operational time t, were in some cases formulated by using a function of the absolute parameter τ. In the following lines, the assertion with respect to a function of t, shall be accordingly formulated.

In the case, where $\overline{N}_\mu(\tau)$, $\mu = 1,2 \ldots m$ constitute independent, elementary compound Poisson processes (cPp : 1) with the absolute parameter τ, $\overline{N}(\tau) = \sum\limits_{\mu=1}^{m} \overline{N}_\mu(\tau)$ constitutes a compound Poisson process (cPp : 1) with a structure function equal to the convolution of the structure functions of the components. In the case, where the components only after the transformation $t = \bar{t}(\tau)$ are compound Poisson processes (cPp : 1) in operational time t, $N(t) = \sum\limits_{\mu=1}^{m} N_\mu(t)$ is a compound Poisson process (cPp : 1), if and only if $\bar{t}_\mu(\tau) = k_\mu\ \bar{t}(\tau)$, where k_μ are independent of τ for every value of μ, and if $\bar{t}(\tau)$ are known functions; this case has in [4] been called a "very special case". If $t_\mu(\tau)$ are for every value of μ a known function of τ but not proportionate to $\bar{t}(\tau)$, $N(t)$ constitutes a cPp i.w.s.

It will be observed here, that it seems not impossible to extend the proposition G (see section 3 here above) to be applied also to the amalgamated process in cases, where the components are cPp : 1 only after the transformation of the parameters to $\bar{t}_\mu(\tau)$, and, even, where these functions, not necessarily, are proportionate.

An amalgamation of a number of independent cPp : 1 shall later in this note be dealt with as a particular case of more general models, to be defined in the next section.

First an extension of the term "convolution" shall be introduced. Let ξ be a discontinuous or continuous variable, random or non-random. Let $\underset{(\xi)}{\Pi^*}$ be an operator, which, applied to a given set of distribution functions, where the components correspond to the ch.f. $\varphi_\xi(\eta)$, defined for every value of ξ, transforms the set into a distribution function, which corresponds to the ch.f. exp $[\int d_\xi \log \varphi_\xi(\eta)]$ the Stieltjes integral being taken over the range of ξ. If, particularly, the set is finite, or enumerable, the transform reduces to the asterisk product $\prod_{i=0}^{m_1}{}^*$, $m_1 \leq \infty$ and, if, in addition, all the elements of the given set are equal, to an asterisk power. In these cases the operator defines an ordinary convolution, and, if ξ is allowed to form a non-enumerable set, the transform may be called a *convolution in the extended sense*. It is seen, that, in all cases, the transform reduces to unity for $\xi = 0$. In the following context it will, for simplicity, be assumed that ξ forms a finite or enumerable set; the extension to non-enumerable sets is straight forward.

If in the distribution function of an aco cPp : 1, as defined in section 1 here above, $\underset{(\xi)}{\Pi^*}{}_2U_\xi(v)$ is substituted for ${}_2U^{m_1*}(v)$, the distribution function obtained defines a distribution, which may be called an *extended aco cPd* : 1. If ξ is, at least, enumerable, the probability distribution of ξ may be denoted $Q_{m_1}(s_1)$, which is a probability distribution of the discontinuous variable m_1. After the transformation to the operational scales, s_1, s_2, where s_2 has been defined in section 1, the mth term of (1a) can, on the particular assumptions for an elementary, extended aco cPd : 1, be written in one of the following alternative forms, where the second form is obtained from the first form by an, evidently permissible, reversion of the integration and summation, and by using the proporties of the Laplace transform of a convolution.

$$P_{m_2}(s_1, s_2) = \int_0^\infty e^{-vs_2}(vs_2)^{m_2}\, d_v \left[\sum_{m_1=0}^\infty Q_{m_1}(s_1) \prod_{i=0}^{m_1}{}^{*}\, {}_2U_i(v) \right] / m_2! \quad (5a)$$

$$P_{m_2}(s_1, s_2) = \sum_{m_1=0}^\infty Q_{m_1}(s_1) \prod_{i=0}^{m_1} \left[\int_0^\infty e^{-vs_2}(vs_2)^{m_2}\, d_v\, {}_2U_i(v)/m_2! \right] \quad (5b)$$

The ch.f. corresponding to (5b) can be written

$$\sum_{m_1=0}^\infty Q_{m_1}(s_1) \prod_{i=1}^{m_1} {}_2\varphi_i(\eta, s_2),$$

where ${}_2\varphi_i(\eta, s_2)$ corresponds to the ith factor of the asterisk product in (5b). In the particular case, where ${}_2\varphi_i(\eta, s_2) = {}_2\varphi(\eta, s_2)$ independently of i, the distribution is a bunch distribution according to section 1, in the elementary case and, under certain conditions, in the non-elementary case (see section 7 here below). The amalgamation of m_1 elementary cPp : 1 for a fixed value of m_1, has, by (5b) in the general case, the following distribution function

$$\int_0^\infty e^{-vs_2}(vs_2)^{m_2} d_v \left[\prod_{i=0}^{m_1}{}^{*}\, {}_2U_i(v) \right] / m_2! \qquad (5c)$$

In fact, (5c) is consistent with the "very special case" referred to in the quotation from [4], on account of the definition of s_2 in section 1, which tacitly implies that the expected number of events in the ith component can be written $\gamma_{i1}\, s_2$, where γ_{i1} is the mean of ${}_2U_i(v)$ for each value of i. This can also be expressed in terms of the intensities of the components. Supposing that this intensity for ith component can be written ${}_2\varkappa_i\ w_i(\tau)$, if ${}_2\varkappa_i$ is independent of τ and $w_i(\tau)$ of i, the change of the variables of integration leads to modified structure functions ${}_2U_i({}_2\tilde{\varkappa} v/{}_2\varkappa_i) = {}_2\overline{U}_i(v)$ say, for every value of i, so that the convolution is in the form of (5c) with ${}_2\overline{U}_i$ instead of ${}_2U_i$. The extension of the amalgamation to cases, where the operational parameters $s_{2i} = \int_0^{\tau_i} {}_2\varkappa_i w_i(u) du$, implying a dependence on i for both τ_i and $w_i(\tau)$, shall be shortly referred to later in this note.

7. The models of grouping in general

A model called the *simple grouping* shall be defined as follows. Consider a group of a population, called a *main group*, defined as a finite or enumerable set of *sub-groups*. Each sub-group is associated with a random function ${}_iY(s_{2i})$, $i = 1,2 \ldots m_1$, where, in the general case, m_1 is a random variable. ${}_iY(s_{2i})$ is either equal to ${}_iN(s_{2i})$, the elementary case, or to ${}_iX(s_{2i})$, the non-elementary case. ${}_iN(s_{2i})$ has a probability distribution ${}_iR_{m_{2i}}(s_{2i})$, and ${}_iX(s_{2i})$ the change distribution ${}_iV(x, s_{2i})$. Each sub-group is supposed to be *homogeneous*, taken to mean, that it can be arbitrarily divided into any number of minor groups, such that all the minor groups have the same volume, and the same probability distribution of the number of events, and, in the non-elementary case, the same change distribution, for the contributions to ${}_iY(s_{2i})$. It is, further, assumed that these contributions are admissible random functions, so that ${}_iY(s_{2i})$ constitutes for each value of i a Markov process, defined by the intensity function ${}_ir_{m_{2i}}(s_{2i})$, say, and by the termination point ζ_i to be defined below, and, in the non-elementary case, by the change distribution ${}_iV(x, s_{2i})$. The process of a sub-group is called a sub-process. These sub-processes are assumed to be mutually independent, and to have non-decreasing, and right continuous trajectories. The process associated with the main group, the *main process*, is constituted by the random function $\bar{Y}(s_1, \bar{s}_2) = \sum_{m_1=0}^{\infty} Q_{m_1}(s_1) \sum_{i=1}^{m_1} {}_iY(s_{2i})$, where $\bar{s}_2 = \sum_{m_1=0}^{\infty} Q_{m_1}(s_1) \sum_{i=1}^{m_1} s_{2i}$, and the term for $m_1 = 0$ shall be equal to $Q_0(s_1)$. In the particular case, where for all values of i $s_{2i} = \gamma_{2i} \int_0^{\tau} w(u)du$ and neither τ nor $w(u)$ depends on i, the main group shall be said to be *stationary* (if the condition is fulfilled for an interval $0 \leqslant \tau \leqslant T$, the main group is stationary on this interval). If the main group is *non-stationary*, it is assumed that the ith sub-group enters only once into the main group, at the point, ${}_i\tau_0$ say, and leaves the main group at the point ${}_i\tau_1$ say, where ${}_i\tau_0$, ${}_i\tau_1$ may depend on $\omega \in {}_i\Omega$, the reference space of the ith sub-process, which, therefore, must be appropriately enlarged. In the calculation of the contributions to the main process on the interval $0 < \tau < T$ it is, further, assumed that each subprocess with ${}_i\tau_1 \leqslant T$, starts at zero, and terminates at ${}_i\tau_1 - {}_i\tau_0$ corresponding to ζ_i on the operational scale for the ith sub-process. That these assumptions with respect to the sub-processes do not restrict the generality of the calculation of the contributions to the main process on (o, T), can be seen from the following arguments. If a sub-group were to enter several times, the group can be divided into a number of new sub-groups, which all fulfil the condition of a single entry. Further, if a sub-process starts at a point before ${}_i\tau_0$, and, if the termination point of the process is later than ${}_i\tau_1 \leqslant T$, the contributions on the interval (o, T) to the random function, and to the operational time of the main process can be written ${}_iy_1 - {}_iy_0$, ${}_is_1 - {}_is_0$, where ${}_iy_j$, ${}_is_j$ refer to the points ${}_i\tau_j$, $j =$ 0,1. These contributions are equal to those, which are rendered by a sub-process, which starts at zero and terminates at ${}_i\tau_1 - {}_i\tau_0$, and such that the random function and the operational time are equal to zero at $\tau =$ 0, and to ${}_iy_1 - {}_iy_0$, ${}_is_1 - {}_is_0$ respectively at ${}_i\tau_1 - {}_i\tau_0$. For the case where ${}_i\tau_1 > T$, the ith sub-process shall not be transferred to another scale.

For a non-stationary main group $s_{2i} = k_i \int\limits_0^{\tau - {}_i\tau_0} w_i(u)\, du$, $i =$ 1,2 . . . m_1, corresponds to τ, if ${}_i\tau_1 \leqslant T$, and with the modification of the integral to $\int\limits_{{}_i\tau_0}^{\tau}$ for ${}_i\tau_1 > T$, so that, in this case, an amalgamation of m_1 sub-processes in the form of cPp: I, for a fixed value of m_1, is a cPp with the structure function of the form $\prod\limits_{i\ 1}^{m_1} {}^* \; U_{2i}(\bar{s}_2 v/s_1)$, which depends on τ so that the amalgamated process is a cPp i.w.s.; this is consistent with the assertion for the general case quoted from [4]. The "very special case" in this assertion implies that $s_{2i} = k_i\, s_2$, and the structure function of the amalgamation of m_1 cPp: I is in the form $\prod\limits_{i=1}^{m_1} {}_2U_i(v/k_i)$, independently of τ. Thus the "very special case" in [4] is characterized by the concept a stationary main group, for a simple grouping, where all the sub-processes are cPp: I. In this case $\overline{Y}(s_i, \bar{s}_2)$ can be reduced to the form $\overline{Y}(s_1, s_2)$; in the elementary case $\overline{Y}(s_1, s_2)$ are distributed with distribution functions in the alternative forms of (5a) and (5b). It may be possible, to extend these forms to include also a non-stationary main group, and, under certain conditions (cf. Theorem 7 below), to the non-elementary case.

By the assumptions made so far, the main group is, in general, *heterogeneous* i.e. ${}_iY(s_{2i})$ are generally dependent of i. In the particular case, where ${}_iY(s_{2i}) = Y(s_2)$ independently of i for $i = 1,2 \ldots m_1$, the main group being *homogeneous*, $\bar{Y}(s_1, s_2)$ is, in the elementary case, and, under certain conditions (cf. section 7 and Theorem 7 here below), in the non-elementary case distributed in a bunch distribution.

An *iterated grouping* shall be defined as follows. A *head group* of a population is defined as an, at least, enumerable set of main groups, each defined as in the simple grouping. The relation between the head group and the main groups shall be the same as the relation between the main group and its sub-groups in the simple grouping. As, generally, a main group is heterogeneous, the iterated grouping could be interpreted as a simple grouping, if this were to include heterogeneous sub-groups.

In the following theorem it will be referred to a condition defined in the following lines.

> *Condition A.* In a Markov process, defined by the intensity function $p_n(t)$ for $t < \zeta$, which is continuous in t for every fixed n, $\zeta = \zeta(\omega)$ being the termination point of the process, the conditional probability, $f_{n,\nu}(t, \Delta t)$, for the occurrence of ν events on $(t, t+\Delta t)$ relative to the hypothesis that n events have occurred on the interval $(0, t)$ shall satisfy relations in the form given for $g_{n,\nu}(t, \Delta t)$ in Theorem 1, for $\nu = 0$, 1 and > 1, respectively, for $t+\Delta t < \zeta$, and shall be equal to zero, for $\nu > 0$, and for $t+\Delta t \geq \zeta$.

A Markov process with never decreasing, right continuous trajectories, which satisfies this condition belongs to the class of pure birth processes.

For the following theorem two sequences of index vectors $\{M_j\}$ and $\{S_k\}$, will, further, be introduced. For each fixed value of m_1 and m_2, the vector $M_j\ \{{}_i\mu_j,\ i = 1,2 \ldots m_1\}$ shall, for a fixed value of j, be obtained by the choice of m_1 values, in an arbitrary but fixed order, of values for ${}_i\mu_j$, such that, for the fixed value of j, ${}_i\mu_j$ is, for each i, equal to one of the components of the vector

$(0, 1 \ldots m_2)$, and that $\sum_{i=1}^{m_1} {}_i\mu_j = m_2$ for the fixed value of j. This shall be iterated in every possible way including the permutations, and $\sum'_{(j)}$ shall denote the summation over all possible values of j, obtained in this way. Similarly, for each fixed value of m_1 and ν, the vector $S_k = \{{}_i\sigma_k,\ i = 1, 2 \ldots m_1\}$ shall for a fixed value of k, be obtained by the choice of m_1 values, in an arbitrary but fixed order, of values for ${}_i\sigma_k$, such that for the fixed value of k ${}_i\sigma_k$ is, for each i, equal to one of the components of the vector $(0, 1)$, and that $\sum_{i=1}^{m_1} {}_i\sigma_k = \nu$. This shall be iterated in every possible way including the permutations, and $\sum''_{(k)}$ shall denote the summation over all possible values of k, thus, obtained.

THEOREM 3

Consider a simple grouping, as previously defined in this section, on the domaine $(0, s_1) \times (0, \bar{s}_2)$ corresponding to $(0, T)$ on the absolute scale, in the elementary case, where ${}_iY(s_{2i}) = {}_iN(s_{2i}) = {}_iN$ say, and $\bar{Y}(s_1, \bar{s}_2) = \bar{N}(s_1, \bar{s}_2) = \bar{N}$ say. Let the conditional probabilities, as defined in Theorem 2, be denoted ${}_if_{m_{2i},\nu}\,(s_{2i}, \Delta s_{2i})$ and let the indices ${}_i\mu_j$ be defined as in the previous paragraph. Let it, particularly be assumed that ${}_i\tau_1 - {}_i\tau_0$, defined previously, does not depend on ω, which, thus, applies also to the corresponding point ζ_i on the operational scale.

Then,

1) $\bar{N}$ is an admissible random function with the probability distribution, $\bar{R}_{m_1}\,(s_1, \bar{s}_2)$ say, which fulfils the following relation.

$$\bar{R}_{m_2}(s_1, \bar{s}_2) = Q_0(s_1) + \sum_{m_1=1}^{\infty} Q_{m_1}(s_1) \sum_{(j)}{}' \prod_{i=1}^{m_1} {}_iR_{{}_i\mu_j}(s_{2i}) \tag{6a}$$

2) If each sub-process satisfies Condition A, defined here above, the main process satisfies also this condition, and is defined by the intensity function $r_{m_2}(s_1, \bar{s}_2)$ given by the following relation.

$$\bar{r}_{m_2}\,(s_1, \bar{s}_2) = \sum_{m_1=1}^{\infty} Q_{m_1}(s_1) \sum_{(j)}{}' \prod_{i=1}^{m_1} {}_iR_{{}_i\mu_j} \sum_{i=1}^{m_1} {}_ir_{{}_i\mu_j}(s_{2i}) \,/\, \bar{R}_{m_2}(s_1, \bar{s}_2) \tag{6b}$$

3) The sufficient condition for that Condition A being satisfied for the main group, is also a necessary condition.

Proof. If for all values of i, ${}_iN$ are admissible, as assumed in the model, this must also hold for the random function equal to the sum $\sum_{i=1}^{m_1} {}_iN$ for fixed m_1, consequently also for $\overline{N}$, defined as the weighted average of this sum with respect to $Q_{m_1}(s_1)$, $m_1 = 0, 1 \ldots$. The probability for the occurrence of ${}_i\mu_j$ events in the ith sub-process, $i = 1, 2 \ldots m_1$, on the interval $(0, s_{2i})$ for every fixed value of m_1 and j, is given, provided all the sub-groups were to belong to the main group at T, by the product of the independent probabilities in the jth term of (6a), where, by definition, $\sum_{i=1}^{m_1} {}_i\mu_j = m_2$. Should some of the sub-processes terminate earlier than T, the parameters for these groups s_{2i} are, by definition, equal to ζ_i, corresponding to ${}_i\tau_1 - {}_i\tau_0$ and, therefore, the assumption that ${}_i\tau_1$ for all i are $\geqslant T$, may be removed. The event that m_2 events occur on the interval for fixed m_1, can be realized in several ways; thus, by the definition of $\sum\limits_{(j)}'$, this sum applied to the products for each possible value of j, represents the probability of m_2 events for fixed m_1. If also m_1 is allowed to vary, the weighted mean of these probabilites with respect to $Q_{m_1}(s_1)$ will be obtained, so that (6a) holds, and, thus, the assertion 1) has been proved.

If all sub-processes satisfy Condition A, the values of ${}_if_{m_{2i},\nu}$ $(s_{2i}, \Delta s_{21})$ are for $\nu > 1$ of lower order than the order of Δs_{2i}, for sub-processes with $\zeta_i - \Delta s_{2i}$ corresponding to a point $<$ T even equal to zero. The conditional probability for the occurrence of ${}_i\sigma_k$ events in the ith sub-process, $i = 1, 2 \ldots m_1$, on the interval $(s_{2i}, s_{2i} + \Delta s_{2i})$ relative to the hypothesis that ${}_i\mu_j$ events have occurred on the interval $(0, s_{2i})$ can, for every fixed value of m_1, m_2 and j, be written in the form $\sum\limits_{(k)}'' \prod\limits_{i=1}^{m_1} {}_if_{{}_i\mu_j,1}^{{}_i\sigma_k}\, {}_if_{{}_i\mu_j,0}^{(1-{}_i\sigma_k)}$ where ${}_i\sigma_k$ have been defined in the last paragraph before the theorem (the arguments s_{2i}, Δs_{2i} have been left out), and, where the probabilities for more then one change on $(s_{2i}, s_{2i} + \Delta s_{2i})$ have been neglected. By the insertion of the expressions for ${}_if_{{}_i\mu_j,\nu} = 0, 1$ according to Condition A, the terms of this sum, which contain at least one factor ${}_if_{{}_i\mu_j,1}$ for a sub-process, for which $\zeta_i - \Delta s_{2i}$ corresponds to a point τ less than T will vanish, and other terms, containing more than one such factor, for any sub-process, for which $\zeta_i - \Delta s_{2i}$ corresponds to $\tau > T$ will be of lower order than the order of Δs_{2i}. Consequently, if quantities of lower order than the order of Δs_{2i} are neglected, the sum reduces to $\sum_{i=1}^{m_1} {}_ir_{i\mu_j}(s_{2i})$, observing that for values of i for which $\zeta_i - \Delta s_{2i}$ corresponds to $\tau < T$, ${}_ir_{i\mu_j}(s_{2i}) = {}_ir_{i\mu_j}(\zeta_i) = 0$. Then, the absolute probability of the composite event, that m_2 events occur on $(0, \bar{s}_2)$, and $m_2 + 1$ events on $(0, \bar{s}_2 + \Delta\bar{s}_2)$ in the main process, is given by the numerator, and, by the assertion 1), the probability for the occurrence of m_2 events on $(0, \bar{s}_2)$ by the denumerator of (6b), which, thus, holds, The proof implies, that, if the Condition A holds for every sub-process, it holds also for the main process. Thus, the assertion 2) has been proved.

If, for $\nu > 1$ and for at least one value of i, ${}_if_{i\mu_j,\nu}\,(s_{2i}, \Delta s_{2i})$ and ζ_i - Δs_{2i} corresponding to $\tau > T$, should be of the same or higher order than the order of Δs_{2i}, this should imply that the contribution to the conditional probability for ν events in the main process, should be of this order. This leads to that, in this case, Condition A does not hold for the main process. In fact, this affords en indirect proof of the assertion 3) of the theorem, for a fixed value of T. By letting T decrease until all the sub-processes have $\zeta_i - \Delta s_{2i}$ corresponding to $\tau > T$, the proof can be extended to include anyone sub-process, which does not satisfy the Condition A. Thus, the assertion 3) has been proved.

Remark. By a remark made here above (just before the definition of the iterated grouping), a simple grouping being both stationary and homogeneous leads in the elementary case always to a bunch distribution according to the general definition quoted in section 1 of this note (Thyrion [1], p. 68). Thyrion has, however, ([11], Ch. 2) introduced a specified model of such distributions, here called the *T-model*, which, in fact, is less general, than the model based on the ch.f. in section 1. The T-model has been introduced in order to allow for the occurrence of several, simultaneous changes (cf. [11], p. 49), while by the theorem, the general model can also be applied to cases, where only one change on a small interval has a probability of the same order as the order of the length of the interval, while the probability of more than one change is of lower order. In the T-model the number of bunches is defined as the number of sub-processes, in which already at least one change has occurred. This number may be designated by m, which, by definition, is less than or equal to m_1 in the theorem, the probability distribution of m may be designated by $\tilde{Q}_m(s_1)$. By a similar deduction to that used in the theorem. Thyrion obtained for the T-model in the particular case, where ${}_iR_{m_2}(s_{2i}) = R_{m_2}$ independently of s_{2i} and i, an expression for $\bar{R}_{m_2}(s_1)$ for the main process of the T-model, in the form of (6a) with two important modifications, one implying the elimination of all factors in the product appearing in (6a), for which ${}_i\mu_j = 0$, and the other a truncation of the sum over m_1 on account of the fact that terms in this sum for which $m_1 > m_2$ will for the T-model vanish. In fact, by the definition of m, the non-occurrence in $m_1 - m$ sub-processes, with fixed m_1 and m has the probability 1. Therefore, one may for the T-model substitute the conditional probability for ${}_i\mu_j$ events, $i = 1,2 \ldots m_1$ for every fixed value of m_1 and m relative to the hypothesis that no events have occurred in $m_1 - m$ sub-processes, for the absolute probabilities used in the proof of (6a). By the substitution in this proof of $\tilde{Q}_m(s_1)$ for $Q_{m_1}(s_1)$, and by the application of the combinatorial methods, as used by Thyrion for the case where ${}_iR_{m_{2i}}(s_{2i}) = R_{m_2}$, it can be proved that Thyrion's form for (6a), called the *T-form* here below is a particular case of (6a). Further, the T-form can be extended to cases, where ${}_iR_{m_{2i}}(s_{2i})$ depend on s_{2i} and i.

For a stationary and homogeneous main group an extension of the bunch distribution in the T-form to the non-elementary case was indicated by Thyrion in [11], and, further, analyzed by him [12], for the case, where ${}_iR_{m_{2i}}(s_{2i}) = R_{m_2}(s_1)$ i.e. dependent on the same parameter as the parameter in $\tilde{Q}_m(s_1)$. In [12] he states, that the distribution functions in this case can be written $\sum\limits_{m_2=0}^{\infty} \bar{R}_{m_2}(s_1) V^{m_2*}(x)$ if $\bar{R}_{m_2}(s_1)$ is an elementary bunch distribution, for the particular case, where $\tilde{Q}_m(s_1)$ is a Poisson distribution, and, that this condition should also be a necessary condition. According to Arfwedson, the condition is necessary for an arbitrary change distribution, only if $R_{m_2}(s_1) = R_{m_2}$ independently of s_1, in the opposite case, a function in the form of the elementary Poisson bunch distribution i.e. a function of the form $B_n(t) = \sum_{m=0}^{\infty} e^{-t}\, t^m R_n^{m*}(t)/m!$ shall be a probability distribution, only if certain conditions are fulfilled. These conditions are, $R_0(t) \geqslant 0$, and two other, given in the form of two inequalities for $b_n = B_n(t)/B_0(t)$, namely $b_2 \geqslant b_1^2/2$ and $b_3 \geqslant b_1 b_2 - b_1^3/3$. He adds, however, the remark, that, if $R_0(t)$ is negative, it can be eliminated by a transformation of the parameter. Such an elimination leads to the T-form. Arfwedson seems in the introduction to have based the Poisson bunch distribution on the T-model, as the bunch is associated with insurance claims which occur in a bunch, exemplified by an air craft accident, but his formal developments are more general, as for the T-model $R_0(t)$ is always $= 0$. In one of the examples, earlier introduced by Arfwedson, $R_0(t)$ was assumed to be positive, and the deduction of this model was based on the assumption, that the probability of multiple changes on a short interval is of the same order as the order of the length of this interval. Thyrion has for the same example deducted distribution in the T-form, eliminating $R_0(t)$ [1, 12]; Arfwedson has remarked, that the two solutions are consistent, [13]. Thus, the distribution according to (6a) is in the T-form, if $R_0(t) = 0$, must be reduced to this form, if $R_0(t) < 0$, or may be, alternatively, given with $R_0(t) > 0$, or, in the T-form, by the elimination of $R_0(t)$. A bunch distribution in the T-form does, thus, not necessarily, be applicable to the T-model, as, (6a) can be given in the T-form, even if Condition A is fulfilled. In the example just mentioned $R_{m_2}(t) = R_{m_2}$ independently of t, so that Condition A is not fulfilled. Later in this note (the remark to Theorem 7). an example will be given, where the bunch distribution function may be transformed into the T-form, and where Condition A holds. Thus, it may be said, that in a way, the T-form is more general than (6a) with positive ${}_iR_0(s_{2i})$, even if the grouping model is more general than the T-model.

8. *Particular results for the simple grouping*

For a fixed value of m_1, the amalgamation of m_1 sub-processes in the form of cPp : 1 for each i, leads to a cPp : 1 or a cPp i.w.s. depending on whether the main group being stationary or non-stationary. In this section it will be assumed that the main group is stationary, and that all the sub-processes are cPp : 1.

Then, the following recurrence relation similar to that, given by Lundberg for a cPp : 1 ([3], (93)), holds for the probabilities of the main process.

$$\bar{r}_{m_2}(s_1, s_2) = (m_2+1)\, \bar{R}_{m_2+1}(s_1, s_2) \,/\, s_2\, \bar{R}_{m_2}(s_1, s_2). \tag{7a}$$

In fact, by the insertion of recurrence relations according to ([3], (93)) for each sub-process into (6b) for the case considered here, (7a) is directly obtained. By the differentiation of $\bar{R}_{m_2}(s_1, s_2)$ with respect to s_2 the following relation (7b) is obtained, if the dependence of s_1 on s_2 is neglected. In fact, if also this dependence is taken into account, it will not alter the result in (7b), provided that $Q_{m_1}(s_1)$ is a probability distribution in a process, and satisfies the Chapman-Kolmogoroff equation. This will first be proved on the assumption, that also $Q_{m_1}(s_1)$ satisfies Condition A. If $q_{m_1}(s_1)$ denotes the intensity function in this case, the derivative of $\bar{R}_{m_2}(s_1, s_2)$, with respect to s_1, if the dependence between s_2 and s_1 is not taken into account, can be written

$$-\, q_0(s_1)\, Q_0(s_1) - \sum_{m_1=1}^{\infty} q_{m_1}(s_1)\, Q_{m_1}(s_1) \sum_{(j)}' \prod_{i=1}^{m_1} {}_iR_{i\mu_j}(s_2) +$$

$$+ \sum_{m_1=1}^{\infty} q_{m_1-1}(s_1)\, Q_{m_1-1}(s_1) \sum_{(j)}' \prod_{i=1}^{m_1-1} {}_iR_{i\mu_j}(s_2)$$

for a fixed value of m_2. This expression is equal to zero, which may be seen by the insertion of the variable of summation $\mu = m_1 - 1$ into the second term. By using Chapman-Kolmogoroff's equation, this may be extended to the case, where the process defined by $Q_{m_1}(s_1)$ satisfies this equation, but not Condition A. Therefore, the following relation holds under these conditions.

$$\partial \bar{R}_{m_2}(s_1, s_2)/\partial s_2 =$$
$$-\, \bar{r}_{m_2}(s_1, s_2)\, \bar{R}_{m_2}(s_1, s_2) + \bar{r}_{m_2-1}(s_1, s_2)\, \bar{R}_{m_2-1}(s_1, s_2) \tag{7b}$$

It will be proved, that the function $R_{n_2, m_2}(t_1, t_2, s_1, s_2)$ defined by the following relation satisfies (7b) with the substitution of this function for $\bar{R}_{m_2}(s_1, s_2)$.

$$\chi_{m_2}(s_2) = \bar{R}_{n_2,\, m_2}(t_1, t_2, s_1, s_2) =$$

$$= \binom{m_2}{n_2} \left(\frac{t_2}{s_2}\right)^{n_2} \left(1 - \frac{t_2}{s_2}\right)^{m_2 - n_2} \bar{R}_{m_2}(s_1, s_2) \,/\, \bar{R}_{n_2}(t_1, t_2) \qquad (7c)$$

In fact,

$$\chi'_{m_2}(s_2) \,/\, \chi_{m_2}(s_2) =$$
$$- m_2/s_2 + (m_2 - n_2) \,/\, (s_2 - t_2) - \bar{r}_{m_2}(s_1, s_2) + m_2/s_2,$$

and $\chi_{m_2-1}(s_2) \,/\, \chi_{m_2}(s_2) = (m_2 - n_2) / \ [(s_2 - s_1)\ \bar{r}_{m_2-1}(s_1, s_2)\ ]$.

From these relations (7b) is directly obtained. This is a modification of a part of the proof of Theorem 6 in [3] (p. 73-75). By using the remaining parts of this proof, it can be proved that the solutions of (7c) satisfy the fundamental conditions for the conditional probabilities of a generalized (cf. section 4 here above) Markov process with the absolute probabilities $\bar{R}_{m_2}(s_1, s_2)$.

9. *Remarks on the iteration of a grouping and on an extension to the case where* ${}_i\tau_1$ and ${}_i\tau_0$ *depend on* ω

In the proof of Theorem 3 the homogeneity of the sub-groups, which is assumed for the simple grouping, has not been used. Consequently, the relations (6a), (6b) hold for a head group, if ${}_iR_{i\mu_j}(s_{2i})$, ${}_ir_{i\mu_j}(s_{2i})$ are associated with the ith main group. As the relations hold for the main group, expressions in the form of (6a), (6b) may be inserted for the functions just mentioned, respectively. Also the remark to the theorem, and the result of the previous section may be modified, so that the iterated grouping will be included. It is, further, seen that the grouping may be iterated several times. Further, if it is assumed that s_{2i}, and the operational parameter corresponding to ${}_i\tau_0$ are, for $i = 1, 2 \ldots m_1$ to be considered sequences of different values of random variables, σ and σ_0 say, the probabilities and numerator of the intensity function can be expressed for every fixed value of m_1, σ, σ_0 in the form of a weighted mean of $\underset{(j)}{\Sigma'}$ in (6a), (6b) with the substitution of σ and $m_1\sigma - \sigma_0$ for s_{2i} with the weight functions equal to the probabilities for $\sigma + \sigma_0$ being $\leq m_1\sigma$ and $> m_1\sigma$ respectively. Then, (6a) and the numerator in (6b) will be modified into the appropriate means of these expressions with respect to the joint distribution of m_1,

σ and σ_0. The iterated grouping without and with chance variation in s_{2i} in the case, where all the sub-groups are cPp : 1, can be explained in terms of iterated aco cPp : 1 and modified such concepts.

10. *Remarks on the application of the grouping models*

In the application of stochastic models to problems of the type met with in the insurance field and to similar problems in other fields, it is generally aimed at a classification of the statistical results, which ensures the relative homogeneity of the sub-groups, as much as this is, in practice, feasible. For life insurance rating the experience is, generally, grouped with regard to age, and, eventually, to sex and/or to duration of insurance, either after sorting out risks considered more dangerous than normal risks (by medical examination and/or declarations by the insured), or with neglect of the existence of such risks, on account of their small weight in the population. Even a finer grouping may be applied, such as a grouping with respect to main groups of death causes; in fact, a finer grouping would, in principle, be preferable. Morbidity statistics is often given in finer groupings. Also for property insurance, and for liability insurance, it is aimed at a homogeneity of the sub-groups, which, however, in practice, is feasible only to a very small extent; material deviations of the risks within a statistical group are often neglected. Further, the classification of the statistics is, generally, deemed to be too elaborate for the application to the actual table of rates, i.e. several statistical groups are pooled to form a single tarif group. All this affords examples of the application of simple grouping, where the heterogeneity of the sub-groups, the statistical groups, and of each main group, tarif group, is neglected. In practice, however, these neglected deviations are so large, that the method commonly used, gives a very coarse description of reality. In fact, the individual properties of each unit insured are of material influence on the risk, e.g. in motor insurance, the skill and experience of the ordinary driver, and the driving properties of the vehicle insured. Therefore, systems for the correction of errors implied in the rating, such as the distributions of dividends, and systems of bonus-malus are often applied. The problem of finding rational systems of this kind entails, particularly, for property and liability insurance, many difficulties. With regard to bonus-malus systems different authors have argued for quite different principles, and even reached different conclusions from the same statistical experience. A particular difficulty is implied in cases, where the risk premium has a steadily increasing time trend, as for motor insurance in most countries, which gives uncertain estimates of the reserves needed, so that the rebatement of the premiums may often lead to larger decrease of the reserves than anticipated. This has in motor insurance caused heavy losses, in situations, where too liberal rebates have been granted. For the judgement of the experience a posteriori well-known statistical methods have been applied. In such applications it is of utmost importance to apply a rational classification, with still higher demands on the homogeneity of the sub-groups, unless it is feasible to evaluate the results by using models based on more or less heterogeneous sub-groups. Therefore, the iteration of the simple grouping suitably modified with regard to the phenomenon considered, seems to be apt for experience rating, which also applies to the rating a priori, to the estimation of the necessary risk reserves, reserves for outstanding liabilities, and for unearned premiums, and to the choice of reinsurance policy. In fact, head groups, main groups, and sub-groups must all be considered heterogeneous. Thus the system of amalgamation of risk processes is, in actual practice, at least as complicated as the model for iterated grouping defined in the section 7 here above, the modifications dealt with in the previous section, seem to be more appropriate. In principle, each group of several insurance treaties, is essentially heterogeneous. At any rate, in theory, one should associate each lowest group of a classification with a single point of a multi-dimensional factor space, the components of the coordinate vector representing all factors of possible influence of the risk, and each component being sufficiently differentiated. The number of points, thus defined, is naturally very high, so that the exit and the entrance of such groups may be considered governed by chance, which is accounted for in the grouping models by allowing the exit and the entrance times to be random variables, which leads to the dependence of ω for the termination points $\zeta_i = \zeta_i(\omega)$, and to the introduction of probability measures for the times here concerned. The whole discussion in this section leads to the conclusion, that the grouping models, as defined in section 7, modified as indicated in this and in the previous section here above, give a very realistic description of reality. This applies not only to the properties dealt with in Theorem 3, but also to the T-model introduced by Thyrion, and discussed by Arfwedson, particularly, in the case, where the model allows for several simultaneous events. In section 8, the conditions for the application of relations known from well-known models used earlier have been given. For the application of the grouping models it seems to be of interest to study the conditions, under which these results may be transformed into other forms. To this effect an analysis will be made in the following sections.

11. *A theorem on ch.f. dependent on a parameter*

For the analysis just mentioned, the following theorem will be needed, for this theorem the following definitions are introduced.

Let $\chi(\eta)$ be a function of ηi, η being a real variable, and i the imaginary unit, where $\chi(\eta)$ is bounded, and a continuous function of η. Let J be an operator defined by $J\chi = \int_0^A\int_0^A \chi(\eta - \xi) e^{ix(\eta-\xi)}\, d\eta d\xi$ $= I(x, A)$ say. The sufficient and necessary conditions for $\chi(\eta)$ being a ch.f. of a distribution are, that $\chi(0) = 1$, and, that $I(x, A)$ is real and non-negative for all real x and all $A > 0$ (Cramer, [14], p. 91). If $\chi(\eta)$ fulfils these conditions, it shall be said, that $\chi(\eta)$ $\subset \mathfrak{C}$.

Let $v \in V$ be a non-negative real parameter and $\lambda \in \Lambda$ a transform of v, such that the mapping λD_v being the domaine on Λ corresponding to $D_v \subset V$, is defined by the transformation $\lambda = \lambda(v)$ taking any point $v \in D_v$ with one-to-one correspondence to one point $\lambda \in \lambda D_v$; the mapping of Λ is analogously defined by the inverse $\bar{v} = v(\lambda)$ of $\lambda(v)$. If there exists a transform $\bar{\chi}(\eta, \bar{v})$ of $\bar{\chi}(\eta, v) \subset \mathfrak{C}$, which is obtained by the mapping of V onto Λ, such that $\chi'(\eta, \bar{v}) =$ $\partial[\lambda(\bar{v})\ \bar{\chi}(\eta, \bar{v})] / \partial\lambda$ belongs to $\mathfrak{C}$ on $\lambda D v \subset \Lambda$, it shall be said, that $\bar{\chi}(\eta, \bar{v}) \subset \mathfrak{C}_\lambda(\lambda D_v)$. Further, if a transform, $\rho = \rho(v)$ say, fulfils the conditions, $\rho(0) = 0$, and, in addition, has derivatives of any finite order, such that $(-1)^\nu \rho^{(\nu+1)}(v) > 0$, the transform shall be called a ρ-*transform.*

THEOREM 4

Let $t \in T$, $s \in S$, and $u \in U$ be non-negative, real parameters such that $s = s(t) = \log u$, and that s is a ρ-transform of t. Let $\bar{t} = t(s)$, and $\tilde{t} = t(u)$ define the mappings of T on S and U, respectively. Let $\bar{\varphi}_t \subset \mathfrak{C}_t$ (0, ∞), i.e. $\bar{\varphi}_t$ is a solution of a differential equation given in the remarks to the Theorems 1 and 2 for a cPd : 1, defined by $\overline{\psi}_t = \exp\{-s[t(1-\bar{\varphi}_t)]\}$ say. Let $\bar{\varphi}_t$ and φ_t correspond to $W(x, t)$ and $V(x, t)$ respectively as defined in the remark just quoted.

Let $\bar{I}(x, A, \bar{t}) = J\ \bar{\varphi}_{\bar{t}}^{\nu}$ and $I_\nu(x, A, \bar{t}) = J(\bar{\varphi}_{\bar{t}}^{\nu}\,\varphi_{\bar{t}})$: here x assumes any real value, and A any positive value, and J has been defined before the theorem. Further, let $\bar{L}_1\,\bar{\varphi}_{\bar{t}} = \sum\limits_{\nu=1}^{\infty} \bar{l}_\nu \bar{\varphi}_{\bar{t}}^{\nu}$, $L_0\,\bar{\varphi}_{\bar{t}} = \sum\limits_{\nu=0}^{\infty} l_\nu \bar{\varphi}_{\bar{t}}^{\nu}$,

$$\bar{l}_\nu = -k_\nu(\bar{t}, s),\ l_\nu = k_\nu(\bar{t}, s'),\ \text{where}\ k_\nu(t, s) = \frac{(-t)^\nu s^{(\nu)}(t)}{\nu!\ s(t)},$$

$$\text{and } Q(x, A, \bar{t}) = \sum_{\nu=0}^{\infty} [\,l_\nu I_\nu(x, A, \bar{t}) - l_{\nu+1}\,\bar{I}_{\nu+1}(x, A, \bar{t})\,].$$

It is assumed, that $Q(x, A, \bar{t})$ is non-negative on a domaine $sD_t \subset S$.

Then, $\bar{L}_1\bar{\varphi}_{\bar{t}} \subset \mathfrak{C}_s\,(sD_t)$, and $\overline{\psi}_{\tilde{t}} \subset \mathfrak{C}_u(uD_t)$.

Proof. Let ${}_1h_s = \partial[s(\bar{t})\,\bar{L}_1\,\bar{\varphi}_{\bar{t}}]\,/\partial s$ and ${}_2h_u = \partial[u(\tilde{t})\,\overline{\psi}_{\tilde{t}}]/\partial u$ respectively. For the calculation of these functions t is substituted for $\bar{t}$ and $\tilde{t}$ respectively in the expressions within the square brackets, and the expressions obtained differentiated with respect to t, and divided by $s'(t)$ and $u'(t)$ respectively, by using the expression $t\bar{\varphi}_t' = \varphi_t - \bar{\varphi}_t$ according to the remark mentioned in the theorem. Thereinafter, $\bar{t}$ and $\tilde{t}$ respectively are substituted for t. Thus, the following relations are obtained for ${}_1h_s$.

$${}_1h_s = \varphi_{\bar{t}} L_0\,\bar{\varphi}_{\bar{t}} - L_0\,\bar{\varphi}_{\bar{t}} + 1 = \varphi_{\bar{t}} \sum_{\nu=0}^{\infty} l_\nu \bar{\varphi}_{\bar{t}}^{\nu} - \bar{\varphi}_{\bar{t}} \sum_{\nu=0}^{\infty} l_{\nu+1}\,\bar{\varphi}_{\bar{t}}^{\nu}.$$

By the second membrum ${}_1h_s = 1$ for $\eta = 0$ (i.e. $\bar{\varphi}_{\bar{t}} = \varphi_{\bar{t}} = 1$). If the operator J, defined before the theorem, is applied to the third membrum, using the fact, that under mild regularity conditions $W(x, t)$ satisfies a differential equation of the same form as for $\bar{\varphi}_t$, with the substitution of $V(x, t)$ for φ_t, then, $J_1h_s = Q(x, A)$, as $\bar{\varphi}_{\bar{t}}^{\nu} \subset \mathfrak{C}$ is real, and, on sD_t, assumed to be non-negative. Thus, ${}_1h_s \subset \mathfrak{C}$ on sD_t. It has earlier been proved, [11-13], that

$\bar{L}_1\bar{\varphi}_{\tilde{t}} \subset \mathfrak{C}$ on account of the elimination of the negative term, obtained by the Taylor expansion of log $\bar{\psi}_{\tilde{t}}$, by the mapping on S. Thus, it has been proved, that ${}_1h_s \subset \mathfrak{C}_s\ (sD_t)$.

By using the Taylor expansion of d log $\bar{\psi}_t/dt$ in the expression of ${}_2h_u$ before the transformation of t, obtained according to the description above, by the elimination of the negative term according to the method in [11-13], and by the mapping on U, the relation, where $\tilde{s} = s(\tilde{t})$, ${}_2h_u = \bar{\psi}_{\tilde{t}}\ {}_1h_{\tilde{s}}$ is obtained. Thus, ${}_2h_u$ is a product of a ch. f. corresponding to a cPd : 1, as proved in [12], and of ${}_1h_s$, proved above to belong to $\mathfrak{C}$ on sD_t. Thus, both $\bar{\psi}_{\tilde{t}}$ and ${}_2h_u$ belong to $\mathfrak{C}$ on uD_t and, thus, it has been proved that $\bar{\psi}_{\tilde{t}} \subset \mathfrak{C}_u(uD_t)$.

Remark. 1) The ch.f. $\bar{\psi}_t$ corresponds to a cPd: 1, fulfilling the condition, that $s(t) = -\log P_0(t)$ is a ρ-transform of t. As in the simple grouping with stationary main group and with all the sub-processes being cPp : 1, the amalgamation for fixed m_1 is a cPp : 1, it follows, that, if the sub-processes fulfil the condition just mentioned, the amalgamation fulfils the same condition, as $s(t) = \sum_{i=1}^{m_1} s_i(t)$. As $\bar{L}_1\bar{\varphi}_{\tilde{t}}$, in this case, is a weighted average of the same functions for the sub-processes, with $s_i(t)/s(t)$ as weight functions, the theorem holds also for $\bar{\psi}_t = \exp\left\{\sum_{i=1}^{m_1} -s_i\,[t(1-\bar{\varphi}_t)\,]\right\}$ provided that φ_t is the same for all the sub-processes. This applies also to a convolution of cPd : 1 fulfilling these conditions. If the mean function of the change distribution is a never-decreasing function of t, which is often realized in most branches of insurance, it will often occur that $W'(x, t) > 0$ for sufficiently great values of x; also the variance being very often a never-decreasing function of t. As, however, $I(x, A, \bar{t})$ is the mean of an essentially decreasing function of the variable of integration, for great values of this variable, the terms for low values of ν in $Q(x, A, \bar{t})$ may be < 0. In many cases the sum is, however, positive as $l_\nu - l_{\nu+1}$, as a rule, increases with ν.

2) It shall here be, particularly, assumed that $\varphi_t = \varphi$ independently of t. In this case, $Q(x, A, \bar{t})$ reduces to $\sum_{\nu=0}^{\infty} (l_\nu - l_{\nu+1})\ \bar{I}_\nu\ (x, A)$, which is non-negative, if $l_{\nu+1}/l_\nu \leq 1$ for all ν. (This is a sufficient, but

not a necessary condition). It will, in addition, be assumed that $s(t) = \sum_{i=1}^{m_1} s_i(t_i)$, $t_i = p_i t$, and that $s_i'(t_i) = p_i(1 + c_i t_i)^{-a_i}$, p_i, a_i, c_i being positive constants. For $m_1 = i = 1$, $s(t)$ and $\bar{L}_1\varphi$ fulfil the following relations, quoted from Thyrion [15].

$$s(t) = p[1 - (1+ct)^{1-a}] / c(a-1) \text{ for a} \neq 1, \text{ and}$$
$$s(t) = p \log (1 + ct)/c \text{ for } a = 1,$$
$$\text{so that } s(t) \text{ is a } \rho\text{-transform} \tag{8a-b}$$

$$\bar{L}_1\varphi = p \sum_{\nu=1}^{\infty} \binom{1-a}{\nu} (-q\varphi)^\nu / [c(1 + c\bar{t})^{a-1} s(\bar{t})] \text{ for a} \neq 1, \text{ and} \tag{9a}$$

$$\bar{L}_1\varphi = p \sum_{\nu=1}^{\infty} (q\varphi)^\nu/\nu \, cs(\bar{t}) \text{ for } a = 1; \; q = \frac{c\bar{t}}{1 + c\bar{t}} \tag{9b}$$

For $m_1 > 1$, $s_i(t)$ and ${}_i\bar{L}_1\varphi$ shall be defined by the same relations after the addition of the index i to the symbols s, $\bar{L}_1$, a, c, p and q, for all values of i. (If the definitions are extended to include $a_i = 0$, $\bar{\psi}_t$ defines a Hofmann probability distribution for $m_1 = 1$, and an extended such distribution for $m_1 > 1$.) In this case $l_{\nu+1}/l_\nu =$ $(a-1)q/(\nu+1)$ $(a-1)q$ for $a \neq 1$, and equal to $q < 1$ for $a = 1$, if $m_1 = 1$; for $m_1 > 1$, the same relations hold, if $\bar{q}$, $\bar{a}$, $\bar{c}$, are substituted for q, a, c, respectively, where $\bar{q} = \sum_{i=1}^{m_1} {}_iw_\nu q_i / \sum_{i=1}^{m_1} {}_iw_\nu = \frac{\bar{c}\bar{t}}{1 + \bar{c}\bar{t}}$, and

$$\bar{a} = \sum_{i=1}^{m_1} a_i \, {}_iw_\nu \, q_i / \bar{q} \sum_{i=1}^{m_1} {}_iw_\nu; \; {}_iw_\nu = \binom{1-a_i}{\nu} (-q_i)^\nu$$

and $q_i^\nu/\nu!$ for $a_i \neq 1$ and $a_i = 1$ respectively.

Then, by the theorem $\bar{L}_1\varphi$ and $\bar{\psi}_{\bar{t}}$ in this case belong to $\mathfrak{C}_s(sD_t)$ and $\mathfrak{C}_u(uD_t)$ respectively, where D_t is equal to $\{0,\infty\}$ for $\bar{a} < 2$, and, at least, equal to $\{0, [\bar{c}(\bar{a}-1)]^{-1}\}$, where, for $m_1 = 1$, the bars may be removed.

12. *The Thyrion transform of the ch.f. defining a cPd* : 1

Thyrion [1, 15], transformed the ch. f. in the form of $\bar{\psi}_t$, with $s(t)$ being a ρ-transform, as defined in Theorem 4, in the particular case, where $\varphi_t = \exp(i\eta)$ independently of t (the cPd : 1 defined by $\bar{\psi}_t$ is, then, said to be in the "canonical form") to a ch.f. defining a generalized Poisson distribution on the transformed space. Lundberg ([3], p. 57), transformed an elementary Polya process to a time-homogeneous and space-heterogeneous, elementary cPp : 1, on the transformed space. Ammeter [16] transformed a non-elementary Polya process, $\varphi_t = \varphi$ independently of t, in [15] extended to include a Hofmann probability distribution, $\varphi_t = \varphi$. In [17] the present author extended the transform to include non-elementary cPp : 1 or cPp : 2 with φ_t dependent on t. In [12] Thyrion analyzed the transform of a general non-elementary, cPd : 1 in the canonical form, $\varphi_t = \varphi$ independently of t. Arfwedson remarked, [13], that in order to obtain a non-negative value for the probability of non-occurrence in the generalizing distribution, it was necessary to add the condition $s(t) \leqslant t$ for $t > 0$, but that, if this condition is not fulfilled, the transform could be obtained by the elimination of the negative term through parameter transformation, which has been referred to in the remark to Theorem 3 here above (if $s(t) = t$, ψ_t reduces without transformation to $\exp[-t + t\varphi]$, so that the untransformed distribution is a Poisson probability distribution). In fact, in the case, where $s(t) > t$, the transform must be given in the T-form, defined in the remark just mentioned. If a bunch distribution is taken in the sense of the T-model, it seems irrelevant to add the condition, that $s(t) \leq t$. Arfwedson referred to Thyrion's alternative expressions for the transformed distribution functions in the following form, where, however, the notations have been adapted so as to avoid confusion with the notations of this note. The development in [18] lead to similar alternative expressions, in a more general case (cf. also (5a) and (5b) here above).

$$G(y, t) = \sum_{n=0}^{\infty} B_n(t)\ V^{n*}(y), \text{ where } B_n(t) = \sum_{m=0}^{\infty} e^{-t} t^m R_n^{m*}(t)/m! \quad (10a)$$

$$G(y, t) = \sum_{m=0}^{\infty} e^{-t} t^m K^{m*}(y, t)/m!, \text{ where } K(y, t) = \sum_{n=0}^{\infty} R_n(t)\ V^{n*}(y) \quad (10b)$$

As was remarked in section 7 here above, the condition for the simultaneous validity of (10a) and (10b), is that $B_n(t)$ is a bunch distribution, in this case a Poisson bunch distribution. Arfwedson has proved, that this is not a necessary condition unless, either

$R_n(t) = R_n$, or $V(y) = \varepsilon(y - c_1)$, independently of t. If, however, the distribution before the transformation is in the canonical form, the conditions for $B_n(t)$ being a bunch distribution are fulfilled, where however, if $s(t) > t$, $B_n(t)$ must be given in the T-form. Thyrion has proved, that, if $s(t)$ is a ρ-transform, $\exp[-s(t)]$ is completely monotonic, and equal to unity for t tending to zero. As this condition is not necessary, there might, as remarked by Arfwedson, exist cPd : 1, which may not be transformed to a bunch distribution, even if he has proved that every cPd : 1 leads to an expression in the form of $B_n(t)$, which satisfies the second of Arfwedson's conditions for $B_n(t)$ being a bunch distribution. This applies, not necessarily, to the third condition. Arfwedson has not, however, been able to find an example of a cPd : 1, for which $B_n(t)$ after the transformation is not a probability distribution, and, thus, not a bunch distribution.

Thyrion has treated the particular case of the transformation for $\varphi_t = \varphi$ independently of t and $s(t)$, $\bar{L}_1\varphi$ in the forms of (8a-b) and (9a-b) respectively and with $m_1 = 1$. In the remark 2) to the Theorem 4 the conditions, in this case for $m_1 \geq 1$, for $\bar{L}_1\varphi_t$ and $\bar{\psi}_t$ belonging to $\mathfrak{C}_s(sD_t)$ and $\mathfrak{C}_u(uD_t)$, respectively, were deducted by using the theorem mentioned. It might be remarked here, that a direct deduction of these conditions for the particular case concerned, without using the theorem, can be made by an easy calculation, which does not imply the J-transform, but is based on Newton's binomial formula.

Under the conditions of the remark 2) to Theorem 4, D_t is an interval with the lower limit zero. In the general case of the theorem. the interval defining D_t may start at a point $t_0 > 0$. As, however, a conditional cPp : 1 defined on the interval (t_0, t_1) may be transformed to an absolute cPp : 1, defined on the interval $(0, t_1 - t_0)$, i.e. with starting point at t_0 (cf. [3], p. 91), it does not restrict the generality to assume, that the ch.f. corresponds to a ch.f. $\subset \mathfrak{C}_s(sD_t)$ with $sD_t = \{0, \zeta_s\}$ say, and $uD_t = \{0, \bar{\zeta}_u\}$ say, where $\zeta_s = \log \bar{\zeta}_u = s(t_1 - t_0)$.

THEOREM 5

Let $X(t)$ and $Y(t)$ be defined as in the remark to the Theorems 1 and 2 here above, i.e. both functions are distributed with the same cPd : 1, and $Y(t)$ constitutes on $(0, \zeta)$ a cPp : 1 defined by the intensity function $\pi_n(t)$. Let $\bar{\pi}_n(\tau)$, and $\bar{P}_n(\tau)$ be obtained from $\pi_n(t)$, and the distribution of the number of events, $P_n(t)$, by the transformation of t to the absolute scale τ, and let $\varkappa w(\tau)$ be the mean $\sum_{n}^{\infty} \bar{\pi}_n \bar{P}_n(\tau)$. Assume, particularly, that $P_n(t)$ is in Thyrion's canonical form, so that the ch.f. of $Y(t)$ is in the form of $\bar{\psi}_t$ in Theorem 4, with $s(t)$ being a ρ-transform; the notations D_t, ${}_1h_s$ shall also be defined as in Theorem 4, and let $ds/d\tau = \varkappa_0 w(\tau)$, and the mapping of T onto S defined as in the previous section.

Then, this mapping transforms $Y(t)$ into the random function $Z(s)$ say, with the ch.f. exp $[-s + s\, \bar{L}_1\, \bar{\varphi}_{\bar{t}}]$, in the notations of Theorem 4, distributed with a non-elementary Poisson bunch distribution according to (10a-b), with s, and the distibution function corresponding to $\bar{L}_1\, \bar{\varphi}_{\bar{t}}$ substituted for t, and $K(y, t)$, respectively. If $\bar{\pi}_n(\tau)/n$ converges uniformly on finite intervals of τ, $Z(s)$ constitutes on the domaine, $(0, \zeta_s)$ say, which corresponds to the intersection of D_t and $(0, \zeta)$, a generalized, non-elementary Poisson process, which has the intensity, with respect to τ, $\varkappa_0 w_0(\tau) = \varkappa \bar{w}(\tau)\, [\bar{s}'(t)]_{t\ =\ \bar{t}(\tau)}$ and the change distribution defined by ${}_1h_s$.

Proof. The equivalence in the sense of section 1 between $X(t)$, $Y(t)$ and $Z(s)$, and the existence of a probability distribution in the form of $B_n(t)$ in (10a) is a consequence of the discussion here above of the Thyrion transform of a cPd : 1 in the canonical form. As, by Theorem 4, $\bar{L}_1\, \bar{\varphi}_{\bar{t}} \subset \mathfrak{C}_s(0, \zeta_s)$, the distribution functions of $Z(s)$ are on the domaine $(0, \zeta_s)$, according to Corollary 1 in section 5 here above, in the form of the absolute distribution functions of a Poisson process, as defined in the assertion. On account of the existence of the translator operators before the mapping, $Z(s)$ constitutes, if the assumption of convergence is satisfied, (cf. [3], (64)), the Poisson process defined in the assertion, which, thus, has been proved.

Remark. The theorem holds for a convolution of cPd : 1, which satisfies the conditions in the remark 1) to the Theorem 4. In the particular case, dealt with in the remark 2) to the same theorem, i.e. where $\bar{\psi}_t$ defines an extended Hofmann distribution with

$\varphi_t = \varphi$, independently of t, sD_t is equal $(0,\infty)$ for $\bar{a} \leq 2$, (cf. [3], the Theorems 2 and 7 A3)), and corresponding to $D_t = \{0,[\bar{c}(\bar{a}-1)]^{-1}\}$ for $\bar{a} > 2$, in the notations of the remark just quoted.

THEOREM 6

Let $Z(s)$ with the ch.f. $\exp[-s + s\,\bar{\chi}_s]$ constitute a Poisson process on $(0, \zeta_s)$. If there exists a non-negative, real parameter t, such that $\bar{\chi}_s$ can be written in the form of $\bar{L}_1\,\bar{\varphi}_t$ according to Theorem 4, with $s(t)$ being a ρ-transform, the mapping of $S = \{s\}$ onto $T = \{t\}$ transforms $Z(s)$ into a random function, which, at least on the domaine of T corresponding to $(0, \zeta_s)$, has the properties of $Y(t)$ in Theorem 5.

Proof. By assumption, $\bar{\chi}_s \subset \mathfrak{C}_s(0,\zeta_s)$, then, the reversion of the proof in Theorem 5 is self-evident, which proves Theorem 6.

Remark. In [4] the following statement with respect to the relation between $Y(t)$ and $Z(s)$ of the preceding theorems, in the particular case, where $Y(t)$ constitutes a non-terminating Polya process with $\varphi_t = \varphi$ independently of t, has been given. This statement purports a point of view on this relation, which is principally different from the interpretation in the preceding theorems. It is, in fact, said in [4] (p. 18) "As both the Poisson parameter and the generalizing distribution," in the notations of this note equal to $s(t)$, and defined by $\bar{L}_1\,\bar{\varphi}_t$ respectively, "depend on t, different times t correspond to different generalized Poisson processes." This implies that $Z(s)$, being defined on S, is related to T (in this case $(0, \zeta)$ and sD_t are equal to T and S respectively, by assumption, and by the remark to the Theorem 5 above). In the author's opinion it is more natural to refer $Z(s)$ to S, where it is defined; this leads to the following statement, which seems to be more elucidating. Different "times" t correspond to different "times" s, $Z(s)$ constitutes on S one, and only one generalized Poisson process, as defined in Theorem 5 for the general case (the differences between $\bar{w}(\tau)$ and $w_0(\tau)$, and between $\bar{L}_1\,\bar{\varphi}_t$ and $\bar{\varphi}_t$ seem not to make the interpretation of the relation in [4] preferable. For comparison, two simple examples of similar transformations will be given here below. It will be seen that these examples were interpreted in principal agreement with the point of view expounded in this note.

In these examples the intensity functions $w(\tau)$, and $p_n\, v(t)$ (i.e. of a Polya process, [3], Theorem 10a), were transformed, by the mappings of $\{\tau\}$ on $\{t\}$, $t = \int_0^\tau w(u)\; du$, and of $\{t\}$ on $\{s\}$, $s = -\log\, P_0(t)$, to $\varkappa$, and p_n/p_0, respectively. The result of the first transformation was by Cramér interpreted in terms of the transformed process on the transformed space: "The occurrence of the claims will constitute a stochastic process of the type known as a Poisson process," ([6], p. 19). Lundberg draws from the second example the conclusion: "Thus, in order to reduce a process with the intensity function $p_n\; v(t)$ to a time homogeneous process with the intensity function p_n/p_0 we have to take the function (66)," which defines s as here above, "as independent time parameter." Thus, both Cramér and O. Lundberg relate the transformed processes to the parametric spaces upon which they are defined, i.e. to T and to S respectively, and not, as in [4], to the parametric spaces before the mapping; both authors describe the transformation in each case in terms of one and only one process on the transformed space. The fact, that by Theorem 5 the introduction of a generalizing distribution into the transform of the second example leads to a Poisson process, does not motivate the principally different points of view in [4] and [3].

For the following theorem, the remark to the Theorems 1 and 2, and the deduction of the assertion 2) of the Theorem 4 shall be modified. Assume, in the remark referred to, that the change distribution is defined by the ch.f. ζ_v, where $v = v(t)$ is a function of the parameter t of the cPp : 1. If, in this case, $\tilde{\zeta}_v$ is defined as $\tilde{z}_t$ in the remark, the differential equation for $\tilde{\zeta}_v$ can either be deducted from the forward differential equation of the process, as referred to in the remark, or by a direct transformation of the variable in the equation for $\tilde{z}_t$. The differential equation for $\tilde{\zeta}_v$, then, takes the form $t\tilde{\zeta}_v'v'(t) = \zeta_v - \tilde{\zeta}_v$, the solution of which, after the transform of v according to $v = v(t)$, leads, after conversion, to $W(x, t)$ in (1a), as in the remark quoted. Let now in Theorem 4, $s(t)$ be equal to $s_2(t_2)$ say, and let $t_2^* = t_2^*(t_1)$ be t_2 as a function of another parameter, t_1 say, and let $v(t_1) = u(t_2^*)$, where u is defined in Theorem 4. Let ${}_3h_v$ be defined by the modified equation, corresponding to ${}_2h_u$ in Theorem 4, as defined by the original equation. By a deduction of ${}_3h_v$ by the method used for ${}_2h_u$ in Theorem 4, the elimination of the negative term leads to the relation $u'(t_2) = u(t_2)\, s_2'(t_2)\, v'(t_1)$ or $1/v = s_2' [t_2^*(t_1)]$, were s_2 is differentiated with respect to t_1. Then, ${}_3h_v = \tilde{\psi}_{\tilde{t}_2 1} h_{\tilde{s}}$, where, in this case, $\tilde{t}_2 = \tilde{t}_2(v)$, $\tilde{s} = \tilde{s}(v)$, and ${}_3h_v \subset \mathfrak{C}$ on vD_{t_2}. Consequently, $\overline{\psi}_{t_2 *} \subset \mathfrak{C}_{t_1}\ (t_1 D_{t_2})$. Here vD_{t_2}, $t_1D_{t_2}$ are the mappings of $D_{t_2} \subset T_2$ on $V = \{v\}$ and $T_1 = \{t_1\}$ respectively.

THEOREM 7

Let $\overline{Y}(t_1, t_2)$ be defined on the rectangle $T_1 \times T_2$, where $T_1 = (0, \zeta_1)$, $T_2 = (0, \zeta_2)$ by the ch.f. ${}_1P_0[t_1(1 - {}_2\overline{\psi}_{t_2})]$, where ${}_2\overline{\psi}_{t_2}$ is the ch.f. of the random function $Y_2(t_2)$ with the properties of $Y(t)$ in Theorem 5, the functions ${}_2\pi_{m_2}(\tau)$, ${}_2\bar{P}_{m_2}(\tau)$, and $\varkappa_2\overline{w}_2(\tau)$ being defined as these functions without index in Theorem 5, and where ${}_1P_{m_2}(t_1)$ is the probability distribution in the form of a cPd : 1, for a random function $M_1(t_1)$, not necessarily, admissible. Let ${}_1\pi_{m_1}(t_1)$ be defined by $-\,{}_1p_0^{(m_1+1)}(t_1) \,/\, {}_1P_0^{(m_1)}(t_1)$, as in Theorem 1, let ${}_1\pi_{m_1}(\tau)$, ${}_1P_{m_1}(\tau)$ be the functions obtained by the transformation of t_1 to τ in ${}_1\pi_{m_1}(t_1)$ and ${}_1P_{m_1}(t_1)$ respectively, and let $\varkappa_1\overline{w}_1(\tau)$ be the mean of ${}_1\pi_{m_1}(\tau)$ with respect to ${}_1\bar{P}_{m_1}(\tau)$. Finally, let D_{t_2} and ${}_3h_v$ be defined as in the modified Theorem 4, here above.

Then,

1) $\overline{Y}(t_1, t_2)$ is admissible and constitutes on $T_1 \times T_2$ a cPp : 2, the absolute probabilities being cPd : 2, and, simultaneously, non-elementary, compound Poisson bunch distributions in the alternative forms given in (10a-b) with the substitution of ${}_1P_{m_1}(t_1)$ for the Poisson expressions and ${}_2P_{m_2}(t_2)$ for $R_n(t)$, so that $R_0(t) > 0$, and the function in the form of $B_n(t)$ being a probability distribution.

2) If ${}_j\pi_{m_j}(\tau)/m_j$ for $j = 1, 2$ are uniformly convergent on finite intervals of τ, t_2 is a function of t_1, $t_2^* = t_2^*(t_1)$ say, with one-to-one correspondence, which defines a mapping of T_2 onto T_1 according to the definition before Theorem 5; this mapping transforms $\overline{Y}(t_1, t_2)$ into a random function, $\overline{Z}(t_1)$ say, with the ch.f.

${}_1P_0\,[t_1(1-\overline{\psi}^*_{t_2})]$, which on the intersection of T_1 with the domaine corresponding to the intersection of T_2 and D_{t_2} constitutes a cPp : 1, defined by the intensity function ${}_1\pi_{m_1}(t_1)$, and by a change distribution corresponding to ${}_3h_{v(t_1)}$.

Proof. By the section 7 $\overline{Y}(t_1, t_2)$ is admissible, and constitutes, on account of its ch.f., a cPp : 2 on $T_1 \times T_2$. It shall first be proved, that, if ${}_1P_{m_1}(t_1)$ is a Poisson distribution, Arfwedson's two remaining conditions for the function in the form of $B_n(t)$ in (10a), in this case called $B_{m_2}(t_1, t_2)$, being a probability distribution, are satisfied. By using the well-known relations for the probabilities of n changes in a cPp : 1 for $n > 0$ to the derivatives for $n = 0$ (cf. [3], (78)), and by the assumption, that ${}_2P_0(t_2) = \exp\,[-s(t_2)]$, the probabilities of m_2 changes in the convolution of m_1 ${}_2P_0(t_2)$ are easily calculated, for $m_2 = 0, 1, 2, 3$ and for a fixed value of m_1. $B_{m_2}(t_1, t_2)$ are, then, the means of the expressions obtained, and found to be functions or the derivatives of $s_2(t_2)$, and the moments about zero of a Poisson distribution of mean $t_1\,{}_2P_0(t_2)$. By the transformation of these moments into semi-invariants, and by the insertion of these means into the inequalities for $b_{m_2}\,(t_1, t_2)$, defined according to the remark in Theorem 3, which are conditions for $B_{m_2}(t_1, t_2)$ being a probability distribution, these inequalities are, by an easy calculation, reduced to $-s_2'' + s'^2 \geq 0$, and $s_2''' - 3s_2'\, s_2'' + (s_2')^3 \geq 0$, which, by the assumption that $s_2(t_2)$ is a ρ-transform, are satisfied. Thus, $B_{m_2}(t_1, t_2)$ is a probability distribution, if ${}_1P_{m_1}(s_1)$ is a Poisson distribution; this applies to every value in the argument of the structure function, if $P_{m_1}(s_1)$ is a cPd : 1. Thus, for the general case, $B_{m_2}(t_1, t_2)$ is the mean with respect to the structure function of probability distributions, so that $B_{m_2}(t_1, t_2)$ is a probability distribution, also if ${}_1P_{m_1}(s_1)$ is a cPd : 1, and, thus, the assertion 1) has been proved.

As by Theorem 1, ${}_1\pi_{m_1}(t_1)$ is the intensity function of a cPp : 1, even if $M_1(t_1)$ is not admissible, the absolute distribution functions of this cPp : 1 are equal to ${}_1P_{m_1}(t_1)$, so that the means of these functions are equal to the means of $M_1(t_1)$. Thus, the theorem for the expected number of changes in a cPp : 1 ([3], (63)), may be applied both to $M_1(t_1)$ and $Y_2(t_2)$, so that $t_j = \overline{t}_j(\tau) = \varkappa_j \int_0^{\tau} \overline{w}_j(u)\,du$ for $j = 1, 2$, if the condition of convergence in the assertion 2) is satisfied. Therefore, t_2 is a function of t_1 with one-to-one correspondence, which defines the mapping of T_2 onto T_1. The equivalence between $\overline{Y}(t_1, t_2)$ and $\overline{Z}(t_1)$ in the sense of section 1, follows from the discussion before Theorem 5. On account of the modifications of the remark to the Theorems 1 and 2, and of Theorem 4, the distribution functions of $\overline{Z}(t_1)$, are, on the domaine defined in the assertion 2), in the form of the absolute distribution functions of a cPp : 1. As $\overline{Z}(t_1)$ is admissible on account of the existence of the translator operators before the mapping, $\overline{Z}(t_1)$ constitutes the cPp : 1 defined in the assertion 2), which, thus, has been proved.

Remark. By the remark to the Theorem 3 here above, a bunch distribution with $R_0(t) > 0$ may be transformed into the T-form by the elimination of $R_0(t)$; in fact, by the mapping of T_2 onto $\Theta = \{0\}$, where $\theta = 1 - {}_2P_0(t_2)$, the bunch distribution of assertion 1) in the theorem is transformed into the T-form. The distribution obtained, satisfies, by Theorem 3, Condition A, so that the probability of multiple changes in $\overline{Y}(t_1, t_2)$ on an interval of small length, Δt_2, is of lower order than the order of Δt_2.

THEOREM 8

Let the random function $\overline{Z}(t_1)$ with the ch.f. ${}_1P_0\ [t_1(1 - {}_1\overline{\chi}_{t_1})\ ]$ on the domaine $(0, \zeta_{t_1})$ constitute a cPp: 1. If there exist a non-negative real parameter t_2, which is a function of t_1, $t_2^* = t_2^*(t_1)$, such that ${}_1\overline{\chi}_{t_1}$ can be written in the form of ${}_2\overline{\psi}_{t_2*}$ as defined in Theorem 4, the transformation of $\overline{\psi}_{t_2*}$ into $\overline{\psi}_{t_2}$, transforms $\overline{Z}(t_1)$ into a random function with the properties of $\overline{Y}(t_1, t_2)$ in Theorem 7, at least on the domaine corresponding to $(0, \zeta_{t_1})$.

Proof. By assumption ${}_1\overline{\chi}_{t_1} \subset \mathfrak{C}_{t_1}\ (0, \zeta_{t_1})$ the remainder of the proof is a consequence of a reversion of the proof of Theorem 7.

Remark. If in a homogeneous main group of a simple grouping both $Q_{m_1}(s_1)$ and $R_{m_2}(s_2)$ are cPd : 1, and if, in addition, $R_{m_2}(s_2)$ is in Thyrion's canonical form, the Theorem 7 holds for the main group. By the additional assumption, that also $Q_{m_1}(s_1)$ is in the canonical form, the cPp : 1 of assertion 2) in Theorem 7 satisfies the conditions of the process constituted by $Y_2(t_2)$. If these conditions are satisfied for all the main groups in an iterated grouping, where the number of such groups is distributed with a cPd : 1, a modified version of Theorem 7 holds for the head group. This modification implies, that the assumption of convergence shall be extended to the function in the form of $\pi_n(t)$ of Theorem 1, associated with the distribution of the number of main groups, and that the bunch distribution and the cPp of the assertion 1) are of the order 3, and a cPp : 2 is substituted for a cPp : 1 in the assertion 2).

Let the distribution functions of $Y(t)$ and $Z(t)$, of the Theorem 5 and 6 be denoted $F(y, t)$ and $G(z, s)$ respectively, then, $G(y, s) = F(y, t)$ for corresponding values of s and t, where one and only one value of s corresponds to a given value of t, and one and only one value of t to a given value of s. This coincides with Thyrion's definition for the equivalence between two random functions [1], which was quoted in section 1, and used in the preceding theorems. This does not, however, imply that the processes constituted by $Y(t)$ and $Z(s)$ are equivalent, one being a cPp : 1, and the other a Poisson process, even in the elementary case generalized, by a variable with the ch.f. $\bar{L}_1 \exp(i\eta)$; in the non-elementary case also the change distributions are different. This applies also to $\bar{Y}(t_1, t_2)$ and $\bar{Z}(t_1)$ of Theorem 7, being equivalent in the sense of section 1, and one constituting a cPp : 2, the other a cPp : 1, which also is generalized even in the elementary case, and the change distributions, in the non-elementary case, are defined by φ_{t_2} and ${}_2\psi t_2^*$ respectively. Similar view-points may be expounded with respect to the extension in the remark here above, and were earlier indicated in a heuristic deduction of a particular case of assertion 2) in Theorem 7 by the present author ([18], p. 62).

Literature Cited

(References are given in the order of first appearance in the body of this note)

[1] Thyrion, P., "Note sur une transformation des distributions généralisées", *Bull. Ass Roy. Act. Belg*, 1963.

[2] Rényi, A., *et al*, "On composed Poisson distributions I-II", *Acta Math. Acad. Sci. Hung.* 1-3, 1951-1952.

[3] Lundberg, O, *On Random Processes and Their Application to Sickness and Accident Statistics*, Uppsala, 1940, and (2nd ed.) 1964.

[4] Jung, J. and Lundberg, O., "Problems connected with compound Poisson processes", *The Lundberg Symposium*, Stockholm, 1968.

[5] BUHLMANN, H., "Kollektive Risikotheorien", *Bull. des actuaires suisses*, 65, 1965.
[6] CRAMÉR, H., "Collective risk theory, a survey from the point of view of the theory of stochastic processes", *Skandia Jubilee Volume*, 1955.
[7] PHILIPSON, C., "A method of estimating grouped frequencies", *Skand. Akt. Tidskr.*, 1959.
[8] DYNKIN, E. B., *Markov Processes* I-II, Berlin-Göttingen-Heidelberg, 1965.
[9] JUNG, J., "A theorem on compound Poisson processes with time-dependent change variables", *Skand. Akt. Tidskr.*, 1963.
[10] PHILIPSON, C., "A generalized model of the risk process and its application to a tentative evaluation of outstanding liabilities", *Astin Bull* III: 3, 1965.
[11] THYRION, P., "Notes sur les distributions par grappes", *Bull. Ass. Roy Act. Belg*, 1960.
[12] THYRION, P., "Extension de la théorie collective du risque", *The Lundberg Symposium*, Stockholm, 1968.
[13] ARFWEDSON, G., "Multipla försäkringsfall, kommentar till ett arbete av P. Thyrion" (Multiple insurance claims, comments on an article by P. Thyrion), *Försäkr. tekn. forskn. nämnden* (Council for Insurance Research) 25, Stockholm, 1968.
[14] CRAMÉR, H., *Mathematical Methods of Statistics*, Uppsala, 1945, and Math Ser., Princeton, 1946.
[15] THYRION, P., "Sur une propriété des processus de Poisson généralisés", *Bull Ass. Roy Act Belg.*, 1959.
[16] AMMETER, H., "A generalization of the collective theory of risk in regard to fluctuating basic probabilities", *Skand. Akt. Tidskr*, 1948.
[17] PHILIPSON, C., "Sur la transformation d'un processus de Poisson composé (selon les méthodes de P Thyrion et F. Esscher)", *Bull. Ass. Roy. Act. Belg*, 1964.
[18] PHILIPSON, C., "Quelques processus applicables dans l'assurance et dans la biologie", *Bull. Ass. Roy. Act. Belg.*, 1963.