A Review of the Collective Theory of Risk

Part I. Comments on the development of the theory

By Carl Philipson, Stockholm

1. Introduction and definitions

Let the random variable ξ be distributed with distribution function $G(\xi)$ of the continuous, discontinuous or mixed type, and the random variables x_ξ be distributed with the distribution functions $V_\xi(x)$ with the corresponding characteristic functions $\psi_\xi(\eta) = \int_{-\infty}^{+\infty} e^{\eta i x}\, dV_\xi(x)$, η being a real variable, and i the imaginary unit, and with the generating functions $\hat{\varphi}_\xi(z)$ defined by $\psi_\xi(-i \log z)$. If the range of x is restricted to the right semi-plane the lower limit of the integral in the definition of $\psi_\xi(\eta)$ can be replaced by zero.

A random variable x having the characteristic function $\hat{\varphi}_1[\psi_2(\eta)]$ is said to be (equivalent to) a variable x_1 *generalized* by the *generalizing* variable x_2 (Fr.: *variable x_1 généralisée par la variable généralisante x_2* [348, 180][1]).

If the distribution function $V_\xi(x)$ for different values of ξ are mutually independent, the distribution function corresponding to exp $[\int \log \psi_\xi(\eta)\, dG(\xi)]$, the integral being a Stieltjes integral taken over the range of ξ, shall be denoted $\Pi^*_{(\xi)} V_\xi(x)$. If ξ is allowed to take also non-integral values, this function shall be called the *convolution in the extended sense* of $V_\xi(x)$ over the range of ξ. For the opposite case, see the next paragraph.

If ξ assumes only integer values, the asterisk product of $V_\xi(x)$ over the range of ξ is the convolution of $V_\xi(x)$ for $\xi = 1, 2, \ldots, n$ and defines the distribution function of $\sum_{\xi=1}^{n} x_\xi$. The convolution of $V_1(x)$ and $V_2(x)$ can be written in the following well-known form, where, if, particularly, $V_2(0) = 0$, the limits of the integral may be replaced by zero and x.

$$\prod\nolimits^*_{(\xi=1,2)} V_\xi(x) = V_1(x) * V_2(x) = \int_{-\infty}^{+\infty} V_2(x-z)\, dV_1(z). \tag{1 a}$$

If, in addition, x_ξ are identically distributed with the distribution function $V(x)$ independently of ξ, the convolution takes the form $V^{n*}(x)$, where $V^{0*}(x)$ shall be taken equal to unity, and $V^{1*}(x) = V(x)$.

If n is a random variable assuming only integer values, and distributed with the probability distribution $\bar{Q}_n(\tau)$, τ being a parameter or parameter vector, and, if the variables $\bar{x}_{\tau_1}, \bar{x}_{\tau_2}, \ldots, \bar{x}_{\tau_n}$ are mutually independent, and identically distributed with the distribution $V(x)$ independently of τ_ν, then the distribution function of the sum of these variables for all $\tau_\nu \leq \tau$ is for each *given value of* τ defined by $\bar{F}(x, \tau)$ with the

[1] Numbers in square brackets refer to the list of literature in Part II, this journal 1968, 3–4. The sign § followed by a number refers to a section of this Part I.

corresponding characteristic function $\bar{\varphi}(\eta, \tau)$, given by the following relations (1 b), (1 c), where $\psi(\eta)$ corresponds to $V(x)$.

$$\bar{F}(x, \tau) = \sum_{n=0}^{\infty} \bar{Q}_n(\tau)\, V^{n*}(x), \tag{1 b}$$

$$\bar{\varphi}(\eta, \tau) = \sum_{n=0}^{\infty} \bar{Q}_n(\tau)\, \psi^n\, (\eta). \tag{1 c}$$

It is here anticipated, that the sum of $\bar{x}_{\tau_\nu}$, all $\tau_\nu \leqslant \tau$, is a random function of τ, $\bar{X}(\tau)$ say, which fulfils the conditions for the probabilities being well-defined (see § 2, below). (1 b) defines, then, a stochastic process constituted by the discontinuous random function $\bar{X}(\tau)$ with a discontinuous or continuous parameter or parameter vector τ. In this context τ shall always be used for the original parameter measured on the absolute scale (or absolute scales), which is (are) independent of any properties of $\bar{Q}_n(\tau)$; a function of τ shall always be denoted by a bar.

In the particular case, where $\bar{Q}_n(\tau)$ denotes the *probability distribution of the number n of claims* occurring in a group of insurances, when the parameter passes the domain from zero to τ in the parametric space, and, where $V(x)$ is taken to mean the conditional distribution function of the size of one claim at any parameter point, relative to the hypothesis that one claim has occurred at the parameter point, here called the *claim distribution*, the random function $\bar{X}(\tau)$ represents the total amount of the claims paid for in the group, when the parameter passes the domain considered. Then, the process constituted by $\bar{X}(\tau)$ is called the *risk process*, and if, particularly, $V(x) = \varepsilon(x-k)$, where k is a given constant and $\varepsilon(y)$, here and in the following context, the unity distribution, being equal to zero for negative and to unity for non-negative arguments, the process is said to be *elementary* and, in the opposite case, *non-elementary* [230]. In a non-elementary process $\bar{X}(\tau)$ is, always, a variable n generalized by the generalizing variable x, wherefore some authors have used the term a generalized process (Ge.: *ein verallgemeinerter Prozess*) for the non-elementary process. In French only the term *un processus généralisé* is used. The term generalized is, however, a wider concept, as in some cases, referred to in the following context, the characteristic functions of an elementary process can be written in the form $\varphi_1[\varphi_2(\eta)]$, so that the random function of an elementary process also may be a generalized variable, therefore the present author prefers the use of the terms in [230]. Some other authors have for a non-elementary process used the term a compound process (Ge.: *ein zusammengesetzter Prozess*), which here shall be kept for another purpose (see the next paragraph), cf. for the terminology [82, 152, 185, 268, 348].

Let $\bar{P}_n(\tau)$ be a probability distribution defined by the following Stieltjes integral. The distribution defined by $\bar{P}_n(\tau)$ is, in this context, called *a compound Poisson distribution* (Ge.: *eine zusammengesetzte Poisson-Verteilung*; Fr.: *une distribution de Poisson composée*). It might be remarked here, that in French the term composée is used also for a composed Poisson distribution to be defined in § 4 here below.

$$\bar{P}_n(\tau) = \int_0^{\infty} e^{-v\bar{t}(\tau)}\, [v\bar{t}(\tau)]^n\, d_v\, \bar{U}(v, \tau) \,/\, n!, \tag{1 d}$$

where $\bar{t}(\tau)$ will be defined below, and, where $\bar{U}(v, \tau)$ for every fixed value of τ is the distribution function of the non-negative variable v, fulfilling certain conditions for $v = 0$. It is here called the *risk distribution* (by certain authors also called the structure function). If in (1 b) $\bar{P}_n(\tau)$ is substituted for $\bar{Q}_n(\tau)$, the process defined by the expressions obtained is called a compound Poisson process, cPp, (Ge.: *ein zusammengesetzter Poisson-Prozess* [185]; Fr.: *processus de Poisson composé* [348]). Other authors use the terms mixed Poisson process (Ge.: *ein gemischter Poisson Prozess*) or weighted Poisson process (Ge.: *ein gewichteter Poisson-Prozess* [82]). All quotations given previously in this paragraph concern a particular case of (1 d), where $\bar{U}(v, \tau)$ is equal to $U(v)$ independently of τ. To differentiate between this case, and the general case, where $\bar{U}(v, \tau)$ may or may not depend on τ, the processes will be denoted with addition of the words *in the narrow sense* (i.n.s.), and *in the wide sense* (i.w.s.) for the particular and the general case respectively (Fr.: *au sens restreint, au sens large*, respectively). The *Poisson process* is a cPp i.n.s. where, particularly, $U(v) = \varepsilon(v - \gamma_1)$, γ_1 a given positive constant, and a *Polya process* a cPp i.n.s., where, particularly, $dU(v)$ is represented by a Pearson Type III frequency curve beginning at origo.

It is often advantageous to transform the parameter τ (see the second paragraph of § 4 here below), for a cPp i.n.s., by the relation $t = \bar{t}(\tau)$, and for a cPp i.w.s. by this transformation in the Poisson expression, and by the substitution of $U(v, s)$ for $\bar{U}(v, \tau)$, where the relation between s and τ shall be determined with regard to the form of $\bar{U}(v, \tau)$. After the transformation the functions appearing in (1 d), and in (1 b), (1 c) after the insertion of (1 d) will be denoted without a bar as functions of t or t, s. The transformation of τ leads to simple expressions even if the assumptions are extended by the assumption, that the claim distribution depends on the parameter point for the occurrence of the claim, denoted, after the transformation of τ, by $V(x, t)$ with the corresponding characteristic functions $\psi(\eta, t)$. If, particularly, t is one-dimensional, and $V(x, t)$ is continuous in t, this assumption leads to the following expressions (1 e), (1 f), which define a cPp i.w.s. with the claim distribution $V(x, t)$.

$$F(x; t, s) = \sum_{n=0}^{\infty} P_n(t, s)\, W^{n*}(x, t), \tag{1 e}$$

$$\varphi(\eta; t, s) = P_0\{t[1 - \chi(\eta, t)], s\}, \tag{1 f}$$

where $W(x, t) = 1/t \int_0^t V(x, u)\,du$, and $\chi(\eta, t) = 1/t \int_0^t \psi(\eta, u)\,du$, which under mild conditions of regularity are consistent expressions with $\chi(\eta, t)$ being the characteristic function corresponding to $W(x, t)$. The modifications of (1 e), (1 f) in cases, where t is a vector, and where $V(x, t)$ is discontinuous in t, are self-evident.

The *conditional probability of the occurrence of ν claims*, if t is one-dimensional, in the interval (t_1, t_2) relative to the hypothesis, that n claims have occurred in the interval $(0, t_1)$, $t_1 < t_2$, is denoted $P_{n,\,n+\nu}(t_1, t_2)$. If τ, $\tau + d\tau$ and 1 are substituted for t_1, t_2 and ν, the conditional probability in this case reduces for a cPp i.n.s. and for a wide set of cPp i.w.s. to $\bar{p}_n(\tau)d\tau + o(d\tau)$. Here $\bar{p}_n(\tau)$ is called the *intensity function* of the process.

If t is one-dimensional and $\bar{p}_n(\tau)/n$ is uniformly bounded for all n, the mean of $P_n(t)$ is given by the following expression, $\gamma_\nu(s)$ being the νth semi-invariant of $U(v, s)$.

$$\gamma_1(s)\, t = \gamma_1(s)\, \bar{t}(\tau) = \int_0^\tau \left[\sum_{n=0}^{\infty} \bar{p}_n(u)\, \bar{P}_n(u) \right] du. \tag{1 g}$$

The conditional mean of ν, i.e. with respect to $P_{n,n+\nu}(t_1, t_2)$ can be written in the following form for a cPp i.n.s.

$$(t_2 - t_1) p_n(t_1). \tag{1 h}$$

By the normalization of $U(v, s)$, so that, in a cPp i.n.s., the mean becomes equal to unity and, in a cPp i.w.s., the mean becomes equal to s, (1 g) and (1 h) will be simplified.

2. The risk process

In [111] Cramér has treated the risk theory from the point of view of the theory of stochastic processes. He has *inter alia* proved for the risk process *in its classical form*, to be defined below, that the sum of the claim amounts in the interval $(0, t)$ of the one-dimensional, transformed parameter, $X(t)$, is—if confined to a *restricted space* to be defined below—a random function associated with well-defined probabilities, induced by the measures of the variables $\omega = \{t_\nu, x_\nu;\ \nu = 1, 2, \ldots\}$, where t_ν, x_ν for the νth claim are the parameter point of occurrence, and the amount of the claim respectively; ω is said to belong to a *reference space* of an enumerable number of dimensions. Thus, each point ω represents the actual development of the claims in one particular case, which corresponds to one and only one individual function $X(t)$ called a *sample function* or a *realization* of the process. Thus, the reference space is mapped on the restricted functional space according to $X(t, \omega) \to X(t)$, where $X(t, \omega)$ for any fixed value of ω is a sample function, and for any fixed value of t, $= t_0$ say, represents, for different values of ω, different values of the random variable $X(t_0)$ distributed with the distribution function $F(x, t_0)$. The *restriction of the functional space* is defined by allowing only for such sample functions, which are relevant in the risk theory, i.e. step-functions with the discontinuity points t_ν, defined here above, and in the intervals between consecutive such points of constant value. The *classical form of the risk process* is taken to mean a Poisson process with a claim distribution equal to $V(x)$ independently of the parameter. A great part of [111] concerns this form, Cramér has, however, indicated an extension to the Polya process with reference to [230], and another extension to a Poisson process with a claim distribution dependent on t with reference to [145].

Later in this review (§ 5) it shall be referred to further extensions of the classical form. The discussion reviewed in the previous paragraph has been extended [230] to include all processes, which fulfil the conditions for the validity of Markov's differential equations (l.c., p. 33), by Ove Lundberg and to include a wide set of cPp i.w.s. by the present author [303].

Owing to the strong connections between the risk theory and the stochastic process theory, which besides in [111] have been elucidated in [230, 40], references to some studies into the general stochastic process theory have been included in the list of literature in Part II of this review, even if these studies do not particularly deal with the risk process [39, 97, 112–113, 115, 117–120, 135–138, 152, 174, 203–204, 234, 364]. Some investigations into pure mathematics, the results of which have been used by authors dealing with the risk theory, have also been included in the reference list [4, 36, 45, 63, 149, 259, 326]. Of all the items in the list of literature, in total 365 items, thus, 26 items do not directly concern contributions to the risk theory.

The collective theory of risk was, originally, created by Filip Lundberg. A great part of his contributions were published before 1930, thereinafter, he published two papers a few years later, his first paper was published in 1903 [223–229]. Cramér reviewed and developed his theory in 1919, 1926 and 1946 [98–99, 107]. According to Cramér, Lundberg anticipated ideas, which later were propounded in the general theory of stochastic processes; the modern development of the general theory started in the early thirties with two important papers by Kolmogoroff [203–204], and was developed by Bartlett, Cramér, Doob, Feller, Gnedenko, Khintchine and many others. As Cramér stated in [107], Filip Lundberg's theory is to be considered an important particular case of the general theory of stochastic processes, the early contributions to the risk theory can, therefore, be considered an auspicious pioneer work for the knowledge of stochastic processes, accomplished a long time before the general principles of the theory of such processes had been established. On the other hand, the modern development of the general theory has deepened our understanding of the problems involved in the risk theory, and facilitated the rigorous deduction of the results in this theory, by giving more satisfactory tools for the solution of such problems (cf. [40, 111, 107]).

In [82] Bühlmann acknowledges Filip Lundberg's contributions, by using modern terms, saying that he investigated stochastic processes with independent increments, and with sample functions—see the first paragraph of this section—in the form of step-functions, a long time before such processes had been rigorously deducted. Bühlmann places in this sense Bachelier beside Lundberg; Bachelier had in 1906 introduced a mathematical theory for the Brownian movement of molecules (Théorie des probabilités continues, *Journ. Math. Pures et Appliquées*). At the Astin Colloquium in Arnhem, 1966, Borch pronounced in an oral contribution, that Bachelier had stimulated a continued study of the ideas propounded by him, and that this stimulation had led to numerous new important contributions to the field of these ideas. Borch added, that this should only to a limited extent apply to Filip Lundberg.

It is true, that Lundberg based most of his developments on assumptions, which lead to the classical form of the risk theory (see § 4 here below). This approximation of the reality has been used by several other authors, and has, in fact, led to very remarkable results. A great part of these results, particularly with regard to the ruin theory (see §§ 9 and 10 below), should have been very difficult to reach, if more realistic models had been introduced from the beginning. A certain criticism of this simplification of the distribution functions defining the risk process has been given by Almer [3–7]. Among other critiques the works by Giovanni and Giuseppe Ottaviani, Campagne, Tedeschi and de Finetti may be mentioned in the first hand; their criticism was mainly directed against the criterion for the decisions by an insurance company especially with regard to the reinsurance policy. Also Borch criticised this criterion, and suggested new methods for the formulation of the decision problems (see § 10 here below). With respect to the distribution functions of the risk process, numerous papers have been based on more realistic assumptions than those leading to the classical form (see § 5 here below). In a few cases such assumptions have been applied to the ruin theory and to the decision criterion of this theory. As is seen above, 339 items in the reference list deal with the risk theory, and all these items are, more or less, based on the fundamental ideas introduced by Filip Lundberg, and later developed by him and by his followers. This statement holds even for the critiques of his theory, also for the papers by Borch, as in these papers the fundamental ideas have been accepted, though some parts of the theory have been modified. It might here be remarked that Borch has in some of his papers used distribution functions of the total claim cost, which have been based on an even less realistic model, than the classical form of the risk theory. In the opinion of the reviewer the discussion of this paragraph affords a strong argumentation for the statement, that Filip Lundberg's contributions have to a very wide extent stimulated the continued study of the risk theory, and that this has led to numerous valuable contributions to the problems within the scope of this theory.

3. Earlier reviews of the risk theory and the list of literature in Part II of this review

Besides Cramér's survey in [111], which has been referred to in the previous section, a reference list was published by Ammeter in 1956 [15], and a survey of the risk theory and other problems within non-life insurance in 1959 by Ammeter, Depoid & de Finetti [19]. General surveys of the risk theory were given by Wilhelmsen in 1955 [358], by Segerdahl in 1959 [323], by Philipson in 1961 [281], by Thyrion in 1963, 1965, 1967 [349, 351–352], and, in 1967, by Bühlmann [82], and by Kupper [210]. Further, a book on the risk theory, not included in the list of literature, is under preparation by Beard, Pentikäinen & Pesonen.

In addition to these reviews the Astin Bulletin for the whole time of its existence, and *Skandinavisk Aktuarietidskrift* from the year 1961 inclusive have been consulted for all contributions to the risk theory published in these journals during the periods mentioned. Further, the reference lists with items from other journals published in the Scandinavian journal in the same period have been read through, as well as reference lists published by different authors. The reviewer has endeavoured to render a list of literature as complete as possible; he regrets, if, notwithstanding his endeavours, imoortant papers should have been neglected.

4. Fundamental assumptions for the risk process

The deduction of a Poisson process is, generally, based on three assumptions, namely (i) homogeneity with respect to the parameter, (ii) homogeneity in space, and (iii) rarity of multiple events. (i) is often formulated stationarity of the increments, (ii) independency of the increments, and (iii) the probability of the occurrence of more than one event in an infinitesimal domain of the parametric space, $d\tau$, is of lower order than the order of $d\tau$ [111, 114]. Other formulations of (iii) are found in [2, 191, 307, 310], published by Rényi *et alia*.

It shall be remarked here, that there exist cases, for which (ii) and (iii) hold, while (i) does not hold with respect to τ. This implies that $\bar{p}_n(\tau)$ is equal to $\bar{p}(\tau)$ independently of n, in this case (1 g) reduces to $\gamma_1 \int_0^\tau \bar{p}(u)du$, so that the probability of one claim in an interval of length dt is equal to $\gamma_1 dt$, and the process with the transformed parameter t fulfils (i) with respect to t. Then, the process before the transformation may be called a Poisson process, heterogeneous with respect to τ. Such processes have been included in the definition of the classical form of the risk theory [111].

Bühlmann [82], seems to have anticipated a theorem (according to a letter from Bühlmann to the reviewer, it should be proved in [83]), which should imply, that (iii) is a consequence of (i), (ii) and of the properties of the sample functions in the restricted space; (iii) should, thus, according to Bühlmann, not be necessarily included in the conditions for the process being a Poisson process. In [114] Cramér gives an example of a case, where (iii) is not fulfilled which leads to a cPp i.n.s. with a risk distribution of the discontinuous type. In this case, however, (ii) is not strictly fulfilled. Cramér refers in [114] to a case, where the probability distribution of the length of the mutually independent time intervals between consecutive discontinuity points is given in a general form, which for the Poisson process is exponential. This process has been called a "process of limited after effects" and has been introduced by C. Palm. According to Goldmann (*Ann. Math. Soc. 38*, 3, 1967), there exist processes with Poisson-distributed number of events, for which (ii) does not hold, so that they are not Poisson processes (cf. also [89, 91, 201, 313, 329, 341]).

According to Bühlmann, who has published his thesis [77] on exchangeable variables, a theorem is given by de Finetti for such variables (by de Finetti called *numeri aleatori equivalenti* [154], cf. also [172]), which should lead to the following fundamental assumption for general cPp i.n.s. (instead of (i) and (ii)). For an arbitrary number of non-overlapping parameter intervals of the same length the amount of the claims occurring in each of the intervals can be arbitrarily exchanged without change in the probability distributions of the process. It seems likely, that it should be possible to find a similar condition leading to a cPp i.w.s. Cramér remarks in [114] that, if (i) is given up, the theory of so-called harmonizable processes defined by a spectrum distribution with correlated increments may lead to better understanding of the risk process, in this case with non-stationary increments. As the Polya process can be deducted both from the Polya-Eggenberger urn scheme [230], and from the Lexis urn scheme [9, 230] such a process may be the consequence of heterogeneity either in space or in time or in both space and time. An analysis of the effects of these different types of heterogeneity has been given in [230].

Rényi *et alia*, quoted here above, have introduced the concept *composed Poisson processes*, which fulfil the condition (ii), and are, therefore, principally different from the *compound Poisson processes*, which, with exception of the Poisson process, have dependent increments. The composed Poisson processes have been discussed in [289, 301] by the reviewer. The characteristic functions defining a cPp can be transformed into a form similar to the form of these functions for a composed Poisson process.

5. Extensions of the classical form of the risk process

The claim distribution of the classical form is, by definition, independent of t. Very often, however, the claim distribution actually varies with t, as proved by extensive statistics. Esscher extended the Poisson process with regard to a claim distribution, dependent on the parameter point for the occurrence of the claim, in 1932 [145]. This was extended to a Polya process in 1957 [272], and to a general cPp in 1965 [299]. The last-mentioned result was reached independently of a theorem published by Jung in 1963, [195], according to which a symmetric function of the increments of a random function attached to a cPp i.n.s. with t-dependent claim distribution is distributed with a distribution function of similar form, as that obtained in [299], and given in (1e) here above.

According to [19], Dubois was the first author, who accounted for a dependency between the events in non-overlapping time intervals, for his calculation of the variance of $F(x, t)$, in a paper published in 1936 [139] (cf. also [142]). In [19] the compound Poisson distribution with t-independent risk distribution was said to have been used by Ammeter with reference to his paper in 1949 [10]. No reference is made in [19] to [230], which in the first edition was published in 1940 by Ove Lundberg. [230] contained a systematic study of the cPp i.n.s. with particular consideration of the Polya process with references to [305, 144, 176, 250, 203]. Similar models were introduced by Ammeter in 1948 [9], independently of Ove Lundberg, and generalized by him in 1949 [10]. In 1954 [13] Ammeter gave Lundberg credit for his priority.

Both Ove Lundberg and Ammeter [230, 9] deducted the limiting distribution, as t tends to infinity, for the variable x/t, if x is distributed in a compound Poisson distribution with t-independent risk distribution $U(v)$, and found this limiting distribution to be in the form of $U(v)$, provided that $c_1 \neq 0$. Lundberg has also for this case given an asymptotic expansion of $F(x, t)$ for a Polya process in terms of $U(v)$, and for $c_1 = 0$ in an Edgeworth series for $F(x, t)$ of a cPp i.n.s. Ammeter proved that, if, for a Polya model, $t\gamma_2/\gamma_1^2$ remains finite, when t tends to infinity, the limiting distribution is in the form of the normal distribution function, and he has also expanded $F(x, t)$ in this case in terms of the normal distribution function after the transformation according to Esscher (see § 7 here below). The reviewer [293, 303] proved that the intensity function of a cPp i.w.s. is in principle proportional to the volume of the population, as the intensity function of a main heterogeneous group is equal to the weighted sum of the intensity functions of homogeneous sub-groups. Consequently, an increase in t may by (1g) be due either to an increase in the volume of the population or to an increase in τ (or both). The limiting distribution, when t tends to infinity only due to the increase in volume, is in the form of the normal distribution, and, when the increase of t is only due to an increase in τ, in the form of $U(v, s)$. The condition in the former case can be replaced by a condition of boundedness for the functions $t^{\nu-1}\gamma_\nu(s)/\gamma_1^\nu(s)$, which for the Polya process reduces to Ammeter's condition for this limit passage. An asymptotic expansion of $\psi(\eta; t, s)$ in an Edgeworth series has been given for this case [293].

Arfwedson [35] extended the Poisson process by the omission of (iii) in § 4, and found, that the extension rendered the same result at the end of time-intervals of finite length as Ammeter's model in the case, where $t\gamma_2/\gamma_1^2$ is bounded even for infinite values of t. It has been proved [271], that this model can be interpreted as a transform of either a sequence of Polya processes or of Poisson processes defined only for discrete parameter points.

Hofmann [185] introduced a wide subset of cPp i.n.s. by defining $P_0(t)$ as the solution of the following differential equation, where k, q and a are constants $k>0$, $q>0$, $a \geqslant 0$.

$$ky' + q(1+t/k)^{-a}y = 0. \tag{5a}$$

The present author introduced the *extended Hofmann processes* by defining $P_0(t)$ as a product of the solutions of equations in the form of (5a) with, not necessarily, different values of k, q, a [280, 282, 290, 297].

The study of the cPp i.n.s. [278–279] led the present author to the introduction of the cPp i.w.s., as defined in § 1 [284, 290–291, 293, 296]. A wide sub-set of these processes was introduced in [303, 304] under the name of cPp of the order ν, $\nu = 1, 2, \ldots$, (cPp: ν) which will be defined below. Pesonen and Jung have discussed the cPp i.w.s. in recent manuscripts to the Lundberg symposium.

Some of the processes exemplified by Bartlett [40], and the processes studied by Matern [238] are cPp i.w.s. This can also be said of Ammeter's model with bounded $t\gamma_2/\gamma_1^2$. Thyrion introduced [345, 347–348] a very wide class of distributions, the *distributions in bunches* (*par grappes*), and *in bunches of bunches* (*par grappes de grappes*) defined by characteristic functions in the following general form.

$$\varphi(\eta) = \hat{\varphi}_1\{\hat{\varphi}_2[\hat{\varphi}_3 \ldots (\varphi_\nu(\eta))]\} \quad (\nu \geqslant 2), \tag{5b}$$

where $\hat{\varphi}_j(z)$ are generating functions of integer valued variables, $j = 1, 2, \ldots, \nu-1$. Originally $\varphi_\nu(\eta)$ was the characteristic function of such a variable, but may also be any characteristic function for $\nu \geqslant 2$. If particularly, $\varphi_\nu(\eta)$ is the characteristic function of a τ-dependent Poisson variable, or of a generalized such variable, (5b) can be considered the characteristic functions defining a cPp, which, if at least one of the $\hat{\varphi}_j(z)$ depends on t, is a general cPp i.w.s. If in (5b) all $\hat{\varphi}_j(z)$, $j = 1, 2, \ldots, \nu-1$ are in the form ${}_jP[t_j(1-z)]$, and, if $\varphi_\nu(\eta)$ defines a cPp i.n.s., (5b) defines a cPp: ν. This implies, that $X(t)$ is a generalized variable even for an elementary process (cf. the remark to the terminology in § 1 here before).

Extensions to processes with parameters of more than one dimension have been dealt with in [40, 238, 270, 273, 298, 342].

In [299] the *extended risk process* was introduced, taken to mean a process, where the occurrence of the accidents and the extent of the damage caused by them, as well as the development of the actual payments for a claim during the period, when it is outstanding, is accounted for. For the deduction the theory of cPp i.w.s. was used (cf. also [300]).

Almer introduced [3–7] a very general model for the risk process. His fundamental assumption can be formulated by saying that "behind" the risk process, there exists another process constituted by a large, but finite number of risk situations, called *risk elements*. Each risk element is supposed to be associated with a certain probability of inducing a claim, and a certain claim distribution. In [275] some of the deductions were based on this model. In [298] Almer's model was modified by the present author, by the assumption that the occurrence of a risk element was associated with a change in a random function of a two-dimensional parameter (time and geographical space). This random function was, further, assumed to be subject also to changes caused by changes in environmental conditions, and the occurrence of an accident was supposed to be correlated with the random function just defined. The extent of the damage of one accident could be correlated either with the same, or with a similar random function. Also this theory could be interpreted in terms of the cPp i.w.s.

6. Particular forms of the claim distribution

Cramér used in [111] for the numerical comparison of different approximations for $F(x, t)$ in one example an exponential distribution, and in another the form $k_1 e^{-\beta_1 x} + k_2(x+b)^{-\beta_2}$ for $x \leq 500$, and equal to zero for $x > 500$, as derived by Esscher from Swedish non-industry fire experience 1948–1951. For the deduction of the ruin function (see § 9 here below) a form indicated by Täcklind [354], defined in the following sentence, was used in [111]. The claim distribution was in this case supposed to be arbitrary in the negative semi-plane, with the restriction of having a finite, absolute mean over this domain, and given by an *exponential polynomial*

$\sum_{n=1}^{r} l_n(1-e^{-\beta_n x})$ in the positive semi-plane.

Almer, who mainly dealt with non-life insurance, for which $V(0)=0$, proved for non-negative values of x, that an upper and lower approximation in the form of exponential polynomials can, with any desired precision, be found for any distribution function [3]. He, further graduated extensive statistics for different time periods mainly from Swedish motor insurance, with exponential polynomials containing three terms for certain periods, and four terms for other periods of time, giving deviations of at most one to two per cent. Almer used this form for the deduction of approximation formulae for $F(x, t)$ (see § 7), this was also done by Hovinen [186–188], and Pesonen [268] both for such deductions and for their numerical investigations. Bohman and Esscher [67] modified this form, by replacing, for higher values of x, the exponential polynomial with actually found frequencies in discrete points spread out over small intervals about such points. This modified form was numerically compared with Swedish experience in life insurance 1957–1961, in third party liability motor insurance 1957, and in fire insurance 1948–1957, differentiated with regard to industry and non-industry. The agreement between graduated and actual values was very satisfactory.

Benckert [48] studied the application of a log-normal distribution to the claim distribution. A more systematic treatment of the effect of different forms of $V(x)$ on the risk process with particular regard to excess of aggregate loss reinsurance was given by Benktander & Segerdahl [50], and by Benktander [51], cf. also [170, 182]. Finally, Thyrion [350] introduced a general class of functions, the compound exponential functions, which is a particular case of (7h) of the next section, with $V_\xi(x) = 1 - e^{-\xi(x-c)}$. A great part of the forms for $V(x)$ used by other authors belong to the class defined by Thyrion. The exponential polynomials are compound exponentials, with $c = 0$ and ξ integer valued. If $V_\xi(x) = 1 - e^{-\xi(x-c)/a}$, and the distribution function of ξ defined by an incomplete Γ-function, $V(x)$ used in [236] is obtained, if $c = 0$, and the Pareto distribution analysed in [50], if $c \neq 0$. Thyrion also proved that this class also contained the functions $[V(x)]^a$, and $V[g(x)]$ for $g(0) = 0$, $g(\infty) = 1$; $(-1)^n g^{(n)}(x) \leq 0$, where $V(x)$ is a compound exponential. A particular form is obtained by taking $g(x) = k_1 x^{k_2}$, so that $V(x)$ has the moments $\mu_r = k_1^{-r/k_2}\Gamma(r/k_2 + 1)$, which, eventually, can be used in life insurance technics.

7. Transforms of approximations for $F(x, t)$ and $F(x; t, s)$

Ove Lundberg proved that an elementary Polya process being heterogeneous with respect to τ and t, is homogeneous with respect to a transformed parameter equal to $-\log P_0(t)$ (230). Ammeter [9] transformed $F(x, t)$ for a non-elementary Polya process, by using this transform of the parameter, and further, a transform of $V(x)$. A similar transform was used by Campagne [91]. These transforms are particular cases of the transformation of $F(x, t)$ for cPp i.n.s. given by Thyrion [345, 347–348]. The characteristic function corresponding to the transform is in the form of the characteristic function of a generalized Poisson variable (even for the transform of an elementary process) by a generalizing variable with the probability distribution $q_n(t) = [1/(-\theta(t))]\cdot[(-t)^n/n!]\theta^{(n)}(t)$, where $\theta(t) = \log P_0(t)$, in the elementary case. It is proved that $q_n(t)$ can be obtained by a truncation of a distribution defined by (1 d) with $\bar{U}(v, \tau) = U(v)$, independently of τ, if $\theta(t)$ is bounded. The characteristic function corresponding to the transform can be written in the form

$$\exp\left\{\theta(t)\left[1 - \sum_{n=1}^{\infty} q_n(t) z^n\right]\right\}, \quad \text{where} \quad z = e^{ki\eta},$$

and equal to $\psi(\eta)$, in the elementary and non-elementary case respectively. (7 a)

(7a) has been extended [297] to include i.a. non-elementary cPp i.w.s., where $\bar{U}(v, \tau)$ defines a generalized extended Hofmann-process (see § 5 here above). Thyrion gave in the papers quoted numerous examples of this and other transforms. One of these transforms was obtained by the substitution of the distribution function of a generalized Poisson variable with the mean of the number of events given by $-\lim_{t\to\infty}\theta(t)$, supposed to be finite, for $U(v)$ of a cPp i.n.s.

Esscher introduced in 1932 a transform of $F(x, t)$ for a Poisson process [145], which was in 1963 extended by him to any distribution function $F(x)$, for which the corresponding characteristic function $\varphi(\eta)$ is known [146]. Let the mean of $F(x)$ be denoted μ_1, and its transform, here called the *Esscher transform*, be denoted $\mathbf{F}(x)$, and the distribution function of the standardized variable, $z=(x-\boldsymbol{\mu}_1)/\sqrt{\boldsymbol{\mu}_2}$, $\mathbf{F}_0(z)$, here $\boldsymbol{\mu}_k$ is the kth semi-invariant of $\mathbf{F}(x)$. h is supposed to be a quantity, $-H_1<h<H_2$, such that the integral representing $\varphi(-ih)$ converges. The transform is, then, defined by the following relation

$$d\mathbf{F}(x)=e^{hx}\,dF(x)/\varphi(-ih) \tag{7b}$$

and, if h is the single root in the interval $-H_1<h<H_2$ of the equation

$$x=\boldsymbol{\mu}_1=\frac{\partial\varphi(-ih)}{\partial h}, \tag{7c}$$

the following relations hold.

$$F(x)=\begin{cases} J(-\infty,\ 0;\ x) & \text{for}\quad x\leqslant\mu_1,\\ 1-J(0,\ +\infty;\ x) & \text{for}\quad x>\mu_1,\end{cases} \tag{7d}$$

where $J(u_1, u_2; x)=e^{-hx}\varphi(-ih)\displaystyle\int_{u_1}^{u_2} e^{-uh\sqrt{\boldsymbol{\mu}_2}}\,d\mathbf{F}_0(u)$.

By a suitable choice of $\mathbf{F}_0(z)$ (7d) can be used for the approximation of $F(x)$.

Before the publication of [146] Esscher prepared a manuscript (not published), which dealt with the particular case, where $F(x)$ had the form of $F(x, t)$ of a cPp i.n.s. The transform, $\mathbf{F}(x, \mathrm{t})$ say, could be written in the same form as $F(x, t)$ with the substitution of the following expressions for t, $U(v)$, $V(x)$ respectively.

$$\mathrm{t}=t\int_0^\infty e^{hu}\,dV(u);\quad \mathrm{U}(v)=\frac{\displaystyle\int_0^v e^{-u(t-\mathrm{t})}\,dU(u)}{\displaystyle\int_0^\infty e^{-u(t-\mathrm{t})}\,dU(u)};\quad \mathrm{V}(x)=\frac{\displaystyle\int_0^x e^{hu}\,dV(u)}{\displaystyle\int_0^\infty e^{hu}\,dV(u)}. \tag{7e}$$

The limits of $\mathbf{F}_0(z, \mathrm{t})$ for z and t tending to infinity were deducted in the general case in [65] and [292] respectively, and the limit, when t tends to infinity, for a Polya process in [67]. With regard to the limits obtained, $\mathbf{F}_0(z, \mathrm{t})$ was chosen in the form of a normal distribution function, and of an incomplete Γ-function in the application of (7d) for the approximation of $F(x, t)$ defining a Poisson and a Polya process [67]. Esscher's method of approximation was modified by Pesonen [266] in such a way, that the solution of (7c) should be independent of $\boldsymbol{\mu}_1$, and determined only by x for an actual case.

Bohman introduced another method [63, 67], the C-method, for the approximation of a distribution function $F(x)$ corresponding to a given characteristic function $\varphi(\eta)$. Let, for $\nu=1, 2$, $\chi_\nu(\eta)=C(\eta)+(-1)^\nu 0.42iC'(\eta)$, where $C(\eta)=0$ for $|\eta|\geqslant 1$ and equal to $(1-|\eta|)\cos \pi\eta+(1/\pi)\sin|\pi\eta|$ for $|\eta|<1$, and $\boldsymbol{\varphi}_\nu(\eta)=\chi_\nu(\eta/T)\ \varphi(\eta)$. Then, $\boldsymbol{\varphi}_\nu(\eta)$ correspond to the "improper" distribution functions $\mathbf{F}_\nu(x)$, taken to mean that $d\mathbf{F}_\nu(x)$ are, not necessarily, non-negative for all x, and that the integrals $\int_{-\infty}^{+\infty} d\mathbf{F}_\nu(x)$ are, not necessarily, equal to unity. It has been proved, that the following inequalities and "conversion"-formulae hold.

$$\mathbf{F}_1(x)\leqslant F(x)\leqslant \mathbf{F}_2(x), \tag{7f}$$

$$\mathbf{F}_\nu(x)=\tfrac{1}{2}-\int_{-T}^{T}\frac{e^{-iux}}{2\pi iu}\boldsymbol{\varphi}_\nu(u)\,du \quad (\nu=1,2), \tag{7g}$$

which may be evaluated by using numerical integration according to Simpson's rule. An approximation of $F(x, t)$ is, then, given by the arithmetical mean of (7g) for $\nu=1, 2$, and the approximation error implied, by the sum of half the difference of (7g) for $\nu=2$ and 1, and the error involved in the numerical integration. In [67] $F(x, t)$ of a Poisson, and of a Polya process with the claim distributions refferred to in § 6 here above were evaluated for different values of t and of x according to the C-method and according to other methods, the Esscher method inclusive. The results from the other methods were compared with those from the C-method with due regard to the estimated errors in the last-mentioned results. Further, according to similar formulae the corresponding stop loss risk premiums were evaluated by different methods and compared. In [147] Esscher gave numerical illustrations for other cPp i.n.s.

A transform of $F(x, t)$ with the claim distribution defined by the following relation (7h) shall be given for the general case in the form of (7i) below [304]. This transform may be called the *convolution transform*.

$$V(x,t)=\int V_\xi(x,t)\,dG(\xi), \tag{7h}$$

where the integral shall be taken over the range of ξ, ξ being a random variable distributed with a distribution function $G(\xi)$ of the continuous, discontinuous or the mixed type. Let the distribution functions defining Poisson processes with $V(x, t)$, $V_\xi(x, t)$ as claim distributions be designated by ${}_\varepsilon F(x, t)$, ${}_\varepsilon F_\xi(x, t)$ respectively, then, the convolution transform of ${}_\varepsilon F(x, t)$ can be written $\Pi^*_{(\xi)}{}_\varepsilon F_\xi(x, t\,dG(\xi))$, where $\Pi^*_{(\xi)}$ has been defined in § 1, for ξ integer valued $\Pi^*_{(\xi)}$ reduces to the convolution of a number of distribution functions. From this relation, the following expression for the convolution transform of $F(x; t, s)$ for a cPp i.w.s. with $V(x)$ defined by (7h), is immediately obtained.

$$F(x;t,s)=\int_0^\infty\left[\prod_{(\xi)}^{*}{}_\varepsilon F_\xi(x, vt\,dG(\xi))\right]d_v\,U(v,s). \tag{7i}$$

The results in (1e), (1f) are particular cases of (7h). In the case, where $V(x) = V_1(x) = 1 - e^{-x}$, $G(\xi) = \varepsilon(\xi - 1)$, Cramér has given an exact expression for ${}_\varepsilon F(x, t)$ in the form of a Bessel function [111]. Esscher deducted in [145] a relation, which leads to the convolution transform of ${}_\varepsilon F(x, t)$ in the particular case, where $V(x)$ is defined by (7h) with ξ integer valued. Almer [5–6] used this transform for the case where $V_\xi(x)$ are different exponential functions for different integer values of ξ, and inserted thereinafter, Bessel functions according to Cramér for ${}_\varepsilon F_\xi(x, t)$. He suggested, then, that these Bessel functions should be approximated by a few terms of their expansions according to Hankel. The calculation of the convolution of these approximate expressions can be performed without material computation work. Pesonen [268] derived approximation formulae for ${}_\varepsilon F(x, t)$ with $V(x)$ in the form of exponential polynomials in a similar way as Almer. Bohman and Esscher discussed a convolution transform of ${}_\varepsilon F(x, t)$, where $V(x)$ was given as the sum of two exponential terms, and one term defined by the unity distribution $\varepsilon(x - a)$, which, evidently, is a particular case of (7h). As an indication for further work, it shall here be remarked that the approximation methods introduced by Almer and Pesonen might probably be extended by using (7i) to $F(x; t, s)$ of a cPp i.w.s. with $V(x)$ defined by (7h).

Pesonen [265, 268], and Hovinen [186, 188] used *inter alia* a *Monte Carlo method* for the approximation of ${}_\varepsilon F(x, t)$ with $V(x)$ in a given form. This implies the simulation of a random sample of x for given values of t, where x is distributed with the distribution function $F(x, t)$; the estimate of the error involved in the approximation can also be calculated. Numerical calculations were made according to this method, and to other methods, and the results of the latter methods were compared with those obtained by the Monte Carlo method with due regard to the approximation error of the Monte Carlo method. A systematic description of the investigations made by the Finnish school will be found in the book on the risk theory under preparation, quoted in § 3 here above.

Approximations to $F(x, t)$ with the claim distributions referred to in § 6 here above were calculated by Cramér [111] using the Esscher method, and an Edgeworth series. The papers [8–13, 189, 199, 218, 260, 262] deal also with numerical illustrations of approximation methods. In [291] the reviewer derived an expansion of $F(x, t)$ for a cPp i.w.s., and, particularly, for a Poisson, and for a Polya process, where $V(x)$ was given in the form of (7h) with $V_\xi(x)$ being exponentials for certain values of ξ, and in the form $\varepsilon(x - a)$ for other values of ξ. The expansions were intended for direct computation of a sufficient number of terms in an electronic computer. So far, the program for the calculation has been considered too complicated for practical use.

8. Applications of $F(x, t)$ to rating problems, and to other problems

If the risk process is considered for smaller groups of insurances, which to a certain extent are homogeneous, $F(x, t)$ for each such group gives valuable information as regards the rating a priori. The rating involves the estimation of the risk premiums from the statistics of such groups, and the decision to which extent the division into groups in the statistics shall be kept also in the tarif. As, however, it must as a rule be assumed, that the claim frequency and the empirical claim distribution of each subgroup depend on time, the risk premium of each group depends on time and must be predicted for the period during which it shall be applied. The principles for the application of statistical results to practice have been expressed by Wold in 6.4 of [359]. The risk premiums for life insurance are dependent on age attained and on calendar time. So far, it seems, that only the classical form of the risk theory has been applied to life insurance. Recent investigations (e.g. T. Larsson, *Mortality in Sweden*, Stockholm and New York, 1965) have, however, led to the conclusion, that the mortality intensities of non-overlapping time intervals often are mutually dependent. Further, the claim distributions—here called the distributions of the risk sums, not to be confounded with the risk distribution defined in § 1, ($\bar{U}(v, \tau)$ of (1 d))—depend as a rule on time as being subject to variation with changes in the economical and social conditions. Therefore,—in the opinion of the reviewer—similar view-points shall be applied to the risk process of life insurance as those discussed in numerous papers for non-life insurance. The Swedish table of premiums for life insurance has also been based on a predicted mortality.

It should, thus, be allowed for the variation of both $U(v, s)$ and $V(x, t)$ with the parameters. It follows, that the risk premium, upon which the tarif rates are based, will as a rule differ from the risk premium for a later tarif applied to the same group. The risk premiums used in the tarif, which may be called *applied risk premiums*, constitutes, therefore, random processes with discontinuous time parameters, defined by sample functions in the form of step functions with discontinuity points at each change of the tarif rates. If also the security loading in the premiums is based on a prediction (of some measure of the variation of the risk premium) such a loading as applied in the tarif, is attached to a similar process.

Statistical aspects have been considered by Beard [41], and with particular regard to mortality by the same author [44]. Large claims were separately treated by Beard [43], Depoid & Duchez [129], Franckx [168] and Gumbel [179]. Distribution functions of the sum of claims, the largest claim excluded, were given in [26, 168], and with the exclusion of the r largest claims in [27]. Almer (3) introduced two particular statistical methods. One of these implies a separate calculation of the risk measures: the risk premium, the claim frequency and the claim distribution for three different claim groups according to size. This method was called *excess claims analysis*. The other enables the estimation of the separate effects on the risk measures of the components in a parameter vector, the method was called *factor analysis* (cf. [275, 287]). If $F(x, t)$ refers to an insurance without a clause for self-retention (deductable), the risk premium for the corresponding insurance with such a clause, can be calculated by the formulae $\int_{(x>s)} (x-s) d_x F(x, t)$, where s is the size of the self-retention, and, if $F(x, t)$ relates to the total of the policies of a certain line underwritten by the company, ceded or not ceded, the risk premium for an excess over s of aggregate loss reinsurance can be calculated by the same formula [11, 274, 276, 283].

Besides the rating *a priori*, it has been customary in certain branches to account for the actual experience *a posteriori*, either by *experience rating* or by the *distribution* *of dividends.* A method, particularly used in motor insurance, is the system of *no claim bonus*, often combined with some penalty (*malus*) for a large number of claims [20, 25, 54, 79, 128, 132, 165–167, 171, 178, 181, 244–245, 248, 277, 288, 314, 344, 346].

The technical reserves have been considered *i.a.* in [46, 86, 217, 247, 263, 299, 337]. More or less general applications to different problems in non-life insurance are found in [1, 13–14, 16, 28, 37, 87, 130, 142–143, 196, 214, 232, 243–244, 270, 273, 278, 287, 312, 314, 325]. In the list of literature some applications of similar models, as those used for the risk process, to fields outside insurance have been included. [175, 235–236, 285] deal with the recovery of the human eye after dazzling and [238, 342] with ecological problems. [220, 300] deal with computer failures, in [220] a branching Poisson process, and in [300] a branching cPp were used as models. In [300] the relations of these models to cPp i.w.s., and to the model used in [299] for the extended risk process (see § 5 here above) were established.

Problems involved in experience rating have lately been subject to a large interest. Ammeter treated this problem for the risk process in its classical form [17, 21, 24]. Bühlmann gave at the Astin Colloquium in Lucerne a distribution-free method for experience rating [81], he proved that the best estimate, in a certain sense, of the risk premium *a posteriori* is a conditional mean. He stated, further, that the credibility method of estimation, implying the weighting of the results of the actual experience and the results obtained by other experience, e.g. from earlier statistics, could be explained in terms of the conditional mean. The practice of experience rating used by American investigators[1] (*e.g.* [205]) is based on a particular case of Bühlmann's method. Delaporte proposed in [122–127] the use of the conditional mean as defined by (1 h), on the assumption of a risk process in the form of a modified Polya process, for experience rating in motor insurance instead of a bonus-malus system. Ove Lundberg developed [231] a theory for experience rating based on a general cPp i.n.s., the conditional mean was also determined according to (1 h) [230]. Derron considered [133] the betterment of credibility by the exclusion of the largest claim, and used for his deductions [26, 168]. Also Bichsel [55–56] treated at the Colloquium referred to the experience rating based on the theory of cPp i.n.s., and allowed to a certain extent for the random variation of the risk with time. The present author stressed—in the oral discussion at the Colloquium—the necessity of allowing also for a systematic variation with time. In [302] the theory was, therefore, extended to include cPp i.w.s., and the relation between the theory based on such processes, and Bühlmann's general theory was established. Also the connections with modern Bayesian theory (e.g. *Robbins, Rev. de l'Inst. Int. Stat. 31*, 1963) and with the general decision theory were discussed in [302]. For the estimation of the parameters appearing in the conditional mean it was in [302] referred to Anscombe (*Biometrika*, 1950), and to Grenander [177]. Other papers in the list of literature of this review dealing with estimations are, *i.a.*, [216, 274–276, 283].

Bohman discussed the experience rating for a company, which aims to increase the volume of its business [68].

---

[1] For literature the reader is referred to L. H. Longley-Cook, *An Introduction to Credibility Theory*, Casualty Act. Soc., New York, 1962.

9. The ruin functions

An approach to a model for the gain of an insurance company, accumulated during a period of time from zero to t on the transformed scale (see § 1), can be based on the difference $Y(t)-Z(t)$, where $Y(t)$ and $Z(t)$ represent the accumulated revenue and cost respectively. If revenue and cost due to other items than those concerning the pure risk business are neglected, the gain of this business on the own account of the company is given by such a difference; a modified such difference by neglecting also the reinsurers' payments for claims and the cost of the reinsurance, gives the gross gain of the pure risk business, for certain problems also the gain of the reinsurance may be of interest. Finally, it is for each problem to be decided whether the payment of dividends and such alike shall be disregarded; also the interest on revenue and cost may or may not be considered. These different definitions of the difference, defining the gain, will here be called *modifications*. In all the modifications, we have to deal with random functions attached to stochastic processes. That also $Y(t)$ in a very general approach must, in all modifications, be considered a random function is a direct consequence of what has been said in the first two paragraphs of the previous section. Tarifs for the direct insurance are often subject to amendment at intervals of a few years, this does, even *a fortiori*, apply to the reinsurance premiums. The applied, continuous risk premiums and their security loading cannot, therefore, be the same for longer periods. There are also other causes, than those considered in § 8, for the variation with time of the continuous premiums collected. Firstly, the market conditions, including both the competition between the insurers and between the reinsurers, and the different interests of a cedent and his reinsurers with respect to the reinsurance premiums, and other conditions for the reinsurance, shall be observed. As far as the reinsurance cost is concerned, the reader is referred to the remarks to this effect given in [287]. Secondly, the premiums may be subject to changes due to the provisions by law or by the authorities. If the *risk reserve* at the time point t, $Q(t)$ say, is defined as the sum of the initial risk reserve, $Q(0)=u$ say, and the accumulated gain, it fulfils the relation

$$Q(t)=u+Y(t)-Z(t), \tag{9a}$$

which can be modified as was pointed out here above. As far as the present author knows, our knowledge of the process constituted by $Y(t)$ is very restricted, at any rate insufficient for the use of this general model.

At the present stage of our knowledge the ruin theory must be based on very restrictive assumptions. Such assumptions are for example, that the continuous risk premium and the continuous security loading may be considered constant for the whole length of the period considered, equal to c_1 the mean of the claim distribution assumed to be independent of time, and $c_1\lambda$ the continuous security loading subject to the same condition. This is consistent with the classical form of the risk theory. For a cPp i.n.s. with t-independent claim distribution c_1 is, it is true, a constant, but the continuous security loading, if based on the standard deviation of the accumulated claims, dependent on time. It has been found that the ruin theory, so far, as a rule, based on these restrictive assumptions, has in spite of this simplification entailed many difficulties.

In some cases, the theory has been extended by assuming λ to be a function of $Q(t)$ at least for a part of the future. It is easily seen, however, that in a very realistic model even the relation between the security loading and the risk reserve may be changed, so that the extension does not completely eliminate the restriction on λ. As an example, the security loading for the Swedish third party liability motor insurance, being compulsory, and, therefore, strictly controlled by the authorities, was up to 1955 5% and after 1955 3% of the tarif premium; these loadings were determined by the Registrar General, and applicable to all companies regardless of their risk reserves.

The following context is divided into three parts A, B and C, where A and B deal with the development up to the publication of [111], [111] inclusive. The context of A and B is a review of a summary of this development given in [111], A refers to the theory based on a constant λ, and B to the theory based on a security loading being a function of $Q(t)$. C refers to the development after the time considered in A and B.

A. In this case the risk reserve is defined by the following relation

$$Q(t) = u + (c_1 + \lambda)t - X(t), \tag{9b}$$

where c_1 and λ are assumed to be constants, λ essentially positive, while c_1 may be positive or negative. $Q(t)$ is attached to a random process, which is a transform the risk process, as treated in the previous sections, and $X(t)$ the accumulated claim cost for a risk process in its classical form, in one modification c_1, λ, $X(t)$ refers to the business on the own account of the company. If, at some time t, $Q(t)$ becomes negative, it is said that *ruin occurs at t*. This is equivalent with the following definitions of the occurrence of the ruin at t, namely, that ruin is said to occur at t, if a sample function of $Q(t)$ crosses the horisontal axis, or, if a sample function of $X(t)$ crosses the line $x = -u - (c_1 + \lambda)t$, at the time point t. The ruin functions are defined as probabilities of the occurrence of ruin at *some point t* fulfilling certain conditions. In [111] Cramér rigorously defined these probabilities, [110], by proving that the events concerned here, have well-defined probabilities in a discussion similar to that reviewed in § 2 here above, according to which $X(t)$ of the restricted space are associated with probabilities induced from a reference space of an enumerable set of dimensions [111]. This proof was given for *the mixed case* taken to mean the case, where $0 < V(0) < 1$, the cases, where $V(0) = 0$ and $V(0) = 1$, are referred to as *the positive case* and *the negative case* respectively. For the definition of the ruin functions (9b) may be considered as a function either of the continuous parameter t, or of the discontinuous parameter $t = rh$, where h is a given positive quantity (e.g. the length of a business period), and $r = 1, 2, \ldots$. The symbol generally used for the ruin functions, ψ, shall be used in this section, and be reserved for this purpose (it shall not be confounded with ψ used in the previous sections and defined in § 1 here above). The *ruin functions*, for the case of t being continuous, are the probabilities for the occurrence of ruin at some $t > 0$, and at some t in the interval $0 < t < T$, where T is a given value, these functions are denoted $\psi(u)$ and $\psi(u, T)$ respectively. The corresponding probabilities, for the case of t being equal to rh, are denoted $\psi_h(u)$, $\psi_h(u, T)$ respectively. Let further, $\Pi(s) = 1 + (c_1 + \lambda)s - \int_{-\infty}^{+\infty} e^{sx} dV(x)$, where $s = \sigma + i\eta$, and the integral is the complex

Fourier transform of $V(x)$, for $\sigma = 0$ being the characteristic function corresponding to $V(x)$.

The ruin functions $\psi(u)$ and $\psi_h(u)$ were introduced by Filip Lundberg in 1926–1928 [226], and, further, developed by him in 1932, 1934 [228–229]. He obtained for each of these functions in the *positive case* an inequality, and an asymptotic relation for large values of u, given in the following lines.

$$0 \leqslant \psi_h(u) < \psi(u) \leqslant e^{-Ru}, \tag{9c}$$

$$\psi(u) \sim C e^{-Ru}; \quad \psi_h(u) \sim C_h e^{-Ru}. \tag{9d}$$

Apart from C_h being dependent on h, R, C and C_h are constants depending only on λ and $V(x)$. In anticipation of the results obtained later, (9c) holds also for the *mixed case*, and, for the *negative case*, if $c_1 + \lambda < 0$, where in the last inequality the sign of equality holds. (In [306] Prabhu has proved the last assertion for a general additive process with stationary increments.) Under an additional assumption (see below) (9d) holds also in the *mixed case*.

Cramér proved in his papers of 1926 and 1930 [99, 101] that $\psi(u)$ in the *positive case* is a solution of an integral equation of Volterra type, which can be solved by complex Fourier transformation. He gave an explicit expression for $\psi(u)$ in this case, and a proof of (9d).

In 1941 [105] Cramér proved that $\psi(u)$ for the *mixed case* satisfied an integral equation, not of the Volterra type. Segerdahl gave for this case the first rigorous proofs of (9c–d). He studied also the *positive* and the *negative cases*, and proved for these cases a number of important results, some of which had, without complete proofs, been stated by Filip Lundberg. Täcklind [354] showed, that the integral equation in the *mixed case* satisfied by $\psi_h(u)$ can be solved by a method due to Wiener and Hopf [259] involving a complex Fourier transformation combined with arguments from the theory of analytic functions. From this solution a proof of (9d) for $\psi_h(u)$ was obtained; he also proved that $\psi_h(u)$ tends to a definite limit, for which an explicit expression was given, as h tends to zero. Cramér [110] gave a probabilistic definition for $\psi(u)$, and applied the Wiener-Hopf method directly to the integral equation for $\psi(u)$ in the *mixed case*.

Certain preliminary results with respect to $\psi(u, T)$, and $\psi_h(u, T)$ were given by Filip Lundberg [226] and by Segerdahl [317]. In [317] the moments of the time, when ruin occurs for the first time were calculated. The problem was thoroughly investigated by Saxén [315–316], mainly for the *negative case*, and, by Arfwedson, for the *positive case*. For these cases explicit expressions for $\psi(u, T)$ were given. Arfwedson obtained also a number of results concerning the asymptotic properties of this function.

In [111] Cramér rigorously proved, that the limits of $\psi_h(u)$, $\psi_h(u, T)$, as h tends to zero, are equal to $\psi(u)$, $\psi(u, T)$ respectively. He, further, introduced the basic assumptions with respect to $V(x)$ that the means of $|x|$ over the negative axis, and of $e^{\sigma x}$ over the positive axis for some $\sigma > 0$ are finite. For the proof of (9d) in the *mixed case* an additional assumption was made, implying that, for some $\sigma > R$, the mean of $e^{\sigma x}$ over the positive axis is finite. The relations between R and certain other constants with $\Pi(s)$ were analysed. R was defined as the least upper bound of q, subject to the conditions that for $0 < \sigma \leq q$ the complex Fourier transform of $V(x)$ is analytic and regular, and $\Pi(\sigma) > 0$. For the *mixed case* the integral equations satisfied by the ruin functions and by their complex Fourier transforms and by certain other functions, are discussed by means of the Wiener-Hopf method [259]. One of these equations, satisfied by the complex Fourier transform of $\psi(u, T)$, leads to explicit expressions for $\psi(u)$, $\psi(u, T)$, and to certain results for the asymptotic properties of these functions. Some results of such properties due to Segerdahl [317] were proved; also an inequality for $\psi(u) - \psi(u, T)$ was deducted. In the *positive case* (9d) was proved, and the inequality for the difference $\psi(u) - \psi(u, T)$ was strengthened; an asymptotic relation for the difference for this case (stated without proof by Arfwedson, later proved by him) was proved in [111].

In the *mixed case* the form of $V(x)$ indicated by Täcklind (see § 6 here above) the following expression for $\psi(u)$ was obtained in [111].

$$\psi(u) = \sum_{u\ 1}^{N} C_n\, e^{-R_n u}, \tag{9e}$$

where $N = r$, the number of terms in the exponential polynomial, if $c_1 + \lambda > 0$, and $N = r + 1$, if $c_1 + \lambda < 0$, C_n is given by an expression analogous to the expression for C in the equality for $\psi(u)$ under more general conditions (5.10, [111]), and R_n are the zeros of $\Pi(s)$ in a rectangle formed by $\sigma \pm iT$, $\Sigma \pm iT$ ($\Pi(s)$ is on the contour of this rectangle positive for sufficiently great values of Σ and T). In the positive case $C_n = \lambda/[-\Pi'(R_n)]$, $(N = r)$, and, if, particularly, $r = 1$, $C_1 = 1/(1+\lambda)$, and $R_1 = 1 - C_1$.

B. In this case a never increasing function $\lambda[Q(t)]$ is substituted in (9b) for λ, if $Q(t) < a$, where a is equal to a finite constant, or to infinity. This problem was already treated by Filip Lundberg in 1926–1928 [226], and, further, by Laurin [215], Täcklind [354] and Davidson [121]. For the particular case, where $V(0) = 0$, $V(x) = 1 - e^{-x}$, and a is given by a finite constant, Davidson gave the following relation for $\psi(u)$, where $H(u)$ is defined by

$$H'(u) = \exp\left\{-\int_0^u \frac{\lambda(v)}{1+\lambda(v)}\, dv\right\},$$

$$\psi(u) = 1 - \frac{H(u) + H'(u)}{H(a) + \dfrac{1+\lambda}{\lambda} H'(a)}. \tag{9f}$$

If in (9f), $a = 0$, i.e. $\lambda(Q(t))$ is equal to λ independently of $Q(t)$, the assumptions in B, reduce to the assumptions in the last paragraph of A, with $r = 1$. In fact, for $a = 0$, (9f), reduces to (9e) with

$C_1 = 1/(1+\lambda)$, and $R_1 = 1 - C_1$.

C. After the publication of [111] Arfwedson published the second part of [35] with numerous results in the ruin theory. Segerdahl published also new studies into this theory particularly dealing with the time point at which ruin occurs for the first time [321, 322]. This has also been treated by Prabhu [306], who used queuing theory in his developments. Arfwedson has recently given some notes on [306], unpublished, showing the relation between the proofs of [306] and [35]. In [323] Segerdahl gave, for a great number of particular cases, explicit expressions for $\psi(u)$ and for $\psi(u, T)$, in one of these cases the interest accrued on $Q(t)$ was accounted for; he referred also to cases treated by Arfwedson [31, 33, 35], where λ was allowed to take zero or even negative values. Segerdahl, further, derived an expression for $\psi_h(u)$ under the asumptions of one of Ammeter's models described in § 5 here above, including the assumption that $t\gamma_2/\gamma_1^2$ is bounded even for $t \to \infty$. Ammeter has, however, in [9] arrived to a similar expression for $\psi_h(u)$ where this last-mentioned assumption seems not to have been used. As far as it is known to the reviewer, no other deduction of the ruin functions, based on other forms of the risk process than the classical form, have been published so far. Almer [3] indicates, however, that the deduction of approximate expressions for $\psi(u, T)$ should be possible under wider assumptions, if based on his approximations of $F(x, t)$. In 1966–1967 Segerdahl lectured on the risk theory at Stockholm University; in these lectures [111] was reviewed with new proofs for particular cases; one of the problems treated in the lectures was, further, studied by Thorin [343]. Segerdahl discussed also in these lectures by the methods used in [111] a rigorous extension of the ruin theory to a Polya process, which will be published later.

An interesting contribution to the ruin theory for the classical form of the risk process was published in 1966 by Beekman [47]. According to his developments $1-\psi(u, T)$ could for the *mixed case*, be determined by the conversion of a double Laplace transform of the probability for the occurrence of the event Max $(-Q(t)) < \alpha$, by a conversion method described by Widder (*The Laplace transform*, Princeton University Press, 1946). In the *positive case*, the probability mentioned, with $c_1 + \lambda$, particularly, replaced by zero, is equal to $F(\alpha, t)$, which, thus, can be deducted either by a limit passage of the said probability, or by the conversion of the corresponding double Laplace transform. The theory has been illustrated by a few simple numerical examples; the application to more realistic models shall be subject for future research. Beekman's paper has been discussed by Thorin in a recent manuscript to the Lundberg Symposium.

10. Application of the ruin theory and of other theories to decision problems, and references to studies into reinsurance problems

In a great part of the literature criteria for decision problems in insurance companies, particularly for decisions related to the reinsurance policy, have been based on some ruin function implying that a decision shall be chosen, which entails a reduction of the ruin function to a fixed predetermined level. It seems evident, that it must be considered more realistic for this purpose to use $\psi(u, T)$ than $\psi(u)$. In many cases such criteria are to be applied to other decision problems such as those regarding the magnitude of the risk reserve and of the security loading. In fact, these decisions are connected with the choice of reinsurance policy, and ought, therefore, to be simultaneously considered. Also the choice of a system for the distribution of dividends is connected with the decisions, just mentioned. A very interesting application of the ruin functions is the solvency control of insurance companies according to the Finnish Act of Insurance, originally suggested and drawn up by Pentikäinen. The eager interest for suitable methods for the approximation of $F(x, t)$ and $\psi(u, T)$, shown by the authors of the Finnish school, is a consequence of the legal provision just mentioned [186–188, 199, 264, 265, 268, 269].

Such criteria, based solely on the ruin theory as reviewed in the previous section, has been subject to criticism by two groups of critiques, referred to in § 2 of this review. The first group can here be exemplified by Campagne [87], Campagne & Driebergen [88], de Finetti [155–161], Giuseppe [254–257] and Giovanni Ottavian [258] and Tedeschi [333–335]. One of the arguments given in some of these papers, is that it seems unnatural, that the criterion of the previous paragraph becomes gradually more and more severe (quoted from [19]). The papers of this group were published from 1940 to 1957. In [158, 161] de Finetti suggested that the reduction oi the ruin function to a fixed level should be combined with an auxiliary condition, which implied a maximisation of future gains.

Even if it should become possible to give a realistic definition of the gains according to the view-points on this problem given in the introduction of the previous section, it is uncertain, whether a discussion based only on the gains will be found sufficient in all cases. Business enterprises in general have very often other aims besides pure profitableness; this seems to be particularly true for insurance companies (cf. *e.g.* [68]). In the preference theory of economics tools for measuring the preference have been given, which have been called utility functions. By the application of this theory combined with the theory of games it is possible to account for different aims of the company and for that part of the variation of the continuous premiums collected which is connected with the competition in the markets and between the interests of the cedent and his reinsurers, as referred to in the first paragraphs of § 9 here above. The application of the theories just mentioned has been introduced by the second group of critiques e.g. Borch [69–75], Kahn [197], Ohlin [253] and Wolff [361]. These papers were published from 1962 to 1967. Particularly Borch's contributions are to a wide extent based on the theories mentioned, as given by Neumann-Morgenstern (*Theory of Games and Economic Behaviour*, Princeton, 1944). As was mentioned in § 2 of this review, Borch has, however, in some papers used an unrealistic model for the distribution of the claim cost, without considering the extensive research on such models accomplished before Borch's first contribution was published, and referred to in the preceding sections of this review.[1]

The last remark applies to the last term, $Z(t)$, in (9a). With respect to the middle term, $Y(t)$, it is evident, that in Borch's approach only a part of the variation in the

[1] Segerdahl kindly drew the reviewer's attention to a paper on the same topic by Klinger in 1965 [199*], where the stringent developments lead to some results, later published also by Borch.

continuous premium collected is accounted for. It is for example difficult to see any possibility of accounting for the influence of the provisions by law and by the authorities on the rating. Further, the part of the variation in the continuous premium collected, which was referred to in the introduction of § 8 in this review, is not accounted for in Borch's models. This variation is due to the fact, that, at least, if the risk distribution and the claim distribution are dependent on time, every new tarif must be based on new statistics, and on new predictions for the $Z(t)$-process. For an ideal decision theory it seems, thus, necessary to combine Borch's ideas with a deeper study of the $Z(t)$-process and the dependence of $Y(t)$ on the trends in the risk measures, which determine the $Z(t)$-process. Such studies must be based on the ideas, which led to the extensions of the classical form, and which were reviewed in § 5 here above.

Many items in the reference list deal with reinsurance problems [8, 11–12, 16, 21–22, 42, 49–53, 58–59, 67, 76, 84–85, 90, 98, 134, 139–141, 147, 155, 157–158, 160, 182, 190, 197, 200, 211–213, 218, 222–226, 246, 249, 253, 255–256, 258, 261–262, 267, 276, 283, 311, 320, 328, 338–340, 353, 357, 361–363]. A part of these papers have been commented upon earlier in this review. Many of the papers referred to are based on the classical form of the risk process, in some papers, however, e.g. [12, 16, 362], a Polya process has been used in the model. Modern forms of reinsurance have been discussed in [52, 338] and other papers.

*

Bühlmann suggested in [82], that—with a new terminology—the risk theory should be divided into three such theories with regard to the different approaches in the theory of decisions in an insurance company as described in the first, second and third paragraph respectively of § 10 here above. Bühlmann's argumentation for such a terminology does not seem the reviewer very convincing. In all the approaches the claim cost ought to be based on the knowledge of the $Z(t)$-process gained so far. Arguments for an extension of this experience, and for a study of the influence of this process on the $Y(t)$-process have been given in the fourth paragraph of § 10. Therefore, the decision theory must be based on all contributions to the risk theory reviewed in this paper, and on a further study of the risk process. Most of the development of our knowledge of the risk process is made under the assumptions leading to the classical form, and as far as the ruin functions are concerned, on the restrictive assumption, that $Y(t)$ is proportional to t eventually modified by the assumption that the security loading depends on the magnitude of the risk reserve. This has led to possibilities for the approximation, and for the application of $F(x, t)$ under very realistic particular assumptions with respect to the form of $V(x)$. For much wider, and, in fact, very mild conditions with respect to this function, and under the assumption just mentioned for $Y(t)$, the probabilities for the occurrence of ruin for the classical form have been completely treated as is seen from § 9 here above. The assumptions of the classical form involve, in fact, that the part of the variation in the continuous premium collected, which was described in the introduction in § 8, should be reduced to a variation due to the sampling errors of the estimated premiums. The assumptions of A and B in § 9 are, thus, connected with the assumptions for the classical form, as far as the influence of the provisions by law and authorities, and of the competition is neglected. From the development of the classical form two lines of development have branched out, one refers to the generalization of the fundamental assumptions, as reviewed in § 4, followed by the extensions reviewed in § 5, and referred to in several remarks in §§ 6–10. The other line refers to the extensions of the decision theory, as reviewed in the second and third paragraph of § 10. The fourth paragraph of § 10 points finally to a union of these two lines in future research. It is evident, that a division of the risk theory according to Bühlmann's suggestions implies the necessity of using force against the strong connections between the three different view-points in the decision theory, and the remaining part of the risk theory.

The development of the risk theory in its classical form has been accomplished by Filip Lundberg, and by Cramér and many others. The first extension of this form was introduced by Ove Lundberg—Filip's son—and by Ammeter, who were followed by many others. The new ideas in the decision theory were introduced by de Finetti, and by Borch, and studied by other authors. These lines of development are, however, all based on the fundamental conception of the collective risk theory, which was created by Filip Lundberg.