MODERN GENERAL RISK THEORY

Bertil Almer [1])

Risk Elements — Definitions and General Properties.

1. *Introduction*

Modern life is characterized by risks of different kind: some threatening all persons and some restricted to the owners of property, motor cars, etc., while still others are typical for some individuals or for special occupations. The corresponding accidents, losses or claims will occur suddenly and unexpectedly and may involve considerable financial loss. It is quite evident that modern life is a fit subject for risk theory, and that some results in the pure mathematic theory might have applications in the study of problems in real life.

In practice, however, we can identify risk theory with insurance risk theory or with the application of the theory of probability on insurance risk problems. This general definition has the advantage, that it covers a wide field of different risks and risk problems as specified in the insurance texts—and a great collection of risk situations = claims occurred (with corresponding loss amounts) is available in the claims acts. In fact, I believe that any actuary or mathematician, starting his researches in risk theory or in risk statistics, should begin his studies by a series of actual claims acts.

2. *Definitions and some general properties*

Detail studies of the claims acts in any chosen branch will prove that almost all claims are the result of well-defined risk situations of short duration—and that the actual accident is usually caused or generated by the combination of a primary cause and some unlucky circumstances.

As an example, we might choose a road accident—started by a small driving error, made worse by a bad patch of road, and made critical by a meeting car, or a car on the cross-road, passing that

very second or those fifty feet which represent all the difference between collision or no-collision. This example illustrates that some types of risk situations might occur very often without any accident because the unlucky circumstances have not combined into a critical risk situation—and also the great influence of the time factor and of the "meeting traffic" in motor car insurance.

Another example might also be of some interest: the fire risk in a one-family house. For our purpose, the heating arrangement represents a very small risk; still we will divide that risk, say, into half-day elements to avoid the introduction of continuous risk elements. The residue is made up of lighted matches, cigarettes, candles and eventually a fire-place in the sitting-room, and an electric oven (or cokes oven) in the kitchen, all representing risk situations of short duration (or divided into short sub-elements). At the same time it is evident, that the "tariff risk", calculated by some parameters describing house and furniture, is most insufficient to characterize the complete risk, which depends more on the family's manner of living (high-life?, non-smoker?, late hours?), carelessness and similar factors.

This is a great problem in most insurance branches, because no tariff parameters can express the exact individual risk, and all control by the most advanced statistics can only prove that tariff premiums are correct in the mean.

The problem is closely associated with another problem, which is analyzed in this paper, that of "stochastic disorder". In short: when authorities declare that "premiums should be exactly determined by statistics even for the greater part of the individual risks" I have proved that this should be translated into "(as an ideal) tariff premiums (diminished by standard additions) should in the long run prove approximatively equal to risk premiums for any combination of risk groups of sufficient magnitude."

Starting from the advanced analysis of claims acts—including the much more frequent analogous risk situations without claims—we will define our risk elements as follows.

Definitions:

Claim = effect of risk element (i), ruled by chance and including some form of damage, and loss of a certain amount of money, which can be insured $(= x_i)$.

Risk element (i) is characterized by $[p_i \mid \psi_i(x) \mid \theta_i(y)]$, where

probability of a claim occurring $= p_i$ $(0 \leq p_i << 1)$

"claims curve" = probable risk amount distribution $\Psi_i(x)$ for $x_i \geq x$, with $\psi_i(x)\, dx$

probable risk amount, mean $= \int_0^\infty x\, \psi(x)\, dx = x_i$

contribution to total claims amount $= y$ distributed as $\theta_i(y)$ $= p_i\, \psi_i(y)$, mean $= \rho$

set or system of risk elements $= \{p_i \mid \psi_i(x) \mid \theta_i(y)\}$ $i = 1, 2, 3, \ldots N$

risk object = object insured—generally the insurance will cover damage (from fire etc.) on object insured or damage caused by the (object) insured on other objects (= third party).

In analysis, we should observe the following properties of risk elements or sets of risk elements:

a. Claims acts seem to indicate that the total number of risk elements in one (insurance) year is finite but very large $= N$. This small simplification is accepted here, but of no great importance because

all distribution functions are of limited variation (à variation bornée) and completely determined by the values in the enumerable rational points of the plane.

all Stieltjes integrals on (simple or) positive functions can be uniformly approximated by finite sums.

b. Individual risk elements and individual claims are characterized by an almost unlimited number of parameters and other markings, which may be used to form "statistical groups" in different analyses—among others

a fixed point on the line of time—eventually followed by a short period—representing the time when any particular risk element (i) is active or the occurrence of any particular claim

insurance parameters: tariff, subtariff, geographical district, occupation classes, profession, age, sex, bonus class, etc.; sums insured, HP, maker, speed.

claims parameters (including insurance parameters and further):

type of accident or claim, causes, other circumstances, weekday and time; claims amount, "excess classes" (= classes according to claims amount, say 0-1.000-10.000-100.000-), days of illness, grade of disablement, etc.

c. Both risk objects and claims are sometimes divided into subelements, defined by some parameters or markings, for instance

traffic: collisions, personal damage, disablement, amount $x \leq z$, some particular causes, amount minus loss excess, catastrophes, multiple claims,

accident: disablement claims, claims from chosen causes,

fire: catastrophe claims, "contagion" from neighbouring risks (as a cause or a consequence),

but also in principle non-actualized claims, say

traffic: latent claim with big danger and strong reactions on the driver's future driving, similar to the effect of a bad catastrophe,

sickness: acute claim = illness in bed versus latent claim = risk for contagion.

d. Conversely, we can form combined risk elements, representing the total risk for an insurance in a calendar year. This is quite good, if the total claims frequency p_i in the year does not exceed some pro mille (life, fire) and if we have a good control of the risk for more than one claim (Cf. ASTIN 1961).

e. Generally, detail studies of the individual risk elements in a system are impossible to perform, and even the combined risk elements above, which are identified with individual insurances, will only lead to an apparent simplification. Our real problem is to study the

risk a priori as the origin of claims, governed by chance,

risk a posteriori or claims occurred,

and the corresponding

probabilities, claims frequencies and claims amounts,

risk groups or statistical groups (tariff groups),

and to decide whether the risk system is large enough to give good statistical figures or not.

In risk analysis and statistical studies the following system of figures and symbols is utilized

Risk elements (*i*) *a priori*	*Risk groups a priori*
Claims probability $= p_i$	Number of claims $= n = \Sigma p_i$ $(i = 1, 2, 3, \dots N)$
Claims amount $= x_i$	Actual claims $= x_j$ $(j = 1, 2, \dots n)$
Claims curve $= \psi_i(x)$	Claims curve $\bar{\psi}(x) = \Sigma p_i\, \psi_i(x)/n$
Mean value $= m_i$	Mean value $= \Sigma p_i m_i/n = m = \bar{x}$

Risk groups a priori	*Risk groups a posteriori*

Number of insurances $= N(I)$, or sums insured $= S(I)$, premiums $= P(I)$

Number of claims $= n$	Number of claims (occurred) $= \nu$
Claim frequency $= f = n/N(I)$	Claim frequency $= \varphi = \nu/N(I)$
Individual claims $= x_j$	Individual claims $= \xi_j$
Sum total of claims $= y = \sum_{j=1}^{n} x_j$	Sum total of claims $= \eta = \sum_{j=1}^{\nu} \xi_j$
Mean value of claims $= m = \bar{x}$	Mean value of claims $= \xi = \eta/\nu$

Mean value of sum total $= n.m = \int_{0}^{\infty} y\theta(y)dy = \bar{y}$

Risk premium $= r = y/N(I)$	Risk premium $= \rho = \eta/N(I)$

Relative risk premium $^0/_{00} = y/S(I)\cdot 1000$
Loss percent $= 100.y/P$ (calculated or tabulated)

The "claims curve" $\bar{\psi}(\xi)$ is defined by the distribution of ξ_j according to magnitude, and is constructed simply by counting the number of claims above a variable limit z.

3. *Error formulas—the road from risk elements to homogeneous binominal distributions and from homogeneous distributions to poisson series.*

In the general risk theory we should avoid the introduction of restrictive hypotheses as much as possible. So far we have only made use of:

a. Risk elements are conform (or similar) to risk situations, as these are described in the claims acts of different insurance branches.

b. Risk elements are formed by a definite small probability p_i for the effect "claim".

c. In case a claim has occurred, there is further a definite probability—expressed by the "claims curve" $\psi_i(x)$—for the claims amount.

d. All risk elements and all claims are actualised at definite points on the line of time, and characterized by a series of parameters or markings giving them a definite place in most statistic classifications.

Returning to the claims acts in any specified branch we should observe, that the sequence of claims represents a mixture of:

different tariffs and risk classes,

risk situations where p_i is comparatively large (say $1\ \%_0$) or very small—or else where large claims are menacing or even completely destructing valuable objects, and, as a contrast, situations leading to small scratches and discolourings;

in short that the corresponding stochastic variables (ν, ξ, η) will move in a most irregular way, without any contact or conformity to homogeneous systems or simple binominal distributions.

In order to get a full control of this complicated problem, we should try to obtain

a. a simplified or homogenized risk system with the same $[n \mid \psi(x)]$ and approximately the same properties.

b. exact error limits or at least some estimate—also an estimate of the effective difference between the general system of risk elements and the corresponding Poisson system with the same values for

$$\text{risk group} = (\alpha, \beta, \gamma, \delta) \text{ or total}$$
$$\text{number of claims a priori} = n$$
$$\text{claims curve} = \bar{\psi}(x)$$
$$\text{total claims amount a priori} = y$$

This problem is completely solved by the first error theorem (Cf. New York 1957).

Definition: For any general system of risk elements $\{p_i \mid \psi_i(x) \mid \theta_i(y)\}$ and for any risk group $\{\alpha, \beta, \gamma, \delta \mid n, \overline{\psi}(x), y\}$, where

$$n = \Sigma p_i \ (i = 1, 2, 3, \ldots N) \quad 0 \leqq p_i < 1$$

$$\overline{\psi}(x) = \Sigma p_i \psi_i(x)/n \quad x > 0 \quad \Psi(x) = \int_x^\infty \overline{\psi}(x)dx \leqq 1$$

$$y = \Sigma \ p_i \ m_i$$

we form the sum of squares

$$\overset{N}{\underset{i=1}{\Sigma}} \ p_i^2 = n^2.Q \leqq n^2/L \quad (L = \text{integer})$$

and define

$Q = $ (exact) measure of inhomogenity (or asymmetry) for $(\alpha, \beta, \gamma, \delta)$

$L = $ reduced number of risk elements or measure of inhomogenity (for all groups and amounts)

$N = $ number of risk elements (finite).

$Q(x) = \Sigma \ [p_i \ \Psi_i(x)]^2 : [\Sigma \ p_i \ \Psi_i(x)]^2 \leqq 1/L$
    $= $ measure of inhomogenity for amounts $\geqq x$.

Theorem 1: The generating function for a general system of independent risk elements = general binominal system is expressed by

$$\overset{i=1}{\underset{N}{\Pi}} \ (1 - p_i + t.p_i) = \Pi \ (1-p_i). \ \Pi \left(1 + \frac{p_i t}{1-p_i}\right) =$$

$$= \sum_{k=0}^{N} A_k t^k = \Pi \ (1 - p_i) \left[\sum_{k=0}^{N} B_k.t^k\right] = \Pi \frac{1}{1+q_i} . \ \Pi \ (1 + t.q_i)$$

where $\dfrac{p_i}{1 - p_i} = q_i; \quad 1 - p_i = \dfrac{1}{1 + q_i} ; \quad \Pi \ (1 - p_i) = \Pi \left(\dfrac{1}{1 + q_i}\right)$

and $\quad \Pi \ (1 - p_i) \ e^{\Sigma \log(1-p_i)} = e^{-\Sigma p_i - \frac{1}{2} \Sigma p_i^2 - \frac{1}{3} \Sigma p_i^3 - \cdots} = e^{-n - \frac{1}{2}n^2 Q - \cdots}$

$$\sim e^{-n - \frac{n^2}{2L} + 0 \left(\frac{n^3}{L}\right)}$$

where the coefficient of t^k is a symmetric function of the probabilities and expresses the probability that the stochastic variable ν (number of claims a posteriori) will take the value k; compare with Poisson: $e^{-n}. \ \dfrac{n^k}{k!}$

Theorem 2: Under the sole conditions

$$\Sigma\, p_i = n \ (i = 1, 2, 3, \ldots N) \quad 0 \leq p_i < 1$$
$$\Sigma\, p_i^2 = n^2 . Q \leq n^2/L \quad (L = \text{integer})$$

the symmetric elementary functions S_μ where $\mu \leq N$,

$$S_1 = \Sigma p_i \qquad S_2 = \underset{i<j}{\Sigma p_i p_j} \qquad S_3 = \underset{i<j<k}{\Sigma p_i p_j p_k}$$

$$S_\mu = \Sigma\ p_{i_1}\, p_{i_2}\, p_{i_3} \cdot\cdot\ \cdot p_{i_\mu}\ (i_1 < i_2 < i_3 \ldots < i_\mu)$$

take their maximum values $S_\mu{}^{\max}$

$$S_1{}^{\max} = n; \ S_2{}^{\max} = \frac{N(N-1)}{1 \cdot 2}\, \frac{n^2}{N^2}; \ S_3{}^{\max} = \frac{N(N-1)(N-2)}{1.2.3}\, \frac{n^3}{N^3};$$

$$S_\mu{}^{\max} = \binom{N}{\mu}\, \frac{n^\mu}{N^\mu}$$

or $S_1{}^{\max}=n; \ S_2{}^{\max} = \frac{1(1-1/N)}{2!}n^2; \ S_3{}^{\max} = \frac{1(1-1/N)(1-2/N)}{3!}n^3 \ldots$

$$S_\mu{}^{\max} = \prod_{k=1}^{\mu-1} \left(1 - \frac{k}{N}\right)\frac{n^\mu}{\mu!}$$

at the same time for the system with maximum symmetry

$$p_i = n/N \ (i = 1, 2, 3, \ldots N)$$

and their minimum values $S_\mu{}^{\min}$

$$S_1{}^{\min} = n; \ S_2{}^{\min} = \frac{L(L-1)}{1.2}\, \frac{n^2}{L^2}; \ S_3{}^{\min} = \frac{L(L-1)(L-2)}{1.2.3}\, \frac{n^3}{L^3};$$

$$S_\mu{}^{\min} = \binom{L}{\mu}\, \frac{n^\mu}{L^\mu}$$

or $S_1{}^{\min}=n; \ S_2{}^{\min} = \frac{1(1-1/L)}{2!}n^2; \ S_3{}^{\min} = \frac{1(1-1/L)(1-2/L)}{3!}n^3 \ldots$

$$S_\mu{}^{\min} = \prod_{k=1}^{\mu-1} \left(1 - \frac{k}{L}\right)\frac{n^\mu}{\mu!}$$

at the same time for the system with maximum asymmetry, where

$$p_i = n/L \ (i = 1, 2, 3, \ldots L); \ p_i = 0 \ (i > L)$$

The proof by the "three point test" is very simple: we "freeze" all variables but three and reduce the problem to

$$x + y + z = A \qquad x^2 + y^2 + z^2 = B$$

$$xy + xz + yz = \tfrac{1}{2}\sqrt{A^2 - B} \qquad xyz = Z = \quad \begin{cases} \text{maximum} \\ \text{minimum} \end{cases}$$

with only one "degree of freedom"—the resulting condition $x = y < z$ or $x = y > z$ is permutated—which leads to the statement above.

Theorem 3: The series in symmetric functions, representing maximum symmetry

$$1 + \sum_{\mu=1}^{N} B_\mu t^\mu = 1 + \sum_{\mu=1}^{N} \binom{N}{\mu} \cdot \left(\frac{n}{N}\right)^\mu t^\mu =$$

$$= 1 + \sum \frac{1\left(1 - \frac{1}{N}\right)\left(1 - \frac{2}{N}\right)\left(1 - \frac{3}{N}\right) \cdots \left(1 - \frac{\mu-1}{N}\right)}{\mu!} n^\mu t^\mu$$

and (in complete form)

$$1 + \sum_{\mu=1}^{N} A_\mu t^\mu = 1 + \sum_{\mu=1}^{N} \left(1 - \frac{n}{N}\right)^N \binom{N}{\mu} \left(\frac{n}{N}\right)^\mu t^\mu$$

will generally converge very quickly with N to the "exponential distribution" ("Poisson")

$$1 + \sum_{\mu=1}^{\infty} e^{-n} \cdot \frac{n^\mu \cdot t^\mu}{\mu!}$$

and all terms or probabilities in any considerable distance from the maximum term with $\mu = n$ will be exceedingly small.

The convergence of the system with maximum asymmetry will greatly depend on the reduced number of elements L.

Convergence will be retarded in particular if there exists a small group of "very bad risks" i.e. with large values of p_i. For instance p_i might represent millionth parts, with a very small group with $p_i \sim 1/10$ generating most of the number of claims n.

Theorem 4: The claims curve $\overline{\psi}(x)$ for the maximum system and

the minimum system and also for the resulting Poisson system is not changed, owing to definitions and the properties of

$$Q(x) \leq 1/L.$$

(*First*) *Error Theorem*: All systems of independent risk elements

$$\{p_i \mid \psi_i(x) \mathbin{|} \theta_i(y)\} \qquad [n \mid \overline{\psi}(x) \mid N \mid L \mid y]$$

with the same characteristic values (a priori):

number of elements = N

number of claims = n

claims curve = $\overline{\psi}(x)$ — or $n.\overline{\Psi}(x)$ = number of claims $\geqq x$

reduced number of elements = L = measure of inhomogenity, will define a stochastic variable v (number a posteriori) with values equal to coefficients of the generating function in t falling between:

homogenized system = system of maximum symmetry

and

reduced homogenized system = system of maximum asymmetry.

Note: This apparent anomaly—both series having the same sum of coefficients—is made possible by the small supplementary factors

$$e^{-n-\frac{1}{2}}\,\frac{n^2}{N}\cdots \text{ and } e^{-n-\frac{1}{2}}\,\frac{n^2}{L}\cdots \text{ ; compare } e^{-\frac{1}{2}\left(\frac{n^2}{L}-\frac{n^2}{N}\right)} =$$

$$= e^{-\frac{1}{2}\,\frac{N-L}{NL}\cdot n^2} \sim e^{-\frac{1}{2}\,\frac{n^2}{L}} \text{ and } A_\mu \text{ resp. } B_\mu$$

adjusting the limits given to exact values.

Further, the difference between the system of maximum symmetry and the corresponding infinite Poisson series will probably always be negligeable in insurance problems. —

The lower limit is more important and if strong asymmetry is to be feared, the eventual methods to determine L should always be carefully analyzed. Numerical error expressions are easily calculated from the generating function compared with note above, simply by applying the Stirling formula and by concentrating to the neighbourhood of $\mu = n$ where coefficients attain their maximum value.

Corollary: In the insurance theory asymmetry is generally the

effect of a small set of very bad cases and when the group is growing we might assume that new risks represent in average fairly good risks.

In this case the adjustment

$$e^{-\frac{1}{2} n^2 \frac{N-L}{LN}} \curvearrowright e^{-\frac{1}{2}\frac{n^2}{L}} \rightarrow 1 \text{ with increasing } n,$$

proves that the original system of risk elements will rapidly converge to the ordinary Poisson distribution when n (number of claims, $\curvearrowright$ number of insurances) increases.

Note: The corollary proves the importance of good estimates of L (degree of inhomogenity).

Many methods have been tested in practice, but generally either the statistic material is far too meagre, or else the conditions are too restrictive. Some examples:

a. Analysis of accident insurance (about 1910). Claims frequency φ varies from 50 to 300 $^0/_{00}$ with the mean 100 $^0/_{00}$. For ten years old insurances the claims frequency was calculated for first and second five-year period. Criticised because the personal risk will change in ten years.—The same method has been proposed for motor car, where the result should be still better as many large groups will have $\varphi = 300$ $^0/_{00}$, $\varphi = 500$ $^0/_{00}$ or even larger. Unfortunately bonus experience proves that the personal risk will generally be reduced to one third or less in some few years,—and then is subject to jacks or jerks, probably depending on children growing up or changes in the manner of living.

b. General study of technical risk character, compared with values of φ for different statistical groups; probably one of the best methods, but very difficult to work through.

c. By analyzing the relative frequency of insurances with no, one, two, three, four, etc., claims (accumulation curves). This important method assumes that claims can be repeated without change in the primary generating force or frequency and without any vacant period, in short that claims can be described by the "Poisson process". The same hypothesis is standard in many statistical researches, often with excellent results, but sometimes with considerable differences. Often these differences will take the

form of a "stretched" curve, where the frequency of "many claims" is increased compared with "no claims". Evidently, the same result is obtained by an exponential binomial (the sum of two Poisson terms) and practical experience proves that at least in the insurance theory, an exponential trinomial will always be sufficient, and—what is more important—easy to determine owing to the quick change in curvature of the Poisson curve for changes in frequency.

Poisson curve $e^{-n} \mid \frac{x}{1} e^{-n} \mid \frac{x^2}{2!} e^{-n} \mid \ldots$

Real curve $e^{-\alpha} + e^{-\beta} \mid \frac{\alpha}{1} e^{-\alpha} + \frac{\beta}{1} e^{-\beta} \mid \ldots$
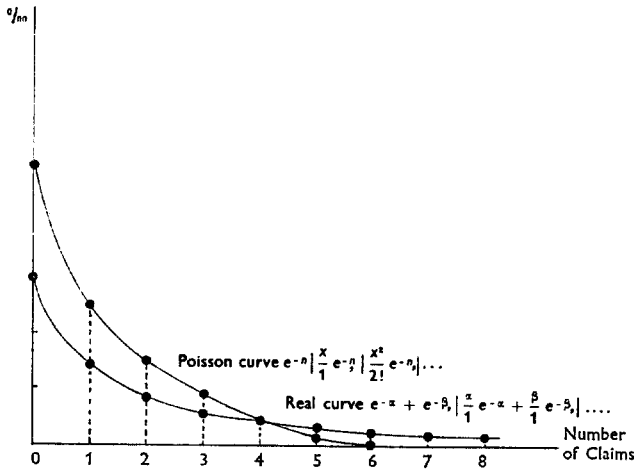
Fig. 1

Definition: Accumulating curve i.e. curve expressing the relative frequency of risk objects = insurances within a risk group $(\alpha, \beta, \gamma, \delta)$ with number of claims $\nu = 1, 2, 3, 4, 5, 6, 7, \ldots$ respectively.

In insurance risk statistics we should distinguish between branches with

very small claim frequency $\varphi \sim 1\text{-}10 \, {}^0/_{00}$ — fire, burglary
moderate claim frequency $\varphi \sim 100 \, {}^0/_{00}$ — accident, third party
large claim frequency $\varphi \sim 50\text{-}500 \%$ (or more) — motor car, "traffic".—

In the first case the number of insurances with multiple claims should be negligeable, in the second case some large risk groups supposed to be fairly homogeneous will furnish test material, and in the last case we will find many risk groups proving, that it is

next to impossible to define homogeneous groups by ordinary tariff parameters.

Some typical figures for motor car were given in my New York (1957) paper, for instance (mean claims frequency = Poisson = 139,2 $^0/00$)

Group	Statistics	Poisson	Calculated by 76,75% 100 $^0/00$ + 23% 250 $^0/00$ + 0,25 % 2.000 $^0/00$
0 claims	874 $^0/00$ *)	869	874
1 claim	112	122	115
2 claims	10,8	8,5	9,7
3 claims	2,1	0,4	1,0
4 claims (or more)	0,2	0,00	0,3

*) Risk group (homogeneous?) ∞ 20.000 insurances

Theorem 5: In insurance risk statistics no tariff parameters will produce real homogeneous risk groups and for motor car the proportion 75 % standard risks, 24 % poor risks, 1 % very bad risks, seems to dominate (extra good risks, say, $\varphi = 50$ $^0/00$ are impossible to discover by existing methods).—The accumulating curve can always be reduced to an exponential trinomial, at the same time utilized for smoothing out the curve.—The corresponding measure of inhomogenity (L), however will prove that for practical purpose and not too small values of n we should always take the homogeneous Poisson and reserve "the inhomogeneous Poisson system" for very specialized researches.

4. *Error theorems—stochastic error of* $(n \mid y \mid r) \infty (\nu \mid \eta \mid \rho)$

Definition: Relative error in the form $(1 \pm 1/\sqrt{n})$ or $(1 \pm v/\sqrt{n})$ and $(1 \pm 1/\sqrt{\nu})$ or $(1 \pm v/\sqrt{\nu})$ is introduced for many reasons:

a. only these two forms are necessary to characterize any risk group,

b. relative error can be used for both total claims amount and risk premium,

c. in statistics relative error is utilized in the same sense as the "standard error" $\bar{x} \pm \sigma$ as a special form of "confidence interval",

d. even when the exact distribution of ν, η, ρ is unknown or undefined the following theorem should be quite correct.

Theorem 6: In any series of stochastic experiments either simultaneous or following each other in time, the difference between "true value" or "value a priori" and "results occurred" or "value a posteriori" should be less than the relative error for at least 50 %, and exceed double the relative error for at most 5 %, (Cf. all ordinary distribution), of the total number of stochastic experiments.

Second error Theorem (7): Standard formulas for relative error in insurance statistics, when the system has converged to Poisson $(\sigma = \sqrt{n})$

$\nu \sim n$. $(1 \pm 1/\sqrt{n})$ and $n =$ unknown $\sim \nu$. $(1 \pm 1/\sqrt{\nu})$

$\eta \sim y$. $(1 \pm \nu/\sqrt{n})$ and $y =$ unknown $\sim \eta$. $(1 \pm \nu/\sqrt{\nu})$

where

$v =$ "Modified Coefficient of Variance" $= \sqrt{1 + V^2} =$

$= \sqrt{\int_0^\infty x^2 \, \psi(x)dx} : \int_0^\infty x\psi(x)dx;$ and $\int_0^\infty x\psi(x)dx = m = \bar{x};$

$\int_0^\infty \psi(x)dx = 1.$

The first formula is found in many text books; the second is a modern product of extensive statistical researches—in Sweden (1954) and other countries—and proving that generally

$v \sim 6$ for "traffic", third party, fire
$v \sim 3$ for most other branches
$v \sim 1$ for short intervals
$v^2 \sim 2$ for excess intervals in motor car 0-1.000-20.000-60.000 Sw.Cr.

My formulas with number of claims and not insurances, seem to be a good simplification—also number of claims is printed in most statistical tables.

Corollary: From these formulas and with our previous notations we might derive:

Claims frequency $= \varphi = \nu/N(I) \sim n/N(I)$. $(1 \pm 1/\sqrt{n})$

$f = n/N(I) \sim \varphi$. $(1 \pm /1\sqrt{\nu})$

Risk premium $= \rho = \eta/N(I) \sim r \cdot (1 \pm v/\sqrt{n}) = y/N(I) \cdot (1 \pm v/\sqrt{n})$;
$r = y/N(I) = \rho \cdot (1 \pm v/\sqrt{\nu})$

Reactions on tariff construction—and tariff work generally.—

When first published these formulas acted as a small bomb on tariff work. Most tariffs are constructed in one of two ways:

a. Tariff centers round some principal tariff classes. For each class premiums are decided, "according to" statistics with standard additions for costs, risk development and catastrophe risk. Minor tariffs should also use statistics and small subtariffs should be formed by a system of additions or factors.

b. Professions, industries or similar subgroups are referred to one standard scale, say, with premium differences of 20 %, and placed on that scale according to risk statistics.

In both cases we need statistics with a relative error of about ± 10 % or less. However, this would correspond to a risk group with 2.500 claims (or for some branches 600 claims)—and that's the bomb.

In practice and for one year the total number of claims for the cooperating group of tariff companies in Sweden might attain

in motor car and traffic insurance	150.000
accident or third party	50.000
fire and most other branches	10.000

At the same time there is a strong tendency for the greater part of the claims to accumulate in a small number of tariff groups. Evidently, the high-flying hopes of a tariff, where most (individual) premiums are determined by statistics will never be realized—on the other hand the effect of the error theorems will be that in "practical risk theory" we will concentrate on methods to extract the maximum effect from our risk statistics.

(In my New York paper (1957) and in some ASTIN papers I have analyzed some methods, which might work in that direction) [1]).

Finally, we should observe that requisitions to the statistical department of a company from tariff men are often impossible to

---

[1]) *Excess claims analysis:* Risk premiums calculated in excess classes (0-1.000-20.000-60.000) *Composite factor analysis.* Method to express risk premiums as product of, say, tariff factor, district factor and yearly factor.

realize—except by pure guesswork—and experienced actuaries should help their young colleagues by this declaration, not once but many times.

5. *Error theorems—the problem of prognosis.*

The scientific risk theory is based on theory of probability and all errors or distributions are given in the stochastic sense. For a large risk class with some 100.000 insurances we might assume that the number of claims occurred is exactly say 10.000.

Using the modern form of error formulas, expressed by a factor with n or ν instead of number of insurances (claims frequency $= \varphi$ and "no-claims frequency" $= 1 - \varphi$), we will find an error interval

$$n \backsim 10.000 \cdot (1 \pm 1/\sqrt{10.000}).$$

We have the somewhat astonishing result, that the unknown number of claims a priori (probability? or sometimes defined by an integral) is almost exactly known, as an error of $\pm 1\%$ being negligeable in practical risk statistics. Even the corresponding risk premium a priori is almost exact, say

$$\mu = 50 \cdot (1 \pm 6/\sqrt{10.000}) = 50 \pm 3 = \text{interval } 47\text{-}53,$$

or well within our critical "10 % limit".

(I have been proud to declare to outsiders, that sometimes even the risk premium (or frequency) a priori is accessible to modern statistical methods).

The same fact, however, applied to some very stabilized system of insurances (tariff)—or to general systems (or branch totals) when introducing claims frequencies and risk premiums a priori— will also prove that differences in the statistical figures from year to year are only partially due to pure "stochastic oscillations" and even that "variations" in the primary risk or risk a priori will generate the greater part of the differences. In motor car insurance we have analyzed these variations and obtained some results, e.g. the influence of the weather and the influence of "new insurances from the last two or three years". These being quite arbitrary —even the strictest deterministic philosopher can only declare, that they are determined by chance in another sphere, without

any stochastic relations to insurances or risk theory—we have full right to insist on the term "arbitrary functions".

Prognosis:

All prognosis in risk statistics and general statistics is based on a series of statistics for previous years and the principle of continuity. As a rule quotients are better stabilized than absolute figures, and results are often much improved, if we find the right type of quotient. Technically our best figures (for large tariffs) should be expressed by primitive curves, drawn by hand as a sort of mean curves. Trend should always be included as a hypothesis —at least in the linear form.
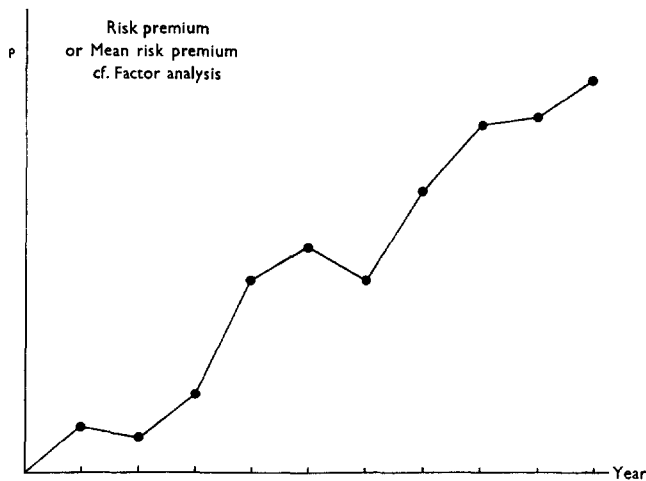

Fig. 2

In the construction of extrapolation curves we should observe that the sequence of risk elements is not stabilized, as some sort of population and that the "population" of insurances only repeats risk objects but neither risk situations nor risk elements or claims.

More precisely, our studies of the probable oscillations of risk figures according to chance have proved that these should be small —much smaller than those often registered in practical statistics —and consequently that effective variations of the primary risk a priori is by far the most important factor in the analysis of risk differences between consecutive years.

Finally, we might observe, that increasing risk figures will always go by leaps and jumps, something like: after two or three years "unaltered", a large increase which soon ceases—very uncomfortable for the tariff actuary.

Third error theorem (8): In actual statistics and in real life the sequence of risk elements (and the Poisson expression, where $[n \mid \bar{\psi}(x)]$ is often smoothened out further to a continuous parameter function in time) will always represent a complex of risk situations with most varying characteristics and depending on arbitrary functions ("aging" of risks and climatic factors being among the most prominent), and prognosis is considerably more uncertain than it should be, if chance was the principal cause of error.

Corollary: In principle at least the error theorem is applicable to most statistical research work—including sampling—with the general effect, that ordinary error limits should be tested anew (for eventual modifications).

The reason is, that in one particular field we have proved, that the effect of the variations between years is more important than the oscillations by chance—and that similar irregularities might be quite common, once we really look for them.

Also there is a great difference between genuine populations, made up by individuals with stabilized properties and with a small proportion of enterings and exits, and nominal populations, for example one year's claims, generated by a fluctuating risk system.

Note 1: "Research work in medicine analysis of road accidents by cases from a single hospital etc. might be valuable—but formal error limits are not". "In sampling we must always be on guard against the feeling, that the active probability is equalized, all times we have found no good reason for differences".

Note 2: In some discussions this theorem has been described as a small bomb—in actualizing problems from the next chapter.

RISK THEORY—RELATIONS TO THE THEORY OF PROBABILITY.

1. *Introduction*

Even the most primitive forms of insurance, for instance burial associations in old Rome, general contributions in case of fire in the Middle Ages, early forms of marine insurance by groups of

individual merchants, and pension funds (based on actual mortality) in our times are based on some general concepts, say:

a universal risk for damage and pecuniar loss,
an estimate of the mean risk for any large group,
extra loss (except the mean risk) for large groups is not very probable.

Evidently, these concepts belong to the theory of probability and when this theory was worked out in the early XIXth century, the insurance theory got a mighty instrument both for exact definitions and for risk analysis.

For a long period, however, actuarial problems were concentrated, say, to general risk structure in life insurance (in particular the dependence on age and sex, on entrance year etc.), and to a small volume of insurance risk statistics, both in life insurance and in some branches of non-life insurance. In most cases the number and classification of insurances were based either on lists of new insurances, alterations and cancellings, or on a manual counting of some form of insurance cards (sorted before on risk classes); claims and eventual claims amounts were calculated directly from claims cards or claims acts.

For these problems probability is very well described by a system of urns containing a (large) number of white and black balls, and defining:

Probability a priori $= p = n/N$
Stochastic experiments $=$ drawing of balls according to specified rules.

Stochastic variable $= \nu =$ number of successes ($=$ white balls) for an experiment. (A better model is obtained if the number of balls drawn is determined by a first drawing from an urn with numbered balls, and all succeeding drawings are made from a new urn, where every ball carries a number $=$ amount).

Similar models were often used in the analysis of insurance risk problems and risk theory even in the late twenties (Analysis by Wold on "rain insurance" to describe and calculate premiums for different "periods under risk", by the Polya-Eggenberger urn system).

2. *Modern Probability Theory*

In the modern theory of functions of one (or more) real variables the Lebesgue measure and Lebesgue integrals take a central place, simply because they lead to a "closed" system and to the proof of some theorems which could not be proved before.

In order to utilize the Lebesgue system most modern theories will define probability as an abstract number p where $0 \leq p < 1$ —given a priori—and a stochastic (or random) variable ξ by a general distribution function $F(x)$, increasing from 0 to 1 in the interval $-\infty$ to $+\infty$ (or by a frequency function $f(x) = F'(x)$, if this function exists). The old limit definition $p = \lim n/N$, where n = favourable cases and N = total cases = number of stochastic experiments has almost disappeared as a formal definition of p.

On the other hand, there was no intention to sacrifice precision and to introduce probabilities without contact with real problems (Cf. Keynes' purely abstract theory), and the conditions $p = n/N$ for populations and $p = \lim n/N$ for a series of identical stochastic experiments are still conserved, we might say as test conditions, where differences will invalidate the previous probabilities a priori.

By these conditions it is possible to divide probability theory into two principal classes:

a. Population and sampling theory characterized by $p = n/N$ or $\Sigma\, p_i = 1$ or both.

In demographic statistics—and generally in descriptive statistics —analysis is concentrated on a population of individuals, which are classified according to many different parameters, separately or in combination. All classifications will lead to definite frequencies

$$p = n/N \text{ or } p_i = n_i/N \qquad \Sigma n_i = n \text{ and } \Sigma p_i = 1$$

Generally the population is stabilized with a moderate number of exits and enterings during one observation period (year). Also, more important figures and frequencies are taken from official registers or reports.

On the contrary, minor questions are often analyzed by sampling and this stabilized population represents an ideal subject for sampling, as all probabilities a priori and all distribution functions exist in the form of definite frequencies. Evidently the hypothesis or definition of probabilities as abstract numbers existing a priori

is well founded. In many cases sampling can be defined by similar stabilized populations, in other cases stabilization might be less advanced, and should be taken as an hypothesis or idealization of the actual statistic material (in such cases error formulas should be taken with some care).

As a counter-example we might choose the insurance claims in some branch during one year, these being generated by a fluctuating and instable chain of risk situations, and varying from year to year.

b. Stochastic (or random) experiments, which can be repeated indefinitely under exactly the same conditions, characterized by $p = \lim n/N$ (number of experiments decided by operator).

The principal idea in this system of random experiments—defining (according to classic theory) or verifying (in modern theory)—one (or more) numerical probabilities by the quotient of $n =$ number of successes and $N =$ number of experiments, is evidently, that the difference expressed by the error factor

$$(1 \pm pq/N)$$

tends to zero, when N tends to infinity (which might be taken as a full proof of the limit formula).

Applications in a wide range to games of chance, to technology and to biology are without any immediate interest for our problems. Some of these have the character of sampling to determine the primary probability by the experimental frequency, others are more concerned with results or differences.

Our principal object, however, was the study of how the modern theory of probability could be applied to modern risk theory—how to state our exact definition of probability—and generally to form our theoretical and statistical model by means of the probability theory.

And there comes the real surprise:

Theorem 9: Standard definitions in the theory of probability are

$$p = n/N \text{ or } p = \lim n/N$$

not applicable to risk theory according to the "third error theorem".

In other words claims represent an ever-changing series of risk situations and claims characters and a priori probabilities or characteristics n and $\bar{\psi}(x)$ are "distributed" and vary from period

to period and from year to year with a "trend"—all making any time definition to join probabilities as abstract numbers a priori with statistic results a posteriori more or less void or vacant.

3. *Stochastic disorder—a neglected attribute to the results of statistics and stochastic experiments. Insurance and claim characteristics.*

One of my first contacts with randomized systems—free for speculative analysis—was the kinetic theory of gases, and the numerical expression for entropy "constant plus the logarithm of the probability of the actual state" is still an excellent example of a measure for stochastic disorder, and for that tendency in stochastic systems to sink into a bottom state of undefinedness. Gas molecules will even give us one of the best definitions of stochastic disorder: by placing or projecting our cards or figures or results on well-defined gas molecules and waiting a few minutes we will get a system with perfect disorder and all systems to follow will represent new forms with comparatively high probability.

Elementary analysis of stochastic disorder is founded on:

Definition: Stochastic disorder = sequence of results from (constant and) completely independent random experiments.

Test method (also supplementary definition): All properties of the sequence should be identical in any arithmetical series with indices

$$i = kt + l \qquad (l = 1, 2, 3, \ldots k \; ; t = 0, 1, 2, \ldots)$$

Note 1: Test by arithmetic series compensates eventual variations in risk (claims frequency) in the analyzed year, also when analyses goes from tariff to tariff or from district to district (not to be recommended; time is the best order).

Note 2: Standard test utilizes the successive arithmetic series with the same difference = k; generally several test series with different differences, say values from k_0 to $k_0 + j$, $= (k_0 + j/2)$ $(j + 1)$ different test series with practically the same difference.

Theorem 10: Normal insurance praxis leads to a series of claims (and claim numbers) in perfect stochastic disorder, the sequence being generated by risk elements with a small probability for claim and a much larger for no-claim. Any change to no-claim should have exchanged all individual claims from that point in all arith-

metic series simultaneously. Generally policy numbers—limited to last three figures—can also be considered to give insurances in perfect disorder, rearranged by many years' cancellations.

Theorem 11: Different arithmetic series with objects in perfect stochastic disorder will generate a perfect system of stochastic experiments, with the sole restriction that the number of claims is fixed and that "superexcess claims" (very large claims) are restricted to "the selection of the year". (In principle interest is here concentrated on risk theory).

Theorem 12: These series will follow ordinary stochastic rules and if we make a correction for above small defects, they can be used for an independent determination of the modified coefficient of variance in the relative error $(1 \pm v/\sqrt{n})$; also for a preliminary analysis of the distribution of v and consequently of the distribution $\theta(y)$ of total claims amount y.

Theorem 13: Test on abstract probabilities. The existence according to the modern theory of probability of abstract probability numbers a priori is tested by the demand that for any arithmetic series or a complex of series and for any set of systematic grouping or classifications determined a priori, differences between the sum of abstract numbers a priori and corresponding claims result a posteriori should follow our standard rule for relative errors, and differences should be distributed conformly to $\theta(y)$.

Note 1: Further studies in stochastic disorder will prove both, that we have found a convergence test with the same "power" as the more conventional limes test and also some new statistic methods of analysis of some importance. Consequently, I demand the same right to speak of probabilities a priori as any competing system.

Note 2: Collective risk theory has chosen another method to solve the problem. Restriction to a "stationary stochastic process" conserves the right to form limit expressions, and the convergence to Poisson might be proved by the methods in this paper, but is generally introduced either as a hypothesis (or "idealization"), or else by forming the mean values for not to small periods (year?) of both abstract probabilities a priori and claims occurred a posteriori.

SPECIAL PROBLEMS:

Inhomogeneous risk groups:

In principle all risk classes (α, β, γ, δ) are inhomogeneous and for motor car we have proved the proportion 75 % ordinary risks (25 % very good?) 24 % poor risks and about 1 % very bad risks.

Every time different groups are blended, homogenity might increase or diminish but the risk figures (mean values) are governed by the measure of inhomogenity L (or first error theorem) and generally homogenity will advance (L proportional to volume) even if the above proportion is maintained. The latter, however, is only of interest in the study of individual insurances.

Otherwise they are easily expressed by exponential polynomials or in a more restrictive sense by "Polya-Eggenberger" or by the negative binominal distribution.

Theorem 14: When joining similar risk groups, L will increase in proportion to n, proving that the effects of inhomogenity— expressed by exponential polynomials in n or by the negative binominal—will disappear long before that point ($n \backsim 100$) where risk figures can be utilized. And if not, well, then it is time to return to the primary binominal elements.

Conditional primary probabilities—"ramified" risk lines. (New York 1957)

By supernatural sight, we should have the power to follow the risk line $\mathfrak{L}$ from the beginning to the end of the year, and to study risk elements (i) and claims occurred (ξ_j). So far all (N) risk elements on the line $\mathfrak{L}$ are independent.

Further, however, we might find (ramification) points on the risk line, where the risk in the future was subject to random decision with at least two alternatives. Assuming there are always only two alternatives we have:

unaltered risk line $\mathfrak{L}$ with the probability $1 - P_j$;

a new risk line $\mathfrak{L}_j$ (with some new and some old risk elements) with the probability P_j, drawn from the ramification point to the end of the year.

We will now extend our risk system with a series of new risk

lines $\mathfrak{L}_1$ $\mathfrak{L}_2$ $\mathfrak{L}_3$ . . . in finite number, all starting from $\mathfrak{L}_0$ and with the respective probabilities P_1 P_2 P_3 . . .

However, the system is still incomplete as there might be new ramification points on these new lines. By systematic extension we finally arrive at a complete system, still in a finite form.
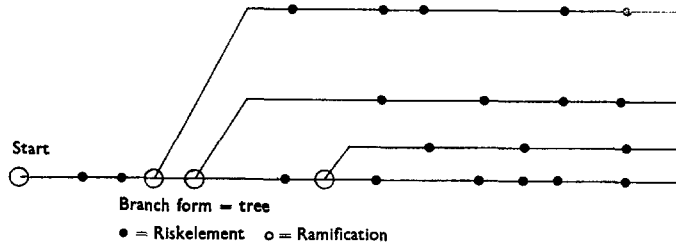

Fig. 3

A simple counting will always give us the combined probability to arrive at some chosen part λ of the branch system. This ramified branch form or tree represents the first expression of a complete system of conditional lines of risk elements.
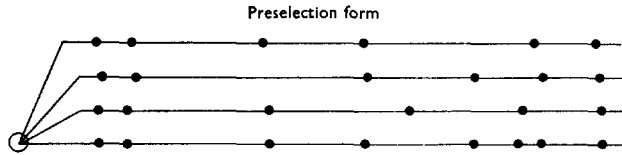

Fig. 4

The second system or preselection is obtained if we "take a sharp knife and cut through all risk lines lengthwise from a ramification point backwards to the starting point of the year, leaving the same risk elements on both lines and conserving the probability of the ramification point". Continuing in this way we get a bundle of parallel risk lines, each with a final series of risk elements, which are independent simply because all random decisions are moved to one starting point and expressed by the relative probability to choose any specified risk line.

So far we have only admitted two-way choice, but we have only to interpolate several ramification points between two risk elements to get room for a many-way choice. In this way we have

proved the following important theorem for finite systems (As to different types of conditional probabilities, affecting future risk elements, vide New York 1957 and other papers).

Uniformisation theorem (15): In a finite system of risk elements (i)

$$\mathfrak{L}_0 \colon \{ p_i \mid \psi_i(x) \mid \theta_i(y) \} \qquad i = 1, \ 2, \ 3, \ldots N$$

the risk from a specified point in time (T) might depend on a precise random choice

$$R_j \colon (P_j{}^{(1)} . \mathfrak{L}_j{}^{(1)} \mid P_j{}^{(2)} . \mathfrak{L}_j{}^{(2)} \mid P_j{}^{(3)} . \mathfrak{L}_j{}^{(3)} \mid (1 - P_j{}^{(1)} - P_j{}^{(2)} - P_j{}^{(3)}) . \mathfrak{L}_0)$$

representing the tree-form of the risk system. This system can always be transformed to a preselection at the starting point $T_0 \ (\leqq T \leqq T_1)$ between the same total number of alternatives

$$\{ P_k \mid \mathfrak{L}_k \} \qquad k = 1, 2, 3, \ldots K \qquad \Sigma \, P_k = 1$$

followed each by one single risk line made up of independent risk elements.

All finite risk systems with locally well-defined ramifications can be transformed into both tree-form and preselection form. In particular all risk systems made up by risk elements (insurances), with well-defined rules for more than one claim can be reduced to this form.

(In order to apply this theory to Poisson we must make an easy extension to enumerable systems).

In this problem both exponential polynomials in n and the two-constant negative binominal distribution may be used for increased applicability as the convergence to simplified forms is not evident.

THE TRANSFORMATION PROBLEM

Risk statistics will furnish a series of figures for large and moderate statistical risk groups such as $n, f, m, r - v, \varphi, \xi, \rho$ and also $\psi(x)$ (relative frequency of claims with different magnitude).

$\psi(x)$ represents no "distribution" in the ordinary sense of the word and different parts of this "claims curve" are absolutely independent and permit no conclusion from one part of the curve to another; also with increasing material the curve will only slowly change its form and does not converge to the normal distribution (it is "additive" and not "multiplicative").

The important real distributions of the stochastic variables $(\nu, \eta) \curvearrowright (n, y)$ can never be determined by real statistics owing to the third error theorem. In most cases the hypothesis that ν (n) is distributed by the simple Poisson should be very good, but η (y) is more difficult.

Many authors will try expressions derived from the normal distribution or the Edgeworth series, and apply higher moments (3, 4, 5), which is not very good, as the mean value is not stabilized, and these higher moments are very sensitive to a small part of $\psi(x)$, so that the expressions are determined from a few very large claims, also most formulas are constructed as asymptotic formulas.

Consequently we have here a neat transformation problem

$$[n, m, \psi(x)] \to \theta \ (n \mid M = n.m \mid y)$$

which in principle should be quite simple (by calculating some asterisk powers), but in practice represents an intricate calculation problem.

In my researches I have always concentrated on exponential polynomials from two reasons:

1a. The claims curve is logarithmically convex = convex against the system of exponential curves

$$\alpha \, . \, \beta e^{-\beta x}$$

b. The convexity is weak and an exponential monomial (and still better binomial) will give a very good approximation for quite large intervals (Note: quite good also for Bessel functions) and furnish a very good smoothing out.

c. An exponential binomial as approximation for both functions leads to quick and simple numerical integration of asterisk products for many types of functions.

d. An exponential trinomial (or tetranomial) will express the whole claims curve from $+ \infty$ to o with an effective maximum error of, say, 2 or 3 %, and the corresponding claims amounts by a "may be correct (exact)" series of values.

2. Poisson + exponential monomials

$$\alpha \, x^r \, e^{-\beta x}$$

represent the characteristics (according to the theory of differential

equations) or elementary standard solutions both to the problem of asterisk powers, and the transformation problem.

I will now simply apply to [1] giving some actual examples of asterisk multiplication and Laplace transforms

$$\mathcal{L}(f(x)) = \int_0^\infty e^{-sx} f(x)dx$$

Theorem 16: Assuming that the stochastic variable (number of claims) ν is Poisson distributed with the a priori value (or mean value) n, the probability for exactly k claims is expressed by

$$e^{-n} \cdot \frac{n^k}{k!} \quad \text{where} \quad \sum_{k=1}^\infty e^{-n} \cdot \frac{n^k}{k!} = e^{-n}(e^n - 1) = 1 - e^{-n},$$

where e^{-n} (no claims) is negligeable and

$$\sum_{k=1}^\infty k\left(e^{-n} \cdot \frac{n^k}{k!}\right) = n \sum_{k=1}^\infty e^{-n} \frac{n^{k-1}}{(k-1)!} = n \cdot e^{-n} \cdot e^n \cdot = n \text{ mean value}$$

Assuming further that the claims curve $\psi(x)$ is expressed by the exponential monomial e^{-x} the asterisk powers or claims curves in case there are exactly k claims is expressed by

$$k = 1 : e^{-x}, \ k = 2 : xe^{-x}, \ k = 3 : \frac{x^2}{2!}e^{-x}, \ k = 4 : \frac{x^3}{3!}e^{-x}, \ k : \frac{x^{k-1}}{(k-1)!}e^{-x}$$

Consequently, the claims distribution for k claims and for all separate values of k taken with their respective probabilities is

$$k \text{ claims} : e^{-n} \cdot \frac{n^k}{k!} \cdot e^{-x} \frac{x^{k-1}}{(k-1)!}$$

$$\text{All values} \quad \sum_{k=1}^\infty e^{-n-x} \frac{n^k}{k!} \cdot \frac{x^{n-1}}{(k-1)!} = H(nx)$$

Note 1: $H(nx) \backsim e^{+2\sqrt{nx}}$ for $nx \to \infty$ and $-n-x+2\sqrt{nx} =$
$$= -(\sqrt{n}-\sqrt{x})^2 = \text{convergence}$$

Note 2: Hyperbolic Bessel function $I_1(2\sqrt{\mu}) = \sqrt{\mu} \sum_{k=0}^\infty \frac{\mu^k}{k!(k+1)!}$

Private form $B_1(x) = \frac{I_1(2\sqrt{\mu})}{\sqrt{\mu}}$; $H(nx) = \frac{n}{\sqrt{nx}} \cdot e^{-n-x} I_1(2\sqrt{\mu})$

Finally we have proved that the elementary transformation problem

$$(n \mid \psi(x) = e^{-x}) \to \theta(y)$$

is resolved by above hyperbolic Bessel function $I_1(2\sqrt{nx})$ and that more precisely

$$\theta(n \mid \psi(x) = e^{-x} \mid y) = n \cdot e^{-n-y} \cdot \sum_{k=0}^{\infty} \frac{(ny)^k}{k!(k+1)!} =$$

$$= \frac{n}{\sqrt{ny}} \cdot e^{-n-y} \cdot I_1(2\sqrt{ny})$$

(claims curve, relative frequency of η around the mean value)

Corollary: In an exponential polynomial—say tetranomial—we will often put

$$
\begin{array}{llll}
n\alpha_1 & n\alpha_2 & n\alpha_3 & n\alpha_4 \quad \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1 \\
& & \Psi(x) = \alpha_1 e^{-\beta_1 x} + \alpha_2 e^{-\beta_2 x} + \alpha_3 e^{-\beta_3 x} + \alpha_4 e^{-\beta_4 x} \\
m_1 & m_2 & m_3 & m_4 \quad \alpha_1 m_1 + \alpha_2 m_2 + \alpha_3 m_3 + \alpha_4 m_4 = x \\
& = 1/\beta_1 & = 1/\beta_2 \;\; = 1/\beta_3 \;\; = 1/\beta_4 \\
& & \psi(x) = \alpha_1 \beta_1 e^{-\beta_1 x} + \alpha_2 \beta_2 e^{-\beta_2 x} + \alpha_3 \beta_3 e^{-\beta_3 x} + \alpha_4 \beta_4 e^{-\beta_4 x}
\end{array}
$$

leading to four similar Bessel functions

$$n\alpha\beta \, e^{-n\alpha - \beta y} \cdot B_1(n\alpha\beta y) = \frac{n\alpha\beta}{\sqrt{n\alpha\beta y}} \cdot e^{-n\alpha - \beta y} \cdot I_1(2\sqrt{n\alpha\beta y})$$

Theorem 17: (my "first solution", Colloquium Leppilampi 1956, New York 1957): Systematic researches (in traffic, motor car and other branches) have proved that an exponential trinomial or tetranomial will always be sufficient to express

$\Psi(x)$ = distribution of number of claims $\geq x$ $\quad \psi(x) = -\Psi'(x)$

$\int_x^{\infty} x \, \psi(x) dx$ = *distribution of claims amounts* slightly smoothened out and with a maximum relative error of some percent.

Consequently, the transformation problem is resolved with a small relative error by three asterisk multiplications.

$$\left[\frac{n\alpha_1\beta_1}{\sqrt{n\alpha_1\beta_1 y}} \cdot e^{-n\alpha_1 - \beta_1 y} \cdot I_1(2\sqrt{n\alpha_1\beta_1 y}) \right] * \ldots * \ldots$$

$$* \left[\frac{n\alpha_4\beta_4}{\sqrt{n\alpha_4\beta_4 y}} \cdot e^{-n\alpha_4\beta_4 y} \cdot I_1(2\sqrt{n\alpha_4\beta_4 y}) \right]$$

Actual integrations are simplified by using a special integrating method, similar to the methods by Simpson and Gauss but based on exponentials. Making a small table for, say, $n = 100, 200, 500, 1,000, 2,000, 5,000, 10,000, 100,000$ conserving some values of y from the integration process, we obtain a complete solution of the transformation problem and not only an asymptotic solution slowly converging (*Cf.* next theorem).

Up to this point analysis has been made without detail reference to exponential polynomials.—Now it is necessary to combine different elements. A good example from practical analysis is given by motor car and traffic 1950, but quite similar to modern figures (values in the unit 1,000 Sw.Cr. ∞ m = mean value of individual claims)

$$\Psi(x) = 0.75 \cdot e^{-4x} + 0.18 \cdot e^{-x} + 0.05 \cdot e^{-0.3x} + 0.02 \cdot e^{-0.05x}$$
$$\psi(x) = -\Psi'(x)$$

Assuming $n = 10.000$ the mean value of $y = M = 10.000.000 = 10.000$ units m (Excess mean values approximately $0 .. 1.000 : \infty$ 300, $1.000 ... 20.000 : \infty$ 2.500, $20.000 : \infty$ 40.000).

Second solution: Integrating by asymptotic formulas and by the properties of one "normal distribution" serving as reference.

(Only some head points of the proof are given)

1. $\theta(n \mid \Psi(x) = \alpha_1 e^{-\beta_1 x} + \alpha_2 e^{-\beta_2 x} + \alpha_3 e^{-\beta_3 x} + \alpha_4 e^{-\beta_4 x} \mid y)$

2. Elementary solution θ $(n \mid \psi(x) = \beta e^{-\beta x} \mid y) =$

$$= n\, e^{-n-\beta y} \sum_{k=0}^{\infty} \frac{(n\beta y)^k}{k!(k+1)!} = \frac{n}{\sqrt{n\beta y}} \cdot e^{-n-\beta y} \cdot I_1(2\sqrt{n\beta y})$$

3. Asymptotic formula (according to Hankel)

$$I_1(\mu) \infty \frac{e^{\mu}}{\sqrt{2\pi\mu}} \left(1 - \frac{0.375}{\mu} + \dots\right)$$

4. The following important lemma is proved:
Assume $n_1 m_1 + n_2 m_2 = \overline{n.m}$ $(\overline{n} = n_1 + n_2)$ and $f_1(x) * f_2(x) = F(x)$
Then $f_1(z) * f_2(z) = F(z)$ is valid with centralized variables

We now introduce the centralized variable
$y = n.m + z$; $\beta = 1/m$; $\beta y = n + \beta z$; $2\sqrt{n\beta y} = 2n.\sqrt{1 + \beta z/n}$

5. Elementary solution $\theta(n \mid \beta e^{-\beta x} \mid y)$

$$\sim \frac{n}{n\sqrt{1+\beta z/n}} \cdot \frac{e^{-(\sqrt{n+\beta z} - \sqrt{n})^2}}{\sqrt{2\pi.\, 2n\,\sqrt{1+\beta z/n}}} \left(1 - \frac{0,375}{2n.\sqrt{1+\beta z/n}} + \ldots\right)$$

6. First form of exponential $e^{-n}\left(\frac{\beta z}{2n} - \frac{\beta^2 z^2}{8n^2} + \ldots\right)^2$ proves that it is dominated by $e^{-\beta^2 z^2/4n}$ and converges although in another form $e^{+\beta^3 z^3/8n^2}$ is made free

7. Second form of exponential $e^{-\beta^2 z^2/4n} \cdot e^{+\beta^3 z^3/8n^2 - + \ldots}$ proves that it can be divided into a strict normal term $e^{-\beta^2 z^2/4n}$ and the series (factor):

$$1 + \frac{\beta^3 z^3}{8n^2} - (\ )\,\frac{\beta^4 z^4}{16n^3} + (\ )\,\frac{\beta^5 z^5}{(\ )n^4} \ldots$$

We have now the full material to prove the following theorem although there is still some lengthy detail work.

Theorem 18: Applying the Hankel asymptotic formulas to the elementary solution of the transformation problem by a Bessel function $I_1(2\sqrt{n\beta y})$ this is transformed to the product of a "normal distribution" $e^{-\beta^2 zz/4n}$ and a set of short series with no, one or two terms to obtain the order $(\beta z/n)^2$ or $\beta z/n^2$. However, all groups of the order $1/n^2$ have the maximum error $O(1/n^3)$ and for groups with at least 100 claims this signifies $1/1.000.000 = $ "numerically exact" in further work.

So far I have not mentioned the last point in the proof: integrals are taken from $0 \ldots + \infty$ in y and $-n.m \ldots + \infty$ in z. In this point, $(-nm)$ however, the exponential is extremely small, say, 10^{-100} for some actual forms and should be taken to $-\infty$ without changing anything. By this simple device all integrals

$$\int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2}.\, x^a dx \text{ can be calculated from the Laplace transforms}$$

$$\int_0^\infty e^{-xs}\,\psi(x)dx.$$

We will now finish by the following two theorems on exact solutions of the transformation problem.

Theorem 19: Assuming that number of claims $n \geqq$ 10,000 and $\varphi(x)$ in the form of an exponential tetranomial are given exactly, we can form a numerically exact expression for the distribution function

$$\theta(n \mid \psi(x) = \alpha_1\beta_1 e^{-\beta_1 x} + \alpha_2\beta_2 e^{-\beta_2 x} + \alpha_3\beta_3 e^{-\beta_3 x} + \alpha_4\beta_4 e^{-\beta_4 x} \mid y)$$

as the product of an exponential e^{-ay^2} with a short asymptotic series—valid from that point upwards so long as $\psi(x)$ is not changed.

Theorem 20: Under the same conditions we can form another expression as a short series in

$$I_1, \ I_2 \ I_3 \ I_4$$

All coefficients in both series can be calculated directly from n and the coefficients of the exponential tetranomial.

Detail analysis will probably prove that the "normal formula" is only valid from $n = $ 10.000, but the Bessel series from, say, $n = $ 1000.

Summary of the transformation problem.

The transformation problem dealt with in the last part of this context is an intricate problem: statistics will give us frequencies and risk premiums and the "claims curve$=\psi(x)$" (relative frequency of individual claims x of different magnitude); the still more important question of the distribution $\theta(n \mid \psi(x) \mid y)$ of total claims amount y cannot be solved by statistics, simply because risk will vary between risk classes and between years. The important problem is to calculate $\theta(n \mid \psi(x) \mid y)$ from $n = $ number of claims and $\psi(x)$.

After some discussions with Swedish colleagues I will give some further remarks on the integration method.

The new method to determine $\theta(n \mid \psi(x) \mid y)$ is based on the following technical system:

1. By risk statistics we have determined $\psi(x) = $ relative frequency of claim amounts and $n = $ a priori value of number of claims. We want to apply the method from $n = $ 10.000 (or $n = $ 1.000), where n should follow Poisson "exactly".

2. From $\psi(x)$ we calculate an exponential polynomial, representing $\psi(x)$ almost exactly, or an upper and a lower limit curve $\psi+$ and $\psi-$. To have full control of later figures, it is important that the maximum relative error should be small.

Note. Some authors prefer modified Poisson systems. In most cases the general method can still be applied—mutatis mutandis—so long as the material includes exact statistical values of $\psi(x)$ etc.

3. Poisson + $\psi(x)$ (= exponential monomial) leads to a simple elementary solution

$$\theta(n \mid \psi(x) \mid y) = n \cdot e^{-n-\beta n} \sum_{k=0}^{\infty} \frac{(n\beta y)^k}{k!(k+1)!} =$$

$$= \frac{n}{\sqrt{n\beta y}} \ e^{-n-\beta y} \ I_1(2\sqrt{n\beta y})$$

4. Now it is time to introduce the Hankel asymptotic formulas

$$I_1(\mu) \cong \frac{e^\mu}{\sqrt{2\pi\mu}} \left(1 - \frac{0,375}{\mu} + \frac{0,11719}{\mu^2} - \frac{0,10253}{\mu^3} + \frac{0,14420}{\mu^4} - \cdots\right)$$

Instead of normalizing we will now put

$$y = n \cdot m + z \quad \beta = 1/m \quad \beta y = n + \beta z \quad 2\sqrt{n\beta y} = 2n\sqrt{1 + \beta z/n}$$

and conserve the mean value $= n \cdot m$ (of total claims amount $= y$ or η)

$$\theta \ (n \mid \psi(x) = \beta e^{-\beta x} \mid y) = \frac{n}{n\sqrt{1+\beta z/n}} \ \frac{e^{-(\sqrt{n+\beta z}+\sqrt{n})^2}}{\sqrt{2\pi} \cdot 2n\sqrt{1+\beta z/n}} \times$$

$$\times \left[1 - \frac{0,375}{2n\sqrt{1+\beta z/n}} + \frac{0,11719}{(2n\sqrt{1+\beta z/n})^2} - \cdots\right]$$

By this method the elementary solution is divided into three components

a. An auxiliary function = exact normal distribution, in z
b. Supplementary terms to the auxiliary function
c. Asymptotic terms representing the difference between asymptotic formula and exact Bessel function.

The auxiliary exponential factor has the form

$$e^{-n\left(\frac{\beta z}{3n} - \frac{\beta^2 z^2}{8n^2} + \cdots\right)^2} \sim e^{-\frac{\beta^2 z^2}{4n}} \ e^{+\frac{\beta^3 z^3}{8n^2}}$$

Combining all facts it is easy to prove

a. All integrals can be taken from $-\infty$ to $+\infty$ instead of $-n$ to $+\infty$
b. Assuming as before that total $n > 1.000$ (and for all partial

groups $n > 100$) all adjustment series can be limited to two terms and an error term in $O(z^3/n^3)$ and combined to one adjustment of the form

$$1 + az/n + bz^2/n^2 + O(z^3/n^3)$$

c. The asterisk product of two exponential monomials is effectively reduced to a problem of the form

$$e^{-t^2/2} (1 + \bar{a}_1 t + \bar{b}_1 t^2) * e^{-t^2/2} (1 + \bar{a}_2 t + \bar{b}_2 t^2)$$

Finally the complete problem with an exponential trinomial or tetranomial is solved by successive asterisk multiplication.