Zeitschrift: Mitteilungen / Vereinigung Schweizerischer Versicherungsmathematiker = Bulletin / Association des Actuaires Suisses = Bulletin / Association of Swiss Actuaries

Herausgeber: Vereinigung Schweizerischer Versicherungsmathematiker

Band: 59 (1959)

Artikel: The theory of random processes and actuarial statistics : dependent and independent probabilities

Autor: Klinken, J. van

DOI: https://doi.org/10.5169/seals-966816

The Theory of Random Processes and Actuarial Statistics

Dependent and Independent Probabilities

By J. van Klinken, Amsterdam

Summary

In this paper the use of the theory of stochastic processes for some fundamental actuarial concepts is stressed. This relates especially to the subject of "dependent" and "independent" probabilities, important in the actuarial theory of the social insurance. Apart from this also some questions of interval estimation for the number of claims in the near future, regression and comparison of risks are dealt with. Special attention is drawn to some approximative random processes, f. i. the Poisson-process which makes it possible to simplify the statistical calculations.

Contents

1. Introduction

In more advanced textbooks on actuarial mathematics usually the reader still finds that the author starts with some considerations on the so-called continuous method. Intensities are introduced. With these intensities functions $l_{(x)+t}$ are formed and afterwards ratio's of function values in distinct points of time are interpreted as probabilities. For instance $q_{(x)+t} = \frac{l_{(x)+t} - l_{(x)+t+1}}{l_{(x)+t}}$ is interpreted as the probability to die. This method of treating some fundamental actuarial concepts (see f. i. the well-known treatise of E. Zwinggi, Versicherungsmathematik, Basel, 1945, 1. Teil, 2. Kapitel) has some drawbacks. In the first place this method ist not quite consistent. Anyone who knows something about probability theory and more especially about the theory of random processes feels that here in a strange and inconsistent way deterministic and probabilistic elements are mingled. If it concerns only the traditional theory of mortality rates and life annuities this criticism seems rather too strong and in a certain sense superfluous. By this last remark is meant that several authors on statistical subjects, at any rate indicate that the function $l_{(x)+t}$ ist to be interpreted as an expectation and consequently the ratio's $\frac{l_{(x)+t+1}}{l_{(x)+t}}$ as probabilities of a binomial distribution. If we have several groups with transitions between the groups and also the possibility of leaving a group without transition into another group, the problems are not so trivial. The mingling of deterministic and probabilistic approaches leads to unclear concepts and interpretations. This relates especially to the so-called theory of "dependent" and "independent" or "partial" and "absolute" probabilities. With regard to the problems there are mainly two situations involved:

a) There is a group with several possibilities of leaving.

b) Transitions between two or more groups.

An example of a) is f.i. the decrease of a group of widows by remarrying and dying; for b) we have as example the development of the insured into active insured and those who are ill etc. In my opinion this theory of dependent and independent probabilities can be treated adequately if there is made use of the theory of random processes. And this is especially true if we want to use intensities and still interpret certain ratio's as probabilities. We stressed here the probabilistic approach. Before using mortality rates, accident probabilities, the actuary is confronted with the problem of estimating and comparing these values.

In statistical text-books one finds several methods dealing with these problems. Now in actuarial practice most variables are approximately Poisson-variables. Unfortunately, the statistical methods on the Poisson-model are usually missing or very incompletely dealth with. However, for the actuary who uses statistical methods, the methods based on the Poisson-model are of fundamental importance.

If the intensities or probabilities have been chosen we may calculate the development of the several groups of insured in the future. Again in certain applications it has sense to consider this development as a stochastic process. If this is true, the actuary may wish to have interval estimates for this development. If we take as example the development of a group of insured in premium payers and invalidity pensionholders we may wish to have interval estimates for the number of pensionsholders and if possible also for the corresponding present value of these pensions. The last problem is difficult to solve. If the Poisson-model is justified something can be said.

Next we intend to treat the problems pointed out. In this no claim is made on important new discoveries in mathematical statistics, only the use of some techniques is stressed which, in my opinion, are less wellknown and which have some importance for the actuary, who wants to make some use of mathematical statistics in his work. The following may be partially seen as a complement of the theory as treated f. i. in the book of E. Zwinggi, already cited.

2. The stochastic process

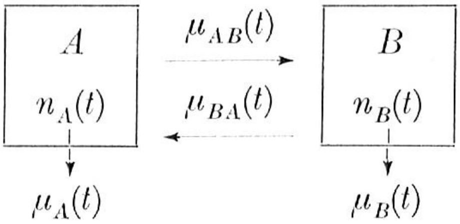
In this section we suppose the reader has some acquaintance with the theory of random processes as treated for instance in [1].

First we remark that in actuarial practice most processes are time-dependent, that means the intensities are functions of the time. We always have groups in which aging plays a role. This is a distinction compared with many physical examples.

a) The general random process

As general process we introduce here the following model. We suppose two groups A and B with transitions between the groups determined by the transition intensities $\mu_{AB}(t)$ and $\mu_{BA}(t)$. Further there is a decrease in each group determined by the intensities $\mu_A(t)$ and $\mu_B(t)$.

Figure 1 marks the situation



As concrete example we think of the development of a group insured in active insured and those who are ill, e.g. the insured in virtue of the dutch invalidity act or a similar insurance. Let be:

$n_A(t)$ the number of active insured,
$n_B(t)$ the number of insured who are ill,
$\mu_{AB}(t)$ the intensity of falling ill,
$\mu_{BA}(t)$ the revalidating intensity,
$\mu_A(t)$ the death rate of valid insured,
$\mu_B(t)$ the death rate of those who are ill.

We have two stochastic variables, $n_A(t)$ and $n_B(t)$. This process is very important for the actuary; apart from the example just given, there are many other applications. That we give it the name "general" means, that we suppose all four intensities to be different from zero, not very small and functions of the time. In the following sections we shall see that in some practical applications simplifications can be made. This is important, as the "general" process is rather difficult to deal with. As stochastic differential equation we obtain:

$$\begin{aligned} P'_t(n_A, n_B) &= -\left[(\mu_{AB}(t) + \mu_A(t))\, n_A(t) + (\mu_{BA}(t) + \mu_B(t))\, n_B(t)\right] P_t(n_A, n_B) \\ &+ \mu_{AB}(t)\,(n_A(t)+1) P_t(n_A+1, n_B-1) + \mu_{BA}(t)\,(n_B(t)+1) P_t(n_A-1, n_B+1) \\ &+ \mu_A(t)\,(n_A(t)+1) P_t(n_A+1, n_B) + \mu_B(n_B(t)+1)\, P_t(n_A, n_B+1)\,. \end{aligned}$$

In this the usual conditions are made that subsequent time intervals are independent and the probability of more than one change in $(t, t+h)$ is $0(h)$. If we define the generating function of $\{P_t(n_A, n_B)\}$ as

$$G_t(u,v) = \sum_{n_A, n_B} u^{n_A} v^{n_B} P_t(n_A, n_B)$$

we obtain the following partial differential equation for $G_t(u,v)$

$$\frac{\partial G_t(u,v)}{\partial t} = \left[\mu_{AB}(t)\,(v-u) + \mu_A(t)\,(1-u)\right] \frac{\partial G_t(u,v)}{\partial u} + \\ + \left[\mu_{BA}(t)\,(u-v) + \mu_B(t)\,(1-v)\right] \frac{\partial G_t(u,v)}{\partial v}.$$

The solution of this equation by quadratures is generally impossible. By imposing by-conditions on the intensities we may obtain approximative solutions which have practical meaning. We shall inquire this in the following two sub-sections. Here we only indicate the solution for the case the intensities are constants. Then we have to solve the subsidiary system.

$$dt = \frac{du}{\mu_{AB}(u-v)+\mu_A(u-1)} = \frac{dv}{\mu_{BA}(v-u)+\mu_B(v-1)}.$$

The equation

$$\frac{du}{\mu_{AB}(u-v)+\mu_A(u-1)} = \frac{dv}{\mu_{BA}(v-u)+\mu_B(v-1)}$$

is of the homogeneous type if we introduce the transformation $\bar{u} = u-1, \bar{v} = v-1$. It can easily be integrated. We can therefore follow the common procedure to solve a simultaneous system. If we only want the functions $E_t(n_A)$ and $E_t(n_B)$, a different approach is possible. By multiplying the equation for $P_t'(n_A, n_B)$ by $n_A(t)$ and $n_B(t)$ respectively, and summing for these variables, we find the equations

$$\begin{aligned} E_t'(n_A) &= -(\mu_{AB}(t)+\mu_A(t))\,E_t(n_A)+\mu_{BA}(t)\,E_t(n_B), \\ E_t'(n_B) &= \mu_{AB}(t)\,E_t(n_A)-(\mu_{BA}(t)+\mu_B(t))\,E_t(n_B). \end{aligned}$$

Supposing again the intensities are constants, we can easily solve this system by following the method of d'Alembert. This gives the solution

$$\begin{aligned} E_t(n_A) &= c_1 e^{r_1 t} + c_2 e^{r_2 t}, \\ E_t(n_B) &= \bar{c}_1 e^{r_1 t} + \bar{c}_2 e^{r_2 t}, \end{aligned}$$

in which r_1 and r_2 are the roots of

$$r^2 - (\mu_{AB}+\mu_A+\mu_{BA}+\mu_B)\,r + [(\mu_{AB}+\mu_A)(\mu_{BA}+\mu_B) - \mu_{AB}\mu_{BA}] = 0$$

and
$$\begin{aligned} \bar{c}_1 &= (r_1-\mu_{AB}-\mu_A)\,c_1, & n_A(0) &= c_1 + c_2, \\ \bar{c}_2 &= (r_2-\mu_{AB}-\mu_A)\,c_2, & n_B(0) &= \bar{c}_1 + \bar{c}_2. \end{aligned}$$

If the intensities are not constants, the method leads to a differential equation of Riccati, which is in general not soluble by quadratures. However, we may suppose that $E_t(n_A)$, $E_t(n_B)$ and the intensities are power series. Solution by series leads to recurrence relations for the coefficients. Suppose $E_t(n_A) = \sum_\nu {}_\nu n_A t^\nu$, $E_t(n_B) = \sum_\nu {}_\nu n_B t^\nu$ the intensities linear functions: $\mu_{AB}(t) = \mu_{AB} + \mu'_{AB}(t)$ etc.

As recurrence relations we obtain:

$$ {}_{\nu}n_A = \tfrac{1}{\nu}\left[\left\{\mu_B\,{}_{\nu-1}n_B - (\mu_{AB}+\mu_A)\,{}_{\nu-1}n_A\right\} + \left\{\mu'_B\,{}_{\nu-2}n_B - (\mu'_{AB}+\mu'_A)\,{}_{\nu-2}n_A\right\}\right], $$
$$ {}_{\nu}n_B = \tfrac{1}{\nu}\left[\left\{\mu_A\,{}_{\nu-1}n_A - (\mu_{BA}+\mu_B)\,{}_{\nu-1}n_B\right\} + \left\{\mu'_A\,{}_{\nu-2}n_A - (\mu'_{BA}+\mu'_B)\,{}_{\nu-2}n_B\right\}\right]. $$

Starting from $n_A(0) = {}_0n_A$ and $n_B(0) = {}_0n_B$ we can compute $E_t(n_A)$ and $E_t(n_B)$ some time ahead. The convergence of the series is only good for rather small values of t. In order to compute $E_t(n_A)$ and $E_t(n_B)$ for larger values of t we have to divide the interval $(0,t)$ into subintervals.

The solutions with constant intensities are important for the derivation of the intensities for smaller periods, f. i. a year, from observational data. From these intensities special probabilities can be calculated needed by the actuary for the determining of annuities etc. This derivation will be our object in section 2.

b) The Poisson-approximation

In subsection a) we introduced the example of a group falling apart into two groups, the group of active insured and the group of those who are ill. Now it depends entirely on the criterion of the act or policy if any one insured has to be considered as ill. If we take the Dutch Invalidity Act the figures show that the level of invalidating is about 0,001—0,01, whereas the rates of revalidating are high, about 0,1—0,5. Renewed invalidating after foregoing revalidating may be neglected if we timely restrict the extrapolation to not too large values. μ_A the death intensity of the valid insured is also very low, while μ_B the death intensity of the invalid is somewhat higher. The foregoing implies that $n_A(t)\,\mu_{AB}(t) = \mu(t)$ the total invalidating intensity may be considered as not stochastic. The numbers which invalidate in $(0,T)$ are given by a Poisson-distribution with parameter $\int\limits_0^T \mu(t)\,dt$. If we put $\mu_{BA}(t) + \mu_B(t) = \lambda(t)$ we obtain as approximative differential equation for $P_t(n_B)$

$$ \begin{aligned} P'_t(n_B) &= -\left[\mu(t) + \lambda(t)\,n_B(t)\right]P_t(n_B) + \\ &\quad + \mu(t)\,P_t(n_B-1) + \lambda(t)\,(n_B(t)+1)\,P_t(n_B+1)\,. \end{aligned} $$

The partial differential equation of the generating function of $\{P_t(n_B)\}$ is

$$ \frac{\partial G_t(s)}{\partial t} = (1-s)\left[-\mu(t)\,G_t(s) + \lambda(t)\,\frac{\partial G_t(s)}{\partial s}\right]. $$

The subsidiary system is

$$dt = \frac{ds}{(s-1)\,\lambda(t)} = \frac{dG_t(s)}{(s-1)\,\mu(t)\,G_t(s)}\,.$$

Solving this simultaneous system we find as solution

$$G_t(s) = e^{-(1-s)\int\limits_0^t e^{-\int\limits_\tau^t \lambda(\varrho)\,d\varrho}\mu_\tau\,d\tau}\left[1-(1-s)\,e^{-\int\limits_0^t \lambda(\tau)\,d\tau}\right]^{n_B(0)}.$$

The second factor is the generating function of the binomial distribution with parameters

$$p_t = e^{-\int\limits_0^t \lambda(\tau)\,d\tau} \qquad \text{and} \qquad N = n_B(0)\,.$$

It arises from the fact that at time 0 there are already $n_B(0)$ invalids. The first factor is the generating function of the Poisson-distribution with parameter

$$\int\limits_0^t e^{-\int\limits_\tau^t \lambda(\varrho)\,d\varrho}\,\mu(\tau)\,d\tau\,.$$

If we consider $\lambda(t)$ and $\mu(t)$ as constants we obtain the solution

$$G_t(s) = e^{-(1-s)\frac{\mu}{\lambda}(1-e^{-\lambda t})}\,[1+(s-1)\,e^{-\lambda t}]^{n_B(0)}\,.$$

$n_B(t)$ is the sum of two independently distributed variables i. e. the Poisson-variable with parameter $\frac{\mu}{\lambda}(1-e^{-\lambda t})$ and the binomial variate with probability $e^{-\lambda t}$. Expectation and variance of $n_B(t)$ are the sums of these values for the variables apart. Hence

$$E_t(n_B) = \frac{\mu}{\lambda}\,(1-e^{-\lambda t}) + n_B(0)\,e^{-\lambda t},$$

$$\mathrm{Var}_t(n_B) = \frac{\mu}{\lambda}\,(1-e^{-\lambda t}) + n_B(0)\,e^{-\lambda t}(1-e^{-\lambda t})\,.$$

Especially $\frac{\mu}{\lambda}(1-e^{-\lambda})$ can be interpreted as the probability of falling ill and remaining ill until the end of the year.

c) Other approximative random processes

In applications it frequently occurs that μ_A and μ_B are very small, whereas μ_{AB} and μ_{BA} are both large. As example we may take the transitions between the group of premium payers and the remaining group, insured in virtue of the Dutch Invalidity Act. Once accepted for the insurance one remains insured one's whole life. Surrender did occur formerly, but can be neglected nowadays. Transitions $A \to B$ are only for a small fraction caused by sickness; they are mostly due to the fact that the insured becomes independent or earns an income higher than the level above which premiums are to be paid. μ_A and μ_B are the death-rates of the premium payers and the remaining insured. From statistical investigations it turns out that μ_B is somewhat larger than μ_A, however, both are very small compared with μ_{AB} and μ_{BA}. If we have only the intention to compute the development over a rather short time interval, we may put as first approximation $\mu_A = \mu_B = 0$ and further assume that μ_{AB} and μ_{BA} are constants. Hence the number of the total group is a constant N, and we only have one random variable f. i. $n_B(t)$ to take into account. As stochastic differential equation for $P_t(n_B)$ we obtain

$$P_t'(n_B) = -[(N-n_B)\mu_{AB} + n_B\mu_{BA}]P_t(n_B) + \\ + (N-n_B+1)\mu_{AB}P_t(n_B-1) + (n_B+1)\mu_{BA}P_t(n_B+1).$$

Hence as partial differential equation for the generating function of $\{P_t(n_B\}$ we get

$$\frac{\partial G_t(s)}{\partial t} = \mu_{AB}N(s-1)\,G_t(s) + \{\mu_{AB}s(1-s) + \mu_{BA}(1-s)\}\frac{\partial G_t(s)}{\partial s},$$

$$G_0(s) = s^{n_B(0)}.$$

The solution is

$$G_t(s) = (1+(s-1)_1p_t)^{N-n_B(0)} \cdot (1+(s-1)_2p_t)^{n_B(0)}$$

in which

$${}_1p_t = \frac{\mu_{AB}(1-e^{-(\mu_{AB}+\mu_{BA})t})}{\mu_{AB}+\mu_{BA}}, \quad {}_2p_t = \frac{\mu_{AB}+\mu_{BA}e^{-(\mu_{AB}+\mu_{BA})t}}{\mu_{AB}+\mu_{BA}}.$$

The generating function of the binominal distribution with parameters p and n is $[1+(s-1)p]^n$. Hence $n_B(t)$ is the sum of two independent binomial variates with probabilities ${}_1p_t$ and ${}_2p_t$.

We have in conclusion.

$$E_t(n_B) = (N - n_B(0))\,{}_1p_t + n_B(0)\,{}_2p_t\,,$$
$$\mathrm{Var}_t(n_B) = (N - n_B(0))\,{}_1p_t(1 - {}_1p_t) + n_B(0)\,{}_2p_t(1 - {}_2p_t)$$

and

$$n_B(t) \frown N(E_t(n_B))_1 \sqrt{\mathrm{Var}_t(n_B)}\,,$$

($n_B(t)$ has approximately a normal distribution).

If μ_{AB} and μ_{BA} are not constants but functions of the time we obtain a similar result. We get the same partial differential equation for $G_t(s)$, except that μ_{AB} and μ_{BA} are now to be replaced by $\mu_{AB}(t)$ and $\mu_{BA}(t)$.

The subsidiary system is

$$dt = \frac{ds}{(s-1)\,(\mu_{AB}(t)\,s + \mu_{BA}(t))} = \frac{dG_t(s)}{\mu_{AB}(t)\,N\,(s-1)\,G_t(s)}\,.$$

The first equation of this system yields a differential equation of Riccati which in this case can be integrated by quadratures. We have

$$\frac{ds}{dt} = \mu_{AB}(t)\,s^2 + (\mu_{BA}(t) - \mu_{AB}(t))\,s - \mu_{BA}(t)\,.$$

A particular solution is $s = 1$. Hence the transformation $\bar{s} = s - 1$ reduces this equation to one of the Bernouilli type. If this equation is integrated and the solution is $s - 1 = f(c_1, t)$ in which c_1 is a constant of integration, he have to integrate

$$\mu_{AB}(t)\,N f(c_1, t)\,dt = \frac{dG_t(s)}{G_t(s)}\,.$$

We shall find two solutions

$$c_1 = f_1(s, t, G)\,,$$
$$c_2 = f_2(s, t, G)\,.$$

The solution of the partial differential equation is now given by $F(c_1, c_2) = 0$ and the boundary condition $G_0(s) = s^{n_B(0)}$. The expressions for c_1 and c_2 are rather complicated. Therefore we shall follow another way to determine the function $G_t(s)$. If we multiply the differential equation for $P_t(n_B)$ by n_B and sum over n_B, we get the equation

$$E_t'(n_B) = \mu_{AB}[N - E_t(n_B)] - \mu_{BA}\,E_t(n_B)\,.$$

The solution is

$$E_t(n_B) = (N - n_B(0)) \int_0^t \mu_{AB}(\tau)\, e^{-\int_\tau^t (\mu_{AB}(\varrho) + \mu_{BA}(\varrho))\, d\varrho}\, d\tau +$$

$$+ n_B(0) \left[e^{-\int_0^t (\mu_{AB}(\tau) + \mu_{BA}(\tau))\, d\tau} + \int_0^t \mu_{AB}(\tau)\, e^{-\int_\tau^t (\mu_{AB}(\varrho) + \mu_{BA}(\varrho))\, d\varrho}\, d\tau \right].$$

We put

$$_1p_t = \int_0^t \mu_{AB}(\tau)\, e^{-\int_\tau^t (\mu_{AB}(\varrho) + \mu_{BA}(\varrho))\, d\varrho}\, d\tau$$

and

$$_2p_t = e^{-\int_0^t (\mu_{AB}(\tau) + \mu_{BA}(\tau))\, d\tau} + \int_0^t \mu_{AB}(\tau)\, e^{-\int_\tau^t (\mu_{AB}(\varrho) + \mu_{BA}(\varrho))\, d\varrho}\, d\tau.$$

Now intuition tells us that again the solution is given by

$$[1 + (s-1)\, {_1p_t}]^{N - n_B(0)} \cdot [1 + (s-1)\, {_2p_t}]^{n_B(0)}.$$

That this supposition is not wrong can easily be verified by substituting this expression in the partial differential equation for $G_t(s)$. (Take e. g. first $n_B(0) = 0$ and apply the differentiation rule

$$\frac{d}{dx} \int_0^x f(x,y)\, dy = \int_0^x f_x(x,y)\, dy + f(x,x), \; f(x,y)$$

an arbitrary function). Hence again $n_B(t)$ has approximately a normal distribution with

$$E_t(n_B) = (N - n_B(0))\, {_1p_t} + n_B(0)\, {_2p_t},$$

$$\mathrm{Var}_t(n_B) = (N - n_B(0))\, {_1p_t}(1 - {_1p_t}) + n_B(0)\, {_2p_t}(1 - {_2p_t}).$$

So far we only considered the situation that the size of the total group of insured is a constant. In reality there is a slight decrease by dying etc. If we want to consider this decrease stochastic, the problem is rather difficult to solve. Now the rate of decrease is small compared with that of the transitions. Consequently the stochastic effect of this decrease is small. We may take this decrease into account by considering it to be not stochastic and by supposing $N = N(t)$ is a fixed slowly decreasing function of time. For instance $N(t)$ is a linear function $N(1 - \alpha t)$ in which α is a known constant.

A crude correction for $E_t(n_B)$ will be

$$\approx -\tfrac{1}{2}\alpha N \int_0^t \mu_{AB}(\tau)\, e^{-\int_\tau^t (\mu_{AB}(\varrho)+\mu_{BA}(\varrho))d\varrho}\, d\tau = -\tfrac{1}{2}\alpha N\, {}_1p_t .$$

Likewise we may obtain a correction for $\mathrm{Var}_t(n_B)$ by considering the differential equation for $E_t(n_B^2)$, and applying the formula

$$\mathrm{Var}_t(n_B) = E_t(n_B^2) - (E_t(n_B))^2,$$

Multiplying the equation for $P_t(n_B)$ by n_B^2 and summing over n_B gives

$$E_t'(n_B^2) = -2\,(\mu_{AB}(t)+\mu_{BA}(t))\, E_t(n_B^2) + (2\mu_{AB}(t)\, N(t) - \mu_{AB}(t) + \mu_{BA}(t)) \cdot$$
$$\cdot\, E_t(n_B) + N(t)\,\mu_{AB}(t)\,, \qquad E_0(n_B^2) = n_B^2(0)\,.$$

This is a linear equation which can easily be solved.

3. Dependent and independent probabilities

a) Transitions between groups

Over a small period, for instance a year, we may consider μ_{AB} and μ_{BA} as constants. Further we shall neglect the decrease of the total group of insured. Afterwards we'll make a correction. Suppose we have k observations

$$({}_jn_A(0)\,,\ {}_jn_B(0)\,,\ {}_jn_B(1))\,, \qquad j = 1, \ldots, k\,.$$

We saw that ${}_jn_B(1)$ is the sum of two independent binomial variates with probabilities

$${}_1p = \frac{\mu_{AB}\,(1-e^{-(\mu_{AB}+\mu_{BA})})}{\mu_{AB}+\mu_{BA}}\,, \qquad {}_2p = \frac{\mu_{AB}+\mu_{BA}\, e^{-(\mu_{AB}+\mu_{BA})}}{\mu_{AB}+\mu_{BA}}\,.$$

If ${}_1p$ and ${}_2p$ are rather small, we obtain approximately maximum likelihood estimates for ${}_1p$ and ${}_2p$ by determining ${}_1\hat{p}$ and ${}_2\hat{p}$ from

$$\sum_{j=1}^{k} \frac{[{}_jn_B(1) - \{{}_jn_A(0)\,{}_1p + {}_jn_B(0)\,{}_2p\}]^2}{{}_jn_B(1)} = \text{Minimum} \qquad [\text{see } 3\,b)].$$

However, a moment's reflection tells us that this method only gives valid results if the ratio's $\dfrac{{}_jn_B(0)}{{}_jn_A(0)}$ do not differ too less. If this last condition is fulfilled and the observations are not very small, the method may give reliable results. The estimates $\hat{\mu}_{AB}$ ans $\hat{\mu}_{BA}$ are now to be computed from

$$\begin{cases} \hat{\mu}_{AB} + \hat{\mu}_{BA} = -\log({}_2\hat{p} - {}_1\hat{p}), \\ \hat{\mu}_{AB} = \dfrac{-{}_1\hat{p}\log({}_2\hat{p} - {}_1\hat{p})}{1 - {}_2\hat{p} + {}_1\hat{p}}. \end{cases}$$

In practice, the condition, that the $\dfrac{{}_jn_B(0)}{{}_jn_A(0)}$ show large differences not owing to random effects, is often not fulfilled. In this situation the method cannot be applied. But it might happen that it is possible to split up the numbers ${}_jn_B(1)$ into two parts ${}_{j,A}n_B(1)$ and ${}_{j,B}n_B(1)$, ${}_{j,A}n_B(1)$ being the number of insured at time 1 in B and at time 0 in A, and ${}_{j,B}n_B(1)$ being the number of insured in the B group, who, one year earlier, were also in the B-group. Then we have

$${}_1p \approx \frac{{}_{j,A}n_B(1)}{{}_jn_A(0)}, \qquad {}_2p \approx \frac{{}_{j,B}n_B(1)}{{}_jn_B(0)}.$$

If we have k observations, we again obtain the estimates ${}_1\hat{p}$ and ${}_2\hat{p}$ from the minimum χ^2-method:

$$\sum_{j=1}^{k} \frac{[{}_{j,A}n_B(1) - {}_jn_A(0)\,{}_1p]^2}{{}_{j,A}n_B(1)} = \text{minimum},$$

$$\sum_{j=1}^{k} \frac{[{}_{j,B}n_B(1) - {}_jn_B(0)\,{}_2p]^2}{{}_{j,B}n_B(1)} = \text{minimum}.$$

Or if we apply the modified minimum χ_2-method, we obtain the equations [2]

$$\sum_{j=1}^{k} \frac{[{}_{j,A}n_B(1) - {}_jn_A(0)\,{}_1p]\,{}_jn_A(0)}{{}_jn_A(0)\,{}_1p} = 0,$$

$$\sum_{j=1}^{k} \frac{[{}_{j,B}n_B(1) - {}_jn_B(0)\,{}_2p]\,{}_jn_B(0)}{{}_jn_B(0)\,{}_2p} = 0.$$

This is equivalent with maximum likelihood estimation:

$${}_1\hat{p} = \frac{\sum_{j=1}^{k} {}_{j,A}n_B(1)}{\sum_{j=1}^{k} {}_jn_A(0)}, \qquad {}_2\hat{p} = \frac{\sum_{j=1}^{k} {}_{j,B}n_B(1)}{\sum_{j=1}^{k} {}_jn_B(0)}.$$

${}_1p$ and ${}_2p$ are commonly indicated with "dependent" probabilities. In connection with this notation $1-e^{-\mu_{AB}}$ and $1-e^{-\mu_{BA}}$ are called "independent" probabilities. They represent the "real" transition probabilities.

If the decrease of the total group cannot be neglected so f. i. $N(t)=N(1-\alpha t)$, we find that the correction for $E(n_B(1))$ is

$$-\alpha N \mu_{AB}\left(\frac{e^{-(\mu_{AB}+\mu_{BA})}+(\mu_{AB}+\mu_{BA})-1}{(\mu_{AB}+\mu_{BA})^2}\right)\sim-\tfrac{1}{2}\alpha N \mu_{AB}.$$

Neglecting random effects

$$\frac{{}_An_B(1)}{n_A(0)}\approx\frac{\mu_{AB}\,(1-e^{-(\mu_{AB}+\mu_{BA})})}{\mu_{AB}+\mu_{BA}}-\tfrac{1}{2}\alpha\mu_{AB},$$

$$\frac{{}_Bn_B(1)}{n_B(0)}\approx\frac{\mu_{AB}+\mu_{BA}\,e^{-(\mu_{AB}+\mu_{BA})}}{\mu_{AB}+\mu_{BA}}-\tfrac{1}{2}\alpha\mu_{AB}.$$

From these observed frequencies $\hat{\mu}_{AB}$ and $\hat{\mu}_{BA}$ can be computed. We haven again

$$\mu_{AB}+\hat{\mu}_{BA}\approx-\log\left(\frac{{}_Bn_B(1)}{n_B(0)}-\frac{{}_An_B(1)}{n_A(0)}\right).$$

Next $\hat{\mu}_{AB}$ and $\hat{\mu}_{BA}$ seperately can be derived from the expression for $\dfrac{{}_An_B(1)}{n_A(0)}$. In conclusion we may again determine the independent probabilities as $1-e^{-\hat{\mu}_{AB}}$ and $1-e^{-\hat{\mu}_{BA}}$.

b) Groups with several possibilities of leaving

As most simple example we consider a group which decreases in two ways. For instance group B, the group of the insured who are ill. Group B diminishes by revalidating and dying, determined by the intensities $\mu_{BA}(t)$ and $\mu_B(t)$. Both $\mu_{BA}(t)$ and $\mu_B(t)$ are rather large. We suppose that at time $t=0$, there are N insured who are ill. Further $n_1(t)$ is the number of those who revalidated in $(0,t)$ and $n_2(t)$ the number of deaths. As stochastic differential equations we have

$$\begin{aligned}\frac{\partial P_t(n_1,n_2)}{\partial t}=&-\{\mu_{BA}(t)+\mu_B(t)\}\{N-n_1(t)-n_2(t)\}P_t(n_1,n_2)\\&+\mu_{BA}(t)\{N-n_1(t)-n_2(t)+1\}P_t(n_1-1,n_2)\\&+\mu_B(t)\{N-n_1(t)-n_2(t)+1\}P_t(n_1,n_2-1).\end{aligned}$$

The solution is of course the trinomial distribution

$$P_t(n_1, n_2) = \frac{N!}{n_1(t)!\, n_2(t)!\, (N - n_1(t) - n_2(t))!} \cdot \left\{ \int_t^0 \mu_{BA}(\tau)\, e^{-\int_0^\tau \{\mu_{BA}(\varrho) + \mu_B(\varrho)\} d\varrho} d\tau \right\}^{n_1(t)}$$

$$x \left\{ \int_0^t \mu_B(\tau)\, e^{-\int_0^\tau \{\mu_{BA}(\varrho) + \mu_B(\varrho)\} d\varrho} d\tau \right\}^{n_2(t)} \cdot \left\{ 1 - e^{-\int_0^t \{\mu_{BA}(\tau) + \mu_B(\tau)\} d\tau} \right\}^{N - n_1(t) - n_2(t)} .$$

For the expectations we find

$$E_t(n_1) = N \int_0^t \mu_{BA}(\tau)\, e^{-\int_0^\tau \{\mu_{BA}(\varrho) + \mu_B(\varrho)\} d\varrho} d\tau ,$$

$$E_t(n_2) = N \int_0^t \mu_B(\tau)\, e^{-\int_0^\tau \{\mu_{BA}(\varrho) + \mu_B(\varrho)\} d\varrho} d\tau .$$

Both of them are functions of $\mu_{BA}(t)$ and $\mu_B(t)$.

$\frac{E_t(n_1)}{N}$ and $\frac{E_t(n_2)}{N}$ are therefore called dependent probabilities. $n_1(t)$ and $n_2(t)$ are not independently distributed. We could also have arrived at the expressions for $E_t(n_1)$ and $E_t(n_2)$ by multiplying the equation for $\frac{\partial P_t(n_1, n_2)}{\partial t}$ with $n_1(t)$ and $n_2(t)$ respectively and summing for these variables. We then find

$$E_t'(n_1) = \mu_{BA}(t) \{ N - E_t(n_1) - E_t(n_2) \} ,$$
$$E_t'(n_2) = \mu_B(t) \{ N - E_t(n_1) - E_t(n_2) \} .$$

Solving this system leads to the expressions above. If t is restricted to small values, wo may again consider $\mu_{BA}(t)$ and $\mu_B(t)$ as constants. In this case we simply get

$$q_1 = \frac{E_t(n_1)}{N} = \frac{\mu_{BA}}{\mu_{BA} + \mu_B} (1 - e^{-(\mu_{BA} + \mu_B)t}) ,$$

$$q_2 = \frac{E_t(n_2)}{N} = \frac{\mu_B}{\mu_{BA} + \mu_B} (1 - e^{-(\mu_{BA} + \mu_B)t}) .$$

The independent probabilities q_1' and q_2' are obtained from the artificial assumption that the stochastic differential equation for $P_t(n_1, n_2)$

is given by

$$\frac{\partial P_t(n_1, n_2)}{\partial t} = [\mu_{BA}(t)\,(N-n_1(t)) + \mu_B(t)\,(N-n_2(t))]\,P_t(n_1, n_2)$$
$$+ \mu_{BA}(t)\,(N-n_1(t)+1)\,P_t(n_1-1, n_2)$$
$$+ \mu_B(t)\,(N-n_2(t)+1)\,P_t(n_1, n_2-1)\,.$$

That means we suppose that the probability of a change

$$n_1(t) \to n_1(t)+1 \qquad \text{in } (t, t+h)$$

is given by $\mu_{BA}(N-n_1(t))\,h + 0(h)$ and not by $\mu_{BA}(N-n_1(t)-n_2(t))\,h + 0(h)$ and equally the probability of a change $n_2(t) \to n_2(t)+1$ is $\mu_B(N-n_2(t))\cdot$ $\cdot\, h + 0(h)$ and not $\mu_B(N-n_1(t)-n_2(t))\,h + 0(h)$.

The solution is

$$P_t(n_1, n_2) = \frac{N!}{n_1(t)\,(N-n_1(t))!}\left(1-e^{-\int_0^t \mu_{BA}(\tau)\,d\tau}\right)^{n_1(t)}\left(e^{-\int_0^t \mu_{BA}(\tau)\,d\tau}\right)^{N-n_1(t)}$$
$$\cdot\,\frac{N!}{n_2(t)\,(N-n_2(t))!}\left(1-e^{-\int_0^t \mu_B(\tau)\,d\tau}\right)^{n2(t)}\left(e^{-\int_0^t \mu_B(\tau)\,d\tau}\right)^{N-n_2(t)}.$$

$n_1(t)$ and $n_2(t)$ are independent binomial variates with means

$$E_t(n_1) = N\left(1-e^{-\int_0^t \mu_{BA}(\tau)\,d\tau}\right),$$
$$E_t(n_2) = N\left(1-e^{-\int_0^t \mu_B(\tau)\,d\tau}\right).$$

Hence we have as independent probabilities

$$q_1' = 1-e^{-\int_0^t \mu_{BA}(\tau)\,d\tau},$$
$$q_2' = 1-e^{-\int_0^t \mu_B(\tau)\,d\tau}.$$

In some applications it may occur that μ_{AB} and μ_B are both small f. i. $< 0{,}01$. In this case we obtain as approximative equation for $P_t(n_1, n_2)$

$$\frac{\partial P_t(n_1, n_2)}{\partial t} = -(\mu_{BA}(t) + \mu_B(t))\,N P_t(n_1, n_2) +$$
$$+ \mu_{BA}(t)\,N P_t(n_1-1, n_2) + \mu_B(t)\,N P_t(n_1, n_2-1)\,.$$

The solution of the system is

$$P_t(n_1, n_2) = \frac{e^{-N\int_0^t \{\mu_{BA}(\tau)+\mu_B(\tau)\}\,d\tau} \left\{N\int_0^t \mu_{BA}(\tau)\,d\tau\right\}^{n_1(t)} \cdot \left\{N\int_0^t \mu_B(\tau)\,d\tau\right\}^{n_2(t)}}{n_1(t)!\; n_2(t)!}$$

$$= \frac{e^{-N\int_0^t \mu_{BA}(\tau)\,d\tau} \left\{N\int_0^t \mu_{BA}(\tau)\,d\tau\right\}^{n_1(t)}}{n_1(t)!} \cdot \frac{e^{-N\int_0^t \mu_B(\tau)\,d\tau} \left\{N\int_0^t \mu_B(\tau)\,d\tau\right\}^{n_2(t)}}{n_2(t)!}.$$

From the right side we see that $n_1(t)$ and $n_2(t)$ are independently distributed as Poisson-variables with expectations

$$N\int_0^t \mu_{BA}(\tau)\,d\tau \quad \text{and} \quad N\int_0^t \mu_B(\tau)\,d\tau.$$

From the foregoing it is clear that the problem of dependent and independant probabilities arises from the fact that sometimes $n_1(t)$ and $n_2(t)$ cannot be considered as Poisson-variables. In this case $n_1(t)$ and $n_2(t)$ have a covariance different from zero and we have to learn the values $\mu_{BA}(t)$ and $\mu_B(t)$ separately. In actuarial statistics it is not customary to publish the intensities and therefore q_1' and q_2' are computed. Another remark concerns the computation of q_1, q_1', q_2 and q_2'. The custom is to compute the independent probabilities q_1', q_2' from the observations and afterwards the dependent probabilities as

$$q_1 \approx q_1'(1 - \tfrac{1}{2}q_2'),$$
$$q_2 \approx q_2'(1 - \tfrac{1}{2}q_1').$$

q_1' and q_2' are computed after the well-known formula's:

$$q_1' \approx \frac{n_1}{\frac{1}{2}(B+E+n_1)} \quad \text{and} \quad q_2' \approx \frac{n_2}{\frac{1}{2}(B+E+n_2)},$$

(B and E are the numbers of insured in B at the beginning and the end of the years of observation; active insured who invalidate in the course of the year are treated separately). This method is rather inaccurate and cumbersome. Therefore we prefer the following calculation.

We have

$$\sigma = q_1 + q_2 = 1 - e^{-\int_0^t \{\mu_{BA}(\tau) + \mu_B(\tau)\} d\tau} \approx \frac{n_1 + n_2}{\frac{1}{2}(B + E + n_1 + n_2)}$$

and

$$q_1 : q_2 = E_1(n_1) : E_1(n_2) \approx n_1 : n_2 .$$

Hence the formula's for q_1 and q_2 are:

$$q_1 \approx \frac{n_1}{\frac{1}{2}(B + E + n_1 + n_2)}, \qquad q_2 \approx \frac{n_2}{\frac{1}{2}(B + E + n_1 + n_2)} .$$

If $\mu_{BA}(t)$ and $\mu_B(t)$ are constants, we compute $\mu_{BA} + \mu_B$ from $q_1 + q_2 = 1 - e^{-(\mu_{BA} + \mu_B)}$ with the aid of a table of the e-function. After that μ_{BA} and μ_B are computed from

$$\mu_{BA} \approx \frac{n_1}{n_1 + n_2} (\mu_{BA} + \mu_B)$$

and

$$\mu_B \approx \frac{n_2}{n_1 + n_2} (\mu_{BA} + \mu_B) .$$

In the end we determine the independent probabilities q_1' and q_2' as $q_1' = 1 - e^{-\mu_{BA}}$ and $q_2' = 1 - e^{-\mu_B}$.

4. Estimates and tests

a) The comparison of risks, tests based on the Poisson-model

In many actuarial applications the Poisson-process can be used as an approximative solution. In connection with this it is important to have tests for Poisson-variables. For instance we may want to compare accident risks of certain enterprises, the risks of subsequent years etc. The numbers of accidents may be $d_1, \ldots, d_k$. We now suppose that the d_i are approximately independent Poisson-variables with means λ_i. If the number of man-years of the i^{th} enterprise is w_i and this value indicates the size of the enterprise, we may say that there are no differences with regard to accident risk if $\lambda_1 : \lambda_2 : \lambda_k = w_1 : w_2 : w_k$. To investigate this hypothesis of equal risk, we shall give here two tests. We are not going to give complete proofs and information, these can be obtained from [3]. It is easy to prove that the conditional simultaneous distribution of the d_i, under the condition that $d = \sum_{i=1}^{k} d_i$ is fixed, is the multinomial distribution

$$P[d_1, \ldots, d_k | d] = \frac{d!}{\prod_{i=1}^{k} d_i!} \prod_{i=1}^{k} \tilde{w}_i^{d_i}$$

in which

$$\tilde{w}_i = \frac{w_i}{\sum_{i=1}^{k} w_i}, \qquad i = 1, \ldots, k.$$

According to this result, we see immediately that if the hypothesis of equal risks is true, $T_1 = \sum_{i=1}^{k} \frac{(d_i - d\tilde{w}_i)^2}{d\tilde{w}_i}$ has approximately a $\chi^2_{(k-1)}$-distribution. If we only want to test against a specific trend, f. i. the d_i showing a definite increasing risk: $\frac{\lambda_1}{w_1} < \frac{\lambda_2}{w_2} < \ldots < \frac{\lambda_k}{w_k}$ there is still another suitable statistic.

We may use $\sum_{i=1}^{k} id_i$ as criterion. Again if $d = \sum_{i=1}^{k} d_i$ is fixed, this quantity has a normal distribution with expectation $d \sum_{i=1}^{k} i\tilde{w}_i$ and variance

$$d\left[\sum_{i=1}^{k} i^2 \tilde{w}_i - \left(\sum_{i=1}^{k} i\tilde{w}_i\right)^2\right].$$

Hence

$$T_2 = \frac{\sum_{i=1}^{k} i(d_i - \tilde{\omega}_i)}{\sqrt{d\left[\sum_{i=1}^{k} i^2 \tilde{w}_i - \left(\sum_{i=1}^{k} i\tilde{w}_i\right)^2\right]}} \sim N(0,1).$$

With this statistic T_2 it is possible to test one-sidedly against an increasing or decreasing trend.

Both tests are essentially conditional tests, $d = \sum_{i=1}^{k} d_i$ is fixed. This condition is easily removed. We may consider d as stochastic, the distributions of T_1 and T_2 remain the same. If T_1^* is T_1 in which d is stochastic and ${}_\alpha\chi^2$ is the value with $P[\chi^2 > {}_\alpha\chi^2] = \alpha$ we have

$$\begin{aligned} P[T_1^* > {}_\alpha\chi^2] &= \sum_d P[T_1 > {}_\alpha\chi^2 | d]\, P(d) \\ &\approx P[\chi^2 > {}_\alpha\chi^2] \sum_d P(d) = P[\chi^2 > {}_\alpha\chi^2]. \end{aligned}$$

If the values d_i are very small, we may make use of exact tests. Further there are some methods of combining independent tests etc. We shall not enter into these matters; many details and numerical solutions can be found in [3]. We have given these tests here only because tests based on the Poisson-model are not usually dealt with in text-books or if so, very inadequately; however, for the actuary who wishes to make use of mathematical statistics, they are very important.

b) Regression and least squares in the case for Poisson-variables

This too is certainly a question of interest for the actuary. We shall only indicate the problem. Suppose the d_i are independent Poisson-variables with means λ_i. Further the λ_i follow the functional relationship

$$\lambda_i = w_i f(t_i, \alpha_1, \ldots, \alpha_h) = w_i f_i, \qquad i = 1, \ldots, k.$$

As system of equations to determine the maximum likelihood estimates $\hat{\alpha}_1, \ldots, \hat{\alpha}_h$, we obtain

$$P(d_1, \ldots, d_k) = e^{-\sum_{i=1}^{k} w_i f_i} \prod_{i=1}^{k} \frac{(w_i f_i)^{d_i}}{d_i!},$$

$$\frac{\partial \log P}{\partial \alpha_j} = -\sum_{i=1}^{k} w_i \frac{\partial f_i}{\partial \alpha_j} + \sum_{i=1}^{k} \frac{d_i w_i \frac{\partial f_i}{\partial \alpha_j}}{w_i f_i}$$

$$= \sum_{i=1}^{k} (d_i - w_i f_i) \frac{\partial \log f_i}{\partial \alpha_i} = 0, \qquad j = 1, \ldots, h.$$

These equations are identical with those of the modified minimum χ^2 method [2].

The system cannot be solved by elementary methods. An approximative solution will be obtained by solving the system

$$\sum_{i=1}^{k} \frac{(d_i - w_i f_i)}{d_i} w_i \frac{\partial f_i}{\partial \alpha_j} = 0, \qquad j = 1, \ldots, h.$$

If $f(t_i, \alpha_1, \ldots, \alpha_h)$ is a polynomial, we have a simultaneous system of linear equations. We obtain a similar result by determining $\hat{a}_1, \ldots, \hat{a}_h$ from

$$\sum_{i=1}^{k} \frac{(d_i - w_i f_i)^2}{d_i} = \text{minimum}.$$

This is also often called the minimum χ^2-method.

From the foregoing follows that, in this particular example of Poisson-variables, the minimum χ^2-method approximatively leads to maximum likelihood estimates for $\alpha_1, \ldots, \alpha_h$. The actuary mostly considers variables which nearly follow Poisson-distributions, therefore the minimum χ^2-method is to be preferred to the classical least squares method

$$\sum_{i=1}^{k} (d_i - w_i f_i)^2 = \text{minimum}.$$

We stress this method here because a treatment starting with the Poisson-model cannot be found in the commonly used text-books.

5. Extrapolation

a) Interval estimates of future numbers

In the foregoing stochastic processes are considered mainly in connection with the computation of certain intensities and probabilities. The problem of prediction was largely neglected. In this section at last we are going to make some remarks on this subject. In pratice, extrapolation with the intention to obtain interval estimates has only sense for some years ahead, as the intensities determining the process cannot be predicted with accuracy. The influence of errors in the intensities is much larger than the effect of random effects. This is especially true in case the numbers of insured are large. Therefore we shall restrict the extrapolation to only a few years. In this situation we may often use, in the example of insured who are seriously ill, the Poisson-approximation of 2b) with advantage. If we consider in addition the intensities as constants, we have

$$E_t(n_B) = \frac{\mu}{\lambda}(1 - e^{-\lambda t}) + n_B(0)\, e^{-\lambda t}.$$

$$\mathrm{Var}_t(n_B) = \frac{\mu}{\lambda}(1 - e^{-\lambda t}) + n_B(0)\, e^{-\lambda t}(1 - e^{-\lambda t}).$$

If the supposition that μ is a constant is too crude, we may put $\mu(t) = \mu(1 + \beta t)$. It is easy to prove that the correction for $E_t(n_B)$ amounts to

$$\Delta E_t(n_B) = \frac{\beta\mu}{\lambda^2}(e^{-\lambda t} + \lambda t - 1).$$

Because we have here a Poisson-distribution the correction for $\mathrm{Var}_t(n_B)$ is also $\Delta E_t(n_B)$.

If we do not only consider serious illness, but also not serious illness and small accidents, the Poisson approximation is too crude. In this case μ_{AB} and μ_{BA} are large while μ_A and μ_B are both small.

Now we may estimate $E_t(n_A)$ and $E_t(n_B)$ some years ahead with the formula's

$$E_t(n_A) = c_1 e^{r_1 t} + c_2 e^{r_2 t},$$
$$E_t(n_B) = \bar{c}_1 e^{r_1 t} + \bar{c}_2 e^{r_2 t}, \qquad [\text{see 3a})],$$

or, if we do not consider the intensities as constants, by the power series solution. The variance of $n_B(t)$ can be crudely estimated with the formula

$$\mathrm{Var}_t(n_B) \approx n_A(0)\, {}_1p_t(1-{}_1p_t) + n_B(0)\, {}_2p_t(1-{}_2p_t)$$

in which

$${}_1p_t = \frac{\mu_{AB}(1-e^{-(\mu_{AB}+\mu_{BA})t})}{\mu_{AB}+\mu_{BA}}, \quad {}_2p_t = \frac{\mu_{AB}+\mu_{BA}\, e^{-(\mu_{AB}+\mu_{BA})t}}{\mu_{AB}+\mu_{BA}}.$$

As already remarked, this has only sense for values of t restricted to a small period. If the intensities are strongly age-dependent this last remark has to be stressed.

b) Interval estimates for the present value in the case of the Poisson-model

In subsection a) we considered the problem of interval estimates for $n_B(t)$. There is also the problem of interval estimates for the present value of the pensions etc. to be paid by the insurer to these $n_B(t)$ insured. This problem is in general a difficult one. If $n_B(t)$ may be seen approximately as a Poisson-variable a solution can be given. We shall indicate the distribution of the present value of payments falling due in the period $(0,T)$. For simplicity we suppose $n_B(0) = 0$, and all payments equal to the money-unity. The number of pension-holders who fell ill at time τ, $n_{(\tau)+t-\tau}$, has a Poisson-distribution with parameter $M_{\tau,t} = \mu_\tau e^{-\int_0^{t-\tau} \lambda_{(\tau)+\varrho} d\varrho}$, μ_τ the total intensity to fall ill at τ and $\lambda_{(\tau)+\varrho}$ the intensity of leaving the group of insured who are ill at time $\tau+\varrho$ by revalidating etc. The present value of the pensions corresponding with n_τ is

$$w_{\tau,T} = \int_0^{T-\tau} e^{-(\tau+\varrho)\delta} n_{(\tau)+\varrho}\, d\varrho,$$

(δ intensity of interest).

$w_{\tau,T}$ is a random variable. We find

$$E(w_{\tau,T}) = \int_0^{T-\tau} e^{-(\tau+\varrho)\delta} M_{\tau,\varrho}\, d\varrho\,.$$

Var $(w_{\tau,T})$ is somewhat difficult to determine. Let $\tilde{w}_{\tau,T}$ mean the present value of one pension and $f(\tilde{w}_{\tau,T})$ the corresponding distribution. We shall now follow a heuristic way. We suppose that $\tilde{w}_{\tau,T}$ and $w_{\tau,T}$ can only assume positive integral values. The generating function of $w_{\tau,T}$ is then given by $G_{\tau,T}(s) = e^{-\mu_\tau + \mu_\tau g_{\tau,T}(s)}$ in which $g_{\tau,T}(s)$ is the generating function of the distribution $f(\tilde{w}_{\tau,T})$. For Var $(w_{\tau,T})$ we now have

$$\mathrm{Var}\,(w_{\tau,T}) = G''_{\tau,T}(1) + G'_{\tau,T}(1) - (G'_{\tau,T}(1))^2 = \mu_\tau\, E\,(\tilde{w}^2_{\tau,T})\,.$$

There remains to compute $E\,(\tilde{w}^2_{\tau,T})$

We have

$$E\,(\tilde{w}^2_{\tau,T}) = \int_\tau^{T-\tau} \left(\int_0^{t-\tau} e^{-(\tau+\varrho)\delta}\, d\varrho \right)^2 e^{-\int_0^{t-\tau} \mu_{(\tau)+\varrho} d\varrho}\, \mu_{(\tau)+t-\tau}\, dt + \left(\int_0^{T-\tau} e^{-(\tau+\varrho)\delta}\, d\varrho \right)^2 e^{-\int_0^{T-\tau} \mu_{(\tau)+\varrho} d\varrho}\,.$$

(The second term on the rigt side corresponds with those insured who are still ill at time T.)

The n_τ are approximatively independent Poisson-variables. Hence, if w_T is the total present value, so $w_T = \int_0^T w_{\tau,T}\, d\tau$ we have

$$E\,(w_T) = \int_0^T E\,(w_{\tau,T})\, d\tau = \int_0^T \mu_\tau\, E\,(\tilde{w}_{\tau,T})\, d\tau$$

and

$$\mathrm{Var}\,(w_T) = \int_0^T \mathrm{Var}\,(w_{\tau,T})\, d\tau = \int_0^T \mu_\tau\, E\,(\tilde{w}^2_{\tau,T})\, d\tau$$

in which $E(\tilde{w}_{\tau,T})$ and $E(\tilde{w}^2_{\tau,T})$ have the values indicated above. In pratice the $w_{\tau,T}$ are uniformly bounded; the central limit theorem may be applied.

Hence
$$\frac{w_T - E\,(w_T)}{\sqrt{\mathrm{Var}\,(w_T)}} \sim N(0,1)\,.$$

Numerical integration yields $E\,(w_T)$ and Var (w_T), the $N\,(0,1)$ table at last the interval estimation.

We have to remark here that the formula's for $E(w_T)$ and $\operatorname{Var}(w_T)$ can be considered as a generalisation of the theorem of Campbell. As know this theorem describes random fluctuation in an electron stream. Suppose that the number of electrons arriving at the anode in $(0,t)$ follows the Poisson-distribution with parameter λt, whereas the effect of an electron arriving at t_ν at the anode is $f(t_\nu)$, and $F = \sum_\nu f(t_\nu)$ the total random effect. The theorem now tells us that

$$E(F) = \lambda \int_0^\infty f(t)\, dt, \qquad \operatorname{Var}(F) = \lambda \int_0^\infty f^2(t)\, dt, \quad [4].$$

The analogy between this physical example and our actuarial one is evident.

Literature

[1] *W.Feller*, An Introduction to Probability Theory and its Applications, Volume One, New York, 1950, chapter 17.

[2] *H.Cramer*, Mathematical Methods of Statistics, Princeton 1946, pp. 424–427. *E.Zwinggi*, Versicherungsmathematik, Basel, 1945, Bl. 174, 175.

[3] *J.van Klinken*, De Toepassing van het Poisson-proces in de Actuariële Statistiek, Academisch Proefschrift, Amsterdam, Bl. 25–56.

[4] Selected Papers on Noise and Stochastic Process, Dover Publications, New York, 1945, pp. 146–150.

[5] *W.Saxer*, Versicherungsmathematik, Springer-Verlag, Berlin. Erster Teil, 1955, S. 6–30, 225–229, Zweiter Teil, 1958, S. 53–95.

Zusammenfassung

Der Autor betont die Wichtigkeit des Gebrauchs der Theorie stochastischer Prozesse für die Umschreibung gewisser grundlegender Begriffe in der Versicherungsmathematik. Dies gilt insbesondere hinsichtlich der in der mathematischen Theorie der Sozialversicherung wichtigen Begriffe der sogenannten abhängigen und unabhängigen Wahrscheinlichkeit. Gesondert werden einige Fragen der Intervallschätzung für die Zahl künftiger Schadenereignisse behandelt, sowie der Regression und der Risikovergleichung. Besondere Aufmerksamkeit wird sodann einigen genäherten Zufallsprozessen gewidmet, so vor allem dem Poisson-Prozess, welcher die statistischen Berechnungen zu vereinfachen gestattet.

Résumé

L'auteur souligne l'importance de l'usage de la théorie des processus stochastiques pour quelques conceptions de la mathématique actuarielle. Il le démontre spécialement à l'égard du sujet des probabilités dépendantes et indépendantes, très importantes dans la thorie mathématique de l'assurance sociale. En outre, sont traités quelques aspects de l'estimation d'intervalle des événement futurs, ainsi que de la régression et de la comparaison de risques. Une attention toute spéciale est attribuée à quelques processus stochastiques approximatifs, avant tout au modèle de Poisson, puisqu'il rend possible une simplification des calculs statistiques.

Riassunto

Viene sottolineata l'importanza dell'uso della teoria dei processi stocastici per alcuni concetti fondamentali nella matematica attuariale. Questo vale specialmente per i soggetti delle probabilità dipendenti e indipendenti, aventi parte importante nella teoria matematica dell'assicurazione sociale. A parte sono trattate alcune questioni sulla stima degli intervalli per il numero dei danni futuri, come pure sulla regressione e sul confronto dei rischi. Particolare attenzione è stata data ad alcuni processi stocastici e, in modo approfondito, al modello di Poisson, il quale rende possibile la semplificazione dei calcoli statistici.