m3f22chVIII.tex

Chapter VIII. INSURANCE MATHEMATICS

1. Insurance Background

The idea of insurance is simple: it is the *spreading*, or *pooling*, of risk. The relevant theory is that of *collective risk*.

History.

Insurance can be traced back to antiquity (Greek and Roman times). Like much else, it disappeared, to be re-developed in Renaissance Italy (Genoa, 14th C.). It received a great impetus in the UK from the Great Fire of London in 1666; fire insurance had started there by 1681. Property insurance had begun in London by 1710, and in Philadelphia (Benjamin Franklin) in 1752.

Shipping insurance grew in London around Edward Lloyd's coffee house in the 1680s. He died in 1710; Lloyd's of London had developed by 1774.

John Graunt (1620-74) published his *Bills of Mortality* in 1662 (breaking down London deaths by cause, age etc.). This was followed by the first life table (Edmund Halley, 1693). Mutual life insurance had begun by 1762. One of the earliest such companies is Scottish Widows (1815) (founded to look after the widows of Presbyterian ministers who died in office, and had to leave the manse – the minister's house).

At a national level, national insurance began in Germany with Bismarck in the 1880s. It developed here with e.g. Lloyd George (pre-WWI), Beveridge and the Beveridge Report (1942), and the founding of the Welfare State post-WWII.

Limited liability.

Lloyd's of London pre-dates limited liability (which developed in the mid-19th C.). The Lloyd's participants, or *names*, had unlimited liability, and were liable for the full extent of losses, irrespective of their investment or their assets. This changed, following the Lloyd's scandal of the 1990s.

Insurance is now done (and most was before the Lloyd's scandal) by limited liability companies. So for these, the possibility of *ruin* is crucial. Not only would this wipe out the company, its assets and expertise, the jobs of its employees etc., but it would leave policy-holders without cover.

Reinsurance.

Because a run of large claims could bankrupt an insurance company, companies seek to lay off large risks – to reinsure – insure themselves – with larger, specialist reinsurance companies.

The question arises as to where reinsurane companies re-reinsure themselves ... This raises the modern form of Juvenal's question (*Satires*, c. 80 AD): *Quis custodiet ipsos custodes* — Who guards the guards? Who polices the police? Reinsurers reinsure insurers, but — who reinsures the reinsurers? — etc.

Regulation.

It is in the interest of some industries to agree to cover each other's liabilities in the event of a bankruptcy. For instance, this happens with *travel firms*. If a travel firm goes bust, leaving large numbers of people stranded abroad, or unable to travel on a foreign holiday booked and paid for, this would destroy public confidence in the whole industry – *unless* other firms, by prior agreement, step in to cover. This is what happens, and works well.

As motor insurance is compulsory by law, motor insurance companies are regulated by the state, and again, this provides a degree of protection in case of bankruptcy.

The actuarial profession.

People involved in the insurance industry have been known as *actuaries* from the early days of insurance. Companies offering insurance employ actuaries, and these need to be qualified. Actuaries become qualified by passing exams set by the *Institute of Actuaries*. London is an important centre for the actuarial/insurance industry, and so is Edinburgh. The mathematics involved is interesting, and useful. Those taking this course would be well advised to consider an actuarial career as one of their career possibilities.

Life v. non-life.

The usual way the modern insurance industry splits is between life and non-life. Life insurance is payable on death, and/or as an annuity ceasing on death. Life insurance is often combined with a mortgage (so that the mortgage is paid if one dies before it expires). Naturally, assessing premiums here depends on a detailed knowledge of mortality rates over ages, etc. The relevant mathematics is largely *Survival Analysis* – hazard rates, etc. Much use is made here nowadays of *martingale methods*. Non-life splits again into categories: motor; house; (house) contents (these are the only three kinds of insurance ordinary people take out); (personal) accident (the next commonest); travel; commercial property; industrial; ... There are even catastrophe insurance, weather insurance etc. nowadays.

2. The Poisson process and compound Poisson process (continued).

Time-dependent rates.

The parameter λ is called the *rate* or *intensity* of the Poisson process. Think of it as the rate at which accidents happen (or telephone calls arrive at an exchange), or the intensity of a bombardment, etc. The work of Ch. I extends to include time-dependent intensities. We say that $\{N(s), s \geq 0\}$ is a *Poisson process* with *rate* $\lambda(r)$ if
(i) $N(0) = 0$,
(ii) for $s < t$, $N(t) - N(s)$ is Poisson with mean $\int_s^t \lambda(r) dr$, and
(iii) $N(t)$ has independent increments.

Limit Theory.

For independent, identically distributed (iid for short) random variables $X_1, X_2, \cdots$, the *sample mean* (a *statistic*: a function of the data – random, as the data is, but known, after sampling, when you have the data) is

$$\overline{X} := \frac{1}{n} \sum_1^n X_k.$$

The *mean*, or *population mean*, $E[X]$ is defined as in Measure Theory. One would expect that $\overline{X}$ would tend to $E[X]$ as the sample size n increases. This is exactly right. By Kolmogorov's Strong Law of Large Numbers of 1933 (SLLN, or just LLN for short), convergence takes place *with probability one* (*almost surely*, or a.s. for short):

$$\overline{X} \to E[X] \qquad (n \to \infty) \qquad a.s.$$

For renewal theory (in particular, for the Poisson process), this gives another LLN.

Theorem (LLN for Renewal Theory; Doob, 1948). For X_i (positive) iid with mean μ, the renewal process $N = (N(t))$ satisfies

$$\frac{N(t)}{t} \to \frac{1}{\mu} \qquad (t \to \infty) \qquad a.s.$$

Proof. By definition of $N(t)$ and $S_n := \sum_1^n X_k$,

$$S_{N(t)} \leq t < S_{N(t)+1}.$$

So as soon as $N(t) > 0$,

$$\frac{S_{N(t)}}{N(t)} \leq \frac{t}{N(t)} < \frac{S_{N(t)+1}}{N(t)+1} . \frac{N(t)+1}{N(t)} .$$

As $t \to \infty$, $N(t) \to \infty$ a.s. So by the LLN the left tends to μ a.s.; so does the first term on the right, while the second term on the right tends to 1. This gives

$$t/N(t) \to \mu \qquad (t \to \infty) \qquad a.s.$$

The result follows by inverting this. //

The Conditional Mean Formula

Theorem (Conditional Mean Formula. For $\mathcal{B}$ any σ-field,

$$E[E[X|\mathcal{B}]] = E[X].$$

Proof. Take $\mathcal{C}$ the trivial σ-field $\{\emptyset, \Omega\}$. This contains no information, so an expectation conditioning on it is the same as an unconditional expectation. The first form of the tower property now gives

$$E[E[X|\mathcal{B}] \; |\{\emptyset, \Omega\}] = E[X|\{\emptyset, \Omega\}] = E[X]. \qquad //$$

The Conditional Variance Formula

Theorem (Conditional Variance Formula).

$$var(Y) = E[var(Y|X)] + var(E[Y|X]).$$

Proof. Recall $varX := E[(X - EX)^2]$. Expanding the square,

$$varX = E[X^2 - 2X.(EX) + (EX)^2] = E(X^2) - 2(EX)(EX) + (EX)^2 = E(X^2) - (EX)^2.$$

Conditional variances can be defined in the same way. Recall ('taking out what is known') that $E(Y|X)$ is constant when X is known ($= x$, say), so can be taken outside an expectation over X, E_X say. Then

$$var(Y|X) := E(Y^2|X) - [E(Y|X)]^2.$$

Take expectations of both sides over X:

$$E_X var(Y|X) = E_X[E(Y^2|X)] - E_X[E(Y|X)]^2.$$

Now $E_X[E(Y^2|X)] = E(Y^2)$, by the Conditional Mean Formula, so the right is, adding and subtracting $(EY)^2$,

$$\{E(Y^2) - (EY)^2\} - \{E_X[E(Y|X)]^2 - (EY)^2\}.$$

The first term is $varY$, by above. Since $E(Y|X)$ has E_X-mean EY, the second term is $var_X E(Y|X)$, the variance (over X) of the random variable $E[Y|X]$ (random because X is). Combining, the result follows. //

Interpretation.
$varY$ = total variability in Y,
$E_X var(Y|X)$ = variability in Y not accounted for by knowledge of X,
$var_X E(Y|X)$ = variability in Y accounted for by knowledge of X.
In words:
variance = mean of conditional variance + variance of conditional mean, with these interpretations. This is extremely useful in Statistics, in breaking down uncertainty, or variability, into its contributing components. There is a whole area of Statistics devoted to such Components of Variance.

Compound Poisson Processes

We now associate i.i.d. random variables X_i with each arrival and consider

$$S(t) = X_1 + \ldots + X_{N(t)}, \qquad S(t) = 0 \text{ if } N(t) = 0.$$

Thus $S(t)$ is a *random sum* – a sum of a random number of random variables.

A typical application in the insurance context is a Poisson model of claim arrivals with random claim sizes. The claims arrive at the epochs of a Poisson process with rate λ. The claims are independent (different motor accidents are independent; so are different house-insurance claims for fire damage, burglary etc.). Then the claim-total mean is the claim-number mean times the claim-amount mean. This is a special case of *Wald's identity* (below).

Theorem. (i) For N Poisson distributed with parameter λ and $X_1, X_2, \ldots$ independent of each other and of N, each with distribution F with mean μ, variance σ^2 and characteristic function $\phi(t)$, the compound Poisson distribution of

$$Y := X_1 + \ldots + X_N$$

has CF $\psi(u) = \exp\{-\lambda(1 - \phi(u))\}$, mean $\lambda\mu$ and variance $\lambda E[X^2]$.
(ii) For $N = (N_t)$ a compound Poisson process with rate λ and jump-distribution F with mean μ and variance σ^2, N_t has CF

$\psi(u) = \exp\{-\lambda t(1 - \phi(u))\}$, mean $\lambda t \mu$ and variance $\lambda t E[X^2]$.

Proof. This is Prob/Soln 1 Q2,3. For the mean and variance, we give a second proof below.

Given N, $Y = X_1 + \ldots + X_N$ has mean $NEX = N\mu$ and variance $N\ var\ X = N\sigma^2$. As N is Poisson with parameter λ, N has mean λ and variance λ. So by the Conditional Mean Formula,

$$EY = E[E(Y|N)] = E[N\mu] = \lambda\mu.$$

By the Conditional Variance Formula,

$$\begin{aligned} var\ Y &= E[var(Y|N)] + var\ E[Y|N] \\ &= E[Nvar\ X] + var([N\ E[X]) \\ &= E[N].var\ X + var\ N.(EX)^2 \\ &= \lambda[E[X^2] - (E[X])^2] + \lambda.(E[X])^2 \\ &= \lambda E[X^2] = \lambda(\sigma^2 + \mu^2). \end{aligned}$$

(ii) Apply (i): N_t has mean λt and variance λt. //

In the insurance context (below), the Poisson points represent the claim arrivals, so the Poisson rate λ is the rate at which claims arrive; μ is the mean claim size. So $\lambda\mu$ has the interpretion of a *claim rate* – rate at which money goes *out* of the company in claims.

Just as the mathematics of the Black-Scholes model (Ch. VII) is dominated by *Brownian motion*, that of insurance is dominated by the *Poisson* and *compound Poisson* processes. These are the basic prototypes, and all we have time to cover in detail in this course. However, these are models, of reality, and reality is always more complicated than any model! Box's dictum (George Box, British statistician, 1919-2013): *All models are wrong. Some models are useful.* In more advanced work, more complicated and detailed models are needed. So there is plenty of scope for useful applications in the real world of any probability or statistics you know, or will learn! At the end of the course (VIII.5), we discuss briefly some generalisations. But to note for now: the principal weakness of our assumptions here is the *independence* of claims. This is reasonable under normal conditions, but not during a crisis. Think of natural disasters such as major hurricanes, etc.

§3. Renewal theory

Renewal Processes

Suppose we use components – light-bulbs, say – whose lifetimes $X_1, X_2, \ldots$ are independent, all with law F on $(0, \infty)$. The first component is installed new, used until failure, then replaced, and we continue in this way. Write

$$S_n := \sum_1^n X_i, \qquad N_t := \max\{k : S_k \le t\}.$$

Then $N = (N_t : t \ge 0)$ is called the *renewal process* generated by F; it is a *counting process*, counting the number of failures seen by time t. Note that

$$S_{N(t)} \le t.$$

Note. For stochastic processes, notations such as N_t and $N(t)$ are used interchangeably.

Renewal processes are often used, but the only ones we need here are the Poisson processes – those for which the lifetime law is exponential.

The renewal function

We saw above that

$$N_t/t \to 1/\mu \qquad (t \to \infty), \ a.s.$$

If we apply the expectation operator $E[.]$ formally, this suggests that

$$E[N_t]/t \to 1/\mu \qquad (t \to \infty).$$

This is indeed true, but although its conclusion seems weaker than that of the a.s. result, its proof if harder (though not as hard as that of the SLLN!).

Theorem (Renewal Theorem; Feller 1941; Doob, 1948). If the mean lifetime length μ is finite, the renewal function $E[N_t]$ satisfies

$$E[N_t]/t \to 1/\mu \quad (t \to \infty).$$

Proof. The conclusion with $\ge$ in place of $=$ does indeed follow from the a.s. result by taking expectations. This is by *Fatou's lemma*, which we quote from Measure Theory. [For proof, see e.g. a book on Measure Theory, or my homepage, Stochastic Processes, I.5 Lecture 8.] For the $\leq$ part, choose $a > 0$, and truncate the X_n at level a:

$$\tilde{X}_n := min(X_n, a).$$

Write $\tilde{N}_t$, $\tilde{\mu}$ for the 'tilde' analogues of N_t, μ. By Wald's identity,

$$E[\tilde{X}_1 + \cdots + \tilde{X}_{\tilde{N}_t}] = E[\tilde{X}].E[\tilde{N}_t] = \tilde{\mu}.E[\tilde{N}_t].$$

Now $\tilde{N}_t \geq N_t$ (because of the truncation, there will be more renewals if anything), and $\tilde{S}_{\tilde{N}_t - 1} + \tilde{X}_{\tilde{N}_t} \leq t + a$ (the 't' from the first term, the 'a' from the second). So

$$\begin{array}{rcll} E[N_t]/t & \leq & E[\tilde{N}_t]/t & (N_t \leq \tilde{N}_t) \\ & = & \tilde{\mu}^{-1} E[\tilde{X}_1 + \cdots + \tilde{X}_{\tilde{N}_t}]/t & (\text{above – Wald's identity}) \\ & = & \tilde{\mu}^{-1} E[\tilde{S}_{\tilde{N}_t}]/t & (\text{ definition of } \tilde{S}_n) \\ & \leq & \tilde{\mu}^{-1} & (S_{N(t)} \leq t, \text{ and similarly for } \tilde{S}_n,\ \tilde{N}_t). \end{array}$$

So

$$\limsup E[N_t]/t \leq \tilde{\mu}^{-1}.$$

Now let $a \uparrow \infty$: $\tilde{\mu} \to \mu$, giving the $\leq$ part and the result. //

With F the lifetime distribution function – that of each X_i – the distribution function of $S_n := X_1 + \cdots + X_n$ is $F * \cdots * F$ (n Fs), the n-fold convolution of F with itself, written F^{*n}. Define

$$U(t) := \sum_{n=0}^{\infty} F^{*n}(t).$$

This is called the *renewal function* of F. For, it gives the mean number $E[N_t]$ of renewals up to time t. This gives the reformulation of the Renewal Theorem below.

Theorem (Renewal Theorem, second form; Feller, 1941, Doob, 1948). The renewal function gives the mean number of renewals:

$$U(t) = E[N_t].$$

So if the mean lifetime is μ,

$$U(t)/t \to 1/\mu \qquad (t \to \infty).$$

Proof.

$$\begin{aligned} E[N_t] &= \sum_0^\infty nP(N_t = n) \\ &= \sum n[P(N_t \geq n) - P(N_t \geq n+1)] \\ &= \sum P(N_t \geq n), \end{aligned}$$

by *partial summation* (or *Abel's lemma*). [This is the discrete analogue of integration by parts. See e.g. a book on Analysis, or my homepage, M3P16 Analytic Number Theory, I.3.] But $\{N_t \geq n\} = \{S_n \leq t\}$, so

$$E[N_t] = \sum P(S_n \leq t) = \sum F_n^*(t) = U(t),$$

giving the first part; the second part follows from the result above. //

The renewal theorem

Renewal theory needs a distinction between two cases. If the X_i are integer-valued (when so are the S_n), or are supported by an arithmetic progression (AP), we are in the *lattice case*, otherwise in the *non-lattice case.*

The next result looks like a differenced form of the last one. It is due to David Blackwell (1919-2010) in 1953. We state it for the non-lattice case and $\mu < \infty$, but it extends to the lattice case and $\mu = \infty$ also.

Theorem (Blackwell's renewal theorem, 1948). In the non-lattice case,

$$U(t+h) - U(t) \to h/\mu \quad (t \to \infty) \quad \forall h > 0.$$

This famous result has a number of different proofs, but we do not include one here (my favourite is only a few lines, but needs a prerequisite beyond our scope here).

Blackwell's theorem has a number of variants. The one we need (which we also quote) is due to W. L. Smith and W. Feller. Recall the *Riemann integral* (defined for functions on a finite interval), and the *Lebesgue integral* which generalises it (defined for functions on e.g. the line, plane etc.). We need a new concept.

Definition. Divide the line into intervals $I_{n,h} := (nh, (n+1)h]$. For a function z on $\mathbb{R}$ and $x \in I_{n,h}$, write

$$\overline{z}_h := sup\{z(y) : y \in I_{n,h}\}, \qquad \underline{z}_h := inf\{z(y) : y \in I_{n,h}\}.$$

Call z *directly Riemann integrable (dRi)* if $\int \overline{z_h} := \int_{-\infty}^{\infty} \overline{z}_h(x)dx$ is finite for some (equivalently, for all) $h > 0$, and similarly for $\int \underline{z_h}$, and

$$\int \overline{z}_h - \int \underline{z}_h \to 0 \qquad (h \to 0).$$

This is the same as Riemann integrability if z is supported on some finite interval, but for z of unbounded support is stronger than Lebesgue integrability: z is dRi iff it is Lebesgue integrable, and both $\int \overline{z}_h$ and $\int \underline{z}_h$ have a common limit $\int z$ as $h \to 0$. Condition dRI will hold whenever we need it. We quote that dRi needs z bounded and a.e. continuous (w.r.t. Lebesgue measure), and that this plus z of bounded support implies dRi. Also, z non-increasing and Lebesgue integrable imples dRi.

The *renewal equation* for F and z (both known) is the integral equation

$$Z(t) = z(t) + \int_0^t Z(t-u)dF(u) \quad (t \geq 0): \quad Z = z + F * Z. \qquad (RE)$$

Here F (for us, the lifetime distribution above) and z are given, and (RE) is to be solved for Z.

Theorem (Key Renewal Theorem; W. L. Smith, 1954, 1955). If z in (RE) is dRi, then for U the renewal function of F as above,

$$\lim_{t\to\infty} Z(t) = \lim_{t\to\infty}(U * z)(t) = \frac{1}{\mu}\int_0^\infty z(x)dx.$$

The proof of the Key Renewal Theorem from Blackwell's Renewal Theorem is not long or hard, but as it is Analysis rather than probability or insurance mathematics, we omit it. For a proof, see e.g. [RSST, 6.1.4 p216-219.

§4. The Ruin Problem.

Consider the cash flow of an insurance company. The premium income comes in from the policy holders at constant rate, c say (to a first approximation: the company hopes to attract more policy holders, and premium rates will typically vary on renewal – but are constant during the lifetime of the policy). So income over time t is ct. If the company has initial capital u, its capital at time t is thus $u + ct$. Meanwhile, claims occur. We model these as occurring at the instants of a Poisson process of rate λ, the claims being independent and identically distributed (iid) with claim distribution F, with CF ϕ, mean μ and variance σ^2. So the number of claims over the interval $[0, t]$ is $N(t)$, which is Poisson distributed with parameter λt: $N(t) \sim P(\lambda t)$. So by the Theorem of VIII.2 above, the total claim has mean $\lambda\mu t$. Thus cash comes in at rate c, but goes out at rate $\lambda\mu$. This simple argument suggests – what is indeed true – that a *necessary* condition for the company to avoid bankruptcy is

$$c > \lambda\mu \qquad (NPC)$$

(NPC for 'net profit condition, below): money should come in *faster* than it goes out. The proof is by the Strong Law of Large Numbers (LLN, as above). In the critical case $c = \lambda\mu$ the company is 'balanced on a knife-edge', and will soon go bankrupt.

The company thus *must* have $c > \lambda\mu$, so we assume this from now on. But, any insurance company has only finite funds; it may face arbitrarily severe runs of bad luck; combining these, bankruptcy is always a possibility. (Indeed, this is true for all companies, not just insurance companies! This is why bankruptcy needs to be recognised as a possibility, and governed by bankruptcy law. This varies from time to time and from country to country – a very interesting and important subject, but not one we can pursue here.)

Clearly the company's best defence against bankruptcy is to have a large cash reserve u, to act as a buffer, or 'insurance policy', against such runs of bad luck. Clearly the probability of ruin – ruin probability – decreases with u. How fast? The classical *ruin problem* is to investigate this question.

Note. We may if we wish take $c = 1$ for convenience. This (slightly) simplifies the formulae. It amounts to changing from real time to *operational* or *business* time – looking at the situation in the time-scale most natural to it. Recall that there are *no natural units of time or space* (except the Planck scale, at subatomic level, for those with a background in Physics!): time is measured in seconds, minutes, hours, days (60 s to the m, 60 m to the h, 24 h to the day – pre-decimal), and length in metres (metric system – mm, cm, m, km) or inches/feet/yards/miles (Imperial measure) – neither is natural, both are conventional.

The Net Profit Condition (NPC)
With c the premium rate, X_i the claim sizes and W_i the inter-claim waiting times, write

$$Z_i := X_i - cW_i.$$

Then

$$E[Z_i] := E[X_i] - cE[W_i] = \mu - c/\lambda.$$

The first term on the right measures money *out* (of the company), the second measures money *in*. To avoid bankruptcy we need ('more in than out')

$$E[Z_i] := E[X_i] - cE[W_i] = \mu - c/\lambda < 0 : \qquad c > \mu\lambda. \qquad (NPC)$$

This is the *net profit condition* (NPC) above. For as we have seen, $\lambda\mu$ is the claim rate (rate at which cash goes *out* to claims); c is the *premium rate* (rate at which cash comes *in*, through premiums); we need (NPC) – 'more in than out' for survival.

Safety loading and premium calculation
The first duty of any company is to stay solvent – to avoid bankruptcy. To do this, an insurance company has to have its premium rate $c > \mu\lambda$ so as to satisfy (NPC). But, like any other business, the insurance business is competitive. If premiums are too *low*, the firm goes bankrupt (above) because its premium income fails to meet its outgoings on claims. But if premiums are too *high*, the firm will not be competitive with other firms; over time, it will lose market share to them, and will eventually go bankrupt (or otherwise go out of business – e.g., be taken over) as premium income declines to be too small to meet overheads. So the firm needs to take a policy decision as to how much to charge in premiums. This is measured by the *safety loading (SL)*, ρ, defined by

$$c = (1+\rho)\frac{E[X_i]}{E[W_i]} = (1+\rho)\lambda\mu : \qquad \rho := \frac{c - \lambda\mu}{\lambda\mu}. \qquad (SL)$$

Thus $\rho > 0$ in (SL) is equivalent to (NPC).

5. Lundberg's inequality

Before, we used the characteristic function (CF), defined for a random variable X by $\phi(t) := E[e^{itX}]$, for t real. But we now find it convenient to use real numbers, and switch to the *moment-generating function (MGF)*,

$$M(s) := E[e^{sX}].$$

This is certainly defined for $s = 0$: $M(0) = E[e^0] = E[1] = 1$. But it may not be defined (finite) for all (or even any) $s \neq 0$. (Example: the *exponential* distribution $E(\lambda)$ with parameter λ has MGF $\lambda/(\lambda - s)$, but this is only finite for $s < \lambda$.) We now assume the *small claim condition (SCC)*,

$$M(s) := E[e^{sX_1}] < \infty \quad \forall s \in (-s_0, s_0), \quad \text{for some } s_0 > 0. \qquad (SCC)$$

Note. 1. This condition implies that the MGF is *holomorphic (analytic)* in a neighbourhood of the origin in the complex s-plane (indeed, in an open vertical strip in the s-plane containing the origin. This means that the MGF behaves very well here, and we can use the methods of Complex Analysis (M2P3). The bigger this strip, the better; what limits the strip is the *tail* $1 - F$ of the claim-size law F – that is, the *large claims* (hence the name 'small-claims condition'). So, the bigger $s_0 > 0$, the better.
2. (SCC) implies that the tail of X_1 decays exponentially. For (*Markov's Inequality*): for $s \in (0, s_0)$ and $x > 0$,

$$M(s) = E[e^{sX_1}] \geq E[e^{sX_1}; X_1 > x] \geq e^{sx} E[1; X_1 > x] = e^{sx} P(X_1 > x):$$

$$P(X_1 > x) \leq e^{-sx} M(s) \qquad \forall x > 0.$$

Differentiating the MGF twice (and writing X for X_1 for convenience):

$$M(s) = E[e^{sX}], \qquad M'(s) = E[Xe^{sX}], \qquad M''(s) = E[X^2 e^{sX}] \geq 0.$$

Also, the MGF M(s) is smooth (we can differentiate it as often as we like, where it is defined). So its graph has a tangent, and as $M'' \geq 0$, the *tangent is increasing* – the *graph bends upwards*. Such functions are called *convex*. Also, as $M(0) = 1$, the graph goes through 1 at the origin. Now smooth convex functions can intersect any line at most twice (e.g., a parabola may not cut a line, or can cut it once (double point of contact), or twice, but not more).

The crucial assumption is that $M(s)$ cuts the line $y = 1$ *twice*, once (necessarily) at the origin and once at a positive point r.

Definition.

The *Lundberg coefficient* (or *adjustment coefficient*) r, which we assume to exist in what follows, is the point $r > 0$ (we assume r exists; it is then unique) such that $r = s$ satisfies

$$M_{Z_1}(s) := E[\exp\{s(X_1 - cW_1)\}] = 1. \qquad (LC)$$

The right is (writing X, W for X_1, W_1) $M_X(s).M_W(-cs)$. Now as $W \sim E(\lambda)$, W has Laplace-Stieltjes transform (LST) $E[e^{-tW}] = M_W(-t) = \int_0^\infty e^{-tx}.\lambda e^{-\lambda x}dx = \lambda/(\lambda + t)$. So the defining property of the Lundberg (adjustment) coefficient is (writing M for M_X for short)

$$M(r).\frac{\lambda}{\lambda + cr} = 1 : \qquad M(r) = \frac{\lambda + cr}{\lambda} = 1 + \frac{cr}{\lambda}. \qquad (LC')$$

As s increases, $M(s) \uparrow \infty$; $M(s)$ may well increase to $+\infty$ as s increases to some finite limit, s_1 say. So (from a graph of $M(.)$): the bigger r is ($r < s_0 \leq s_1$), the better. Also, the bigger u is, the better. So, the bigger ru is, the better. So the ruin probability $\psi(u)$ with initial cash reserve u decreases with ru. How nice it would be if this decrease were exponential (fast, and easy!) Fortunately, this is just what happens, as the next two classical results show.

Theorem (Lundberg's Inequality, 1903, 1926). Assuming that the Small-Claims Condition (SCC) holds and that the Lundberg coefficient r in (LC) exists, the ruin probability $\psi(u)$ with initial capital u and over all time satisfies

$$\psi(u) \leq e^{-ru}.$$

Proof. Write

$$S_n := Z_1 + \cdots + Z_n, \qquad Z_i := X_i - cW_i.$$

Then $S = (S_n)$ is a *random walk*, with *step-lengths* $Z_i := X_i - cW_i$. As the ruin probability increases with time, the ruin probability $\psi(u)$ is the increasing limit of the ruin probability $\psi_n(u)$ with just the first n claims X_i and waiting times W_i involved:

$$\psi_n(u) = P(\max_{1 \leq k \leq n} S_k > u) = P(S_k > u \text{ for some } k \in \{1, \cdots, n\}\ ). \qquad (RW_n)$$

We prove that

$$\psi_n(u) \leq e^{-ru} \qquad \forall n \in \mathbb{N}, u > 0. \tag{$*$}$$

The result follows from this by letting $n \to \infty$; we prove $(*)$ by induction (on n). Note that letting $n \to \infty$ in (RW_n) above,

$$\psi(u) = P(\sup_n S_n > u). \tag{RW}$$

The induction starts, by Markov's Inequality:

$$\psi_1(u) \leq e^{-ru} M_{Z_1}(r) = e^{-ru},$$

by definition of the Lundberg coefficient: $M_{Z_1}(r) = 1$.

Assume that $(*)$ holds for n, and write F for F_{Z_1}, the distribution function of Z_1. Then

$$\begin{aligned}
\psi_{n+1}(u) &= P(\max\{S_k : 1 \leq k \leq n+1\} > u) \\
&= P(Z_1 > u) + P(Z_1 \leq u, \max\{Z_1 - (S_k - Z_1) : 2 \leq k \leq n+1\} > u) \\
&= p_1 + p_2,
\end{aligned}$$

say.

We now make our first use of the *renewal argument*, which will allow us to reduce the proof of our main results to an application of the Key Renewal Theorem. The idea is to *condition* on the value of the first claim Z_1, and let the process 'renew itself' with the first claim, starting afresh thereafter. So, starting the random walk after $Z_1 = x$ in the p_2-term above and conditioning on the value x of Z_1,

$$p_2 = \int_{(-\infty,u]} P(\max_{1 \leq k \leq n}(x + S_k) > u) dF(x).$$

In full, this is a use of the Conditional Mean Formula. For an event A, the random variable I_A (its indicator function: 1 if $\omega \in A$, 0 if not) has mean

$$E[I_A] = P(A).$$

Then conditioning on information $\mathcal{B}$ (size of first claim here),

$$P(A) = E[I_A] = E[E[I_A|\mathcal{B}]].$$

Now

$$p_1 = \int_{(u,\infty)} dF(x) \leq \int_{(u,\infty)} e^{r(x-u)} dF(x),$$

as $r > 0$, while

$$\begin{aligned} p_2 &= \int_{(-\infty,u]} P(\max_{1\leq k\leq n}(x+S_n)>u)dF(x) \\ &= \int_{(-\infty,u]} \psi_n(u-x)dF(x) \\ &\leq \int_{(-\infty,u]} e^{r(x-u)}dF(x) \quad \text{(by the induction hypothesis).} \end{aligned}$$

Combining the domains $(-\infty, u]$ and (u, ∞) of integration here,

$$p_1 + p_2 \leq \int_{-\infty}^{\infty} e^{r(x-u)}dF(x) = e^{-ru}\int e^{rx}dF(x) = e^{-ru}M(r) = e^{-ru},$$

as $M(r) = 1$ by definition of the Lundberg coefficient r, completing the induction. //

Example: Exponential claims.

Recall the exponential distribution $E(\lambda)$ with parameter λ, which has mean $1/\lambda$ and MGF $\lambda/(\lambda - s)$. With the arrival process Poisson with rate λ as above (so the inter-claim waiting times are $E(\lambda)$), consider now the simplest case, when the claim sizes are also exponential, $E(\gamma)$ say. So W_i has MGF $\gamma/(\gamma - s)$, cW_i has MGF $\gamma/(\gamma - cs)$, and $Z_i = X_i - cW_i$ has MGF

$$M_Z(s) = M_X(s)M_{cW}(-s) = \frac{\gamma}{\gamma - s}.\frac{\lambda}{\lambda + cs}.$$

As usual, we assume the Net-Profit Condition (NPC):

$$E[X]/E[W] = \lambda/\gamma < c.$$

Then the Lundberg coefficient r is the (unique, positive) root of

$$M_Z(r) = \frac{\gamma}{\gamma - r}.\frac{\lambda}{\lambda + cr} = 1.$$

This is a quadratic,

$$Q(r) := -[(cr+\lambda)(-r+\gamma) - \lambda\gamma] = cr^2 + (\lambda - c\gamma)r = r(cr + \lambda - c\gamma) = 0,$$

with positive root

$$r = \gamma - \frac{\lambda}{c} > 0,$$

by (NPC). In terms of the safety loading ρ,

$$c = \frac{E[X]}{E[W]}(1+\rho) = \frac{\lambda}{\gamma}(1+\rho).$$

So in terms of the safety loading ρ rather than the premium rate c,

$$r = \gamma \frac{\rho}{(1+\rho)},$$

and the Lundberg inequality is

$$\psi(u) \leq \exp\{-u\gamma\rho/(1+\rho)\}.$$

This is nearly exact: in this case, there is a constant C with

$$\psi(u) = C \exp\{-u\gamma\rho/(1+\rho)\}.$$

Note. This example is unusually simple: in general, there is no closed form for r, and we have to find it by numerical methods. This is typically the case for solutions of transcendental (rather than algebraic) equations.

6. The ruin problem and the renewal equation

Here and in §7 we follow Mikosch [Mik, p.166-171]. First, note that F has mean

$$\mu := \int_0^\infty x dF(x) = -\int_0^\infty x d(1-F)(x).$$

Integrating by parts, the integrated term vanishes, giving

$$\mu = \int_0^\infty (1 - F(x)) dx.$$

Thus $(1 - F(x))/\mu$ is a probability density on $(0, \infty)$, of G, say:

$$dG(x) = \frac{1 - F(x)}{\mu} dx$$

(the notation F_I, for 'integrated tail of F', is also used).

With initial capital u, write $\psi(u)$ for the probability of ruin as above, $\phi(u) := 1 - \psi(u)$ for the probability of non-ruin. Then by (RW),

$$\psi(u) = P(\sup S_n > u), \qquad \phi(u) = P(\sup S_n \le u).$$

The key to the relevance of renewal methods here – the *renewal argument* we used before – is that the capital process *renews itself at the time of the first claim*: if this is at time $W_1 = s$ and of size $X_1 = x$, it begins again, with initial capital $u + cs - x$ (of course if this is negative, the company goes bankrupt when it receives its first claim!). We can *condition* (as above) on the time W_1 (density $\lambda e^{-\lambda s}$) and size X_1 (distribution F) of first claim.

$$\begin{aligned}
\phi(u) &= P(S_n \le u\ \forall n \ge 1) \\
&= P(Z_1 \le u, S_n - Z_1 \le u - Z_1\ \forall n \ge 1) = E[I(...)] \\
&= E[E[I(Z_1 \le u, S_n - Z_1 \le u - Z_1\ \forall n \ge 2|Z_1)]] \text{ (Conditional Mean Formula)} \\
&= E[I(Z_1 \le u)E[I(S_n - Z_1 \le u - Z_1\ \forall n \ge 2|Z_1)]] \text{ (taking out what is known)} \\
&= E[I(Z_1 \le u)P(S_n - Z_1 \le u - Z_1\ \forall n \ge 2|Z_1)] \quad (E[I(.)] = P(.)) \\
&= E[I(Z_1 \le u)P(T_n - Z_1 \le u - Z_1\ \forall n \ge 1|Z_1)],
\end{aligned}$$

writing $T_n := Z_2 + \cdots + Z_{n+1} = S_{n+1} - Z_1$. But T_n is independent of Z_1, and given $Z_1 - z = x - cw$, T_n has the same law as S_n. Recall $X_1 \sim F$, $W_1 \sim E(\lambda)$ with density $\lambda e^{-\lambda w}$. So doing the conditioning,

$$\begin{aligned}
\phi(u) &= E[I(X_1 - cW_1 \le u)P(T_n \le u - (X_1 - cW_1)|Z_1)] \\
&= \int_0^\infty \lambda e^{-\lambda w} dw \int_0^\infty dF(x) I(x - cw \le u) P(S_n \le u - (x - cw)\ \forall n \ge 1) \\
&= \int_0^\infty \lambda e^{-\lambda w} dw \int_0^\infty dF(x) I(x - cw \le u)\phi(u + cw - x)
\end{aligned}$$

(this is the renewal argument again). Thus $\phi(u)$ satisfies a linear integral equation, which we shall show is 'almost' of renewal-equation type (the key is to make it exactly of renewal type).

The limits are $0 < w < \infty$, $0 < x < u + cw$:

$$\phi(u) = \int_0^\infty \lambda e^{-\lambda w} dw \int_0^{u+cw} dF(x) \cdot \phi(u + cw - x).$$

Write $z := u + cw$, and change from w to z: limits $0 < x < z$, $u < z < \infty$, $dw = dz/c$, $w = (z-u)/c$, $-\lambda w = \lambda u/c - \lambda z/c$:

$$\phi(u) = \frac{\lambda}{c} e^{\lambda u/c} \int_u^\infty dz e^{-\lambda z/c} dw \int_0^z dF(x) \cdot \phi(z-x). \qquad (*)$$

Write

$$g(z) := \int_0^z \phi(z-x) dF(x) :$$

then $(*)$ becomes

$$\phi(u) = \frac{\lambda}{c} e^{\lambda u/c} \int_u^\infty e^{-\lambda z/c} g(z) dz.$$

So ϕ is differentiable, as the exponential and the integral are. So differentiating $(*)$,

$$\phi'(u) = \frac{\lambda}{c}\phi(u) - \frac{\lambda}{c} e^{\lambda u/c} \cdot e^{-\lambda u/c} \int_0^u \phi(u-x) dF(x)$$

(the first term from differentiating the exponential, the second from differentiating the integral):

$$\phi'(u) = \frac{\lambda}{c}\phi(u) - \frac{\lambda}{c}\int_0^u \phi(u-x)\ dF(x).$$

Now integrate this:

$$\phi(t) - \phi(0) - \frac{\lambda}{c}\int_0^t \phi(u) du = -\frac{\lambda}{c}\int_0^t du \int_0^u dF(x) \cdot \phi(u-x).$$

Integrating by parts,

$$\int_0^u \phi(u-x) dF(x) = \phi(0)F(u) - \int_0^u \phi'(x-u) F(x) dx$$

(as $F(0) = 0$). Combining,

$$\phi(t) - \phi(0) = \frac{\lambda}{c}\int_0^t \phi(u) du - \frac{\lambda}{c}\phi(0)\int_0^t F(u) du + \frac{\lambda}{c}\int_0^t du \int_0^u dx \phi'(x-u) F(x).$$

The limits here are $0 < x < u < t$. So interchanging the order of integration, the limits become $u \in (x,t)$, $x \in (0,t)$. This gives

$$\phi(t) - \phi(0) = \frac{\lambda}{c}\int_0^t \phi(u) du - \frac{\lambda}{c}\phi(0)\int_0^t F(u) du - \frac{\lambda}{c}\int_0^t F(x)[\phi(t-x) - \phi(0)] dx.$$

The $\phi(0)$ terms (2nd and 4th on RHS) cancel, and the first integral on RHS is $\int_0^t \phi(t-x)dx$, giving

$$\phi(t) - \phi(0) = \frac{\lambda}{c}\int_0^t \phi(t-x)[1-F(x)]dx = \frac{\lambda}{c}\int_0^t \phi(t-x)\overline{F}(x)dx,$$

or by (SL) (§4),

$$\begin{aligned}\phi(t) - \phi(0) &= \frac{1}{(1+\rho)\mu}\cdot\int_0^t \phi(t-x)\overline{F}(x)dx, \\ &= \frac{1}{(1+\rho)}\cdot\int_0^t \phi(t-x)dG(x),\end{aligned}$$

recalling G (the integrated tail distribution at the beginning of §6).

By the NPC (§4), $c > \lambda\mu$, so $E[Z] = E[X] - cE[W] = \mu - c/\lambda < 0$. So by LLN, $S_n := \sum_1^n Z_k \to -\infty$ (as $n \to \infty$, a.s.), so $\sup_n S_n < \infty$ a.s. So the non-ruin probability $\phi(u) \uparrow 1$ as $u \to \infty$. This allows us to find $\phi(0)$ above:

$$\phi(u) - \phi(0) = \frac{1}{(1+\rho)}\int_0^\infty I(x<u)\phi(u-x)dG(x).$$

Letting $u \uparrow \infty$, Lebesgue's monotone convergence theorem (we quote this from Measure Theory) allows us to interchange limit and integral here:

$$1 - \phi(0) = \frac{1}{(1+\rho)}\int_0^\infty 1dG(x) = \frac{1}{(1+\rho)}: \qquad \phi(0) = \frac{\rho}{(1+\rho)}.$$

Combining, we obtain the integral equation for the non-ruin probability $\phi(u)$:

$$\begin{aligned}\phi(u) &= \frac{\rho}{(1+\rho)} + \frac{1}{(1+\rho)}\cdot\int_0^u \phi(u-x)dG(x) \\ &= \frac{\rho}{(1+\rho)} + \frac{1}{(1+\rho)}\cdot\int_0^u \phi(u-x)\frac{(1-F(x))}{\mu}dx.\end{aligned}$$

We re-write this as the corresponding integral equation for the ruin probability $\psi(u) = 1 - \phi(u)$:

$$\psi(u) = \frac{1}{(1+\rho)}\int_u^\infty \frac{(1-F(x))}{\mu}dx + \frac{1}{(1+\rho)}\cdot\int_0^u \psi(u-x)\frac{(1-F(x))}{\mu}dx \quad (**)$$

(as $(1-F(x))/\mu$ is a probability density, so integrates to 1).

7. Cramér's estimate of ruin

The above integral equation $(**)$ for $\psi(u)$ is of renewal-equation type, *except* that, as $(1-F(x))/\mu$ is a probability *density*, the factor $1/(1+\rho)<1$ turns it into a *sub-probability* (or *defective*) density.

Next, from the existence of the Lundberg coefficient $r>0$ in (LC), (LC'),

$$M(r) := \int_0^\infty e^{rx} dF(x) = -\int_0^\infty e^{rx} d(1-F)(x) = 1 + \frac{cr}{\lambda}.$$

Integrating by parts (as above), the integrated term is 1, giving

$$\int_0^\infty (1-F(x))e^{rx}dx = \frac{c}{\lambda}, = (1+\rho)\mu,$$

by (SL). So

$$\frac{\lambda}{c}(1-F(x))e^{rx} = \frac{1}{(1+\rho)\mu}(1-F(x))e^{rx}$$

is a probability density on $(0,\infty)$.

The following result was obtained by Cramér in 1930, by complex-variable methods (Cauchy's theorem). Complex-variable methods turn out not to be natural here. The right tools are real analysis (direct Riemann integrability, key renewal theorem) and probability theory (renewal theory); the link was made by W. Feller, and is in his book (1966, 2nd ed. 1971).

Theorem (Cramér's estimate of ruin, 1930).

For the Cramér-Lundberg model, under the Net Profit Condition (NPC) and the Lundberg condition (LC), with r the Lundberg coefficient and $\psi(u)$ the probability of ruin with initial capital u,

$$e^{ru}\psi(u) \to C: \qquad \psi(u) \sim Ce^{-ru} \qquad (u\to\infty),$$

where the constant C is given by

$$C = \frac{c-\lambda\mu}{\lambda r \int_0^\infty x e^{rx}(1-F(x))dx}.$$

Proof. Multiply $(**)$ by e^{ru}, and regard it as an integral equation in $\psi(u)e^{ru}$:

$$[\psi(u)e^{ru}] = e^{ru}\int_u^\infty \frac{(1-F(x))}{(1+\rho)\mu}dx + \int_0^u [\psi(u-x)e^{r(u-x)}]\frac{e^{rx}(1-F(x))}{(1+\rho)\mu}dx.$$

This *is* now an integral equation of renewal type (RE). So by the Key Renewal Theorem, its solution $\psi(u)e^{ru}$ has a *limit*, C say, as $u \to \infty$, giving the first (and more important) part.

To identify the limit C: from the Key Renewal Theorem, C is the integral of the first (z-) term on the right, divided by the mean of the probability distribution in the convolution. The integral here is

$$\begin{aligned}\int_0^\infty e^{ru}du \int_u^\infty (1 - F(x))dx &= \frac{1}{r}\int_0^\infty [\int_u^\infty (1 - F(x))dx]d(e^{ru}) \\ &= \frac{1}{r}[e^{ru}\int_u^\infty (1 - F(x))dx]_0^\infty + \frac{1}{r}\int_0^\infty e^{ru}(1 - F(u))du \\ &= -\frac{\mu}{r} + \frac{c}{r\lambda} = \frac{c - \lambda\mu}{\lambda r},\end{aligned}$$

by the calculation above. So, in the notation of the Key Renewal Theorem,

$$\int_0^\infty z(x)dx = \frac{\lambda}{c}.\frac{c - \lambda\mu}{\lambda r}.$$

The mean of this density (the 'μ' term in the Key Renewal Theorem) is

$$\frac{\lambda}{c}.\int_0^\infty xe^{rx}(1 - F(x))dx.$$

So C is their ratio:

$$C = \frac{c - \lambda\mu}{\lambda r \int_0^\infty xe^{rx}(1 - F(x))dx}. \qquad //$$

Note. In addition to the Key Renewal Theorem, the crux in the above is the *change of measure*

$$F = F(dx) \mapsto \frac{\lambda}{c}(1 - F(x))e^{rx}dx.$$

This is also called *exponential tilting* and the *Esscher transform*, after the Swedish actuary Fredrik Esscher in 1932. (It also occurs in *large deviations*, important in many areas of probability, statistics and statistical mechanics.) This change-of-measure technique is of course also related to that in *Girsanov's theorem* in mathematical finance (Ch. VII).

8. More on insurance.

Non-life insurance: regression and covariates

House insurance

If one insures a house's *contents*, one of the the principal risk factors the insurance company will consider (and the easiest one to measure) is the risk of *burglary*. This varies greatly according to the nature of the area: affluent areas have more to attract a burglar, but tend to have better burglar alarms; poorer areas tend to have higher crime rates, etc. If one insures a house as a *building*, the principal risk factor is *subsidence*. This depends largely on the geological conditions in the area (and so are indicated by the postal code), but also on the quality of the building at the time the area was developed (which can be assessed from past claims). Risk of *fire* is important in both, but harder to assess (it depends on people not leaving chip-pans on the cooker when called to the door or the phone, etc.). These subsidiary bits of information are called *covariates*; the way to use them is called *regression*. This kind of statistics is very useful in the actuarial/insurance profession.

Motor insurance

Motor insurance rates vary widely. Of course, the most important single thing is the claims record of the insuring motorist – a good record is worth money, in a no-claims bonus. But, the type of car is also relevant (sports cars are penalised); so is the type of driver (young men are penalised), the annual mileage, the type of use (private or for hire), etc.

Life insurance

Eventual death is certain, so life insurance is largely a matter of covariates such as: age, sex, medical record, profession etc. The tools involved come under Survival Analysis: hazard rates, etc. Following the introduction of the *proportional hazards model* by Cox in 1972, martingale methods have been widely used. This is a very interesting and useful area.

To give some flavour of Survival Analysis: suppose that a person survives for time t. What is the chance that he dies by time $t + dt$? With T as the lifetime, with distribution function F on $(0, \infty)$, density f and tail $\overline{F}(x) = 1 - F(x)$, this is

$$\begin{aligned} P(T \leq x + dx | T > x) &= P(x < T \leq x + dx)/P(T > x) \\ &= (F(x + dx) - F(x))/(1 - F(x)) \\ &\sim f(x)dx/(1 - F(x)) \\ &= h(x)dx, \end{aligned}$$

say, where $h(x)$ has the interpretation of a *hazard rate*. So

$$h(x) = f(x)/(1 - F(x)).$$

Integrating,

$$1 - F(x) = \exp\{\int_0^x h(u)du\} : \qquad F(x) = 1 - \exp\{\int_0^x h(u)du\}.$$

The simplest case is *constant* hazard rate, λ say, leading to the *exponential* distribution $E(\lambda)$, and so to the *Poisson process* $Ppp(\lambda)$ of VII.2:

$$h(x) \equiv \lambda, \qquad F(x) = 1 - e^{-\lambda x}, \quad (x > 0) : \qquad F = E(\lambda).$$

Now hazard rates vary according to many factors, or covariates: age (older people die out faster than younger ones); medical history; weight, smoking status, occupation, marital status (married people live longer!), etc. So applicants for life insurance will be asked to fill out a form detailing the covariates the insurance company deems relevant; assessing the premium depending on these covariates involves regression, as with the non-life examples above.

Reinsurance

Reinsurers play a major role, in the modern economy, beyond insuring insurers. Reinsurance companies act as de facto *regulators*: they monitor insurers and put a price on their heads. The government need have no say, as 'it's money that talks here'. A good reinsurance premium implies confidence, and makes it easier for the primary insurer to raise capital on the open market. Insurers hold, to cover losses, a mix of cash reserve, investment reserve and reinsurance. (It used to be that the reinsurance pot was biggest, but that is changing as investment becomes more affordable.) The basic fact is that the balance of the three sources of capital is important, and precarious: the reinsurance company watches the cash position of the client like a hawk.

Lender of last resort

Companies may fail, and disappear (leaving debts behind them, as well as lost jobs, etc.). But countries cannot disappear (even though sovereign states have on occasion defaulted on debt, split up, etc.). The ultimate underpinning (in so far as there is one) here is provided by the state, in the form of the central bank – the Bank of England (BoE) in the UK, the Federal Reserve Bank (Fed) in the USA, the European Central Bank (ECB) in the EU, and indeed the World Bank at UN level. The phrase 'lender of last resort' is used to convey this.

Postscript to Ch. VIII, Insurance Mathematics

As noted in VIII.1, the actuarial profession regulates itself carefully. The Institute of Actuaries sets professional exams, which intending actuaries must pass in order to become qualified. In order to earn exemption by passing a course at university, the university course (particularly its syllabus) must be accredited (validated) by the Institute. (The situation is similar in the accountancy profession.)

The two main centres for actuarial work in the UK are London and Edinburgh. In London, the City University was an early centre, followed later by the London School of Economics (LSE). The LSE's Risk and Stochastics MSc has now become a major producer of actuaries. In Edinburgh, a similar role has long been played by Heriot-Watt University.

As a glance at the skyline in the City of London reveals, London is a major world financial centre. The financial services industry is one of the UK's major industries (thirty years ago manufacturing industry predominated – recall that the UK pioneered the Industrial Revolution – but this is no longer so). Most of the leading UK Mathematics Departments have MSc programmes in Financial Mathematics. I think it is fair to say that UK academia provides well for the needs of the financial services industry. I think it is also fair to say that it provides less well for the needs of the actuarial profession and the insurance industry. This is a great pity (recall from VIII.1 the UK's historic leading role here).

I am very pleased that Insurance Mathematics is included in the syllabus for this course. I would urge anyone taking this course who does not already have a clear career path mapped out ahead of them to consider actuarial work (which I would probably have gone into myself had I not been sucked into academia). The work is very useful, and very interesting.

It is worth noting that the boundary between the mathematics of finance (Ch. I-VII) and insurance (Ch. VIII) has become quite blurred in recent years. This is partly because, following the Crash of 2008 and a number of major defaults, default in finance is seen as analogous to death in life insurance or a claim in non-life insurance. The two areas are no longer separate, as they once were, and the trend towards further interaction will no doubt continue. So it does not have to be an 'either or' choice for you!

NHB