6. The ruin problem and the renewal equation.

With initial capital u, the company's capital at time t is

$$C_t = u + ct - \sum_1^{N_t} X_i.$$

The probability of (ultimate) ruin and of survival are

$$\psi(u) = P(\inf_{0 < t < \infty} C_t < 0 | C_0 = u),$$

$$\overline{\psi}(u) := 1 - \psi(u) = P(C_t \geq 0 \ \forall t | C_0 = u).$$

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:

$$\overline{\psi}(u) = \int_0^\infty \lambda e^{-\lambda s} ds \int_0^{u+cs} dF(x) \overline{\psi}(u + cs - x).$$

Change variable from s to $t := u + cs$: the limits $0 < x < u + cs$, $s > 0$ become $0 < x < t$, $t > u$:

$$\overline{\psi}(u) e^{-\lambda u/c} = \frac{\lambda}{c} \int_u^\infty \lambda e^{-\lambda t/c} dt \int_0^t dF(x) \overline{\psi}(t - x).$$

This shows that $\overline{\psi}$ is differentiable (as the exponential and the integral are). Differentiating w.r.t. u:

$$e^{-\lambda u/c}(\overline{\psi}'(u) - \frac{\lambda}{c}\overline{\psi}(u)) = -\frac{\lambda}{c} e^{-\lambda u/c} \int_0^u \overline{\psi}(u - x) dF(x) :$$

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

Integrate over $u \in [0, t]$, and write

$$h(y) := \int_0^{t-y} \overline{\psi}(u) du \quad (0 \leq y \leq t), \quad 0 \quad (y > t).$$

The first term on the right integrates to $h(0)$, so

$$\begin{aligned}
\overline{\psi}(t) - \overline{\psi}(0) - \frac{\lambda}{c} h(0) &= -\frac{\lambda}{c} \int_0^t du \int_0^u \overline{\psi}(u - x) dF(x) \\
&= -\frac{\lambda}{c} \int_0^t du h(x) dF(x) \quad \text{(def. of } h(.) \text{ on } [0, t]) \\
&= -\frac{\lambda}{c} \int_0^\infty du h(x) dF(x) \quad (h(.) = 0 \text{ on } [t, \infty]) \\
&= \frac{\lambda}{c} \int_0^\infty du h(x) d(1 - F)(x).
\end{aligned}$$

Integrating by parts, the integrated term on the right cancels with the last term on the left:

$$\overline{\psi}(t) - \overline{\psi}(0) = -\frac{\lambda}{c} \int_0^\infty h'(x)(1 - F(x)) dx :$$

$$\overline{\psi}(t) = \overline{\psi}(0) + \frac{\lambda}{c} \int_0^\infty \overline{\psi}(t - x)(1 - F(x)) dx.$$

This integral equation for $\overline{\psi}$ translates into one for ψ itself:

$$\psi(u) = \frac{\lambda}{c} \left(\int_u^\infty (1 - F(x)) dx + \int_0^\infty \psi(u - x)(1 - F(x)) dx \right). \quad (*)$$

Note that F has mean

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

Integrating by parts (rather as above), the integrated term vanishes, giving

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

Thus $(1 - F(x))/\mu$ is a probability density on $(0, \infty)$, and the integral equation $(*)$ above is of *renewal-equation* type. This is crucial: it reduces the proof of

2

the main result (Cramér's estimate of ruin, below) to an application of the Key Renewal Theorem.

Multiplying both sides of $(*)$ by e^{ru} gives

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

$$(**)$$

7. Cramér's estimate of ruin.

Theorem (Cramér's estimate of ruin).

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

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

Proof. From the existence of the Lundberg coefficient $r > 0$ in (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 (again as above!), the integrated term is 1, giving

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

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

is a probability density on $(0, \infty)$, which shows that $(**)$ is 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

3

of the first (z-) term on the right, divided by the mean of the probability distribution in the convolution. The integral term is λ/c times

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

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

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

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

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

So C is their ratio:

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

Note. 1. The argument above draws on several sources: [Mik 4.2.2, 166-171], [AA, 4.5a p90], [A, IV.2 Ex. 2.3; IV.4, 5], [RSST, 5.3.2, 5.4.2].
2. 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. VI).

Filip Lundberg
   Filip Lundberg (1876-1965) was a Swedish actuary and pioneer of the

theory of collective risk. His work in actuarial mathematics goes back to 1903, long before probability theory as we know it existed. He is credited by Cramér (1969, 1976) as initiating the theory of collective risk, in a series of papers in the late 1920s. Here, as in the work of Cramér below, one sees the modern formulation: the income stream of an insurance company, from premiums, is deterministic and linear; the outgoings, to meet claims, form a compound Poisson process, from the claims process (a Poisson process, of rate or intensity λ say) and the claim-size distribution (F say). Given the company's initial capital, u say, one studies the dependence of the probability of ruin (clearly positive) as a function of u and the current time, obtaining the familiar exponential estimate.

Lundberg may be regarded as having introduced the *Poisson process*, the foundation stone of actuarial mathematics. But one must bear in mind that the very term stochastic process is anachronistic here: the term was coined by Khinchin in the 1920s, and the necessary mathematical underpinning had to wait for Kolmogorov's *Grundbegriffe* of 1933.

Cramér (1969) draws attention to the implications of Lundberg's work for *reinsurance*. This field is of ever-growing importance, as the financial world becomes larger and more complicated, as it poses in modern form Juvenal's famous question (VII.1): *quis custodiet ipsos custodes*? Who guards the guards? Who insures the insurers? Who reinsures the reinsurers?

Harald Cramér

Harald Cramér (1893-1985) was a Swedish mathematician and probabilist of great distinction. In his personal recollections (Cramér, Half a century with probability, *Annals of Prob.*(1976)) he writes, of the period after he obtained his PhD (in 1917, in analytic number theory, under Marcel Riesz): "For a young Swedish mathematician of my generation, who wanted to find a job that would enable him to support a family, it was quite natural to turn to insurance. It was a tradition for Swedish insurance companies to employ highly qualified mathematicians as actuaries ..." (he continues to describe how his actuarial and insurance work led him into probability theory). It is by no means unusual for people to be drawn into a field for such reasons (Doob in probability in the US, and Bartlett and Cox in statistics in the UK, come to mind). In 1929 Cramér became the first holder of the chair in Actuarial Mathematics and Mathematical Statistics at the University of Stockholm – an important event in the development of actuarial mathematics in Scandinavia, and indeed more generally.

The Cramér estimate of ruin (above) of 1930 is perhaps Cramér's most prominent contribution to actuarial and insurance mathematics, and with it the now-standard *Cramér-Lundberg model* in insurance, as we will now call the model above.

§8. Complements

More general processes

The classical Cramér-Lundberg model above is the basic prototype in insurance mathematics, but it is by no means the only one, and is not general enough for all purposes.

1. Non-homogeneous Poisson processes.

These we have met before. Here the Poisson rate $\lambda(t)$ may vary with time. Matters become more complicated, but the theory may be carried through much as before.

2. Cox processes.

These were introduced by D. R. (Sir David) Cox (1924 -) in 1955, under the name *doubly stochastic Poisson process* or *mixed Poisson process*. Here the Poisson rate is *random*. This makes things more flexible and realistic, as well as more complicated.

Perhaps the most important case of a Cox process is where the rate has a *Gamma* distribution, when it is called a *Pólya process*. Recall that the Gamma distribution is the prototype of an error (or noise) distribution on the positive half-line, just as the Normal is on the line. For background here, see Generalised Linear Models (GLMs) in regression, in statistics.

3. Lévy processes.

The compound Poisson process models a situation where we can clearly identify the jumps. But what matters to the company is the flow of cash. For a large company, claims of small (or even ordinary) size may be so numerous as to be treated as 'small change'; it is the *large claims* that predominate, as these can be lethal. Allowing for this, it makes sense to generalise to *Lévy processes* (named after the great French probabilist Paul Lévy (1886 - 1971) for his pioneering work on them in the 1930s). These are stochastic processes with *stationary independent increments*. By the *Lévy-Khintchine formula* and the *Lévy-Itô decomposition*, they may be decomposed into three independent components: (i) a linear deterministic drift (trivial); (ii) a Brownian-motion component; (iii) a sum of jumps (any of these may be absent). The jumps case splits, into (a) only finitely many jumps in finite time (*finite ac-*

tivitiy, FA – the *compound Poisson* case above); (b) infinitely many jumps in finite time (*infinite activity, IA*). The theory can be extended to the Lévy case; for details, see e.g. [Kyp].

4. Gerber-Shiu theory.

This theory (Hans Gerber and Elias Shiu, 1997 and 1998) looks at the financial situation of a company *at ruin* or bankruptcy. This is an important matter!:

(i) The size of the cash reserve just before failure governs how much in the pound (dollar, euro, ...) the creditors will receive.

(ii) The *overshoot* – amount of the deficit which triggers failure – will be used by the liquidators, creditors, regulators etc. to determine whether or to what extent the company was negligent. This has important legal implications. Never forget that it is *illegal* under the Companies Act to trade while insolvent – or to enter into a transaction without the capacity to carry it through. A transaction needs two counter-parties, each willing to trade, and each able to do so. Each has to trust the other here, and inability to complete a deal is a breach of trust here. See e.g. [Kyp, Ch. 10].

Related problems and processes
Fluctuation theory.

The ruin problem above involves the infimum (minimum) over time of the cash balance of the company, or equivalently the supremum (maximum) of the liability. Thus for a process $X := \{X_t\}$, the *supremum* and *infimum* processes are relevant:

$$\overline{X}(t) := \sup\{X_s : s \in [0,t]\}, \qquad \underline{X}(t) := \inf\{X_s : s \in [0,t]\}.$$

Related to these is the *reflected* process, $\overline{X} - X$:

$$(\overline{X} - X)(t) := \overline{X}(t) - X(t) \geq 0.$$

The study of these and related functionals is called *fluctuation theory* – cf. the title of Kyprianou's book [Kyp], where ruin problems are indeed studied.

Queues and dams.

Other areas of Applied Probability involve such functionals and processes, which have to be non-negative, for example, *storage* processes [Kyp Ch. 4] (one cannot store a negative quantity of a commodity, etc.). The classical example here is a *dam*, whose reservoir may run dry – become empty – but which cannot store a negative amount of water. Here the input process of

water is often modelled by a Lévy process. Now consider a *queue* – for simplicity, a *single-server* queue. The server is initially idle (say). This *idle period* ends when the first customer arrives for service; the server then works non-stop to serve him, and continues in the same way with any customers who arrive during this *busy period*, and so on. The analogue of the content of the dam (or storage model) is the *workload* facing the server – alternatively, from the customer's point of view, the *virtual waiting time* – the time a customer arriving at time t would have to wait to begin service (a.s. no customer *does* arrive at any given t – but busy facilities such as restaurants, exhibitions etc. often post how long one *would* have to wait if one *did*).
Random walks.

The subject lurking in the background here is that of *random walks* – sums $S_n = X_1 + \cdots + X_n$ of iid random variables. These have an extensive and interesting *fluctuation theory*, developed in the 1950s by Spitzer, Baxter, Sparre Andersen and others.
Duality.

The link between ruin problems and queues etc. lies in *duality* for random walks. In brief, this involves reversing both time and space – looking at the steps of a random walk backwards in time and 'upside down' (see e.g. [Kyp, 3.2]). One can often then transfer from one of the above problem areas to another. This is the most efficient, attractive and modern way to handle the material. Before, results that are now handled easily by duality as above had to be discovered at least twice, and it to be noticed that the relevant distributions were the same. Duality enables one to proceed 'pathwise' – by looking at the random quantities themselves, not just their distributions.

In the ruin problem here, the Net Profit Condition means: to avoid ruin, 'more money in than out'. In queues, this corresponds to the *stability condition*: the server can handle work faster than it comes in (mean service time less than mean inter-arrival time – without this, the server is overwhelmed and the queue 'blows up'). For such a *stable* queue, the workload (or virtual waiting time) has a limit distribution as time increases – the queue settles down to a steady state. The most important case is the $M/G/1$ queue (M: Markov arrival process – Poisson; G: general service-time distribution; 1: single server). Here the limiting waiting-time distribution is given by the classical *Pollaczek-Khintchine formula*. This corresponds to our main result above, the *Cramér estimate of ruin*. For background and details, see [Kyp, 1.3.1, 1.3.2].

Stochastic calculus for jump processes

In Ch. V we developed stochastic (Itô) calculus based on Brownian motion, and applied it in Ch. VI to mathematical finance (Black-Scholes theory). It turns out that this calculus can be extended to the processes with jumps relevant here in Ch. VII on insurance, where the jumps represent the claims. This is technically easier (at least for the Poisson process), but actually came later. It was developed in the context of queueing theory, where the jumps represent customers arriving (or departing). For details, see e.g.
D. Applebaum, *Lévy processes and stochastic calculus*, 2nd ed., CUP, 2009
P. Brémaud, *Point processes and queues: martingale dynamics*, Springer, 1981.

Recall that the essence of Brownian-based stochastic calculus is captured in the simple equation

$$(dB_t)^2 = dt.$$

The essence of Poisson-based stochastic calculus is similarly captured in

$$(dN_t)^2 = dN_t.$$

For, the change dN_t in a Poisson process $N = (N_t)$ at time t is 0 or 1, and the above expresses that these are the only roots of $x^2 = x$, i.e. $x^2 - x = x(x-1) = 0$.

The context of Lévy processes in [App] is the simplest natural one containing both the Brownian and the Poisson/compound Poisson cases. But the natural context for stochastic integration is (a lot) more general still – that of *semi-martingales*. These are processes expressible as the sum of a local martingale and a process of (locally) finite variation (FV). The theory here was developed by Paul-André Meyer (1934-2003) and the French (Strasbourg, Paris) school – the 'general theory of processes'.

9. 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*. The areas of statistics involved are 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, but not one we can pursue further here.

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. VII, Insurance Mathematics

As noted in VII.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 VII.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-VI) and insurance (Ch. VII) 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