Notes on Risk Assessment, I: Basic concepts

Jonathan Rougier, University of Bristol

Version 1.2, compiled October 20, 2016.

These notes represent an applied statistican's overview of risk assessment, presented as non-technically as possible. Large-scale risk assessment is not something that humans do well. But one source of difficulty is mitigable, namely ambiguity of meaning. So I provide precise definitions for words that are much in use, and handy names for some of the recurring concepts. My intention is to present a 'controlled vocabulary' which will help people involved in risk assessment to communicate with one another. The controlled vocabulary is given in an index at the end of the document. I welcome comments and suggestions (see my email address at the top of this page).

Contents

1 The Risk Manager and her environment

To start with, here are some outline definitions. I will respect the conventional distinction (e.g., HSE, 2001, pp. 5–6) between:

Hazard The thing that has the potential to cause harm; and

Risk (Informal) The chance of harm, arising from the hazard.

Examples of UK hazards would be 'Icelandic volcanic eruptions', or 'flu-like pandemic'. A more prosaic hazard would be a falling tree. Somewhere in-between we have a flooding river. Risk needs to be precisely defined (see section 3), but that will require some preliminary spade-work; so, for the time being, accept that it has something to do with chance and harm. I will also introduce

Policy An action designed to reduce the risk of a hazard.

Putting these definitions together:

> *A 'Risk Manager' accepts a portfolio of hazards (maybe just a single hazard), and chooses between policies to minimise risk.*

As part of this process, the Risk Manager might need to rank hazard/policy combinations. In some applications, such as strategic planning over several hazards, the only policy for each hazard might be the 'do-nothing' policy.[1]

[1] Not literally do nothing, but 'do nothing in addition to what is currently being done'.

In the narrative of risk assessment, it is helpful to introduce some additional players (taken from Smith, 2010, ch. 1). The 'Client', who hires the Risk Manager to operate on her behalf, and the 'Auditor', whom the Client hires to check the Risk Manager's recommendations. Finally, there are the 'Experts', whom the Risk Manager consults; these might include domain experts, who specialise in the hazard, and statisticians, who specialise in the assessment of uncertainty and risk. These players are shown in Figure 1. I say more about Experts in section 5.

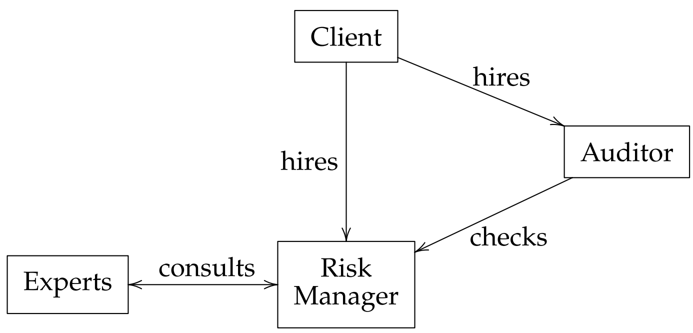

Figure 1: The players in risk assessment.

For example (not hypothetical), the Client might be an investment company, considering bidding for land with a licence to build a wind-farm. The Risk Manager will often be a large consultancy hired to assess the various options, such as different configurations of turbines. The Auditor might be a smaller consultancy called in to do 'due dilligence' on the resulting report. The Experts would be engineers, but also maybe statisticians, to forecast the wind at the site, and economists, to forecast the price of electricity. 'Harm' in this case is negative profit.

In my own thinking about risk assessment, I put a lot of emphasis on the Auditor. This is because large-risk management is almost invariably a 'wicked problem'.[2] One complication is that the Client is often an agent for a range of stakeholders with different and conflicting objectives (e.g. for national-scale risk assessment). Another is that the harm from a hazard is incurred in the future, but the costs of any action are incurred now. In the light of these difficulties, the successful implementation of a risk management policy requires some degree of shared ownership, so that those stakeholders who perceive themselves to be disadvantaged by the policy can nevertheless accept it. Thus the Risk Manager should expect her recommendation to be audited for *transparency* and *defensibility*, and should plan every aspect of the process accordingly. These notes outline a framework that, if conscientiously implemented, would satisfy me, were I the Auditor.

[2] See, e.g., Funtowicz and Ravetz (2006) and Conklin (2009).

2 *Harm versus loss*

In risk assessment, it is important to recognise that very few things that matter to humans can be quantified. In Statistics we try to maintain a careful distinction between a 'surrogate endpoint' and a 'true endpoint' (Cox and Donnelly, 2011, sec. 4.4). For example, cholesterol levels in the blood are a surrogate endpoint; quality of life is a true endpoint, which is affected by heart disease, for which cholesterol is a risk factor. Unlike quality of life, though, cholesterol is easily measured, and a treatment like a statin is initially assessed in terms of its ability to lower the amount of bad (LDL) cholesterol in the blood. Quality of life is also affected by side-effects, and for statins this has been a contentious issue.

It is the same with assessing 'harm' for the purposes of risk assessment and risk management. Quantitative risk assessment requires measurable quantities, which can only be surrogates for other quantities that really matter. And so I define

Loss A surrogate endpoint (quantifiable) for harm.

In the case of risk assessment, one hopes that the policy recommendation based on loss is so clear that it may be accepted for harm as well. Otherwise, Experts and the Risk Manager must together decide how to adjust the recommendation to close the gap between what is feasible to solve and what really matters. So, for example, a national-scale Risk Manager might report:

> According to what we can quantify, the recommended policy is to do A. But when we incorporate the negative effects of 'sense of outrage' the recommended policy is to do B.

In this decision it should be clear to everybody (and notably the Auditor) why initially A was chosen, and how incorporating 'sense of outrage' moves the recommendation away from A towards B. Precisely why B was chosen, and not some other policy on the trajectory away from A might be less clear.[3]

[3] I return to this topic at the end of section 6.

When choosing a measure for loss, it is very helpful if it is an 'extensive quantity', which means capable of being accumulated by adding (Cox and Donnelly, 2011, sec. 4.4). There is an obvious advantage to a loss which can be added: for a particular hazard event, the accumulated losses across different zones are simply the sum of the losses for each zone. This allows the loss assessment for a hazard event to be modularised. Likewise, when there is a non-negligible probability of more than one hazard event in a time-interval, the total loss from the hazard is the sum of the losses from each of the events. This notion of total loss will be developed in section 3. An example of a non-extensive quantity would be particulate concentration at ground level. This is a surrogate endpoint for health, but it is not suitable for adding across times or locations. To make it extensive, it could be converted into loss of QALYs (quality-adjusted life years).

Of the possible extensive measures of loss, monetary value is the most obvious contender. There are two reasons for this. First, most policies involve monetary costs in implementation; if loss is also measured in monetary terms then cost and loss can simply be added, which greatly simplifies the overall assessment. But there is also a deeper reason. Many policies are effective across several hazards. The benefit of such a policy, in terms of a reduction in risk, should therefore be accumulated across hazards. This is only possible if loss has a common and extensive measure. Monetary loss is common to almost all hazards. Since monetary value is extensive, it is the natural choice for 'the' surrogate for harm.[4]

[4] On this basis, I would not discount back to present value those 'monetary' losses which occur in the future, because they often represent harm, for which discounting might be inappropriate. Mind you, if increasing the discount rate to a little above zero made a different to the recommended policy, then one could hardly consider the recommendation to be robust.

It might be distasteful to assign a monetary value to a human life, or to a QALY, in order to assess all risks in terms of monetary value. We must always remember that this is a surrogate endpoint, chosen to provide the most effective *quantitative* assessment of the harm that arises under different policies. So we do not forget that there is a further stage in which the recommendations based on monetary loss are re-evaluated to account for the difference between loss and harm.

3 Risk and 'riskiness'

Throughout these notes I assume that probability is an appropriate measure for the quantifiable aspects of uncertainty. What this means, in effect, is that uncertainty is processed according to the rules of the probability calculus. There are many justifications for these rules, and statisticians, philosophers, and physicists sometimes disagree on this topic. However, there is little disagreement about the rules themselves. See Hacking (2001) for an insightful and non-technical introduction.

Fix a specified hazard and policy, and a specified time-interval in the future; this time-interval will be denoted $(a, b]$. During that time-interval, the number of hazard events which will occur is a random quantity N. Each hazard event has the capacity to create a loss, which will depend on magnitude, timing, and other factors as well. Represent the loss of the ith hazard event in the time-interval as the random quantity L_i, assumed to be extensive (see section 2). The 'total loss' in the time-interval is then

$$T = \sum_{i=1}^{N} L_i, \qquad N > 0, \tag{1}$$

and $T = 0$ otherwise. Technically, I should index both T and N by the time-interval, e.g. write $T(a, b]$ and $N(a, b]$. Practically, this would also be a good idea, to cement the notion that in risk assessment there is always a time-interval involved. However, this also seems a bit precious, and cluttered. Therefore I write simply T and N, suppressing *but not ignoring* the underlying time-interval $(a, b]$. See Figure 2.

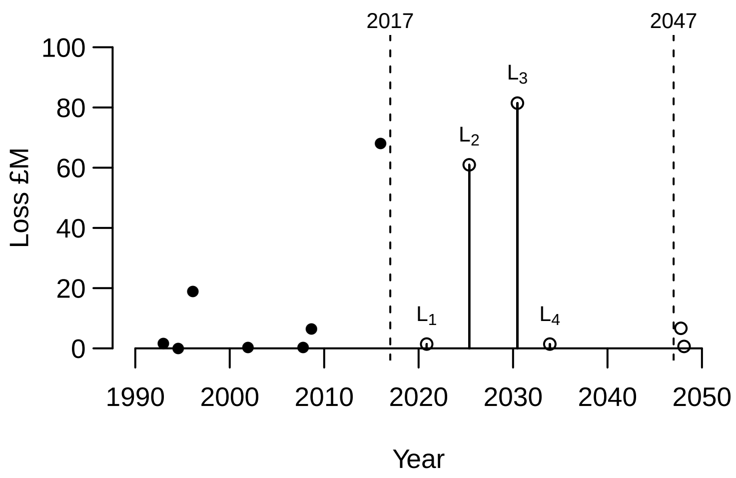

Figure 2: Hypothetical hazard/policy, 'actual' and simulated losses, and total loss for the 30-year time-interval from $a = 2017$ to $b = 2047$; $N = 4$ and $T = L_1 + L_2 + L_3 + L_4$.

Risk assessment for this combination of hazard, policy, and time-interval is synonymous with computing the probability distribution of T, represented in terms of the 'distribution function' F_T, where[5]

$$F_T(t) := \Pr(T \leq t), \qquad t \in \mathbb{R}. \tag{2}$$

The distribution function F_T does not exist 'in nature'; it is a human construct only partly linked, through our beliefs, to things that have been recorded about the past. This is why the specification of F_T is so challenging. For this section, I will assume that this task has been accomplished, and F_T exists as a function.[6] The 'survival function' of T is defined as

$$\bar{F}_T(t) := 1 - F_T(t) = \Pr(T > t), \qquad t \in \mathbb{R}, \tag{3}$$

and is also known as the 'probability of exceedance'. Now we can define the 'risk curve'.

[5] The symbol ':=' is read as 'is defined as', and 'Pr' as 'the probability of'; $\mathbb{R}$ is the set of all numbers. Thus in (2) the function F_T is defined for all numbers, and for each number t, $F_T(t)$ is the probability that the total loss is less than or equal to t.

[6] See section 4 and section 5 for two approaches to specifying F_T.

Definition 1 (Risk, risk curve)**.** The risk curve for a specified hazard and policy is the survival function of the total loss over a specified time-interval, denoted $\bar{F}_T$. Risk is synonymous with the risk curve.

So 'risk' is not a value, but a mathematical object, $\bar{F}_T$. Figure 3 shows an example of a risk curve. According to its definition, a risk curve must have the following properties:

1. If $t < 0$, then $\bar{F}_T(t) = 1$, because total loss is non-negative.
2. $\bar{F}_T$ is decreasing,[7] because if $t < t'$ then $T > t'$ implies that $T > t$.
3. There is a value $u > 0$ such that for all $t \geq u$, $\bar{F}_T(t) = 0$, because total loss cannot be infinite.

[7] In mathematics 'decreasing' typically means 'never increasing'. So $\bar{F}_T$ can have flat segments. If it had no flat segments it would be 'strictly decreasing'.

The value $1 - \bar{F}_T(0)$ is the probability of no hazard events in the time-interval $(a, b]$, and the value $\bar{F}_T(0)$ is the probability of at least one event in the time-interval.

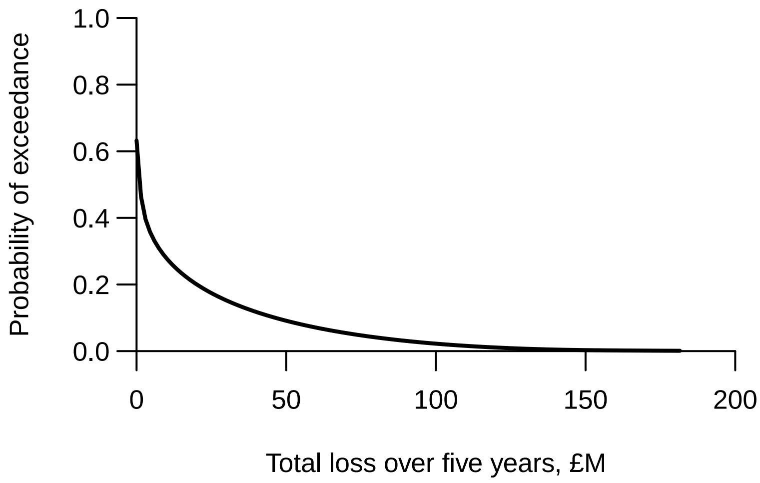

Figure 3: Hypothetical hazard/policy, risk curve for time-interval of 5 years. The probability of no hazard events in five years is 0.37.

The fact that risk is a function and not simply a number indicates that comparing risk across hazards, or within a hazard across different policies, is more complicated than just comparing the values of numbers. We do not expect to be able to order every combination of hazard/policy from best to worst, although, as will be discussed in section 6, we do expect to be able to 'partially order' hazard/policies.

In the meantime, though, no one can deny the usefulness of a single *summary* measure of risk, i.e. a property of the function $\bar{F}_T$ that can be represented by a single number, with the broad interpretation that larger values correspond to larger risks. For practical and mathematical reasons, I propose the following definition.

Definition 2 (Riskiness)**.** The 'riskiness' of a risk curve $\bar{F}_T$ is either one of these two equivalent properties:

A. The mathematical expectation of total loss, denoted $\mathbb{E}(T)$, or

B. The area under the risk curve (non-negative part).

It is a mathematical result (quite a sophisticated one, but see the end of this section) that these two are equivalent. The second property of Definition 2 is the intuititive one, and the one that allows us to assess riskiness from the risk curve by eye; see Figure 4. But the first property is the one that allows us to do useful calculations. The next section shows how this definition of riskiness sometimes decomposes in a way that will be familiar to almost everyone involved in risk assessment.

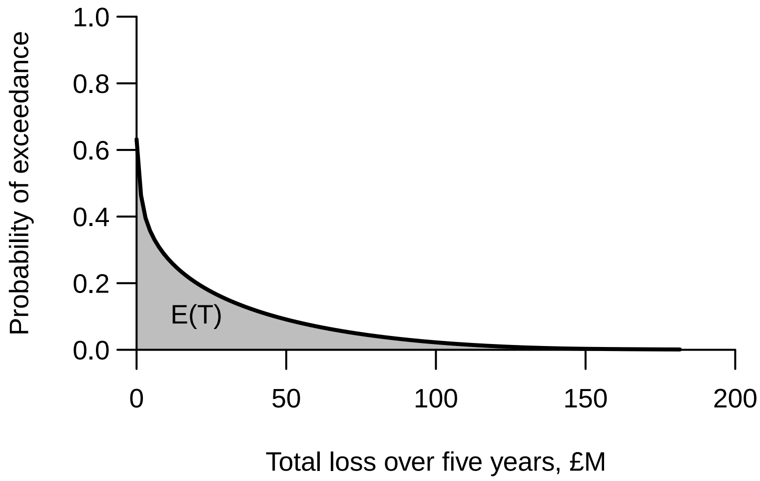

Figure 4: Same as Figure 3, but with the riskiness shaded in; see Definition 2. By eye, the riskiness is about £100M $\times$ 0.4/2 = £20M. The true value is £18.9M.

* * *

It is a mysterious result that the expected total loss is equal to the area under the risk curve. But a simple example can illustrate why it might be true. Suppose that, for a specified hazard/policy and specified time-interval, $\Pr(N = 0) = 0.3$, $\Pr(N = 1) = 0.7$, and $\Pr(N > 1) = 0$. Also suppose that if there is an event, then the loss L_1 is exactly £13M. So the expected total loss is

$$\mathbb{E}(T) = 0.3 \times \text{£0M} + 0.7 \times \text{£13M} = \text{£9.1M}.$$

The risk curve is shown in Figure 5, and it is obvious in this case that the risk curve is a step function, and the area under the risk curve is $0.7 \times$ £13M $=$ £9.1M.

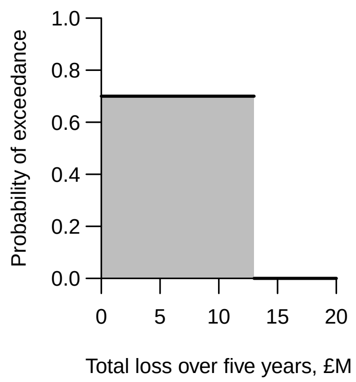

Figure 5: Risk curve for the special case given in the text, with the area under the curve shaded in.

4 *Rare events*

Recollect that N is the number of events in the time-interval $(a, b]$, for a specified hazard and policy. A hazard/policy with 'rare events' is one where $\Pr(N > 1) \approx 0$; note that this is always with respect to a specified time-interval. For concreteness, I define the 'rare event condition' as

$$\Pr(N > 1) < \frac{1}{20}. \tag{4}$$

Under the rare event condition,

$$\Pr(N = 0) = 1 - p, \quad \Pr(N = 1) \approx p, \quad \Pr(N > 1) \approx 0, \tag{5}$$

where $p := \Pr(N > 0)$. This value p will feature repeatedly below. I emphasise that the hazard must be specified, the policy must be specified, and the time-interval must be specified, before the rare event condition can be assessed.

Under the rare event condition, there is a large probability (equal to $1 - p$) there there will be no hazard event in $(a, b]$, and therefore no loss; there is a small probability (approximately p) that there will be one event; and there is a negligible probability that there will be more than one event. If there is an event, then its loss is a random quantity, and represented by the 'single-event loss' L_1, drawn from some distribution function F_L. If the Experts decide, *a priori*, that the rare event condition holds for a specified hazard/policy and time-interval, then the Experts just have to choose p and the distribution function F_L.

A very attractive feature of the rare event condition is that it simplifies computing (approximately) the risk curve and the riskiness, summarised by the following result.

Result 1 (Rare event approximation)**.** *Under the rare event condition in* (4)*, with p and F_L specified,*

1. $\bar{F}_T(t) \approx p \cdot \bar{F}_L(t)$ *for* $t \geq 0$*, and* 1 *otherwise, and*
2. $\mathbb{E}(T) \approx p \cdot \mathbb{E}(L_1)$*.*

Many readers will recognise that the second expression in Result 1 states

$$\text{riskiness} \approx \text{probability} \times \text{impact}$$

where we now have precise definitions for all three terms: 'riskiness' is expected total loss, 'probability' is the probability of at least one event in the specified time-interval, and 'impact' is the expected loss for a single event; finally, the approximation holds under the rare event condition in (4).

This decomposition of riskiness into 'probability × impact' is a highly successful meme. It is the basis of many heuristic treatments of risk assessment, and also of a visual tool, termed the *risk matrix*, discussed in section 6. It it is a decomposition of the summary measure 'riskiness', *not* of risk itself: remember that 'risk' is synonymous with the whole of the risk curve.

5 *The Lundberg assumptions*

We cannot simply ordain that the rare event condition in (4) holds: there are lots of situations where the probability of more than one event is non-negligible, for a specified hazard/policy and specified time-interval. At a national scale, for example, the probability of more than one UK large flood in five years is non-negligible under the do-nothing policy. In these risk assessments the Risk Manager and her Experts need to work harder in order to compute a distribution function for total loss, F_T, and a risk curve.

Or do they? In fact, there is a simple extension of the the rare event condition, which uses the same basic components, which were p and F_L (see section 4). Naturally, some simplifying assumptions are required. I will refer to them as the 'Lundberg assumptions', as they can be traced back to the Swedish actuary Filip Lundberg, writing at the beginning of the Twentieth Century.[8]

[8] See, e.g., Klüppelberg et al. (2014, chs 1,4).

Definition 3 (Lundberg assumptions)**.**

(i) There is a magnitude threshold $\nu \geq 0$ below which hazard events do not generate appreciable losses.

(ii) Hazard events above magnitude ν follow a homogeneous Poisson process with rate λ /yr. This is the 'homogeneity assumption'.

(iii) The losses for hazard events are independent and identically distributed (IID) with marginal distribution function F_L. This is the 'IID assumption'.

The Lundberg assumptions are able to reuse the single-event loss distribution F_L by treating the event losses as independent of their times of occurrence, and of each other (the IID assumption). Clearly this is a dubious assumption. If losses are incurred, then the capacity for another hazard event to generate losses is changed. This could go either way. For example, if a first earthquake destroys a building, then the building cannot be destroyed for a second time, so the loss from a second earthquake is zero. But if the building is rebuilt and then destroyed again, the second loss would typically be larger than than the first loss. But the new building may not be not be destroyed, if it has been rebuilt to improved building codes. The IID assumption is therefore a "don't really know which way it will go" assumption. It may or may not be applicable to a particular hazard/policy.

But even where the IID assumption is not applicable, it might still be useful.[9] Anyone who successfully designs or constructs something complicated—e.g. a machine, a piece of music or art, a computer programme, a building—will likely use the strategy of 'get something working' (GSW).[10] Risk assessment is very complicated, and an assumption like the Lundberg IID assumption is a GSW assumption for the Experts.

[9] At this point one is obliged to write: "All models are wrong, but some are useful" (George Box).

[10] In writing, this is known as a 'shitty first draft'.

The other major Lundberg assumption is the homogeneity assumption, number (ii).[11] Another dubious assumption. Some hazards are self-exciting, and some are self-quenching (they need to recharge). Some hazards are both, on different time-scales. For example, earthquakes come in swarms, i.e. they are locally self-exciting. But according to the predominant model of earthquakes, they occur in response to a build-up of stress, and thus, once they have occurred and the stress is discharged, it takes a while for the stress to build-up again. The same could be said for volcanoes,

[11] Assumption (i) is innocuous, because we could take $\nu = 0$; its purpose is to provide slightly greater generality to (ii).

where the magma chamber must recharge; or wildfires, where the vegetation must grow back.[12] Again, though, the homogeneity assumption is a GSW assumption.

[12] See Rougier et al. (2013, chs 8,11,12) for more details.

The purpose of the Lundberg assumptions is therefore to get the Experts up and running. Once they have chosen a single-event loss distribution function F_L and a large-event rate λ, the Lundberg assumptions give them a risk curve for any specified time-interval (see below, Figure 6). They can then spend the rest of their resources improving it, by thinking harder about F_L and λ, and then by relaxing the IID assumption and/or the homogeneity assumption. As it happens, these are both challenging assumptions to relax, from a statistical point of view. Which is why Experts would be well-advised to start with the Lundberg assumptions, to have something in the bank in case they run out of resources while trying to relax them.

* * *

Back to implementing the Lundberg assumptions. Recollect from section 4 that $p := \Pr(N > 0)$. Under the homogeneity assumption,

$$p = 1 - e^{-\lambda \cdot (b-a)}, \tag{6}$$

from the Poisson distribution. This expression can be inverted to convert from p to λ, if p is specified. From this point of view, the Lundberg assumptions are a direct extension of the rare event model, because they simply 'repurpose' p and F_L. Alternatively, λ can be specified directly, as the rate of a homogeneous Poisson process for large-magnitude events, from which p could be inferred, if required. If λ is specified directly, then a quick calculation shows that the rare event condition in (4) holds when $\lambda \cdot (b-a) < 0.36$. So if $\lambda = 0.2$ then the rare event condition would hold for one year, but not for two years.

There is an algorithm, 'Panjer recursion', for converting F_L and λ into F_T under the Lundberg assumptions; see `https://en.wikipedia.org/wiki/Panjer_recursion`.[13] Figure 6 shows risk curves for different time-intervals computed using the Lundberg assumptions. Naturally, as the time-interval extends, the risk curve rises.

[13] Or Ross (1996, sec. 2.5) for the mathematics.

One thing that does not survive the transition from the rare event condition to the Lundberg assumptions is the riskiness approximation in Result 1. The Lundberg version for the time-interval $(a, b]$ is

$$\mathbb{E}(T) = \lambda \cdot (b-a) \cdot \mathbb{E}(L_1) \tag{7}$$

(an exact result). The rare event condition is the special case when $p \approx \lambda \cdot (b-a)$. So in general, riskiness does not decompose as 'probability $\times$ impact'. Given the *de facto* ubiquity of this decomposition, this is worth stressing:

> *The decomposition 'riskiness $\approx$ probability $\times$ impact' only holds in the case where the probability of more than one hazard event in the specified time-interval is negligible.*

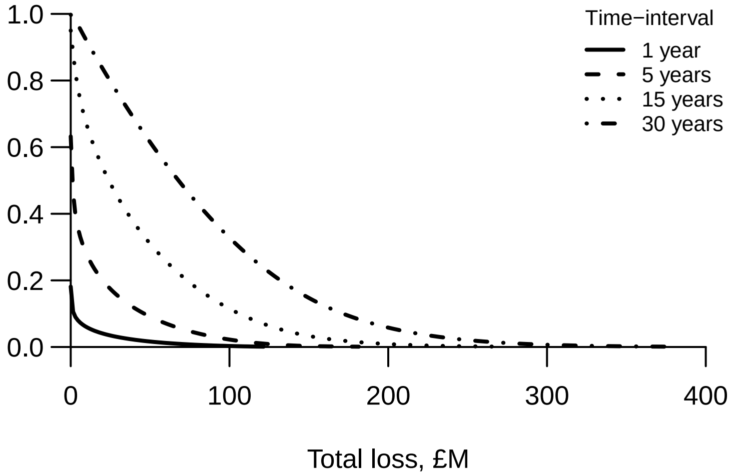

Figure 6: Hypothetical hazard/policy, risk curve for different time-intervals computed using the Lundberg assumptions with the same λ and F_L, using Panjer recursion.

6 *Comparing risk curves, risk matrices*

The Risk Manager will often want to compare risk curves, possibly across hazards, e.g. to rank them, or possibly across policies for the same hazard, e.g. to choose a particular policy. As a framing device, we assume that Clients are risk averse, which allows us to say that a Client 'prefers one risk curve to another' rather than the clunky 'finds one risk curve less risky than another'. Each Risk Manager needs to have an understanding of her Client's preferences, in order to prepare an appropriate ranking, or to make an informed choice. There is a lot of theory about this, the theory of 'utility functions' (see, e.g., Smith, 2010, ch. 3), but I regard this theory as a bit abstruse for actual use. So instead, let's try to make progress with some relatively simple ideas.

Suppose, to start with, that the Risk Manager decided to summarise each risk curve in terms of a single value, its riskiness (see Definition 2). Could this be effective? Definitely not. Most Clients tend to be sensitive to large losses.[14] The righthand tail of the risk curve is not well-constrained by the riskiness. So another possibility is to replace the riskiness with a high quantile of the total loss distribution, like the 95th percentile.[15] But is it defensible to neglect the more probable outcomes and only look at the improbable ones?

[14] See the end of the section for my explanation.

[15] The 95th percentile of total loss is the value t satisfying $F_T(t) = 0.95$.

Figure 7 illustrates these two summaries. In the first, A and B both have the same expected total loss (riskiness), but I would expect most Clients to find A more risky than B because of its much longer righthand tail. In the second, A and C both have the same 95th percentile, but I would expect most Clients to find A less risky than B because it has a much lower expected total loss. In fact I am confident about the following assertion:

> *Given any single summary of the risk curve, I can construct two risk curves where the value of that summary is the same, but where a typical Client will strongly prefer one risk curve to the other.*

What this means is that when comparing risk curves, the Risk Manager will need to use two or more summary values, in order to

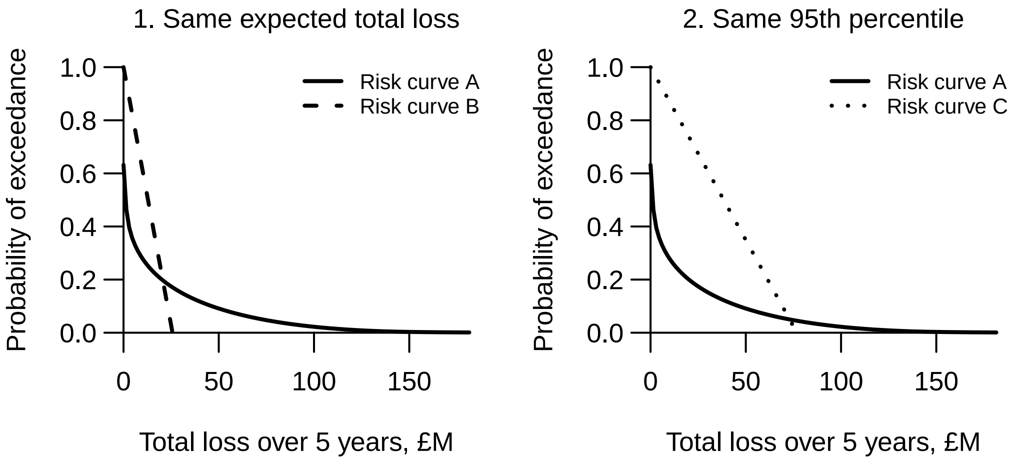

Figure 7: Two examples of risk curves that agree on one criterion but disagree on another. In the first, risk curve *A* has a much larger 95th percentile than *B*, while in the second, risk curve *C* has a much larger expected total loss (riskiness) than *A*.

represent her Client's preferences. But once each risk curve is summarised by more than one value, we lose the ability to completely order all of the risk curves. Instead, we have what is termed in mathematics a 'partial order'. Suppose that there are two summary measures for each risk curve, and in both cases larger values are less preferred. If risk curve *A* is smaller than risk curve *B* in both values, then we can say that *A* is preferred to *B*. But if risk curve *A* is smaller in one value and larger in the other, then the situation is ambiguous.

There is a powerful visual tool for capturing this idea of a partial order based on two summary values, termed a 'risk matrix'. A risk matrix is a graph where the horizontal axis (x-axis) represents one summary value, the vertical axis (y-axis) the other, and each risk is represented by a symbol: see Figures 8 and 9. If the symbol for risk curve *A* is to the southwest of risk curve *B*, then *A* is preferred to *B*, being smaller in both values. So a risk matrix is actually a visual tool for representing a partial ordering. We can see at a glance whether *A* is preferred to *B*, or *B* to *A*, or whether it is ambiguous; similarly for every other risk.

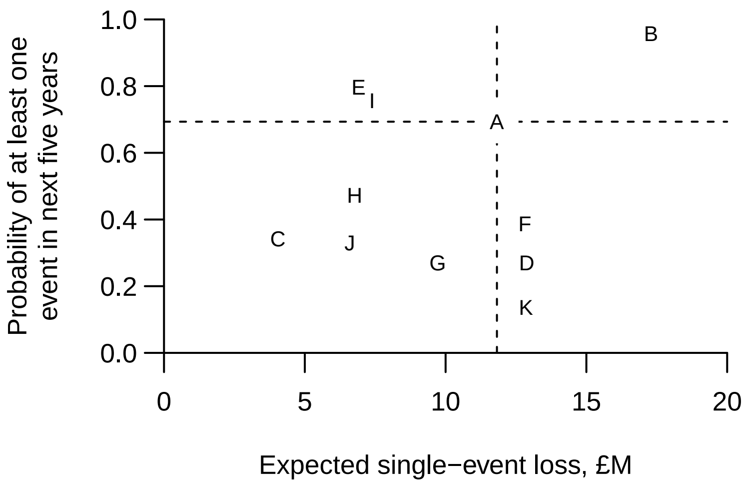

Figure 8: 'Vanilla' risk matrix, see Definition 4. According to the partial ordering induced by these two statistics, *A* is less risky than *B*, but more risky than *C*, *G*, *H*, and *J*. It is ambiguous whether *A* is less or more risky than *D*, *E*, *F*, *I*, and *K*. The dashed lines are just to highlight *A*.

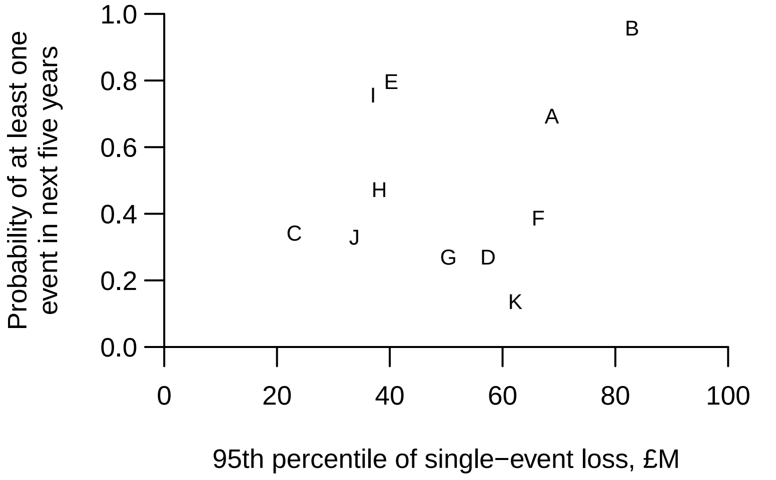

Figure 9: 'RWC' risk matrix, see Definition 4. Based on the same risk curves as in Figure 8. In this partial ordering, A is less risky than B, more risky than C, D, F, G, H, J, and K, and it is ambiguous whether A is more or less risky than E and I.

I am going to define two types of risk matrix, according to the two summary values used.

Definition 4 (Types of risk matrix)**.**

Vanilla x-axis is $\mathbb{E}(L_1)$, the expected loss from the first event, and y-axis is p, the probability of at least one event in the time-interval (see section 4).

Reasonable Worst Case (RWC) x-axis is the 95th percentile of L_1, and y-axis is p, as for Vanilla.

In both types of risk matrix, these two values correspond to 'impact' (x-axis) and 'probability' (y-axis), as discussed in section 4. The vanilla risk matrix has a special property. Under the Lundberg assumptions,[16] the riskiness of each risk curve is exactly determined by the x and y values. Therefore riskiness in a vanilla risk matrix can be represented using contour lines; see Figure 10. More generally, riskiness, or any other third summary value, could be encoded in the colour of each symbol, but perhaps this is too much information for a visual tool.

[16] Including the rare event condition as a special case.

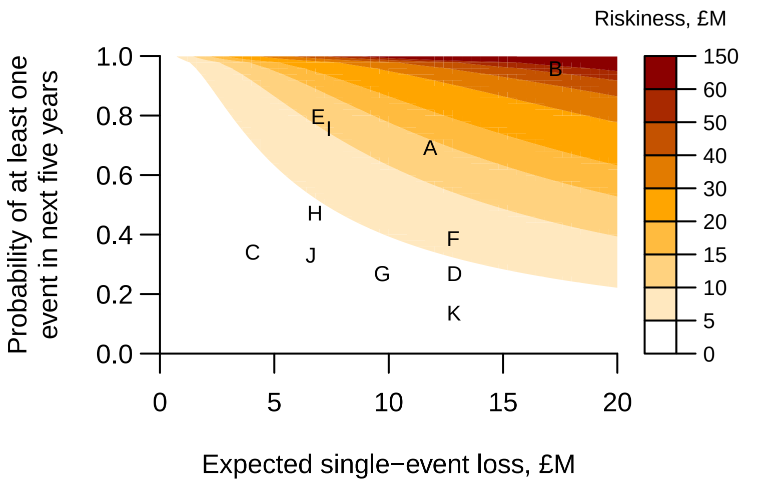

Figure 10: 'Vanilla' risk matrix, with contours showing riskiness.

There are two further points to be made about risk matrices. First, a risk matrix represents a 2D summary of each risk curve, and the partial order it displays might not reflect all of a Client's preferences. The Risk Manager must understand that the Client will struggle to articulate his preferences, and she (the Risk Manager) should be careful not to railroad the Client into the wrong set of preferences through only presenting him with a restricted set of summaries. It is a fortuitous outcome if the Client's preferences can be represented by some 2D summary and visualised in a risk matrix, but this needs to be discovered by the Risk Manager through a careful elicitation exercise.

Second, a risk matrix only makes sense if all risk curves are summarised in exactly the same way. This is why I have given careful definitions of the axes in Definition 4. In many cases different risk curves are assessed by different groups of people (e.g. for different hazards). Ambiguity about the definitions of the two axes leads to incomparable positioning of the symbols in the risk matrix, which severely degrades its usefulness. The safe way to proceed is to ensure that (i) each group assesses the risk curve for its hazard(s), and (ii) the same summaries are then applied to each risk curve to locate the symbols. This procedure will ensure that all of the symbols in the risk matrix are comparable. It also means that the Risk Manager has the whole of the risk curve, if she needs it, which she may well do in order to reflect the Client's preferences.

One comment on current practice. Risk matrices are widely used. In practice, precise definitions of the two axes are seldom given, and the cells of the risk matrix are given fanciful colours, sometimes labelled as, say, 'low', 'medium', 'high', 'critical', from bottom-left to top-right. A risk matrix is only useful if the time-interval is defined, both axes are clearly defined (ideally as summaries of a risk curve), and there are no fanciful and possibly misleading colour schemes and labels. These are basic things that the Auditor should check. Were I the Auditor, I would be very happy to see a risk matrix such as Figure 10. Poor examples are easily found online, e.g. by googling "image risk matrix".

* * *

Why are Clients so often sensitive to the righthand tail of the risk curve? For some Clients, e.g. insurance companies, this is because a loss over some high threshold triggers an irreversible event, like insolvency. For other Clients, my suggestion is that large losses have the potential to create disproportionately large harms; see Figure 11.

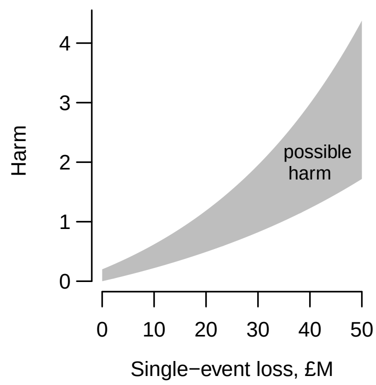

Figure 11: Large losses have the potential to create disproportionately large harms. The units of harm are notional.

It is much harder for the Experts to foresee all of the harm that arises from an event with a very large loss, due to a lack of precedents and a failure of imagination. As discussed in section 2, it is the harm that the Client really cares about, not the loss. If a risk curve has a long righthand tail, summarised for example by a large 95th percentile, then it has more capacity to do harm. In the interests of transparency, I would prefer the initial assessment across hazard/policies to be in terms of loss. But I can imagine a Risk Manager reporting

> In terms of loss, risk curve *A* is slightly preferred to risk curve *B*. But because risk curve *A* has a much longer tail than *B*, overall *B* is preferred to *A*, because *A* has the capacity to produce much more harm than *B*.

Further reading

Anyone interested in how humans think, and in particular how they think about complicated things like uncertainty and risk, should read Kahneman (2011). Tetlock and Gardner (2015) provide lots of practical advice about forecasting. Woo (2011) provides a quirky and engaging tour of catastrophes. And, for how it all goes wrong, read *The Big Short* by Michael Lewis (Penguin, 2011), or watch the film (2015).

References

Conklin, J. (2009). Building shared understanding of wicked problems. *Rotman Magazine*. Available at http://www.cognexus.org/Rotman-interview_SharedUnderstanding.pdf. 3

Cox, D. and Donnelly, C. (2011). *Principles of Applied Statistics*. Cambridge University Press, Cambridge, UK. 3, 4

Funtowicz, S. and Ravetz, J. (2006). Post-Normal Science: Environmental policy under conditions of complexity. Available at http://www.nusap.net/sections.php?op=viewarticle&artid=13. 3

Hacking, I. (2001). *An Introduction to Probability and Inductive Logic*. Cambridge University Press, Cambridge, UK. 4

HSE (2001). Reducing risks, protecting people. Technical report, Health and Safety Executive. Available online, http://www.hse.gov.uk/risk/theory/r2p2.pdf. 2

Kahneman, D. (2011). *Thinking, Fast and Slow*. Penguin Books Ltd, London, UK. 15

Klüppelberg, C., Straub, D., and Welpe, I., editors (2014). *Risk: A Multidisciplinary Introduction*. Springer, Heidelberg, Germany. 9

Ross, S. (1996). *Stochastic Processes*. John Wiley & Sons, Inc., New York, USA, second edition. 10

Rougier, J., Sparks, R., and Hill, L., editors (2013). *Risk and Uncertainty Assessment for Natural Hazards*. Cambridge University Press, Cambridge, UK. 10

Smith, J. (2010). *Bayesian Decision Analysis: Principle and Practice*. Cambridge University Press, Cambridge, UK. 2, 11

Tetlock, P. and Gardner, D. (2015). *Superforecasting: The Art & Science of Prediction*. Random House Books, London, UK. 15

Woo, G. (2011). *Calculating Catastrophe*. Imperial College Press, London, UK. 15

Controlled vocabulary