Quantifying confidence in probability assessments

Jonty Rougier

School of Mathematics
University of Bristol

ADMLC Meeting, ESA, May 2017

A modest proposal

A modest proposal

- Risk assessments often close with the question:

How confident are you in your assessment?

Not confident at all [1] [2] [3] [4] [5] Very confident

I think it would be inadvisable to base an action on the answer to this question.

A modest proposal

- Risk assessments often close with the question:

How confident are you in your assessment?

Not confident at all 1 2 3 4 5 Very confident

I think it would be inadvisable to base an action on the answer to this question.

- My modest proposal is to replaced this by:

How confident are you in your assessment?

Unlikely to change much over the next five years 1 2 3 4 5 Likely to change a lot over the next five years

A modest proposal (cont)

This proposal, which is not original, has several advantages:

1. It operationalizes the notion of 'confidence', which makes assessment easier for the experts, and permits retrospective evaluation.

A modest proposal (cont)

This proposal, which is not original, has several advantages:

1. It operationalizes the notion of 'confidence', which makes assessment easier for the experts, and permits retrospective evaluation.

2. It aligns more closely with the needs of policymakers, for whom the pressing question is often "Do we act now, or do we delay for another cycle?"

A modest proposal (cont)

This proposal, which is not original, has several advantages:

1. It operationalizes the notion of 'confidence', which makes assessment easier for the experts, and permits retrospective evaluation.

2. It aligns more closely with the needs of policymakers, for whom the pressing question is often "Do we act now, or do we delay for another cycle?"

3. It can be quantified using the experts' assessment of the relevance of the historical record. Notably the length of the relevant record compared to the prospective period.

‘Likelihood’ (UK NRA definition)

The ‘likelihood’ $\mathcal{L}$ of a hazard class is the probability of at least one major event happening in the next five years. Or, if N is the number of major events, $\mathcal{L} = \Pr\{N(0, 5] > 0\}$.

‘Likelihood’ (UK NRA definition)

The ‘likelihood’ $\mathcal{L}$ of a hazard class is the probability of at least one major event happening in the next five years. Or, if N is the number of major events, $\mathcal{L} = \Pr\{N(0,5] > 0\}$.

1. If h are the records from the relevant historical period, then the likelihood is currently

$$\mathcal{L}(h) = \Pr\{N(0,5] > 0 \mid h\}.$$

‘Likelihood’ (UK NRA definition)

The ‘likelihood’ $\mathcal{L}$ of a hazard class is the probability of at least one major event happening in the next five years. Or, if N is the number of major events, $\mathcal{L} = \Pr\{N(0,5] > 0\}$.

1. If h are the records from the relevant historical period, then the likelihood is currently

$$\mathcal{L}(h) = \Pr\{N(0,5] > 0 \mid h\}.$$

2. k years into the future, when we will have additional records f, the likelihood will be computed as:

$$\mathcal{L}(h, f; k) = \Pr\{N(k, k+5] > 0 \mid h, f).$$

'Likelihood' (UK NRA definition)

The 'likelihood' $\mathcal{L}$ of a hazard class is the probability of at least one major event happening in the next five years. Or, if N is the number of major events, $\mathcal{L} = \Pr\{N(0,5] > 0\}$.

1. If h are the records from the relevant historical period, then the likelihood is currently

$$\mathcal{L}(h) = \Pr\{N(0,5] > 0 \mid h\}.$$

2. k years into the future, when we will have additional records f, the likelihood will be computed as:

$$\mathcal{L}(h, f; k) = \Pr\{N(k, k+5] > 0 \mid h, f).$$

3. So currently, the future likelihood is a random quantity $\mathcal{L}(h, F; k)$, and we can compute its distribution function

$$F_{\mathcal{L}}(\ell \mid h; k) := \Pr\{\mathcal{L}(h, F; k) \leq \ell \mid h\}.$$

‘Likelihood’ (UK NRA definition) (cont)

4. Confidence (proposed definition) can be approximated by the 5th and 95th percentiles of $F_{\mathcal{L}}$. Call this a 90% k-year prospective interval.

‘Likelihood’ (UK NRA definition) (cont)

4. Confidence (proposed definition) can be approximated by the 5th and 95th percentiles of $F_{\mathcal{L}}$. Call this a 90% k-year prospective interval.

5. Informally, a hard likelihood is a likelihood with a small prospective interval, while a soft likelihood is one with a large prospective interval. E.g., “The likelihood is a hard 0.15.”

'Likelihood' (UK NRA definition) (cont)

4. Confidence (proposed definition) can be approximated by the 5th and 95th percentiles of $F_{\mathcal{L}}$. Call this a 90% k-year prospective interval.

5. Informally, a hard likelihood is a likelihood with a small prospective interval, while a soft likelihood is one with a large prospective interval. E.g., "The likelihood is a hard 0.15."

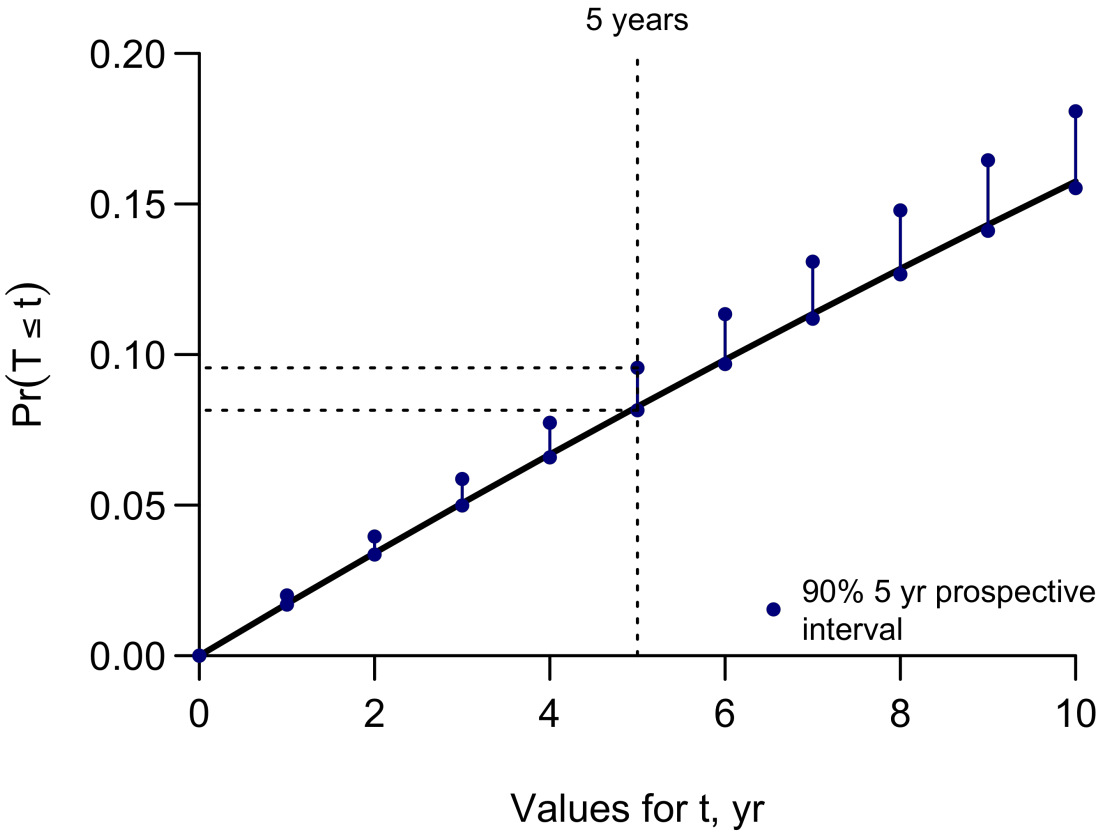
The proposed definition is generic, but the calculation becomes nearly trivial under the model that large events follow a Poisson process. In this case, in the simplest treatment, the experts need only decide how far back to go while not violating homogeneity.

Illustration: large explosive Icelandic eruptions

There have been 5 recorded $M5+$ eruptions since 1700 CE.

Illustration: large explosive Icelandic eruptions

There have been 5 recorded $M5+$ eruptions since 1700 CE.



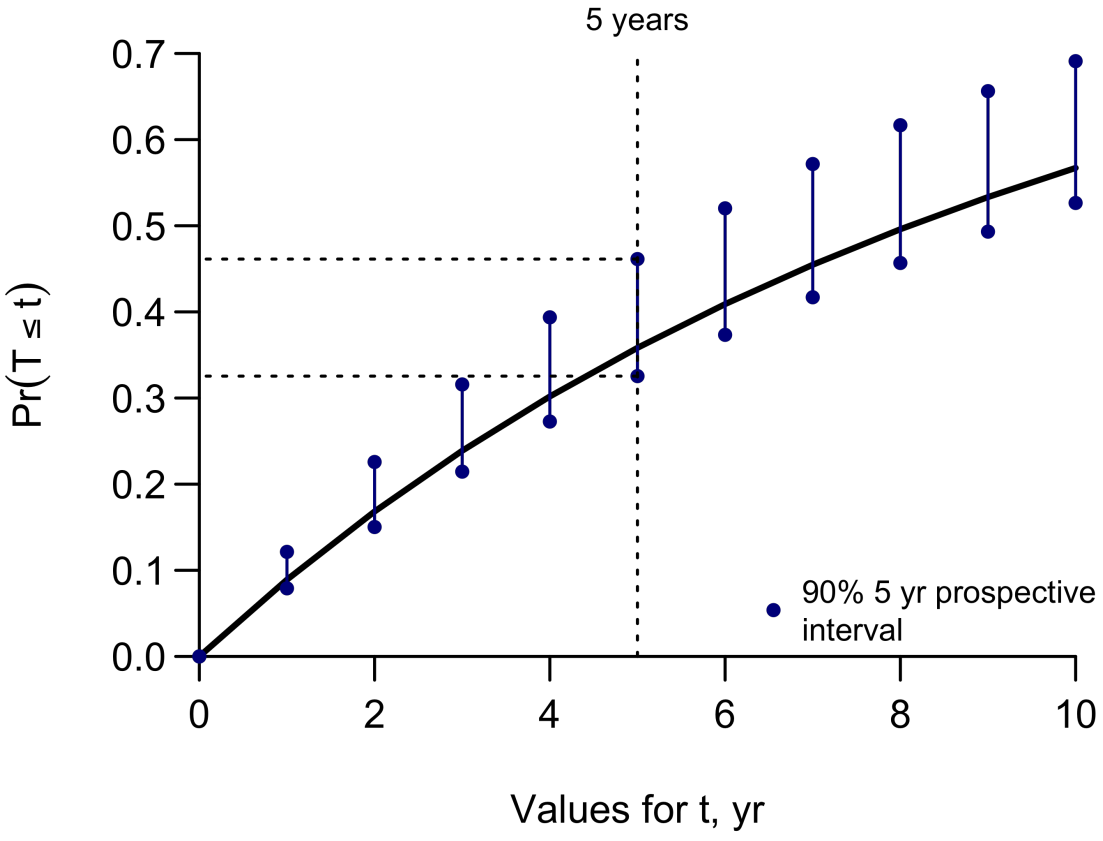
That looks like a firm-ish 0.083 (90% 5-year PI: 0.081, 0.096).

Illustration: UK oil refinery/facility fires

Three large fires since 1980 (Grangemouth, 1987; Buncefield, 2005; Pembroke, 2011).

Illustration: UK oil refinery/facility fires

Three large fires since 1980 (Grangemouth, 1987; Buncefield, 2005; Pembroke, 2011).



Soft-ish 0.36 (90% 5-year PI: 0.33, 0.46). Illustration only!!

Appraisal

1. Operationalizing 'confidence' is a good idea.

2. Using the quantiles of $F_{\mathcal{L}}$ seems to be a natural approach.

3. Stationary process modelling reduces the assessment process to two values: the length of the relevant historical record, and the number of events during that time.

4. This can adapt to non-stationary processes simply by reducing the length of the relevant past, but of course the intervals are wide.