Formalizing Life Table in Isabelle/HOL

Yosuke ITO

Sompo Himawari Life Insurance Inc.

The 19th Theorem Proving and Provers meeting (TPP 2023)
October 31st, 2023

Disclaimer

1. The contents presented here are solely the speaker's opinions and do not reflect the views of Company.
2. There are some inaccuracies in explaining actuarial mathematics due to the priority on intuitive understanding.
3. This presentation is a progress report on the library "Actuarial Mathematics" in the Archive of Formal Proofs of Isabelle. The future version might be different from the one presented below.

Self Introduction

Professional Experience

- Sompo Himawari Life Insurance Inc., December 2020 – Present.
  - Aggregates the business results of life insurance products.
- Meiji Yasuda Life Insurance Company, April 2014 – November 2020.
  - Revised the reinsurance contracts.
  - Determined the prices of life insurance products.
  - Attended the approval negotiations with Financial Services Agency.
  - Qualified as an actuary (Fellow of the Institute of Actuaries of Japan).
  - Aggregated the business results of group life insurance.
  - Calculated retirement benefit obligations of client enterprises.
  - Validated the financial soundness of Employees' Pension Plans.

Education

- Nagoya University
  - Master of Mathematical Sciences, March 2014.
- The University of Tokyo
  - Bachelor of Science, Mathematics Course, March 2012.

Contents

Contents

What Is Life Table? I

1. A life table, a.k.a. mortality table, “shows the rate of deaths occurring in a defined population during a selected time interval, or survival rates from birth to death” [2].
2. A life table depends on its statistical population; we choose an appropriate life table on a case-by-case basis.
3. For example, the Ministry of Health, Labour and Welfare produces Life Tables of the people in Japan.
   - The 23rd Complete Life Table (2020) `https://www.mhlw.go.jp/toukei/saikin/hw/life/23th/`

What Is Life Table? II

1. In life tables, the following symbols are generally used:
   - l_x, the number of people alive at age x;
   - ${}_nd_x$, the number of deaths between ages x and $x+n$; $d_x := {}_1d_x$;
   - ${}_np_x$, the probability that a person aged x lives for another n years;
   - ${}_nq_x$, the probability that a person aged x dies within n years; $q_x = {}_1q_x$;
   - μ_x, the force (instantaneous rate) of mortality at age x;
   - T_x, the person-years lived above age x;
   - ${}_nL_x$, the person-years lived for n years from age x; $L_x := {}_1L_x$;
   - $\mathring{e}_x$, the complete expectation of life at age x; $\mathring{e}_0$ is the average longevity.
2. These symbols and relevant formulas are important in actuarial mathematics, e.g.

$$
{}_np_x = l_{x+n}/l_x, \qquad {}_nq_x = {}_nd_x/l_x,
$$

$$
\mu_x = -\frac{1}{l_x}\frac{dl_x}{dx}, \qquad T_x = \int_0^\infty l_{x+t}\,dt,
$$

$$
\mathring{e}_x = \int_0^\infty {}_tp_x\,dt = \frac{1}{l_x}\int_0^\infty t l_{x+t}\mu_{x+t}\,dt.
$$

Contribution

I updated the entry “Actuarial Mathematics” in the AFP (Archive of Formal Proofs) compatible with Isabelle2023.
`https://www.isa-afp.org/entries/Actuarial_Mathematics.html`

1. Formalized the theory of the life table in Isabelle/HOL based on the measure-theoretic probability theory.
2. Showed some practical applications: verifying examination problems in life insurance mathematics.
3. Newly formalized missing important theorems in analysis and probability theory.

The formalization is mainly based on the online document “A Reading of the Theory of Life Contingency” by Marcel B. Finan [1], but not the same.

Contents

What Is Actuarial Mathematics?

1. Actuarial mathematics is a branch of applied mathematics, which is used to evaluate financial risks of undesirable events.
2. It is related to
   - calculus,
   - probability theory,
   - statistics,
   - financial theory.
3. The traditional actuarial roles are considered as
   - determining the prices of insurance products,
   - estimating the liability of a company associating with the insurance contracts.
4. Recently, the risk management skill of actuaries is required in a wider range of businesses.

Pricing a Term Life Insurance I

"Term life insurance provides a death benefit that pays the beneficiaries of the policyholder throughout a specified period of time" [2].

Assumption

1. amount insured: \$10000
2. entry age: 30 years old
3. policy period: 1 year
4. annual mortality rate: 1%
5. annual interest rate: 2%

The expected payment after 1 year is $\$10000 \times 1\% = \$100$.
If the insurance company earns a 2% investment yield annually, the required amount for this insurance can be discounted:

$$\frac{\$100}{1 + 2\%} \approx \$98.$$

Pricing a Term Life Insurance II

Definition

The present value of a term life insurance on a person aged x payable at the end of the year of death within n years is written as $A^1_{x:\overline{n}|}$ per unit insurance amount:

$$A^1_{x:\overline{n}|} := \sum_{k=1}^{n} \frac{{}_{k-1}p_x \cdot q_{x+k-1}}{(1+i)^k},$$

where i denotes the annual interest rate.

In the example above, $A^1_{30:\overline{1}|} \approx 0.0098$.

International Actuarial Notation I

There are various actuarial symbols used worldwide since at least 20th century [3, Perryman 1949].

a_x = an annuity, first payment at the end of a year, to continue during the life of (x).

$\ddot{a}_x = 1 + a_x$ = an 'annuity-due' to continue during the life of (x), the first payment to be made at once.

A_x = an assurance payable at the end of the year of death of (x).

Note. $e_x = a_x$ at rate of interest $i = 0$.

A letter or number at the lower left corner of the principal symbol denotes the number of years involved in the probability or benefit in question. Thus:

${}_np_x$ = the probability that (x) will live n years.

${}_nq_x$ = the probability that (x) will die within n years.

Note. When $n = 1$ it is customary to omit it, as shown on page 2, provided no ambiguity is introduced.

${}_nE_x = v^n {}_np_x$ = the value of an endowment on (x) payable at the end of n years if (x) be then alive.

If the letter or number comes before a perpendicular bar it shows that a period of deferment is meant. Thus:

${}_{n|}q_x$ = the probability that (x) will die in a year, deferred n years; that is, that he will die in the $(n+1)$th year.

${}_{n|}a_x$ = an annuity on (x) deferred n years; that is, that the first payment is to be made at the end of $(n+1)$ years.

${}_{n|t}a_x$ = an intercepted or deferred temporary annuity on (x) deferred n years and, after that, to run for t years.

International Actuarial Notation II

1. In life insurance mathematics, the relations between the actuarial symbols are well examined, e.g.

$$A_{x:\overline{n}|} = 1 - d\ddot{a}_{x:\overline{n}|},$$

where

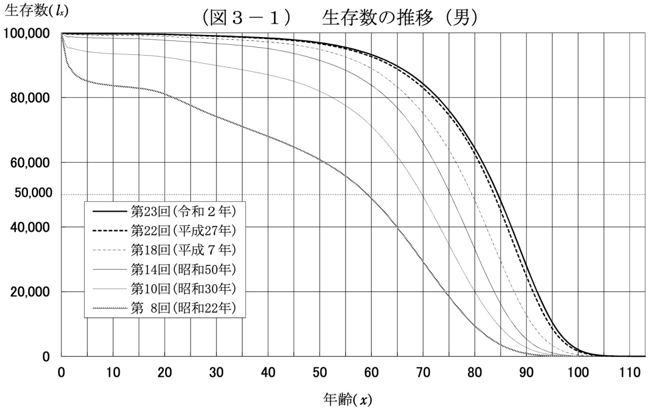
$$\begin{aligned}
v &:= \frac{1}{1+i} \quad \text{(the present value of 1 to be paid 1 year from now)}, \\
d &:= 1 - v \quad \text{(the annual effective discount rate)}, \\
\ddot{a}_{x:\overline{n}|} &:= \sum_{k=0}^{n-1} {}_{k}p_x v^k \quad \text{(the present value of a life annuity due)}, \\
A_{x:\overline{n}|} &:= A^{1}_{x:\overline{n}|} + {}_{n}p_x v^n \quad \text{(the present value of an endowment)}.
\end{aligned}$$

2. Actuaries use these symbols and formulas efficiently to calculate prices of products, reserves of the company, etc.

Contents

Properties of Life Table

Graph of l_x of the 23rd Male Complete Life Table (continuous line) [4]:



1. $\lim_{x\to\infty} l_x = 0$.
2. l_x is non-increasing.

Life Table to Tail Distribution

Definition (Life Table)

A life table is a function $l : \mathbb{R} \to \mathbb{R}$ which satisfies all the following properties:

1. $l_0 > 0$,
2. $l_x = l_0$ for all $x < 0$,
3. $\lim_{x \to \infty} l_x = 0$,
4. l is non-increasing,
5. l is right-continuous.

$\overline{F}(x) := l_x / l_0$ becomes the tail distribution of a random variable X:

$$\overline{F}(x) = P(X > x).$$

X is nothing other than *the age at death*. In survival analysis, the tail distribution is also called the *survival function*.

Survival Model

1. When X is a random variable representing the age at death, actuarial statistics can be written as follows:

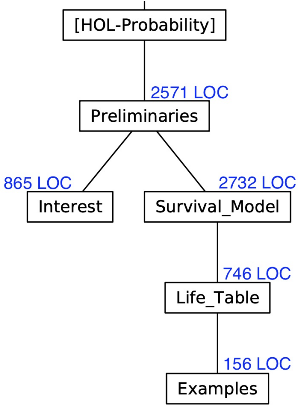
$$
{}_n p_x = P(X > x + n \mid X > x), \quad {}_n q_x = P(x < X \leq x + n \mid X > x),
$$
$$
\mu_x = \lim_{h \to +0} \frac{P(x < X \leq x + h \mid X > x)}{h}, \quad \mathring{e}_0 = E[X].
$$

2. Therefore, the life table can be formalized in the following order.

Survival Model. Introduce the age-at-death random variable X, and define actuarial statistics in the context of probability theory. Many important formulas are proved without a life table.

Life Table. Define a life table axiomatically, and reduce it to the survival model. Formulas including the life table are derived by easy translation from the survival model.

Overview of the AFP Entry "Actuarial Mathematics"



1. Theories are developed as generally as possible.
2. This seems to be the first Bourbaki-Style formulation of actuarial mathematics (still developing).

The survival model and the life table are formalized by using locale.

```
locale prob_space_actuary = MM_PS: prob_space 𝔐 for 𝔐
  — ‹Since the letter M may be used as a commutation function,
     adopt the letter @{text "𝔐"} instead as a notation for the measure space.›
locale survival_model = prob_space_actuary +
  fixes X :: "'a ⇒ real"
  assumes X_RV[simp]: "MM_PS.random_variable (borel :: real measure) X"
    and X_pos_AE[simp]: "AE ξ in 𝔐. X ξ > 0"

locale life_table =
  fixes l :: "real ⇒ real" ("$l'_" [101] 200)
  assumes l_0_pos: "0 < l 0"
    and l_neg_nil: "⋀x. x ≤ 0 ⟹ l x = l 0"
    and l_PInfty_0: "(l ⟶ 0) at_top "
    and l_antimono: "antimono l"
    and l_right_continuous: "⋀x. continuous (at_right x) l"
```

Applications I

You can verify examination problems and solutions of life contingencies.

Question (MCQ No.3 of Exam LTAM, SoA, Spring 2022)

You are given the following survival function for a newborn:

$$S_0(x) = (1 - 0.01x)^{0.5}, \quad 0 \leq x \leq 100.$$

Calculate $1000\mu_{25}$.

Answer

6.7.

```
lemma SoA_LTAM_2022_Spring_MCQ_No3:
  assumes "⋀x::real. 0 ≤ x ⟹ x ≤ 100 ⟹ ccdf (distr 𝔐 borel X) x = (1 - 0.01*x).^0.5"
  shows "¦1000*$μ_25 - 6.7¦ < 0.05"
```

Applications II

Question (No.2-1-1 of Life Insurance Math, IAJ, 2016; my translation)

Suppose $\mathring{e}_x = ax + b$ for $0 \leq x \leq -b/a$, where a $(-1 < a < 0)$ and b are constants independent of x. Find l_x.

Answer

$$l_x = l_0 \cdot \left(\frac{b}{ax+b}\right)^{(a+1)/a}.$$

```
lemma IAJ_Life_Insurance_Math_2016_2_1_1:
  fixes a b :: real
  assumes "-1 < a" "a < 0" "0 < b" "-b/a ≤ $ψ" and
    "⋀x. 0 < x ⟹ x < -b/a ⟹ l differentiable at x" and
    "⋀x. 0 ≤ x ⟹ x < -b/a ⟹ l integrable_on {x..} ∧ $e`∘_x = a*x + b"
  shows "⋀x. 0 ≤ x ⟹ x < -b/a ⟹ $l_x = $l_0 * (b / (a*x + b)).^((a+1)/a)"
```

There are some implicit assumptions (highlighted in yellow) in exam problems.

Contents

Interchange of Differentiation and Lebesgue Integration

Theorem

Let Ω be a measure space and a, b be real numbers. Suppose $f : (a,b) \times \Omega \to \mathbb{R}$ and $g : \Omega \to \mathbb{R}$ satisfy the following conditions:

1. *$f(r,x)$ is an integrable function of $x \in \Omega$ for each $r \in (a,b)$,*
2. *for almost all $x \in \Omega$, the partial derivative $f_r(r,x)$ exists for all $r \in (a,b)$,*
3. *g is an integrable function,*
4. *for almost all $x \in \Omega$, we have $|f_r(r,x)| \le g(x)$ for all $r \in (a,b)$.*

Then we have $\frac{d}{dr} \int_\Omega f(r,x)\,dx = \int_\Omega f_r(r,x)\,dx$ for all $r \in (a,b)$.

```
proposition interchange_deriv_LINT_general:
  fixes a b :: real and f :: "real ⇒ 'a ⇒ real" and g :: "'a ⇒ real"
  assumes f_integ: "⋀r. r∈{a<..<b} ⟹ integrable M (f r)" and
    f_diff: "AE x in M. (λr. f r x) differentiable_on {a<..<b}" and
    Df_bound: "AE x in M. ∀r∈{a<..<b}. ¦deriv (λr. f r x) r¦ ≤ g x" "integrable M g"
  shows "⋀r. r∈{a<..<b} ⟹ ((λr. ∫x. f r x ∂M) has_real_derivative
    ∫x. borel_measurize M (λx. deriv (λr. f r x) r) x ∂M) (at r)"
```

Tail Expectation Formula

Theorem

Let $X : \Omega \to \mathbb{R}$ be a random variable, and define $F(u) := P(X \leq u)$. If X is integrable on Ω, we have $E[X] = \int_0^\infty (1 - F(u))\,du - \int_{-\infty}^0 F(u)\,du$.

```
proposition expectation_tail:
  assumes [measurable]: "integrable M X"
  defines "F u ≡ cdf (distr M borel X) u"
  shows "expectation X = (LBINT u:{0..}. 1 - F u) - (LBINT u:{..0}. F u)"
```

Conditional Probability Space

Definition (Conditional Probability Space)

Let $\mathfrak{M} := (\Omega, \mathcal{F}, P)$ be a probability space, and suppose $A \in \mathcal{F}$ satisfies $P(A) > 0$. If we define

1. $\mathcal{F}_A := \{C \cap A \mid C \in \mathcal{F}\}$,
2. $P_A(B) := P(B \mid A)$ for $B \in \mathcal{F}_A$,
3. $\mathfrak{M} \downharpoonright A := (A, \mathcal{F}_A, P_A)$,

then $\mathfrak{M} \downharpoonright A$ becomes the probability space.

```
definition cond_prob_space :: "'a measure ⇒ 'a set ⇒ 'a measure" (infix "⇂" 200)
  where "M⇂A ≡ scale_measure (1 / emeasure M A) (restrict_space M A)"

lemma cond_prob_space_correct:
  assumes "A ∈ events" "prob A > 0"
  shows "prob_space (M⇂A)"
```

Complementary Cumulative Distribution Function

Definition (Complementary Cumulative Distribution Function)

For a random variable X, define $\overline{F}(x) := P(X > x)$ for $x \in \mathbb{R}$. The function $\overline{F}$ is called the *complementary cumulative distribution function* or the *tail distribution*.

```
definition ccdf :: "real measure ⇒ real ⇒ real"
  where "ccdf M ≡ λx. measure M {x<..}"
    — ‹complementary cumulative distribution function (tail distribution)›
```

Lemma

In any probability space, we have $F(x) + \overline{F}(x) = 1$.

```
lemma add_cdf_ccdf: "cdf M x + ccdf M x = measure M (space M)"
```

Hazard Rate

Definition (Hazard Rate [2])

- The hazard rate refers to the rate of death for an item of a given age (x).
- The hazard rate for any time can be determined using the following equation:

$$h(t) = f(t)/R(t)$$

  $F(t)$ [*sic*] is the probability density function (PDF), ... $R(t)$, on the other hand, represents the survival function, ...

```
definition hazard_rate :: "('a ⇒ real) ⇒ real ⇒ real"
  where "hazard_rate X t ≡
    Lim (at_right 0) (λdt. 𝒫(x in M. t < X x ∧ X x ≤ t + dt ¦ X x > t) / dt)"
    — ‹Here, X is supposed to be a random variable.›

lemma hazard_rate_density_ccdf:
  assumes "distributed M lborel X f"
    and "⋀s. f s ≥ 0" "t < u" "continuous_on {t..u} f"
  shows "hazard_rate X t = f t / ccdf (distr M borel X) t"
```

Contents

Future Works

1. Upgrade the theory of the life table.
2. Formalize the remaining part of actuarial mathematics in Isabelle/HOL: the pricing model, the evaluation model, etc.
3. Give the probabilistic foundation to the previous work `coq-actuary`: `https://github.com/Yosuke-Ito-345/Actuary`

Future Applications

1. Prevent errors in examination papers of actuarial mathematics.
2. Detect errors in actuarial documents, e.g. statement of calculation procedures.
3. Formally verify programs in actuarial software, e.g. MG-ALFA.

References

[1] Marcel B. Finan.
A reading of the theory of life contingency models: A preparation for exam MLC/3L.
https://faculty.atu.edu/mfinan/actuarieshall/mlcbook2.pdf, 2011.

[2] Investopedia.
Dictionary.
https://www.investopedia.com/financial-term-dictionary-4769738, 2023.

[3] Francis S. Perryman.
International actuarial notation.
In *Proceedings of the Casualty Actuarial Society*, Vol. 36, pp. 123–131, 1949.
Also available as https://www.casact.org/sites/default/files/database/proceed_proceed49_49123.pdf.

[4] 厚生労働省.
第 23 回生命表（完全生命表）の概況.
https://www.mhlw.go.jp/toukei/saikin/hw/life/23th/dl/23th-11.pdf, 2022.

Q&A

Any Questions?
(English Available)

Contents

Life Tables in Japan

The Ministry of Health, Labour and Welfare produces 2 kinds of Life Tables in Japan.

1. *Complete Life Tables* are made every 5 years based on the Annual Vital Statistics and the Population Census.
2. *Abridged Life Tables* are made every year based on the Provisional Annual Vital Statistics and the Population Estimates.

`https://www.mhlw.go.jp/english/database/db-hw/vs02.html`

Specific Life Tables

Definition (Smooth Life Table)

A life table l is *smooth* when

1. l is continuous, and
2. there exists a finite subset $S \subseteq \mathbb{R}$ such that l is differentiable on $\mathbb{R} \setminus S$.

```
locale smooth_life_table = life_table +
  assumes l_piecewise_differentiable[simp]: "l piecewise_differentiable_on UNIV"
```

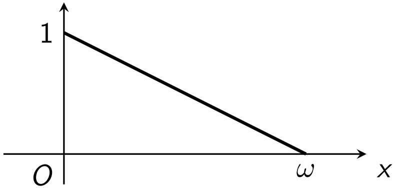
Definition (Finite Life Table)

A life table l is *finite* when $l_x = 0$ for some x.

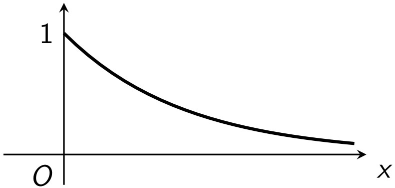
```
locale finite_life_table = life_table +
  assumes l_finite: "∃x::real. $l_x = 0"
```

Examples of Survival Model

De Moivre's Law: $s(x) = 1 - \frac{x}{\omega},\ 0 \leq x \leq \omega.$



Constant Force Model: $s(x) = e^{-\mu x},\ 0 \leq x.$



Ultimate Age

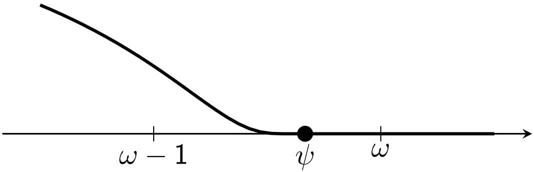
Definition

Let s be a survival function.

1. Define $\psi := \inf\{x \in \mathbb{R} \mid s(x) = 0\}$. This is my original notation.
2. When s is finite, define $\omega := \min\{x \in \mathbb{N} \mid s(x) = 0\}$. This notation is internationally used.

```
definition death_pt :: ereal ("$ψ")
  where "$ψ ≡ Inf (ereal ` {x ∈ ℝ. ccdf (distr 𝔐 borel X) x = 0})"
    — ‹This is my original notation,
       which is used to develop life insurance mathematics rigorously.›

definition ult_age :: nat ("$ω")
  where "$ω ≡ LEAST x::nat. ccdf (distr 𝔐 borel X) x = 0"
    — ‹the conventional notation for ultimate age›
```



Extension of Almost-Everywhere-Defined Function

In measure theory, we often extend an almost-everywhere-defined function to a measurable total function. I coined the function `measurize` to formalize this operation.

```
definition measurize :: "'a measure ⇒ 'b measure ⇒ ('a ⇒ 'b) ⇒ 'a ⇒ 'b" where
  "measurize M N f = (SOME g. g ∈ M →M N ∧ (∃S∈(null_sets M). {x ∈ space M. f x ≠ g x} ⊆ S))"
  — ‹The term "measurize" is what I coined in this formalization.›

definition borel_measurize :: "'a measure ⇒ ('a ⇒ 'b::topological_space) ⇒ 'a ⇒ 'b" where
  "borel_measurize M f = measurize M borel f"

lemma
  fixes f g
  assumes "g ∈ M →M N" "S ∈ null_sets M" "{x ∈ space M. f x ≠ g x} ⊆ S"
  shows measurizeI: "AE x in M. f x = measurize M N f x" and
    measurizeI2: "AE x in M. g x = measurize M N f x" and
    measurize_measurable: "measurize M N f ∈ measurable M N"
```