MA671 course notes

Krit Boonsiriseth

Updated November 9, 2018

Abstract

Course notes for Hotchkiss class MA671 (Topics in Advanced Mathematics). The "Topic" for this course is Actuarial Mathematics. Proofs in here are very non-rigorous; you can figure out the rest on your own.

Contents

1 Probability and Statistics

1.1 Random variables and distributions

Refer to MA655 Probability and Statistics handout. I'll just put very important parts of the handout here.

Definition 1.1.1. A *probability measure* $\Pr : \mathcal{E} \to [0,1]$ is a function which assigns to each event $E \in \mathcal{E}$ a real number $\Pr(E)$ from 0 to 1, subject to the following conditions:

(i) $\Pr(\Omega) = 1$

(ii) For countably many pairwise disjoint events $\{E_j\}$, $\Pr(\bigcup E_j) = \sum \Pr(E_j)$.

Problem 1.1.2. Show the following properties of probability measures:

(a) $\Pr(\emptyset) = 0$

(b) If $E, F \in \mathcal{E}$ such that $E \subset F$, then $\Pr(E) \leqslant \Pr(F)$.

(c) If $E, F \in \mathcal{E}$, then $\Pr(E \cup F) = \Pr(E) + \Pr(F) - \Pr(E \cap F)$.

Solution. Basically use property (ii) of probability measures.

(a) $\Pr(\emptyset) = \Pr(\emptyset) + \Pr(\emptyset)$.

(b) $\Pr(F) = \Pr(E) + \Pr(F - E) \geqslant \Pr(E)$.

(c) (I'll try to format this better)

$$\begin{aligned}\Pr(E \cup F) &= \Pr(E \cap F) + \Pr(E - F) + \Pr(F - E) \\ &= (\Pr(E \cap F) + \Pr(E - F)) + (\Pr(E \cap F) + \Pr(F - E)) - \Pr(E \cap F) \\ &= \Pr(E) + \Pr(F) - \Pr(E \cap F) \blacksquare\end{aligned}$$

Definition 1.1.3. An n-dimensional *random variable* is a function $X : \Omega \to \mathbb{R}^n$.

Problem 1.1.4. Consider the random experiment "tossing two fair dice." Let X_1 and X_2 be the values of the first and second die (as naturally defined). Furthermore, let $Y = X_1 + X_2$.

(a) What are all possible values y of Y?

(b) Write out the sample space for this experiment.

(c) Define a probability measure for this sample space.

(d) Determine the events $\{Y = y\}$ for $y = 2, 3, \ldots, 12$.

(e) Determine $\Pr(\{Y = y\})$ for $y = 2, 3, \ldots, 12$.

Solution. (a) The possible values are $2, 3, \ldots, 12$

(b) 36 tuples of dice faces (graphics omitted due to some complications in the link file generated)

(c) Define $\Pr(E) = |E|/36$

(d) Basically, make a table

(e) $\Pr(\{Y = y\})$ follows this table:

y	2	3	4	5	6	7	8	9	10	11	12
$\Pr(\{Y = y\})$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

1.2 Distributions

Definition 1.2.1. A *Bernoulli random variable* X is a random variable that takes on the two values $\{0, 1\}$. It possesses a *Bernoulli distribution* $B(p), p \in [0, 1]$, which means that $\Pr(X = 1) = p$ and $\Pr(X = 0) = 1 - p$.

Definition 1.2.2. Suppose that X is a discrete random variable defined on a sample space Ω. The *probability mass function (p.m.f.)* $f_X : \mathbb{R} \to [0, 1]$ for X is defined as

$$f_X(x) = \Pr(X = x).$$

The set A where f_X can take nonzero values is called the *support* of f_X.

Definition 1.2.3. The *Binomial distribution* $B(n, p)$ is just the distribution of the sum of n Bernoulli random variables each with a Bernoulli distribution $B(p)$.

Observation 1.2.4. If $X \sim B(n, p)$ (meaning that X follows $B(n, p)$), then

$$f_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{if } x \in \{0, 1, \ldots, n\} \\ 0 & \text{otherwise} \end{cases}$$

Problem 1.2.5. The probability that I dial a wrong number is 0.005 when I make a telephone call. In a typical week I will make 75 telephone calls.

(a) Describe the distribution of the random variable T that indicates the number of times I dial a wrong number in a week

(b) In a given week, find the probability that I dial no wrong numbers.

(b') In a given week, find the probability that I dial more than two wrong numbers.

Solution. (a) It is the binomial distribution $B(75, 0.005)$.

(b) $f_T(0) = 0.995^{75} \approx 0.6866$.

(b') $\sum_{k>2} f_T(k) = 1 - f_T(0) - f_T(1) - f_T(2) = 1 - 0.995^{75} - 75 \cdot 0.005 \cdot 0.995^{74} - \binom{75}{2} \cdot 0.005^2 \cdot 0.995^{73} \approx 0.0064$. ■

Definition 1.2.6. The *Geometric distribution* Geo(p) is the distribution of the number of trials before the first success when we conduct independent Bernoulli trials $B(p)$. It is a matter of convention whether we count the number of failures before, or the number of trials until, the first success.

Problem 1.2.7. At a luxury ski resort in Switzerland, the probability that snow will fall on any given day in the snow season is 0.15

(a) If the snow season begins on November 1st, find the probability that the first snow will fall on November 15.

(b) Given that no snow fell during November, a tourist decides to wait no longer to book a holiday. The tourist decides to book for the earliest date for which the probability that snow will have fallen on or before that date is greater than 0.85. Find the exact date of the booking.

Solution. (a) $f_X(15) = (1 - 0.15)^{14}(0.15) \approx 0.0154$

(b) Chance of no snow on or before December t is $(1 - 0.15)^t$. We want to find minimal $t \in \mathbb{Z}_{\geqslant 0}$ such that $(1 - 0.15)^t < 0.15$. Since $0.85^{11} > 0.15 > 0.85^{12}$, the date of the booking is December 12th. ■

1.3 Continuous random variables

Before class we read about the *probability density function (p.d.f)* which is basically a continuous version of p.m.f and also the *cumulative density function (c.d.f)* which is basically the sum (or integral) of p.m.f (or p.d.f) up to a value x. Here are more formal definitions:

Definition 1.3.1. A random variable X has *probability density function* $f_X : \mathbb{R} \to [0, +\infty)$ if

$$\Pr(a \leqslant X \leqslant b) = \int_a^b f_X(x)\, \mathrm{d}x.$$

Again, the set A where f_X can take nonzero values is called the *support* of f_X.

Definition 1.3.2. The *cumulative distribution function (c.d.f)* of a random variable is defined as $F_x(x) = \Pr(X \leqslant x)$.

Observation 1.3.3. We can write c.d.f in terms of p.m.f or p.d.f as

$$F_x(x) = \begin{cases} \sum_{t \leqslant x} f_X(t) & \text{for } X \text{ discrete} \\ \int_{-\infty}^{x} f_X(t)\, \mathrm{d}t & \text{for } X \text{ continuous} \end{cases}$$

Some Q&As to start the class

Q: So what happens when a random variable is both discrete and continuous (like $X = 1$ with probability 0.5 and $2 \leq X \leq 3$ with probability 0.5 distributed uniformly)
A: It works if we move from Riemann integrals to Lebesgue integrals (advanced stuff)

Q: Difference between $\mathcal{E}$ and Ω?
A: $\mathcal{E}$ is the power set of Ω (= the set of all subsets of Ω.)

After this we re-did the problems from last class (because one-thirds of our members (=1 person) were missing in the previous class), but the difference is that now we have calculators!

Solution (to Problem 2.6). (a) It is the binomial distribution $B(75, 0.005)$.

(b) `BinomPdf(75, 0.005, 0)`

(b') `BinomCdf(75, 0.005, 3, 75)`

1.4 E and Var

Example 1.4.1. If $F_x(x)$ is a c.d.f of a discrete random variable taking values in $\{0, 1, \ldots, n\}$ then for all $a \leqslant b \in \{0, 1, \ldots, n\}$ then $\Pr(X < a) = F_X(b-1)$ and $\Pr(a \leqslant X \leqslant b) = F_X(b) - F_X(a-1)$.

Definition 1.4.2. Let X be a random variable. Then the *expected value* $\mathrm{E}(X)$ of X is defined as

$$\mathrm{E}(X) = \begin{cases} \sum_{x \in X(\Omega)} x f_X(x) & \text{for } X \text{ discrete} \\ \int_{-\infty}^{\infty} x f_X(x)\, \mathrm{d}x & \text{for } X \text{ continuous} \end{cases}$$

Definition 1.4.3. The *variance*, used to measure how spread out the values of X is defined as

$$\mathrm{Var}(X) = \mathrm{E}\left((X - \mathrm{E}(X))^2\right).$$

The above definition of variance is nice in terms of explaining why it is the variance, but a bit complex. It is actually true that

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2$$

but this requires two facts: $\mathrm{E}(aX + b) = a\mathrm{E}(X) + b$ and $\mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y)$, both of which will be shown later.

Also note that the variance is "quadratic" in X, so we also have the standard deviation which is in the same unit as X:

Definition 1.4.4. The *standard deviation* σ_X of X is defined as $\sigma_X = \sqrt{\mathrm{Var}(x)}$.

Problem 1.4.5. Show that $\mathrm{E}(aX + b) = a\mathrm{E}(X) + b$ for all $a, b \in \mathbb{R}$.

Solution. Since $\Pr(a\ell + b \leqslant aX + b \leqslant ar + b) = \Pr(\ell \leqslant X \leqslant r)$, from definition of p.d.f. we have $f_{aX+b}(ax + b)\,\mathrm{d}(ax + b) = f_X(x)\,\mathrm{d}x$, so

$$\mathrm{E}(aX + b) = \int_{-\infty}^{\infty} (ax + b) f_{aX+b}(ax + b)\,\mathrm{d}(ax + b) = \int_{-\infty}^{\infty} (ax + b) f_X(x)\,\mathrm{d}x = a\mathrm{E}(X) + b. \blacksquare$$

Observation 1.4.6. Also we can note that if $Y = g(X)$ then $\mathrm{E}(Y) = \int g(x) f_X(x)\,\mathrm{d}x$. ■

Problem 1.4.7 (Homework)**.** Let X be the continuous random variable with p.d.f.

$$f_X(x) = \lambda e^{-\lambda x}.$$

for $x \in [0, \infty)$ and $\lambda > 0$. Find $\mathrm{E}(X)$ and $\mathrm{Var}(X)$.

Solution.

$$\mathrm{E}(X) = \int_0^{\infty} x \cdot \lambda e^{-\lambda x}\,\mathrm{d}x = -e^{-\lambda x}\left(x + \frac{1}{\lambda}\right)\Big|_0^{\infty} = \frac{1}{\lambda}.$$

$$\mathrm{E}(X^2) = \int_0^{\infty} x^2 \cdot \lambda e^{-\lambda x}\,\mathrm{d}x = -e^{-\lambda x}\left(x^2 + \frac{2x}{\lambda} + \frac{2}{\lambda^2}\right)\Big|_0^{\infty} = \frac{2}{\lambda^2}$$

Therefore,

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{1}{\lambda^2}. \blacksquare$$

1.5 Independent variables

Definition 1.5.1. Two random variables X and Y are called *independent* if observing one variable yields no additional information about the value of the other, that is, $\Pr(Y = y | X = x) = \Pr(Y = y)$ for all x, y.

Since $\Pr(E|F) = \frac{\Pr(E \cap F)}{\Pr(F)}$, for independent events E and F we have

$$\Pr(E \cap F) = \Pr(E)\Pr(F).$$

Theorem 1.5.2. *If X and Y are independent random variables then* $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$.

Theorem 1.5.3. *For* any *random variables X and Y,* $\mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y)$.

To prove this (for the discrete case), we need another theorem:

Theorem 1.5.4 (Law of Total Probability). *If $\{B_n : n = 1, 2, 3, \ldots\}$ is a finite or countably infinite partition of a sample space, then for any event A of the same probability space,*

$$\Pr(A) = \sum_n \Pr(A|B_n)\Pr(B_n).$$

This sounds pretty intuitive and the proof is not given in class.

Proof (of Theorem 5.3; discrete case only). $\mathrm{E}(X+Y) = \sum_{x,y}(x+y)\Pr(X = x \text{ and } Y = y)$

$$\begin{aligned}
&= \sum_x x \sum_y \Pr(X = x \text{ and } Y = y) + \sum_y y \sum_x \Pr(X = x \text{ and } Y = y) \\
&= \sum_x x \sum_y \Pr(X = x|Y = y)\Pr(Y = y) + \sum_y y \sum_x \Pr(Y = y|X = x)\Pr(X = x) \\
&= \sum_x x\Pr(X = x) + \sum_y y\Pr(Y = y) \quad \text{[from Thm 5.4]} \\
&= \mathrm{E}(X) + \mathrm{E}(Y). \quad \blacksquare
\end{aligned}$$

1.6 Two-dimensional random variables

Two-dimensional random variables: $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$. If X is discrete, the p.m.f. is

$$f_X(x_1, x_2) = \Pr(X_1 = x_1, X_2 = x_2).$$

If X is continuous, $A \subseteq \mathbb{R}^2$, the joint p.d.f. of X_1, X_2 is $f_X(x_1, x_2)$ such that

$$\Pr(X \in A) = \int_A f_X(x_1, x_2)\, \mathrm{d}^2(x_1, x_2).$$

The c.d.f. of X is $F_X(x_1, x_2) = \Pr(X_1 \leqslant x_1, X_2 \leqslant x_2)$.

The *marginal distribution* of X_1, X_2 can be obtained by summing / integrating "away" the other variable:

$$f_{X_1}(x) = \sum_{x_2} f_X(x, x_2) \text{ or } \int_{x_2=-\infty}^{\infty} f_X(x, x_2)\, \mathrm{d}x_2.$$

$$f_{X_2}(x) = \sum_{x_1} f_X(x, x_1) \text{ or } \int_{x_1=-\infty}^{\infty} f_X(x, x_1)\, \mathrm{d}x_1.$$

See Nov '01 Course 1 exam #28 for example.

1.7 Covariance and correlation

Recall

$$\mathrm{Var}(x) = \mathrm{E}\left[(X - \mathrm{E}(X))^2\right].$$

The *covariance* $\mathrm{Cov}(X, Y)$ is defined as

$$\mathrm{Cov}(X, Y) = \mathrm{E}\big((X - \mathrm{E}(X))(Y - \mathrm{E}(Y))\big)$$

Similar to $\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2$,

$$\mathrm{Cov}(X, Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y).$$

We also define the *correlation* $\rho(X,Y)$ as

$$\frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}.$$

This looks strikingly similar to the definition of the angle between two vectors.

Problem 1.7.1. Open-ended: what are the parallels between covariance of 2 random variables and the dot product between 2 vectors in $\mathbb{R}^n$? What can linear algebra tell us about probability?

Example 1.7.2. Let X be a 2×1 discrete random vector and denote its components by X_1 and X_2. Let the support of X be

$$R_X = \left\{ \begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}2\\0\end{bmatrix}, \begin{bmatrix}0\\0\end{bmatrix} \right\}$$

and its joint probability mass function be $p_X(x) = 1/3$ for $x \in R_X$, and 0 otherwise.

We can compute $\mathrm{E}(X_1) = 1, \mathrm{E}(X_2) = 1/3, \mathrm{E}(X_1X_2) = 1/3, \mathrm{Cov}(X_1, X_2) = 0, \rho(X_1, X_2) = 0$ so X_1 and X_2 are *uncorrelated*.

Definition 1.7.3. Given random vector $[X\ Y]^T$ (discrete),

$$p_{X|Y=y}(x) = \Pr(X = x \mid Y = y) = \frac{\Pr(X = x \text{ and } Y = y)}{\Pr(Y = y)} = \frac{p_{X,Y}(x,y)}{p_Y(y)}.$$

The continuous version is similar.

2 Econ 101

Now let's talk about money

— Dr. Weiss, 10 Sep 2018

⋆ ⋆ ⋆

2.1 Interest

So we are turning to some Econ stuff I don't really understand that well. Let's start with *interest* which compensates for time-value of the money.

Definition 2.1.1.
- The *effective interest rate* i is the amount of interest earned per base period, usually a year.
- The *nominal interest rate* $i^{(m)}$ is interest converted m times per base unit of time. It is related to i by the formula

$$1+i=\left(1+\frac{i^{(m)}}{m}\right)^m.$$

- The *discount factor* $v=\frac{1}{1+i}$ is the amount you need to invest to get 1 dollar at the end of the year (so basically a dollar n years in the future is worth v^n dollars now).
- The *force of interest* $\delta=\lim_{m\to\infty} i^{(m)}$ is basically the "continuous" interest. It is related to i by

$$1+i=e^\delta.$$

Example 2.1.2 (Investing with interest)**.** Let F_0 be an initial amount invested; $r_k, k\geqslant 1$, additional amounts invested at the end of each year k. Let F_k be the amount at the end of year k. Then, for all $k\geqslant 1$,

$$F_k=(1+i)F_{k-1}+r_k,$$

where, by induction, we have

$$F_n=(1+i)^nF_0+\sum_{k=1}^n(1+i)^{n-k}r_k.$$

Arranging in terms of v instead, we have

$$v^nF_n=F_0+\sum_{k=1}^n v^kr^k$$

where v^nF_n is the present value of the final balance, F_0 is the initial amount, and the $\sum$ is the present value of future payments.

Example 2.1.3 (Continuous model)**.** Consider a model where
- payments are made at a rate $r(t)$, and
- interest credited continuously at force $\delta(t)$.

Over an infinitesimal time period dt, the balance grows as $dF(t) = F(t)\delta(t)dt + r(t)dt$ or

$$\boxed{F'(t) = F(t)\delta(t) + r(t).}$$

Assuming $r(t) \equiv 0$, this is a homogeneous differential equation $F' = F\delta$ which can be solved to get

$$\ln(F(t)) = \int \delta(t)\ \mathrm{d}t + c = \int \delta + c$$

so

$$F(t) = c \cdot e^{\int \delta(t)\ \mathrm{d}t}.$$

Variation of the constant allows us to solve the original equation: set $c = C(t)$ so

$$F(t) = C(t)e^{\int \delta(t)dt} \text{ and } F'(t) = C'(t)e^{\int \delta(t)\ \mathrm{d}t} + C(t)\delta(t)e^{\int \delta(t)\ \mathrm{d}t}.$$

Compare this to the boxed equation above to get

$$r(t) = C'(t)e^{\int \delta(t)\ \mathrm{d}t}.$$

Therefore, since $C(0) = F(0)$,

$$C(t) = F(0) + \int_0^t r(u)e^{-\int_0^u \delta(s)\ \mathrm{d}s}\ \mathrm{d}u.$$

Let's continue from the formula we derived in the previous class. If $\delta(t) \equiv \delta$ is constant then

$$F(t) = F(0)e^{\delta t} + \int_0^t r(u)e^{\delta(t-u)}\ \mathrm{d}u.$$

If $r(t) \equiv r$ is also constant, then

$$F(t) = F(0)e^{\delta t} + r\frac{e^{\delta t} - 1}{\delta}$$

Definition 2.1.4. If we receive "interest" dC at the *beginning* of the period on an investment C instead, we call d the *discount* rate.

Investing dC under the same conditions yields an interest of d^2C and on and on, so at the end of the year we will have

$$C + dC + d^2C + \cdots = \frac{1}{1-d}C.$$

If i is the effective interest rate then $1 + i = \frac{1}{1-d}$, so

$$d = \frac{i}{1+i} \text{ and } i = \frac{d}{1-d}.$$

2.2 Perpetuities and annuities

Definition 2.2.1.
- *Perpetuity due* is when you receive \$1 at the beginning of each period: time $t = 0, 1, 2, 3, \ldots$
- *Immediate perpetuity* is when you receive \$1 at the end of each period: time $t = 1, 2, 3, 4, \ldots$

We can calculate the present value of perpetuity due as

$$\ddot{a}_{\overline{\infty}|} = 1 + v + v^2 + \cdots = \frac{1}{1-v} = \frac{1}{d},$$

and the present value of immediate perpetuity as

$$a_{\overline{\infty}|} = v + v^2 + \cdots = \frac{1}{i}.$$

Scenario 2.2.2. m payments of $\frac{1}{m}$ per year at nominal rate $i^{(m)}$ or $d^{(m)}$ You'll get present value of the immediate perpetuity

$$a^{(m)}_{\overline{\infty}|} = \frac{1}{m}\left(v^{\frac{1}{m}} + v^{\frac{2}{m}} + \cdots\right) = \frac{1}{m} \cdot \frac{v^{\frac{1}{m}}}{1 - v^{\frac{1}{m}}} = \frac{1}{i^{(m)}}$$

Connection between $i^{(m)}$ and $d^{(m)}$:

- You receive interest of $\frac{d^{(m)}}{m} \cdot C$ at the beginning of the pay period and return the capital C at the end. At the end of the period, you will get $\frac{1}{1-\frac{d^{(m)}}{m}}C$.
- With a nominal interest rate $i^{(m)}$ you would get $C + \frac{i^{(m)}}{m}C$.

These are equal, so

$$\frac{1}{1 - \frac{d^{(m)}}{m}} = 1 + \frac{i^{(m)}}{m}.$$

By some algebraic manipulations,

$$d^{(m)} = m\left(1 - (1+i)^{-\frac{1}{m}}\right) \text{ and } \frac{1}{d^{(m)}} = \frac{1}{m} + \frac{1}{i^{(m)}}.$$

Exercise 2.2.3. Present value of perpetuity due $\ddot{a}^{(m)}_{\overline{\infty}|}$ is equal to $\frac{1}{d^{(m)}}$.

Solution. Present value of perpetuity due is

$$\frac{1}{m}\left(1 + v^{\frac{1}{m}} + v^{\frac{2}{m}} + \cdots\right) = \frac{1}{i^{(m)} v^{\frac{1}{m}}} = \frac{1 + \frac{i^{(m)}}{m}}{i^{(m)}} = \frac{1}{i^{(m)}} + \frac{1}{m} = \frac{1}{d^{(m)}}. \blacksquare$$

Note that when we take m to ∞ we have

$$\lim_{m\to\infty} d^{(m)} = \delta = \lim_{m\to\infty} i^{(m)}$$

so

$$\bar{a}_{\overline{\infty}|} = \text{ present value of continuous perpetuity } = \frac{1}{\delta}.$$

Scenario 2.2.4. Perpetuity with varying payments $r_0, r_1, \ldots$

$$\ddot{a} = r_0 + r_1 v + r_2 v^2 + \cdots$$

represents sum of constant payment perpetuities: $r_i - r_{i-1}$ made at time i.

Therefore

$$\ddot{a} = \frac{r_0}{d} + \frac{r_1 - r_0}{d}v + \frac{r_2 - r_1}{d}v^2 + \cdots.$$

Example 2.2.5. Payment of \$1, increasing by 1 each year is worth

$$\ddot{a} = \frac{1}{d} \cdot \frac{1}{1-v} = \frac{1}{d^2}.$$

Example 2.2.6. Continuous payment, increasing exponentially, with continuous compounding of δ.

$$\ddot{a} = \int_0^\infty e^{\lambda t} e^{-\delta t} = \frac{1}{\delta - \lambda} \text{ for } \delta > \lambda.$$

Definition 2.2.7. *Annuity* is just like perpetuities, but with n payments instead of ∞.

Similar to perpetuities, the present value of annuity due is

$$\ddot{a}_{\overline{n}|} = 1 + v + v^2 + \cdots + v^{n-1} = \frac{1 - v^n}{d}$$

and the present value of immediate annuity is

$$a_{\overline{n}|} = v + v^2 + v^3 + \cdots + v^n = \frac{1 - v^n}{i}.$$

Since annuities have an ending, they also have a final value

$$\ddot{s}_{\overline{n}|} = (1+i)^n \ddot{a}_{\overline{n}|} = \frac{(1+i)^n - 1}{d} \quad \text{and} \quad s_{\overline{n}|} = \frac{(1+i)^n - 1}{i}.$$

2.3 Debt

Repayment of debt: **debt = present value of payments repaying the debt.**

Assume that repayment starts at end of first period:

$$S = vr_1 + v^2 r_2 + \cdots + v^n r_n.$$

Let S_k be the debt after k^{th} payment, then

$$S_k = (1+i)S_{k-1} - r_k \quad \text{and} \quad r_k = iS_{k-1} + (S_{k-1} - S_k).$$

Here, iS_{k-1} represents interest payment, and $S_{k-1} - S_k$ represents principal payment.

Example 2.3.1. When paying loans for a house, the interest payments actually represents the *rent* you are paying the bank. However, paying for a house in cash will also cause you to lose liquidity of assets, exposing you to risk. Moral of the story: deciding how to use money is hard.

Another formula of debt is the retrospective formula (which is just equivalent to previous formulæ)

$$S_k = (1+i)^k S - \sum_{j=1}^{k} (1+i)^{k-j} r_j.$$

Here, $(1+i)^k S$ is the interest accrued on the debt and $\sum_{j=1}^{k} (1+i)^{k-j} r_j$ is the accumulated value of the payment stream.

There is also the prospective formula, looking ahead instead of backwards:

$$S_k = vr_{k+1} + v^2 r_{k+2} + \cdots + v^{n-k} r_n.$$

In real life, most loans have constant payments (r is constant), so

$$S = r a_{\overline{n}|}.$$

3 Life Insurance

3.1 Introduction

Notation 3.1.1 (Life insurance notations).

- (x) denotes a person with age x years, also called *life aged* x.
- T is a random variable for future lifetime.
- $G(t) = \Pr(T \leqslant t)$ is the c.d.f. of T.
- $g(t) = G'(t)$ is the p.d.f. of T.
- ${}_tq_x = G(t)$ is the probability that (x) will die within t
- ${}_tp_x = 1 - G(t)$ is the probability that (x) will survive at least t years.
- ${}_{s|t}q_x$ is the probability that (x) survives s years but dies within another t years.
- ${}_tp_{x+s}$ is the probability that (x), having attained age $x+s$, will survive another t years.
- ${}_tq_{x+s}$ is the probability that (x), having attained age $x+s$, will die in the next t years.

Observation 3.1.2. We have

$${}_tp_{x+s} = \Pr(T > s+t | T > s) = \frac{Pr(T > s+t) \text{ and } Pr(T > s)}{Pr(T > s)} = \frac{1 - G(s+t)}{1 - G(s)}$$

and

$${}_tq_{x+s} = 1 - {}_tp_{x+s} = \frac{1 - G(s+t)}{1 - G(s)}.$$

Definition 3.1.3. The *expected remaining lifetime* of a life aged x is $\mathrm{E}(T)$, denoted by

$$\mathring{e}_x = \int_0^\infty t g(t)\, \mathrm{d}t.$$

In fact we have

$$\mathring{e}_x = \int_0^\infty {}_tp_x\, \mathrm{d}t,$$

provable as follows:

$$\int_0^\infty {}_tp_x\, \mathrm{d}t = \int_0^\infty \int_t^\infty g(s)\, \mathrm{d}s\, \mathrm{d}t = \int_0^\infty \int_0^s g(s)\, \mathrm{d}t\, \mathrm{d}s = \int_0^\infty s g(s)\, \mathrm{d}s.$$

Next we'll do exercises.

Exercise 3.1.4. C.1.1.3 in [Gerber]

Solution. Calculating the value on January 1, 1990 gives

$$x + vx + \cdots + v^9 x = 15000(v^{10} + \cdots + v^{14}).$$

This reduces to

$$x = \frac{15000 v^{10}}{1 + v^5} = 4793.75. \blacksquare$$

Exercise 3.1.5. C.1.1.10 in [Gerber]

Solution. The value from (i) is $120a^{(12)}_{\overline{\infty}|} = \frac{120}{i^{(12)}}$. The value from (ii) is $365.47a^{(12)}_{\overline{n}|} = \frac{365.47}{i^{(12)}}(1-v^n)$. Comparing these gives

$$v^n = 245.47/365.47.$$

The value from (iii) is $17866v^n$ while the value from (iv) is simply X, so $X = 17866v^n = 12000.02$. ■

Exercise 3.1.6. C.1.2.2 in [Gerber]

Solution. We have $i^{(4)} = 0.12$ so $v = \frac{1}{1.03^4}$. Calculating the present value gives

$$100000 = \left(v^{\frac{1}{2}} + 2v + 3v^{\frac{3}{2}} + 4v^2 + 5v^{\frac{5}{2}} + 6v^3\right) X$$

and a calculator will tell us that this is

$$100000 = 16.3184X$$

so $X = 6128.05$. ■

Exercise 3.1.7. C.1.2.7 in [Gerber]

Solution. It reduces to solving linear equations

3.2 Force of mortality

Definition 3.2.1. The *force of mortality* of (x) at age $x + t$ is defined by

$$\mu_{x+t} = \frac{g(t)}{1 - G(t)} = -\frac{\mathrm{d}}{\mathrm{d}t} \ln\left(1 - G(t)\right).$$

This doesn't tell us much about what μ_{x+t} really is. In more understandable terms, $\mu_{x+t}\ \mathrm{d}t$ is the probability of dying between t and $t + \mathrm{d}t$, provided you make it to t.

Now the interesting question is how does μ_{x+t} behave, and there has been lots of models for this:

- De Moivre postulated that there is a limiting age ω, and that for any $0 < t < \omega - x$,

$$\mu_{x+t} = \frac{1}{\omega - x - t}.$$

- Gompertz postulated that μ_{x+t} is exponential, which is

$$\mu_{x+t} = Bc^{x+t}.$$

- Makeham improved the guess to

$$\mu_{x+t} = A + Bc^{x+t}.$$

- Weibull argues that it is polynomial, not exponential, so

$$\mu_{x+t} = k(x+t)^n.$$

Let's try to find some of the constants that work best.

For De Moivre we found that $\omega = 2\mathrm{E}(T)$. We then tried to find the constants for Gompertz's model, assuming $\mathrm{E}(T) = 80$ and $\sigma(T) = 15$. Unfortunately I lost the results.

3.3 Life tables

In a life table, we have:

- $T(x)$ is the future life time of a life aged (x).
- $K(x)$ is the curtate future lifetime (the number of completed years of life).

We have

$$\Pr(K = k) = {}_k p_x q_{x+k} = \Pr(k \leqslant T \leqslant k+1,$$

$$e_x = \sum_{k=1}^{\infty} k {}_k p_x q_{x+k} = \sum_{k=1}^{\infty} {}_k p_x$$

Define $S(x) = T(x) - K(x) \in [0, 1)$. Since life tables only give us $K(x)$, assumptions for $S(x)$ are used if $T(x)$ is needed

- Linearity of ${}_u q_k$ for $u \in (0, 1)$, so ${}_u q_k = u q_x$.
- μ_{x+u} is constant for $u \in (0, 1)$ as $\mu_{x+1/2}$; this gives ${}_u p_x = p_x^u$.

We can interpret $\mu_{x+t}\,\mathrm{d}t$ as the probability of surviving to $x+t$ and dying within $\mathrm{d}t$. Also, using the life table, we have

$$\mu_{x+t} \approx \frac{\ell_{x+t}}{\ell_x} q_{x+t}.$$

3.4 Life insurances and life annuities

Let's calculate how much should a life insurance be worth.

Model 3.4.1 (Single-premium whole life insurance)**.** You pay me r today, and I will pay $1 at the end of your year of death.

- For this to be fair, r has to be the expected value of $1 payable at time $K+1$, that is,

$$r = A_x = \mathrm{E}(v^{K+1}).$$

- Now a reminder: if X has p.d.f. $f_X(x)$ then

$$\mathrm{E}(g(X)) = \int g(x) f_X(x)\,\mathrm{d}x.$$

- Whoops, we actually have a discrete variable here, so we have

$$\mathrm{E}(v^{K+1}) = \sum_{k=0}^{\infty} v^{k+1} \Pr(K = k) = \sum_{k=0}^{\infty} v^{k+1} {}_k p_x q_{x+k}.$$

On the actual life table, we got $(r, \sigma(r)) = (786.74, 1048.23)$ for males and $(589.80, 754.49)$ for females.

Model 3.4.2 (Single-premium life annuity)**.** You pay me r today, and I will pay $1 at the start of each year until your death

- This is $\ddot{a}_x$ which is the expected present value of an annuity due of $1 payable at times $0, 1, 2, \ldots, k$ which has present value $\ddot{a}_{\overline{k}|} = \frac{1 - v^{k+1}}{d}$.

- Therefore,

$$\ddot{a}_x = \mathrm{E}\left(\frac{1-v^{K+1}}{d}\right) = \frac{1-\mathrm{E}(v^{K+1})}{d} = \frac{1-A_x}{d}.$$

This gives an identity

$$d\ddot{a}_x + A_x = 1,$$

which means that a payment of d while alive plus a payment of 1 at death is worth $1 today. It is also possible to calculate the variance of $\ddot{a}_{\overline{K}|}$:

$$\mathrm{Var}\left(\ddot{a}_{\overline{K}|}\right) = \left(\frac{1-v^{K+1}}{d}\right) = \frac{1}{d^2}\mathrm{Var}(v^{K+1})$$

These single-premium models do not make much sense; if you can afford $\ddot{a}_x$ at present you probably do not need a life insurance anyway. A more realistic model is a life insurance with periodic premiums.

Model 3.4.3. The loss L to the insurance company where Π is the periodic premiums and C is the death benefit is

$$L = \Pi\ddot{a}_{\overline{K}|} - Cv^{K+1}.$$

Let's first find the equilibrium point: $\mathrm{E}(L) = 0$:

$$0 = \Pi\mathrm{E}(\ddot{a}_{\overline{K}|}) - C\mathrm{E}(v^{K+1}) = \Pi\ddot{a}_x - CA_x,$$

so the periodic premium should be

$$\Pi = \frac{CA_x}{\ddot{a}_x} = dC\frac{A_x}{1-A_x}$$

This gives a pretty good model. Some possible extensions include

- Expense and risk loading: to cover the cost of hiring people and the risk of high-valued insurances.
- Re-insurance: insure yourself with another company.
- Multiple lives: family insurances, etc.
- Deferred annuities: payments start later.
- Variable annuities: payments change over time.

Exercise 3.4.4. Price a life insurance policy for your entire family.

Solution. Using a $100k policy, my age = 18, my mom's age = 50, my dad's age = 52 gives

$$\$406.63 + \$1400.86 + \$1956.91 = \$3764.40 \quad \blacksquare$$

3.5 Recursive formulas

There is one tool that can help simplify calculations: recursive formulas. For single-premium models, we have

$$A_x = vq_x + vA_{x+1}p_x.$$

Here, vq_x stands for the expected money you will earn from dying this year, and $vA_{x+1}p_x$ is the cost of life insurance for $(x+1)$ if you make it another year.

For annuities, the formula is even simpler:

$$\ddot{a}_x = 1 + v\ddot{a}_{x+1}p_x.$$

Here 1 is the $1 you receive this year, and $vA_{x+1}p_x$ is expected (factoring in your chance to live) the cost of life annuity starting next year.

Project for this week: Single-premium insurance product of a joint annuity with reduced survivor benefit plan: Life annuity for two lives (x) and (y), pays $1 until the first death, then two thirds until survivor's death

3.6 Utility theory

Risk-based pricing: higher-valued insurances carry much more risk, so the pricing should reflect the risk. To model this, we will consider *Utility Theory*.

Main idea: most of us are risk-averse (we don't want to take risk) to an extent that depends on your existing wealth: we are more squeamish about risking higher losses even if the expected loss is the same. (Ex: high-risk low-probability events)

This is modeled by a *utility function* $u(w)$, which measures the utility you derive from wealth w (which can be positive or negative.) If faced with two risks X, Y, you'd be indifferent between them if

$$\mathrm{E}(u(X)) = \mathrm{E}(u(Y)).$$

(Note that this is invariant under a linear transformation of $u(w)$.)

Clearly, the utility function is very subjective, and differs very much for different people. You might want to ask: how to ball-park $u(w)$ for yourself?

Example 3.6.1. Suppose that you're indifferent to not making or losing money ($u(0) = 0$) and losing $\$1000$ has utility -1000. Now we play a game:

I make you toss a coin, if it shows up as heads, you lose $1000, and tails you keep everything. How much are you willing to pay (Π) to insure against this loss? Then,

$$u(-\Pi) = \mathrm{E}(u(-\Pi)) = \mathrm{E}(u(X)) = -500.$$

Now play this game again and again with varying probabilities and amounts to get more points on the curve. This applies to positive wealth too:

How much would you wager (w) for the chance to win $1000 on a coin flip. Then,

$$\frac{1}{2}u(1000 - w) + \frac{1}{2}u(-w) = 0.$$

In general, $u'(w) > 0$ (more wealth is better) and $u''(w) < 0$ (diminishing utility at the margin).

A simple model of this is an exponential function:

$$u(w) = \frac{1}{a}(1 - e^{-aw})$$

where a controls your risk aversion.

Model 3.6.2 (Insurance with risk loading)**.** We want $\mathrm{E}(u(L)) = u(0)$ where L is the loss variable. This translates to

$$\sum_{k=0}^{100} u(\Pi - Cv^{k+1}) \Pr(K = k) = 0.$$