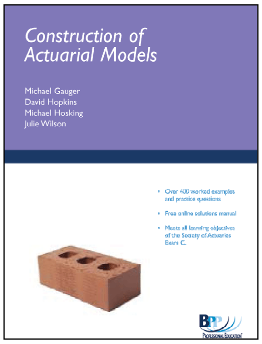



# Construction of Actuarial Models

By J. Wilson, M. Hosking, D. Hopkins, and M. Gauger

Published by BPP Professional Education

## *Solutions to practice questions – Chapter 5*

### *Solution 5.1*

Possibilities include:

- the time till you complete the actuarial exams
- the time till the next major earthquake occurs.

---

### *Solution 5.2*

The **survival function** for the time to an event is the probability that the event has not occurred by a give time.

An **empirical estimate** of this function is a numerical estimate obtained directly from a data set of observed times.

---

### *Solution 5.3*

Left and right censoring are both special cases of interval censoring. The general point is that data are censored whenever we are prevented from quantifying precisely the (complete) lifetimes of individuals.

---

### *Solution 5.4*

The more data we can base our estimate upon, the less error will be involved. In statistical terms, the standard error of the estimator will be smaller, the larger the sample size upon which it is based.

In other words, the more data we use, the closer our estimate is likely to be to the true (but unknown) parameter that we are estimating.

---

### *Solution 5.5*

If the individuals experience independent mortality, the number of deaths will have a binomial distribution with parameters 20 and 0.015. In particular, the mean number of deaths will be $20 \times 0.015 = 0.3$ and the standard deviation will be $\sqrt{20 \times 0.015 \times 0.985} = 0.544$ .

## *Solution 5.6*

(a) The empirical estimate of $S(3.1)$ is $\hat{S}(3.1)$: the observed proportion surviving from time 0 to time 3.1. We look up the value directly from the table, which is 0.4.

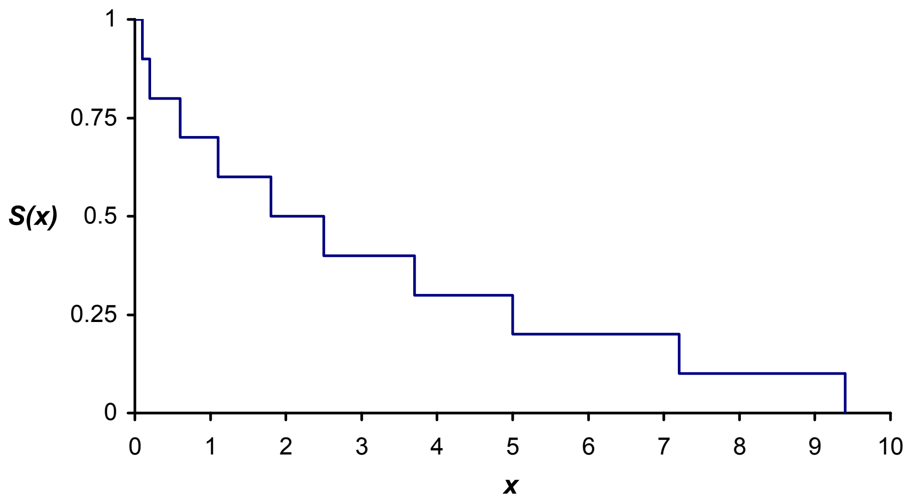
(b) The empirical estimate of $F(7.5)$ is:

$$\hat{F}(7.5) = 1 - \hat{S}(7.5) = 1 - 0.1 = 0.9$$

---

## *Solution 5.7*

$\hat{S}(x)$ is plotted against $x$ in the following diagram.




---

## *Solution 5.8*

We can work backwards from the distribution function $F(x)$ to the probability function $f(x)$, if the answer is not intuitively obvious. Here we have:

$$f(x) = \begin{cases} 0.1 & x = 0.1,\ 0.2,\ 0.6,\ 1.1,\ 1.8,\ 2.5,\ 3.7,\ 5.0,\ 7.2,\ 9.4 \\ 0 & \text{elsewhere} \end{cases}$$

## *Solution 5.9*

The expression $0.092x - 0.048$ is the equation of the straight line joining the points $(1, 0.044)$ and $(2, 0.136)$.

The completed table looks like this:

| $x$ **(hours after 9am)** | $F(x)$ |
|---|---|
| $0 \le x < 1$ | $0.044x$ |
| $1 \le x < 2$ | $0.092x - 0.048$ |
| $2 \le x < 3$ | $0.136x - 0.136$ |
| $3 \le x < 4$ | $0.14x - 0.148$ |
| $4 \le x < 5$ | $0.13x - 0.108$ |
| $5 \le x < 7$ | $0.11x - 0.008$ |
| $7 \le x < 10$ | $0.0793x + 0.207$ |

---

## *Solution 5.10*

The table gives us an estimate of the probability that a given ticket will be sold by 3:00pm (*ie* $x = 6$), which is:

$$\hat{F}(6) = 0.11 \times 6 - 0.008 = 0.652$$

---

## *Solution 5.11*

The estimates are:

- $P\left(X > 6.25 \middle| X > 6\right) = \dfrac{1 - (0.11 \times 6.25 - 0.008)}{1 - (0.11 \times 6 - 0.008)} = \dfrac{0.3205}{0.348} = 0.921$
- $P\left(X > 8 \middle| X > 6\right) = \dfrac{1 - (0.0793 \times 8 + 0.207)}{1 - (0.11 \times 6 - 0.008)} = \dfrac{0.1586}{0.348} = 0.456$

---

## *Solution 5.12*

Assuming uniform ticket sales within each time interval:

$$1 \times 0.11 + 1 \times 0.0793 = 0.1893$$

## *Solution 5.13*

Revision! The conditional probability that a "life" aged 1 hour survives to age $x$ (*ie* for a further $x-1$ hours) is:

$$ {}_{x-1}p_1 $$

In terms of International Actuarial Notation all we can write here is:

$$ \frac{{}_{x}q_0}{{}_{4}q_0} $$

---

## *Solution 5.14*

There is **left truncation** at time 1 (March 1), because we do not know how many patients died in the first month. It is therefore impossible for us to estimate the probability of surviving over the first month.

There is **interval censoring** over all times between time 1 (March 1) and time 11 (December 31), with monthly intervals. So we only know individual lifetimes (between these times) to the nearest month.

There is **right censoring** at time 11 (involving the three people still alive at that time), because we do not know how long they survive beyond time 11.

---

## *Solution 5.15*

We can calculate the proportion of those alive at time 1 who survive to time $x$ $(1 < x \le 11)$, as follows.

| **Lifetime** $= x$ | **Number surviving to time $x$** $= n_x$ | **Proportion of those alive at time 1 surviving to time $x$** $= \frac{n_x}{n_1}$ |
|---|---|---|
| 1 | 15 | 1.0000 |
| 2 | 14 | 0.9333 |
| 3 | 14 | 0.9333 |
| 4 | 12 | 0.8000 |
| 5 | 9 | 0.6000 |
| 6 | 8 | 0.5333 |
| 7 | 8 | 0.5333 |
| 8 | 8 | 0.5333 |
| 9 | 5 | 0.3333 |
| 10 | 3 | 0.2000 |
| 11 | 3 | 0.2000 |

We have therefore been able to estimate values of $\frac{S(x)}{S(1)} = P(X > x | X > 1)$ for all integer values of $x$ between 1 and 11 inclusive.

## *Solution 5.16*

Now:

$$S(6) = S(1)\,P\left(X > 6 \middle| X > 1\right)$$

$$\therefore \qquad \hat{S}(1) = \frac{S(6)}{\hat{P}\left(X > 6 \middle| X > 1\right)} = \frac{0.5}{0.5333} = 0.9375$$

Now:

$$\hat{S}(4) = \hat{S}(1)\,\hat{P}\left(X > 4 \middle| X > 1\right) = 0.9375 \times 0.8 = 0.75$$

## *Solution 5.17*

There is **interval censoring** at yearly intervals throughout the age range, as we only know lifetimes to the nearest year.

## *Solution 5.18*

We need to build up a table of cumulative numbers of deaths up to each age (this will be each *integer* age because of the interval censoring).

For example, one male dies aged 0 last birthday: so the number of deaths up to age 1 is 1. The cumulative number of deaths stays at 1 for all integer ages up to and including age 32. By age 33 there has been a total of 2 deaths, so the count increases to 2 from age 33 onwards. By continuing in similar fashion we obtain the following table.

| **Inclusive integer range of age** $= x$ | **Number observed to die by age *x*** $= d_x$ | **Proportion of lives observed to die by age *x*** $= \frac{d_x}{l_0}$ |
|---|---|---|
| [1, 32] | 1 | 0.04 |
| [33, 58] | 2 | 0.08 |
| [59, 62] | 3 | 0.12 |
| 63 | 4 | 0.16 |
| [64, 65] | 5 | 0.20 |
| [66, 68] | 7 | 0.28 |
| 69 | 8 | 0.32 |
| [70, 71] | 10 | 0.40 |
| 72 | 13 | 0.52 |
| [73, 75] | 15 | 0.60 |
| [76, 77] | 19 | 0.76 |
| [78, 81] | 22 | 0.88 |
| [82, 89] | 24 | 0.96 |
| 90 | 25 | 1.00 |

We have therefore been able to estimate values of $F(x)=\Pr(X \le x)$ for all integer values of $x$ up to age 90 inclusive.

---

## *Solution 5.19*

From the data we estimate:

$$\frac{\hat{F}(70)}{\hat{F}(90)} = 0.4$$

Using the additional information:

$$\frac{1-S(70)}{\hat{F}(90)} = \frac{1-0.62}{\hat{F}(90)} = 0.4$$

$$\therefore \quad \hat{F}(90) = \frac{0.38}{0.4} = 0.95$$

So:

$$\begin{aligned}\hat{S}(80) &= 1-\hat{F}(80) \\ &= 1-\frac{\hat{F}(80)}{\hat{F}(90)}\hat{F}(90) \\ &= 1-0.88 \times 0.95 \\ &= 0.164\end{aligned}$$

---

## *Solution 5.20*

False.

Double censoring is where we have both left censoring and right censoring, *ie* for some data values we only know that the exact value is below a certain value and for some data values we only know that the exact value is above a certain level.

Interval censoring is where, for some data values we only know that the exact value falls in a certain range, *ie* it lies between two values.