ON THE BALDUCCI HYPOTHESIS

Ho Kuen Ng

Abstract

This article investigates the simplicity of the Balducci hypothesis, and compares the fractional-age death probability given by three widely used assumptions.

1. Introduction

There are three widely used assumptions for fractional-age mortality, namely, uniform distribution of deaths, the Balducci hypothesis, and constant force of mortality [1, 2]. Respectively, they state that, for any x and any $0 \leq t \leq 1$,

(uniform distribution of deaths) $\quad {}_tq_x = tq_x$

(the Balducci hypothesis) $\quad {}_{1-t}q_{x+t} = (1-t)q_x$

(constant force of mortality) $\quad \mu_{x+t} = \text{constant}$.

An important application of these assumptions is in the construction of mortality tables. It has been observed that the Balducci hypothesis is the most practical of the three, giving expressions most amenable to calculations [1]. The first part of this article will study this simplicity, showing that the Balducci hypothesis is the necessary and sufficient condition for the exposure equation to be linear.

Most textbooks just observe that the three assumptions give very similar numerical values of fractional-age death probability, and verify such observation by a few numerical examples. In the second part of this article, we will give a mathematical justification of the above claim by calculating the absolute difference among the three assumptions.

2. The criteria of 'simplicity'

As mentioned in the introduction, one of the important applications of an assumption for fractional ages is in the construction of mortality tables. Suppose that, in a mortality study, A lives are under observation at age x, m lives enter the study at ages $x + r_i$, $i = 1, 2, ..., m$, n lives withdraw from the study at ages $x + s_j$, $j = 1, 2, ..., n$, where r_i and s_j are between 0 and 1, and D deaths are observed between ages x and x+1 among those in the study. We have the equation [1]

$$Aq_x + \sum_{i=1}^{m} {}_{1-r_i}q_{x+r_i} - \sum_{j=1}^{n} {}_{1-s_j}q_{x+s_j} = D.$$

Let us call this the exposure equation.

Agreeing that linear equations are the simplest equations, we postulate two criteria of simplicity:

(1) When regarded as an equation in q_x, the exposure equation is linear.

(2) When regarded as an equation in r_i, $i = 1, 2, ..., m$, s_j, $j = 1, 2, ..., n$, the exposure equation is linear.

Theorem: A necessary and sufficient condition for the criteria of simplicity is the Balducci hypothesis.

Proof:

(sufficiency) By the Balducci hypothesis, we have ${}_{1-r_i}q_{x+r_i} = (1 - r_i)q_x$ and ${}_{1-s_j}q_{x+s_j} = (1 - s_j)q_x$. The exposure equation becomes

$Aq_x + \sum_{i=1}^{m}(1 - r_i)q_x - \sum_{j=1}^{n}(1 - s_j)q_x = D$, i.e. $(A + \sum_{i=1}^{m}(1 - r_i) - \sum_{j=1}^{n}(1 - s_j))q_x = D$. Conditions (1) and (2) then follow.

(necessity) With m = 1, n = 0, and r_1 = r which is any number between 0 and 1, the exposure equation becomes $Aq_x + {}_{1-r}q_{x+r} = D$, which must satisfy conditions (1) and (2). By (1), we have ${}_{1-r}q_{x+r} = q_x f(r) + g(r)$, where f and g depend on r only. Thus the exposure equation is $Aq_x + q_x f(r) + g(r) = D$. By (2), we can write $f(r) = \alpha r + \beta$, $g(r) = \gamma r + \delta$, where $\alpha, \beta, \gamma, \delta$ are constants. Then ${}_{1-r}q_{x+r} = q_x(\alpha r + \beta) + \gamma r + \delta$. For r = 0, $q_x = q_x\beta + \delta$. For r = 1, $0 = q_x\alpha + q_x\beta + \gamma + \delta = q_x\alpha + \gamma + q_x$. Therefore $q_x\alpha + \gamma = -q_x$. Then ${}_{1-r}q_{x+r} = q_x\alpha r + q_x\beta + \gamma r + \delta = q_x\beta + \delta + r(q_x\alpha + \gamma) = q_x + r(-q_x) = (1 - r)q_x$, which is the Balducci hypothesis.

3. Error bounds

Since the assumption of uniform distribution of deaths is usually regarded as the most reasonable one, we will use it as a measuring standard and calculate the error bound of the other two assumptions relative to it.

Let ${}_tq_x^{(1)}$, ${}_tq_x^{(2)}$, ${}_tq_x^{(3)}$ be the probability of death calculated according to the assumptions of uniform distribution of deaths, Balducci, and constant force of mortality respectively.

From [1, 2], we have

${}_tq_x^{(1)} = tq_x$,

${}_tq_x^{(2)} = \frac{tq_x}{1-(1-t)q_x}$,

${}_tq_x^{(3)} = 1 - e^{-\mu t} = 1 - (1-q_x)^t$.

We first compare ${}_tq_x^{(1)}$ and ${}_tq_x^{(2)}$, for any x such that $0 < q_x < 1$, and any $0 \le t \le 1$.

${}_tq_x^{(2)} - {}_tq_x^{(1)} = \frac{tq_x}{1-(1-t)q_x} - tq_x = \frac{t(1-t)q_x^2}{1-(1-t)q_x} \ge 0.$

We denote this difference by f(t). We note that f(t) is continuously differentiable in (0, 1), and that

$$f'(t) = \frac{(1-(1-t)q_x)(1-2t)q_x^2 - t(1-t)q_x^3}{(1-(1-t)q_x)^2}$$

.

Setting $f'(t) = 0$ and restricting to the interval (0, 1), we see that f(t) has only one critical point, namely at $t_0 = 1 - \frac{1-\sqrt{1-q_x}}{q_x}$. As $f(0) = f(1) = 0$, and $f(t) \ge 0$, t_0 must give a maximum in the interval [0, 1]. Since $f(t_0) = (1-\sqrt{1-q_x})^2$, we have

$(1-\sqrt{1-q_x})^2 \ge {}_tq_x^{(2)} - {}_tq_x^{(1)} \ge 0$, providing upper and lower bounds for the difference ${}_tq_x^{(2)} - {}_tq_x^{(1)}$.

Using the binomial series, we have $(1-\sqrt{1-q_x})^2 = 1 + 1 - q_x - 2\sqrt{1-q_x}$

$= 2 - q_x - 2(1 + \frac{1}{2}(-q_x) + \frac{1}{2}(\frac{1}{2})(-\frac{1}{2})(-q_x)^2 + \sum_{k=3}^{\infty} \frac{1}{k!}(\frac{1}{2})(\frac{1}{2}-1)(\frac{1}{2}-2)...(\frac{1}{2}-k+1)(-q_x)^k)$

$= \frac{1}{4}q_x^2 - 2\sum_{k=3}^{\infty} \frac{1}{k!}(\frac{1}{2})(\frac{1}{2}-1)(\frac{1}{2}-2)...(\frac{1}{2}-k+1)(-q_x)^k$

$= \frac{1}{4}q_x^2 - 2(\frac{1}{2})(\frac{-1}{2})(\frac{1}{2})\sum_{k=3}^{\infty}(\frac{1}{3.4...k})(\frac{1}{2}-2)...(\frac{1}{2}-k+1)(-q_x)^k$

$\le \frac{1}{4}q_x^2 + \frac{1}{4}\sum_{k=3}^{\infty} q_x^k = \frac{1}{4}q_x^2 + \frac{1}{4}q_x^3(\frac{1}{1-q_x}) = \frac{1}{4}q_x^2(1 + \frac{q_x}{1-q_x}) \le \frac{1}{2}q_x^2$ if $q_x \le \frac{1}{2}$.

Thus we conclude that the maximum difference between ${}_tq_x^{(2)}$ and ${}_tq_x^{(1)}$ is $O(q_x^2)$, which is of an order of magnitude lower than that of q_x.

Now we compare ${}_tq_x^{(3)}$ with ${}_tq_x^{(1)}$, for any x with $q_x < 1$, and any $0 \le t \le 1$. Again we make use of the binomial series.

$${}_tq_x^{(3)} = 1 - e^{-\mu t} = 1 - (1 - q_x)^t$$

$$= 1 - (1 + t(-q_x) + \tfrac{1}{2!}t(t-1)(-q_x)^2 + \textstyle\sum_{k=3}^{\infty} \tfrac{1}{k!}t(t-1)(t-2)...(t-k+1)(-q_x)^k)$$

$$= tq_x - \tfrac{1}{2}t(t-1)q_x^2 - \textstyle\sum_{k=3}^{\infty} \tfrac{1}{k!}t(t-1)(t-2)...(t-k+1)(-q_x)^k$$

$$= {}_tq_x^{(1)} - \tfrac{1}{2}t(t-1)q_x^2 - \textstyle\sum_{k=3}^{\infty} \tfrac{1}{k!}t(t-1)(t-2)...(t-k+1)(-q_x)^k .$$

It follows that ${}_tq_x^{(3)} - {}_tq_x^{(1)} = -\frac{1}{2}t(t-1)q_x^2 - \sum_{k=3}^{\infty} \frac{1}{k!}t(t-1)(t-2)...(t-k+1)(-q_x)^k$ and $|{}_tq_x^{(3)} - {}_tq_x^{(1)}| \le \frac{1}{2}|t(t-1)|q_x^2 + \frac{1}{2}|t(t-1)||\sum_{k=3}^{\infty}(\frac{1}{3.4...k})(t-2)...(t-k+1)(-q_x)^k|$.

Note that $|t(t-1)| \le \frac{1}{4}$. We then have $|{}_tq_x^{(3)} - {}_tq_x^{(1)}| \le \frac{1}{8}q_x^2 + \frac{1}{8}|\sum_{k=3}^{\infty} q_x^k| = \frac{1}{8}q_x^2 + \frac{1}{8}q_x^3(\frac{1}{1-q_x})$ $= \frac{1}{8}q_x^2(1 + \frac{q_x}{1-q_x}) \le \frac{1}{4}q_x^2$ if $q_x \le \frac{1}{2}$.

From the foregoing calculations, it also follows that $|{}_tq_x^{(2)} - {}_tq_x^{(3)}| \le \frac{3}{4}q_x^2$ if $q_x \le \frac{1}{2}$.

In the following tables, we list the absolute differences of ${}_tq_x^{(1)}$, ${}_tq_x^{(2)}$ and ${}_tq_x^{(3)}$ for various q's and t's, and compare them with the upper bounds which we have derived. We see that our bounds are much larger than the actual differences, and that they remain valid even beyond the predicted range of $q_x \le \frac{1}{2}$, the reason being that we have used quite conservative estimates in our calculations. But the main point is that for $q_x \le \frac{1}{2}$, which covers most of the human life span, the error introduced in substituting one assumption for another is of a lower order of magnitude than q_x .

Reference:

[1] Batten, R. W., Mortality Table Construction, 1978, Prentice-Hall.

[2] Bowers, N. L., Jr., et al, Actuarial Mathematics, 1987, Society of Actuaries.

ABSOLUTE DIFFERENCE BETWEEN UNIFORM DISTRIBUTION OF DEATHS AND THE BALDUCCI HYPOTHESIS

T =	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	.5×Q×Q
Q = 0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.005000
Q = 0.1	0.000000	0.000969	0.001739	0.002258	0.002553	0.002632	0.002500	0.002165	0.001633	0.000909	0.000000	0.020000
Q = 0.2	0.000000	0.004390	0.007619	0.009767	0.010909	0.011111	0.010435	0.008936	0.006667	0.003673	0.000000	0.045000
Q = 0.3	0.000000	0.011096	0.018947	0.023924	0.026341	0.026471	0.024545	0.020769	0.015319	0.008351	0.000000	0.080000
Q = 0.4	0.000000	0.022500	0.037647	0.046667	0.050526	0.050000	0.045714	0.038182	0.027826	0.015000	0.000000	0.125000
Q = 0.5	0.000000	0.040909	0.066667	0.080769	0.085714	0.083333	0.075000	0.061765	0.044444	0.023684	0.000000	0.180000
Q = 0.6	0.000000	0.070435	0.110769	0.180345	0.135000	0.128571	0.113684	0.092195	0.065455	0.034468	0.000000	0.245000
Q = 0.7	0.000000	0.119189	0.178182	0.201765	0.202759	0.188462	0.163333	0.130253	0.091163	0.047419	0.000000	0.320000
Q = 0.8	0.000000	0.205714	0.284444	0.305455	0.295385	0.266667	0.225882	0.176842	0.121905	0.062609	0.000000	0.405000
Q = 0.9	0.000000	0.383684	0.462857	0.459730	0.422609	0.368182	0.303750	0.233014	0.158049	0.080110	0.000000	0.500000

ABSOLUTE DIFFERENCE BETWEEN UNIFORM DISTRIBUTION OF DEATHS AND CONSTANT FORCE OF MORTALITY

T =	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	.25×Q×Q
Q = 0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.002500
Q = 0.1	0.000000	0.000481	0.000852	0.001114	0.001268	0.001317	0.001260	0.001098	0.000834	0.000467	0.000000	0.010000
Q = 0.2	0.000000	0.002067	0.003648	0.004752	0.005390	0.005573	0.005310	0.004612	0.003488	0.001948	0.000000	0.022500
Q = 0.3	0.000000	0.005039	0.008850	0.011477	0.012960	0.013340	0.012656	0.010944	0.008241	0.004582	0.000000	0.040000
Q = 0.4	0.000000	0.009800	0.017120	0.022083	0.024807	0.025403	0.023978	0.020632	0.015460	0.008554	0.000000	0.062500
Q = 0.5	0.000000	0.016967	0.029449	0.037748	0.042142	0.042893	0.040246	0.034428	0.025651	0.014113	0.000000	0.090000
Q = 0.6	0.000000	0.027556	0.047447	0.060342	0.066855	0.067544	0.062920	0.053447	0.039590	0.021617	0.000000	0.122500
Q = 0.7	0.000000	0.043432	0.073997	0.093155	0.102199	0.102277	0.094407	0.079488	0.058322	0.031617	0.000000	0.160000
Q = 0.8	0.000000	0.068660	0.115220	0.142966	0.154694	0.152786	0.139269	0.115869	0.084094	0.045076	0.000000	0.202500
Q = 0.9	0.000000	0.115672	0.189043	0.228813	0.241893	0.233772	0.208811	0.170474	0.121511	0.064107	0.000000	0.250000

ABSOLUTE DIFFERENCE BETWEEN THE BALDUCCI HYPOTHESIS AND CONSTANT FORCE OF MORTALITY

T =	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	.75×*Q*×*Q*
Q = 0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.007500
Q = 0.1	0.000000	0.000508	0.000887	0.001144	0.001285	0.001315	0.001240	0.001067	0.000799	0.000442	0.000000	0.030000
Q = 0.2	0.000000	0.002323	0.003972	0.005016	0.005519	0.005538	0.005124	0.004324	0.003178	0.001726	0.000000	0.067500
Q = 0.3	0.000000	0.006057	0.010097	0.012447	0.013382	0.013131	0.011890	0.009826	0.007078	0.003768	0.000000	0.120000
Q = 0.4	0.000000	0.012700	0.020528	0.024584	0.025719	0.024597	0.021736	0.017550	0.012366	0.006446	0.000000	0.187500
Q = 0.5	0.000000	0.023942	0.037217	0.043022	0.043573	0.040440	0.034754	0.027337	0.018794	0.009571	0.000000	0.270000
Q = 0.6	0.000000	0.042878	0.063322	0.070003	0.068145	0.061027	0.050764	0.038748	0.025904	0.012851	0.000000	0.367500
Q = 0.7	0.000000	0.075757	0.104185	0.108610	0.100559	0.086184	0.068927	0.050765	0.032841	0.015803	0.000000	0.480000
Q = 0.8	0.000000	0.137054	0.169224	0.162488	0.140690	0.113880	0.086613	0.060973	0.037851	0.017532	0.000000	0.607500
Q = 0.9	0.000000	0.268012	0.273814	0.230917	0.180716	0.134410	0.094939	0.062540	0.036538	0.016002	0.000000	0.750000