

APPROXIMATE RELATIONSHIPS AMONG SOME  
ESTIMATORS OF MORTALITY PROBABILITIES

Norman L. Johnson

*Department of Statistics  
University of North Carolina  
Chapel Hill, North Carolina 27514*

Institute of Statistics Mimeo Series No. 1098

December, 1976

---

This research was supported by the Army Research Office under Grant  
No. DAH G29-74-C-0030.

APPROXIMATE RELATIONSHIPS AMONG SOME  
ESTIMATORS OF MORTALITY PROBABILITIES

by

N.L. Johnson\*

University of North Carolina at Chapel Hill

1. Introduction

Kuzma [1967] and Drolette [1975] have compared by means of quite extensive numerical calculations, several estimators of  $q$ , the probability of death in an interval (given alive at the beginning of the interval). It is the purpose of this note to give some approximate algebraic relations among some of these estimators. These confirm, in general terms, the conclusions reached from the numerical investigations. They also provide quite good approximations (to four decimal places when  $q < 0.15$ , apparently) to two rather complicated (in appearance) estimators, in terms of estimators of simpler form. This also gives an improved appreciation of the nature of the more complicated estimators.

2. Notation

We use the following notation (as in Elandt-Johnson [1976]).

$N$  = number alive at beginning of interval

$d$  = number of deaths in the interval

$c$  = number "due to be withdrawn" in the interval

$d'$  = number of deaths among these  $c$  individuals

$w$  = number actually withdrawn.

It will be assumed that there are no "lost to follow up," or new entrants during the interval so that  $w+d' = c$ ;  $d' \leq d \leq N - w$ .

---

\*This research was supported by the Army Research Office, under Contract DAH G29-74-C-0030.

The estimators to be discussed are:

$$\hat{q}^{(2)} = \{N' - (N'^2 - 2Nd)^{1/2}\}N^{-1} \quad (\text{see Elveback [1958]}) \quad (1)$$

where  $N' = N - \frac{1}{2}w + \frac{1}{2}d$ ,

$$\hat{q}^{(3)} = d(N - \frac{1}{2}w)^{-1} \quad (\text{the "actuarial" estimator}), \quad (2)$$

$$\hat{q}^{(4)} = 1 - \frac{1}{4}(2N-c)^{-2} \{[d'^2 + 4(2N-c)(2N-w-2d)]^{1/2} - d'\}^2 \quad (\text{Chiang [1961]}) , \quad (3)$$

$$\hat{q}^{(5)} = d(N - \frac{1}{2}c)^{-1} \quad (\text{see Elveback [1958]}). \quad (4)$$

The estimators  $\hat{q}^{(3)}$  and  $\hat{q}^{(5)}$  are much simpler in form than  $\hat{q}^{(2)}$  and  $\hat{q}^{(4)}$ .

### 3. The Approximations

Using the method outlined in the Appendix, the following approximations are obtained.

$$\hat{q}^{(2)} \doteq \frac{1}{2} \left[ \left(1 + \frac{1}{2}\hat{q}^{(3)}\right)^{-1} + \left\{1 + \left(\frac{1}{2} - R\right)\hat{q}^{(3)}\right\}^{-1} \right] \hat{q}^{(3)} \quad (5)$$

where

$$R = N/N' = \left( \frac{N}{N - \frac{1}{2}w} \right) \left(1 + \frac{1}{2}\hat{q}^{(3)}\right)^{-1} \quad (6)$$

$$\hat{q}^{(4)} \doteq \left\{1 - \frac{1}{4}(d'/d)\hat{q}^{(5)}\right\} \hat{q}^{(5)} . \quad (7)$$

Results of some numerical calculations illustrating the accuracy of these approximations, using data from Table 2 of Chiang [1961] are presented in Table 1.

For larger  $\hat{q}$ , (5) may be improved by addition of  $\frac{1}{8}\hat{q}^{(3)}\{\hat{q}^{(3)}(1 + \frac{1}{2}\hat{q}^{(3)})^{-1}\}^3$ , but values of  $\hat{q}$  (>.20, say) for which this is of any importance are so large that validity of assumptions made in derivation of formulae become of overriding importance (and usually uncertain validity).

Table 1.

Approximations to various estimators of  $q$ 

Interval	$\hat{q}^{(2)}$	Approx. (5)	$\hat{q}^{(4)}$	Approx. (7)	$\hat{q}^{(3)}$	$\hat{q}^{(5)}$
0-1	.2434	.2429	.2425	.2426	.2417	.2436
1-2	.1822	.1821	.1814	.1815	.1810	.1819
2-3	.1033	.1033	.1030	.1030	.1028	.1032
3-4	.0860	.0860	.0858	.0858	.0856	.0859
4-5	.0643	.0643	.0641	.0641	.0641	.0642
5-6	.0584	.0584	.0582	.0582	.0582	.0583
6-7	.0438	.0438	.0438	.0438	.0437	.0438
7-8	.0433	.0433	.0432	.0432	.0431	.0433
8-9	.0337	.0337	.0337	.0337	.0337	.0337
9-10	.0466	.0466	.0465	.04655*	.0464	.0467
10-11	.0439	.0439	.0438	.0438	.0437	.0440
11-12	.0513	.0513	.0511	.0511	.0508	.0513

\* - exactly

Except for the approximation to  $\hat{q}^{(2)}$  in interval 0-1, there is never more than 0.0001 difference between approximation and calculated value (to 4 decimal places). In most cases the values agree to 4 decimal places.

It is interesting to note that for these data  $\hat{q}^{(5)}$  happens to be quite a good approximator for  $\hat{q}^{(2)}$ .

#### 4. Discussion

From (7) it is clear that the approximate value for  $\hat{q}^{(4)}$  is always less than  $\hat{q}^{(5)}$ . This is consistent with numerical results obtained by Drolette [1975]. She also finds that  $\hat{q}^{(2)}$  is greater than  $\hat{q}^{(3)}$ . To see that this is in accordance with our formulae, we note that the right hand side of (5) is

greater than  $\hat{q}^{(3)}$  if and only if  $R > (1 + \frac{1}{2}\hat{q}^{(3)})(1 + \hat{q}^{(3)})^{-1}$ . From the definition of  $R$  in (6), it follows that  $R \geq (1 + \frac{1}{2}\hat{q}^{(3)})^{-1}$  and this is less than  $(1 + \frac{1}{2}\hat{q}^{(3)})$ , so it is possible for the approximate value of  $\hat{q}^{(2)}$  to be less than  $\hat{q}^{(3)}$ , though the least possible value (when  $w = 0$ ) is

$$\begin{aligned} \hat{q}^{(3)} [1 - (\hat{q}^{(3)})^3 \{8 + 4\hat{q}^{(3)} + 2(\hat{q}^{(3)})^2 + (\hat{q}^{(3)})^3\}^{-1}] \\ > \hat{q}^{(3)} \{1 - \frac{1}{8}(\hat{q}^{(3)})^3\} . \end{aligned}$$

While the labor of calculating  $\hat{q}^{(2)}$  and  $\hat{q}^{(4)}$  is not great, the approximations put forward here provide some insight into the nature of these estimators, as well as quite simple formulae for approximate calculation.

In view of the small differences usually found among the numerical values given by the different formulae, and the usually dubious nature of assumptions involved in setting up the likelihood functions underlying  $\hat{q}^{(2)}$  and  $\hat{q}^{(4)}$ , there are good grounds for using simpler formulae (like  $\hat{q}^{(3)}$  or  $\hat{q}^{(5)}$ ) except in very special cases.

#### APPENDIX

For  $|t| < 1$ , we have the expansion

$$\begin{aligned} (1-2t)^{1/2} &= 1 - t - \frac{1}{2}t^2 - \frac{1}{2}t^3 - \frac{5}{8}t^4 - \frac{7}{8}t^5 - \dots \\ &= 1 - \frac{1}{2}t - \frac{1}{2}(t + t^2 + t^3 + t^4 + t^5) - \frac{1}{8}t^4 - \frac{3}{8}t^5 - \dots \\ &\doteq 1 - \frac{1}{2}t - \frac{1}{2}t(1-t)^{-1} . \end{aligned}$$

So

$$1 - (1-2t)^{1/2} \doteq \frac{1}{2}t[1 + (1-t)^{-1}] .$$

This approximation is used in

$$\hat{q}^{(2)} = (N'/N)[1 - \{1 - 2(N/N'^2)d\}^{1/2}]$$

to yield (5).

Similarly, we write (remembering  $d' = c-w$ )

$$\hat{q}^{(4)} = 1 - \frac{1}{4(2N-c)^2} [4(2N-c)^2(1 - \frac{2d-d'}{2N-c}) + 2d'^2 - 4d'(2N-c)\{1 - \frac{2d-d'}{2N-c} + \frac{d'^2}{4(2N-c)^2}\}^{1/2}]$$

and expand the square root, to yield (7).

#### REFERENCES

- Chiang, C.L. [1961]. A stochastic study of the life table and its applications: III. The follow-up studies with the considerations of competing risk, *Biometrics*, 17, 57-78.
- Drollette, M.E. [1975]. The effect of incomplete follow-up, *Biometrics*, 31, 135-144.
- Elandt-Johnson, R.C. [1976]. Various estimators of conditional probabilities of death in follow-up studies: Summary of results, *J. Chron. Dis.*, 29, to appear.
- Elveback, L. [1958]. Actuarial estimation of survivorship in chronic disease, *J. Amer. Statist. Assoc.*, 53, 420-440.
- Kuzma, J.W. [1967]. A comparison of two life table models. *Biometrics*, 23, 51-64.