Statistical approximations to the Multi State model

CHRIS SUTTON, OCTOBER 2023

Last time we covered Markov processes

2-state model

- introduction
- transition intensity
- probabilities
- observational framework
- MLE

general multi-state model

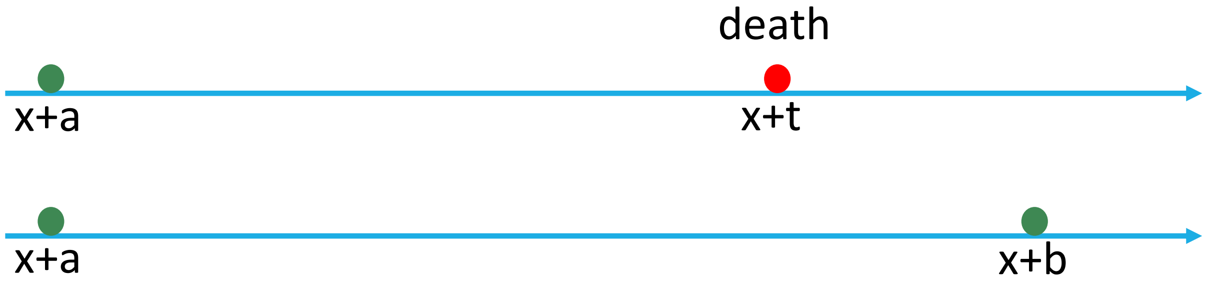
Now we turn to approximations to these models which rely on the assumption of some probability distribution

Topic outline

1. Binomial type models
2. naïve binomial
3. general binomial
4. the actuarial estimate
5. Poisson models
6. Comparing models

2-state model observational framework

- observe a life from age x+a to age x+b (so right censoring taking place)
- two potential outcomes (death during the period or survival to the end)
- two pieces of data collected, both assumed to be samples from random variables (D and V)



D	V
1	t-a
0	b-a

Binomial models

Binomial type models

Much practical actuarial work relies on life tables

- p_x or q_x at integer ages x

Can we find probabilistic models which, with data, will lead to life tables like this?

based on or adapted to a theory of probability; subject to or involving chance variation.

Indian Assured Lives Mortality (2006-08) Ult.

Published Mortality Table, effective 1st April, 2013, within the meaning of regulation 4 of IRDA (Asset, Liabilities and Solvency Margin of Insurers)

Published with the concurrence of IRDA vide it's letter dated 20th February 2013

Age x is defined as age nearest birthday.

Age (x)	Mortality rate(qx)	Age (x)	Mortality rate(qx)
0	0.004445	58	0.009944
1	0.003897	59	0.010709
2	0.002935	60	0.011534
3	0.002212	61	0.012431
4	0.001670	62	0.013414
5	0.001265	63	0.014497
6	0.000964	64	0.015691
7	0.000744	65	0.017009
8	0.000590	66	0.018462
9	0.000492	67	0.020061
10	0.000440	68	0.021819
11	0.000428	69	0.023746
12	0.000448	70	0.025855
13	0.000491	71	0.028159
14	0.000549	72	0.030673
15	0.000614	73	0.033412
16	0.000680	74	0.036394
17	0.000743	75	0.039637
18	0.000800	76	0.043162
19	0.000848	77	0.046991
20	0.000888	78	0.051149
21	0.000919	79	0.055662
22	0.000943	80	0.060558
23	0.000961	81	0.065870
24	0.000974	82	0.071630
25	0.000984	83	0.077876
26	0.000994	84	0.084645
27	0.001004	85	0.091982
28	0.001017	86	0.099930
29	0.001034	87	0.108540
30	0.001056	88	0.117866
31	0.001084	89	0.127963

naïve binomial

observe N (iid) lives age x for exactly 1 year

d = number of death recorded [sample value from a random variable D]

if we assume each life dies with probability q_x and survives with probability $1 - q_x$

then $\hat{q}_x = \frac{d}{N}$ is intuitive and the MLE of q_x

the estimator $\hat{q}_x$ has mean q_x and variance $\frac{1}{N} q_x(1 - q_x)$

this is the **naïve binomial model of mortality**

observations in practice

However,

- we might not observe all lives for the same interval
- there are usually other (non-death) decrements
- there may be increments

general binomial

with these observations in practice and to find an approximation to a multi-state model we need a **general binomial model of mortality**

- in this case to obtain a likelihood function in terms of q_x we need a simplifying assumption for the distribution of T_x in our range x to x+1
- this is usually complicated to implement in practice
- *constructing a likelihood function for the general binomial case is outside the scope of this module*

the "actuarial estimate"

one alternative (to the MLE) estimate of q_x in the general binomial case is called the **actuarial estimate**

this is $\hat{q}_x = \frac{d}{E_x}$

where, E_x is the "initial exposed to risk" given by

$$E_x = E_x^c + \sum_{i=1}^{d} (1-t_i)$$

and $E_x^c = v$ [in our previous 2-state model notation], the observed waiting time

$x+t_i$ is the exact observed age of death of the i^{th} life in year [x, x+1]

approximations

1. If the exact age at death (needed for E_x) is not available, use the approximation
   $E_x \approx E_x^c + \frac{1}{2}d$

2. the actuarial estimate is similar to (but ≠) a moments estimate of q_x under the "Balducci assumption" that the force of mortality is decreasing between integer ages
   - Which of course in practice it does not – this is a weakness of the actuarial estimate approach

Poisson Models

Poisson Models

an alternative model based on a different distribution assumption for T_x

observe N individuals for E_x^c person-years

assume a constant force of mortality μ in the observation window

then the Poisson Model says D follows a Poisson distribution with parameter μE_x^c

that is

$$P[D=d] = \frac{\exp(-\mu E_x^c)\,(\mu E_x^c)^d}{d!}$$

Poisson (continued)

note this cannot be an exact model for mortality as P[D>N] > 0 under this Poisson arrangement

… but it is often a good approximation

The Poisson likelihood function leads to the estimator $\mu = \frac{D}{E_x^c}$ for the constant μ

with $E[\mu] = \mu$ and $Var[\mu] = \frac{\mu}{E_x^c}$

Comparison of models

comparisons

5 models
& estimates

2 types of
framework

3 questions
we will ask

5 models and their associated estimates

2-state model with MLE of μ_x

Poisson model with MLE of μ

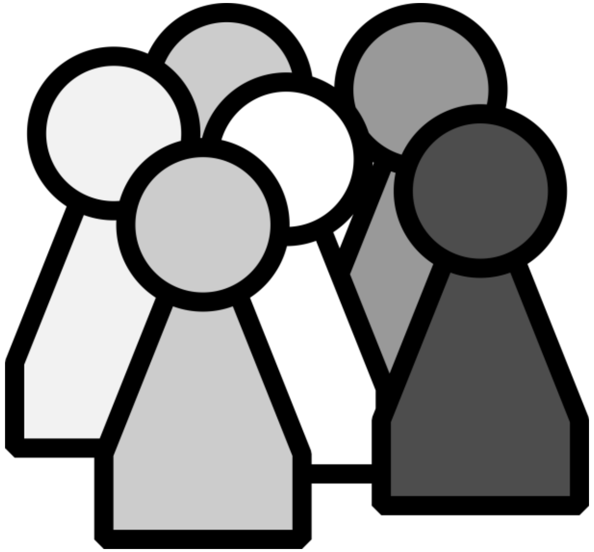
naïve Binomial case MLE of q_x

general Binomial case MLE of q_x

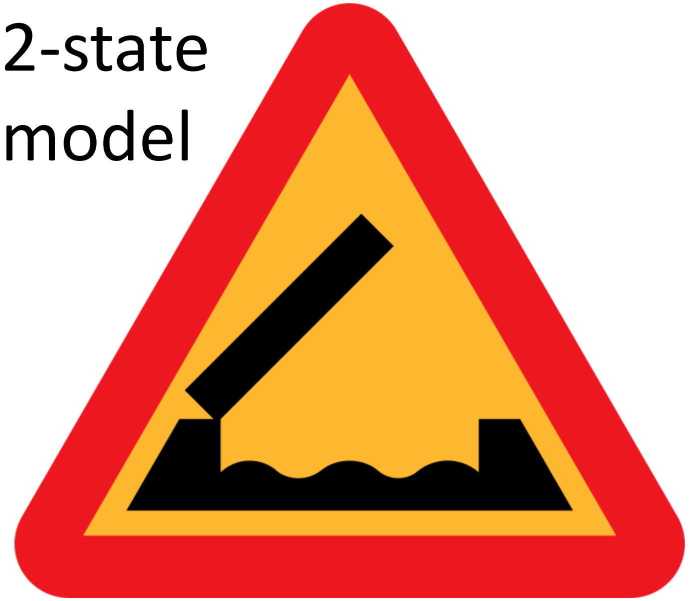
'actuarial estimate' of q_x

2 types of framework

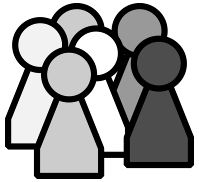
random variable T_x



2-state model



2 types of framework



consider some population

lifetime of an individual in that population is a random variable T_x

we look for that random variable's distribution function $F_x(t)$ or ${}_tq_x$ and survival function $S_x(t)$ or ${}_tp_x$

with ${}_tq_x$ or ${}_tp_x$ we can move to obtain the force of mortality μ_x

consider an individual

that individual may be in one of two states

we seek to understand how they might move between the two states dependent on the transition intensity μ_x

with μ_x we can move to obtain the probabilities ${}_tq_x$ and ${}_tp_x$

3 questions to ask of each model

Q1

- How well does the model represent the underlying process?

Q2

- How easy is it to find and then use the parameters needed in the model?

Q3

- How easy is it to extend the model to more complex processes?

Q1

How well does the model represent the underlying process?

- we are seeking to model the time of death

2-state model represents this precisely and does so by definition

- q_x can be obtained from $1 - \exp(-\tilde{\mu})$
- other models are approximations by comparison

Q1

Binomial type models represent year of death not time of death

- they effectively model the discrete random lifetime K_x rather than T_x
- so binomial estimators have a larger variance than q_x obtained from 2-state model $\tilde{\mu}$
- the approximation is good if μ_x small

The Poisson is also an approximation of the 2-state model unless E_x^c values are fixed (which would be unusual)

- again good approximation if μ_x small

The actuarial estimate assumes μ_{x+t} decreases over the year

Q2 how easy to find parameters needed?

exact date of birth and entry into observation

+

exact date of death or exit from observation

→

2-state μ, Poisson μ, and actuarial estimate q_x can all be calculated exactly

Q2

2-state, Poisson and actuarial estimate parameters can all be found exactly with the full date set data

- although where exact dates not known, all models need some approximations
- can use the "census method" which we'll cover later in this module
- although in cases where some approximation is needed to obtain E_x^c then given that further approximations are necessary to obtain E_x the actuarial method should not be used over the 2-state or Poisson

However, the Binomial q_x parameter will always require further assumptions (e.g. Balducci) to calculate

Q2(b) statistical properties of estimators

Estimator	Consistent?	Unbiased?	Mean available?	Variance available?
2-state MLE	Yes	asymptotically	Yes	Asymptotically (typically need d≥10)
Poisson MLE	Yes	Yes	Yes … tru… ob… le… exp… s…	…terms of …ut using …ed data …o same …ons as 2- …model)
Naïve binomial	Yes		Yes (in terms of true q_x)	Yes (in terms of true q_x)
General binomial	MLE approximate and so less optimal			
Actuarial estimate	close to a moments estimate therefore inferior properties to a MLE			

An estimator is said to be consistent if, as the number of observations increases towards ∞ the sequence of estimates converges on the true value of the parameter

An estimator is said to be unbiased if its expected value equals the true value of the parameter being estimated

Q2(b) statistical properties of estimators

Estimator	Consistent?	Unbiased?	Mean available?	Variance available?
2-state MLE	Yes	asymptotically	Yes	Asymptotically (typically need d≥10)
Poisson MLE	Yes	Yes	Yes (in terms of true μ but using observed data leads to same expressions as 2-state model)	Yes (in terms of true μ but using observed data leads to same expressions as 2-state model)
Naïve binomial	Yes	Yes	Yes (in terms of true q_x)	Yes (in terms of true q_x)
General binomial	MLE approximate and so less optimal			
Actuarial estimate	close to a moments estimate therefore inferior properties to a MLE			

Q3

How easy is it to extend these models to more complex processes?

- with >1 decrement
- with increments

The 2-state model is easily extended to the multi-state model

- estimators have the same form and statistical properties

Poisson extends to multiple decrements but not to increments

Binomial type models are very difficult to extend

comparison questions summary

Historically in human studies where the force of mortality or transition intensity is generally low, actuaries have been comfortable using (initially) Binomial and then Poisson models

Other applications of survival modelling with higher transition intensities will find advantages with Markov process models

In general the best approach is to begin with a model specification which most nearly represents the process being modelled and only make approximations as required for estimation