Binomial Model – Assumptions

- homogeneous population of size N
- lives are independent
- i^{th} life is observed between ages $x+a_i$ and $x+b_i$ for $0 \le a_i < b_i \le 1$
- observations on different lives are made at different points of time
- D_i is indicator random variable for death of i^{th} life and d_i is observed value
- $x+t_i$ is age of death of i^{th} life for $a_i < t_i < b_i$ if the life dies

Binomial Model

$$D_i \sim \text{Bernoulli}\left({}_{b_i-a_i}q_{x+a_i}\right)$$

$$\Pr(D_i = 0) = 1 - {}_{b_i-a_i}q_{x+a_i}$$

$$\Pr(D_i = 1) = {}_{b_i-a_i}q_{x+a_i}$$

$$D = \sum_{i=1}^{N} D_i$$

$$d = \sum_{i=1}^{N} d_i$$

$$\mathrm{E}(D) = \sum_{i=1}^{N} \mathrm{E}(D_i) = \sum_{i=1}^{N} {}_{b_i-a_i}q_{x+a_i} = \sum_{i=1}^{N} \left(1 - {}_{b_i-a_i}p_{x+a_i}\right)$$

$$= \sum_{i=1}^{N} \left({}_{1-a_i}p_{x+a_i} + {}_{1-a_i}q_{x+a_i} - {}_{b_i-a_i}p_{x+a_i}\right)$$

$$= \sum_{i=1}^{N} \left({}_{1-a_i}q_{x+a_i} - {}_{b_i-a_i}p_{x+a_i}\left(1 - {}_{1-b_i}p_{x+b_i}\right)\right)$$

$$= \sum_{i=1}^{N} {}_{1-a_i}q_{x+a_i} - \sum_{i=1}^{N} {}_{b_i-a_i}p_{x+a_i}\,{}_{1-b_i}q_{x+b_i}$$

Binomial Model

– use Balducci assumption :

$$\mathrm{E}(D) = \sum_{i=1}^{N} (1-a_i)\, q_x - \sum_{i=1}^{N} (1-\mathrm{E}(D_i))(1-b_i)\, q_x$$

– use moment matching :

$$\hat{q}_x = \frac{d}{\sum_{i=1}^{N}(1-a_i) - \sum_{i=1}^{N}(1-d_i)(1-b_i)}$$

$$= \frac{d}{\sum_{i=1}^{N}(1-a_i) - \sum_{\text{survivors}}(1-b_i)}$$

$$= \frac{d}{\sum_{\text{survivors}}(b_i-a_i) + \sum_{\text{deaths}}(1-a_i)}$$

– denominator is *initial* exposed to risk E_x

– survivors contribute $b_i - a_i$ (ages $x+a_i$ to $x+b_i$)

– deaths contribute $1-a_i$ (ages $x+a_i$ to $x+1$)

Initial vs Central Exposed to Risk

– $$E_x = \sum_{\text{survivors}} \left(b_i - a_i\right) + \sum_{\text{deaths}} \left(1 - a_i\right)$$

$$= \sum_{\text{survivors}} \left(b_i - a_i\right) + \sum_{\text{deaths}} \left(t_i - a_i\right) + \sum_{\text{deaths}} \left(1 - t_i\right)$$

– sum of first two components is *central* exposed to risk E_x^C

– survivors contribute $b_i - a_i$ (ages $x + a_i$ to $x + b_i$)

– deaths contribute $t_i - a_i$ (ages $x + a_i$ to $x + t_i$)

Initial vs Central Exposed to Risk

- $E_x = E_x^C + \sum_{i=1}^{N} d_i \left(1 - t_i\right)$
- assume deaths occur at age x + 1/2 on average so $t_i = 1/2$
- $E_x \approx E_X^C + d/2$
- reasonable when mortality is low, only one decrement is studied, and investigation period is long
- $\hat{q}_x = \dfrac{d}{E_x} \approx \dfrac{d}{E_x^C + d/2}$

Age Label

- subscript x of E_x and E_x^C refers to those who are ‘aged x last birthday’ within their observation periods
- ‘aged x last birthday’ is an age label
- observed total number of deaths refers to the same age label

Principle of Correspondence

- numerator and denominator of an estimator should have the same age label
- otherwise estimate is not sensible

Binomial Model – Pros & Cons

- convenient way to estimate mortality rate
- only deals with number of deaths but not underlying process
- difficult to extend it to more than one decrement (e.g. illness and death)
- used when information is limited and mortality is low

Binomial Model – Simplified

– suppose all lives are aged exactly x at the start and they are observed for 1 full year

– $D \sim \text{Binomial}\left(N\ ,\ q_x\right)$

– $\Pr\left(D=d\right)=\begin{pmatrix} N \\ d \end{pmatrix} q_x^d \left(1-q_x\right)^{N-d}$

– only one parameter q_x

– maximum likelihood :

$$L=\begin{pmatrix} N \\ d \end{pmatrix} q_x^d \left(1-q_x\right)^{N-d}$$

$$\ln L = \ln\begin{pmatrix} N \\ d \end{pmatrix} + d\ln q_x + \left(N-d\right)\ln\left(1-q_x\right)$$

$$\frac{\partial}{\partial q_x}\ln L = \frac{d}{q_x}-\frac{N-d}{1-q_x} \qquad \frac{\partial^2}{\partial q_x^2}\ln L = -\frac{d}{q_x^2}-\frac{N-d}{\left(1-q_x\right)^2}<0$$

$$\frac{d}{\hat{q}_x}-\frac{N-d}{1-\hat{q}_x}=0 \qquad \hat{q}_x=\frac{d}{N}$$

Binomial Model – Simplified

– $\mathrm{E}\left(\tilde{q}_x\right) = \mathrm{E}\left(\dfrac{D}{N}\right) = \dfrac{Nq_x}{N} = q_x$

– $\mathrm{Var}\left(\tilde{q}_x\right) = \mathrm{Var}\left(\dfrac{D}{N}\right) = \dfrac{Nq_x\left(1-q_x\right)}{N^2} = \dfrac{q_x\left(1-q_x\right)}{N}$

– $\tilde{q}_x$ is normally distributed asymptotically

– mortality rate estimator under original model is roughly treated as normally distributed and its variance is estimated by $\hat{q}_x\left(1-\hat{q}_x\right)/E_x$ with $\hat{q}_x = d/E_x$

Poisson Model – Assumptions

- homogeneous population of size N
- lives are independent
- each life belongs to a certain age group during its observation period
- drop subscript x
- D is total number of deaths and d is observed value
- central exposed to risk E^C is survivors' observation periods plus deaths' observation periods till death
- force of mortality is constant μ for all lives

Poisson Model

- $D \sim \text{Poisson}\left(E^C \mu\right)$
- $\Pr(D = d) = \dfrac{\exp\left(-E^C \mu\right)\left(E^C \mu\right)^d}{d\,!}$
- central exposed to risk is treated as fixed
- maximum likelihood :

$$L = \frac{\exp\left(-E^C \mu\right)\left(E^C \mu\right)^d}{d\,!}$$

$$\ln L = -E^C \mu + d \ln\left(E^C \mu\right) - \ln(d\,!)$$

$$\frac{\partial}{\partial \mu} \ln L = -E^C + \frac{d}{\mu}$$

$$\frac{\partial^2}{\partial \mu^2} \ln L = -\frac{d}{\mu^2} < 0$$

$$-E^C + \frac{d}{\hat{\mu}} = 0$$

$$\hat{\mu} = \frac{d}{E^C}$$

Poisson Model

– $\mathrm{E}(\tilde{\mu}) = \mathrm{E}\left(\dfrac{D}{E^C}\right) = \dfrac{E^C \mu}{E^C} = \mu$

– $\mathrm{Var}(\tilde{\mu}) = \mathrm{Var}\left(\dfrac{D}{E^C}\right) = \dfrac{E^C \mu}{\left(E^C\right)^2} = \dfrac{\mu}{E^C}$

– $\tilde{\mu}$ is normally distributed asymptotically

Poisson Model – Pros & Cons

- sensible way to estimate force of mortality
- Poisson distribution models number of random events within a period of time
- can be extended to more than one decrement
- non-zero probability for more than N deaths, but this is small
- used when mortality is low

Poisson Model – Central Exposed to Risk

- central exposed to risk treated as fixed
- random variable indeed
- acceptable when mortality is low
- or continue investigation until central exposed to risk arrives at pre-specified value
- or replace each death with another independent and identical life

Central Exposed to Risk

- natural quantity and no additional adjustment
- preferred measure to start with
- initial exposed to risk is then computed by $E_x \approx E_x^C + d/2$
- central exposed to risk is the time from Date A to Date B
- Date A is the latest of :

  date reaching age label x

  start of overall investigation

  date of entry
- Date B is the earliest of :

  date reaching age label $x + 1$

  end of overall investigation

  date of exit
- do not count both days
- divide number of days by 365.25 to obtain years

Rate Interval

– a rate interval is a period of one year during which a life's age label value stays the same

– for 'aged x last birthday' the rate interval is $[x, x+1)$

– 'aged x last birthday' is life year rate interval

Important Logic

– under binomial model :

  d refers to those aged x last birthday

  E_x^C refers to those aged x last birthday

  $\hat{q}_x$ estimates q_x (mortality rate at exact age x)

– under Poisson model :

  if d refers to those aged x last birthday

  if E^C refers to those aged x last birthday

  $\hat{\mu}$ estimates $\mu_{x+1/2}$ (force of mortality at exact age x + 1/2)

– rate interval of data is [x , x + 1)

Important Logic

- q-estimate is assigned to START of rate interval of data
- μ-estimate is assigned to MIDDLE of rate interval of data
- this logic can be applied to other age labels and rate intervals

Other Life Year Rate Intervals

– for ‘aged x nearest birthday’ the rate interval is $[\,x-1/2\ ,\ x+1/2\,)$
q-estimate is assigned to age $x-1/2$
μ-estimate is assigned to age x

– for ‘aged x next birthday’ the rate interval is $[x – 1\ ,\ x)$
q-estimate is assigned to age $x – 1$
μ-estimate is assigned to age $x-1/2$

– these are life year rate intervals

Use of Census Data

- imagine we observe a population at every second and count number of lives who are aged x last birthday at every second
- $P_{x,t}$ is number of lives who are aged x last birthday at t
- total time lived by $P_{x,t}$ lives from t to $t + dt$ is $P_{x,t}dt$
- overall investigation is from 0 to $K + 1$
- central exposed to risk is calculated as :

$$E_x^C = \int_0^{K+1} P_{x,t}dt$$

Trapezium Approximation

- in reality an observation is made each year or every few years e.g. policyholders census or country census
- $P_{x,0}$, $P_{x,1}$, $P_{x,2}$, … , $P_{x,K+1}$ are annual census data with age label aged *x* last birthday
- central exposed to risk is approximately calculated as :

$$E_x^C = \int_0^{K+1} P_{x,t} dt \approx \sum_{t=0}^{K} \left(P_{x,t} + P_{x,t+1} \right) / 2$$

- this trapezium approximation can be applied similarly to other periodic data
- it implicitly assumes that events are uniformly spread between two points of time
- it is applied equally to the other two age labels aged *x* nearest birthday and aged *x* next birthday

Principle of Correspondence

- sources of death data and census data may be different
- their age labels may then be different
- by principle of correspondence they have to be the same
- death data carries most information when mortality is low
- so death data determines final age label while census data is adjusted

Principle of Correspondence

- $d_x^{(1)}$, $d_x^{(2)}$, $d_x^{(3)}$ are observed total number of deaths
- $P_{x,t}^{(1)}$, $P_{x,t}^{(2)}$, $P_{x,t}^{(3)}$ are number of lives at time t
- ${}^{(1)}E_x^{\mathrm{C}}$, ${}^{(2)}E_x^{C}$, ${}^{(3)}E_x^{C}$ are central exposed to risk
- (1), (2), (3) denote aged x last birthday, aged x nearest birthday, aged x next birthday
- using trapezium approximation :

$${}^{(1)}E_x^C \approx \sum_{t=0}^{K} \left(P_{x,t}^{(1)} + P_{x,t+1}^{(1)}\right) \Big/ 2$$

$${}^{(2)}E_x^C \approx \sum_{t=0}^{K} \left(P_{x,t}^{(2)} + P_{x,t+1}^{(2)}\right) \Big/ 2$$

$${}^{(3)}E_x^C \approx \sum_{t=0}^{K} \left(P_{x,t}^{(3)} + P_{x,t+1}^{(3)}\right) \Big/ 2$$

Principle of Correspondence

– given $d_x^{(1)}$ but with $P_{x,t}^{(2)}$ or $P_{x,t}^{(3)}$:

$P_{x,t}^{(1)} \approx \left(P_{x,t}^{(2)} + P_{x+1,t}^{(2)}\right)/2$ or $P_{x,t}^{(1)} = P_{x+1,t}^{(3)}$

– given $d_x^{(2)}$ but with $P_{x,t}^{(1)}$ or $P_{x,t}^{(3)}$:

$P_{x,t}^{(2)} \approx \left(P_{x-1,t}^{(1)} + P_{x,t}^{(1)}\right)/2$ or $P_{x,t}^{(2)} \approx \left(P_{x,t}^{(3)} + P_{x+1,t}^{(3)}\right)/2$

– given $d_x^{(3)}$ but with $P_{x,t}^{(1)}$ or $P_{x,t}^{(2)}$:

$P_{x,t}^{(3)} = P_{x-1,t}^{(1)}$ or $P_{x,t}^{(3)} \approx \left(P_{x-1,t}^{(2)} + P_{x,t}^{(2)}\right)/2$