The likelihood of observing d deaths if the true value of the hazard rate is μ is

$$L(\mu) = \frac{(\mu E_x^c)^d e^{-\mu E_x^c}}{d!}$$

which can be maximised by maximising its log

$$\log L(\mu) = d(\log \mu + \log E_x^c) - \mu E_x^c - \log d!$$

Differentiating w.r.t μ

$$\frac{\partial}{\partial \mu} \log L(\mu) = \frac{d}{\mu} - E_x^c$$

which is zero when

$$\hat{\mu} = \frac{d}{E_x^c}$$

The estimator $\tilde{\mu}$ has the following properties

- $E[\tilde{\mu}] = \mu$
- $Var[\tilde{\mu}] = \frac{\mu}{E_x^c}$

The asymptotic distribution of $\tilde{\mu}$ is

$$\tilde{\mu} \sim N\Big(\mu, \frac{\mu}{E_x^c}\Big)$$

Exposed to Risk

Central exposed to risk is the total waiting time which features in both two-state markov model and the poisson model.

The central exposed to risk is a natural quantity intrinsically observable even if the observation may be incomplete in practice.

Homogeneity

The Poisson models are based on the assumption that we can observe groups of identical lives or homogeneous groups.

A group of lives with different characteristics is said to be *heterogenous*

As a result of this heterogeneity, our estimate of the mortality rate would be the estimate of the average rate over the whole group of lives.

Example

consider a country in which 50% of the population are smokers. If $\mu_{40} = 0.001$ for non-smokers and $\mu_{40} = 0.002$ for smokers, then a mortality investigation based on the entire population may lead us to the estimate $\hat{\mu}_{40} = 0.0015$ An insurance company that calculates its premiums using this average figure would overcharge non-smokers and undercharge smokers.

The solution is subdivide our data according to characteristics known, from experience, to have a significant effect on mortality. This ought to reduce the heterogeneity of each class.

Among the factors in respect of which life insurance mortality statistics are often sub-divided are:

- Sex
- Age
- Type of policy
- Smoker/non-smoker status
- Duration in force
- Level of underwriting

Principle of Correspondence

Mortality investigations based on estimation of $\mu_{x+\frac{1}{2}}$ at individual ages brings together two different items of data **deaths and exposures**

These should be defined consistently or the ratios are meaningless.

The principle of correspondence states that:

A life alive at time t should be included in the exposure at age x at time t if and only if, were that life to die immediately he or she would be counted in the death data d_x at age x.

Exact Calculation of E_x^c

The procedure for the exact calculation of E_x^c is obvious:

a. Record all dates of birth
b. Record all dates of entry into observation
c. Record all dates of exit from observation
d. Compute E_x^c

If we add to the data above the cause of the cessation of observation we have d_x as well and we have finished.

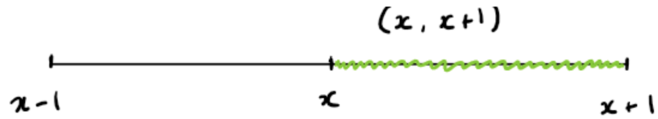
The central exposed to risk E_x^c for a life with age label x is the time from Date A to Date B where

Date A is the latest of:	the date of reaching age label x
	the start of the investigation and
	the date of entry
Date B is the earliest of:	the date of reaching age label $x + 1$
	the end of the investigation and
	the date of exit (for whatever reason)

Age Definitions

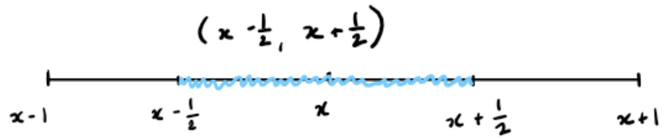
Age last birthday

A life will be considered age x with their real age being in the range $(x, x+1)$



Age nearest birthday

A life will be considered age x with their real age being in the range $x - \frac{1}{2}, x + \frac{1}{2}$



Age Next Birthday

A life will be considered age x with their real age being in the range $(x-1, x)$

Census Approximation to E_x^c

Suppose that we have death data of the form:

d_x total number of deaths $\times$ last birthday during calendar years $K, K+1, ..., K+N$

That is we have over $N+1$ calendar years of all deaths between ages x and $x+1$

However, instead of the times of entry to and exit from observation of each life being known, we have instead only the following census data

$P_{x,t} $ = Number of lives under observation aged x last birthday at time t where $t = 1$ january in calendar years $K, K+1, ..., K+N, K+N+1$

Define $P_{x,t}$ to be the number of lives under observation aged x last birthday, at ant time t. Note that

$$E_x^c = \int_K^{K+N+1} P_{x,t} dt$$

During any short time interval $(t, t+dt)$ there will be $P_{x,t}$ lives each contributing a fraction of a year dt to the exposure.

So integrating $P_{x,t} * dt$ over the observation period gives the total exposed to risk for this age.

Using the trapezium approximation

$$E_x^c = \int_K^{K+N+1} P_{x,t}dt \approx \sum_{t=K}^{K+N} \frac{1}{2}(P_{x,t} + P_{x,t+1})$$

Example

Estimate E_{55}^c based on the following data

Calendar year	Population aged 55 last birthday on 1 January
2005	46,233
2006	42,399
2007	42,618
2008	42,020

$$E_{55}^c = \int_0^3 P_{55,t}\, dt$$

$$E_{55}^c = \frac{1}{2}\left[P_{55,0} + P_{55,1}\right] + \frac{1}{2}\left[P_{55,1} + P_{55,2}\right] + \frac{1}{2}\left[P_{55,2} + P_{55,3}\right]$$

$$= \frac{1}{2}P_{55,0} + P_{55,1} + P_{55,2} + \frac{1}{2}P_{55,3}$$

$$= 0.5 * 46233 + 42399 + 42618 + 0.5 * 42020$$

$$= 129143.5$$

Deaths classified using different definitions of age

Definitions that could be used for year of age include

- $d_x^{(1)}$ total number of deaths at age x last birthday during calendar years $K, K+1, ..., K+N$
- $d_x^{(2)}$ total number of deaths age x nearest birthday during calendar years $K, K+1, ..., K+N$
- $d_x^{(3)}$ total number of deaths age x next birthday during calendar years $K, K+1, ..., K+N$

Rate Interval

A rate interval is a period of one year during which a life's recorded age remains the same.

The rate of mortality q measures the probability of death over the next year of age or more generally over the next rate interval.

The possibilities are:

Definition of x	Rate interval	$\hat{q}$ estimates	$\hat{\mu}$ estimates
Age last birthday	$[x, x+1]$	q_x	$\mu_{x+½}$
Age nearest birthday	$[x-½, x+½]$	$q_{x-½}$	μ_x
Age next birthday	$[x-1, x]$	q_{x-1}	$\mu_{x-½}$

Once the rate interval has been identified (from the age definition used in d_x) the rule is that

- the crude $\hat{\mu}$ estimates μ in the middle of the rate interval
- the crude $\hat{q}$ estimates q at the start of the rate interval.

Graduation and Statistical tests

Graduation refers to the process of using statistical techniques to improve the estimates provided by the crude rates.

The aims of graduation are to produce a smooth set of rates that are suitable for a particular purpose, to remove random sampling errors (as far as possible) and to use the information available from adjacent ages to improve the reliability of the estimates. Graduation results in a "smoothing" of the crude rates.