INSTITUTE OF ACTUARIAL & QUANTITATIVE STUDIES

Class: SY BSc

Subject : Statistical and Risk Modelling 1

Chapter: Unit 1 Chapter 1

Chapter Name: Survival Modelling

Today's Agenda

1. Introduction to Survival Modelling
    1. What is Survival Modelling?
    2. Why do we need Survival Modelling?

2. Lifetable Functions
    1. Understanding distribution and density functions
    2. Understand and apply approximate methods for fractional ages

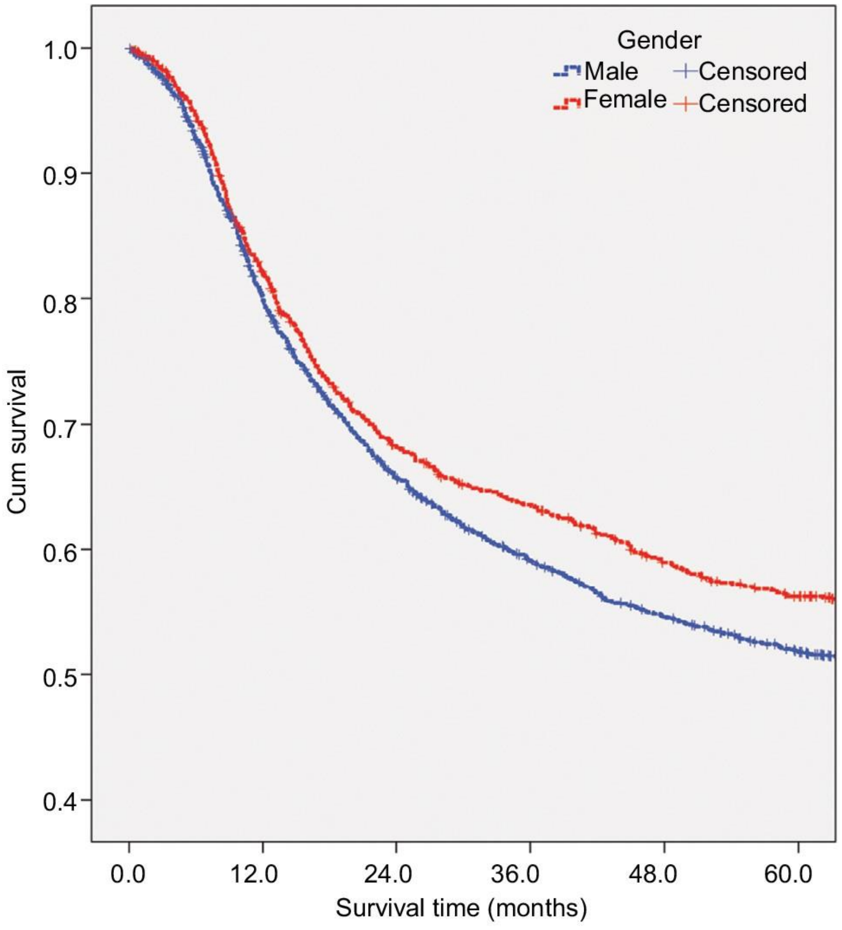
3. Parametric Survival Models

1 Introduction to Survival Modelling

1.1 What is Survival Modelling?

- Survival is the probability of remaining alive for a specific length of time.
- Survival analysis or survival modelling is the use of statistical methods for analyzing the data on the occurrence of an event.
- Events may include death, injury, onset of illness, recovery from illness, death after recovery or transition above or below the clinical threshold of a meaningful continuous variable, etc.

- Example:-Researchers used a SYSUCC medical record database to identify Oral Cavity Cancer (OCC) patients diagnosed from 1960 to 2009. A total of 3,362 previously untreated patients with histologically confirmed OCC were enrolled in this study. They were classified based on age and gender. It was realized survival rates for females were significantly higher than males.

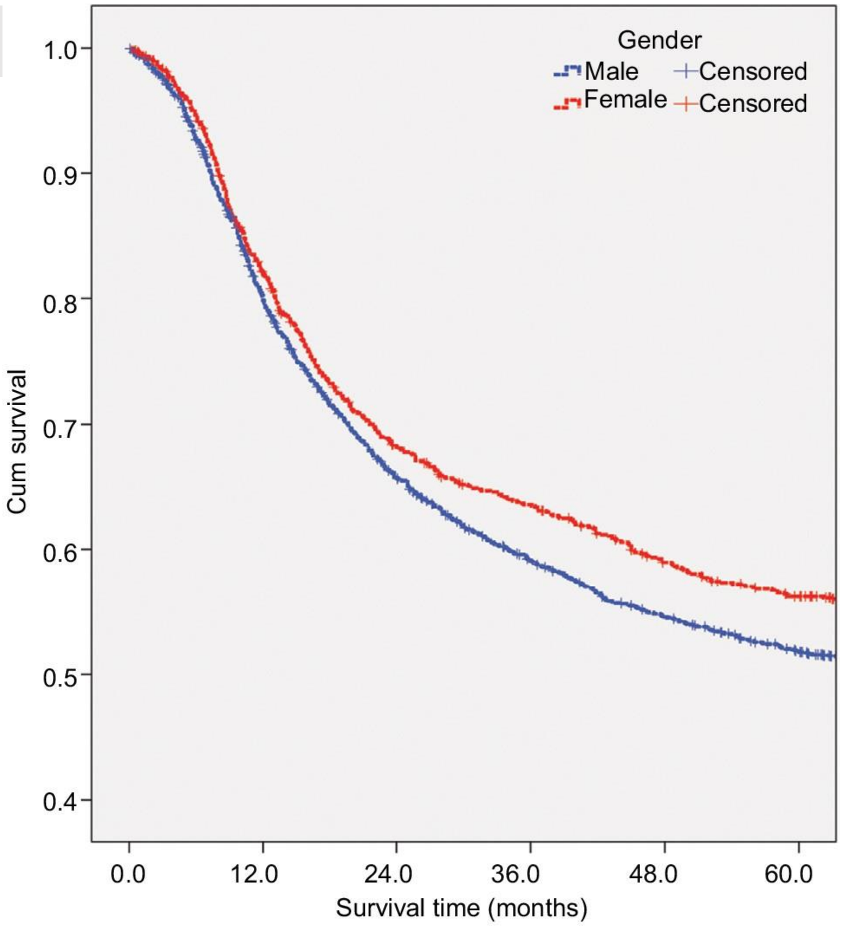


1.1 What is Survival Modelling?

Interpret and comment on the graph alongside

The graphs can be interpreted as follows:

1. All patients who had untreated OCC were enrolled into the study between 1960 and 2009.
2. We assume the study concluded in 2009 and hence the last patient admitted in this study was in 1 Jan 2005 and the study concluded on 31 December 2009. The duration of this study was 60 months and patients were not monitored post 5 years i.e. censored post 5 years.
3. Hence we see that all patients who enrolled were alive on day 1 as can be inferred from the starting point.
4. After 60 months it can be concluded based on the graph that more no. of females were alive than males which leads us to conclude that females had higher survival rates.



1.2 Why survival modelling?

Various institutions like insurance and reinsurance firms and others face the need to do survival analysis in order to plan for future
Life table or actuarial methods are used to develop survival curves, these help to study the patterns in them. The expectation of life is often used as a measure of the standard of living and health care in a given country.

Can you guess which countries had life expectancy of 35 to 40 years and which had life expectancies of 75 to 80 years based on CIA word factbook in 2009?

35 to 40 years – Angola and Zambia;
75 to 80 years – Western countries like USA, Japan etc.

1.2 Why survival modelling?

Did you know 1 in 555k people do not survive Scuba Diving! Can you guess how many people don't survive a dog bite?

1 in 225K do not survive but annually there are millions of reported dog bites.

The models of lifetime or survival learned in this chapter can be applied in several other actuarial contexts. Can you guess a few??

Other actuarial context : disability, sickness/illness, retirement, unemployment and withdrawals (lapse rates)

Self Study: Watch the video about Probability Comparison: Chances Of Survival – YouTube

1.2 Why survival modelling?

Can you name a few other fields and the applications of these models in these fields?

- **Engineering :** lifetime of a machine, lifetime of a light bulb
- **Medical statistics**: time-until-death from diagnosis of a disease, survival after surgery.
- **Finance :** time-until-default of credit payment in a bond, time-until-bankruptcy of a company
- **Space probe** : probability radios installed in space continue to transmit
- **Biology**: lifetime of an organism

Watch the video about uses of survival analysis

2 Life table Distributions & Functions

2.1a Simple Survival Model

We start with the **future lifetime**.
For a person now aged x, it's future lifetime is defined as T_x . For a new born, x = 0, so that we have future lifetime as T . Life aged x is denoted by (x).

T_x = Future lifetime beyond age x of an individual who has survived to age x [measured in years and partial years].

- The total life length of this individual will be x + T_x , i.e. this is the age at which the individual dies [including partial years]. The additional years of life Tx beyond x is unknown and therefore is viewed as a **continuous random variable.**
- We have a maximum age or limiting age. Let ω denote some upper age limit . Therefore, future lifetime is continuously distributed on an interval [0,ω] where 0 < ω < ∞.

2.1.b Distribution of Future Lifetime Random Variable

The future lifetime random variable for life aged x is described as,

For $0 < x < \omega$,
$F_x(t)$ = P [$T_x \leq t$] is the distribution function of T_x
$S_x(t)$ = P [$T_x > t$] = 1 - $F_x(t)$ is the survival function of T_x

$S_x(t)$ is known as the survival function of T_x because it represents the probability of a life aged x surviving to age x + t .
Clearly, $F_x(t)$ = P[$T_x \leq t$] is the probability that someone who has survived to age x will not survive beyond age x + t.

2.1.b Distribution of Future Lifetime Random Variable

Explain $F_{20}(40)$ and compare ${}_{3}p_{22}$ with ${}_{4}p_{21}$

$F_{20}(40)$ is the probability of a person currently aged 20 dies before their 60th birthday. This can also be represented as ${}_{40}q_{20}$

${}_{3}p_{22}$ with ${}_{4}p_{21}$ is the probability that a person aged 22 survives till age 25 and a person aged 21 survives till age 25 respectively. Probability of survival for someone aged 22 till 25 will be higher than someone aged 21 till 25 as the period of survival is already included in later.

2.1c Actuarial notation for Probabilities

Survival Probability

$$ {}_t p_x = 1 - {}_t q_x = S_x(t) $$

- ${}_t p_x$ is the probability that a life now aged *x* is still alive after *t* years.
- p_x is the probability that a life now aged *x* is still alive after 1 year

Mortality Probability

$$ {}_t q_x = F_x(t) $$

- ${}_t q_x$ is the probability that a life now aged *x* dies within *t* years.
- q_x is the probability that a life now aged *x* dies within 1 year.

From above: $ {}_t q_x + {}_t p_x = 1$

2.1d Force of Mortality

Force of mortality is the instantaneous death rate for a life. It is denoted as μ_x = μ(x)= force of mortality at age x, given survival to age x. This is also called the "hazard rate" or "failure rate." It is the continuous equivalent of the discrete quantity q_x.

Force of mortality at age x is defined as: $\mu_x = \lim_{h \to 0^+} \frac{1}{h}$ x $P[T \leq x + h | \, T > x]$

Force of mortality at age x + t, can be defined as: $\mu_{x+t} = \lim_{h \to 0^+} \frac{1}{h}$ x $P[Tx \leq t + h | \, Tx > t]$

Force of mortality at age x + t, can also be defined as

$$\mu_{x+t} = \lim_{h \to 0^+} \frac{1}{h} \text{ x } P[T \leq x + t + h | \, T > x + t]$$

Note: Here we have defined μx using the lifetime function at age x.

2.1e The density function of T_x

- We have already seen that the distribution function of T_x is $F_x(t)$ = P [$T_x \leq$ t]. Now we look at the probability density function.
- Denote this by $f_x(t)$, and recall that: $f_x(t) = \frac{d}{dt} F_x(t)$
- Substituting the formula for $F_x(t)$ and carrying out a systematic derivation gives us a very important result of the survival model.

PDF of $T_x = S_x(t)$ x μ_{x+t}

Substituting the survival probability $S_x(t)$ as ${}_tp_x$ gives us:

$$f_x(t) = {}_tp_x * \mu_{x+t} \qquad ; (0 \leq t < \omega - x)$$

2.1e The probability density function of T_x

The distribution function of T_X is $F_X(t)$, by definition. We also want to know its probability density function (PDF).

Denote this by $f_X(t)$, and recall that:

$$f_X(t) = \frac{d}{dt} F_X(t)$$

Then:

$$\begin{aligned} f_X(t) &= \frac{d}{dt} P[T_x \leq t] \\ &= \lim_{h \to 0^+} \frac{1}{h} \times (P[T_x \leq t + h] - P[T_x \leq t]) \\ &= \lim_{h \to 0^+} \frac{P[T \leq x + t + h \mid T > x] - P[T \leq x + t \mid T > x]}{h} \\ &= \lim_{h \to 0^+} \frac{P[T \leq x + t + h] - P[T \leq x] - (P[T \leq x + t] - P[T \leq x])}{S(x) \times h} \\ &= \lim_{h \to 0^+} \frac{P[T \leq x + t + h] - P[T \leq x + t]}{S(x)xh} \end{aligned}$$

2.1e The probability density function of T_x

Now multiply and divide by $S(x+t)$ and we have:

$$f_x(t) = \frac{S(x+t)}{S(x)} \times \lim_{h \to 0^+} \frac{1}{h} \frac{P[T \leq x+t+h] - P[T \leq x+t]}{S(x+t)}$$

$$= S_x(t) x \lim_{h \to 0^+} \frac{1}{h} P[T \leq x+t+h \mid T > x+t]$$

$$= S_x(t) \times \mu_{x+\tau}$$

or, in actuarial notation, for a fixed age x between 0 and ω :

$$f_x(t) = {}_{\tau}p_x \mu_{x+\tau} \ (0 \leq t < \omega - x)$$

This is one of the most important results concerning survival models.

2.1f Complete Expectation of Life

We now look at the expected value of the random variable Tx.
It is denoted by e_x^o . The symbol reads as '*e*-circle-*x*'. This is the complete expectation of life for a life of age x . It considers complete years as well as partial or fractional years.
We will be Using the result that, ${}_tp_x\,\mu_{x+t} = f_x(t) = \frac{\partial}{\partial t}F_x(t) = \frac{\partial}{\partial t}\ {}_tq_x = \frac{\partial}{\partial t}\left(1 - {}_tp_x\right) = -\frac{\partial}{\partial t}\ {}_tp_x$

Using integration by parts we get,

$$\begin{aligned}
\dot{e}_x &= \int_0^\infty t \text{ x } {}_tp_x\,\mu_{x+t}\,dt \\
&= \int_0^\infty t \text{ x } \left(-\frac{\partial}{\partial t}\ {}_tp_x\right) dt \\
&= \left[t \text{ x } (-{}_tp_x)\right]_{t=0}^{t=\infty} - \int_0^\infty (-{}_tp_x)\,dt \\
&= \int_0^\infty {}_tp_x\,dt
\end{aligned}$$

Complete Expectation of Life = $\dot{e}_x = E(T_x)$

2.1g Curtate Future Lifetime

The random variable K_x is used to represent the curtate future lifetime for a life of exact age x (i.e. the number of complete years lived after age x).
The curtate future lifetime of a life age x is: K_x = $[T_x]$ where the square brackets denote the integer part. [T_x rounded down]

The probability:

$$P[K_x = k] = P[k \leq T_x < k + 1]$$
$$= P[k < T_x \leq k + 1]$$
$$= {}_k p_x q_{x+k}$$

This result is intuitive. If the random variable Kx takes the value k , then a life of exact age x must live for k complete years after age x . Therefore, the life must die in the year of age to (x+k) to (x+k+1).
We also use the symbol $k|q_x$ to represent P[K_x= k]. It is read as 'k deferred q_x', and we can think about this as deferring the event of death until the year that begins in k years from now.

2.1h Curtate Expectation of Life

Curtate expectation of life is denoted by e_x.

Thus, we have $e_x = E(K_x) = \sum_{k=0}^{\infty} k \text{ x } \Pr(K_x = k)$

Carrying the derivation gives the result for the expectation.

Curtate expectation of life: $e_x = E(K_x) = \sum_{k=1}^{\infty} {}_kp_x$

Relationship between complete and curtate expectation assuming linearity between integer ages

They are related by the approximate equation:

$\mathring{e}_x \cong e_x + ½$

To see this, define $\mathbf{J}_x = \mathbf{T}_x - \mathbf{K}_x$ to be the random lifetime after the highest **integer** age to which a life age x survives.

Approximately, $E[\mathbf{J}_x] = ½$, but $E[\mathbf{T}_x] = E[\mathbf{J}_x] + E[\mathbf{K}_x]$ so $\mathring{e}_x \cong e_x + ½$ as stated.

2.1i Life Table Functions

A **life table**(also called a **mortality table** or **actuarial table**) is a table which shows, for each age, what the probability is that a person of that age will die before his or her next birthday ("probability of death"). In other words, it represents the survivorship of people from a certain population
The following are definitions of the standard actuarial life table functions:

Symbol	Definition
	Probability that a person aged x will die within 1 year.
	Number of persons surviving to exact x.
	Number of deaths between exact ages x and x+1.
	Number of persons years lived between exact ages x and x+1.
	Number of person years lived after exact age x.
	Average number of years of life remaining at exact age x.

2.2 Life table Distributions & Functions

It is common for the standard life table functions such as l , or μ to be tabulated at integer ages only. However, the actuary may be required to calculate probabilities involving non-integer ages or durations. In order to do so we can approximate the values for non integer ages $x + t$ where $0<t<1$.

We consider three possible approaches.

2.2a Uniform Distribution of Deaths (UDD)

In this case, we assume that any deaths over the year of age to occur uniformly over the year x to (x +1). This is equivalent to the assumption that the function lx is linear over the interval (x,x+1).
Thus

UDD Assumption
If deaths are uniformly distributed between the ages of x and $x+1$, it follows that ${}_tq_x = tq_x$ for $0 \le t \le 1$.

Thus, under UDD , **the force of mortality is an increasing function** over the year of age x to x+1.

2.2b Constant Force of Mortality (CFM)

In this case, we assume that **the function μ is constant over the year of age x to(x+1)** i.e. for integer x and 0 < t < 1, we have μ_{x+t} = μ = constant.

Under the assumption of a constant force of mortality between integer ages, we find the value of the constant μ using: $p_x = \exp\left(-\int_0^1 \mu_{x+t} dt\right) = e^{-\mu} \quad => \mu = -\ln(p_x)$

Then, for 0 < t < 1, we have: $_tq_x = 1 - {_tp_x}$

$$= 1 - \exp\left(-\int_0^t \mu_{x+s} ds\right)$$

$$= 1 - \exp\left(-\int_0^t \mu ds\right)$$

$$= 1 - e^{-t\mu}$$

CFM Assumption : $_tp_x = e^{-t*u}$; For 0 < t < 1

2.2 Life table Distributions & Functions

Example

Given $p_{90} = 0.75$, calculate $_{\frac{1}{12}}q_{90}$, assuming:

(a)a uniform distribution of deaths between integer ages, and

(b) a constant force of mortality between integer ages.
Solution:

Question

Subject CT4 September 2007 (Q6)

Below is an extract from English Life Table 15 (Males)

Ages	l_x
58	88,792
62	84,173

(i) Estimate l_{60} under each of the following assumptions:

(a) a uniform distribution of deaths between exact ages 58 and 62 years;
and
(b) a constant force of mortality between exact ages 58 and 62 years

(ii) Find the actual value of l_{60} in the tables and hence comment on the relative validity of the two assumptions you used in part (i)

Solution

(i) (a) Assuming a uniform distribution of deaths between ages 58 and 62 implies that half of those who die between those ages die between ages 58 and 60.

Therefore

$$l_{60} = l_{58} - 0.5(l_{58} - l_{62})$$

$$= 88{,}792 - 0.5(88{,}792 - 84{,}173)$$

$$= 86{,}482.5.$$

Solution

(b) ALTERNATIVE 1

Let the constant force of mortality be μ.

Then we have ${}_4p_{58} = \exp\left(-\int_0^4 \mu dx\right) = e^{-4\mu}$.

But ${}_4p_{58} = \dfrac{l_{62}}{l_{58}} = \dfrac{84,173}{88,792} = 0.94798$.

Therefore $e^{-4\mu} = 0.94798$,

so that $-4\mu = \log_e(0.94798) = -0.05342$,

whence $\mu = 0.01336$.

Therefore with a constant force of mortality,

$l_{60} = l_{58}\exp[-2(0.01336)] = 88,792(0.97363)$

so $l_{60} = 86,452$.

ALTERNATIVE 2

Let the constant force of mortality be μ.

Then we have ${}_4p_{58} = \exp\left(-\int_0^4 \mu dx\right) = e^{-4\mu}$.

But ${}_4p_{58} = \dfrac{l_{62}}{l_{58}}$.

Now $l_{60} = l_{58} \cdot {}_2p_{58}$.

and, since ${}_2p_{58} = e^{-2\mu} = \sqrt{e^{-4\mu}} = \sqrt{\dfrac{l_{62}}{l_{58}}}$,

$l_{60} = l_{58}\sqrt{\dfrac{l_{62}}{l_{58}}} = \sqrt{l_{58}l_{62}} = \sqrt{(88,792)(84,173)}$

so $l_{60} = 86,452$

Solution

(ii) The actual value of l_{60} from the tables is 86,714.

This shows that neither assumption is very accurate, but that the uniform distribution of deaths (UDD) is closer than the constant force of mortality.

The UDD assumption is better than the constant force of mortality assumption because UDD implies an increasing force of mortality over this age range, which is biologically more plausible than the assumption of a constant force.

The fact that the actual value of l_{60} is considerably greater than that implied by the UDD assumption suggests that the true rate of increase of the force of mortality over this age range in English Life Table 15 (males) is even greater than that implied by UDD.

2.2c The Balducci Assumption

The Italian actuary Balducci proposed an alternative approach for estimating probabilities at non integer ages and durations.

Balducci Assumption

For integer age x and $0 < t < 1$, the Balducci assumption gives,

$${}_tq_x = 1 - \frac{1 - q_x}{1 - (1 - t) \text{ x } q_x} = \frac{t \text{ x } q_x}{1 - (1 - t) \text{ x } q_x}$$

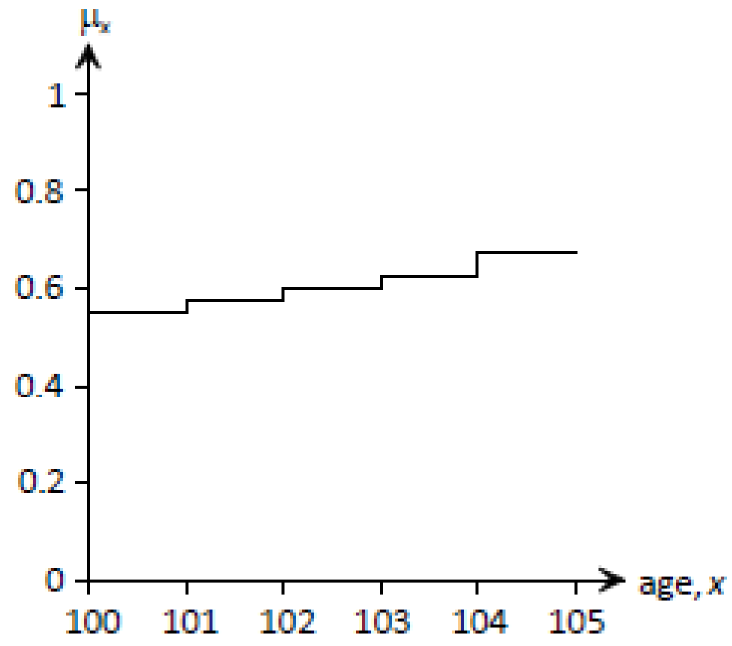
For the Balducci assumption, it can shown that the force of mortality at age (x+t) is given by:

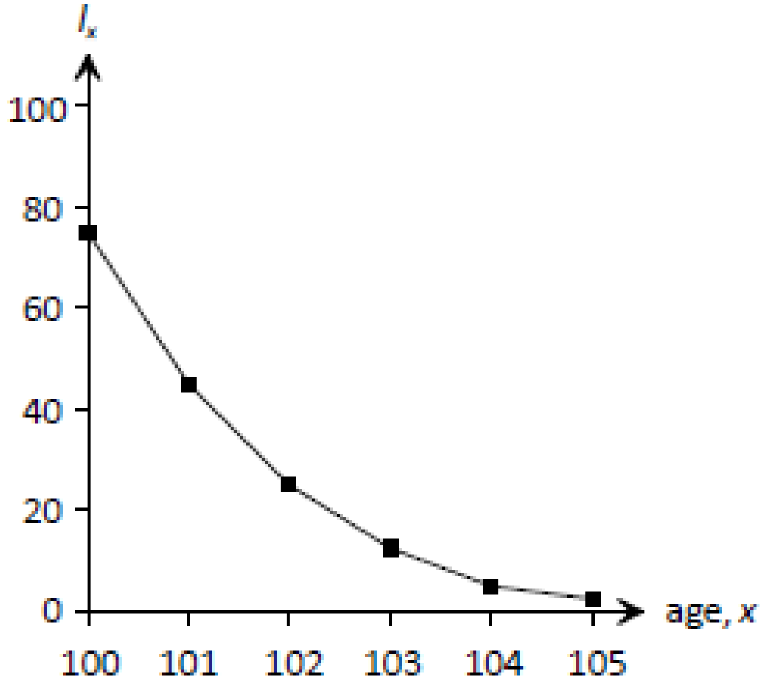
$$\mu_{x+t} = \frac{q_x}{1 - (1 - t) \text{ x } q_x}$$

Thus, under the Balducci assumption, **the force of mortality is a decreasing function** over the year of age x to x +1 . This result is counter-intuitive and inconsistent with the expected pattern for the force of mortality for human populations.

1.2 Life table Distributions & Functions

Discuss which graph corresponds with which fractional age assumption.





2 Life table Distributions & Functions

Important Formulae!

We have already seen the definitions of the survival probability (${}_tp_x$) and mortality probability (${}_tq_x$). We consider the formulas, while one can look out at the derivations if interested.

A formula for ${}_tq_x$:

$$ {}_tq_x = F_x(t) = \int_0^t f_x(s)ds = \int_0^t {}_sp_x\mu_{x+s}\, ds $$

A formula for ${}_tp_x$:

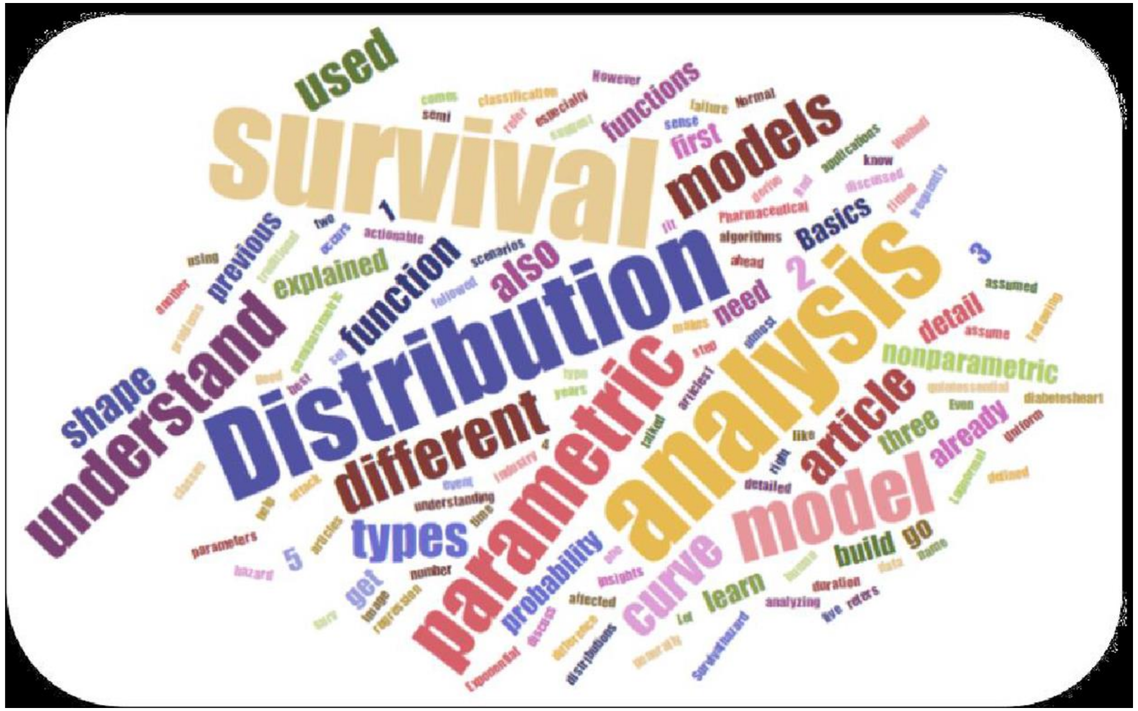
$$ {}_tp_x = \exp\left\{-\int_0^t \mu_{x+s}ds\right\} $$

3 Parametric Survival Models

Parametric modelling requires choosing one or more distributions. The parametric survival models are in regular practical use where the future lifetime random variable is defined using certain statistical distributions with specific defined parameters.

Under parametric models, we make an assumption regarding the underlying distribution, and then try to make inferences about the survival function or the hazard function.

We now look at some parametric models.



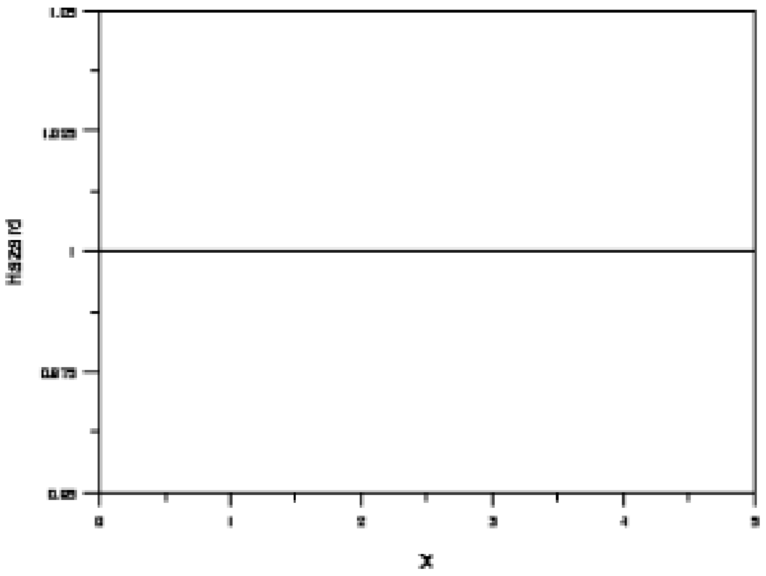
3.1 Parametric Survival Models

Exponential Model

Exponential distribution is one of the common assumption taken in survival models. The hazard function does not vary with time i.e., we assume that the hazard function is constant. This distribution can be assumed in case of natural death of human beings where the rate does not vary much over time.
Assume that T ~ Exp(λ) with the probability density function $f(t) = \lambda * \exp(-\lambda t)$
and hazard function = λ

S(t) = ${}_{t}p_{x}$ = $e^{-\lambda t}$
Therefore, ${}_{t}q_{x}$ = 1 -$e^{-\lambda t}$

The graph of hazard function looks like:



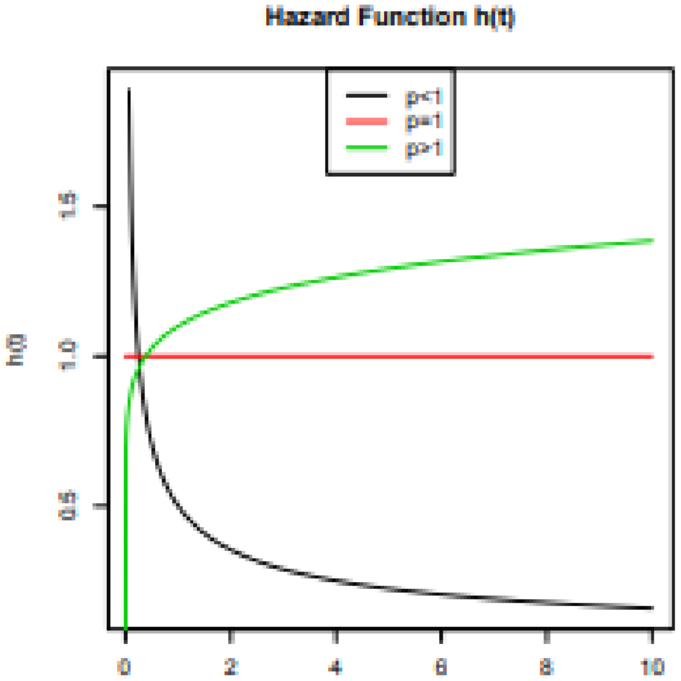
3.2 Weibull Model

Weibull model is an extension to the exponential survival model.
Assuming T ~Weibull(λ, p) with probability density function $f(t) = \lambda p t^{p-1} \exp(-\lambda t^p)$, where p > 0 and λ > 0,

$S_x(t) = {}_tp_x = e^{[-\lambda * t^p]}$
Hazard function is given by
h(t) = $\lambda p * t^{(p-1)}$

p is called shape parameter:

If p > 1 the hazard increases.
If p = 1 the hazard is constant (exponential model).
If p < 1 the hazard decreases.



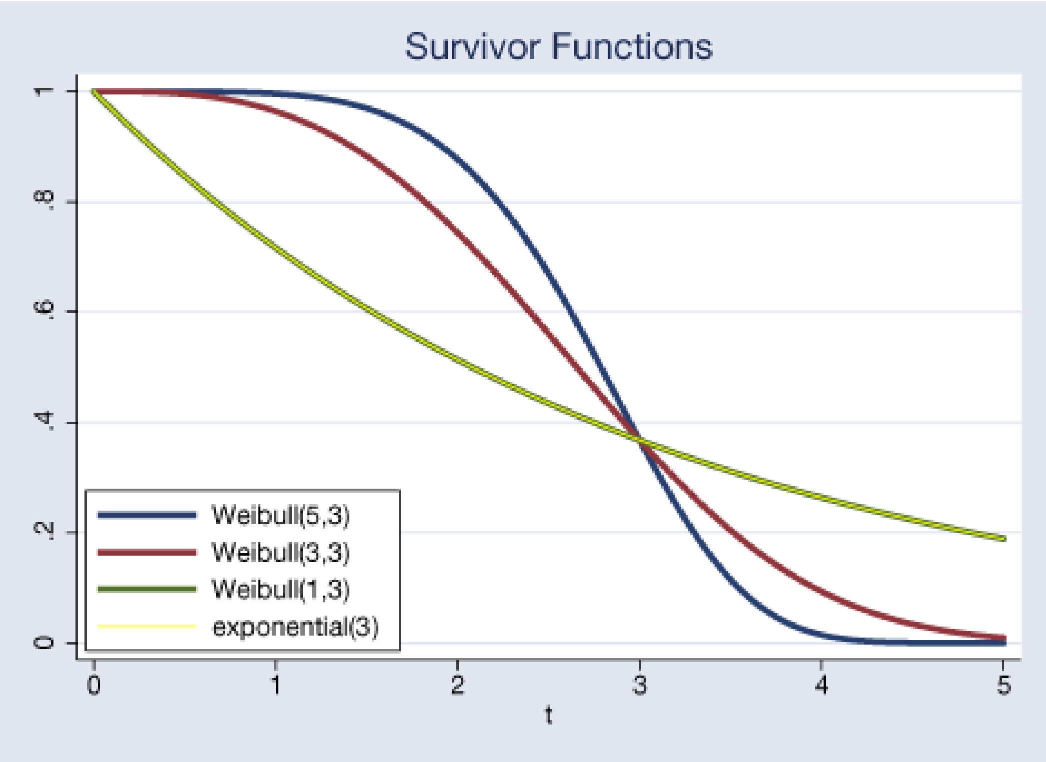
3 Parametric Survival Models

Relationship between Exponential and Weibull Model

We can see how survivor functions for various distributions relate to each other. Recall that the survivor function is 1 minus the cumulative distribution function, S(t) = 1 – F(t).

We plot the survivor function that corresponds to our Weibull(5,3). We add a Weibull(3,3) and Weibull(1,3).

We are also going to plot an exponential(3) with a thin line. Where is this line seen in the graph?
You will see that it falls entirely over the Weibull(1,3) because the Weibull(1,b) is equal to the exponential(b)



3 Parametric Survival Models

Weibull Model usage:

Weibull Model:

Reliability engineers use statistics and mathematical analysis to predict how long their devices will function. By knowing how long a device should work, they can predict warranty periods, plan preventative maintenance, and order replacement parts before they are needed.

(Using example of bulbs)

Comparing and Weighting Two Weibull Models
https://demonstrations.wolfram.com/ComparingAndWeightingTwoWeibullModels/

Question

CT4 April 2009 Q10

Let *Tx* be a random variable denoting future lifetime after age *x*, and let *T* be another random variable denoting the lifetime of a new-born person.

(i) (a) Define, in terms of probabilities, *Sx* (*t*) , which represents the survival function of *Tx*.

(b) Derive an expression relating *Sx* (*t*) to *S*(*t*) , the survival function of *T*. [2]

ii) Define, in terms of probabilities involving *Tx* , the force of mortality, μ_{x+t}.

Question

The Weibull distribution has a survival function given by:
$Sx\,(t) = \exp(-(\lambda t)^{\beta})$

where λ and β are parameters (λ, β> 0).

(iii) Derive an expression for the Weibull force of mortality in terms of λ and β.

(iv) Sketch, on the same graph, the Weibull force of mortality for $0 \le t \le 5$for the following pairs of values of λ and β:

λ=1, β=0.5
λ=1, β=1.0
λ=1 ,β=1.

Solution

(i) (a) $S_x(t) = \Pr[T_x > t]$

(b) EITHER

Since $\Pr[T_x > t] = \Pr[T > x+t \mid T > x] = \dfrac{\Pr[T > x+t]}{\Pr[T > x]}$

and $S(t) = \Pr[T > t]$,

then $S_x(t) = \dfrac{S(x+t)}{S(x)}$.

OR

Since $S_x(t) = {}_t p_x$, then using the consistency principle

${}_{x+t}p_0 = {}_t p_x \cdot {}_x p_0$

Therefore ${}_t p_x = S_x(t) = \dfrac{{}_{x+t}p_0}{{}_x p_0} = \dfrac{S(x+t)}{S(x)}$.

(ii) EITHER

$$\mu_{x+t} = -\frac{1}{\Pr[T_x > t]}\frac{d}{dt}[\Pr(T_x > t)]$$

OR

$$\mu_{x+t} = \lim_{h \to 0^+} \frac{1}{h}\left(\Pr[T_x \le t+h \mid T_x > t\right)$$

Solution

(iii) EITHER

If the density function of T_x is $f_x(t)$, then we can write

$$f_x(t) = S_x(t)\mu_{x+t} = -\frac{d}{dt}S_x(t)$$

Therefore $\mu_{x+t} = -\frac{1}{S_x(t)}\frac{d}{dt}S_x(t)$

If $S_x(t) = \exp\left(-(\lambda t)^\beta\right)$, therefore, we have

$$\mu_{x+t} = -\frac{1}{\exp\left(-(\lambda t)^\beta\right)}\frac{d}{dt}\exp\left(-(\lambda t)^\beta\right)$$

$$\mu_{x+t} = -\frac{1}{\exp\left(-(\lambda t)^\beta\right)}\left(\exp\left(-(\lambda t)^\beta\right)\right)\left(-\lambda^\beta \beta t^{\beta-1}\right) = \lambda^\beta \beta t^{\beta-1}$$

OR

$$S_x(t) = \exp\left[-\int_0^t \mu_{x+s}ds\right] = \exp\left[-(\lambda t)^\beta\right].$$

So

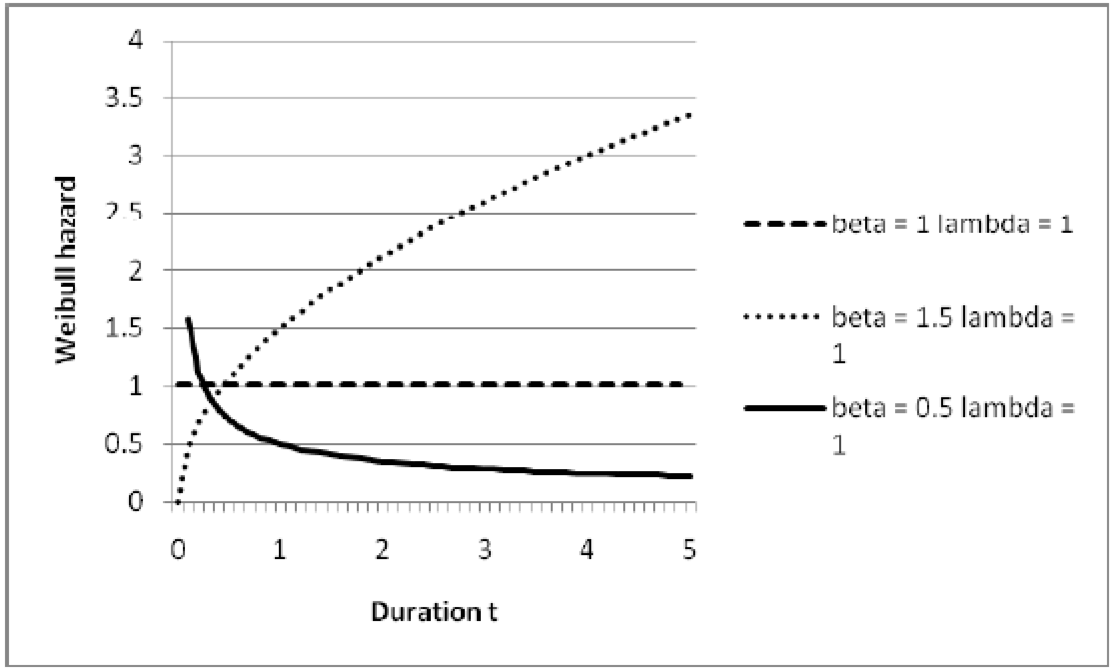
$$\frac{d}{dt}\left[\int_0^t \mu_{x+s}ds\right] = \mu_{x+t} = \frac{d}{dt}\left[(\lambda t)^\beta\right],$$

and hence

$$\mu_{x+t} = \beta\lambda^\beta t^{\beta-1}.$$

Solution

(iv)



4
3.5
3
2.5
2
1.5
1
0.5
0
Weibull hazard
0
1
2
3
4
5
Duration t
beta = 1 lambda = 1
beta = 1.5 lambda = 1
beta = 0.5 lambda = 1

3.3 Gompertz Law of Mortality

This is a defined parametric survival model. The Gompertz function, named after Benjamin Gompertz, is an exponential function, and it is often a reasonable assumption for middle ages and older ages. The function increases exponentially with age. The law of mortality describes the age dynamics of human mortality rather accurately in the age window from about 30 to 80 years of age.

Gompertz Law: $\mu_x = B.c^x$

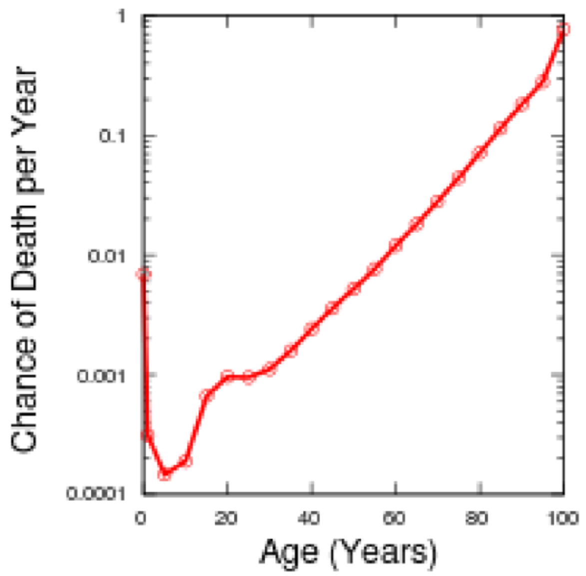
Survival probability tpx can be found using the integral formula for the same.

Under Gompertz law :-

$${}_tp_x = g^{c^x(c^t-1)}$$

where $g = \exp\left(-\frac{B}{\log \quad c}\right)$

Watch video explaining Covid 19 spread by Gompertz Curves



3.4 Makeham Law of Mortality

The Makeham function is an age-independent component (the Makeham term, named after William Makeham). Makeham's Law incorporates a constant term, which is sometimes interpreted as an allowance for accidental deaths, not depending on age.

Makeham's Law: $\mu_x = A + B.c^x$

Survival probability tpx can be found using the integral formula for the same
Under Makeham's law :-

$${}_tp_x = s^t g^{c^x(c^t-1)}$$

where $g = \exp\left(-\frac{B}{\log\ c}\right)$ and $s = \exp(-A)$

3 Relationship between Gompertz and Makeham's law

The addition of the parameter alpha has a certain act on the nature of the model. Firstly it increases the initial starting point at age 0. The second affect is that it makes the demographic slope be more gradual at early ages. This gives it the affect of a gradually increasing in gradient until it converts to the Gompertz function. When the Makeham model converts to the Gompertz function the parameter alpha has little affect to the overall model. From observing figure (1.2) it is clear to see the effect the alpha parameter has on the Makeham model. The initial increase of the parameter alpha from 0 to 0.0001 the model reveals a sharp jump from the initial mortality. As the alpha parameter continues to increase by 0.0001 the jumps between the previous model become small.

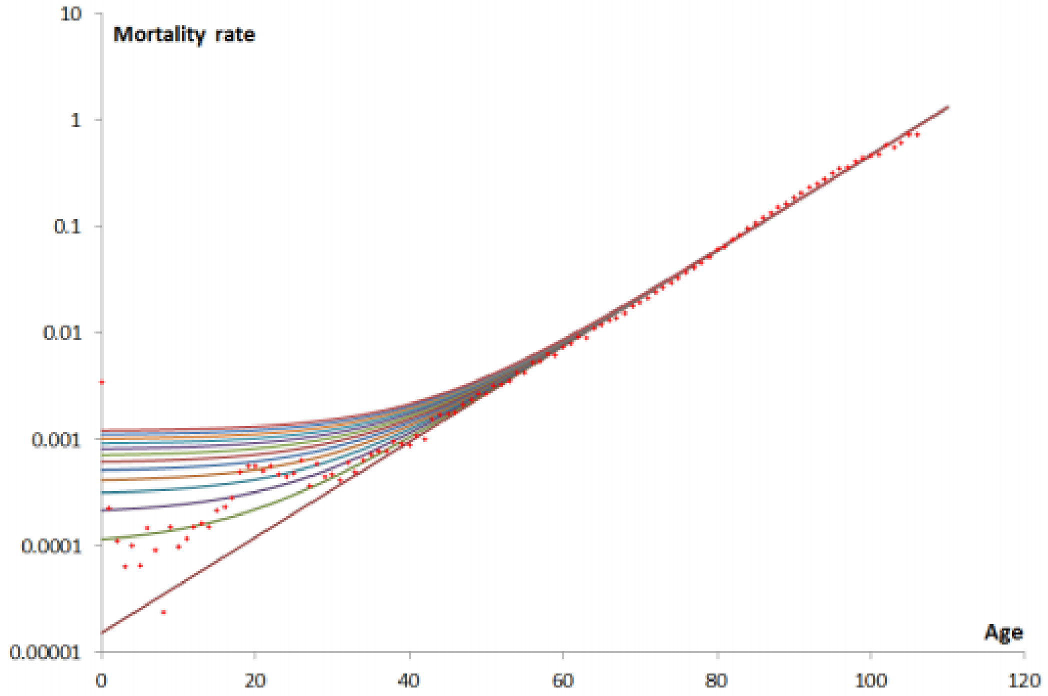


Figure 1.2: *Period data from Sweden 2000. Set in a semi logarithmic scale, the natural logarithm of mortality against age. The Makeham model has been fitted with constant initial mortality and beta. The alpha parameter is increasing from 0 to 0.0012 by a factor of 0.0001.*

Recap

- Survival analysis or survival modelling is the use of statistical methods for analyzing the data on the occurrence of an event.
- Various institutions like insurance and reinsurance firms and others face the need to do survival analysis in order to plan for future.
- T_x = Future lifetime beyond age x of an individual who has survived to age x [measured in years and partial years]. The additional years of life Tx beyond x is unknown and therefore is viewed as a continuous random variable.
- For 0 < x < ω, $F_x(t)$ = P [T_x ≤ t] is the distribution function of T_x
- $S_x(t)$ = P [T_x > t] = 1 - $F_x(t)$ is the survival function of T_x.
- Survival Probability : ${}_tp_x = 1 - {}_tq_x = S_x(t)$: ${}_tp_x$ is the probability that a life now aged *x* is still alive after *t* years.
- ${}_tp_x = \exp\left\{- \int_0^t \mu_{x+s} ds\right\}$ & ${}_tq_x = F_x(t) = \int_o^t f_x(s) ds = \int_o^t {}_sp_x \mu_{x+s}\, ds$
- Mortality Probability : ${}_tq_x = F_x(t)$: ${}_tq_x$ is the probability that a life now aged *x* dies within *t* years.
- ${}_tq_x + {}_tp_x = 1$
- Force of mortality is the instantaneous death rate for a life. It is denoted as μ_x = μ(x)= force of mortality at age x, given survival to age x. This is also called the "hazard rate" or "failure rate." It is the continuous equivalent of the discrete quantity q_x.
- Force of mortality at age x is defined as: $\mu_x = \lim_{h \to 0^+} \frac{1}{h}\, x\, P[T \leq x + h |\, T > x]$
- PDF of $T_x = S_x(t)\, x\, \mu_{x+t}$ & Density function of T_x : $f_x(t) = {}_tp_x * \mu_{x+t}$; $(0 \leq t < \omega - x)$

Continued

- Complete Expectation of Life considers complete years as well as partial or fractional years. It is denoted by $\dot{e}_x$
- $\dot{e}_x = E(T_x) = \int_0^{\infty} t \text{ x } {}_tp_x\, \mu_{x+t}\, dt = \int_0^{\infty} {}_tp_x\, dt$
- The random variable K_x is used to represent the curtate future lifetime for a life of exact age x (i.e. the number of complete years lived after age x).
- The curtate future lifetime of a life age x is: $K_x = [T_x]$ where the square brackets denote the integer part. [T_x rounded down]. The Probability : $P[K_x = k] = P[k \leq T_x < k+1] = {}_k\, p_x q_{x+k}$
- Curtate expectation of life: $e_x = E(K_x) = \sum_{k=1}^{\infty} {}_kp_x$
- Assuming that the function ${}_tp_x$ is linear between integer ages, we have : $e_x^o = e_x + \frac{1}{2}$
- To calculate probabilities involving non-integer ages or durations, we can approximate the values for non integer ages x + t where 0<t<1. The three approaches used are : UDD, CFM & Balducci Assumption.
- UDD Assumption : If deaths are uniformly distributed between the ages of *x* and *x*+1 , it follows that ${}_tq_x = tq_x$ for $0 \leq t \leq 1$. Thus, under UDD , the force of mortality is an increasing function over the year of age x to x+1.
- Under the assumption of a constant force of mortality(CFM) between integer ages, we find the value of the constant μ using: $p_x = exp\left(-\int_0^1 \mu_{x+t} dt\right) = e^{-\mu}$ $i.e.\ \mu = -ln(p_x)$ Therefore : ${}_tp_x = e^{-t*u}$; For 0 < t < 1
- For integer age x and 0 < t < 1, the Balducci assumption gives , ${}_tq_x = 1 - \frac{1-q_x}{1-(1-t) \text{ x } q_x} = \frac{t \text{ x } q_x}{1-(1-t) \text{ x } q_x}$. Under the Balducci assumption, the force of mortality is a decreasing function over the year of age x to x +1

Continued

- For parametric models, we make an assumption regarding the underlying distribution, and then try to make inferences about the survival function or the hazard function.
- Exponential distribution is one of the common assumption taken in survival models. The hazard function does not vary with time,i.e. we assume that the hazard function is constant. Assume that T ~ Exp(λ) with the probability density function $f(t) = \lambda * \exp(-\lambda t)$ and hazard function = λ.
- S(t) = ${}_tp_x$ = $e^{-\lambda t}$ Therefore, ${}_tq_x$ = 1 -$e^{-\lambda t}$
- Weibull model is an extension to the exponential survival model. Assuming T ~Weibull(λ, p) with probability density function $f(t) = \lambda p t^{p-1} \exp(-\lambda t^p)$, where p > 0 and λ > 0.
- $S_x(t)$ = ${}_tp_x$ = $e^{[-\lambda * t^p]}$. Therefore hazard function is given by h(t) = $\lambda p * t^{(p-1)}$
- p is called shape parameter: If p > 1 the hazard increases; If p = 1 the hazard is constant (exponential model) & If p < 1 the hazard decreases.
- The Gompertz function is an exponential function on which the namesake survival model is based. The law of mortality describes the age dynamics of human mortality rather accurately in the age window from about 30 to 80 years of age.
- Gompertz Law: $\mu_x = B.c^x$ & Under Gompertz law :- ${}_tp_x = g^{c^x(c^t-1)}$ where $g = \exp\left(-\frac{B}{\log\ c}\right)$
- The Makeham function is an age-independent component. Makeham's Law incorporates a constant term, which is sometimes interpreted as an allowance for accidental deaths, not depending on age.
- Makeham's Law: $\mu_x = A + B.c^x$ & Under Makeham's law :- ${}_tp_x = s^t g^{c^x(c^t-1)}$

  where $g = \exp\left(-\frac{B}{\log\ c}\right)$ and $s = \exp(-A)$