TMA4275 LIFETIME ANALYSIS

Slides 2: General concepts for lifetime modeling

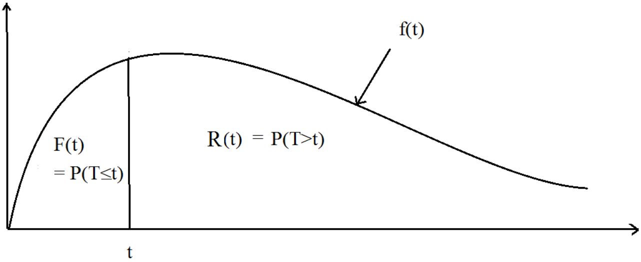
Bo Lindqvist
Department of Mathematical Sciences
Norwegian University of Science and Technology
Trondheim

https://www.ntnu.edu/employees/bo
bo.lindqvist@ntnu.no

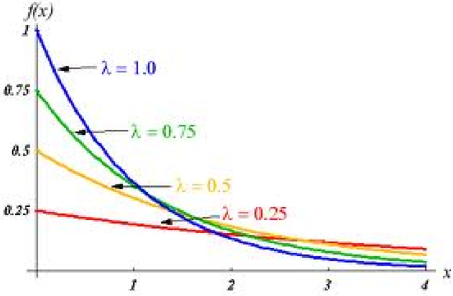
NTNU, Spring 2020

The lifetime T of an individual or unit is a *positive* and *continuously distributed* random variable.

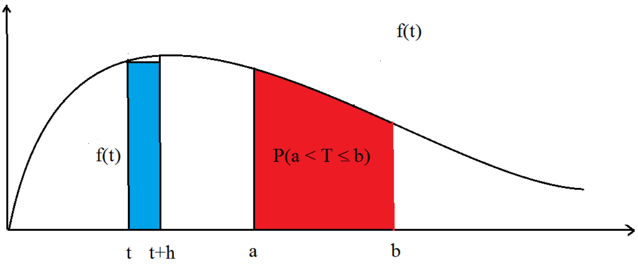
- The probability density function (pdf) is usually called $f(t)$,
- the cumulative distribution function (cdf) $F(t)$ is then given by $F(t) = P(T \leq t) = \int_0^t f(u)du$,
- the reliability (or: survival) function is defined as $R(t) = P(T > t) = 1 - F(t) = \int_t^\infty f(u)du$.



EXAMPLE: EXPONENTIAL DISTRIBUTION



$$
\begin{aligned}
f(t) &= \lambda e^{-\lambda t} \\
F(t) &= 1 - e^{-\lambda t} \\
R(t) &= e^{-\lambda t}
\end{aligned}
$$



$$f(t) = F'(t)$$

$$P(a < T \leq b) = \int_a^b f(u)du = F(b) - F(a)$$

$$P(t < T \leq t+h) = \int_t^{t+h} f(u)du \approx f(t) \cdot h$$

Hence,

$$f(t) \approx \frac{P(t < T \leq t+h)}{h}$$

FROM DENSITY TO HAZARD FUNCTION OF T

From last slide,

$$P(t < T \leq t + h) \approx f(t) \cdot h$$

If we know that the unit is alive (functioning) at time t, i.e. $T > t$, we may be interested in the conditional probability

$$P(t < T \leq t + h | T > t).$$

Using the *conditional probability* formula: $P(A|B) = P(A \cap B)/P(B)$, we get

$$P(t < T \leq t + h | T > t) = \frac{P(t < T \leq t + h)}{P(T > t)} \approx \frac{f(t)h}{R(t)} = \frac{f(t)}{R(t)}h \equiv z(t)h$$

where we define the *hazard function* (also called *hazard rate* or *failure rate*) of T at time t by:

$$z(t) = \frac{f(t)}{R(t)}$$

Formal definition of hazard function is

$$z(t) = \lim_{h \to 0} \frac{P(t < T \leq t + h | T > t)}{h} = \frac{f(t)}{R(t)}$$

Example: For the exponential distribution we have $f(t) = \lambda e^{-\lambda t}$ and $R(t) = e^{-\lambda t}$, so

$$z(t) = \frac{f(t)}{R(t)} = \lambda \quad \textit{(not depending on time!)}.$$

Recall that $z(t) = \lim_{h\to 0} \frac{P(t<T\leq t+h|T>t)}{h}$.

Thus

$$z(t)h \approx P(t < T \leq t+h|T > t) = P(\text{fail in } (t, t+h)| \text{ alive at } t)$$

Suppose a typical T is large compared to time unit. Then for $h = 1$:

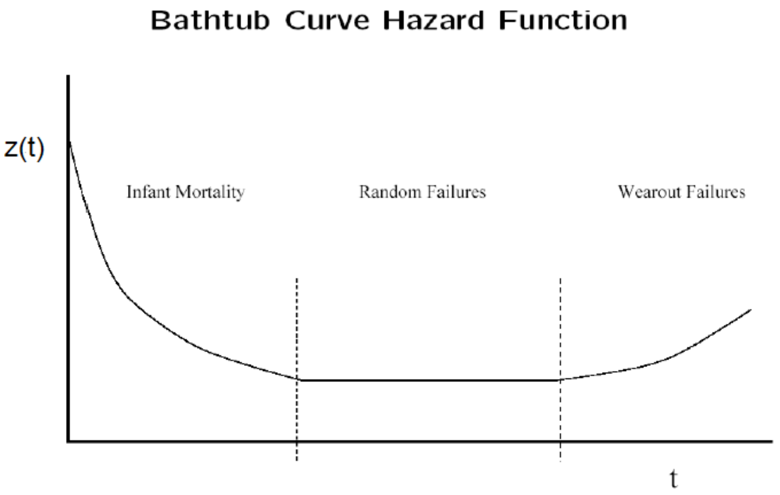
$$z(t) \approx P(t < T \leq t+1|T > t) = P(\text{fail in next time unit } |\text{alive at } t)$$

Thus: Suppose we have n units of age t. How many can we expect to fail in next time unit?

$$e \approx n \cdot z(t)$$

In practice: Ask an expert: "If you have 100 components (of specific type) of age 1000 hours. How many do you expect to fail in the next hour"? Answer is, say, "2". Assuming $e = n \cdot z(t)$ we estimate;

$$\hat{z}(1000) = \frac{2}{100} = 0.02$$

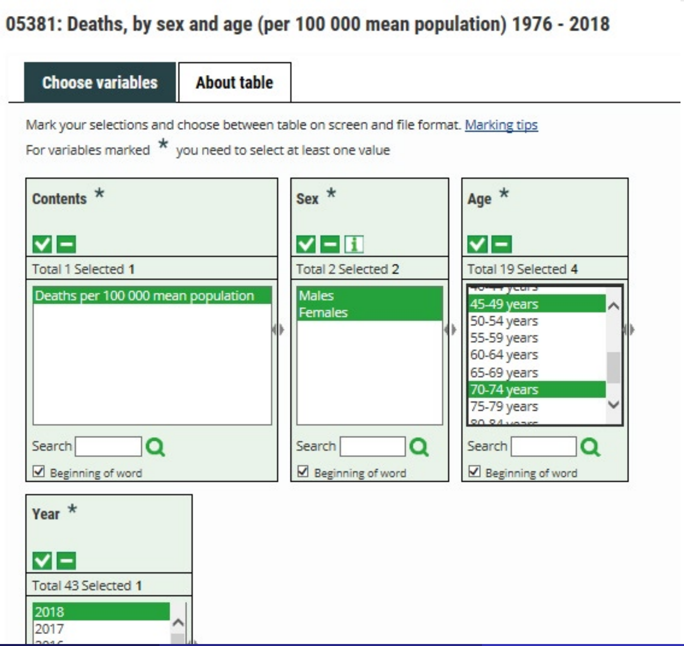

Bathtub Curve Hazard Function
z(t)
Infant Mortality
Random Failures
Wearout Failures
t

Let T be the lifetime of a Norwegian person measured in years.

Let $z_M(t)$ be the hazard function for a male person as a function of the age t, while $z_F(t)$ is the corresponding function for a female.

You may find information about these hazards on the webpages of *Statistics Norway (Statistisk Sentralbyrå)*

`https://www.ssb.no/en/statbank/list/dode/`


05381: Deaths, by sex and age (per 100 000 mean population) 1976 - 2018
Choose variables
About table
Mark your selections and choose between table on screen and file format. Marking tips
For variables marked * you need to select at least one value
Contents *
Total 1 Selected 1
Deaths per 100 000 mean population
Search
Beginning of word
Sex *
Total 2 Selected 2
Males
Females
Search
Beginning of word
Age *
Total 19 Selected 4
45-49 years
50-54 years
55-59 years
60-64 years
65-69 years
70-74 years
75-79 years
Search
Beginning of word
Year *
Total 43 Selected 1
2018
2017

MORTALITY TABLES

	Deaths per 100 000 mean population
	2018
Males	
20-24 years	33
45-49 years	156
70-74 years	1 990
90 years or older	24 230
Females	
20-24 years	6
45-49 years	111
70-74 years	1 331
90 years or older	21 909

Thus we can estimate, e.g.,

$$\begin{aligned} z_M(22) &\approx 33 \cdot 10^{-5} = 0.00033 \\ z_F(22) &\approx 6 \cdot 10^{-5} = 0.00006 \\ z_M(72) &\approx 1990 \cdot 10^{-5} = 0.01990 \\ z_F(72) &\approx 1331 \cdot 10^{-5} = 0.01331 \end{aligned}$$

Since $F(t) = 1 - R(t)$ we get, $f(t) = F'(t) = -R'(t)$, and hence

$$z(t) = \frac{f(t)}{R(t)} = -\frac{R'(t)}{R(t)}$$

Thus we can write,

$$\frac{d}{dt}\big(\ln R(t)\big) = -z(t)$$

$$\Rightarrow \ln R(t) = -\int_0^t z(u)du + c$$

$$\Rightarrow R(t) = e^{-\int_0^t z(u)du + c}$$

Since $R(0) = 1$, we have $c = 0$, so

$$R(t) = e^{-\int_0^t z(u)du} \equiv e^{-Z(t)}$$

where $Z(t) = \int_0^t z(u)du$ is called the *cumulative hazard function*.

Recall from last slide:

- $Z(t) = \int_0^t z(u)du$
- $z(t) = Z'(t)$
- $R(t) = e^{-Z(t)}$

Since $f(t) = F'(t) = -R'(t)$, it follows that

$$f(t) = z(t)e^{-\int_0^t z(u)du} = z(t)e^{-Z(t)} \tag{1}$$

For exponential distribution:

$$Z(t) = \int_0^t \lambda du = \lambda t$$

so (1) gives (the well known formula)

$$f(t) = \lambda e^{-\lambda t}$$

OVERVIEW OF FUNCTIONS DESCRIBING DISTRIBUTION OF LIFETIME T

Function	Formula	Exponential distr
Density (pdf)	f(t)	$= \lambda e^{-\lambda t}$
Cum. distr. (cdf)	F(t)	$= 1 - e^{-\lambda t}$
Rel/surv function	$R(t) = 1 - F(t)$	$= e^{-\lambda t}$
Hazard function	$z(t) = f(t)/R(t)$	$= \lambda$
Cum hazard function	$Z(t) = \int_0^t z(u)du$	$= \lambda t$
	$R(t) = e^{-Z(t)}$	$= e^{-\lambda t}$
	$f(t) = z(t)e^{-Z(t)}$	$= \lambda e^{-\lambda t}$

1. Suppose the reliability function of T is $R(t) = e^{-t^{1.7}}$.
   Find the functions $F(t), f(t), z(t), Z(t)$.
2. Show that if you get to know only one of the functions $R(t), F(t), f(t), z(t), Z(t)$, then you can still compute all the other!