Reliability
Part 1: Reminders of Lifetime Data Analysis

Jean-Yves Dauxois

INSA-Toulouse & IMT, FRANCE

4-8 June 2018

Lifetime and related functions

Lifetime = Time elapsed before the occurrence of an event:

- death, recovery or relapse in Survival Analysis
- Failure of a system or an equipment in Reliability
- Loss of employment in Econometric
- ...

Probabilistic Model: random variable (r.v.) $X \geq 0$, with **cumulative distribution function (c.d.f.)** $F(x) = P(X \leq x)$.

$R(x) = 1 - F(x)$=**Reliability function** (or **Survival Function** in Survival Analysis).

Hazard rate (or risk) function

If X is a continuous r.v., the **hazard rate function** $\lambda(\cdot)$ is defined by:

$$\begin{aligned}\lambda(x) &= \lim_{h\to 0^+} \frac{1}{h} P(X \in [x, x+h[\,|X \geq x) \\ &= \frac{f(x)}{R(x)},\end{aligned}$$

for $x \geq 0$, where $f(\cdot)$ is the probability density function (p.d.f.) of X.

The hazard rate at point x represents the instantaneous probability of failure (or death) at time x given that failure (or death) didn't occur before.

The hazard rate function may have different shapes: the most well known is called the **bathtup curve**.

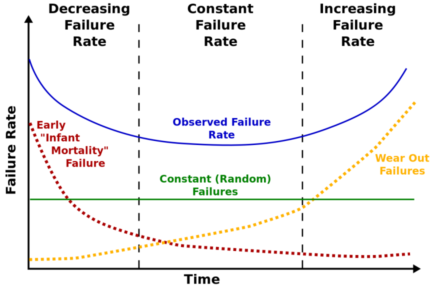

Figure: Bathtup curve (source: Wikipedia).

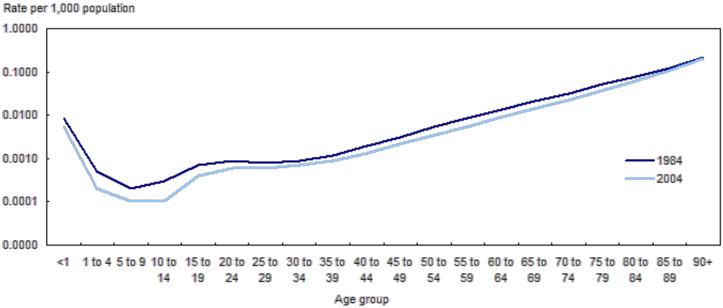

Figure: Age-specific mortality rates, Canada, in 1984 and 2004 (logarithmic scale) (source: http://www.statcan.gc.ca).

Cumulative hazard rate function :

$$\Lambda(x) = \int_0^x \lambda(s)ds, \text{ for all } x.$$

Important relations between these functions : we have, for all x:

$$\begin{aligned} R(x) &= \exp(-\Lambda(x)) \\ f(x) &= \lambda(x) \exp\left(-\int_0^x \lambda(s)ds\right). \end{aligned}$$

If the r.v. X is discrete, i.e. with values in the denumerable set $\{x_1 < x_2 < \cdots < x_n < \cdots\}$, we have:

$$F(x) = \sum_{i : x_i \leq x} p_i,$$

where $p_i = P(X = x_i)$.
The **Hazard rate function** is defined by:

$$\lambda(x_i) = P(X = x_i | X \geq x_i) = \frac{p_i}{R(x_{i-1})}.$$

The **Cumulative hazard rate function** is defined by:

$$\Lambda(x) = \sum_{i : x_i \leq x} \lambda(x_i).$$

We have:

$$R(x) = \prod_{i : x_i \leq x} (1 - \lambda(x_i)).$$

The mathematical expectation of X can be written in terms of the Reliability function:

$$\mathbb{E}(X) = \int_0^{+\infty} x dF(x) = \int_0^{+\infty} R(x) dx.$$

The **Residual Life at time** x, denoted by τ_x, is the r.v with distribution

$$P(\tau_x > y) = P(X - x > y | X > x) = \frac{R(x+y)}{R(x)}.$$

The **Mean Residual Life function** $m(\cdot)$, is defined for $x \geq 0$, by

$$\begin{aligned} m(x) &= \mathbb{E}(\tau_x) = \mathbb{E}(X - x | X \geq x) \\ &= \frac{\int_x^{+\infty} R(s) ds}{R(x)}. \end{aligned}$$

Censoring

Time-to-Failure, hr	Mode of Failure
105	A
125	B
134	A
167	C
212	C
345	A
457	B
541	C
623	B

Figure: Competing risks data (source: ReliaWiki.org).

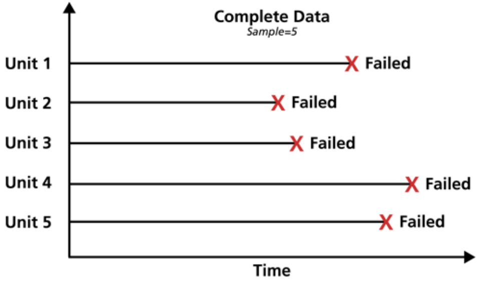

Figure: Complete data (source: ReliaWiki.org).

6-MP dataset

Pair	*Remission Status at Randomization*	*Time to Relapse for Placebo Patients*	*Time to Relapse for 6-MP Patients*
1	Partial Remission	1	10
2	Complete Remission	22	7
3	Complete Remission	3	32^+
4	Complete Remission	12	23
5	Complete Remission	8	22
6	Partial Remission	17	6
7	Complete Remission	2	16
8	Complete Remission	11	34^+
9	Complete Remission	8	32^+
10	Complete Remission	12	25^+
11	Complete Remission	2	11^+
12	Partial Remission	5	20^+
13	Complete Remission	4	19^+
14	Complete Remission	15	6
15	Complete Remission	8	17^+
16	Partial Remission	23	35^+
17	Partial Remission	5	6
18	Complete Remission	11	13
19	Complete Remission	4	9^+
20	Complete Remission	1	6^+
21	Complete Remission	8	10^+

$^+$Censored observation

Figure: Clinical trial: 6-mercaptopurine versus placebo (Freirich et al. *Blood* 21, 1963). Time in months.

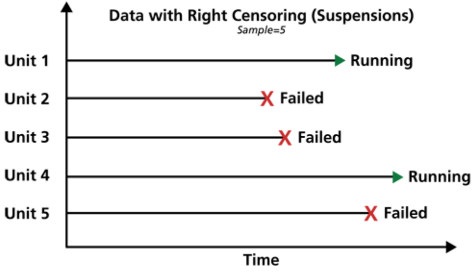

Figure: Right Censored data (source: ReliaWiki.org).

Definition

The lifetime X is said to be **right censored** *(resp.* **left censored***) by C if, instead of observing X, one observe $X \wedge C$ (resp. $X \vee C$) where C is a r.v., $a \wedge b = \min(a, b)$ and $a \vee b = \max(a, b)$.*

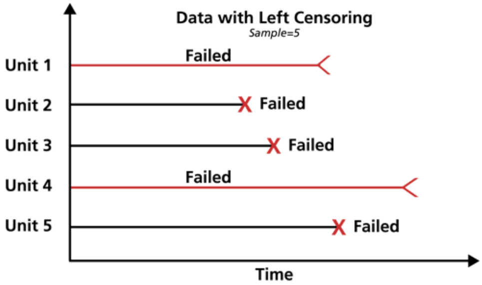

Figure: Left Censored data (source: ReliaWiki.org).

Age	*Number of Exact Observations*	*Number Who Have Yet to Smoke Marijuana*	*Number Who Have Started Smoking at an Earlier Age*
10	4	0	0
11	12	0	0
12	19	2	0
13	24	15	1
14	20	24	2
15	13	18	3
16	3	14	2
17	1	6	3
18	0	0	1
>18	4	0	0

Figure: Marijuana use in high school boys (from Turnbull and Weiss, 1978).

Type I right censoring. The censoring time c is known and fixed. For example c is equal to the end of study time. One only observes the r.v. $(T_i, \delta_i)_{i=1,\ldots,n}$ defined by

$$\left\{\begin{array}{lcl} T_i & = & \min(X_i, c) \\ \delta_i & = & \mathbb{1}_{\{X_i \leq c\}} \end{array}\right., \text{ for } i = 1, \ldots, n \,.$$

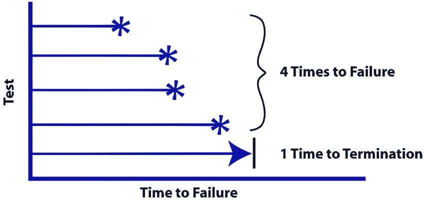

Figure: Type I right Censored data (source: ReliaWiki.org).

Progressive type I right censoring. The censoring times c_i, for $i = 1, \ldots, n$ are known and fixed. One only observes the r.v. $(T_i, \delta_i)_{i=1,\ldots,n}$ defined by

$$\left\{ \begin{array}{lcl} T_i & = & \min(X_i, c_i) \\ \delta_i & = & 1\!\!1_{\{X_i \leq c_i\}} \end{array} \right. , \text{ for } i = 1, \ldots, n \; .$$

Type II right censoring. The censoring time is given by the time of the rth failure observed in the sample. One only observes the r.v. $(T_i, \delta_i)_{i=1,\ldots,n}$ given by

$$\begin{cases} T_i &= \min(X_i, X_{(r)}) \\ \delta_i &= \mathbb{1}_{\{X_i \leq X_{(r)}\}} \end{cases}, \text{ for } i = 1, \ldots, n \,,$$

where $X_{(1)} \leqslant X_{(2)} \leqslant \cdots \leqslant X_{(r)}$ are the r first order statistics.

Random censoring. One observes the r.v. $(T_i, \delta_i)_{i=1,\dots,n}$ given by:

$$\left\{ \begin{array}{lcl} T_i & = & \min(X_i, C_i) \\ \delta_i & = & \mathbb{1}_{\{X_i \le C_i\}} \end{array} \right. , \text{ for } i = 1, \dots, n \, ,$$

where $C_1, \dots, C_n$ are r.v. with c.d.f. $G_1, \dots, G_n$ respectively.

The r.v. $X_1, \dots, X_n, C_1, \dots, C_n$ are generally assumed to be independent.

Example

Consider failure times of a component with different modes of failure.

Time-to-Failure, hr	Mode of Failure
105	A
125	B
134	A
167	C
212	C
345	A
457	B
541	C
623	B

Figure: Competing risks data (source: ReliaWiki.org).

Definition

The lifetime X is said to be **interval censored** *if, instead of observing X, one observes only a (possibly random) interval $[L, U]$ such that $X \in [L, U]$.*

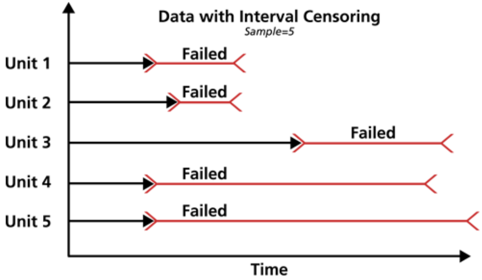

Figure: Interval Censored data (source: ReliaWiki.org).

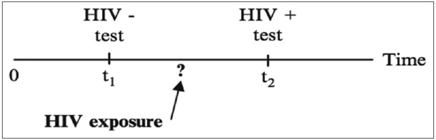

Figure: Example of Interval Censored data (source: http://www.jcrsmed.org).

Truncation

Definition

The lifetime X is said to be **left truncated** *(resp.* **right truncated***) by C if one can observe the r.v. X only when X is greater than (resp. lower) than C.*

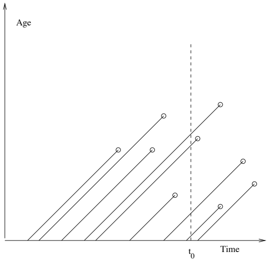

Figure: Example of left truncated data obtained from cross sectional observation. Only the survival time of individuals alived at time t_0 are known.

Exponential distribution $\mathcal{E}(\lambda)$

X has an $\mathcal{E}(\lambda)$ distribution if one of the following (equivalent) equations is fulfilled.

$$\begin{aligned} F(x) &= \begin{cases} 1 - e^{-\lambda x} & \text{for } x \geq 0 \\ 0 & \text{elsewhere} \end{cases} \\ f(x) &= \lambda \exp(-\lambda x), \text{ for } x \geq 0 \\ \lambda(x) &= \lambda, \text{ for } x \geq 0 \\ m(x) &= \frac{1}{\lambda}, \text{ for } x \geq 0. \end{aligned}$$

One can see that:

$$\begin{aligned}\mathbb{E}(X) &= 1/\lambda \\ Var(X) &= 1/\lambda^2.\end{aligned}$$

The exponential distribution is well known for its **memoryless property** :

$$\begin{aligned} P(X \leqslant x + x_0 \ / \ X > x_0) &= P(X \leqslant x), \ \forall (x, x_0) \in \mathbb{R}^+ \times \mathbb{R}^+ \\ \Leftrightarrow \bar{F}(x + x_0) &= \bar{F}(x)\bar{F}(x_0), \ \forall (x, x_0) \in \mathbb{R}^+ \times \mathbb{R}^+ \\ \Leftrightarrow \mathcal{L}(\tau_{x_0}) &= \mathcal{L}(X), \ \forall x_0 \in \mathbb{R}^+. \end{aligned}$$

There is an important link between the exponential distribution and the **Poisson Process**: $\Rightarrow$ recurrent failures on a system with exponential interarrival times.

Weibull distribution $W(\alpha, \beta)$

X has a $W(\alpha, \beta)$ distribution if its c.d.f. is given by:

$$F(x) = 1 - \exp\left(-\left(\frac{x}{\alpha}\right)^{\beta}\right), \text{ for } x \geq 0,$$

where α and β are respectively the scale and the form parameters.

$$\begin{aligned} f(x) &= \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} \exp\left(-\left(\frac{x}{\alpha}\right)^{\beta}\right), \text{ for } x \geq 0 \\ \lambda(x) &= \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1}, \text{ for } x \geq 0. \end{aligned}$$

$(\beta = 1) \Rightarrow$ Exponential distribution with prameter $1/\alpha$.

One can show that:

$$\begin{aligned}\mathbb{E}(X) &= \alpha\,\Gamma(1+1/\beta)\\ \mathrm{Var}(X) &= \alpha^2\left[\Gamma(1+2/\beta)-\Gamma^2(1+1/\beta)\right].\end{aligned}$$

Remarks.

- The r.v. $(X/\alpha)^\beta$ has an exponential distribution with parameter 1.
- The minimum of n i.i.d. r.v. with same Weibull $W(\alpha,\beta)$ distribution has a Weibull $W(\alpha/n^{1/\beta},\beta)$ distribution.

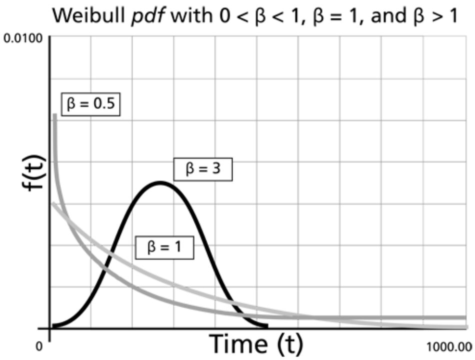

Figure: Weibull p.d.f. (source: ReliaWiki.org).

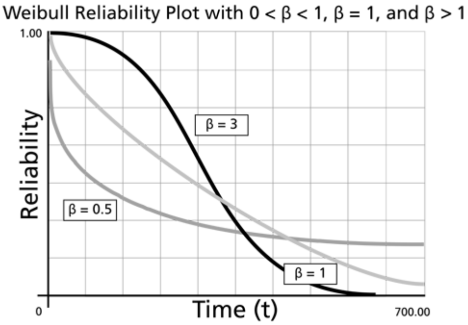

Figure: Weibull Reliability function (source: ReliaWiki.org).

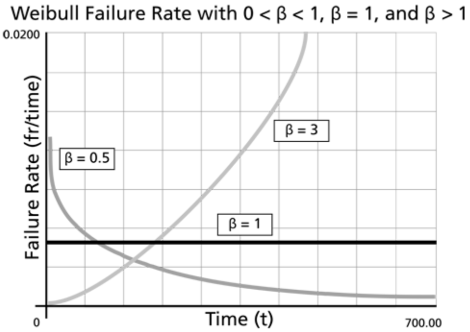

Figure: Weibull hazard rate function (source: ReliaWiki.org).

IFR and DFR distributions

Definition

The distribution has the Increasing Failure Rate (IFR), resp. Decreasing Failure Rate (DFR), property when the hazard rate function $\lambda(\cdot)$ is increasing, resp. decreasing.

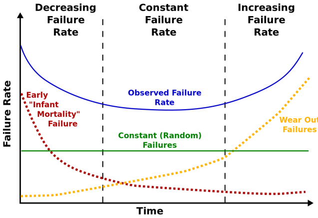

Figure: Bathtup curve (source: Wikipedia).

Proposition

The following assertions are equivalent

i) *The distribution is IFR*

ii) *The cumulative hazard rate function $\Lambda(\cdot)$ is convex*

iii) *The residual lifetime at time x is stochastically greater than the one at time x', for $x \leq x'$, i.e.*

$$\forall (x, x') \in \mathbb{R}^+ \times \mathbb{R}^+ \text{such that } x \leq x', \text{ we have:}$$
$$P(\tau_x > y) \geq P(\tau_{x'} > y), \text{ for all } y > 0.$$

Example

.

Weibull distribution:

- If $\beta = 1$, the hazard rate function is constant (exponential distribution).
- if $\beta > 1$, the distribution is IFR.
- si $\beta < 1$, the distribution is DFR.

Statistical Inference in parametric models

Let us consider a parametric model

$$\mathcal{F} = \{P_\theta, \theta \in \Theta\} \text{ or } \mathcal{F} = \{f_\theta, \theta \in \Theta\} \text{ or } \mathcal{F} = \{F_\theta, \theta \in \Theta\}$$

for the lifetime X.

Example: Exponential model $\{R_\lambda(x) = e^{-\lambda x}; \lambda > 0\}$.

Problem: How to estimate the (possibly) multidimensional parameter θ? Many different methods are available...

You certainly know how to do Maximum Likelihood Estimation (MLE) with a complete sample (i.e. without censoring). Is it the case?

MLE with censored data

It is more difficult when censoring occurs...

Let us first consider the case where the censoring times are deterministic.

Based on the intuitive idea of the likelihood, one can understand that the likelihood of the observations is like:

$$\begin{aligned} L(\text{observations}, \theta) &= \prod_{i \in D} f_\theta(t_i) \prod_{i \in R} R_\theta(t_i) \prod_{i \in L} F_\theta(t_i) \\ &\times \prod_{i \in I} \left(F_\theta(U_i) - F_\theta(L_i)\right), \end{aligned}$$

where D, R, L and I are respectively the sets of individuals where the observation corresponds to a time of death, right censoring, left censoring or interval censoring.

Likelihood with randomly right censored data

Under **assumption of independence between lifetime and censoring time** the likelihood is given by:

$$L((t_1, \delta_1), \ldots, (t_n, \delta_n); \theta) = \prod_{i=1}^{n} \left[f_\theta(t_i)\bar{G}(t_i)\right]^{\delta_i} \left[\bar{F}_\theta(t_i)g(t_i)\right]^{1-\delta_i}.$$

If the distribution of the censoring time doesn't depend on the parameter of interest θ, one can consider the likelihood

$$L((t_1, \delta_1), \ldots, (t_n, \delta_n); \theta) = \prod_{i=1}^{n} (f_\theta(t_i))^{\delta_i} (\bar{F}_\theta(t_i))^{1-\delta_i}.$$

The maximum likelihood estimator is the value of θ which maximizes the likelihood.

Example. Lifetimes with exponential distribution: complete and right censored observations.

Nonparametric inference

Here we do not assume any parametric model for the lifetime.

The aim is to estimate its distribution (the functions of interest) using only the data and without any assumption.

With uncensored observations

Suppose that we observe a sample $X_1, ..., X_n$ of the lifetime X. The empirical cumulative distribution function

$$\widehat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{X_i \le x\}} = \frac{1}{n}\sum_{i:x_{(i)}\le x} \mathbb{1}_{\{X_{(i)} \le x\}}.$$

is an estimator of the c.d.f. $F(\cdot)$.
$\widehat{R}_n(x) = 1 - \widehat{F}_n(x)$ is an estimator of the Reliability function.

It's equivalent to estimate the hazard rate $\lambda(\cdot)$ by

$$\begin{aligned} \widehat{\lambda}_n(x_{(i)}) &= \frac{1}{n-i+1}, \text{ for } i = 1, \ldots, n \\ \text{and } \widehat{\lambda}_n(x) &= 0, \text{ for all } x \text{ where there is no observation.} \end{aligned}$$

The cumulative hazard rate function can be estimated by

$$\widehat{\Lambda}_n(x) = \sum_{i:x_{(i)} \leq x} \frac{1}{n-i+1}.$$

Example. Competing Risks dataset. Estimations of the lifetime c.d.f., hazard rate and cumulative hazard rate functions if we do not take into account the different types of failure.

Time-to-Failure, hr	Mode of Failure
105	A
125	B
134	A
167	C
212	C
345	A
457	B
541	C
623	B

Figure: Competing risks data (source: ReliaWiki.org).

With censored observations

Suppose that we observe a sample of possibly right censored data $(T_i, \delta_i)_{i=1,\ldots,n}$.

Example. Competing Risks dataset. Estimations of the Reliability, hazard rate and cumulative hazard rate functions of the time before failure of type A.

Time-to-Failure, hr	Mode of Failure
105	A
125	B
134	A
167	C
212	C
345	A
457	B
541	C
623	B

Figure: Competing risks data (source: ReliaWiki.org).

Let $n_1 = \sum_{i=1}^n \delta_i$ be the number of uncensored data in the sample. One may want to use

$$\widehat{R}_n^{(1)}(t) = \frac{1}{n_1} \sum_{i=1}^n \mathbb{1}_{\{T_i > t, \delta_i = 1\}}$$

as an estimator of $R(t)$. **But this is not a good idea!** Indeed one can show that

$$\widehat{R}_n^{(1)}(t) \xrightarrow{p.s.} \frac{\int_t^{+\infty} \bar{G}(x)dF(x)}{P(\delta = 1)} = \frac{\int_t^{+\infty} \bar{G}(x)dF(x)}{\int_0^{+\infty} \bar{G}(x)dF(x)} \neq R(t),$$

except if $\bar{G}(x) = 1$, for $x > t$, which means that there is no censoring.

Nelson-Aalen and Kaplan-Meier estimators

First, let us suppose that there is **no ex-aequo** in the sample, as it is the case for the Competing Risks dataset.

Let $T_{(1)} \leq \cdots \leq T_{(n)}$ be the n observed and ordered times and $\delta_{(1)}, \ldots, \delta_{(n)}$ their corresponding indicators. One can estimate the hazard rate function $\lambda(\cdot)$ by:

$$\begin{aligned} \widehat{\lambda}_n(t_{(i)}) &= \frac{\delta_{(i)}}{n-i+1}, \text{ for } i = 1, \ldots, n \\ \text{and } \widehat{\lambda}_n(x) &= 0, \text{ elsewhere.} \end{aligned}$$

The Nelson-Aalen estimator of the cumulative hazard rate function $\Lambda(\cdot)$ is:

$$\widehat{\Lambda}_n(x) = \sum_{i:T_{(i)}\leq x} \frac{\delta_{(i)}}{n-i+1}.$$

The Kaplan-Meier estimator of the Reliability function $R(\cdot)$ is

$$\widehat{R}_n(x) = \prod_{i:T_{(i)}\leq x} \left(1 - \frac{\delta_{(i)}}{n-i+1}\right).$$

One can find in the literature another expression of the Kaplan-Meier estimator:

$$\widehat{R}_n(x) = \prod_{i:T_{(i)}\leq x} \left(1 - \frac{1}{n-i+1}\right)^{\delta_{(i)}} \mathbb{1}_{\{x\leqslant T_{(n)}\}}.$$

In the case where there are some ex-aequo in the sample, one can use the following expression of the Kaplan-Meier estimator.

Let $T_1' < \cdots < T_k'$ be the k different distinct and ordered times observed in the sample. Let M_i be the number of death (or failure) observed at T_i' and Y_i the number of subject at risk at time T_i'. The Kaplan-Meier estimator of $R(\cdot)$ can be written

$$\widehat{R}_n(x) = \prod_{i:T_i' \leq x} (1 - \frac{M_i}{Y_i}).$$

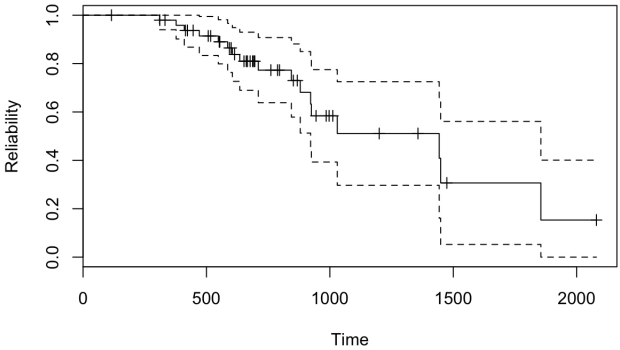

Figure: Example of Kaplan-Meier estimates with pointwise confidence intervals.

Estimation of the mean lifetime

An estimator of the expectation $\mathbb{E}(X)$ one can use:

$$\begin{aligned}\widehat{\mu}_n &= \int_0^{+\infty} t\, d\widehat{F}_n(t) = \sum_{i=1}^{k} T'_i\, \Delta\widehat{F}_n(T'_i) \\ &= \sum_{i=1}^{k} T'_i \frac{M_i}{Y_i} \prod_{j=1}^{i-1}\left(1 - \frac{M_j}{Y_j}\right).\end{aligned}$$

We have:

$$\hat{\mu}_n \xrightarrow{p.s.} \int x dF(x) = \mathbb{E}(X), \text{ when } n \to +\infty.$$